Analyse à l'échelle régionale de l'impact de l'occupation du sol dans les corridors rivulaires sur l'état écologique des cours d'eau

T. Tormos

▶ **To cite this version:**

T. Tormos. Analyse à l'échelle régionale de l'impact de l'occupation du sol dans les corridors rivulaires sur l'état écologique des cours d'eau. Sciences de l'environnement. Doctorat AgroParisTech, Eaux continentales et Société, 2010. Français. NNT : . pastel-00560368v1

HAL Id: pastel-00560368

https://hal.inrae.fr/pastel-00560368v1

Submitted on 15 May 2020 (v1), last revised 28 Jan 2011 (v2)

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.

THÈSE

pour l'obtention du grade de

DOCTEUR D'AGROPARISTECH

Institut des Sciences et Industries du Vivant et de l'Environnement

Spécialité : Eaux Continentales et Société

présentée et soutenue publiquement le *6 Décembre 2010* par

Thierry Tormos

ANALYSE À L'ÉCHELLE RÉGIONALE DE L'IMPACT DE L'OCCUPATION DU SOL DANS LES CORRIDORS RIVULAIRES SUR L'ÉTAT ÉCOLOGIQUE DES COURS D'EAU

Membres du jury :

M. John THOMPSON	Directeur de recherches, CNRS, Montpellier	Président du jury
M. Yves SOUCHON	Directeur de recherches, Cemagref, Lyon	Rapporteur
Mme Laurence HUBERT-MOY	Professeur d'Université, Rennes 2	Rapporteur
M. Etienne MULLER	Ingénieur de recherches, CNRS, Toulouse	Examinateur
M. Philip ROCHE	Directeur de recherches, Cemagref, Aix-en-Provence	Examinateur
M. Hervé PIEGAY	Directeur de recherches, CNRS, Lyon	Examinateur
M. Pascal KOSUTH	ICPEF, Cemagref, Montpellier	Directeur de Thèse

L'Institut des Sciences et Industries du Vivant et de l'Environnement (Agro Paris Tech) est un Grand Etablissement dépendant du Ministère de l'Agriculture et de la Pêche, composé de l'INA PG, de l'ENGREF et de l'ENSIA (decret n° 2006-1592 du 13 decembre 2006)

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

A mon Père

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

REMERCIEMENTS

Peut être, quelque part,
est-il des efforts créatifs accomplis seuls.
Ceci n'en est pas un.

Ces travaux de thèse se sont déroulés au sein l'équipe de recherche « Territoires, Environnement, Télédétection et Analyse Spatiale » (URM TETIS AgroParisTech-Cemagref-Cirad) de Montpellier et en forte collaboration avec l'équipe du « Laboratoire d'Hydroécologie Quantitative » (LHQ Cemagref) de Lyon.

Ils ont été financés en majeure partie par le Ministère de l'Enseignement Supérieur et de la Recherche au travers d'une allocation de recherche accordée par l'école doctorale Terre-Eau-Espace (actuelle école doctorale SIBAGHE) de l'Université de Montpellier et ensuite sur la base des conventions ONEMA/Cemagref avec l'UMR TETIS.

En premier lieu, mes remerciements vont à PASCAL KOSUTH qui a accepté avec enthousiasme de diriger cette thèse. Même si sa disponibilité a été, comment dire, souvent « nocturne », il a su être présent avec rigueur et efficacité dans les moments clés et participer activement aux réflexions de ce travail. Je tiens à le remercier également pour son soutien infaillible, et les moments toujours agréables qui ont été pour moi enrichissants tant sur le plan scientifique que personnel. J'attends avec impatience ces prochains moments.

C'est avec beaucoup de reconnaissance que je remercie ensuite les membres du jury qui ont su faire preuve de patience et de compréhension par rapport aux différents reports d'échéances qu'a connu ce travail : les rapporteurs YVES SOUCHON (Cemagref), LAURENCE HUBERT-MOY (Université de Rennes 2) ; les examinateurs ETIENNE MULLER (CNRS), PHILIP ROCHE (Cemagref), HERVE PIEGAY (CNRS) ; et le président du jury JOHN THOMPSON (CNRS).

Parmi les membres du jury, je tiens spécialement à remercier YVES SOUCHON qui m'a accueilli chaleureusement dans son équipe et qui a toujours été indulgent et compréhensif vis-à-vis de ma situation. Ses orientations, délivrées le plus souvent sous forme de paronymie, de calembours et autres jeux de mots, ont été très précieuses dans les derniers moments de la thèse.

Je tiens à remercier particulièrement JEAN GABRIEL WASSON car sans lui la thèse n'aurait certainement pas vu le jour. Je regrette qu'il ne puisse pas en voir la fin, il nous a malheureusement quitté trop tôt. Instigateur de la mise en place de la directive cadre européenne sur l'eau et fervent porte parole de la protection des corridors rivulaires en France, il a œuvré durant sa vie professionnelle à la compréhension et la protection des écosystèmes aquatiques. En plus des précieuses discussions que nous avons eues, je me suis longuement nourri durant la thèse de ses nombreux articles, actes de colloques et autres rapports, qui m'ont sans aucun doute au fil de leur lecture persuadé de continuer à travailler dans ce domaine. Merci JEAN GABRIEL.

Ensuite je tiens à remercier les nombreuses personnes qui m'ont permis de réaliser cette thèse et m'ont accompagné dans ce périple :

En particulier SYLVIE DURRIEU. Merci à l'attention que tu as portée à ce travail, aux remarques toujours très pertinentes et aux enseignements sur les bases de la télédétection que tu m'as distillées aussi bien lors de mon DEA que tout au long de la thèse.

Je remercie également CHRISTINE ARGILLIER et son équipe. Notre collaboration sur l'application de la télédétection pour la caractérisation de l'hydromorphologie lacustre a été très instructive et a contribué à me donner le recul nécessaire pour approfondir mon travail de thèse. Mais surtout, grâce à eux j'ai pu faire « une » sortie terrain au cours de la thèse (ouf !).

CemOA : archive ouverte d'Irstea / Cemagref

Je tiens à remercier l'ensemble des membres du LHQ (et aujourd'hui mes collègues du pôle Onema/Cemagref d'hydroécologie des cours d'eau) pour leur accueil et leur soutien dans la finalisation de ma thèse. Je remercie particulièrement BERTRAND VILLENEUVE pour m'avoir suivi dès le début de ma thèse et pour m'avoir guidé dans les méandres de la modélisation statistique, ainsi que MARTIAL FERREOL pour ces conseils en statistiques générales, et pour avoir constitué avec moi la paire redoutée et redoutable « Super Poussin – Pyjama Boy » que l'on peut apercevoir sur leurs montures d'acier (heu, plutôt aluminium) le long des quais de la Saône lorsque le temps s'y prête.

Ensuite, spéciale dédicace à GUY ROUSSIN, qui pour une fois ne pourra pas se plaindre d'avoir été (selon lui) oublié dans cet exercice. Je tiens à le remercier (même s'il supporte les couleurs d'une équipe qui n'en est pas une) pour ses conseils avisés sur le développement informatique, pour son soutien « unixien » sur le fameux cluster de la MTD (Maison de la Télédétection), pour sa confiance aveugle (il m'a quand même laissé partir avec plus de 10 000 euros de matériel pour aller récupérer les orthophotos à Toulouse, quelle expérience ça aussi !!), pour ces nombreux cafés qu'il m'a offerts à ma demande, et les nombreux moments de détente que l'on a partagés. J'espère qu'il me pardonnera un jour d'avoir ramené avant lui la coupe du concours de pétanque du Cemagref à la MTD, même si mon départ lui a permis de la remporter à son tour.

Je tiens à remercier amicalement STEPHANE DUPUY. L'exercice à son sujet devient complexe, car il me vient en tête un nombre incalculables de remerciements. Utilisateur aguerri de l'analyse orientée objet, je commencerai d'abord par le remercier pour avoir débloqué un grand nombre d'embuches et pour m'avoir éclairé dans le développement de mes méthodes. Ensuite, je le remercierai pour avoir été constamment à mes côtés, pour m'avoir encouragé et soutenu dans ce périple, avoir pris le temps de relire l'ensemble de la thèse (même si au final les ¾ de la thèse lui sont étrangers), avoir écouté mes états d'âme, et … en ce moment même de corriger ces quelques lignes …bref pour faire court, je lui exprime toute ma reconnaissance pour son amitié.

Dans le bureau à côté, je tiens à remercier GERARD LAINE, le « sage » de la MTD, pour ses conseils avisés en télédétection, ses nombreuses anecdotes, et pour m'avoir fait découvrir une disquette de stockage d'au moins 20 cm de côté (oui oui ça existe encore !) entreposée de façon bien ordonnée avec d'autres reliques appartenant au patrimoine de la télédétection. J'aurai certainement le plaisir de le retrouver dans sa « payotte » à Lez GéGé en passant par le chemin des « pécheresses » pour aller à la fameuse cantine du Cirad. D'ailleurs, je remercie l'équipe pédestre du Cirad de m'avoir accueilli dans leur cantine « Tout ça Tout ça » (comme dit José) grâce au précieux sésame prêté par JEAN-CLAUDE LORENTE.

Je remercie également les personnes qui, au quotidien, ont apporté leur contribution au bon déroulement de mon travail de thèse. Merci à ALAIN DELACOURT, FRANÇOISE BOISSIER et SYLVIE BLIN (promis Sylvie je vais te rendre les ouvrages et autres documents que je t'ai empruntés, si si cette fois-ci c'est vrai !).

Je tiens à remercier mes collègues et amis de bureau. Tout d'abord MAHMOUD pour ses nombreux conseils et son assistance au quotidien : j'ai eu, à n'en pas douter, la chance de partager mon bureau avec un futur « number one ». Ensuite un grand merci à mes acolytes « écothésardes » VALOU, ELO et CLAIRE, pour les séances de décompression « du père Fouras », pour les grands moments de délire et de pétages de plombs qui, dans notre quotidien de « TDM », nous ont fait redescendre sur terre.

Je tiens à remercier mes collocs et amis MAROUAN (dit Marouan le « français ») et ARTHUR (le futur marié), qui embarqués dans la même galère ont su me comprendre et partager mes états d'âmes, puis au quotidien supporter mon chauvinisme phocéen.

Enfin, je tiens à remercier également ma famille qui m'a soutenu affectueusement durant ce long périple, à leurs yeux interminable. Mais ça y est enfin, je vais pouvoir leur annoncer un premier « Scoop » (enfin !).

Mais surtout à Cécile qui a suivi de très près cette thèse …Merci à toi, mais te relâche pas trop un autre périple nous attends ! Bon je crois que c'est bon … Ah oui j'ai oublié Lilette, merci pour tes câlins et ta présence qui a été parfois un peu trop débordante pendant la rédaction du présent rapport (miaou).

CemOA : archive ouverte d'Irstea / Cemagref

RESUME

L'état écologique des milieux aquatiques est influencé par l'occupation des sols environnante à de multiples échelles, du bassin versant amont au corridor rivulaire local. Préserver ou restaurer ces écosystèmes, comme le préconise la Directive Cadre européenne sur l'Eau, implique de comprendre au niveau régional, voire national, les relations entre leur état écologique et les pressions environnantes. Un intérêt particulier doit être porté aux corridors rivulaires pour élaborer des stratégies de préservation et de restauration compte tenu du rôle protecteur et régulateur des formations végétales rivulaires sur les écosystèmes aquatiques et du fait que la gestion du foncier est plus facilement négociable dans ces espaces qu'à l'échelle du bassin versant. La question centrale de cette recherche est la caractérisation de l'influence de l'occupation des sols dans les corridors rivulaires sur l'état écologique des cours d'eau, caractérisé par l'EQR-IBGN (écart normalisé entre l'Indice Biologique Global Normalisé et sa valeur de référence).

Une telle analyse doit reposer sur des méthodologies robustes et fiables de cartographie fine de l'occupation des sols dans les corridors rivulaires, de construction d'indicateurs spatialisés traduisant cette occupation et de modèles pressions/impacts quantifiant le lien entre occupation des sols dans les corridors rivulaires et état écologique des milieux aquatiques. La thèse a été consacrée au développement et à la validation de méthodes adaptées à ces trois phases.

Dans une première phase, une procédure de classification a été conçue, capable d'extraire finement l'occupation du sol le long des cours d'eau sur de grands territoires à partir de l'information de données spatiales multi-source - images aéroportées et satellitaires à Très Haute Résolution Spatiale (THRS, métrique) et données thématiques - mobilisables à des coûts réduits pour les gestionnaires sur l'ensemble du territoire français. Cette procédure repose sur une analyse orientée objet multi-échelle qui combine, à l'aide de règles de classification floues définies par expertise, l'information des orthophotos (0.5 m), des images satellites SPOT 5 XS (10 m) et de données thématiques complémentaires. Son application sur le bassin versant de l'Hérault (1150 Kkm de réseau hydrographique) démontre (i) son efficacité : 89 % de précision globale selon une nomenclature regroupant les principales catégories exerçant des pressions sur le milieux aquatique ; (ii) sa généricité : les règles de classification ont été facilement transposée sur les contextes géographiques contrastés du bassin ; (iii) son applicabilité sur de grands territoires : les temps de traitement sont réduits du fait de la rapidité de transfert des règles de classification et de la possibilité de distribuer les traitement liés à l'analyse orientée objet sur un cluster de calcul.

Une seconde phase a été consacrée à la construction d'indicateurs spatialisés caractéristiques des corridors rivulaires, traduisant les connaissances écologiques sur les mécanismes d'impact des activités humaines et d'influence de la végétation rivulaire sur l'état écologique des cours d'eau. Des méthodologies automatisées ont été développées de manière à disposer de plusieurs emprises spatiales candidates afin d'améliorer l'identification, dans les modèles pressions/impacts, de l'emprise au sein de laquelle une catégorie d'occupation du sol a une influence significative sur la variabilité de la réponse biologique du cours d'eau. Ces méthodes de calcul d'indicateurs de pression liés au corridor rivulaire ont été mises en œuvre sur le même bassin de l'Hérault et ont démontré (i) l'intérêt de la THRS pour quantifier correctement les indicateurs au niveau du corridor rivulaire ; et (ii) la capacité de ces indicateurs à fournir une information cartographique explicite pour guider les gestionnaires dans leurs stratégies de restauration sur le corridor rivulaire.

Enfin une troisième phase a été consacrée à la modélisation régionalisés pressions / impacts liant l'occupation des sols sur le bassin amont, le corridor rivulaire amont et le corridor rivulaire à l'état écologique observé aux stations. Cette approche de construction de modèles pressions/impacts a été mise en œuvre sur la région Nord-Ouest de l'Hydro-écorégion Tables Calcaire (6000 km de réseau hydrographique) où le nombre de stations de mesure de l'Indice Biologique Global Normalisé sur la période de 1992 à 2004 est suffisant (155 stations) pour engager une approche statistique. La comparaison sur cette région de différentes structures de modèles montre l'intérêt de disposer d'une donnée à THRS et d'indicateurs calculés sur différentes emprises spatiales pour améliorer l'identification et la localisation des sources de pressions dans ces espaces. Ces avancées méthodologiques ont permis de démontrer à l'échelle régionale le rôle significatif des formations végétales rivulaires locales sur l'état écologique des cours d'eau, ceci sur une bande de 20 m de part et d'autre du cours d'eau

Parallèlement au développement de ces méthodologies, la propagation des incertitudes de la carte d'occupation des sols aux indicateurs spatiaux des corridors rivulaires, puis des indicateurs jusqu'aux résultats des modèles pressions / impacts a été analysée. La difficulté de spatialisation des incertitudes de la carte d'occupation des sols demeure un frein majeur à ce type d'analyse.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

RESUME EN ANGLAIS (ABSTRACT)

Preservation and restoration of river ecosystems require an improved understanding of the mechanisms through which they are influenced by landscape at multiple spatial scales from the overall upstream watershed to the local river corridor. A key question in this research, relates with the riparian landscape influence considering the role of riparian vegetation for regulating and protecting river ecological status and the relevance of this specific area (compared to watershed) for implementing efficient and realistic preservation and restoration strategies.

Such an analysis should be based on reliable and robust methodologies for (i) fine-scale mapping of land cover in riparian corridors, (ii) construction of spatial indicators reflecting this occupation and (iii) pressures / impacts models quantifying the relationships between land cover in riparian corridor and ecological status of aquatic environments. The thesis was devoted to the development and validation of methods appropriate to these three phases.

In a first phase, a classification procedure was designed, capable of extracting finescale land cover maps along rivers over large territories using information from multi-source spatial data (Very High Spatial Resolution (VHRS, metric) remotely sensed imagery and thematic data) affordable at low costs for managers on the French territory. This classification procedure is based on multi-scale object-oriented analysis that combines, using fuzzy classification rules defined by expertise, information from orthophotos (0.5 m pixels) and satellite images SPOT 5 XS (10 m pixels) and complementary thematic data. Its implementation over the Herault watershed (1150 km of river system) clearly demonstrate the method (i) efficiency : 89% overall accuracy using a nomenclature involving the main land cover categories exerting pressures on the aquatic environment, (ii) genericity : the classification rules were easily transposed to the contrasted geographical contexts of the basin, (iii) applicability on large areas : processing times are reduced because of the rapid transfer of classification rules and the possibility to distribute the processes related to object oriented analysis on a computer cluster.

The second phase of the thesis was devoted to the construction of spatial indicators reflecting ecological knowledge on the mechanisms of human impact and riparian vegetation influence on the ecological status of rivers. Automated methodologies have been developed in order to compute indicators on a wide range of riparian footprints in order to identify, in pressures / impacts models, the footprint in which a land cover category has the most significant influence on the variability of the stream biological response. These methods of calculation of pressure indicators related to the riparian corridor were tested in the Herault watershed. They clearly demonstrated (i) the interest of VHRS to quantify properly these indicators at riparian corridor scale; and (ii) the ability of these indicators to provide explicit mapping information useful to guide managers in their strategies for restoration of riparian corridor.

Finally, in a third phase, these methodologies have been used to develop pressures / impacts models and apply them on the North-West Limestone Tables Hydro-ecoregion (6000 km of river system) where the number of stations measuring EQR-IBGN (the difference between the reference and the Global Biological Index Standardized average over the 1994 to 2012 period) is sufficient (155 stations) to initiate a statistical analysis of the relationships between the biological responses at the stations and the pressures on the basin and riparian corridors. The implementation of these models shows the high value of VHRS maps and of indicators calculated on a spectrum of riparian footprints to improve identification and localization of sources of pressures in these spaces. With these methodological advances the role of local riparian vegetation, on a 20m large strip both sides of the river, on river ecological status has been demonstrated at regional level.

In parallel with the development of these methodologies, we focused on the propagation of uncertainties from the land cover map to the indicators, and from the indicators to the model results. However, difficulties in properly spatializing uncertainties in land cover maps remains a deadlock in completing this study that still has to be lifted.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

LISTE DES ABBREVIATIONS

ACP	Analyse en Composante Principale
AE	Agence de l'Eau
BDC	BDCarthage® (Bases de données IGN)
BDR	Route500® (Bases de données IGN)
BV	Bassin Versant
CLC	CORINE Land Cover
CR	Corridor Rivulaire
CRamont	Corridor Rivulaire Amont
CRlocal	Corridor Rivulaire Local
CSPNB	Conseil Scientifique du Patrimoine et de la Biodiversité
DCE	Directive Cadre Eau
DPSIR	Driving - Pressure - State - Impact - Response
EEA	Agence Européenne de l'Environnement
EQR	Ecological Quality Ratio (écart à la référence)
FNEA	Fractal Net Evolution Approach
HER	HydroEcoRégion
HER1	HydroEcoRégion de niveau 1
HER2	HydroEcoRégion de niveau 2
HER9-NO	Nord Ouest de l'HydroEcoRégion Tables Calcaires (zone de modélisation)
HRS	Haute Résolution Spatiale
IBGN	Indice Biologique Global Normalisé
IGN	Institut Géographique National
IRI	Infra Red Index
LHQ	Laboratoire Hydroécologie Quantitative
LIDAR	Light Detection And Ranging
MEEDDM	Ministère de l'Ecologie de l'Energie du Developpement Durable et de la Mer
MES	Matières En Suspension
MIR	Moyen Infra Rouge
MNT	Modèle Numérique de Terrain
MO	Matières Organiques
NDVI	Nomalized Difference Vegetation Index
ONEMA	Office Nationale de l'Eau et des Milieux Aquatiques
OS	Occupation du Sol
OSCR	Occupation du Sol dans le Corridor Rivulaire
PAN	Panchromatique (image SPOT 5)
PIR	Proche Infra Rouge
PLS	Partial Least Square
RPG	Registre Parcellaire Graphique
TETIS	Territoires, Environnement, Télédétection et Information Spatiale
THRS	Très Haute Résolution Spatiale
XS	Multi-spectrale (image SPOT 5)

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

LISTE DES VALORISATIONS

Articles

Article 1:
Tormos, T.,., Kosuth, P. Durrieu, S, Wasson, J.G. & Villeneuve, B. (SUBMITTED-2010) Object oriented approach for operational broad-scale mapping and quantification of land cover spatial indicators along river corridor. *International Journal for Remote Sensing of Environment*.

Article 2:
Tormos, T., Kosuth, P., Durrieu, S., Villeneuve, B. & Wasson, J.-G. (IN PRESS - 2010) Improving the quantification of land cover pressures on stream ecological status at riparian scale using High Spatial Resolution Imagery. *Physics and Chemistry of the Earth*.

Article 3:
Durrieu, S., **Tormos, T**., Kosuth, P. & Golden, C. (2008) Influence of training sampling protocol and of feature space optimization methods on supervised classification results. In: *International Geoscience and Remote Sensing Symposium (IGARSS)* pp. 2030-2033.

Article 4:
Kosuth, P. & **Tormos, T**. (2010) Riparian vegetation and river ecological status: contribution of high spatial resolution imagery. *World Committee on Space Research 2010 Report, CNES.* pp. 96-97

Communications et posters

Communication 1 :

Tormos, T., Kosuth, P., Durrieu, S., Wasson, J.-G., Pella, H. & Villeneuve, B. (2006a) Contribution of remote sensing for the spatial analysis of land use in river corridors to model relationship between land use pressures and ecological status of rivers. In: *Man and River System II* (Paris, 4-6 Décembre), pp. 249-252.

Communication 2 :

Kosuth, P., **Tormos, T**., Cernesson, F. & Lalande, N. (2010) Assessing the role of riparian vegetation and land use on river ecological status using remote sensing and spatial modelling. In: *SER-Ecological Restoration and Sustainable Development - Establishing Links Across Frontiers* (Avignon, 23 – 27 August 2010).

Communication 3 :

Kosuth, P., **Tormos, T**., Cernesson, F. & Lalande, N. (2010) Construyendo nuevos indicadores de estado ecológico a través de la teledetección. In : *Congreso sobre evaluacion del estado ecologico de los rios en el ambito del sudoeste europeo* (Mérida, 28 – 29 octobre 2010)

Poster 1 :

Tormos, T., Kosuth, P., Durrieu, S., Wasson, J.G., Pella, H., Villeneuve, B. & Perez Correa, M. (2006) Remote-sensing for the spatial analysis of land use in river corridors and its impact on the ecological status of rivers. Vienne, Austia, 2-7 April).

CemOA : archive ouverte d'Irstea / Cemagref

Poster 2 :

Tormos, T., Kosuth, P., Durrieu, S., Wasson, J.G., Pella, H., Villeneuve, B. & Perez Correa, M. (2006) Use of Multi-Platform-sensing for characterisation of land use in river corridors. Communication au congrès International Society for Photogrammetric and Remote Sensing. In: *ISPRS Commission 1* (Paris, 2006).

Poster 3 :
Tormos, T., Kosuth, P., Durrieu, S. & Wasson, J.G. (2006) Indicateurs spatialisés de l'impact de l'occupation du sol dans le corridor rivulaire sur l'état écologique des cours d'eau. In: *Journées IALE France 2006, 3ème journées de l'Association Française de l'Ecologie du Paysage* (Rennes, 10 novembre 2006).

Poster 4 :
Tormos, T., Durrieu, S., Kosuth, P. & Wasson, J.G. (2008) An operational methodology for characterizing land use along river corridor from remote sensing imagery at large scale. In: *EGU Remote sensing of rivers,* (Vienne, Austria, 13-18 April 2008).

Rapports techniques

Rapport technique 1 :
Tormos, T., De Bortoli, J., Kosuth, P., Holley, J.-F. & Argillier, C. (2007) Analyse d'images et prospection terrain pour caractériser l'hydromorphologie lacustre. *Rapport convention MEDD/Cemagref.* Cemagref groupement Montpellier, UMR TETIS & Unité hydrobiologie écosystèmes lacustres, p. 53.

Rapport technique 2 :
Tormos, T., Kosuth, P., Durrieu, S., Wasson, J.G., Villeneuve, B. & Chandesris, A. (2009) Rapport méthodologique sur l'évaluation par télédétection de l'état des corridors rivulaires à large échelle. *Rapport Convention ONEMA/Cemagref.* UMR TETIS Cemagref, p. 60

Articles de vulgarisation

Article de vulgarisation 1 :
Kosuth, P., Tormos, T. (2010) L'espace au service de la Directive Cadre sur l'Eau. *CNES MAG Juillet 2010 n°46 « Espace et Biodiversité »* pp 54-55

CemOA : archive ouverte d'Irstea / Cemagref

TABLE DES MATIÈRES GÉNÉRALE

Remerciements *5*
Résumé *7*
Résumé en anglais (abstract) *9*
Liste des abbréviations *11*
Liste des valorisations *13*

INTRODUCTION GENERALE 17

PREMIERE PARTIE : CONTEXTE DE L'ETUDE ET ÉTATS DES CONNAISSANCES 29

Introduction de la première partie 31
CHAPITRE 1 : ÉVALUATION DE L'ETAT ECOLOGIQUE DES COURS D'EAU FRANÇAIS AU SENS DE LA DCE *33*
CHAPITRE 2 : LE ROLE DES FORMATIONS VEGETALES RIVULAIRES *55*
CHAPITRE 3 : MODELISATION DES RELATIONS PRESSIONS/IMPACTS *73*
CHAPITRE 4 : OBJECTIFS, DEMARCHE ET ZONES D'ETUDES *105*

DEUXIEME PARTIE : EXTRACTION DE L'OCCUPATION DU SOL LE LONG DES COURS D'EAU PAR TELEDETECTION 119

Introduction de la deuxième partie 121
CHAPITRE 5 : CONCEPTION DE LA PROCEDURE DE CLASSIFICATION *123*
CHAPITRE 6 : APPLICATION SUR LE BASSIN VERSANT DE L'HERAULT *180*
Synthèse de la deuxième partie 226

TROISIEME PARTIE : LES INDICATEURS SPATIALISES DU CORRIDOR RIVULAIRE 228

Introduction de la troisième partie 230
CHAPITRE 7 : CONCEPTION DES INDICATEURS SPATIALISES DU CORRIDOR RIVULAIRE *232*
CHAPITRE 8 : ANALYSE ET MISE EN ŒUVRE DES INDICATEURS SUR LE BASSIN VERSANT DE L'HERAULT *258*
Synthèse de la troisième partie 288

QUATRIEME PARTIE : MODELISATION PRESSIONS/IMPACTS REGIONALISEE -PRISE EN COMPTE DES CORRIDORS RIVULAIRES 290

Introduction de la quatrième partie 292
CHAPITRE 9 : MODELES PRESSIONS/IMPACTS : MATERIELS, METHODE ET DEMARCHE *294*
CHAPITRE 10 : MODELES PRESSIONS/IMPACTS : APPLICATION SUR LA ZONE D'ETUDE (HER9-NO) *328*
Synthèse de la quatrième partie 384

CONCLUSION & PERSPECTIVE 386

Liste des figures *400*
Liste des tableaux *406*
Bibliographie *408*
Liste des annexes *426*
Document ANNEXE *427*

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

INTRODUCTION GÉNÉRALE

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Les cours d'eau : des écosystèmes originaux, essentiels et particulièrement menacés

Les fleuves et les rivières s'individualisent comme des ***écosystèmes originaux*** et complexes au fonctionnement tout à fait particulier. La notion classique d'écosystème bien identifié spatialement s'applique mal aux cours d'eau. Ils ne se limitent pas, en effet, à un simple chenal d'écoulement étiré dans l'espace, mais forment un tout indissociable avec le milieu terrestre adjacent. C'est au cours des années 70 que les réflexions scientifiques abordent les rivières comme des systèmes fluviaux dynamiques pour décrire leur fonctionnement en interaction avec leur plaine d'inondation et l'ensemble de leur bassin versant (e.g., Hynes, 1975; Schumm, 1977; Vannote *et al.*, 1980; Amoros and Bravard, 1985).

Dés lors de nouveaux concepts émergent et permettent de démontrer que l'hétérogénéité spatiale et la variabilité temporelle sont des composantes essentielles du fonctionnement des cours d'eau (Amoros and Bravard, 1985). A la vision linéaire d'une rivière, dont le peuplement serait régi essentiellement par la qualité de son eau, s'est progressivement substituée celle d'un paysage aquatique : l'hydrosystème. Défini comme un écocomplexe[1], il est constitué par l'imbrication d'une mosaïque d'écosystèmes d'eau courante (chenal principal), d'eau stagnante (bras morts, plans d'eau), semi-aquatiques et terrestres (marais, forêt alluviales), aussi bien superficiels que souterrains, en interaction dans la plaine alluviale. Cet ensemble interactif de milieux naturels ou anthropiques est caractérisé par des échanges multidirectionnels de matières et d'énergie conceptualisés par (Amoros *et al.*, 1987 ; Ward, 1989) dans une structure à quatre dimensions (Figure 0.1) une dimension longitudinale (amont-aval), transversale (incluant le lit majeur), verticale (sédiments et sous-écoulements), et temporelle (cycles saisonniers et évènements rares).

Le fonctionnement de l'hydrosystème ainsi défini dépend directement ou indirectement du cours actif (flux d'eau) de la rivière qui érode naturellement son lit et ses berges, transporte des sédiments et inonde périodiquement les terrains riverains. Ces processus naturels, fortement dépendant de l'hydrologie, de la géomorphologie, et de la végétation rivulaire, entretiennent les connectivités, créent une mosaïque d'habitats favorables à la faune, et maintiennent ainsi une variabilité spatiale et temporelle capitale pour préserver la diversité du milieu aquatique et l'état du système (Maridet, 1995).

La variété très importante d'habitats offerte par les écosystèmes aquatiques permet donc l'expression d'une grande diversité d'espèces (diversité spécifique). Occupant approximativement 0,8 % de la surface terrestre (Gleick, 1996), ils abritent pas moins de 6 % des 1,8 millions espèces décrites dans la biosphère (Strayer, 2006). Au delà de ce rôle de réservoirs de biodiversité d'intérêt, les systèmes fluviaux sont, par leur caractère dynamique, un vecteur de transfert majeur dans le fonctionnement du cycle hydrologique qui contribue à la régulation des ressources en eau indispensables à la vie et au fonctionnement global de la planète terre. Les cours d'eau constituent donc des ***écosystèmes essentiels*** pour la nature comme pour l'homme.

D'un point de vue historique, cette dépendance de l'homme au fleuve est illustrée par plusieurs grandes civilisations (Gleick, 2008) : le Nil pour l'Egypte, le Tigre et l'Euphrate pour la Mésopotamie. Les cours d'eau sont indispensables à nos cultures : l'eau qui coule apporte des atouts, génère des enjeux et des conflits, éveille et reflète des valeurs culturelles symboliques. Les savoir-faire, les pratiques, les sentiments d'appartenance, les identités territoriales se cristallisent autour des usages de ces paysages. Aujourd'hui, ils sont multiples et apportent de nombreux services à la société, liés notamment à l'approvisionnement en eau potable (boissons, aliments), aux activités domestiques, agricoles et industrielles, à la pêche, à la production d'énergie, à l'évacuation des rejets, aux transports, aux loisirs, au tourisme, à l'épanouissement spirituel, à la santé (…) ; la liste de ces usages est longue et l'eau demeure irremplaçable pour certains d'entre eux.

Au XIX^{e} et XX^{e} siècles, la révolution industrielle, l'explosion démographique et les évolutions de l'agriculture ont accru la demande en eau : des dizaines de milliers d'ouvrages monumentaux ont été érigés pour stocker l'eau, contrôler les inondations, améliorer l'irrigation et produire de l'énergie, ou encore préserver les réserves d'eau potable. Plus de la moitié des cours d'eau accessibles sont aujourd'hui appropriés par l'homme (Allan and Castillo, 2007). L'anthropisation des terres (routes, voies de chemin de fer, digues, barrages, urbanisation, agriculture…) tend à rompre la dynamique naturelle et le degré de connectivité des écosystèmes aquatiques. Elle engendre des perturbations au niveau du fonctionnement global des écosystèmes d'eau courante via des mécanismes nombreux et complexes qui interagissent (Allan, 2004a).

[1] L'écocomplexe est défini comme un ensemble d'écosystèmes interactifs, et non pas seulement juxtaposés en des mosaïques plus ou moins hétérogènes (références).

CemOA : archive ouverte d'Irstea / Cemagref

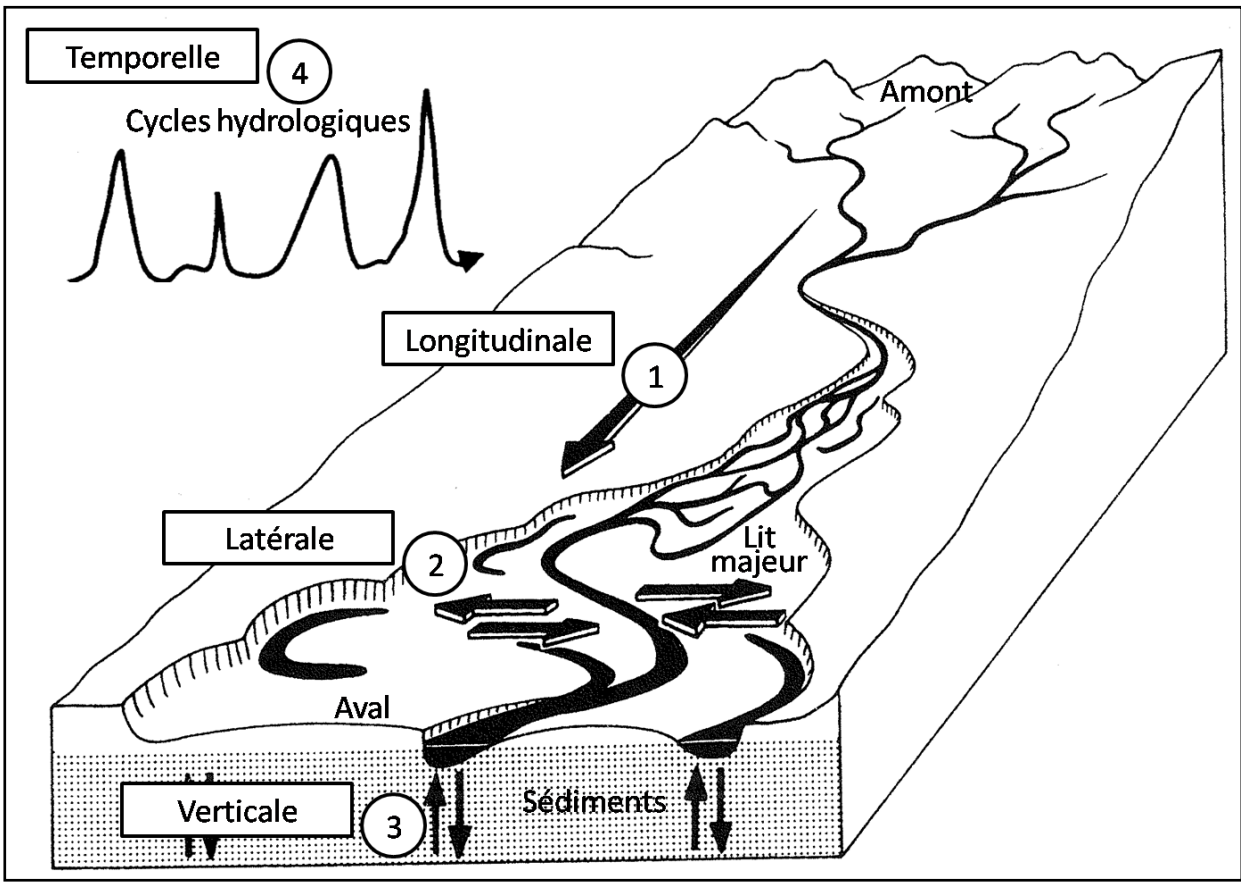


Figure 0.1 : Schématisation des flux selon les quatre dimensions de l'hydrosystème (d'après Amoros and Petts, 1993) :

① La dimension longitudinale. Les rivières, les fleuves constituent des systèmes étirés dans l'espace, mais solidaires dans leur dimension longitudinale, dans la mesure où la dynamique des systèmes situés en aval dépend des processus physiques et chimiques qui se déroulent en amont. Cet axe longitudinal traduit l'évolution des principales caractéristiques hydrauliques, géomorphologiques et écologiques des cours d'eau suivant un gradient plus ou moins continu de la source à l'embouchure.

② La dimension verticale. Il existe des relations étroites entre les eaux de surface et les eaux souterraines qui se manifestent, en particulier, par des échanges d'eau et d'éléments dissous : sels minéraux, nitrates, polluants... Les nappes souterraines assurent le maintien du débit en période d'étiage, alors qu'en période de crue les eaux de surface rechargent les nappes au niveau des zones d'inondations.

③ La dimension transversale. Le fleuve fonctionne avec sa plaine alluviale constituée d'une mosaïque d'écosystèmes. On a mis en évidence le rôle des plaines inondables et des annexes fluviales, tels les bras morts, dans le fonctionnement des eaux courantes, ainsi que le rôle des boisements riverains sur le recyclage, la rétention et le relargage des substances nutritives. Les zones riveraines fournissent aux eaux courantes l'essentiel de leur alimentation en matière organique sous forme de détritus. Pour l'ensemble du bassin versant, l'utilisation des terres (boisements, cultures, urbanisation) a des conséquences sur le ruissellement et le fonctionnement hydrologique ainsi que sur les apports en substances particulaires et dissoutes.

④ La dimension temporelle. Au-delà des cycles annuels, les systèmes aquatiques actuels sont un héritage tant sur le plan géomorphologique que sur le plan écologique, des changements climatiques et tectoniques qui ont eu lieu dans le passé. Pour comprendre leur structure et leur fonctionnement, il est nécessaire de connaître cette histoire ainsi que les conditions de mise en place des flores et des faunes. Dans un passé plus récent, divers types d'aménagements (barrages, endiguements) ou d'activités (prélèvements de granulats) ont souvent modifié le lit des rivières. A chaque échelle de temps se déroulent des phénomènes qui ont une influence sur la dynamique actuelle.

CemOA : archive ouverte d'Irstea / Cemagref

Les atteintes physiques telles que les actions de rupture hydraulique, de chenalisation ou les interventions massives sur le bassin (de type coupes forestières à blanc), qui entraînent une forte érosion suivie d'une sédimentation excessive (Niemi *et al.*, 1990), sont le plus souvent irréversibles, avec des temps de rémanence très longs (Brooks, 1998; Niemi *et al.*, 1990 et Wasson *et al.*, 1998). Les cours d'eau sont donc des ***écosystèmes particulièrement menacés.*** D'une part, leur rôle de collecteur d'un bassin versant fait que toute altération des écosystèmes terrestres sur ce bassin a des répercussions potentielles, à grande distance et à long terme, sur les flux d'eau et d'éléments qui alimentent les rivières, les lacs et à l'aval les estuaires et le milieu marin. D'autre part, la structure linéaire des hydrosystèmes les rend particulièrement vulnérables aux structures transversales (e.g., barrages, écluses) qui interrompent les flux, comme aux contraintes longitudinales trop sévères (e.g., chenaux, digues) (Wasson, 1992).

De ce fait, une grande partie de l'intégrité biologique des cours d'eau a été perdue il y a plusieurs siècles en Europe et en Asie, il y a un siècle en Amérique du Nord (Harmon *et al.*, 1986) et la perte de la biodiversité des écosystèmes aquatiques est plus forte que celle des écosystèmes terrestres (Sala *et al.*, 2000, Dudgeon *et al.*, 2006). Actuellement, la détérioration de ces écosystèmes est telle qu'il existe des incompatibilités entre les différents usages comme par exemple (Wasson, 1992): évacuation des rejets *vs* eau potable, loisirs ; hydroélectricité *vs* halieutisme ; hydraulique agricole *vs* ressource en eau multi-usages ; occupation du lit majeur *vs* épandage des crues ; canal à grand gabarit vs paysage et tourisme. Maintenir l'intégrité du fonctionnement des rivières est un enjeu majeur de notre société afin de garantir pour tous l'agrément d'un cadre de vie, la pérennité et la qualité de la ressource et des services et usages qui y sont associés.

La directive cadre européenne sur l'eau : une politique forte et volontariste soulevant de nouveaux défis pour les scientifiques

L'évolution de la sensibilité aux problèmes environnementaux dans les sociétés industrialisées a entraîné, au cours des deux dernières décennies, l'apparition de nouveaux principes dans le droit de l'environnement (e.g., principes du développement durable, de précaution, de prévention, du pollueur-payeur). Elle indique clairement que le sacrifice de la nature et du cadre de vie pour des raisons économiques à court ou moyen terme est maintenant ressenti comme une erreur historique. Le maintien de la quantité et de qualité de la ressource, des potentialités évolutives des systèmes, de la diversité génétique, sont des impératifs pour préserver le cadre de vie et les possibilités des générations à venir. L'eau et les cours d'eau sont clairement au cœur de ces préoccupations.

La *gestion écologique intégrée*[2] des rivières, dont la finalité peut se résumer en une phrase : « conserver ou réintégrer l'activité humaine dans les équilibres naturels », est devenue un objectif général commun pour les pays de l'Union européenne par le biais de la Directive Cadre sur l'Eau (DCE). Cette directive, entrée en vigueur le 22 décembre 2000 sous l'intitulé « directive 2000/60/CE » du parlement européen et du conseil (European commission, 2000), dépasse les simples préoccupations prioritaires telles que l'alimentation en eau potable ou la protection contre les inondations. Cette directive est une ***politique forte et volontariste***, qui vise à tirer vers le haut la qualité des écosystèmes aquatiques. Elle affirme clairement une volonté de protection et de gestion de l'ensemble des milieux aquatiques (les eaux intérieures de surface, de transition, côtières et souterraines) en vue de prévenir et de réduire leur pollution, protéger leur environnement, améliorer l'état des écosystèmes aquatiques, atténuer les effets des inondations et des sécheresses, ceci afin de garantir de façon durable leur utilisation pour les générations futures.

Pour cela, la DCE a fixé comme objectif à tous les Etats membres d'atteindre à l'horizon 2015 le « bon état écologique » pour tous les milieux aquatiques naturels, de préserver ceux qui sont en « très bon état », et d'atteindre le « bon potentiel » dans les milieux fortement artificialisés. Il s'y ajoute également un objectif « zéro toxiques », non daté mais clairement affiché (Wasson, 2001). L'état se définit d'après la situation la plus déclassante entre un état chimique se rapportant à des normes de concentration de certaines substances particulièrement dangereuses (toxiques), et un état écologique qui repose sur l'évaluation de « la structure et du fonctionnement des écosystèmes aquatiques ».

Cet objectif ambitieux, compte tenu de l'obligation de résultats dans le délai imparti, a ouvert un chantier important et complexe pour les gestionnaires et est à l'origine ***de nouveaux défis pour les scientifiques*** (Roche

[2] Formulation proposée par M. Coulet lors des 3° Entretiens Jacques Cartier, Arles, 1990. La notion de gestion écologique intégrée suppose une prise en compte des écosystèmes alors que le terme gestion intégrée s'applique généralement aux seuls usages de la ressource en eau (Wasson *et al.*, 1993).

CemOA : archive ouverte d'Irstea / Cemagref

et al., 2005). Sa mise en œuvre implique clairement deux phases (Figure 0.2) : d'abord une phase de diagnostic général de l'état des lieux, qui s'appuie sur la mise en place d'un réseau de suivi, puis à moyen terme une phase de restauration des milieux n'atteignant pas le « bon état ». Ces deux phases sont évidemment liées, car le diagnostic doit orienter les actions de restauration. Elles se traduisent par des besoins précis en termes d'outils opérationnels : d'une part des méthodes d'évaluation de l'état écoloqiue (typologie, bioindicateurs, réseaux de suivi…), et d'autre part des modèles reliant les multiples pressions aux impacts sur l'état des écosystèmes pour définir et choisir les actions prioritaires de restauration. Le développement de ces outils se fait sous la responsabilité des états membres, la directive fixant seulement les objectifs et un catalogue des mesures possibles.

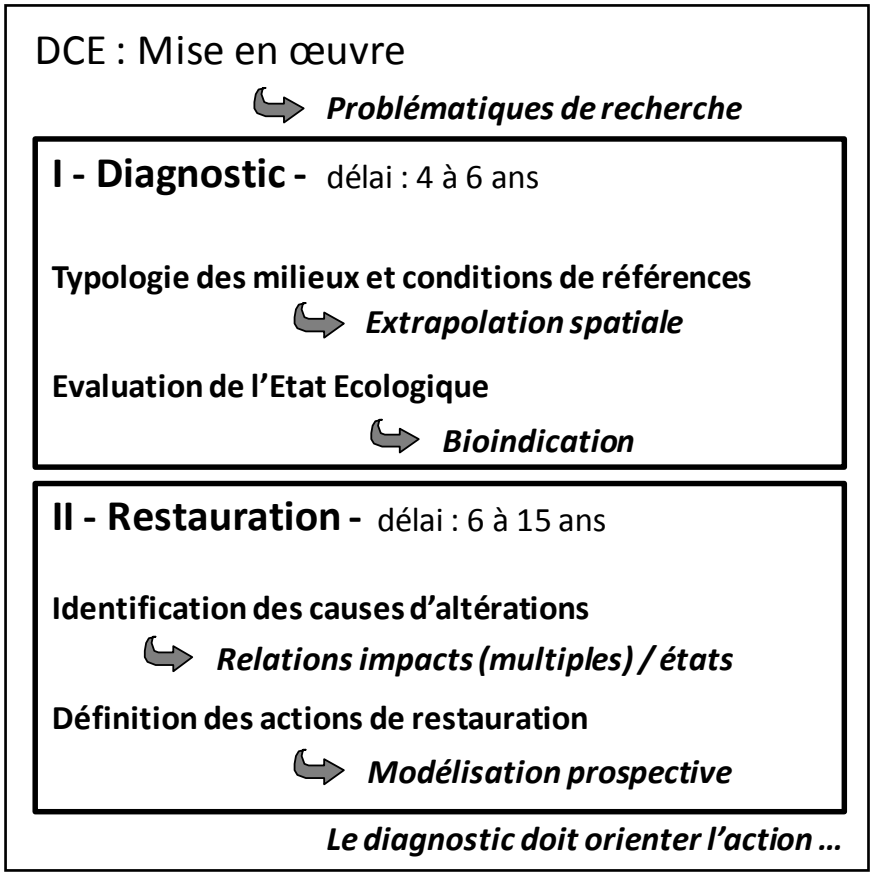


Figure 0.2 : Les principales problématiques de recherches liées à la mise en œuvre DCE à partir de 2000 (Wasson, 2001)

Les corridors rivulaires : espace clé de la restauration

Les corridors rivulaires sont des écosystèmes situés le long des réseaux hydrographiques à l'interface entre les milieux aquatique et terrestre. On entend par « corridor rivulaire » une bande de terrain le long d'un cours d'eau (un espace), dont la végétation naturelle ou « ripisylve » est influencée par la rivière, et interagit avec elle. Sa largeur est variable, elle augmente vers l'aval au fur et à mesure que les débits augmentent et que le chenal et le fond de vallée s'élargissent. De par leurs fonctionnalités d'écotone (positionnement d'interface) et de connexion (structure linéaire), ces corridors sont de véritables carrefours biologiques sur lesquels s'articulent des chaînes alimentaires d'invertébrés, de poissons, de mammifères et d'oiseaux (Piegay *et al.*, 2003). Pour les spécialistes des rivières, ces corridors sont perçus comme un élément clé de la DCE, car ils contribuent de manière très significative à la préservation de la qualité et de la biodiversité aquatique (Naiman *et al.*, 2005).

CemOA : archive ouverte d'Irstea / Cemagref

De par leur influence sur les processus fonctionnels (physiques, biogéochimiques et écologiques), les formations végétales rivulaires constituent un compartiment essentiel des écosystèmes d'eau courante (Figure 0.3). Elles participent ainsi au maintien de la dynamique physique et biologique nécessaire pour la résilience du système face aux perturbations d'origines naturelle ou anthropique. Au niveau des *processus physiques*, la ripisylve est un facteur de contrôle de la dynamique fluviale (moteur du fonctionnement de l'hydrosystème) en ralentissant les crues et en limitant les vitesses de courant dans les zones inondées (Lâchât, 1991). Elle diversifie aussi les habitats aquatiques en modifiant sensiblement la morphologie du lit, en stabilisant les berges (Maridet and Souchon, 1995) et générant un approfondissement du chenal favorable aux poissons. Les arbres tombés dans l'eau (bois-morts) amplifient aussi ce rôle en servant d'abri pour les poissons et en générant des micro-habitats favorables aux invertébrés (Harmon *et al.*, 1986). Au niveau des *processus biogéochimiques*, les bandes végétales rivulaires sont des barrières naturelles vis-à-vis des apports latéraux, par ruissellement et par écoulements souterrains, de polluants diffus qui proviennent des versants aussi bien agricoles (e.g., sédiments fins, nitrates, phosphore, pesticides) (Peterjohn and Correll, 1984) qu'urbains (e.g., écoulement de substances chimiques issus des routes et des parkings) (Paul and Meyer, 2001). Au niveau des *processus écologiques*, la ripisylve est aussi un régulateur thermique (ombrage) et trophique (source de nourritures), facteurs-clés du milieu aquatique (Maridet, 1994).

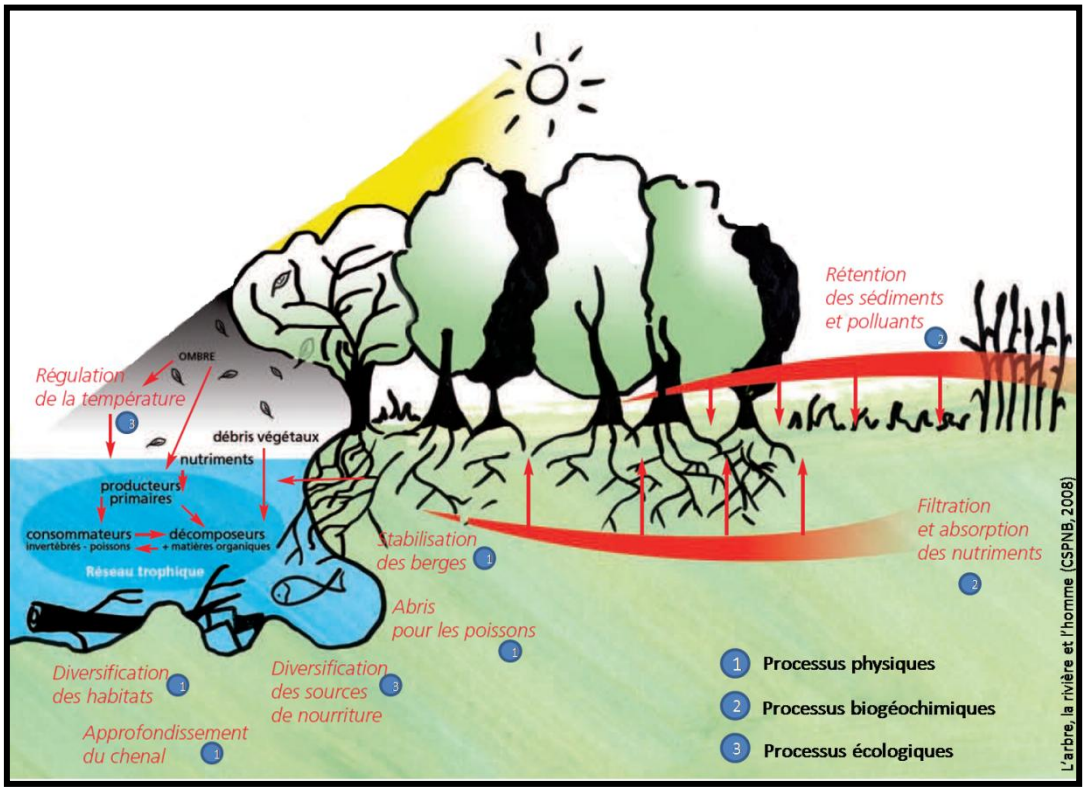


Figure 0.3 : Influence des corridors rivulaires sur les processus fonctionnels de l'écosystème d'eau courante (physiques, biogéochimiques et écologiques) (d'après Maridet and Collin Huet, 1995).

Depuis les années 80, de nombreuses études (Shearer and Xiang, 2007), le plus souvent à l'échelle locale (i.e., la station d'étude), mettent en avant les multiples fonctions de la végétation rivulaire et notamment le rôle de « filtre » vis à vis des pollutions diffuses environnantes qui a popularisé l'utilité de ces écosystèmes (Décamps, 2002).

Les connaissances écologiques, en termes de fonctionnement, sont donc aujourd'hui suffisantes pour affirmer l'importance de ces formations végétales sur l'état écologique des cours d'eau. Avec un foncier plus facilement négociable qu'à l'échelle d'un bassin, les corridors rivulaires sont dès lors vus par de nombreux chercheurs et gestionnaires comme l'***espace clé de la restauration*** du bon état écologique. Ils ont par exemple mené à l'adoption de mesures agri-environnementales comme les bandes enherbées (CSPNB, 2008). Ils sont aujourd'hui intégrés, en France, sous le label de « trame verte et bleue » dans la loi actuellement en cours de préparation suite au Grenelle de l'Environnement.

CemOA : archive ouverte d'Irstea / Cemagref

Pour orienter de façon efficace les gestionnaires dans les stratégies de restauration de ces espaces, il est indispensable de prendre en compte l'ensemble des facteurs qui régissent les processus fonctionnels des écosystèmes d'eau courante, processus en œuvre à des échelles multiples allant du microhabitat jusqu'au bassin versant (Frissell *et al.*, 1986, Allan, 2004b). Cependant, la plupart des connaissances sur ces processus, comme indiqué plus haut, résultent d'études à des échelles locales, le plus souvent au niveau de la station d'étude (quelques centaines de mètres de rivière et son environnement immédiat). Les recherches se sont focalisées le plus souvent sur des sites locaux jugés représentatifs de l'ensemble d'un système (Allan, 2004a). Les échelles plus larges, notamment celles du tronçon de rivière et du bassin versant, ont été moins considérées par la littérature scientifique, alors qu'elles constituent des unités de gestion cohérentes pour les gestionnaires (Wasson, 1992).

Dès lors, *mettre en évidence, au niveau régional, voire national, le rôle des corridors rivulaires sur l'état écologique des cours d'eau constitue un enjeu de recherche majeur pour appuyer les gestionnaires dans la définition des stratégies de restauration de ces espaces*. Cette mise en évidence doit prendre en compte l'ensemble des facteurs naturels et anthropiques agissant sur les processus de fonctionnement de l'écosystème rivière aux différentes échelles. Elle doit particulièrement identifier les caractéristiques des ripisylves (largeur de végétation, type de couvert végétal, …) ayant une influence significative sur l'état écologique des cours d'eau. Une telle recherche implique le développement de modèles étudiant à large échelle (i.e. sur de grands territoires) l'impact de ces différents facteurs sur l'état écologique des cours d'eau (i.e., des modèles pressions/impacts).

Les modèles pressions/impacts : un outil spatialisé de diagnoctic de l'état écologique à largé échelle pour orienter les gestionnaires dans la définition des stratégies de restauration et/ou de préservation

Actuellement, la compréhension des nombreux mécanismes complexes et multiscalaires, à travers lesquels les activités humaines et les formations végétales rivulaires influencent l'état écologique des cours d'eau, est extrêmement limitée (Allan, 2004a). Plusieurs approches intéressantes de modélisation ont émergé, mais aucune n'est pleinement satisfaisante pour guider de façon efficace les gestionnaires qui doivent prendre des décisions sur de grands territoires (Johnson *et al.*, 2007).

Des modèles mécanistes ont été développés pour comprendre et prédire l'influence de la végétation rivulaire sur les flux de matières vers les cours d'eau (e.g., Lowrance *et al.*, 1983; Peterjohn and Correll, 1984). Le « Riparian Ecosystem Management Model » (REMM ; Lowrance *et al.*, 2000; Altier *et al.*, 2002) en est un bon exemple. Il peut être utilisé pour prédire les effets des zones rivulaires sur la chimie des cours d'eau lorsque les apports de sédiments et de nutriments sont connus. Cependant comme les autres modèles mécanistes, l'application du modèle REMM est limitée par sa demande très importante en variables d'entrée et par la complexité de son paramétrage, notamment lorsque l'on change d'échelle (Inamdar, 2006). Dès que les recherches sur ces modèles se sont orientées vers des hydrosystèmes plus vastes, plus de 10 000 km² dans le cas de la Garonne, près de 190 000 km² pour le bassin versant du Rhin, le problème s'est considérablement compliqué (Ruffinoni *et al.*, 2003). Cette approche est donc limitée dans la vision à large échelle des problèmes mais également dans la prédiction des effets sur l'état écologique des cours d'eau (i.e sur la structure et le fonctionnement des communautés biologiques) mesurés au niveau des stations.

Les modèles statistiques offrent une approche alternative intéressante. Comme Wiens, 2002 l'a observé, ces différents modèles sont basés sur les concepts et les outils de l'écologie du paysage. Dans cette approche, les structures spatiales, qui composent le paysage aquatique (l'hydrosystème), dépendent des relations entre les activités humaines et le milieu environnant. Elles sont en interaction constante avec les processus fonctionnels des écosystèmes d'eau courante (Figure 0.4). L'analyse de la composition, de l'organisation et de la connectivité de ces structures à travers différentes échelles spatiales permet de rechercher la gamme d'échelles pour laquelle un ou plusieurs facteurs donnés (naturel ou anthropique) sont pertinents pour expliquer un processus fonctionnel (Burel and Baudry, 1999). L'émergence des Systèmes d'Information Géographique (SIG) et des technologies de traitement d'images (aéroportées ou satellitaires), qui permettent de caractériser et de quantifier les structures spatiales (l'occupation du sol) à différente échelles dans le paysage, a largement contribué au développement de cette approche.

CemOA : archive ouverte d'Irstea / Cemagref

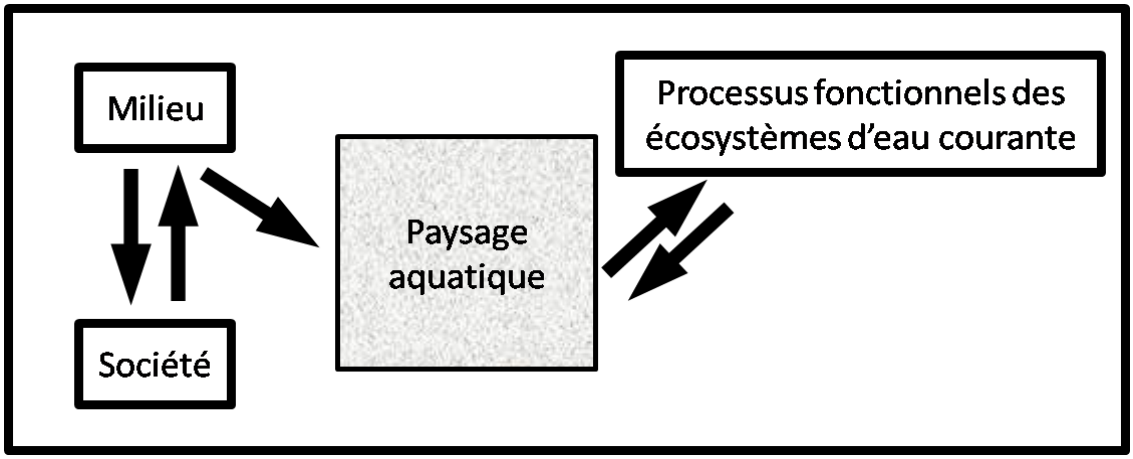


Figure 0.4 : Démarche scientifique dérivée de l'écologie du paysage adoptée dans les modèles pressions / impacts : le paysage aquatique est la résultante de la dynamique du milieu et de la société qui s'y est développée. La structure, l'organisation, la dynamique du paysage aquatique (l'hydrosystème) sont en interaction constante avec les processus fonctionnels qui s'y déroulent (modifiée de Burel and Baudry, 1999).

Ainsi, de nombreux modèles, baptisés modèles pressions/impacts en France (Wasson *et al.*, 2005), étudiant les relations statistiques entre des indicateurs d'occupation du sol (rendant compte de la plupart des pressions s'exerçant sur le milieu aquatique) et les indicateurs de l'état écologique des cours d'eau (mesurés au niveau des stations), à travers différentes échelles spatiales, ont été développés pour mieux comprendre les mécanismes d'impacts (e.g., Allan *et al.*, 1997 ; Sponseller *et al.*, 2001 ; Stewart *et al.*, 2001 ; Frimpong *et al.*, 2005a). Ils constituent donc un ***outil spatialisé de diagnostic de l'état écologique à large échelle*** qui identifie et hiérarchise les facteurs limitants (i.e. dégradant) et protecteurs (i.e. améliorant) de l'état écologique au niveau du tronçon de rivière jusqu'au bassin versant. Cet outil, qui élabore des réponses spatialisées, permet ***d'orienter les gestionnaires dans la définition de stratégies de restauration et/ou de préservation*** efficaces et adaptées sur de grands territoires.

L'imagerie à très haute résolution spatiale et les outils SIG : des potentialités à explorer pour améliorer la prise en compte des corridors rivulaires dans les modèles pressions/impacts

Cependant, en l'état actuel des travaux, les résultats de ces modèles se contredisent en ce qui concerne l'influence des facteurs au niveau des corridors rivulaires (Gergel *et al.*, 2002 ; Allan, 2004a) et sont insuffisants pour guider les gestionnaires dans les stratégies de restauration de ces espaces. L'impact respectif et relatif des pressions s'exerçant à l'échelle du bassin versant ou du corridor rivulaire n'est pas clairement établi ; certains travaux mettent en évidence un effet prépondérant de l'occupation du sol des bassins versants (Roth *et al.*, 1996, Wang *et al.*, 1997), d'autres aboutissent à la conclusion inverse et attribuent l'effet prépondérant à l'occupation des sols des corridors rivulaires (Lammert and Allan, 1999, Meador and Goldstein, 2003, Sponseller *et al.*, 2001).

Deux principales difficultés compliquent le problème. La première réside dans le manque d'information adaptée à cette échelle (Frimpong *et al.*, 2005a, Stewart *et al.*, 2001). Dans la plupart de ces études, la donnée d'occupation du sol utilisée est issue de l'imagerie satellitaire Landsat Thematic Mapper, à 30 m de résolution spatiale, très employée pour inventorier l'occupation du sol sur de grands territoires du fait de sa disponibilité et de son faible coût (Congalton *et al.*, 2002). Or, dans la plupart des cas, les structures linéaires et étroites des écosystèmes riverains sont trop petites pour être détectées par des satellites de moyenne ou haute résolution (Müller, 1997). En conséquence, la donnée géographique dérivée de cette imagerie satellitaire est en limite d'utilisation pour la simple raison qu'on ne peut pas distinguer de façon précise la composition et la configuration des structures rivulaires dont dépendent les processus qui régissent le fonctionnement de ces systèmes.

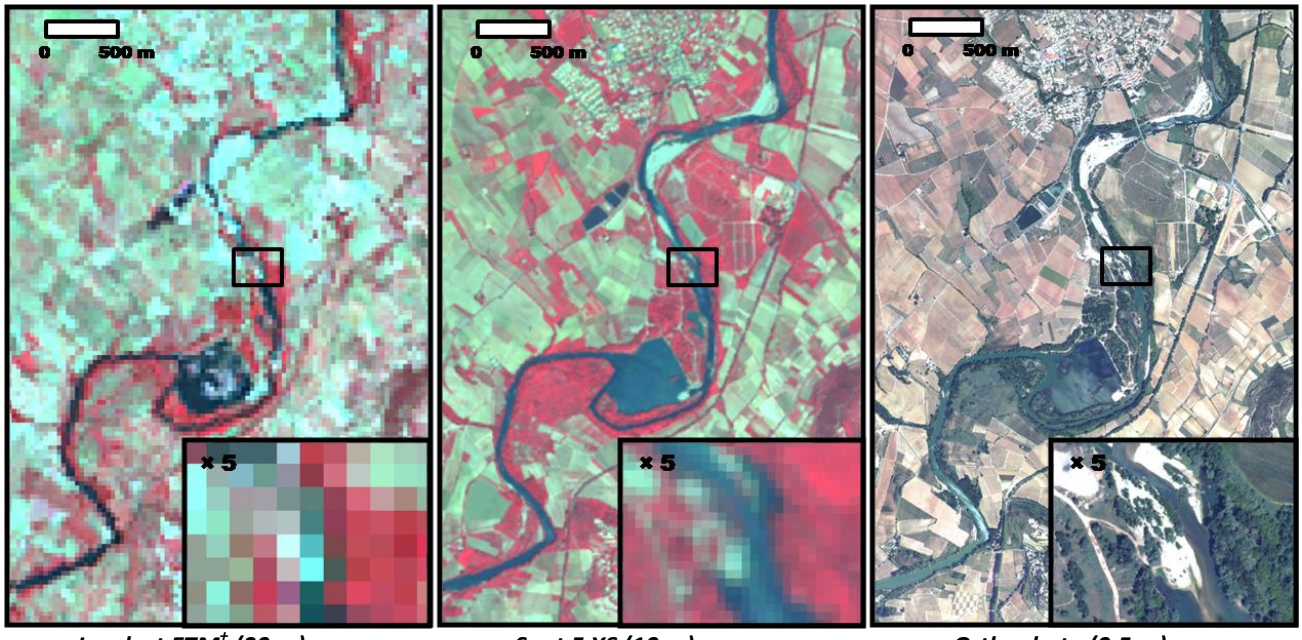
CemOA : archive ouverte d'Irstea / Cemagref

La deuxième est liée au manque de spatialisation des différentes sources de pression[3] (Frimpong *et al.*, 2005b ; Van Sickle and Johnson, 2008). Dans la plupart de ces études, une seule emprise spatiale candidate au niveau du corridor rivulaire (quelquefois deux ou trois) est utilisée pour étudier l'influence de l'occupation du sol dans ces espaces sur l'état écologique. Ceci limite d'une part la compréhension des relations pressions/impacts puisque l'influence d'une catégorie d'occupation du sol donnée est plus ou moins perceptible sur l'état écologique en fonction de l'emprise spatiale qu'elle occupe dans le corridor (i.e., en fonction de sa distance au cours d'eau et à la station écologique), et d'autre part l'information précise sur la localisation des différentes sources de pressions dans ces espaces.

Dès lors, *développer (i) une méthode opérationnelle (efficace et à coût réduit) de caractérisation fine de l'occupation des sols le long des cours d'eau, applicable (reproductible et portable) immédiatement sur l'ensemble des rivières ; et (ii) des techniques permettant d'améliorer la spatialisation des sources de pression s'exerçant sur une station écologique par la définition de jeux d'indicateurs spatialisés adaptés, sont des enjeux de recherche majeurs, pour une meilleure prise en compte des mécanismes influençant l'état écologique des cours d'eau au niveau de ces corridors.*

Aujourd'hui, grâce à l'augmentation du nombre de satellites en opération et aux avancées technologiques, l'imagerie satellitaire (et aéroportée) ne cesse de devenir plus performante en résolution spatiale comme en résolution spectrale. La télédétection à haute (5-20 m) et très haute résolution spatiale (< 5 m) fournit une information plus riche (Figure 0.5) appropriée à la caractérisation des systèmes rivulaires (Müller, 1997).

Dans le même temps, les progrès technologiques sur les Systèmes d'Information Géographique (SIG), offrent aujourd'hui aux scientifiques des outils performants pour mener des analyses spatiales entre différents objets géographiques de natures différentes (e.g, réseau hydrographique, stations de mesures, occupation du sol) et quantifier la composition et/ou la configuration spatiale de l'occupation du sol à l'origine des pressions s'exerçant sur les milieux aquatiques.



Landsat ETM+ (30 m) ***Spot 5 XS (10 m)*** ***Orthophoto (0.5 m)***

Figure 0.5 : Comparaison d'images à moyenne résolution, Landsat ETM+ (30 m), à haute résolution, Spot 5 XS (10 m) et à très haute résolution spatiale, orthophoto IGN (0.5 m) sur une portion du fleuve Hérault au niveau de sa plaine. L'identification des structures rivulaires est nettement plus visible sur les images à haute et très haute résolution spatiale qui fournissent une information plus riche et plus détaillée.

[3] Le terme « source de pression » fait référence à une catégorie d'occupation du sol. Une catégorie d'occupation du sol donnée est susceptible de générer diverses pressions qui peuvent altérer un ou plusieurs processus de fonctionnement de l'écosystème rivière.

CemOA : archive ouverte d'Irstea / Cemagref

Par conséquent, concevoir une méthode de classification semi-automatique à partir de l'imagerie à très haute résolution spatiale afin d'obtenir une cartographie fine et détaillée de l'occupation du sol dans les corridors rivulaires et développer des méthodes automatiques à partir des outils des SIG afin de mieux spatialiser les sources de pression dans ces espaces et construire des indicateurs spatialisés pertinents vis-à-vis des mécanismes d'influence de l'occupation du sol sont ***des potentialités à explorer pour améliorer la prise en compte des corridors rivulaires dans les modèles pressions/impacts***.

Objectifs et structure de la thèse

Cette thèse s'inscrit dans le cadre des recherches menées en amont de la mise en œuvre de la directive européenne sur l'eau. Son objectif est d'améliorer l'étude des relations pressions/impacts au niveau des corridors rivulaires afin d'orienter plus efficacement les gestionnaires dans la définition des stratégies de restauration et/ou de préservation dans ces espaces en déterminant notamment si le rôle que l'on prête aux formations végétales rivulaires (ripisylves) à l'échelle locale sur l'état écologique des cours d'eau est significatif à des échelles de gestion (i.e, plus larges) et si oui dans quelles mesures elles sont influentes (dimension, type de couvert végétal,…). C'est une question de recherche centrale à laquelle il faut répondre afin d'atteindre les objectifs de le DCE à l'horizon 2015.

Cette étude implique clairement trois étapes qui sont :

(1) de développer des méthodes efficaces et opérationnelles de cartographie fine de l'Occupation des Sols dans les Corridors Rivulaires (OSCR) à partir de l'imagerie (aéroportée et/ou satellitaire) à haute et/ou très haute résolution spatiale ;

(2) de construire à partir de ces cartes d'OSCR des indicateurs spatialisés pertinents pour représenter les mécanismes d'influence de la végétation et des activités humaines rivulaires sur l'état écologique des cours d'eau, en analysant l'organisation et la composition de l'OSCR ;

(3) d'étudier les relations statistiques entre les indicateurs spatialisés d'occupation des sols calculés aux échelles du bassin versant et du corridor rivulaire et les indicateurs d'état écologique des cours d'eau mesurés au niveau de la station (i.e. développer des modèles pressions/impacts) afin d'améliorer la compréhension des facteurs régulant l'état écologique et de préciser l'influence des espaces ripariaux sur cet état.

En conséquence, la thèse est structurée en quatre parties :

La ***première partie*** est dédiée au contexte de l'étude, à l'état des connaissances sur le sujet, et vise à expliciter les implications de cette recherche.

La ***deuxième partie*** est consacrée au développement de la méthodologie opérationnelle de cartographie fine de l'occupation du sol dans les corridors rivulaires et à son évaluation et sa validation.

La ***troisième partie*** est vouée à la conception des indicateurs spatialisés de l'occupation du sol et de la végétation rivulaire, à l'analyse de leur sensibilité vis-à-vis de la résolution spatiale de la donnée d'occupation du sol, et à l'évaluation de leur capacité à produire des cartes didactiques pour les gestionnaires.

La ***quatrième partie*** est dédiée aux développements de modèles pressions/impacts afin d'évaluer l'apport de ces indicateurs et d'analyser plus précisément les rôles respectifs de l'occupation des sols sur le bassin versant et l'occupation des sols dans les corridors rivulaires.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Première partie :

Contexte de l'étude et État des connaissances

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Introduction de la première partie

Afin de dégager les objectifs spécifiques de la thèse, cette partie détaille le contexte de l'étude et présente l'état des connaissances. Elle est structurée en quatre chapitres.

Le chapitre 1 offre une vue synthétique des concepts et méthodes utilisées pour la caractérisation et l'évaluation de l'état écologique des cours d'eau au sens de la directive cadre européenne. Il expose les méthodes développées, les choix effectués et les premiers résultats obtenus en France.

Le chapitre 2 vise à mieux éclairer l'intérêt des formations rivulaires. Il donne une vue synthétique de l'état des connaissances sur les mécanismes influençant l'état écologique des cours d'eau. Il souligne l'influence des structures rivulaires, et leur importance dans les stratégies de restauration visant à atteindre les objectifs de la directive cadre.

Le chapitre 3 propose un état l'art des approches et outils de modélisation pressions/impacts qui permettent d'étudier les relations entre les pressions anthropiques et l'état écologique des milieux aquatiques.

Enfin, le chapitre 4 expose les objectifs détaillés de la thèse tirés de l'analyse des trois chapitres précédents. Il présente aussi la démarche et le choix des zones d'études pour atteindre les objectifs fixés.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Chapitre 1
Évaluation de l'état écologique des cours d'eau français au sens de la DCE

SOMMAIRE

Résumé *35*

I. CONCEPTS DE BASE ET ELEMENTS CLES **39**

I.1. BON ETAT ECOLOGIQUE ET BON FONCTIONNEMENT *39*

I.2. ORGANISATION ET ELEMENTS CLES DE L'EVALUATION *40*

II. LA TYPOLOGIE PHYSIQUE DES MILIEUX **41**

II.1. EN FRANCE : UNE APPROCHE REGIONALISEE *41*

II.1.1. Objectifs de la typologie 41
II.1.2. Choix du système 42

II.2. LES ELEMENTS DE LA TYPOLOGIE (CADRE CONCEPTUEL) *43*

II.2.1. Le cadre régional 43
II.2.2. La taille des cours d'eau 44
II.2.3. L'Héritage amont 45

II.3. MISE EN ŒUVRE *45*

III. L'EVALUATION DE L'ETAT ECOLOGIQUE **47**

III.1. NOTIONS ET PRINCIPES *47*

III.1.1. La surveillance biologique des eaux 47
III.1.2. Les conditions de référence 48

III.2. PREMIERE EVALUATION A PARTIR DES MACRO-INVERTEBRES BENTHIQUES *49*

III.2.1. Présentation de l'Indice Biologique Global Normalisé (IBGN) 49
III.2.2. Sélection des sites de référence 50
III.2.3. Écart à la référence et limite de bon état 51
III.2.4. Premiers modèles d'extrapolation spatiale 51

III.3. DEVELOPPEMENT EN COURS *53*

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Résumé

Pour une utilisation durable des services des écosystèmes d'eau courante par la société, la Directive Cadre Eau (DCE) impose aux états membres d'atteindre à l'horizon 2015 le « bon état écologique », i.e. un état dans lequel l'écosystème est capable de maintenir les processus fonctionnels lui permettant de « s'auto-renouveler » et, malgré les pressions, de conserver sa résilience, sa diversité et sa productivité.

L'évaluation de cet état repose sur l'étude des structures biologiques (peuplements) qui répondent à ces processus fonctionnels. Elle implique aussi la définition des conditions de référence, spécifiques à chaque type de cours d'eau compte tenu de leur diversité naturelle, de manière à procurer un étalon de mesure pour cette évaluation. Mais, malgré le fait que la DCE reprenne la plupart des concepts qui sous-tendent la gestion de l'eau en France, de nouveaux outils ont dû être développés pour réaliser la phase de diagnostic de l'état écologique des cours d'eau.

Tout d'abord une typologie des milieux physiques, reflétant le fonctionnement écologique naturel des cours d'eau, a été mise en place à partir d'une approche régionalisée descendante. Cette approche, qui se base sur les facteurs de contrôle globaux (géologie, relief, climat) plutôt que sur leurs conséquences au niveau local, permet de définir une première enveloppe de variabilité des paramètres des milieux : les hydro-écorégions. Dans ce cadre régional, la prise en compte de la dimension longitudinale par l'ordination de Strahler et de l'héritage amont par l'hydrochimie et l'hydrologie a permis de définir 124 types de cours d'eau.

Ensuite, une première évaluation de l'état des milieux a été réalisée à partir de l'Indice Biologique Global Normalisé (IBGN). Pour chaque type, la médiane des valeurs IBGN observées sur les sites de référence soumis à des pressions anthropiques faibles, sélectionnés de deux manières indépendantes (par expertise et par analyse des pressions anthropiques sous SIG), constitue la valeur de référence IBGN. Les notes IBGN sont alors transformées, en EQR (pour Ecological Quality Ratio), une mesure de « l'écart à la référence ».

L'état écologique ainsi mesuré au niveau des stations et les résulats obtenus à partir des premiers modèles d'extrapolation spatiale montrent que la situation générale des cours d'eaux en France métropolitaine n'apparaît pas exagérément dégradée tout en étant clairement préoccupante. Cette première évaluation, qui devra être confirmée et complétée par la prise en compte d'autres structures biologiques (poissons, diatomées et macrophytes) et des paramètres physico-chimiques, montre toutefois qu'il est indispensable de rechercher les causes d'altération des milieux dégradés et de s'intéresser plus précisément au rôle du corridor rivulaire pour orienter les gestionnaires dans leurs stratégies de restauration afin d'atteindre les objectifs de la Directive Cadre.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Ce chapitre a été conçu en partie à partir des travaux du Laboratoire d'Hydroécologie Quantitative (LHQ) du Cemagref de Lyon. Le LHQ fournit un appui scientifique et technique aux politiques publiques pour la gestion durable et équilibrée des écosystèmes d'eau courante et en particulier pour la mise en œuvre de la Directive Cadre Européenne sur l'eau en France (régionalisation, typologie, évaluation et diagnostic de l'état écologique). Ces documents sont accessibles au public sur leur site internet[4]. En mars 2009, une partie du fruit de ce travail a été actée dans le guide technique des règles d'évaluation de l'état des eaux douces de surface de métropole par le Ministère chargé de l'environnement.

CemOA : archive ouverte d'Irstea / Cemagref

[4] http://www.cemagref.fr/le-cemagref/lorganisation/les-centres/lyon/ur-maly/Hydroecologie_Cours_dEau

CemOA : archive ouverte d'Irstea / Cemagref

I. Concepts de base et éléments clés

I.1.Bon état écologique et bon fonctionnement

De nouveaux objectifs environnementaux sont clairement décrits dans l'article 4 de la Directive Cadre Eau (DCE) concernant l'évaluation de la qualité de l'eau et font référence à un « état écologique » (ainsi qu'à un « état toxique ») qui traduit la « structure et le fonctionnement des écosystèmes aquatiques ». Le « bon état » se comprend dès lors comme « bon fonctionnement » (Roche *et al.*, 2005).

Les cours d'eau sont des systèmes de « processus – réponses », contrôlés d'abord par les facteurs physiques, même si des éléments biologiques comme la végétation aquatique et rivulaire contribuent à leur structuration. Ils sont soumis à trois grands types de processus (Wasson *et al.*, 2001) :

- *physiques*, qui déterminent la morphodynamique fluviale et l'habitat aquatique ;
- *biogéochimiques*, qui régissent les flux de matière organique et d'énergie métabolique, et les paramètres physico-chimiques essentiels ;
- *écologiques*, qui conditionnent les possibilités de dispersion et de recolonisation des espèces et la structure des peuplements.

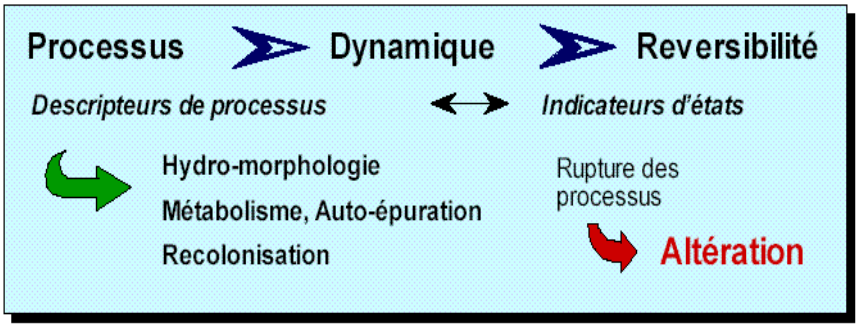


Figure 1.1 : les trois concepts caractérisant le « bon état » : Processus, Dynamique, Réversibilité (d'après Wasson *et al.*, 2001)

Dans un écosystème qui fonctionne « bien », ces processus fonctionnels maintiennent une dynamique physique et biologique (l'état stable est une exception ou une altération), assurant la résilience[5] du système face aux perturbations d'origine naturelle ou anthropique (Figure 1.1). Autrement dit, il est capable de « s'auto-renouveler » et, malgré les stress, de conserver sa résilience, sa diversité et sa productivité, attributs dont dépend une utilisation durable de ses services par la société (Décamps, 2006).

Dès lors, pour poser un diagnostic, la recherche de *descripteurs* de ces processus fonctionnels est à privilégier (Wasson *et al.*, 2001). Cependant l'alternance d'états qui caractérise ces écosystèmes ne rend pas la tâche facile (Décamps, 2006). Le problème est donc d'identifier, aux échelles pertinentes d'espace et de temps, les processus-clés du fonctionnement écologique des milieux, et leurs éventuelles perturbations d'origine anthropique pour orienter la restauration.

[5] Capacité du système à revenir à sa situation initiale après une perturbation (capacité d'adaptation)

CemOA : archive ouverte d'Irstea / Cemagref

I.2. Organisation et éléments clés de l'évaluation

L'évaluation de l'état écologique des milieux aquatiques doit suivre un protocole très précis décrit dans les annexes II et V de la directive (Figure 1.2). Par cette nouvelle définition de la qualité de l'eau, la directive cadre assure la cohérence de la politique de l'eau à l'échelle Européenne. Elle repose sur deux idées novatrices dans l'évaluation de l'état des cours d'eau.

La première idée clé est de faire référence au contexte géographique naturel par une typologie des milieux : la définition des conditions de référence, biologiques mais aussi physicochimiques et hydromorphologiques, doit être établie par type de milieu. La phase d'évaluation commence dés lors par une typologie des milieux sur des critères purement physiques, incluant nécessairement un large cadre géographique. Dans chacun des milieux sont définis des tronçons de cours d'eau ou « masses d'eaux », qui correspondent aux unités à évaluer. Parallèlement, une spatialisation des pressions anthropiques permet de classer les milieux en trois catégories : (1) les milieux de référence, exempts de contamination toxique, et pour lesquels l'impact des activités humaines est supposé non significatif ; (2) des milieux naturels plus ou moins modifiés ou dégradés ; (3) des milieux artificiels fortement modifiés.

La deuxième idée clé est d'évaluer l'état écologique d'un cours d'eau en comparant son état actuel avec les conditions de référence du type de milieu auquel il appartient. Cette évaluation repose principalement sur la bioindication: les peuplements aquatiques (phytoplancton, macrophytes, poissons et invertébrés) sont « juges de paix » (Roche *et al.*, 2005). Si l'écart à la référence mesuré sur les peuplements est « léger » (et si les toxiques ne dépassent pas des seuils qui seront définis sur la base du risque écotoxicologique) les milieux sont considérés en « bon état ». Ils devront faire l'objet d'une surveillance régulière et conserver au moins cet état. Dans le cas contraire, si l'écart est « modéré / fort » (et/ou si les concentrations toxiques dépassent les seuils fixés) les milieux sont classés dans un état « moyen, médiocre ou mauvais ». Ils devront être restaurés pour les ramener au bon état en 2015, sauf dérogations qui devront être soigneusement justifiées en termes de coût et de faisabilité. Cette évaluation concerne bien évidemment les milieux dits «naturels » (objet d'étude de cette thèse), car pour les milieux artificiels ou fortement modifiés, les problèmes et les enjeux sont différents : l'évaluation ne dépend plus uniquement de la qualité et de la biodiversité biologique mais prend en compte aussi le maintien des usages liés à ces milieux pour des enjeux économiques ; on parle alors de « bon potentiel » dont l'objectif est d'atteindre un « potentiel écologique maximum » compte tenu de l'artificialisation du milieu.

Dès lors, les définitions de l'état de référence et du seuil entre le « bon » et le « moyen » état apparaissent comme un élément essentiel dans la phase d'évaluation de l'état des milieux. Sa mise en œuvre amène la communauté scientifique à se poser des questions à la fois conceptuelles (quelle approche typologique ?, qu'est ce que la référence ?, quelle limite de bon état ?) et méthodologiques (quelles sont les données nécessaires et mobilisables, quels outils statistiques ?).

Malgré le fait que la directive cadre reprend la plupart des concepts qui sous-tendent la gestion de l'eau en France, une simple adaptation des textes législatifs ne suffit pas à répondre à ces questions. Les outils existants, comme les Systèmes d'Évaluation de la Qualité des cours d'eau (Agences de l'eau, 1999) ont été adaptés en conséquence. Ce travail a été confié notamment au *Cemagref* et a fait l'objet de nombreux rapports synthétisés dans la suite de ce chapitre.

CemOA : archive ouverte d'Irstea / Cemagref

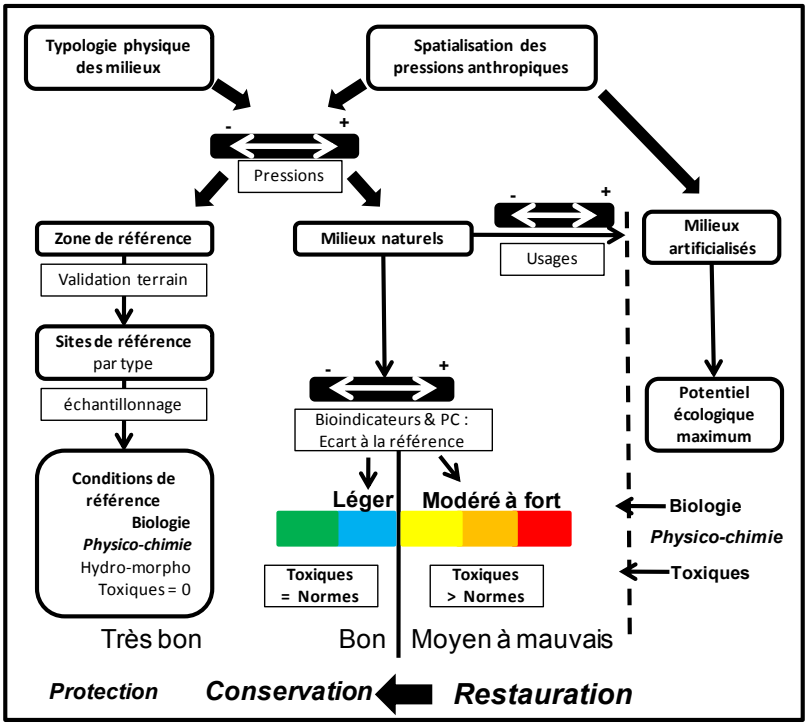


Figure 1.2 : Schéma d'organisation de la phase d'évaluation de la DCE (d'après Wasson, 2001)

II. La typologie physique des milieux

II.1. En France : une approche régionalisée

II.1.1.Objectifs de la typologie

Pour être applicable au concept de « bon état » écologique, la typologie doit être basée sur les processus qui génèrent les différents types de milieux et leurs peuplements associés : plus qu'une classification, elle doit être une typologie fonctionnelle (Wasson *et al.*, 2001). La typologie doit donc refléter le fonctionnement naturel des écosystèmes et leur structure biologique tout en s'appuyant sur des éléments non modifiés par l'activité humaine, puisque l'objectif est de définir des ensembles de milieux qui, en situation non impactée, auraient les mêmes caractéristiques biologiques.

Cette typologie a clairement deux objectifs qui n'ont pas la même échelle de perception (Roche *et al.*, 2005). L'utilité première de la typologie est de permettre la définition de conditions de référence pour chaque type de milieu avec une vision suffisante (qui colle au plus près de la réalité naturelle) pour pouvoir distinguer les sites en bon état des sites altérés, essentiellement sur la base de méthodes de bioindication. La deuxième est d'assurer la délimitation des masses d'eau qui sont les unités d'évaluation de l'état écologique. Cependant, afin d'optimiser les priorités de restauration, la typologie, dans une vision plus large, doit permettre le regroupement des masses d'eau en unité cohérentes de gestion. Si elle reflète correctement le fonctionnement des milieux, des cours d'eau de même type devraient présenter une même sensibilité aux impacts anthropiques. Ainsi pour un type donné, la répartition géographique des différentes catégories d'activités humaines, permettra de délimiter des ensembles de tronçons de cours d'eau souffrant des mêmes « pathologies régionales », et relevant donc d'une même politique de restauration (Wasson *et al.*, 2002).

CemOA : archive ouverte d'Irstea / Cemagref

II.1.2.Choix du système

Pour mettre en place cette typologie la DCE laisse le choix entre deux systèmes : une approche régionalisée, mais basée sur une combinaison imposée de paramètres simples avec des limites figées (système A), ou un système de classification plus souple (système B), ouvrant la possibilité d'une modélisation, avec des paramètres obligatoires (e.g., altitude, géologie, dimension du bassin…) et optionnels (e.g., morphologie, géométrie, substrat, chimie…) et un choix ouvert dans les limites de classes. Les deux systèmes sont basés sur des caractéristiques abiotiques naturelles et requièrent à *minima* une base géographique (incluant la géologie et l'altitude), couplée à la taille du cours d'eau.

Le système A repose sur des bases conceptuelles saines, à savoir que la diversité des types d'écosystèmes d'eau courante provient à la fois des caractéristiques régionales et de la dimension des bassins versant. Cependant, les « écorégions » proposées sont inadaptées car ce sont des régions biogéographiques basées essentiellement sur les insectes aquatiques, sans lien fort avec le fonctionnement des écosystèmes d'eau courante (Wasson *et al.*, 2002). Par exemple, dans ce système il n'apparaît pas en France d'écorégion méditerranéenne, alors que le stress hydrique est une caractéristique structurelle fondamentale (Wasson *et al.*, 2003b). D'autre part le cadre rigide des classes prédéfinies conduit à des incohérences lorsque l'on croise les différents paramètres, ce qui aboutit à masquer les discontinuités naturelles de l'espace et à multiplier le nombre de types. Par exemple, en France, les limites arbitraires d'altitude aboutissent à morceler des ensembles géographiques homogènes (Wasson *et al.*, 2003b).

Pour ces raisons, une grande majorité des pays ont développé des typologies à partir du système B.

Dans ce système s'insère notamment la méthodologie RIVPACS[6]. Cette approche multivariée pionnière fut élaborée par l'intégration de paramètres physiques et chimiques afin de prédire la faune benthique des milieux aquatiques. Actuellement, les systèmes multivariés RIVPACS et ses dérivés (e.g, AusRivAS[7], BEAST[8], STAR[9]) sont utilisés par la plupart des États membres (Davy-Bowker *et al.*, 2006) et les plus employés mondialement (Sandin *et al.*, 2001). Elle permet, par l'entremise de modèles statistiques, d'établir les relations discriminantes entre les variables environnementales (variables chimiques et physiques de l'habitat) et les biotypes de référence (Wright *et al.*, 1998). Elle requière pour cela un nombre important de sites de référence pour obtenir un modèle statistiquement fiable et applicable.

Toutefois, cette approche ne permet pas l'obtention d'une typologie au sens défini précédemment (i.e., basée sur les processus fonctionnels), mais un modèle de référence permettant d'asseoir une méthode de bioindication (Wasson *et al.*, 2003b). Ce modèle qui a l'avantage de prendre en compte une description plus précise de la variabilité longitudinale des paramètres physiques et biologiques le long du réseau hydrographique, peut paraître plus satisfaisant, mais il est beaucoup plus difficile à mettre en œuvre dans un contexte géographique très hétérogène, et demande de nombreuses données. En France, un tel modèle validé n'existe pas à l'heure actuelle (Wasson *et al.*, 2003b). De plus, il n'est pas entièrement compatible avec la DCE car il permet plus l'identification de types de « sites » que de types de « masse d'eau » (unités essentielles dans les stratégies de gestion de la DCE).

Dès lors, en France, un système utilisant les bases conceptuelles du système A (concept d' « éco-régions ») et en cohérence avec les paramètres requis par le système B a été développé pour mieux répondre au double objectif de la typologie physique des milieux (Wasson *et al.*, 2002). Ce système à la fois régionalisé, c'est-à-dire intégrant sur une base spatiale les facteurs de contrôle des écosystèmes aquatiques, et hiérarchisé, c'est-à-dire constitué d'éléments géographiques emboîtés utilisables à plusieurs niveaux de précision selon le besoin, garantit à la fois les aspects scientifiques (refléter le fonctionnement écologique des milieux) et pratiques (opérationnalité et communication pour les politiques de gestion) attendues. Cette approche de régionalisation présente de fortes similitudes avec celle d'Omernik, 1987 largement utilisée par l'US-EPA (United States Environmental Protection Agency) pour la définition des conditions de référence et la bioindication. Elle s'en distingue cependant par une analyse séparée des structures naturelles et anthropiques. La suite de cette partie est dédiée à la présentation synthétique du cadre conceptuel, de la méthodologie et des résultats de cette approche mise en place en France.

[6] River Invertebrate Prediction and Classification System (Wright *et al.*, 1993)
[7] AUStralian River Assessment System (Parsons and Norris, 1996)
[8] BEnthic Assessment of SedimenT (Reynoldson *et al.*, 1995)
[9] Standardisation of River Classifications (Furse *et al.*, 2006)

CemOA : archive ouverte d'Irstea / Cemagref

II.2. Les éléments de la typologie (cadre conceptuel)

Un tronçon homogène sur le plan fonctionnel doit en théorie, selon les lois de la géomorphologie fluviale, présenter des caractéristiques géomorphologiques homogènes (géométrie du lit, pente, sinuosité, style fluvial, etc.) qui intègrent les quatre dimensions de l'hydrosystème (Figure 0.1). Ces caractéristiques résultent de la superposition de deux facteurs structurels : l'hétérogénéité régionale et le gradient amont-aval (Wasson *et al.*, 2001). Trois éléments importants ont été pris en compte pour aboutir à une typologie fonctionnelle intégrant ces deux facteurs : (1) Le cadre régional ; (2) la taille des cours d'eau ; (3) l'héritage amont.

II.2.1.Le cadre régional

Une régionalisation simple a été construite à partir du concept d'« hydro-écorégions » (HER). Il est fondé sur une approche descendante dont le principe de base, énoncé par Lotspeich, 1980 et Lotspeich and Platts, 1982 se résume ainsi : il est plus pertinent de classifier les causes que les conséquences, et donc les facteurs de contrôle globaux plutôt que leurs conséquences au niveau local. Ce concept s'est enrichi au contact des théories de contrôle hiérarchique des hydrosystèmes, et particulièrement celle de l'emboîtement des échelles physiques, du bassin versant jusqu'au micro-habitat, proposée par Frissell *et al.*, 1986 et Naiman *et al.*, 1992. Tous ces auteurs reconnaissent que la géologie, le relief et le climat sont les déterminants primaires du fonctionnement écologique des cours d'eau à l'échelle du bassin.

On peut schématiser comme suit l'enchaînement hiérarchique des facteurs qui déterminent le fonctionnement écologique des cours d'eau (Figure 1.3) (Wasson *et al.*, 2002) : ***à l'échelle locale***, dans le milieu aquatique, la biodiversité et la productivité des communautés vivantes répondent aux « facteurs-clés » que sont l'habitat physique (hydraulique, substrat), le climat aquatique (lumière, température, gaz dissous, hydrochimie), et les ressources trophiques (production endogène et matière organique exogène) ; toutefois l'influence de l'habitat local sur des paramètres essentiels (oxygénation, ressources trophiques) conduit à assigner un rôle prépondérant aux facteurs physiques. Ensuite, ***à l'échelle du tronçon***, la dynamique spatio-temporelle des facteurs-clés dépend du régime hydro-sédimentologique et de la géomorphologie de la vallée ; ces « facteurs de contrôle » définissent la morpho-dynamique fluviale (forme et stabilité du lit), la structure de la végétation rivulaire (qui influe significativement sur le milieu aquatique), et la connectivité du cours d'eau (interactions avec la zone inondable). Finalement, ***à l'échelle régionale***, ces facteurs dépendent des «déterminants primaires» que sont la géologie (nature des roches), le relief (géomorphologie) et le climat (températures et précipitations). Les sols et la végétation du bassin interviennent évidemment dans le bilan hydrique et sédimentologique, mais ils sont en dernier ressort déterminés par les caractéristiques géophysiques et climatiques.

Ce cadre conceptuel justifie une régionalisation des écosystèmes d'eau courante à partir des déterminants primaires géophysiques et climatiques, qui permettent de délimiter les HER. Les hypothèses de base sont les suivantes :

- à l'intérieur d'une même région les cours d'eau présenteront des caractéristiques physiques et biologiques similaires et un même gradient d'évolution longitudinale ;

- les écosystèmes d'eau courante de différentes régions devraient se distinguer sur au moins un paramètre abiotique important, conduisant à des différences significatives et remarquables au niveau des peuplements.

CemOA : archive ouverte d'Irstea / Cemagref

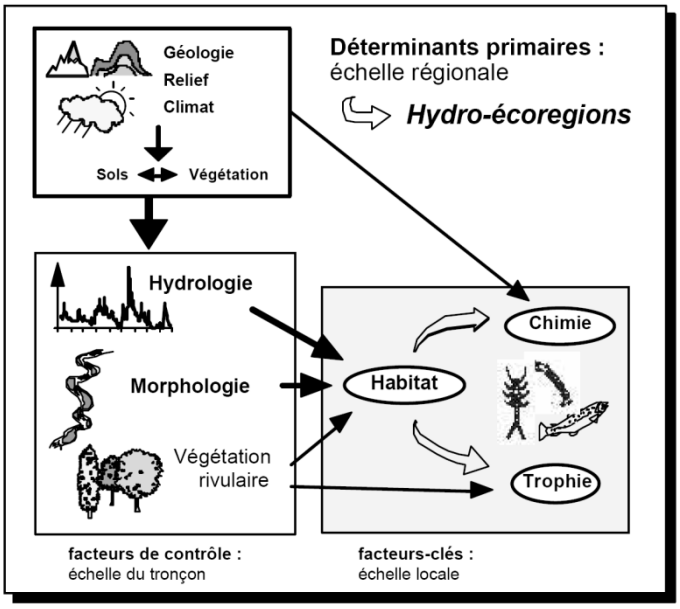


Figure 1.3 : Modèle conceptuel de l'écosystème d'eau courante d'après Wasson, 1993

Les HER, expression synthétique des facteurs déterminant la physico-chimie de l'eau, les régimes hydrologiques et sédimentaires, la morphologie du lit et de sa dynamique, l'environnement végétal, permettent de délimiter ainsi une première « enveloppe de variabilité » des facteurs clés de l'écosystème d'eau courante. Elles constituent le premier niveau de différenciation de la typologie.

II.2.2. La taille des cours d'eau

Les facteurs hydrologiques et morphologiques (facteurs de contrôle) conditionnent très largement la structure physique et le fonctionnement écologique à l'échelle du tronçon. Le gradient longitudinal (amont-aval) de ces facteurs a fait l'objet depuis plus d'un siècle de très nombreux travaux aboutissant à des propositions de zonations (Huet, 1949, Illies and Botosaneanu, 1963), ou des typologies longitudinales (Verneaux, 1976 , Vannote *et al.*, 1980) plus ou moins fonctionnelles (Wasson, 1989).

L'ordination de Strahler, 1957, largement utilisée dans des modèles fonctionnels (e.g., RIVERSTRAHLER, Billen *et al.*, 1994, Meybeck *et al.*, 1998), est une méthode simple, robuste, permettant de proposer des limites typologiques au niveau des discontinuités longitudinales réelles du réseau hydrographique correspondant à des changements de dimension (de taille des cours d'eau).

L'intérêt majeur de l'ordination de Strahler (Figure 1.4) est de prendre en compte la structure et la densité du réseau qui reflète, en première approche, la quantité d'eau transportée en surface durant les épisodes de forts débits. Ces épisodes constituent un des indices les plus significatifs en termes de processus hydrologiques au niveau du bassin (Gregory and Walling, 1973) et un élément majeur dans la structuration morphologique du cours d'eau, en particulier la section à plein bord et les faciès morphodynamiques.

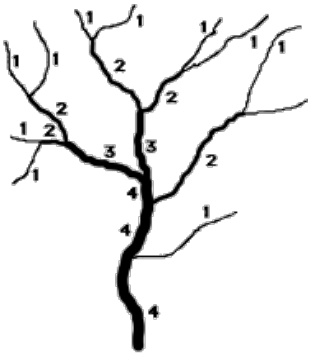


Figure 1.4 : Méthode d'ordination des cours d'eau selon Strahler, 1957. L'ordination consiste à attribuer le rang 1 aux cours d'eau en tête de bassin, puis, en progressant vers l'aval, chaque cours d'eau recevant un cours d'eau du même rang voit son rang augmenter d'une unité.

CemOA : archive ouverte d'Irstea / Cemagref

Cette densité varie fortement selon la situation géographique (bilan hydrique, perméabilité des sols). Il est donc logique que des tronçons de même rang aient des surfaces de bassin versant sensiblement différentes selon les régions. Cependant, cette variabilité de surface étant compensée par l'influence de la lame d'eau écoulée, des cours d'eau de même rang dans différentes régions ont des dimensions relativement comparables. Une analyse effectuée sur le bassin de la Loire (Souchon *et al.*, 2000) montre qu'à rang égal la largeur moyenne du lit, qui reflète mieux la dimension du cours d'eau, est équivalente quelque soit l'hydro-écorégion du bassin.

En revanche, l'ordination ne préjuge en rien de la variabilité hydrologique saisonnière et des autres paramètres de la structure morphologique comme la pente et la géologie du bassin qui conditionnent l'énergie du cours d'eau et la nature des sédiments transportés. Cependant, ces caractéristiques sont bien prises en compte par les HER. Dès lors, le croisement HER x rang de Strahler est supposé avoir une signification réelle en termes de morphodynamique et d'habitat aquatique (Chandesris *et al.*, 2006).

II.2.3.L'Héritage amont

Le classement typologique proposé (HER x rangs) ne prend pas en compte la caractéristique principale des cours d'eau : ils s'écoulent vers l'aval et transfèrent par-là même des caractéristiques issues des facteurs prédominants des HER de leur bassin amont dans l'HER aval où ils se situent. La formulation de cet héritage amont apparaît donc indispensable pour identifier le cas échéant des types de masses d'eau susceptibles de se différencier en raison de l'influence des HER amont. On parle alors de cours d'eau de type « *exogène »* qui transportent une information différente provenant d'une ou plusieurs HER amont (géologie, climat, relief), information qui se dilue progressivement jusqu'à devenir négligeable en ce qui concerne le fonctionnement de l'écosystème aquatique.

Deux facteurs intégrateurs des déterminants primaires du fonctionnement des hydro-systèmes ont été utilisés pour quantifier cet héritage : (1) la chimie de l'eau (facteur intégrateur de la géologie), en utilisant comme traceur l'ion calcium, élément essentiel du fonctionnement de l'écosystème : pH, alimentation en carbone des végétaux, régulation du métabolisme bactérien ; et (2) l'hydrologie (facteur intégrateur du relief et du climat), sur la base du débit moyen annuel naturel qui traduit les transports de matières.

Ces deux paramètres sont complémentaires. Ils sont bien évidemment non influencés de manière significative par des altérations d'origine anthropique, et se répartissent de façon homogène et suffisamment dense sur l'ensemble du territoire étudié pour analyser leurs effets sur les différents types de cours d'eau.

Un cours d'eau sera donc de type exogène dès qu'il sera influencé par l'hydrochimie ou l'hydrologie d'une HER amont.

II.3.Mise en œuvre

Les trois éléments, décrits précédemment, ont été mis en œuvre pour définir la typologie nationale des cours d'eau.

Tout d'abord, la recherche des discontinuités spatiales des déterminants primaires (caractéristiques géophysiques et climatiques) a nécessité l'utilisation de couches (i.e., des cartes dont la validité et la pertinence ont dû être soigneusement évaluées en comparant différentes sources d'informations) rassemblées dans un Système d'Information Géographique (SIG) (Pella *et al.*, 2004). Elle a permis de délimiter des hydro-écorégions (HER) de deux niveaux hiérarchiques. Le niveau HER-1 (Figure 1.5) qui correspond aux grandes structures géophysiques et climatiques, le niveau HER-2 qui correspond à des variations générales à l'intérieur de ces types (carte en ***ANNEXE I***). Chaque polygone est caractérisé par une distribution homogène de paramètres abiotiques (lithologie, altitude, pentes, hydrographie, température, précipitations).

CemOA : archive ouverte d'Irstea / Cemagref

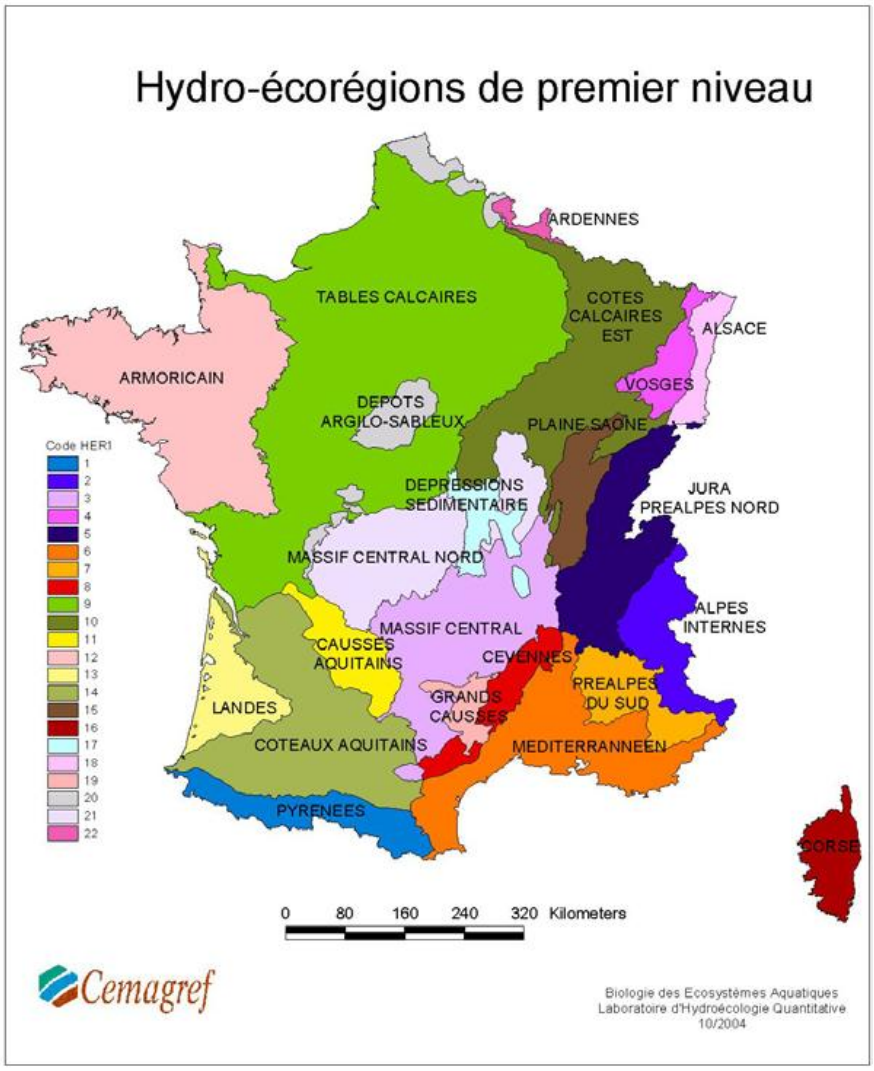


Figure 1.5 : Carte des hydro-écorégions (HER-1) de France métropolitaine d'après Wasson *et al.*, 2002

Ensuite cinq classes de tailles différentes fondées sur le rang de Strahler (Tableau 1-1), avec des ajustements locaux notamment pour les systèmes karstiques, ont été proposées à partir du rang de Strahler pour traduire la dimension longitudinale après analyse globale de l'ordination avec les surfaces de bassin versant (Chandesris *et al.*, 2006).

Tableau 1-1 : Classes de taille pour les cours d'eau en fonction des rangs de Strahler (Chandesris et al., 2006)

Classes de taille	Très petit	Petit	Moyen	Grand	Très Grand
Rangs : bassin Loire Bretagne	**1,2,3**	**4**	**5**	**6**	**7,8**
Rangs : autres bassins	**1,2**	**3**	**4**	**5**	**6,7**

Enfin, le caractère exogène de certains cours d'eau, soumis à l'influence d'une hydro-écorégion située à l'amont, a été traduit par les caractéristiques liées à l'hydrochimie et/ou à l'hydrologie. Pour chacun des paramètres (concentration en ion calcium pour le facteur hydrochimique et débits reconstitués[10] pour le facteur hydrologique), des règles d'identification ont pu être établies en analysant les relations amont-aval

[10] Grâce aux résultats récents d'extrapolation spatiale des données issus de la base nationale des données hydrométriques (Sauquet, 2005), il est possible de connaître la lame d'eau annuelle moyenne générée par chaque zone hydrographique, ce que l'on peut traduire par un débit naturel (module annuel) reconstitué à l'exutoire de chacune des zones hydrographiques.

CemOA : archive ouverte d'Irstea / Cemagref

entre des groupes d'HER homogènes (Chandesris *et al.*, 2006). Ces groupes (pour chacun des facteurs) ont été constitués au préalable à partir de l'observation des valeurs de paramètres au niveau des stations de type endogène (cours d'eau dont la totalité du bassin versant se situe dans une même HER). Ainsi un type exogène est identifié si, pour les cours d'eau situés en région calcaire et provenant d'une région non calcaire, la concentration est inférieure à 60 mg/l de Ca++ ; et/ou si la quantité d'eau « supplémentaire » représente 50 % du débit théorique aval (l'influence sera plus forte si la quantité d'eau « supplémentaire » représente 200% du débit théorique aval).

Ainsi pour chaque hydro-écorégion, les masses d'eau seront affectées à l'une des cinq classes de taille et affiliées à l'appartenance éventuelle à un type exogène.

Une validation de la pertinence de la typologie résultante a été réalisée en analysant des données d'invertébrés benthiques (issues de stations de référence). Dans l'ensemble une bonne correspondance entre les types et la faune d'invertébrés est observée (Wasson *et al.*, 2002, Wasson *et al.*, 2004a). Cette validation a permis (1) de simplifier la typologie en regroupant dans certaines HER, des classes de taille de cours d'eau pour lesquelles les différences sur les invertébrés n'étaient pas perceptibles ; et (2) de prendre en compte une variabilité résiduelle à l'intérieur de certaines HER de niveau 1, en isolant des « sous-types biologiques » correspondant à des HER de niveau 2 aux caractéristiques particulières.

Au final, la combinaison du cadre des hydro-écorégions, des classes de tailles ou rangs de Strahler (plus la simplification), et des types « exogènes », a abouti à une typologie physique des cours d'eau comprenant 124 types au niveau de la France métropolitaine.

L'approche par HER constitue un premier niveau de différenciation robuste et scientifiquement fondé : les analyses statistiques des données d'invertébrés (Wasson *et al.*, 2002) montrent que l'effet « écorégion » explique une part prépondérante de la variation des métriques utilisées pour la bioindication, notamment la richesse taxonomique et la distribution des groupes de taxons indicateurs qui déterminent la note de l'IBGN. Dans ce cadre, d'autres facteurs physiques significatifs décrivant la géométrie des chenaux pourront être pris en compte pour aboutir à une sectorisation en tronçons géomorphologiquement homogènes (Valette *et al.*, 2008) afin d'évaluer l'état hydromorphologique des cours d'eau.

De plus, ce système régionalisé et hiérarchisé, constitué d'éléments géographiques emboîtés utilisables à plusieurs niveaux de précision selon les besoins, permet aussi d'établir une typologie-cadre cohérente et opérationnelle pour la gestion. La transcription cartographique des problèmes facilitera la communication avec les décideurs publics (Roche *et al.*, 2005).

III. L'évaluation de l'état écologique

III.1.Notions et principes

III.1.1.La surveillance biologique des eaux

Les structures de peuplements répondent aux trois grandes catégories de processus qui déterminent le fonctionnement du cours d'eau (processus physiques, biogéochimiques et écologiques cf. I.1). L'utilisation de la bioindication (à l'échelle de la communauté) se révèle alors comme l'outil adéquat pour évaluer l'état écologique des cours d'eau au sens de la directive.

Ces indicateurs sont construits selon le principe biocénotique : à un milieu donné correspond une biocénose[11] particulière résumant l'expression de l'ensemble des facteurs écologiques qui conditionnent le système. Les perturbations naturelles ou anthropiques, qui se traduisent par l'évolution de certains de ces facteurs, provoquent alors des modifications plus ou moins marquées des communautés vivantes hébergées. Ainsi, l'analyse de la composition d'un répertoire faunistique considéré isolément, permet de définir l'état du milieu tandis que l'analyse biocénotique comparative permet d'évaluer l'effet des changements de l'environnement qui affectent les communautés (Verneaux *et al.*, 1982).

[11] En écologie, une biocénose (ou biocœnose) désigne l'ensemble des êtres vivants coexistant dans un espace défini (le biotope)

CemOA : archive ouverte d'Irstea / Cemagref

Cette surveillance porte, notamment en France, sur des indices définis selon des protocoles normalisés qui ont été développés à partir des communautés : (1) de poissons (IPR : Indice Poisson Rivière NF T90-344, 2004-05-01) ; (2) d'invertébrés (IBGN : Indice Biologique Global Normalisé NF T90-350, 2004-03-01) ; (3) de diatomées (IBD : Indice Biologique Diatomées NF T90-354, 2007-12-01) ; et (4) de macrophytes (IBRM : Indice Biologique Macrophyte en Rivière NF T90-395, 2003-10-01) pour les écosystèmes aquatiques. L'utilisation de plusieurs bioindicateurs est nécessaire à l'évaluation car on sait depuis longtemps qu'ils réagissent à des paramètres et à des gammes d'altération différents (Kolkwich et Marsson, 1902 in Bonada *et al.*, 2006).

Ces indices sont dits « multimétriques » car ils s'appuient sur plusieurs attributs (ou métriques, e.g. composition taxonomique, abondance…) des communautés biologiques pour évaluer l'état d'un secteur de rivière (i.e. la station d'étude). Le choix de ces métriques repose sur leur capacité à répondre aux activités humaines : elles doivent varier de manière prévisible en fonction de ces activités et permettre de distinguer les changements provoqués par les activités humaines du « bruit de fond » des variations naturelles. Cela suppose une bonne connaissance de l'histoire naturelle des milieux aquatiques (Décamps, 2006).

La surveillance biologique, via ces métriques, permet de combler les lacunes des approches habituelles (approches physicochimiques). Elles permettent : (1) de détecter toutes les perturbations, quelles qu'elles soient, y compris celles des substances ou phénomènes inconnus de l'observateur ; (2) de mettre en évidence les effets des mélanges de contaminants (effets synergiques, additifs, cumulatifs et antagonistes) ; et (3) d'intégrer le facteur temps : alors qu'un dosage chimique donne une image ponctuelle de la qualité de l'eau, un organisme qui accomplit tout ou partie de son cycle vital dans le milieu aquatique témoigne des conditions passées du milieu.

Cinq métriques fonctionnelles sont requises par la DCE : la composition taxonomique, l'abondance, le ratio de taxa sensibles, la diversité et l'âge des structures. D'après la Figure 1.6, qui résume la distribution de ces métriques selon le peuplement aquatique étudié, les études sur les poissons et les macroinvertébrés benthiques semblent les mieux adaptées aux exigences de la DCE.

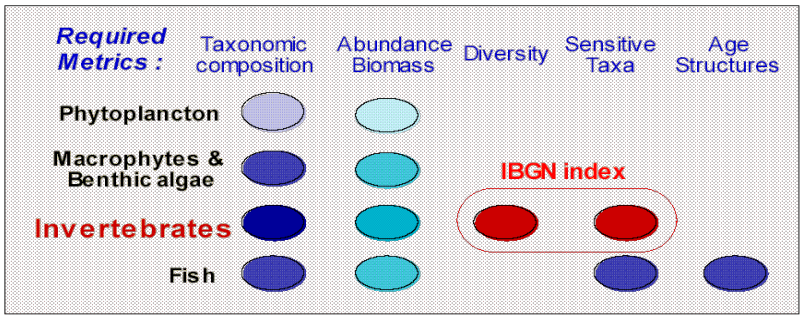


Figure 1.6 : Métriques fonctionnelles requises par la DCE en fonction du peuplement aquatique étudié et métriques prises en compte par l'Indice Biologique Global Normalisé (IBGN) (Wasson *et al.*, 2004b)

III.1.2. Les conditions de référence

La notion d'état (ou de conditions) de référence est essentielle dans la logique de la DCE. Ce n'est pas l'objectif fixé par la directive mais une situation repère, variable selon le contexte naturel, qui constitue le point de calage des métriques de bioindication (Décamps, 2006). Cette donnée fondamentale doit reposer sur des bases scientifiques solides (Wasson *et al.*, 2001).

La référence n'est pas définie dans la DCE par une diversité biologique maximale, rarement observée dans la nature, mais à partir des pressions anthropiques : elle correspond à un niveau de pression anthropique suffisamment faible pour que l'impact au niveau du fonctionnement de l'écosystème et des communautés biologiques soit négligeable (c'est-à-dire difficilement discernable de la variabilité naturelle, ou très localisé). Mettre l'accent sur les processus fondamentaux permet alors de chercher des références dans des bassins où l'occupation humaine n'a pas modifié significativement ces processus : hydrologie et dynamique fluviale non entravées par des ouvrages, flux de matière organique et de nutriments faiblement modifiés, connectivité latérale et longitudinale maintenue, etc. Les conditions de référence sont alors définies par l'enveloppe de variabilité des paramètres biologiques observés dans ces situations. Il s'agit donc bien d'une définition statistique intégrant la variabilité naturelle, à la fois locale et temporelle à court terme (Roche *et al.*, 2005).

En outre, les conditions de référence ne correspondent pas nécessairement à celles d'un passé plus ou moins lointain. Elles peuvent varier dans le temps compte tenu de la dynamique du milieu physique sous

CemOA : archive ouverte d'Irstea / Cemagref

l'influence des variations climatiques ou de la végétation des bassins. L'hypothèse de base est que la biocénose s'adaptera à l'évolution naturelle de ces conditions physiques, si bien évidemment les processus d'ajustement morphodynamiques ne sont pas entravés par des ouvrages structurants (digues, barrages...), et que les dynamiques écologiques sont actives (ce qui suppose le maintien du stock d'espèces et des connectivités du système). Les conditions de référence correspondent alors à une biocénose en adaptation permanente à un milieu physique en « déséquilibre dynamique » (Roche *et al.*, 2005).

A partir des données correspondant aux sites de référence, le guide européen REFCOND de la DCE (Wallin *et al.*, 2003) indique trois méthodes pour définir les valeurs de référence des métriques de bioindication et fixer les limites des cinq classes d'état (très bon, bon, moyen, mauvais, très mauvais). Alors que la deuxième méthode se résume à un avis d'expert, les deux autres se basent sur des grandeurs robustes, comme la médiane de la distribution des données des sites de référence, pour définir les valeurs de référence.

III.2. Première évaluation à partir des macro-invertébrés benthiques

III.2.1.Présentation de l'Indice Biologique Global Normalisé (IBGN)

L'IBGN, mis au point en s'appuyant sur les analyses statistiques de nombreuses données relevant de protocoles antérieurs et standardisé depuis 1992 (norme AFNOR NF-T92), est très largement appliqué en France et a prouvé son efficacité dans l'évaluation globale de la qualité biologique d'un cours d'eau (Usseglio-Polatera *et al.*, 2001). Il est employé dans de nombreux réseaux de mesures tels que le Réseau National de Bassin (RNB), les réseaux complémentaires de bassin, et le Réseau Hydrobiologique Piscicole (RHB) et bien évidemment dans les nouveaux réseaux de surveillance de l'état écologique des cours d'eau.

Cet indice est basé sur l'étude des macroinvertébrés benthiques. Ce sont principalement des vers, des crustacés, des mollusques et des insectes qui vivent au fond des ruisseaux, rivières, lacs ou marais. Ils forment une partie importante des écosystèmes d'eau douce puisqu'ils servent de nourriture à nombre de poissons, d'amphibiens et d'oiseaux et participent à la transformation de la matière organique en la dissociant en éléments plus fins facilitant l'action de dégradation des micro-organismes. C'est un groupe très diversifié, et les organismes le composant possèdent des sensibilités variables à différents stress tels que la pollution ou la modification de l'habitat. Les macroinvertébrés sont les organismes les plus souvent utilisés pour évaluer l'état de santé des écosystèmes d'eau douce compte tenu de la facilité à les prélever et à les reconnaître (Barbour *et al.*, 1999 ; Tachet *et al.*, 2000).

L'IBGN est établi par stations de prélèvement. Une station de prélèvement correspond en général à une section de la rivière dont la longueur est égale à environ dix fois la largeur. L'échantillonneur est équipé d'un filet (vide de maille 500 µm) de type Surber et effectue huit prélèvements par station en favorisant l'échantillonnage des habitats (caractérisés par un couple substrat-vitesse) les plus biogènes (capacité à accueillir une biocénose). Dans l'analyse de ces prélèvements, l'unité taxonomique retenue est la famille, à l'exception de quelques groupes faunistiques pour lesquels c'est l'embranchement ou la classe (notamment pour les oligochètes). Le système indiciaire associé à l'IBGN (noté sur 20) est basé seulement sur 2 métriques : le *Groupe Faunistique Indicateur* (GFI) plutôt qualitatif, qui correspond à la présence de taxa sensibles (présence/absence), mais qui n'indique pas leur pourcentage dans la communauté ; la *richesse ou Variété Taxonomique* (VT), qui est une composante de la diversité mais qui, au sens écologique de la diversité, ne prend pas en compte l'abondance relative des différents taxons.

L'IBGN, rarement appliqué sur les grands cours d'eau (> 5 dans l'ordination de Strahler), intègre, et encore de manière partielle, seulement 2 des 5 métriques requises par la DCE (Figure 1.6), ce qui impose d'emblée une certaine prudence (Wasson *et al.*, 2003a). De plus, ses réponses sont variables selon les régions et de nombreux biais sont possibles (cf. Chapitre 9 :II.1.2).

Néanmoins, l'IBGN apparaît, parmi les bioindicateurs appliqués en France, comme le plus pertinent et le plus robuste pour évaluer le « bon état » au sens de la DCE. Une base de données IBGN nationale (GIRAFE), issue des données réseaux existantes (RNB et RHB) a été constituée au Cemagref, rassemblant 3640 stations et un total de 12318 relevés (ou notes IBGN) couvrant la période 1992-2004 (Figure 1.7). C'est cette base qui a été utilisée pour mener la première évaluation de l'état écologique des cours d'eau en France.

CemOA : archive ouverte d'Irstea / Cemagref

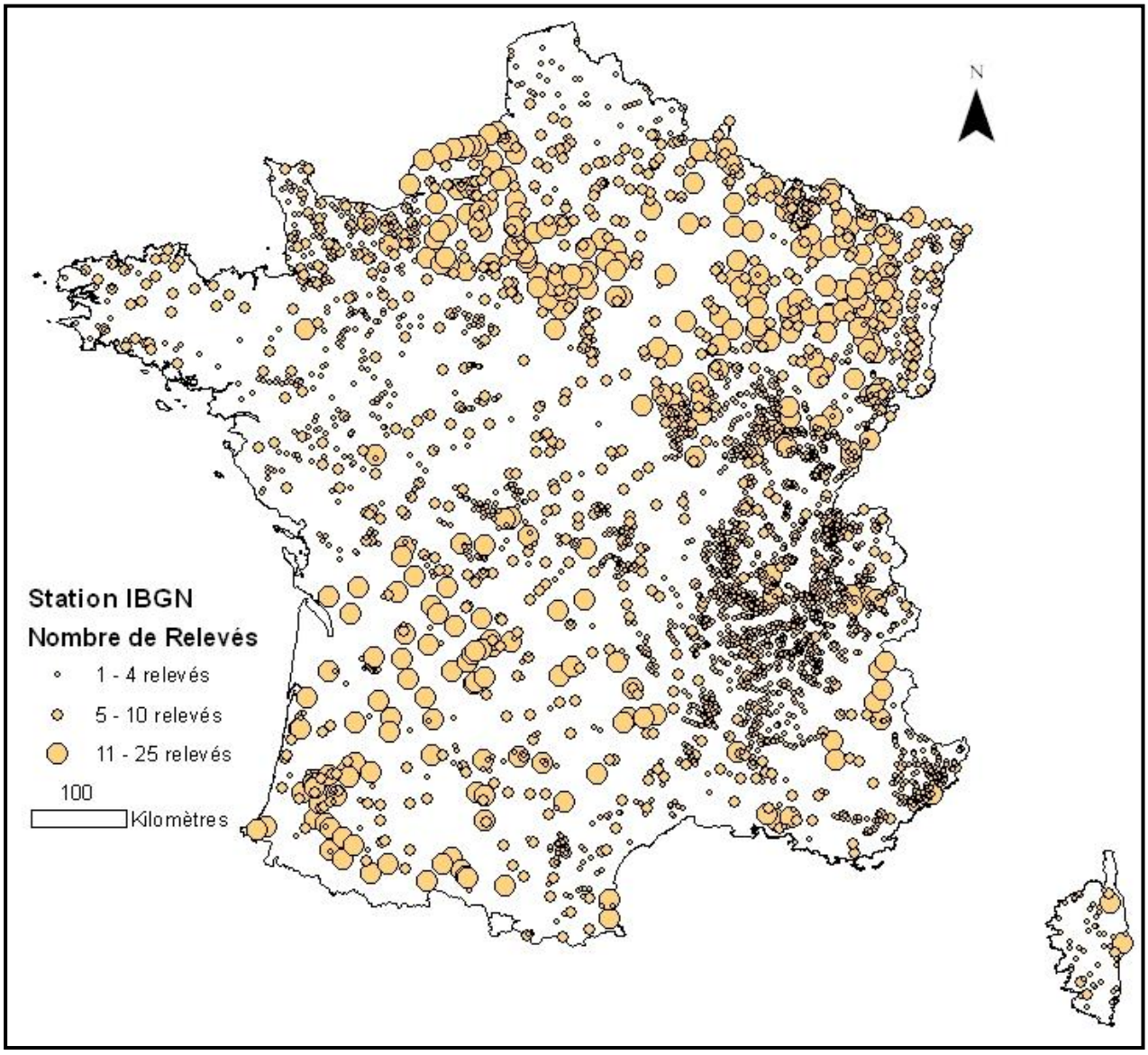


Figure 1.7 : Cartographie des stations IBGN en France d'après la base de données GIRAFE couvrant la période de 1992 à 2004

III.2.2.Sélection des sites de référence

A partir de la base de données GIRAFE, une première sélection des sites de référence pour chaque type a été mise en place en France selon des critères de « très faible pression anthropique ». Deux approches indépendantes d'évaluation des pressions anthropiques ont été confrontées.

La première approche est une sélection à dire d'expert, par les services déconcentrés (DREAL), selon un questionnaire précis (***ANNEXE II***).

La deuxième étudie, par le moyen d'un Système d'Information Géographique (SIG), les données de pressions anthropiques à partir de la base de données d'occupation du sol CORINE Land Cover (issue de l'interprétation des images à moyenne résolution spatiale de 30 m et avec une unité cartographique minimum de 25 ha) calculées au niveau du bassin versant amont de la station et localement au niveau du corridor rivulaire (Pella *et al.*, 2004).

A ce stade, la première approche (expertise) est absolument indispensable. En effet, l'analyse sous SIG des pressions est en limite d'utilisation pour plusieurs raisons : (1) les altérations physiques (aménagements) n'ont pas pu être prises en compte (données non bancarisées à l'échelle nationale) ; (2) la résolution spatiale de la donnée d'occupation du sol utilisée pour quantifier les variables de pressions n'est pas assez fine à l'échelle du

CemOA : archive ouverte d'Irstea / Cemagref

corridor rivulaire ; et (3) il n'était pas possible de modéliser à l'échelle du territoire ni les phénomènes d'autoépuration, ni les pollutions diffuses.

Toutefois la confrontation de ces deux approches permet d'aboutir à une première sélection suffisamment fiable pour formuler des hypothèses de travail. Les données IBGN des sites de référence ont été validées par des données « invertébrés » (non IBGN) indépendantes provenant des stations de référence propres au Cemagref (Wasson *et al.*, 2003a).

III.2.3.Écart à la référence et limite de bon état

La méthodologie pour normaliser les indices IBGN suit les recommandations de la méthode 3 du guide européen « REFCOND ». Pour un type naturel de masses d'eau, la valeur de référence REF est la médiane de la distribution des valeurs IBGN observées sur les sites de référence de ce type. Elle correspond ainsi à la valeur la plus probable de l'IBGN en conditions de référence. Les premiers résultats donnent des valeurs de référence REF de l'indice IBGN variant entre 14 et 19 selon les HER (Roche *et al.*, 2005).

Les notes IBGN sont alors transformées, selon les préconisations de la DCE, en EQR (pour Ecological Quality Ratio), selon la formule :

$$EQR = \frac{IBGN - 1}{REF - 1}$$

Cet EQR-IBGN, détermine ainsi un indicateur d'état écologique indépendant des variations naturelles (typologiques) de l'indice. Il varie entre 0 et 1 (exceptionnellement supérieur à 1 pour des milieux en excellent état). En toute logique, la limite du Bon Etat correspondant à un « écart léger » devrait se traduire par des EQR à peu près équivalents pour tous les types (Wasson *et al.*, 2003a).

La limite inférieure du « très bon état » (couleur bleue) est fixée au 25ème percentile de la distribution des valeurs IBGN observées sur les sites de référence. A noter que par construction, un quart des notes IBGN observées sur les sites de référence sont en-dessous de la limite inférieure du « très bon état ».

Ensuite, en première hypothèse, la limite inférieure du « bon état » (couleur verte) est fixée en divisant en 4 classes égales l'étendue des valeurs IBGN entre la limite inférieure du très bon état et la valeur minimum « réaliste » de l'indice, fixée à 1 pour l'IBGN (Figure 1.8).

Chaque limite de classe est alors exprimée en EQR-IBGN. C'est sur cette base que les valeurs des limites de classes peuvent être comparées, entre différents types de milieux pour un même pays.

La cohérence des valeurs de référence et de limites de classes a été vérifiée et ajustée en analysant, de manière adaptée, les valeurs des métriques de l'IBGN (Groupe Faunistique Indicateur et Variété Taxonomique) (Wasson *et al.*, 2003a) et lors de l'inter-calibration au niveau européen (ECOSTAT GROUP, 2007).

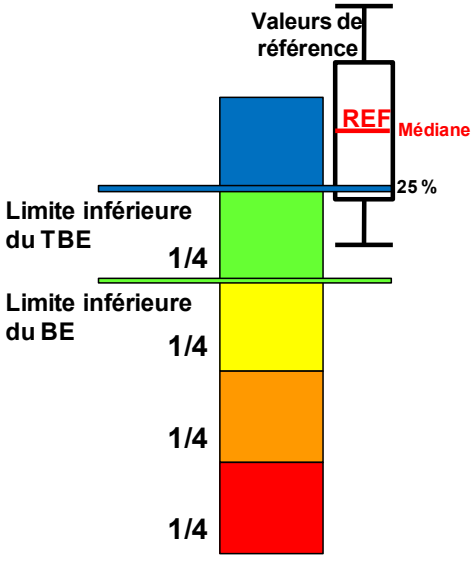


Figure 1.8 ***: Principe de calcul de la limite du bon état pour l'IBGN***

III.2.4.Premiers modèles d'extrapolation spatiale

Afin de représenter à l'échelle nationale l'état écologique actuel probable des cours d'eau, à partir de l'IBGN, des modèles simples d'extrapolation spatiale ont été réalisés (Wasson *et al.*, 2006). L'enjeu est de fournir, au-delà des stations de mesure de l'état écologique, une vision d'ensemble « approchée » et spatialisée de l'état des milieux aquatiques. Le principe de ces modèles est résumé Figure 1.9.

CemOA : archive ouverte d'Irstea / Cemagref

Ces premiers modèles sont limités à l'extrapolation sur le drain principal d'une zone hydrographique[12] et ont un caractère explicatif limité car le seul jeu de variables extrapolables utilisé correspond aux grandes catégories d'occupation du sol des bassins versants.

Afin de simplifier la procédure de modélisation, deux classes seulement ont été retenues pour la variable « état écologique mesuré par l'IBGN : (1) « bon état » (BE), pour les stations classées en bon ou très bon état écologique selon l'IBGN (couleurs bleue et verte) ; (2) « mauvais état » (ME) pour les stations des classes inférieures au bon état écologique selon l'IBGN (couleurs jaune, orange et rouge).

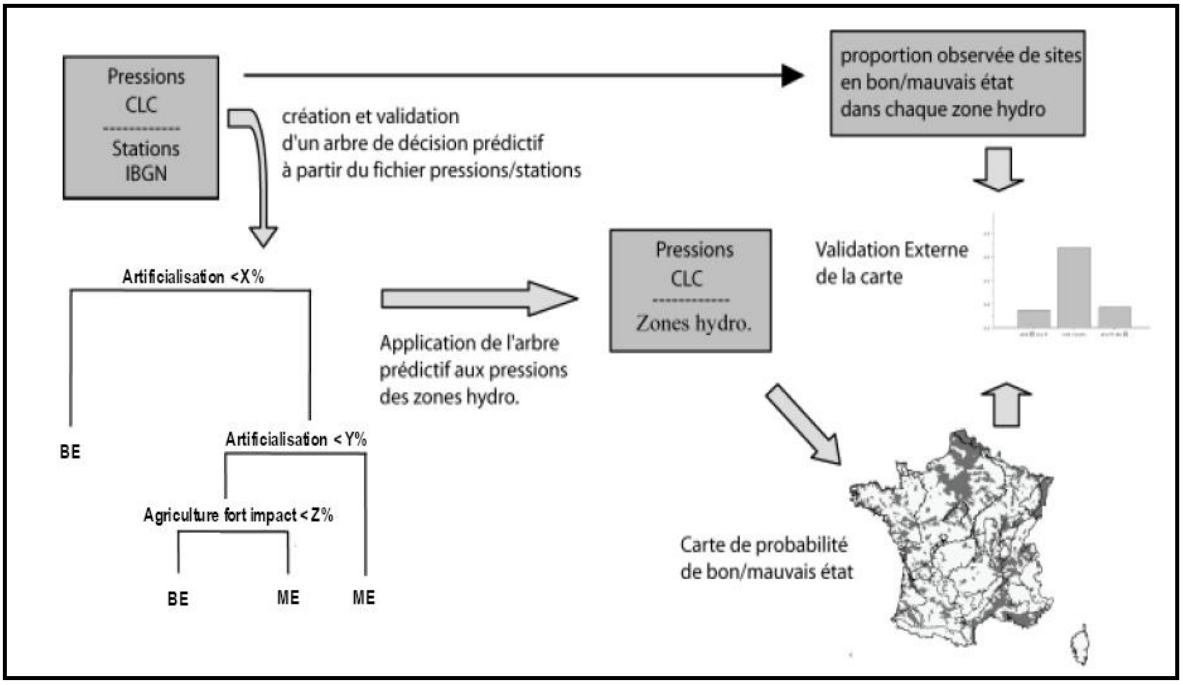


Figure 1.9 : Description du processus de modélisation et d'extrapolation spatiale qui se déroule en quatre temps (Wasson *et al.*, 2006) :
(1) un arbre de décision est développé et validé sur la base des variables d'occupation du sol calculées pour les bassins versants réels des stations IBGN ; (2) cet arbre de décision est ensuite utilisé pour prédire l'état écologique actuel probable (« bon » BE ou « mauvais » ME) des zones hydrographiques, d'après l'occupation du sol du bassin versant cumulée à l'amont de chaque zone hydrographique (ZH); (3) les sorties prédictives du modèle sont utilisées pour représenter sur une carte la probabilité actuelle de non atteinte du bon état pour le drain principal de chaque ZH ; (4) le modèle est validé de façon externe en comparant l'état des ZH « prédit » par le modèle et l'état « observé » sur des stations situées dans les mêmes ZH.

Dans ces modèles d'extrapolation spatiale, deux hypothèses ont été testées pour la limite du « Bon état écologique » mesurée avec l'IBGN :

- Une hypothèse « basse » (relativement arbitraire) correspondant à la limite du guide REFCOND méthode 3 (Wallin *et al.*, 2003).
- Une hypothèse un peu plus ambitieuse, appelée par commodité « Bon État (BE) +1 », a été calculée en relevant de 1 point IBGN (soit environ 6% pour l'EQR-IBGN) les notes de limite du Bon état correspondant à l'hypothèse de base.

[12] La circulaire n°91-50 du 12 février 1991 relative à la codification hydrographique et au repérage spatial des milieux aquatiques superficiels en France métropolitaine, définit la zone hydrographique comme suit : "L'ensemble du territoire français est divisé en zones élémentaires appelées zones hydrographiques. Leurs limites s'appuient sur celles des bassins versants topographiques (en tout ou partie)". Une zone est une partition d'un sous-secteur qui peut en comporter jusqu'à 10. Elle est entièrement comprise dans une limite hydrographique de bassin et sert, avec d'autres éléments, à la délimitation de zones de programmation ou réglementaires diverses comme les Schémas d'Aménagement et de Gestion des Eaux, les zones sensibles, les masses d'eau citées dans la Directive Cadre Européenne du 23 octobre 2000 (Sandre, 2008).

CemOA : archive ouverte d'Irstea / Cemagref

Les premiers résultats sur l'IBGN prédits à l'aide des modèles d'extrapolation spatiale (**Tableau 1-2**) montrent que pour l'ensemble du territoire métropolitain 38% des drains principaux des zones hydrographiques sont en « bon état » écologique suivant l'hypothèse haute et 36% sont en « mauvais état » écologique suivant l'hypothèse basse. Ainsi, 26% de drains principaux sont en situation limite car ils oscillent entre le « bon état » et le « mauvaise état » en fonction de l'hypothèse retenue. Les résultats IBGN au niveau des sites de surveillance (Tableau 1-2) montrent sensiblement la même situation (49% des stations sont en « bon état » écologique suivant l'hypothèse haute, 40% sont en « mauvais état » écologique suivant l'hypothèse basse, et donc 11% sont en situation limite). La situation générale des cours d'eaux en France métropolitaine n'apparaît donc pas exagérément dégradée tout en étant clairement préoccupante.

Tableau 1-2 : Résultats de l'état écologique IBGN au niveau des zones hydrographiques (prédit par les modèles d'extrapolation spatiale) et au niveau des réseaux de suivi selon les deux hypothèses de bon état (d'après Wasson *et al.*, 2006).

	État (modélisés) du drain principal des zones hydrographiques		État observé sur les stations du réseau de suivi	
	BE	ME	BE	ME
Hypothèse haute	**38 %**	**62 %**	**49 %**	**51 %**
Hypothèse basse	**64 %**	**36 %**	**60 %**	**40%**

Quelle que soit la limite qui sera choisie, des mesures généralisées sont donc nécessaires pour consolider cette situation qui semble fragile. Il faut être conscient qu'une dégradation légère mais continue pourrait faire basculer une grande partie du réseau en dessous de la limite fatidique du Bon État (Wasson *et al.*, 2005). Il est donc nécessaire de rechercher les causes d'altération de l'état écologique des cours d'eau en identifiant et hiérarchisant les principales pressions et de s'intéresser de plus près au rôle que jouent les corridors rivulaires pour mettre en place des mesures d'action prioritaires sur ces masses d'eau.

III.3. Développement en cours

Une évaluation de l'état écologique au sens de la DCE a donc été menée en France à partir de l'IBGN (cf. III.2). Cependant, cette étude, nécessaire pour donner une première vision de l'état des cours d'eau aux gestionnaires français, ne répond qu'en partie aux exigences de la DCE. D'une part l'IBGN n'est pas complètement DCE compatible, car il ne couvre pas l'ensemble des métriques exigées, et d'autre part l'évaluation ne rend pas compte de l'ensemble des pressions s'exerçant sur les milieux aquatiques puisque l'état des stations est vu uniquement sous le filtre des macroinvertébrés benthiques.

Des travaux menés notamment par AQUAREF[13] sont en cours à l'échelle de la France, en concertation avec les programmes européens, pour (1) se doter de méthodes de bioindication robustes et opérationnelles pour évaluer l'état écologique, (2) de disposer de métriques biologiques répondant spécifiquement à un grand nombre de pressions anthropiques afin de faciliter la recherche des causes d'altérations (amélioration de leur valeur de diagnostic) pour orienter les stratégies de restauration, et (3) harmoniser les méthodes au niveau européen.

Dès lors, en France de nouveaux réseaux ont été mis en place (circulaire DCE 2006/16[14]), pour le programme de surveillance de l'état des cours d'eau (e.g., Réseau de Contrôle de Référence, RCR : environ 400

[13] AQUAREF, laboratoire national de référence pour la surveillance des milieux aquatiques, est né de la nécessité de renforcer l'expertise française dans le domaine de la surveillance des milieux aquatiques, et notamment la surveillance biologique des cours d'eau, à partir de la mise en réseau des compétences et des capacités de recherche des cinq établissements publics directement concernés : BRGM, CEMAGREF, IFREMER, INERIS, LNE

[14] http://www.ineris.fr/aida/?q=consult_doc/consultation/2.250.190.28.8.2815

CemOA : archive ouverte d'Irstea / Cemagref

sites ; Réseau de Contrôle de Surveillance, RCS : environ 1500 sites…). Ces réseaux on été construits en enrichissant et en recadrant spatialement les réseaux existants afin de recouvrir une certaine diversité de situations.

Ensuite, le protocole IBGN a été, au niveau du prélèvement et du traitement des échantillons, modifié pour améliorer l'information contenue dans cet indice (Usseglio Polatera *et al.*, 2007). Il permettra, en plus de l'IBGN, d'évaluer des métriques plus fonctionnelles utilisant les traits bio/écologiques des organismes et des indices de structure des peuplements qui permettront la mise en œuvre d'un nouvel indice multi-métrique d'évaluation de l'état écologique conforme aux exigences de la DCE.

Une banque nationale pour les éléments de qualité biologiques, hydromorphologiques et physico-chimiques des cours d'eau (et des plans d'eau) est mise en place au niveau national par l'Office National de l'Eau et des Milieux Aquatiques (ONEMA) afin de constituer un outil rationnel et mutualisable de collecte, de validation et de conservation des données accessible à un large ensemble de partenaires et alimentant les portails publics.

Ces développements permettront d'améliorer l'évaluation de l'état écologique des cours d'eau au niveau national. Cependant, des progrès restent encore à faire sur la bioindication, notamment (i) sur la quantification des incertitudes que les différents indices véhiculent afin de mieux appréhender leur limites d'utilisation et d'interprétation ; et (ii) sur leur adaptation aux grands cours d'eau notamment pour les macroinvertébrés aussi bien au niveau du protocole de prélèvement que sur la constitution des indices (Archaimbault and Dumont, 2010).

CemOA : archive ouverte d'Irstea / Cemagref

Chapitre 2 :
Le rôle des formations végétales rivulaires

SOMMAIRE

Résumé *57*

I. UN MILIEU PARTICULIER AU SERVICE DE LA NATURE ET DE L'HOMME 61

I.1. UN COMPARTIMENT COMPLEXE DE L'HYDROSYSTEME *61*

I.2. UN MILIEU FAVORISANT LA BIODIVERSITE *61*

I.3. UN ESPACE UTILE POUR LA SOCIETE *63*

II. FACTEUR DE CONTROLE DE L'ECOSYSTEME RIVIERE 63

II.1. CREATION ET DIVERSIFICATION DES HABITATS *64*

II.2. CONTROLE TROPHIQUE *65*

II.3. REGULATEUR THERMIQUE *65*

III. UNE ZONE TAMPON REGULATRICE DES NUTRIMENTS 67

III.1. DES MECANISMES COMPLEMENTAIRES *67*

III.1.1. Le ralentissement des écoulements 67

III.1.2. L'infiltration dans les zones riveraines perméables 67

III.1.3. Une barrière physique 68

III.2. UN FILTRE SOUS CONDITIONS *69*

III.2.1. Les cheminements de l'eau 69

III.2.2. Variations spatiales de la régulation 69

IV. LES CORRIDORS RIVULAIRES : ESPACE CLE DE LA RESTAURATION MAIS ESPACE FRAGILISE 69

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Résumé

Les ripisylves sont des formations végétales bordant les cours d'eau qui assurent des fonctions multiples et variées au bénéfice de l'homme (e.g., espace de détente et de loisirs) et de la nature. Dans l'hydrosystème, les ripisylves forment le compartiment terrestre. Régies par la dynamique fluviale, elles constituent une mosaïque complexe qui varie dans le temps et dans l'espace le long du réseau hydrographique. Elles changent donc de visage périodiquement (elles rajeunissent) et constituent, à l'échelle du paysage, un linéaire (corridor rivulaire) de plus en plus large vers l'embouchure et une zone d'interface entre les milieux terrestre et aquatique. Cette dynamique spatiale leur permet d'assurer la double fonctionnalité d'écotone et de connexion dans l'hydrosystème qui favorise la biodiversité (terrestre et aquatique) à toutes les échelles (locales et régionales).

A l'échelle du tronçon, ces formations végétales rivulaires sont un compartiment clé de l'écosystème rivière puisqu'elles contrôlent et régulent les processus fonctionnels de cet écosystème. Au niveau physique : la ripisylve diversifie l'habitat aquatique par sa présence au bord du cours d'eau (structures racinaires et/ou végétation surplombant la rivière), et/ou en étant la source de débris ligneux grossiers (bois morts). Au niveau trophique : elle limite la production autotrophe par ombrage, elle injecte dans le système des détritus organiques, et elle filtre les apports en nutriments. Au niveau thermique : elle agit directement en faisant varier la surface exposée aux radiations et à l'énergie entrante.

Ces formations contribuent donc au bon fonctionnement de l'écosystème rivière et assurent ainsi le maintien de la biodiversité aquatique tout en la protégeant. En effet, les ripisylves jouent aussi un rôle de zone tampon (de barrière naturelle), aussi bien latérale que longitudinale, contre les polluants (e.g., nitrates, phosphate, pesticides, matières en suspension) issus des bassins versants agricoles et urbains grâce à des mécanismes complémentaires. La rugosité hydraulique des berges (liée particulièrement à la végétation herbacée et arbustive) favorise le dépôt de sédiments où une quantité importante de nutriments est adsorbée. Ces nutriments ainsi accumulés peuvent être recyclés par le sol et/ou la végétation. La perméabilité des sols riverains permettant les conditions d'anaérobie nécessaires à la dégradation des nutriments par les micro-organismes.

L'ensemble de ces fonctions, démontrées généralement à une échelle locale, est garant de la biodiversité de l'écosystème rivière condition essentielle de sa résilience. Le corridor rivulaire, emprise gérable à l'échelle d'un bassin versant et abritant cette végétation, constitue ainsi un espace clé de la restauration de l'état écologique des cours d'eau qui est déjà fortement fragilisé par la présence caractérisée des activités humaines. Cependant la mise en place d'une politique ambitieuse sur ces espaces nécessite une compréhension de ces mécanismes à des échelles plus larges notamment en se dotant d'indicateurs pertinents permettant de dresser l'état de ces corridors et de suivre les effets des actions entreprises.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

De nombreux documents alimentent la littérature concernant les formations végétales rivulaires. Ce chapitre a été conçu principalement à partir (i) des synthèses bibliographiques de H. Piegay & L. Maridet («Formations végétales arborées riveraines des cours d'eau et potentialités piscicoles », 1994) dans le bulletin français de la pêche et de la pisciculture et de L. Maridet (« Rôle des formations végétales riveraines. Recommandations pour une gestion régionalisée », 1995) réalisées pour le compte de la Direction de l'Eau du Ministère de l'Environnement (convention n° 1/94), (ii) des numéros spéciaux dans la revue Forêt Méditerranéenne coordonnés par H. Décamps (Ripisylves: la biodiversité par l'instabilité, 2003) et du Conseil Supérieur de la Protection de la Nature et de la Biodiversité (L'arbre, la rivière et l'homme, 2008), puis (iii) des ouvrages de H. Piegay et al. («Les forêts riveraines des cours d'eau: Ecologie, fonctions et gestion, 2002) et de Naiman et al. (Riparia : ecology, conservation, and management of streamside communities », 2005).

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

I. Un milieu particulier au service de la nature et de l'homme

I.1. Un compartiment complexe de l'hydrosystème

Plusieurs termes (aussi bien dans la littérature technique et scientifique) identifient les formations végétales rivulaires : ripisylves (terme le plus commun), forêt alluviale ou forêt inondable (pour les grand cours d'eau), corridor forestier ou corridor rivulaire, franges forestières, linéaire arboré, milieux ripuaires ou ripariaux (« riparian » en anglais)… tous désignent une formation végétale multi-stratifiée, où domine généralement l'arbre, en relation avec un cours d'eau mais aussi une zone humide, un marais, un lac… Cette multiplicité des termes démontre l'intérêt que portent beaucoup d'acteurs, qu'ils soient écologues, géographes, naturalistes ou gestionnaires, chacun définissant son propre questionnement (Piégay *et al.*, 2003). Dans la thèse on utilisera plutôt le terme de ripisylves lorsque l'on s'intéressera aux formations végétales bordant le cours d'eau et de corridor rivulaire en ce qui concerne l'espace où se développe cette végétation.

Dans l'hydrosystème, la ripisylve forme le compartiment « terrestre » localisé sur ses marges, ou en son cœur dans le cas des îles. Elle se développe sur les sols alluviaux (mosaïque sédimentaires récentes) situés en fond de vallée. Ces sols sont sous la dépendance des cours d'eau ; la dynamique fluviale (moteur de l'hydrosystème) conditionnent la mise en place de cette mosaïque sédimentaire de façon : (1) mécanique par l'intensité et la périodicité des crues (érosion et/ou dépôt d'alluvion) et par les fluctuations du niveau de la nappe (conditions hydrique des sols) ; et/ou (2) chimique par le biais de sa composition en élément dissous (dépendante de la lithologie du bassin versant).

L'eau dans sa dynamique spatio-temporelle régit donc la distribution des communautés végétales riveraines. Ainsi, il existe de manière générale une variation longitudinale (Rot *et al.*, 2000; Catterall *et al.*, 2001) de la composition des ripisylves qui se traduit notamment par un tri granulométrique des dépôts alluvionnaires de la source (blocs, galets, graviers) à l'embouchure (sables, limons voir argiles) d'un cours d'eau ; se rajoutent bien sûr les facteurs déterminants du relief, du climat et de la géologie. Des variations transversales (Hupp and Osterkamp, 1985, Hupp and Osterkamp, 1996, Rot *et al.*, 2000) s'observent également du chenal principal vers les milieux connexes : la taille moyenne des alluvions diminue proportionnellement à la diminution de l'énergie véhiculée par les eaux d'inondation du chenal vers la limite externe de la plaine d'inondation. Ce schéma général peut être restructuré par des variations locales (modelés du terrain ou action de la végétation qui modifie l'énergie véhiculé par le flux d'eau). Aux gradients longitudinaux et transversaux dans l'espace, il faut ajouter les variations temporelles de l'hydrodynamique et du transport sédimentaire qui règlent les conditions d'érosion et de sédimentation ; et ces variations se traduisent par le décapage plus ou moins important de certains dépôts sédimentaires ou par la superposition progressive de couches sédimentaires de qualité et d'épaisseurs variables qui génèrent ainsi un gradient vertical (Guenat *et al.*, 2003).

Cette dynamique spatiale et temporelle dans les quatre dimensions de l'hydrosystème est à l'origine d'une mosaïque végétale complexe (Piégay *et al.*, 2003) qui est périodiquement renouvelée. Les crues et les phénomènes d'érosion et de transport qui exportent les débris végétaux vers l'aval génèrent des chablis et favorisent la présence de stades de successions variées sur des espaces restreints. Ces rajeunissements sont la condition même de la coexistence de fonctions multiples au sein des ripisylves (Décamps, 2002).

I.2. Un milieu favorisant la biodiversité

A l'échelle du paysage aquatique, la ripisylve forme un corridor (un linéaire) rivulaire continu (Forman, 1983) de largeur variable qui se complexifie vers l'aval au fur et à mesure que les débits augmentent et que le chenal et le fond de vallée s'élargissent. Il constitue aussi une zone écotonale, une interface entre le milieu terrestre et aquatique (Naiman and Décamps, 1990) régulant les flux d'énergie et de matières latéraux entre

CemOA : archive ouverte d'Irstea / Cemagref

ces deux compartiments. Ainsi, par leur positionnement dans l'espace, les corridors rivulaires déterminent deux propriétés écologiques caractéristiques : la fonction d'écotone et la fonction de connexion (Figure 2.1). Cette double fonctionnalité favorise la biodiversité (terrestre et aquatique) à toutes les échelles (locales et régionales).

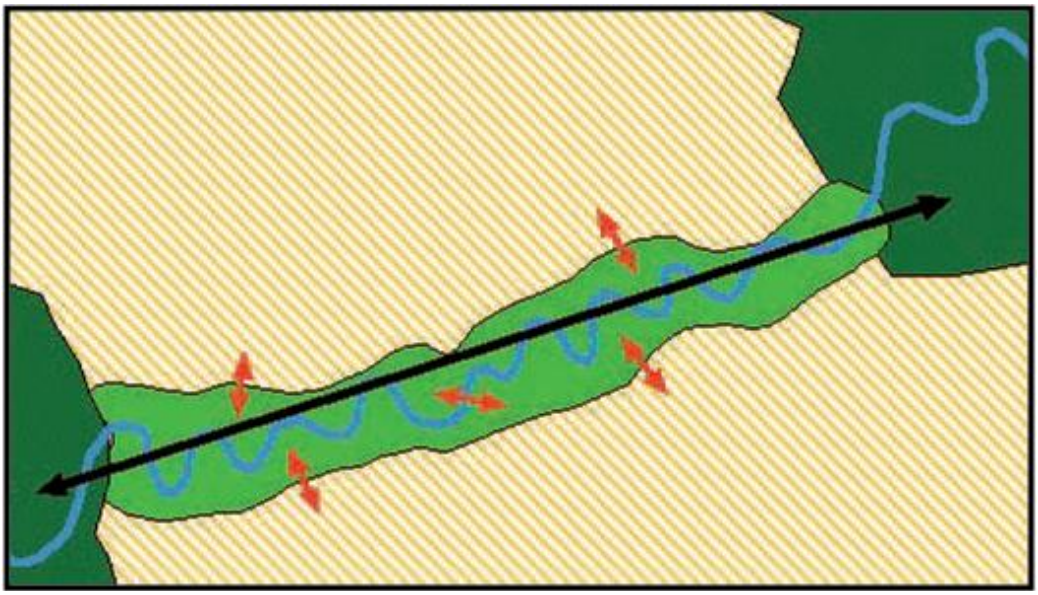


Figure 2.1 : Les corridors rivulaires remplissent deux fonctions essentielles pour la protection de la biodiversité : dans le sens longitudinal (flèche noire), ils mettent en connexion des habitats éloignés, comme deux massifs forestiers ; dans le sens transversal (flèches rouges), ce sont des écotones entre des formations boisées et ouvertes, et entre le milieu aquatique et le milieu terrestre (CSPNB, 2008)

La fonction d'écotone résulte du contact entre deux milieux aux caractéristiques écologiques fortement contrastées, le milieu aquatique et le milieu terrestre. Cet écotone présente un intérêt majeur pour la biodiversité : c'est un site de forte productivité.

Il abrite, à l'échelle locale, une mosaïque changeante complexe composée d'habitats aquatiques, semi-aquatiques et terrestres particulièrement favorable à la biodiversité. D'une part de nombreuses espèces animales (e.g., les annélides, les mollusques, les insectes et les araignées pour les invertébrés ; les tritons et rainettes pour les vertébrés) et végétales (e.g., les saules, les peupliers, les aulnes, pour les bois tendres, le frêne, le pin sylvestre, les érables et les chênes pour les bois durs) trouvent dans ce milieu particulier un habitat à part entière leur permettant d'accomplir tout leur cycle vital. D'autre part, des espèces y trouvent un habitat complémentaire (des refuges, des sites de ponte, d'hibernation ou de nymphose) nécessaire pour accomplir certaines fonctions vitales: c'est la « complémentarité de niche » (CSPNB, 2008).

Cette richesse particulière est favorisée aussi par d'autres facteurs. Pour les animaux, par exemple, l'accès à l'eau est aussi une cause d'attraction importante. Pour les végétaux, la biodiversité est déterminée également par la présence simultanée d'eau, de lumière et de nutriments qui favorise leur croissance; et la variabilité des conditions des cours d'eau de la tête de bassin jusqu'aux grands fleuves élève cette richesse à l'échelle du réseau hydrographique.

Ainsi, on trouve dans le corridor rivulaire de l'Adour près de 1400 espèces de végétaux vasculaires, soit 30 % de la flore de France (Planty-Tabacchi *et al.*, 1996) ; et dans le bassin du Nord-Ouest des États-Unis, les habitats boisés rivulaires hébergent près de 290 des 490 espèces sauvages recensées de mammifères, oiseaux, amphibiens et reptiles de la région, et pour 68 de ces espèces, les aires riveraines sont indispensables à la satisfaction d'un besoin vital pendant tout une partie de l'année (Kelsey and West, 1998). Dans les sols riverains (moins connus) également, on retrouve une diversité apparemment impressionnante (Décamps, 2008) : environ 10 000 espèces de champignons ecto-mycorhiziens, plus de 3 000 espèces de bactéries, plus de 5 000 espèces de nématodes, et des dizaines de milliers d'espèces pour la méiofaune (acariens, collemboles) et la macrofaune (fourmis, termites, lombrics).

La fonction de connexion est déterminée par la structure linéaire qui favorise la dispersion, y compris d'espèces qui ne sont pas inféodées à l'écotone. Les corridors rivulaires mettent en communication des

CemOA : archive ouverte d'Irstea / Cemagref

habitats éloignés qui sont ainsi reliés par ces structures linéaires couvrant potentiellement tout le territoire, et traversent les zones fortement anthropisés. Ils constituent ainsi un élément crucial pour le fonctionnement des « métapopulations », c'est-à-dire de l'ensemble des populations spatialement isolées d'une même espèce et favorisent ainsi la biodiversité à l'échelle régionale (Décamps, 2002).

Par exemple, les populations d'amphibiens sont structurées en petits groupes de reproducteurs (des populations locales) dont la viabilité est très faible. La persistance des populations dépend fortement de la dispersion. Les corridors rivulaires peuvent contribuer de façon décisive à cette dispersion et au maintien de cette biodiversité fragile (CSPNB, 2008). Ils jouent aussi un rôle de fil conducteur pour de nombreuses espèces migratrices qui les utilisent lors de leurs migrations d'automne et de printemps, surtout lorsque les rivières sont orientées Nord-Sud (Blondel, 2003) ; ils sont largement utilisés comme voies de migration par les oiseaux forestiers dans les espaces très ouverts (CSPNB, 2008).

Ce couloir de vie « terrestre », par cette double fonctionnalité (écotone et connexion), favorise et maintient ainsi une biodiversité remarquable à toutes les échelles ; mais il préserve et contrôle aussi en son sein un autre couloir abritant la vie « aquatique » : les rivières.

I.3.Un espace utile pour la société

Les ripisylves assurent des fonctions multiples et variées bénéfiques pour la société qui se répercutent directement au niveau économique, en constituant une matière première pour le bois de chauffage par exemple, mais aussi indirectement (1) en protégeant l'homme et ses activités et (2) en lui procurant un meilleur cadre de vie.

(1) La présence d'arbres à enracinement profond favorise la stabilisation des berges, qui freine le ruissellement de surface et, par la suite, diminue l'érosion des terres et augmente leur résistance aux affouillements. Les ripisylves contribuent ainsi à la protection des routes, des ponts et des bâtiments (Lâchât, 1991). Ce frein à l'écoulement est accentué aussi en surface : les parties aériennes des végétaux permettent d'augmenter la rugosité hydraulique de la surface du sol. Une zone rivulaire composée de végétaux buissonnants et arbustifs denses, dont les parties aériennes possèdent une grande élasticité, est très efficace pour réduire la vitesse du courant et donc la puissance érosive de l'eau lors des crues (Lâchât, 1994). Les ripisylves constituent ainsi des barrières physiques capables de protéger du courant ou de l'érosion les infrastructures et les cultures (tolérant l'immersion) lors des inondations. Par ailleurs, elles protègent aussi les cultures en formant des paravents naturels, évitant ainsi l'action directe du vent et elles servent également d'abris au bétail pour qu'il se protège de la chaleur et du froid (Décamps, 2002).

(2) Le corridor rivulaire boisé est généralement perçu comme une composante forte du paysage visuel (Dupuis and Fischesser, 1997 ; CSPNB, 2008). Dans le paysage structuré par les cours d'eau – physiquement parce qu'ils creusent les vallées et socialement par l'organisation des activités humaines autour d'eux – la ligne de saules, la berge boisée ou le ruban d'une ripisylve naturelle signalent la rivière et constituent l'élément vertical qui tranche dans la monotonie des cultures. Mais les bords de rivière ne sont pas seulement vus, ils sont aussi vécus. Ils forment un lieu de détente, de promenade et de loisirs améliorant le cadre de vie des populations et favorisent ainsi l'activité touristique (Dupuis and Fischesser, 1997).

II. Facteur de contrôle de l'écosystème rivière

Si les conditions hydrodynamiques et morphodynamiques régissent la répartition des communautés végétales, réciproquement la végétation riveraine, par sa présence effective ou par ses apports en débris ligneux grossiers, est elle-même facteur de contrôle de l'écosystème rivière au niveau physique (habitat), trophique et thermique à l'échelle du tronçon (Figure 2.2) assurant la qualité et le bon état de ce système. L'influence des ripisylves sur le fonctionnement écologique varie en fonction de leur morphologie (largeur, pente des rives) et du port des végétaux rivulaires (Maridet, 1995). La végétation sera d'autant plus influente que le cours d'eau est de faible largeur. Ainsi sur les cours d'eau larges (ordre de 5 à 7), seuls les arbres ne pourront influencer que faiblement le fonctionnement, alors que sur les cours d'eau moyens (ordre de 3 à 4) des taillis suffiront et leur influence se fera ressentir jusqu'au milieu du chenal.

CemOA : archive ouverte d'Irstea / Cemagref

II.1.Création et diversification des habitats

La ripisylve complexifie l'habitat aquatique, elle crée un milieu diversifié réduisant ainsi la compétition intra et interspécifique (Gorman and Karr, 1978, Bisson *et al.*, 1981) qui est propice au développement des communautés biologiques. Cette influence se manifeste directement par sa présence au bord du cours d'eau, et/ou indirectement en étant la source de débris ligneux grossiers (bois morts).

A l'échelle de la berge, la végétation riveraine, les aulnes et les saules en particulier (Maridet and Souchon, 1995) influence l'habitat des salmonidés en favorisant les couverts soit directement par les structures racinaires ou par la végétation surplombante (Boussu, 1954, Meehan *et al.*, 1977; House and Boehne, 1986, Hunter, 1991). Ces formations jouent un rôle d'abri face à des conditions climatiques et hydrologiques difficiles (Bisson *et al.*, 1987). Les racines sont bien sûr des caches et des supports de ponte pour les poissons, mais surtout elles stabilisent la berge approfondissant le chenal favorable au poisson (particulièrement aux gros individus) au fil de l'écoulement.

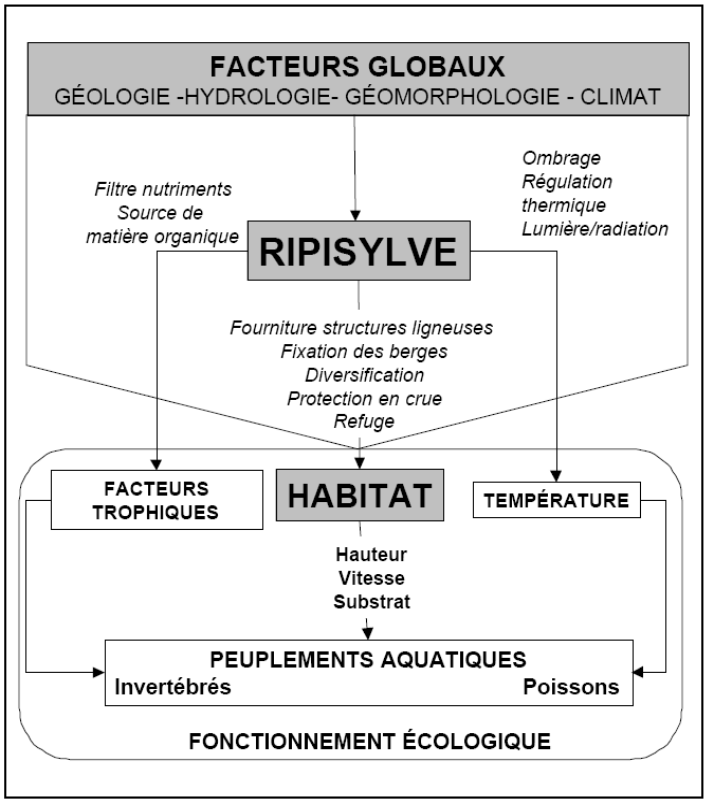


Figure 2.2 : Influence de la ripisylve sur le fonctionnement écologique des écosystèmes d'eau courante. Fonctions assurées par la ripisylve. Effet spécifique sur l'habitat aquatique (d'après Maridet, 1995).

A l'échelle du lit majeur, la connexion forêt alluviale-chenal principal et le gradient latéral de ces formations sont deux fonctions importantes pour la migration des poissons (Sioli, 1982, Naiman and Décamps, 1990) mais aussi pour la reproduction (zone de frayère) (Corp *et al.*, 1992).

La ripisylve agit également directement sur l'éthologie des poissons. De nombreux auteurs ont constaté chez un grand nombre d'espèces de salmonidés, un préférendum en période de repos pour des secteurs ombragés (Brusven *et al.*, 1986, Meehan *et al.*, 1987). Par contre, la recherche de nourriture et l'alimentation sont facilitées dans les secteurs à canopée ouverte (Wilzbach, 1985).

Les débris ligneux et les microenvironnements qui leur sont associés constituent eux aussi un habitat pour beaucoup d'organismes (Harmon *et al.*, 1986, Gregory and Davis, 1992, O'Connor, 1992). Certains en sont strictement dépendants (e.g., organismes creuseurs) , d'autres bénéficient de leur opportunité (e.g., organismes fixateurs).

CemOA : archive ouverte d'Irstea / Cemagref

Ils augmentent la qualité de l'habitat piscicole en complexifiant et diversifiant le milieu (Franklin, 1988, Fauch and Northcote, 1992). En effet, les modifications hydrauliques liées à la présence d'embâcles, caractérisées par des vitesses locales, des sens d'écoulement, des hauteurs d'eau et des compositions granulométriques variables, créent une grande variabilité physique d'habitats (Maridet, 1994).

A l'échelle très locale, dans certaines grandes rivières, la densité de salmonidés et d'invertébrés est la plus importante dans les secteurs où s'accumulent les débris ligneux grossiers (Ward *et al.*, 1982; Sedell *et al.*, 1984).

A l'échelle d'un tronçon de rivière, les structures ligneuses constituent des points durs qui vont faire diverger le courant, créer des bancs, des chenaux annexes, de nouveaux bras, et donc diversifier les écoulements (Bisson *et al.*, 1987, Sedell and Luchessa, 1982, Sedell *et al.*, 1987). Sur les grands cours d'eau, c'est aussi un processus essentiel de structuration de la forêt riveraine qui favorise la diversité végétale. Cette fonction est particulièrement importante sur les rivières actives, à fond mobile de sable ou de graviers, comme il en existe dans de nombreuses régions de piedmont.

II.2.Contrôle trophique

Les ripisylves contrôlent la chaîne trophique à sa base en limitant la production autotrophe par ombrage, en injectant dans le système des détritus organiques, en filtrant les apports en nutriments, et en régulant la température (Figure 2.3).

Les ripisylves agissent directement sur cette chaîne en régulant la quantité de lumière qui pénètre dans le cours d'eau. La lumière et la vitesse du cours d'eau sont les facteurs limitant de la production primaire (autotrophe) de l'écosystème. Ainsi, à court terme, un milieu ouvert est plus productif qu'un milieu fermé (Hawkins *et al.*, 1983, Behmer and Hawkins, 1986).

Proches des rivières, les formations végétales arborées riveraines constituent une source de matière organique allochtone (feuilles, tiges, débris ligneux grossiers…) et des structures de rétention nutritionnelle permettant le stockage, le recyclage et le relargage vers l'aval, des éléments nutritifs (Maridet, 1995). La nature des espèces constitutives de la végétation arborée influence la qualité de ces apports allochtones (e.g., les aiguilles de résineux sont beaucoup moins dégradables que les feuilles tendres d'aulnes et saules) de même que la capacité de rétention du cours d'eau (e.g., les espèces collinéennes, hêtres et chênes, ne sont pas susceptible de créer grâce à leur racines et branches submergées des structures de rétention par rapport aux espèces alluviales, aulne et saule, Maridet, 1994; Maridet and Souchon, 1995).

De plus, les débris ligneux accentuent ce phénomène en formant des pièges potentiels et des structures de rétention des particules détritiques (tiges et feuilles) (Bryant, 1981, Prochazka and Stewart, 1991). Ces amas de débris organiques servent de nourriture, de site de ponte, de croissance, de repos et de refuge pour les communautés benthiques (Meehan *et al.*, 1977, Triska and Cromack, 1980, Angermeier and Karr, 1984, Wallace and Benke, 1984).

Les feuilles qui tombent dans l'eau et les bois morts conditionnent donc la présence d'un grand nombre d'espèces, surtout des invertébrés, qui sans ces apports disparaîtraient de la rivière. La densité et la nature de l'environnement végétal en bordure de cours d'eau contrôlent alors le fonctionnement trophique des milieux aquatiques, en déterminant le type de nourriture disponible. Dans ce sens, la fonction « trophique » de la ripisylve est aussi un facteur important de biodiversité, qui ne se limite pas aux petits cours d'eau : la faune de toutes les zones de bordure des rivières et des fleuves en bénéficie également (CSPNB, 2008).

II.3.Régulateur thermique

La végétation rivulaire contrôle également la température, en faisant varier la surface exposée aux radiations et à l'énergie entrante. Quelques kilomètres de corridor rivulaire boisé, même de largeur limitée, peuvent réduire de 2° à 4°C la température estivale (Figure 2.4). La température est un paramètre clé (facteur limitant) du fonctionnement de l'écosystème rivière : (1) elle conditionne la quantité d'oxygène dans l'eau nécessaire pour la respiration des animaux (et des végétaux la nuit), mais aussi des bactéries et champignons qui décomposent la matière organique ; (2) elle agit directement sur le cycle vital des organismes (Quinn *et al.*, 1992), la croissance et la reproduction des poissons, le développement des insectes étant régulé, entre autres,

CemOA : archive ouverte d'Irstea / Cemagref

par la température ; et (3) elle influe indirectement sur la qualité nutritive de la nourriture disponible (Fuller and Fry, 1991, Maridet and Piegay, 1994).

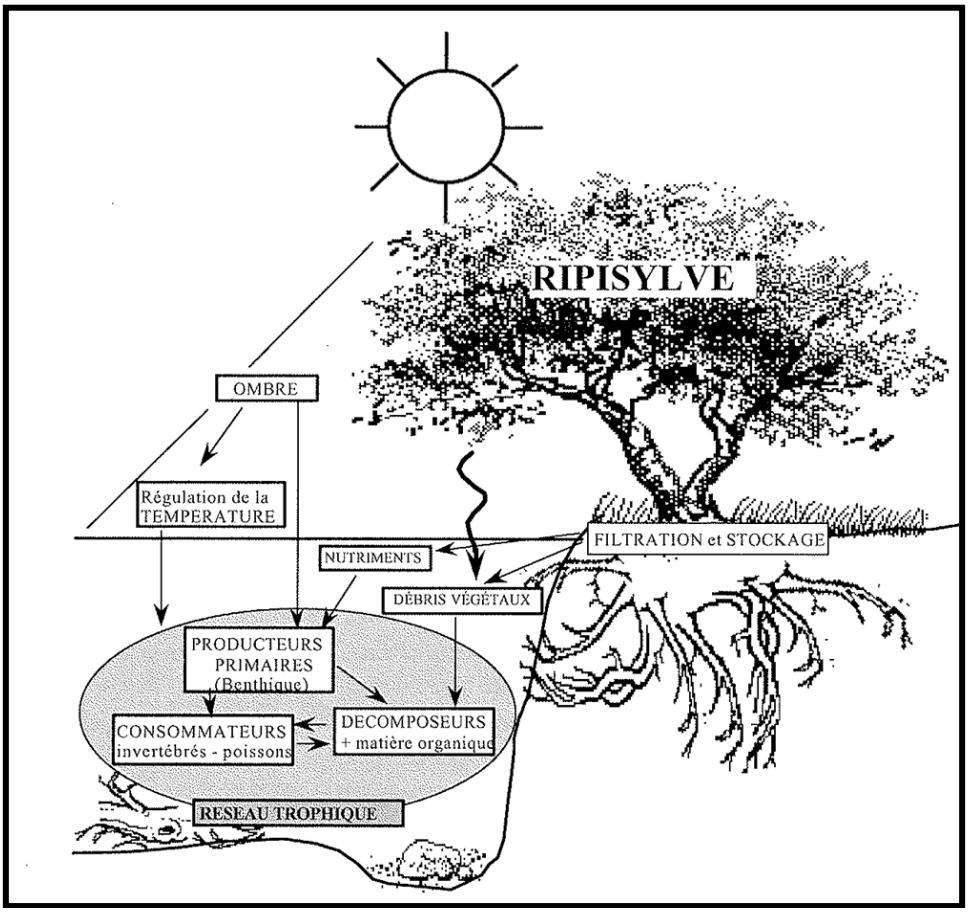


Figure 2.3 : Influence de la ripisylve sur le fonctionnement trophique dans les écosystèmes d'eau courante (d'après Maridet, 1995).

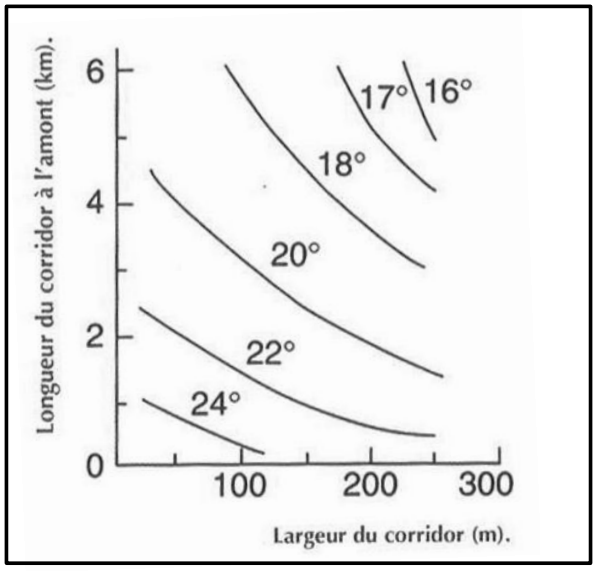


Figure 2.4 : Influence d'un corridor rivulaire boisé sur la température maximale de petits cours d'eau, en fonction de la largeur du corridor et de la longueur protégée à l'amont du site ; d'après l'étude de 40 sites en Ontario, Canada (d'après Naiman *et al.*, 2005).

CemOA : archive ouverte d'Irstea / Cemagref

III. Une zone tampon régulatrice des nutriments

La popularité croissante des ripisylves en matière d'aménagement tient surtout au rôle épurateur qu'on leur prête le long des rivières (Décamps, 2002). La végétation rivulaire a été reconnue tout d'abord comme une zone tampon efficace contre les apports de sédiments dans le milieu des années 60 (Haupt and Kidd, 1965, Wilson, 1967) ; puis dans le milieu des années 70 jusqu'au début des années 80, de nombreux travaux ont démontrés leur efficacité à retenir les nutriments provenant des versants agricoles (e.g., Gambrell *et al.*, 1975, Young *et al.*, 1980; Lowrance *et al.*, 1984). Plus récemment des auteurs ont démontré aussi ce rôle dans les milieux urbains (Booth, 1991 ; Schuler, 1995 ; Paul and Meyer, 2001). Cette capacité de filtre est liée à leur positionnement dans l'espace. A l'interface entre les milieux terrestres et aquatiques, les zones riveraines interceptent pratiquement tous les écoulements superficiels ou peu profonds qui ruissellent vers les cours d'eau (Naiman and Décamps, 1997).

III.1.Des mécanismes complémentaires

L'efficacité des corridors rivulaires repose essentiellement sur deux mécanismes: (i) le ralentissement de l'écoulement et (ii) l'infiltration, la perméabilité des sols étant souvent importante dans les zones riveraines. Ces deux mécanismes se complètent et interagissent, avec une influence variable selon les polluants considérés, dont les dynamiques de transfert ne sont pas identiques. Lorsque l'écoulement provient des ruissellements des versants on parle de filtre « latéral », lorsqu'il est dû aux crues débordantes on parle de filtre « longitudinal ». Un troisième mécanisme, lié à la présence physique de la végétation arborée, œuvre à cette efficacité mais dans une moindre mesure.

III.1.1.Le ralentissement des écoulements

Le ralentissement des écoulements est directement lié à la pente et à la « rugosité » hydraulique des berges et des rives, fortement augmentée par la présence de plantes herbacées et buissonnantes. Ces plantes, par leur position, créent un tapis qui en augmentant la rugosité à la surface du sol entraîne un ralentissement de la vitesse du ruissellement (écoulement de surface) et donc une sédimentation des particules (Maridet, 1994). La longueur de la zone filtrante nécessaire pour obtenir une sédimentation maximale sera d'autant plus importante que les particules seront fines (e.g., alors qu'il faut une distance de 3 m pour les sables, une distance de 122 m est nécessaire pour les argiles, Barling and Moore, 1994).

Ces sédiments fins transportés contiennent des quantités importantes de phosphore, de pesticides (issus des produits phyto-sanitaires) et dans une moindre mesure de l'azote, sous forme organique particulaire ou minérale et de débris organiques. Ces particules sont adsorbées par les sédiments et peuvent ainsi être recyclées (métabolisées) par le sol et par la végétation. Vought *et al.*, 1994 montre que lors du passage des eaux de surface dans les zones riveraines, 66 % des phosphates disparaissent après 8 m et 95 % sont retenus après 16 m, alors que respectivement seulement 20 et 50 % des nitrates sont éliminés (le transfert des nitrates s'effectue majoritairement sous une de ces formes minérale dissoutes).

Au total des quantités massives de matières s'accumulent dans les zones riveraines. Cependant, ces accumulations peuvent menacer à long terme la qualité des eaux. Les flux souterrains peu profonds peuvent les transporter vers les milieux aquatiques sous d'autres formes (oxydées ou organiques), par exemple en conséquence de décompositions accrues de matière organique (Décamps, 2006).

III.1.2.L'infiltration dans les zones riveraines perméables

Ce mécanisme concerne principalement la réduction de l'azote, du fait de son mode de transfert dans le sol (écoulement de subsurface). Il régit l'expression de deux processus permettant l'élimination des nitrates dans les eaux souterraines : la dénitrification microbienne, et le prélèvement par les racines des végétaux (Pinay, 1986 ; Haycock *et al.*, 1993) (Figure 2.5).

CemOA : archive ouverte d'Irstea / Cemagref

En période de hautes eaux, l'engorgement du sol créé des conditions anaérobies propices à la dénitrification microbienne. Le carbone, abondant dans ces zones (feuilles, branches, troncs en décomposition) constitue l'apport énergétique et nutritionnel nécessaire aux bactéries dénitrifiantes. Ces bactéries réduisent les nitrates (NO3) en prélevant l'oxygène des molécules pour leur synthèse carbonée, et élaborent en retour de l'azote gazeux (N2) qui se disperse dans l'atmosphère.

En période de croissance végétale, le processus de dénitrification est relayé par l'absorption directe des nitrates par le système racinaire des plantes. Les nitrates sont alors incorporés et stockés dans les tissus sous forme d'azote réduit, constituant l'essentiel des structures végétales. L'efficacité de ce processus au niveau des ripisylves est expliquée par le grand développement des systèmes racinaires des espèces ligneuses par rapport aux autres types de couverts végétaux et par leur âge compte tenu de leur dynamique.

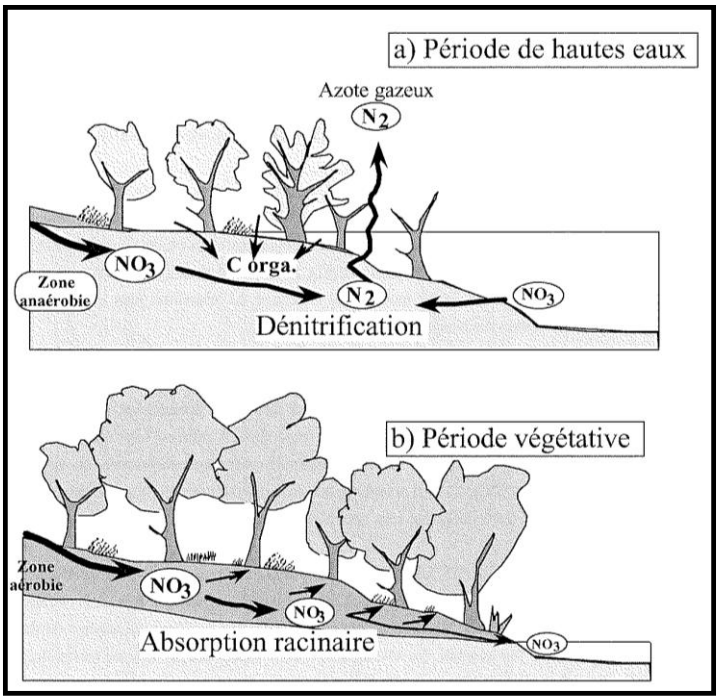


Figure 2.5 : Principaux processus d'épuration des eaux polluées en nitrates par les ripisylves : (a) en période de hautes eaux : dénitrification bactérienne anaérobie ; et (b) en période de croissance végétale : absorption racinaire (d'après Maridet, 1995).

Ainsi, l'épuration des nitrates dans les eaux souterraines peut être quasi totale : on a pu observer, par exemple au bord de la Garonne, une élimination de plus de 99% des nitrates après un transit des eaux dans une forêt alluviale de 30 m de large (Pinay and Décamps, 1988).

Néanmoins, la biomasse végétale finira par s'accumuler dans les sols et des quantités considérables d'azote peuvent être relâchées des zones riveraines vers les rivières sous la forme de matière organique dissoute, une voie dont on connaît mal l'importance et les conséquences (Décamps, 2006).

III.1.3.Une barrière physique

Les ripisylves peuvent aussi constituer une « protection rapprochée » des cours d'eau contre toute émission provenant des champs les plus proches – ne serait-ce que par l'éloignement qu'elles imposent. Par exemple, l'éloignement conditionne une autre fonction importante qui concerne spécifiquement les pesticides : l'interception directe par un écran végétal des gouttelettes issues de la pulvérisation des produits phytosanitaires (la « dérive »). Ou encore, les ripisylves facilitent le contrôle de l'accès du bétail aux berges des cours d'eau, évitant à la fois le piétinement, l'érosion, le rejet direct de déjections dans l'eau, voire les chutes d'animaux susceptibles de s'aventurer sur des berges trop pentues (Décamps, 2002).

CemOA : archive ouverte d'Irstea / Cemagref

III.2.Un filtre sous conditions

III.2.1.Les cheminements de l'eau

L'efficacité de piégeage des sédiments est liée aux caractéristiques topographiques du milieu. Des secteurs vallonnés et accidentés où l'écoulement se concentre dans des chenaux de drainage, augmentant ainsi la vitesse d'écoulement et diminuant le temps de résidence de l'eau dans ces zones, sont reconnus comme inefficaces (Barling and Moore, 1994). Cependant, ces secteurs, inondant les zones rivulaires, favorisent la perméabilité, autre mécanisme responsable de l'efficacité des corridors rivulaires.

Cette hydromorphie détermine les conditions d'anaérobiose propices à la dénitrification microbienne. Celle-ci dépend aussi des modalités du cheminement de l'eau à travers les zones riveraines (Lowrance *et al.*, 1995). Le rôle de filtre vis-à-vis des nitrates n'est tenu efficacement que dans le cas où les écoulements sont peu profonds, baignant ainsi les racines. Il l'est bien moins quand l'eau s'écoule en profondeur, passant sous les racines (cas d'un aquifère profond).

III.2.2.Variations spatiales de la régulation

L'hydromorphologie est un facteur important qui contrôle le pouvoir épurateur de ces milieux puisqu'il conditionne les modalités d'écoulement au sein des ripisylves. La localisation géographique des ripisylves va donc influer sur les mécanismes mis en œuvre, compte tenu du gradient longitudinal de ce facteur (Figure 2.6).

Ainsi, en tête de bassin hydrographique, la majeure partie des apports d'eau entrant dans une ripisylve sera à dominante d'écoulements de subsurface (apport d'azote sous forme de nitrate dissous) provenant de son bassin versant adjacent (filtre latéral).

Dans la partie aval des réseaux hydrographiques, la majeure partie des apports d'eau dans la ripisylve se fera par l'intermédiaire du cours d'eau lors des débordements de crues (filtre longitudinal), même s'il existe des échanges souterrains entre le cours d'eau et la ripisylve à travers la berge (Garcia *et al.*, 1994 ; Pinay *et al.*, 1998). Les apports de nutriments seront alors essentiellement particulaires liés aux sédiments déposés.

Toutes ces fonctions sont d'autant plus efficaces que l'interaction est plus forte entre le cours d'eau et son environnement végétal. La longueur d'interface des grands cours d'eau est très faible comparée à celles de leur petits affluents. Ces petits cours d'eau (<8 m) représentent 80% des 500 000 kilomètres du réseau hydrographique français. Autrement dit, c'est en tête de bassin que les processus régulateurs de la qualité de l'eau sont les plus actifs (CSPNB, 2008).

Les performances des zones tampons rivulaires varient donc notablement en fonction de leurs caractéristiques propres, du polluant incriminé et des nombreux paramètres qui conditionnent l'intensité des écoulements latéraux provenant des versants et longitudinaux dus aux crues de débordements (climat, géologie, pédologie, topographie, usage du sol, etc.) Mais à titre d'ordre de grandeur et dans des conditions moyennes, une largeur de 10 à 20 m permet une épuration correcte, supérieure à 80 % du flux parvenant à la rivière de l'un ou l'autre de ces polluants (CSPNB, 2008).

IV. Les corridors rivulaires : espace clé de la restauration mais espace fragilisé

Par leur multifonctionnalité (Décamps, 2002), les corridors rivulaires fournissent de nombreux bénéfices et services pour la nature comme pour l'homme. De nombreux travaux, généralement menés à l'échelle locale (i.e. la station), montrent que les corridors rivulaires constituent un compartiment-clé de « l'écosystème rivière ». Ils contribuent et veillent au bon fonctionnement de l'écosystème assurant ainsi la biodiversité aquatique qu'ils protègent des flux biogéochimiques provenant de l'extérieur.

Or, la préservation de cette biodiversité est un enjeu important dans l'état écologique des cours d'eau car elle joue un rôle majeur dans le maintien de la résilience des écosystèmes rivières. Ce rôle tient à l'existence de groupes fonctionnels, c'est-à-dire d'ensembles d'espèces aux fonctions écologiques similaires. Par exemple,

CemOA : archive ouverte d'Irstea / Cemagref

parmi les invertébrés herbivores du fond d'un cours d'eau, certains broutent les algues de la surface des pierres, d'autres filtrent celles qui dérivent avec le courant, d'autres déchiquettent les feuilles tombées depuis les rives, s'attaquent aux tiges des macrophytes, etc. Après perturbation, un écosystème se réorganisera d'autant mieux qu'il renferme divers groupes fonctionnels et, au sein de chacun de ces groupes, diverses espèces qui, tout en accomplissant une même fonction, répondent différemment aux perturbations (Décamps, 2006).

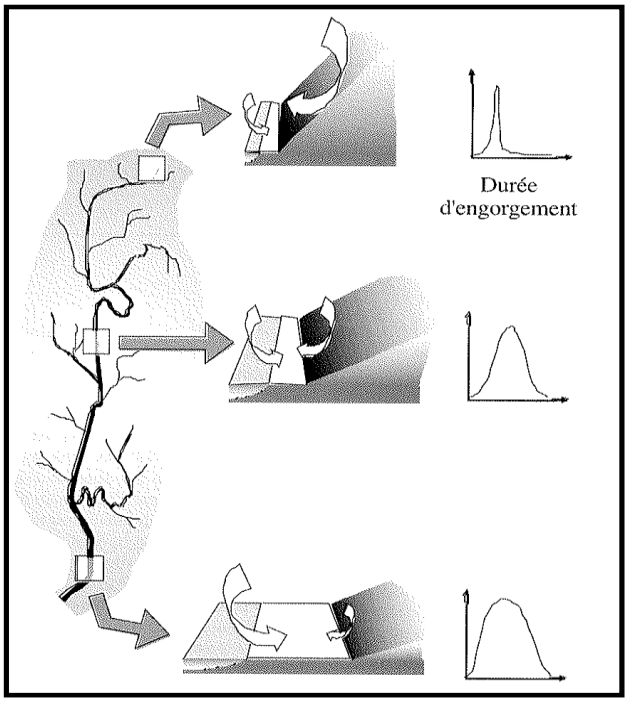


Figure 2.6 : Relations entre les ripisylves, le cours d'eau et le versant en fonction de leur position le long du réseau hydrographique (d'après Tabacchi *et al.*, 1998).

De ce fait, le corridor rivulaire, dont le foncier est plus facilement négociable qu'au niveau du bassin versant, constitue un espace clé de la restauration pour atteindre le « bon état » écologique, objectif de la Directive Cadre Eau à l'horizon 2015. Mais la mise en œuvre d'une politique de gestion ambitieuse nécessite des recommandations régionalisées adaptées qui impliquent la compréhension de ces mécanismes d'influence à de plus larges échelles.

L'espace potentiel, sur lequel une ripisylve est susceptible de se développer, est étroitement dépendant de la dynamique fluviale. Ainsi, sur une rivière qui dispose d'un espace de liberté suffisant et dont le débit n'est pas régulé, la ripisylve ne nécessite aucune intervention particulière pour se maintenir ou se développer et assurer ses fonctions écologiques, environnementales et sociales.

Cependant, l'anthropisation des cours d'eau entrave cette liberté et fragilise la dynamique écologique de cet écosystème (Wasson *et al.*, 1998). Ainsi, l'activité humaine contribue à modifier les types de ripisylves que l'on observe dans les plaines d'inondation. De plus en plus influencées, les ripisylves sont ramenées à l'état de fragments dont la survie dépend de leur aptitude à se renouveler face au stress. L'homme contrôle l'extension spatiale des corridors : il peut les détruire totalement ou permettre au contraire une renaturation par abandon des terrains riverains. Il peut également modifier leur composition floristique au point de privilégier des peuplements monospécifiques cultivés. Les boisements de berge ne sont pas moins soumis à l'action humaine. Les espèces présentes ont souvent été favorisées par l'homme, soit pour alimenter le bétail ou exploiter le bois de chauffe (ex : frêne), soit pour alimenter la filière bois (ex : ligne de peupliers).

De plus, de par leur grande richesse et la forte productivité des sols, régulièrement fertilisés par des dépôts limoneux, les plaines alluviales ont été utilisées et exploitées dès la plus haute antiquité. Avec les progrès techniques, ces activités (agricoles, industrielles, urbaines) se sont intensifiées et l'accumulation de ces interventions sur le milieu a engendré un certain nombre de perturbations, notamment une dégradation croissante de la qualité des eaux souterraines et superficielles. Les exemples sont multiples : les fuites d'engrais

CemOA : archive ouverte d'Irstea / Cemagref

en hiver sur un sol nu peuvent être, en partie, tenues pour responsable de l'enrichissement excessif en nitrates de nombreuses rivières (Corgne, 2004) ; des apports trop importants en sédiment fin, liés souvent à cette anthropisation du bassin versant, sont reconnus comme étant une cause de dégradation des conditions écologiques. Ces particules minérales en suspension ou déposées peuvent être à la source de préjudices graves pour la faune aquatique (séquelles au niveau des branchies, colmatage des frayères et des habitats pour les invertébrés benthiques et les poissons, limitation des apports en oxygène à l'intérieur des sédiments) (Maridet, 1995).

Enfin, c'est prioritairement sur ces territoires que se sont concentrés les infrastructures urbaines (routes, voies de chemin de fer, digues, barrages, terrains de camping et/ou de golf, urbanisation, …) qui influencent aussi, de façon brutale et durable, les évolutions naturelles.

Dès lors, les aspects pratiques de la mise en œuvre d'une politique ambitieuse à l'échelle nationale sur ces espaces impliquent le développement d'indicateurs permettant de dresser un « état des lieux » national et de suivre les effets des actions entreprises.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Chapitre 3 :
Modélisation des relations pressions/impacts

SOMMAIRE

Résumé *75*

I. LES ENJEUX DE LA MODELISATION 77

I.1. PROBLEMATIQUES ET OBJECTIFS *77*

I.2. LA LOGIQUE DPSIR *78*

I.3. LES DIFFERENTES APPROCHES DE MODELISATION *79*

I.4. APPORT DES SYSTEMES INFORMATION GEOGRAPHIQUE *79*

II. LES MODELES MECANISTES 80

II.1. BASES THEORIQUES ET DIFFERENCIATION DES APPROCHES *80*

II.1.1. Le système: le bassin versant 80
II.1.2. Du système au modèle 81
II.1.3. Le couplage modèles de qualité et modèles hydrologiques 81
II.1.4. Les différents types de modèles 82

II.2. APPROCHE DE MODELISATION MECANISTE DE LA QUALITE DE L'EAU AU NIVEAU D'UN BASSIN : L'EXEMPLE DU PROGRAMME PIREN-SEINE *83*

II.3. LES MODELES CIBLES SUR LE ROLE EPURATOIRE DES RIPISYLVES *85*

II.4. LES MODELES « MECANISTES » : UN INTERET LIMITE POUR LA MISE EN ŒUVRE DE LA DCE *87*

III. LES MODELES EMPIRIQUES 89

III.1. PRINCIPES ET CONCEPTS *89*

III.1.1. L'écologie des paysages 89
III.1.2. La hiérarchie des systèmes aquatiques (concept de l'habitat) 90

III.2. REVUE DES MODELES EXISTANTS *92*

III.2.1. Synthèse des travaux internationaux 92
III.2.1.1 Les données de pressions 92
III.2.1.2 Les données biologiques 94
III.2.1.3 Echelles et zones de travail 95
III.2.1.4 Méthodes d'analyses de données 98
III.2.1.5 Synthèse des résultats 98
III.2.1.5.1 Impact de l'agriculture 98
III.2.1.5.2 Impact de l'urbanisation 99
III.2.1.5.3 Bassin versant vs. corridor rivulaire 100
III.2.2. Modèles pressions/impacts en France 100
III.2.2.1 Présentation du modèle pressions/impacts régionalisé 101
III.2.2.2 Principaux résultats 101

IV. VERS DES MODELES ADAPTES A LA DCE : AMELIORER LES MODELES EMPIRIQUES PAR UNE MEILLEURE PRISE EN COMPTE DES CORRIDORS RIVULAIRES. 102

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Résumé

La mise en œuvre de la Directive Cadre sur l'Eau (DCE) exige pour la phase de restauration l'identification des pressions anthropiques significatives et l'évaluation de leurs impacts sur l'état écologique des cours d'eau au niveau des districts hydrographiques. Cette identification passe par des modèles explicatifs spatialisés reliant quantitativement et à plusieurs échelles les impacts à leur causes probables afin de procéder à des évaluations des effets de la mise en œuvre de diverses politiques publiques possibles et notamment au niveau des corridors rivulaires.

Le schéma conceptuel DPSIR (Driving Forces, Pressures, State, Impact, Responses) permet de formaliser les notions mises en jeu. Dans ce schéma, peut-être plus que les éléments, ce sont les liens de causalité qui s'avèrent importants et qui nuancent les approches de modélisation. On distingue deux grands types d'approches: mécaniste et empirique utilisant toutes deux les Systèmes d'information Géographique (SIG) pour étudier ces relations dans l'espace et présenter les résultats sur des cartes didactiques. Alors que les modèles mécanistes cherchent à décrire précisément l'ensemble des processus du système à l'aide de lois mathématiques (mécaniques), les modèles empiriques se contentent, en se basant sur un ensemble de concepts, d'étudier les relations statistiques entre certains éléments du système.

Les progrès réalisés au cours des quinze dernières années dans la connaissance du fonctionnement des écosystèmes (exemple du Programme PIREN-SEINE) permettent aujourd'hui d'expliciter quantitativement à l'aide des modèles mécanistes le lien entre les contraintes qu'exercent sur un grand bassin versant le climat, la topographie, la géologie et les activités humaines, agricoles et urbaines. Cependant, ces modèles s'avèrent limités pour la mise en œuvre de la DCE. L'acquisition d'un grand nombre de données nécessaires pour ces modèles est incompatible avec les contraintes opérationnelles et les sorties de ces modèles ne répondent aujourd'hui que partiellement aux objectifs de qualité: ils *sont capables de simuler le transport, la dispersion d'un certain nombre d'éléments biogéochimiques et de substances toxiques mais ne permettent pas de prédire la présence et la diversité biologique, paramètres clés dans l'évaluation de l'état écologique au sens de la DCE. De plus, il apparaît très difficile de préciser et de quantifier à l'échelle du bassin les processus impliqués au niveau des corridors rivulaires, tant ceux-ci sont imbriqués les uns aux autres et sont contrôlés par de multiples paramètres (exemple du modèle REMM).*

Les modèles empiriques utilisent les outils conceptuels et techniques de l'écologie des paysages pour étudier les relations entre les pressions d'occupation du sol et l'état écologique (réponse biologique) à travers différentes échelles spatiales en se basant sur la hiérarchie emboitée des écosystèmes d'eau courante. Ces modèles, forces motrices/états selon la logique DPSIR, permettent ainsi d'orienter les gestionnaires sur les priorités d'action en identifiant et hiérarchisant les principales structures (sources potentielles de diverses pressions) influençant les paramètres d'évaluation de l'état écologique requis par la DCE. Les travaux existants s'accordent sur l'agriculture et l'urbanisation comme étant les deux plus grandes causes de dégradation mais diffèrent sur l'échelle spatiale d'influence de cette occupation du sol (bassin versant vs. corridor rivulaire) et par conséquent sur le rôle de la végétation rivulaire. Afin de mieux appréhender les relations au niveau du corridor rivulaire, espace clé de la restauration, il ressort de cette synthèse qu'il est nécessaire (i) d'acquérir une information fine de l'occupation du sol le long du cours d'eau afin de caractériser correctement sa composition et sa configuration spatiale ; et (ii) de construire des indicateurs spatialisés à partir de cette information afin de mieux localiser les effets dans ces espaces et de mettre en évidence le rôle de la végétation rivulaire à large échelle.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

I. Les enjeux de la modélisation

I.1. Problématiques et objectifs

La mise en œuvre de la Directive Cadre sur l'Eau (DCE) n'implique pas simplement de constater les dégradations de l'état écologique des milieux aquatiques. Elle oblige les États membres à établir un diagnostic des impacts des activités humaines (Article 5) afin d'orienter les gestionnaires dans leur stratégies de restauration au niveau des unités cohérentes de gestion (bassin versant et tronçon de rivière). En termes de pressions la DCE exige « l'identification des pressions anthropiques significatives et l'évaluation de leurs impacts sur l'état écologique des cours d'eau» (Annexe II). Autrement dit, il est nécessaire d'identifier et de hiérarchiser les causes d'altération des processus fonctionnels des écosystèmes aquatiques (physiques, biogéochimiques et écologiques).

Les cours d'eau, systèmes éminemment ouverts, comme la nature et le fonctionnement des biocénoses qu'ils abritent, sont étroitement liés aux activités humaines ; et comme tout écosystème, leur processus fonctionnels interagissent à des échelles multiples dans l'espace et dans le temps (Lévêque, 2001). Ainsi, le rôle de collecteur d'un bassin versant fait que toute altération des écosystèmes terrestres a des répercussions potentielles, à grande distance et à long terme, sur les flux d'eau et d'éléments qui alimentent les rivières et donc sur les écosystèmes aquatiques situés à l'aval. La structure linéaire des cours d'eau les rend particulièrement vulnérables aux structures transversales qui interrompent les flux, comme aux contraintes longitudinales exercées par leur environnement (Wasson *et al.*, 1993).

Les impacts écologiques ne sont donc pas seulement déterminés par l'identification précise des sources de pollution ponctuelles, mais aussi par l'ensemble des influences anthropiques plus complexes comme les pollutions diffuses, les dysfonctionnements hydrologiques, les variations du flux de sédiments, les modifications hydromorphologiques et les ruptures de connectivité (Borchardt and Richter, 2003). Dans la plupart des cas, ces pressions agissent simultanément (pressions combinées) et sont réparties irrégulièrement sur le territoire. La recherche des causes d'altération de l'état écologique implique donc clairement une analyse multi-échelle des mécanismes d'impact, régionalisée sur de grands territoires.

Compte tenu du délai imparti pour la mise en œuvre de la DCE, cette recherche doit permettre de guider les gestionnaires sur les actions prioritaires à mener en identifiant et hiérarchisant les causes d'altération aussi bien à l'échelle du bassin versant que du tronçon de rivière et en établissant des préconisations sur les espaces rivulaires compte tenu de leur rôle sur l'état écologique des cours d'eau (cf. chapitre 2). Ceci se traduit par des questions clés : sur quelles causes de dégradation agir ? où et à quelles échelles spatiales ? quels couverts végétaux privilégier dans les corridors rivulaires ? de quelles dimensions ?, ceci se traduit également au niveau technique par le besoin d'un outil d'aide à la décision spatialisée (i.e : des résultats visibles sur des cartes).

La première idée qui vient à l'esprit est d'interroger directement les bioindicateurs mesurés par les réseaux de suivi de la qualité de l'état écologique des cours d'eau pour y lire la signature des pressions. Cependant, il n'existe pas de lien de cause à effet assez fort entre la ou les pressions exercées et la réponse du milieu biologique. Même une analyse poussée des peuplements demeure limitée pour identifier les causes (complexes et variées) qui perturbent le milieu (Wasson *et al.*, 2005). Deux facteurs viennent compliquer le problème (Boulton, 1999): (1) la variabilité des réponses biologiques en fonction du contexte naturel ; (2) les interactions entre différents types d'altérations.

Des études au cas par cas n'étant pas envisageables, la définition d'une politique d'action passe alors par le développement d'outils de modélisation au niveau des districts hydrographiques (grands bassins hydrologiques, terminologie de la DCE) analysant (diagnostiquant), quantitativement et à plusieurs échelles, l'impact relatif des différentes pressions anthropiques s'exerçant sur l'état écologique. Ces modèles de diagnostic qui ciblent les priorités de restauration offriront aux acteurs de la gestion de l'eau les moyens de définir des stratégies d'action adaptées et efficaces et de procéder ultérieurement à des évaluations a priori des effets possibles de la mise en œuvre de diverses politiques publiques envisageables.

CemOA : archive ouverte d'Irstea / Cemagref

I.2. La logique DPSIR

L'interaction dynamique entre l'homme et l'état écologique des cours d'eau caractérise ce qu'il est convenu d'appeler des socio-écosystèmes (Walker *et al.*, 2002). Les sociétés humaines utilisent les écosystèmes aquatiques et, par suite, les modifient, localement et globalement. En retour, elles sont capables d'ajuster leurs usages aux modifications de ces écosystèmes qu'elles perçoivent. La résilience de ces socio-écosystèmes passe, notamment, par l'aptitude des sociétés humaines à tirer les leçons d'expériences passées pour anticiper et élaborer des plans pour l'avenir (Décamps, 2006).

Une des difficultés dans la modélisation des relations au sein du socio-écosystème est de conceptualiser la démarche afin de bien clarifier les notions mises en jeu et notamment de pouvoir communiquer dans le cadre de la participation du public. Dans cette optique, le guide IMPRESS européen (MEDD/DE and Aquascop, 2003) s'appuie sur le schéma conceptuel DPSIR (Driving Forces, Pressures, State, Impact, Responses), FPEIR en français (Forces motrices, Pressions, État, Impact, Réponses), mis au point par l'Agence Européenne de l'Environnement. Il s'articule selon ces cinq éléments reliés par des liens de causalité.

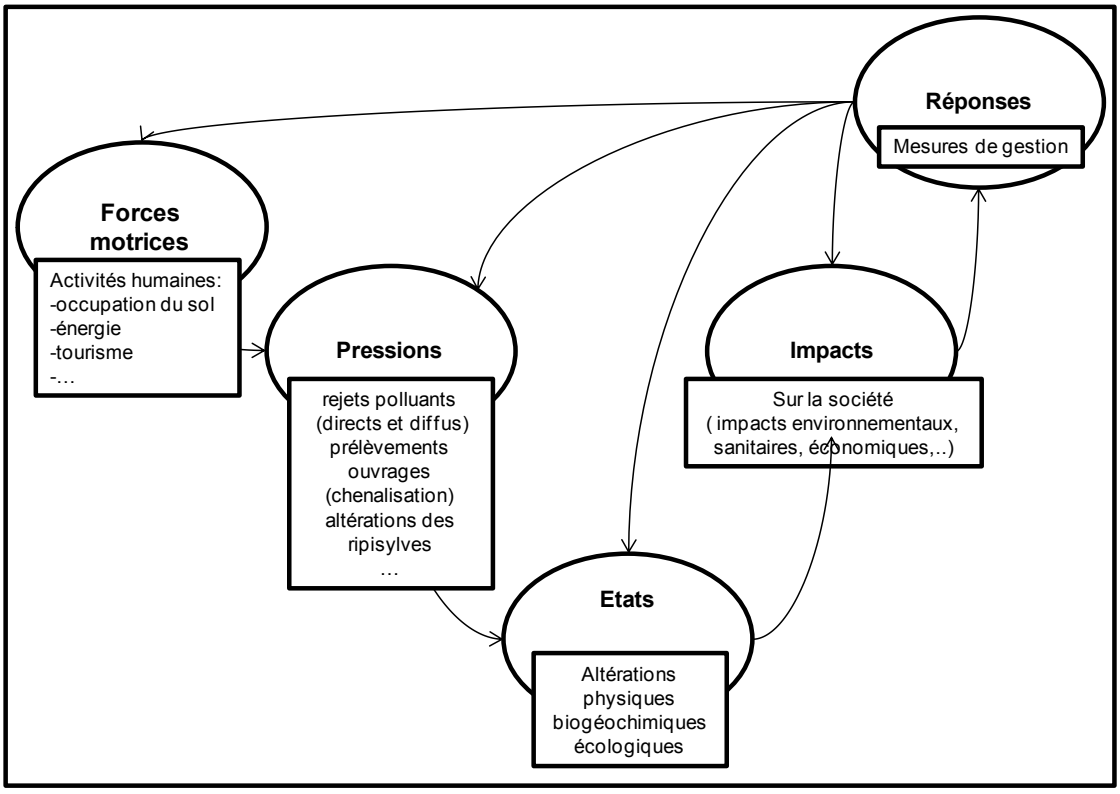


Figure 3.1 : Schéma DPSIR pour la représentation de l'action humaine sur l'état écologique des cours d'eau

Ce schéma adapté à la représentation de l'action humaine sur l'état écologique des cours d'au se décompose de la manière suivante (Figure 3.1): les *forces motrices* (1er élément) que constituent les activités humaines (agriculture, zones urbaines, industries, tourisme, …) génèrent des *pressions* (2ème élément) à l'échelle du bassin versant (pollution diffuses) et du tronçon (rejets directs, prélèvements et ouvrages). Ces pressions modifient dans l'espace et/ou dans le temps les processus fonctionnels au niveau hydro-morphologique (structure, hétérogénéité, connectivité, variabilité de l'habitat physique…), biogéochimique (pollution organique, nutriments, toxiques,...) et écologique, et déterminent in fine les *états* (3ème élément) au niveau biologique (invertébrés, poissons …) et chimique (substance toxiques et paramètres physico-chimiques). Ces altérations d'états engendrent alors des *impacts* (4ème élément) sur la société au niveau environnemental (biodiversité, paysage…), économique (pêche, tourisme…) et sanitaire (qualité de l'eau, maladies…). Ces impacts, en fonction de leur gravité, de leur ampleur sur les différents usages de l'eau, font réagir la société qui apporte ainsi des *réponses* (5ème élément) regroupant l'ensemble des mesures et des instruments politiques.

CemOA : archive ouverte d'Irstea / Cemagref

Ces réponses globalisées et localisées portent sur les 4 autres éléments : mesures préventives dirigées vers les *forces motrices* (agriculture raisonnée, sensibilisation, …), mesures curatives vers les *pressions* et l'*état* (bandes enherbées, stations d'épurations, réméandrement …) et mesures palliatives vers l*es états* et les *impacts* (mélanges des eaux, usines de dénitrification,…).

Cette logique permet d'ordonner les principales composantes (éléments) du système, et de regrouper différentes pressions issues de forces motrices variées dont l'impact s'exercera sur une masse d'eau donnée. Dans ce schéma, peut-être plus que les éléments, ce sont les liens de causalité qui s'avèrent importants et qui nuancent les approches de modélisation.

I.3. Les différentes approches de modélisation

On peut distinguer ainsi deux grandes approches de modélisation ou types de modèle : les modèles mécanistes et les modèles empiriques.

Les modèles mécanistes décrivent, dans un cadre physique défini, les processus (mécanismes) internes du système, relatifs à chacun des éléments du schéma DPSIR et à leurs liens de causalité. Ces modèles se basent sur la connaissance que l'on a de la structure physique du système et des lois de la physique (conservation de la masse, de l'énergie, de la quantité de mouvement, etc.), de la chimie (équilibres chimiques et thermodynamiques, cinétiques chimiques) ou de la biologie (lois de comportement de croissance) qui le régulent. Ils utilisent de bout en bout, la même description fine des processus et adoptent généralement une approche bottom-up[15] dans la compréhension des relations, le fonctionnement à large échelle résultant de l'articulation et de l'aggrégation des fonctionnements à échelle plus fine. La spatialisation joue un rôle important dans la description de la structure du système et de la façon dont les différents sous-systèmes interagissent.

A l'opposé, Les modèles empiriques ne font pas référence aux processus internes du système, qui peuvent être inconnus, mais formalisent les relations statistiques entre variables descriptives de certains éléments du système. La connaissance que l'on a du système peut être synthétisée dans ces variables descriptives (indicateurs). Les modèles empiriques adoptent généralement une approche top-down[16] dans la compréhension des relations, privilégiant celles qui sont statistiquement significatives au niveau du système.

Dans la communauté scientifique (et notamment en France) ces deux approches sont utilisées pour relier un effet mesuré sur la qualité du cours d'eau à une cause anthropique. Dans la suite de ce chapitre, les bases théoriques, les différents modèles développés et les résultats obtenus à partir de ces deux approches sont présentés et analysés en fonction de leur contribution pour la mise en œuvre de la directive cadre eau.

I.4. Apport des Systèmes Information Géographique

Quelque soit l'approche de modélisation (mécaniste ou empirique), la spatialisation est fondamentale dans l'étude des relations pressions/impacts aussi bien pour les scientifiques, dans leurs efforts de compréhension de ces relations (les pressions agissant à divers points et à différentes échelles au sein du bassin versant), que pour les gestionnaires dans la localisation des problèmes et la planification des mesures d'actions.

Les Systèmes Information Géographique (SIG) constituent un outil incontournable dans la modélisation de ces relations (Johnson and Gage, 1997; Ruelland, 2004). Ils connaissent un essor important depuis plusieurs années, en particulier pour l'aménagement du territoire (Johnson, 1990) et dans le domaine de l'eau (Barge and Joliveau, 1996),

[15] L'approche bottom-up (lit. de bas en haut) commence par modéliser finement les processus dans les détails pour aller vers le général.
[16] Une démarche dite descendante ou approche top-dow (lit. de haut en bas) modélise les processus du système dans les grandes lignes, puis itérativement, s'intéresse à des détails de plus en plus fins.

CemOA : archive ouverte d'Irstea / Cemagref

Les SIG sont des outils[17] qui permettent de saisir, stocker, consulter, transformer, combiner, créer et représenter des informations hétérogènes localisées (i.e. géoréférencées dans le même système de coordonnées). Chaque thème d'information est représenté par une couche qui contient la représentation cartographique des objets spatiaux (objets élémentaires de même nature) et la table d'information statistique (attributaire) qui leur est associée.

D'un point de vue scientifique, cet outil permet donc de manipuler et gérer aisément des objets géographiques (informations spatiales) de natures différentes (e.g, réseau hydrographique, stations de mesures, parcelles agricoles) ainsi que les données associées (e.g, types de rivière, diversité biologique, superficies des parcelles) mais aussi de simplifier l'analyse spatiale entre ces informations (par exemple intégration des données sur un objet polygone pour associer le résultat à cet objet : transfert de polluant des parcelles agricoles au niveau d'un bief du réseau hydrographique ou de la station de mesure) et de simuler des relations.

Dans une logique de gestion, les fonctions de transformation et combinaison (agrégation, superposition, interpolation …) sont très intéressantes car elles permettent de construire une même information de façon homogène sur un territoire, dans le mode de représentation propre aux gestionnaires (e.g, qualité de l'eau par bassin versant ou par masses d'eaux), et de fournir des supports didactiques (cartes) qui constituent un outil de dialogue et de confrontation de points de vue entre les différents acteurs (Wasson *et al.*, 1993).

II. Les modèles mécanistes

II.1.Bases théoriques et différenciation des approches

L'objectif opérationnel d'un grand nombre de ces modèles est de quantifier les concentrations ou les flux de matière (eau, nutriments, matières en suspension, …) en tout point (ou bief) d'un cours d'eau, ces variables ayant un impact sur le compartiment biologique ; et si possible de quantifier cet impact en termes de croissance ou de mortalité d'une espèce. Il existe de nombreuses approches de modélisation mécaniste. Elles s'appuient sur des bases théoriques similaires mais se différencient en fonction de leurs hypothèses simplificatrices (sur les processus en jeu) et des dimensions (dans le temps et dans l'espace) de leur mise en œuvre.

II.1.1. Le système: le bassin versant

Le concept de bassin versant[18] s'impose pour la communauté scientifique des hydrologues et des hydrobiologistes comme l'unité fonctionnelle (le cadre physique) d'étude du cycle de l'eau et des flux associés (énergie, matière) dans les approches mécanistes (Payraudeau, 2002). Il est défini comme un système ouvert complexe, puisqu'il échange de la matière et de l'énergie avec son environnement selon de nombreuses boucles de régulation (absentes dans un système simple). Il est par ailleurs très atténuant puisque, soumis à un stimulus, il retrouve rapidement l'état précédant le stimulus.

Comme dans tout système, le fonctionnement d'un bassin versant résulte de la conjonction d'un ensemble de processus. Trois grands types de processus ont été décrits pour expliquer ce fonctionnement global: (1) les processus d'écoulement de stockage/déstockage de l'eau et des matières associées ; (2) les processus de transformation interne liés aux changements de phase pour l'eau, de forme pour un composé chimique (et de comportement pour une espèce au niveau du milieu récepteur) ; et (3) les processus de transfert vers les limites du bassin (atmosphère et exutoire). A noter que le terme de transport est privilégié pour les processus liés au déplacement de polluant sur le bassin qui dépendent des écoulements de surface et de subsurface.

L'imbrication de ces différents processus, dont l'importance varie en tout point du bassin et à chaque instant, rend la dynamique du bassin fortement non linéaire.

[17] Il s'agit là du sens restreint de l'expression SIG. Au sens large les Systèmes d'Information Géographiques (ou SIRS Systèmes d'Information à Références Spatiales) sont un ensemble de moyens humains, techniques, informatiques et organisationnels mis en œuvre pour acquérir, traiter, analyser, stocker et mettre à disposition des acteurs des données et informations pertinentes par rapport à une problématique ou un projet territorial (Bordin, 2002).

[18] Surface drainée par un cours d'eau, en amont d'un point définissant son exutoire

CemOA : archive ouverte d'Irstea / Cemagref

II.1.2. Du système au modèle

Le cycle de l'eau et le cycle des différents éléments biogènes (carbone, hydrogène, oxygène, azote, phosphore, silice, etc.) et toxiques (métaux, médicaments, etc.) à l'échelle du bassin versant sont particulièrement complexes puisqu'ils intègrent des processus physiques, chimiques et biologiques. Alors que chacun de ces processus peut être quantifié et reproduit à l'échelle du laboratoire (afin d'étudier leur cycle complet), leur étude à l'échelle du bassin versant est beaucoup plus complexe compte tenu de leur variabilité dans l'espace et le temps. La représentation du devenir de ces éléments dans les modèles de flux implique souvent un degré de simplification important. Le modèle est donc une représentation simplifiée des processus du système.

De façon schématique on commence par décrire la structure du système en le découpant en sous-systèmes (ou objets) en interaction.

Chaque sous-système est décrit par son état et ses interactions avec les autres sous-systèmes. On différencie ainsi les variables (grandeurs mesurables du modèle qui évoluent au cours du temps) selon :

- les variables d'état qui caractérisent l'état du sous-système (ex. pour une masse d'eau : profondeur, température de l'eau, teneur en oxygène dissous dans la rivière, structure des peuplements…) ;
- les variables d'entrée, générées par les sous-systèmes voisins et par l'extérieur du système (on les appelle alors variables de forçage), qui influencent son évolution. Ce sont par exemple les apports de matière (azote, pesticides, …) liés ou non aux activités anthropiques, les variables de forçage pouvant être les variables atmosphériques (température de l'air, pluie, rayonnement solaire, vent, …);
- les variables de sortie qui caractérisent les sorties du sous -système (flux sortants de matière : eaux, nutriments, …, et d'énergie) soit vers d'autres sous-systèmes, soit vers l'extérieur du système.

Dans un second temps on décrit les processus régissant l'évolution spatiale et temporelle des variables d'état et des variables de sortie d'un sous-système en fonction de ses variables d'entrée, par des lois ou équations dépendant de certains paramètres fixes (qui n'évoluent pas au cours du temps, à l'inverse des variables qui évoluent au cours du temps)

La démarche de modélisation s'organise ainsi généralement en plusieurs étapes : (1) la formalisation du système et sa structuration en sous-systèmes, ce qui inclut la formalisation de l'état et des variables d'entrée et sortie de chaque sous-système, (2) la formalisation des modèles représentant les processus, (3) mise au point du schéma de la résolution numérique du système d'équation, (4) le calage des paramètres des lois, par rapport aux variables d'état observées sur le terrain, et (5) la validation ou évaluation du modèle en comparant les résultats de simulation (après calage) aux situations observées.

II.1.3. Le couplage modèles de qualité et modèles hydrologiques

Dans la pratique on conçoit que l'état écologique d'un milieu aquatique résulte de la conjonction d'un régime hydrologique (flux d'eau et leurs dynamiques temporelles), d'un régime de qualité des eaux (transferts de solutés et de sédiments et leurs dynamiques temporelles) et d'un fonctionnement écologique du milieu (dynamique des peuplements). Les modèles de régime hydrologique, les modèles de qualité des eaux et les modèles de fonctionnement d'écosystèmes aquatiques existent de façon indépendante. Leur couplage reste un point sensible de la démarche de modélisation d'ensemble.

Les flux d'éléments à la sortie du système (exutoire et atmosphère) dépendent d'une part du régime hydrologique (débits et conditions d'écoulement du cours d'eau qui conditionnent la dilution des nutriments, la sédimentation et les temps de rétention), lui-même fortement contrôlé par le climat et la morphologie du bassin, et d'autre part des concentrations d'éléments solutés et sédiments, eux-mêmes dépendants des termes source.

CemOA : archive ouverte d'Irstea / Cemagref

La prédiction de ces flux nécessite le couplage entre un modèle hydrologique (et hydraulique) relatif aux processus de production et de transfert de l'eau et un modèle de qualité relatif aux processus de production et de transport des éléments. Le couplage entre ces modèles ainsi que le niveau d'intégration de l'information spatiale (par les SIG) peut être de différents types : « lâche », « serré », « imbriqué » (Flipo, 2005):

- dans un couplage de type « lâche », les modèles sont totalement dissociés et les données sont transférées d'un modèle à l'autre par l'utilisateur, de manière séquentielle (chaînage de modèles),
- dans un couplage de type « serré », les modèles fonctionnent toujours de manière indépendante mais font partie d'une interface commune,
- dans le couplage « imbriqué », les modèles sont les composantes d'une même application et partagent la même base de données, avec éventuellement des interactions et des boucles de rétroaction.

II.1.4.Les différents types de modèles

On trouve dans la littérature de nombreuses approches de modélisation résultant des couplages entre ces modèles. Ces approches se différencient selon leurs hypothèses simplificatrices (notamment sur les processus) et leurs modes de formalisation de l'espace et du temps (Gineste, 1998; Payraudeau, 2002). Au sein de la famille des modèles « mécanistes » on distingue les modèles « physiques » des modèles « conceptuels ».

Les modèles « physiques » relient les variables d'état et de sortie par des relations dynamiques (ex. flux de matières reliés aux gradients de potentiel et aux résistances du milieu) et des relations de conservation (bilan de matière, d'énergie, de quantité de mouvement). Ils permettent, en précisant les conditions initiales et les conditions aux limites, de simuler l'évolution du système en tout point et à tout instant selon une discrétisation spatio-temporelle fine (e.g, couplage du modèle hydrologique SHETRAN et de qualité NITS, Birkinshaw and Ewen, 2000). Cette approche essaye de décrire précisément les mécanismes en jeu pour fournir une représentation au plus près du phénomène.

Les approches « conceptuelles », globales ou semi-distribuées, sont moins complexes que les approches « physiques » en termes de géométrie, de paramétrisation, de représentation des processus et de temps de calcul. Le fonctionnement du bassin est par exemple assimilé à celui d'un ensemble de réservoirs (les sous-bassins) interconnectés censés représenter plusieurs niveaux de stockage d'eau ou de masse de polluants.

Ces modèles reposent sur (1) une relation de conservation de bilan qui relie les variations de niveau de chaque réservoir aux flux entrant et sortant, (2) des relations de flux (type loi de vidange) qui caractérisent chaque réservoir et déterminent la dynamique des flux sortants (e.g, Modèle nitrate, MA, 1991).

Dans une approche semi-distribuée l'espace est discrétisé en sous-unités que l'on considère « homogènes » en termes de caractéristiques physiques et/ou de fonctionnement hydrologique (Ambroise, 1991). C'est chaque unité et non plus l'ensemble du bassin qui est représentée par un ensemble de réservoirs interconnectés. Cette structure permet de tenir compte de la répartition spatiale des caractéristiques hydrologiques de chaque sous-unité et des variables d'entrée (e.g, modèle SWAT, Arnold *et al.*, 1994; couplage du modèle hydrologique MODCOU et de qualité STICS, Ledoux *et al.*, 2007).

On notera que certaines approches ne sont pas totalement mécanistes mais font appel aux méthodes empiriques soit directement (intégration d'un modèle empirique) soit indirectement (calage des paramètres).

Ces modèles varient aussi en fonction de leur dimensionnement dans le temps et dans l'espace. La dynamique de ces flux peut être analysée aussi bien à l'exutoire de parcelles expérimentales sur des pas de temps très court que sur des bassins de plusieurs milliers de km² à des pas de temps plus long. Par rapport aux objectifs de modélisation (cf. I.1), on s'intéressera dans la suite de cette partie, à l'aide d'un exemple (programme PIREN-SEINE) aux modèles appliqués à la gestion de la qualité de l'eau des bassins versants de plusieurs centaines à plusieurs milliers de km². On peut citer par exemple le modèle SENEQUE (Billen *et al.*, 1994 ; Garnier *et al.*, 1995), PEGASE (Smitz *et al.*, 1997) et WATERWARE (Jamieson and Fedra, 1996). On étudiera aussi de plus près à l'aide d'un autre exemple (modèle REMM) les modèles qui se sont focalisés sur le rôle des corridors rivulaires.

CemOA : archive ouverte d'Irstea / Cemagref

II.2. Approche de modélisation mécaniste de la qualité de l'eau au niveau d'un bassin : l'exemple du programme PIREN-SEINE

Le PIREN-Seine est un groupement de recherche dont l'objectif est de développer, à partir de mesures de terrain et de modélisations, une vision d'ensemble du fonctionnement du système formé par le réseau hydrographique de la Seine, son bassin versant et la société humaine qui l'investit, fournissant ainsi un outil d'aide à la décision concernant l'impact des activités anthropiques sur la qualité de l'eau (PIREN Seine, 2007). Le fonctionnement écologique de l'ensemble du système fluvial et sa modélisation sont basés sur l'étude fine des processus physiques, chimiques et biologiques des milieux. Les modèles développés par le PIREN-Seine simulent les variations écologiques et biochimiques de l'hydrosystème, depuis les ruisseaux jusqu'à l'entrée de l'estuaire.

Ces modèles s'appuient sur le module écologique RIVE (Billen *et al.*, 1994 ; Garnier *et al.*, 1995). A partir d'une profonde connaissance des cycles biogéochimiques du carbone, de l'azote, du phosphore, de la silice et de l'oxygène, ce modèle décrit conceptuellement, de manière assez complexe, les flux de ces éléments entre les différents compartiments du milieu aquatique (colonne d'eau, vase et périphyton[19]). Plus d'une quinzaine de variables d'état (concentrations de nutriments inorganiques et organiques dissous, biomasses phytoplanctoniques, biomasse bactérienne hétérotrophe et autotrophe, etc.) et plusieurs dizaines de paramètres sont nécessaires pour décrire les cinétiques des processus chimiques, microbiologiques et planctoniques. Ces données et résultats de modélisation permettent de décrire la qualité de l'eau par la concentration (ou flux) en oxygène, en nutriments (NH_4^+, NO_3^-, PO_3^{4-}, SiO_2,…), en matières en suspension et en carbone dissous et particulaire. Les compartiments biologiques sont représentés par des espèces d'algues, des organismes zooplanctoniques, des bactéries hétérotrophes ainsi que des bactéries nitrifiantes. Les bactéries fécales et les macrophytes sont également pris en compte en tant que variables d'état.

Ce modèle biogéochimique est ensuite couplé avec un modèle hydrologique qui diffère selon la zone hydrographique du bassin de la Seine (environ 78 000 km²), compte tenu du gradient physique amont/aval des cours d'eau. Quatre modèles ont été nécessaires (Even *et al.*, 2007) pour couvrir l'ensemble du bassin de la Seine : (1) le modèle RIVERSTRAHLER (SENEQUE) pour le réseau de drainage amont (Billen *et al.*, 1994 ; Garnier *et al.*, 1995) ; (2) le modèle ProSe pour le secteur fluvial à l'intérieur et à l'aval de l'agglomération de Paris (Even *et al.*, 1998); (3) le modèle SiAm1D/ et SiAm3D pour la zone estuaire (Thouvenin *et al.*, 1999 ; Le Hir *et al.*, 2001); et (4) le modèle SiAm3D/Elise pour la baie de la Seine (Elise, plus adapté aux processus biogéochimiques dans cette partie de la Seine notamment dans le compartiment benthique, Cugier and Le Hir, 2002, remplace le modèle de qualité RIVE).

Ainsi, ces modèles fournissent une image globale de la qualité de l'eau (et non pas une description fine des évolutions temporelles à court terme). Le modèle SENEQUE (« la Seine en équations »), par exemple, permet de calculer au pas de temps décadaire les principales variables de la qualité physico-chimique et microbiologique à l'échelle d'un bassin versant sur un tronçon de cours d'eau. Il fonctionne aujourd'hui grâce à l'imbrication de plusieurs modules (Figure 3.2).

Au niveau hydrologique, un premier module commence par calculer les écoulements, superficiels et profonds, dans chacun des sous-bassins unitaires de SENEQUE. Ce module est contrôlé par quatre paramètres climatiques calibrés empiriquement. Des études sont en cours pour le remplacer par le schéma de surface CaB pour augmenter le déterminisme physique de SENEQUE vis−à−vis du climat (Ducharne *et al.*, 2001). Ensuite, un module de transfert permet de transformer ces écoulements depuis le bassin versant en débits dans le réseau hydrographique selon un algorithme lagrangien. Une paramétrisation hydraulique, basée sur la formule de Manning permet de calculer vitesses et profondeurs des écoulements en rivière à partir de ces débits et de la morphologie du réseau hydrographique (pente, largeur, section mouillée). Une caractéristique du modèle

[19] Le périphyton est « une communauté complexe de microbiota (algues, bactéries, champignons, animaux, détritus organiques et inorganiques) qui est attachée au substrat qu'il soit organique ou inorganique, vivant ou inerte »

CemOA : archive ouverte d'Irstea / Cemagref

SENEQUE est d'idéaliser cette morphologie, dans les sous-bassins amont, en fonction des ordres de Strahler. A noter que les aménagements hydrauliques, comme les barrages de navigation ou les barrages réservoirs, sont pris en compte dans ce module de transfert (données morphologiques).

Au niveau de la qualité des eaux, à partir des données d'occupation du sol, des pratiques agricoles et des résultats du module hydrologique, la composition sous racinaire en éléments est décrite soit en fonction de relations empiriques (observations lysimétriques) soit par modélisation (modèle STICS, INRA, Brisson *et al.*, 1998) et la composition phréatique est définie à partir du couplage des modèles MODCOU et STICS (Ledoux *et al.*, 2007). Avant que ces compositions soient transmises au modèle écologique RIVE, le rôle épurateur des corridors rivulaires est pris en compte dans le modèle. Associé à l'inventaire typologique des zones humides riveraines (ZHR), un indice synthétique de rétention des nitrates et de régulation des crues a été défini pour chaque classe typologique (Gaillard *et al.*, 2005). L'estimation de cette rétention riparienne dans SENEQUE s'opère à l'échelle de chaque bassin versant élémentaire, où le potentiel de rétention est estimé au prorata des superficies et en fonction des capacités de rétention définies pour chacune des classes présentes. Ce taux de rétention peut être aussi calé empiriquement lors des différentes simulations (Billen and Garnier, 1999). Ainsi, à partir des apports diffus « épurés » par les ripisylves et des apports ponctuels, le modèle écologique RIVE permet de simuler les transferts de matière vers le cours d'eau et l'évolution de la biomasse planctonique et des concentrations pendant le transfert amont/aval. Les principales annexes hydrauliques de surface (notamment les étangs et les barrages réservoirs) sont aussi prises en compte pour décrire le fonctionnement biogéochimique du cours d'eau.

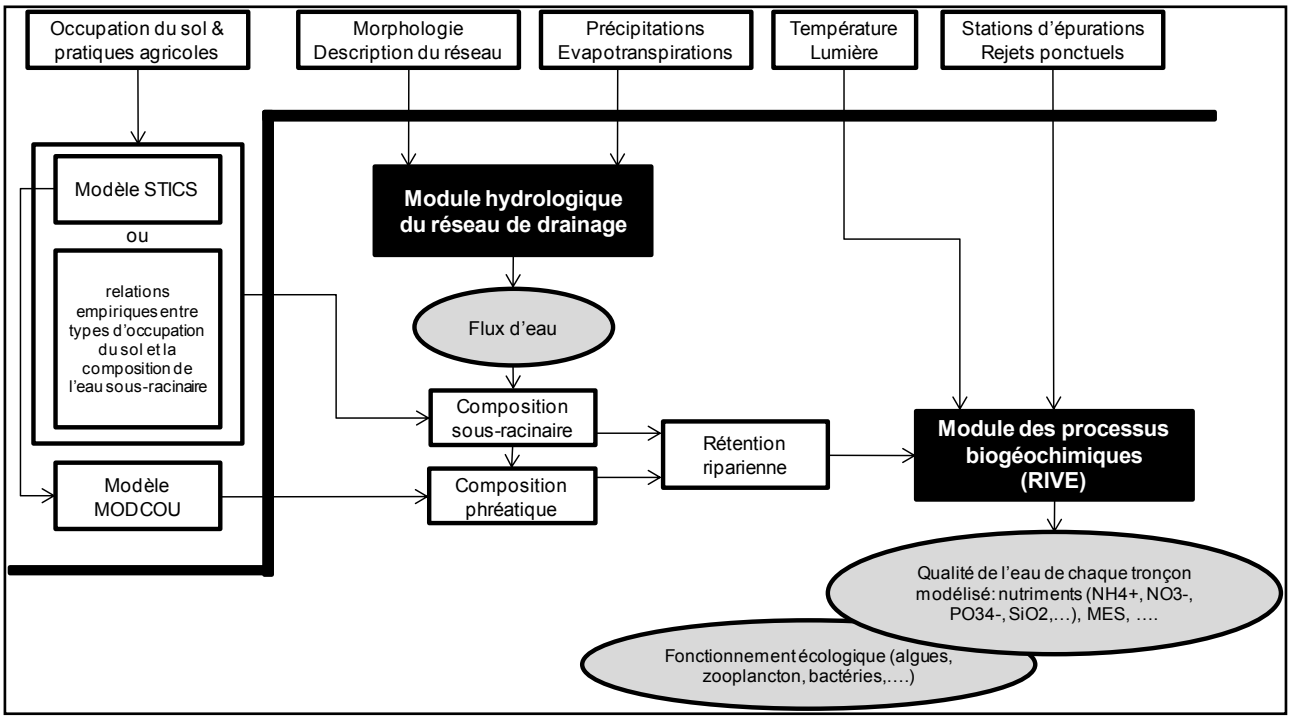


Figure 3.2 : Principes du modèles SENEQUE Ruelland, 2004.

Le modèle SENEQUE est couplé à un SIG. Ce couplage vise à rassembler au sein d'une interface conviviale commune les diverses fonctions d'un système intégré d'information et d'aide à la décision sur la qualité de l'eau. Ce système permet à la fois de visualiser et d'intégrer les bases de données géographiques, d'en extraire les informations pour un projet de modélisation défini, de piloter le déroulement de ce projet, de simuler des scénarios prospectifs et enfin de visualiser les résultats de la modélisation sous forme graphique et cartographique (Ruelland, 2004 ; Ruelland *et al.*, 2007).

CemOA : archive ouverte d'Irstea / Cemagref

II.3. Les modèles ciblés sur le rôle épuratoire des ripisylves

C'est au milieu de la décennie 1980 que des travaux de recherche fondamentale ont débuté sur les capacités de prélèvement direct par les végétaux de l'azote circulant dans l'eau phréatique. Les premiers résultats ont été publiés dès 1983 (Lowrance *et al.*, 1983; Lowrance *et al.*, 1984 ; Peterjohn and Correll, 1984; Jacobs and Gilliam, 1985). Ces premiers travaux évaluent les capacités de stockage des arbres et montrent que la part d'azote prélevée par les végétaux d'une ripisylve contribue pour un tiers à la diminution de la teneur en azote de la nappe. Dès lors, les études se multiplient, se diversifient et se complexifient (Shearer and Xiang, 2007) cherchant à identifier les processus. De nombreux laboratoires de recherche se sont intéressés à ces questions, d'autant plus que l'extension et la fréquence des pollutions diffuses, particulièrement des pollutions azotées dans les milieux aquatiques, inquiètent sérieusement les responsables de la santé publique (Ruffinoni *et al.*, 2003).

Le Riparian Ecosystem Management Model (REMM) (Lowrance *et al.*, 2000 ; Altier *et al.*, 2002) développé par le département de recherche agricole des Etats-Unis (USDA-ARS) constitue un bon exemple de modèle mécaniste. Il permet de quantifier l'azote et le phosphore retenus dans la zone rivulaire en considérant les conditions spécifiques du site, et guide ainsi les gestionnaires dans les mesures à prendre pour améliorer la capacité épuratoire de ces zones. Il peut donc être appliqué par exemple pour déterminer la largeur de buffer nécessaire suivant les conditions de la zone et les charges des intrants, déterminer son efficacité en augmentant ces charges, évaluer l'influence des types de végétation sur cette efficacité, analyser l'impact de la suppression des ripisylves…

Ce modèle est adapté à l'échelle de la parcelle agricole ou d'un versant (en plaine, piémont ou vallée) et donne une réponse au pas de temps journalier (avec des simulations pouvant aller jusqu'à 100 ans). Dans ce modèle, la zone rivulaire est caractérisée en utilisant une séquence de trois réservoirs. Chaque réservoir est supposé représenter une zone de végétation différente et est divisé (verticalement) en trois types de sols (Figure 3.3).

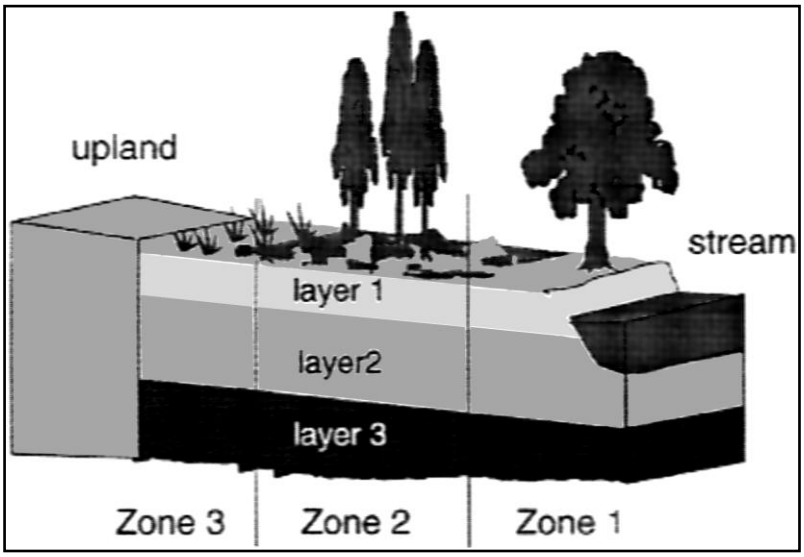


Figure 3.3 : Caractérisation de la zone rivulaire dans le Riparian Ecosystem Management Model (REMM). Ce schéma a été conçu à partir du modèle de buffer en trois bandes d'après Welsch, 1991.

Dans ce cadre physique, un grand nombre de processus sont décrits (une cinquantaine) à travers plusieurs modèles. On peut en distinguer quatre grands types (Williams, 2007): (1) les processus hydrologiques (interception par la canopée de la végétation et la litière, évapotranspiration et transpiration, infiltration et drainages verticaux, écoulements de surface et de subsurface) ; (2) les processus du transport de sédiments (érosion causée par l'impact des gouttes de pluie, érosion diffuse, érosion concentrée et érosion de subsurface...) ; (3) les processus de la dynamique des nutriments pour l'azote, le phosphore et le carbone

CemOA : archive ouverte d'Irstea / Cemagref

(prélèvement racinaire, absorption/désorption, minéralisation et immobilisation …) ; et (4) les processus de croissance des végétaux arborés (conifères et bois durs) et herbacés (photosynthèse, respiration, allocation du carbone…). Compte tenu du nombre important de processus modélisés, un grand nombre de variables d'entrée sont nécessaires si on veut utiliser l'ensemble des capacités de ce modèle (Inamdar, 2006).

D'autres modèles mécanistes ont été développés. On trouve dans la littérature de nombreuses publications cherchant à quantifier par modélisation la réduction de la charge en nutriments (pourcentage calculé par la différence entre l'entrée et la sortie du système rapportée aux apports d'entrée) et évaluer cette capacité en fonction de la taille de la zone rivulaire (largeur et longueur) et du type de couvert. Le Tableau 3-1 présente les résultats d'une synthèse bibliographique d'après Ruffinoni *et al.*, 2003.

Cependant, malgré l'abondance de ces recherches, la littérature est souvent inadaptée pour guider de façon objective les gestionnaires sur les stratégies d'action pour améliorer la capacité épuratrice de ces zones (Correll, 2000). A partir de résultats issus d'un grand nombre de publications évaluant l'efficacité des zones rivulaires (par modélisation ou expérimentalement), Mayer *et al.*, 2007 ont cherché à définir empiriquement les relations entre la largeur du buffer et le taux d'élimination de l'azote en fonction du type de végétation, et du mode d'écoulement sur (Figure 3.4).

Tableau 3-1 : Synthèse bibliographique des recherches sur le pouvoir épurateur des zones alluviales vis-vis des nutriments, azote et phosphore d'après Ruffinoni *et al.*, 2003.

Couvert végétal	Élément	Compartiment étudié	Largeur de la zone d'influence	Réduction de la charge %	Auteurs
Ripisylve	Azote	Eau de nappe	25 m	68	Lowrance et al., 1984
Ripisylve	Azote	Eau de nappe	19 m 50 m	93 99	Peterjohn et Corell, 1984
Ripisylve	Azote	Eau de nappe	47 m 16 m	98 93	Jacobs et Gilliam, 1985
Végétation riveraine	Azote	Eau de nappe	19 m	50	Schnabel, 1986
Ripisylve / marais	Azote	Eau de nappe	16 m	99	Cooper et Thomsen, 1988
Aulnaies	Azote	Eau de nappe	10 m	70	Pinay et Labroue, 1986
Ripisylve	Azote	Eau de nappe		50 à 60 90	Cooper, 1990
Végétation riveraine	Azote	Eau de nappe		99	Fustec et al., 1991
Végétation riveraine	Azote Phosphore	Sols alluviaux			Ruffinoni, 1991
Peupliers **Prairies**	Azote	Eau de nappe	5 m 17 m	100 84	Haycock et Pinay, 1993
Ripisylve	Azote	Eau de nappe	60 m	95	Jordan et al., 1993
Forêt alluviale	Azote Phosphore	Eau de nappe	5 à 100 m	100	Sanchez-Perez, 1992
Ripisylve **Forêt alluviale** **Prairie humide**	Azote	Eau de nappe	50 à 380 m	10 à 95	Ruffinoni, 1994
Ripisylve	Azote	Eau de nappe ruissellement		80	Vought et al., 1995
Ripisylve	Azote	Sols alluviaux			Pinay et al., 1995
Ripisylve	Phosphore	Sols alluviaux			Fabre et al., 1996
Ripisylve **Forêt alluviale** **Prairie humide**	Azote Phosphore	Eau de nappe	2 à 100 m	37 à73 62 à 76	Takatert et al., 1999

CemOA : archive ouverte d'Irstea / Cemagref

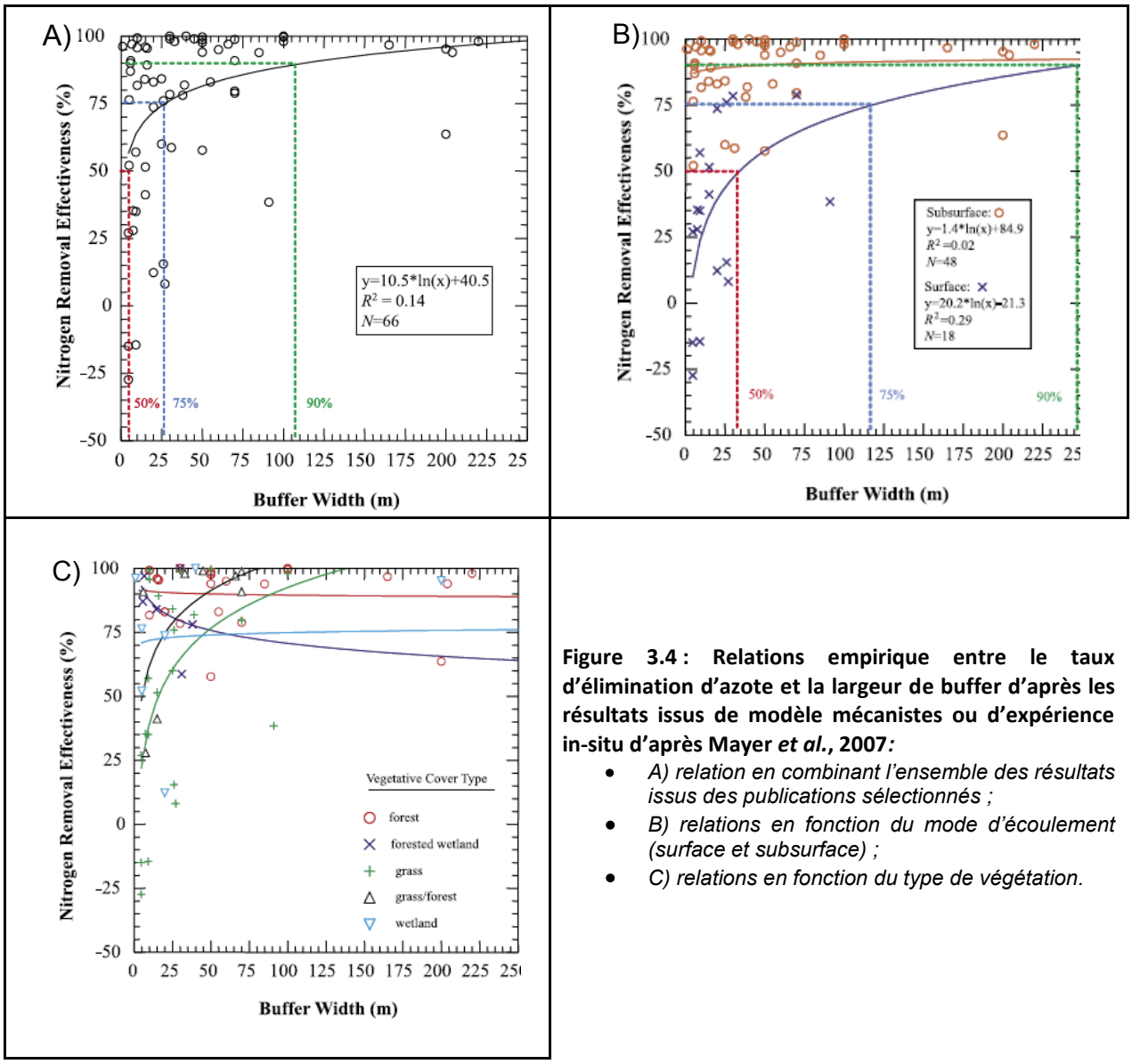


Figure 3.4 : Relations empirique entre le taux d'élimination d'azote et la largeur de buffer d'après les résultats issus de modèle mécanistes ou d'expérience in-situ d'après Mayer *et al.*, 2007*:*

- *A) relation en combinant l'ensemble des résultats issus des publications sélectionnés ;*
- *B) relations en fonction du mode d'écoulement (surface et subsurface) ;*
- *C) relations en fonction du type de végétation.*

D'après cette analyse, une petite, mais significative, proportion de la variabilité du taux d'élimination d'azote est fonction de la largeur de la bande rivulaire (R^2 = 0.14 et N =66). On observe une diminution drastique dès les premiers mètres (50 % pour 5 m et 75 % pour 25 à 30 m) et presque complète (90 %) pour une bande de 100 m. Cette influence diffère en fonction du type d'écoulement (plus efficace en écoulement de surface) et du type de végétation (les formations herbacées sont moins efficaces que les arborées) pour l'azote.

II.4. Les modèles « mécanistes » : un intérêt limité pour la mise en œuvre de la DCE

Au vu des exemples de modélisation présentés précédemment, et notamment du programme PIREN-SEINE, des progrès ont été réalisés au cours des quinze dernières années dans la connaissance du fonctionnement des écosystèmes et permettent aujourd'hui d'expliciter quantitativement sous forme de modèles mathématiques les contraintes que le climat, la topographie, la géologie et les activités humaines, agricoles et urbaines, exercent sur un grand bassin versant (au sens de la DCE sur un district hydrographique). Ces modèles peuvent être utilisés de manière opérationnelle par les gestionnaires pour certaines applications de gestion. C'est le cas, par exemple, du Syndicat Interdépartemental pour l'Assainissement de l'Agglomération Parisienne (SIAAP) qui a fait usage un grand nombre de fois du modèle ProSe pour gérer la qualité de l'eau dans Paris et son

CemOA : archive ouverte d'Irstea / Cemagref

agglomération. Les simulations de ce modèle ont amélioré la stratégie d'installation de diffuseurs à oxygène dans le fleuve pour prévenir des anoxies (Wasson *et al.*, 2003b). Ces modèles permettent ainsi de prendre en compte l'ensemble des éléments de la logique DPSIR (cf. I.2) pour un impact donné (dans cet exemple, l'anoxie du fleuve).

Cependant, compte tenu du délai imparti pour la mise en œuvre de la DCE et de la nouvelle définition de la qualité de l'eau (cf. Chapitre 1) ces approches sont inadaptées pour deux raisons essentielles.

La première raison est liée aux difficultés d'appliquer ces modèles sur de grands territoires et de les transférer sur une diversité de contextes socio-économiques (Wasson *et al.*, 2003b).

D'un point de vue technique, un grand nombre de données (particulièrement celles concernant le sous-sol) sont inaccessibles et leur acquisition est généralement irréalisable dans un contexte opérationnel. Il est alors nécessaire de calibrer (et valider) certains paramètres ce qui engendre rapidement des problèmes numériques associés (équifinalité des modèles, sensibilité aux paramètres) qui s'avèrent parfois insolubles (Beven, 1989). De plus, les résultats sont imprécis et incertains compte tenu des difficultés de gestion des incertitudes dans ces modèles (représentations mathématiques des processus géochimiques, transport des particules solides, spatialisation des paramètres, . . .) (Roche and Zimmer, 2006).

D'un point de vue conceptuel, aussi bien la discrétisation en maille que les lois physiques utilisées sont inadaptées pour modéliser les écoulements sur de grands territoires (Beven, 1989). Par exemple la loi de Darcy, largement utilisée dans ces modèles, a été développée pour des milieux homogènes et isotropes, et son extension à une grille kilométrique est loin d'être évidente (Wasson *et al.*, 2003b).

Dans le cas du programme PIREN-SEINE, ces difficultés sont atténuées compte tenu de la relative homogénéité des conditions naturelles (climat, géologie, relief) du bassin versant de la Seine (Guerrini *et al.*, 1998). Mais la France est un milieu diversifié ce qui complexifie l'utilisation de tels modèles pour mener à bien la DCE dans le délai imparti.

La deuxième raison est que ces modèles ne répondent que partiellement aux objectifs de qualité de la DCE (Wasson *et al.*, 2003b ; Even *et al.*, 2007). Ces modèles sont capables aujourd'hui de simuler le transport, la dispersion d'un certain nombre de paramètres biogéochimiques et de substances toxiques (métaux, médicaments, HAP) évaluant l'état chimique de la rivière mais ne permettent pas de prédire la présence et la diversité biologiques, paramètres clés de l'évaluation de l'état écologique au sens de la DCE.

Les modèles cherchant à expliquer les liens entre l'état chimique et l'état biologique, tels que les modèles éco-toxicologiques (e.g, modèle de relations doses-effets Magaud *et al.*, 1997) développés à partir d'expériences en laboratoire sont en limite d'utilisation (Figure 3.5). Ils ne permettent pas de simuler la biodisponibilité des polluants en conditions réelles (à partir des sorties des modèles biogéochimiques) et la description de l'état écologique (sortie de ces modèles) est insuffisante puisqu'elle reste à l'échelle de l'individu et non pas à celle de la population et de l'écosystème.

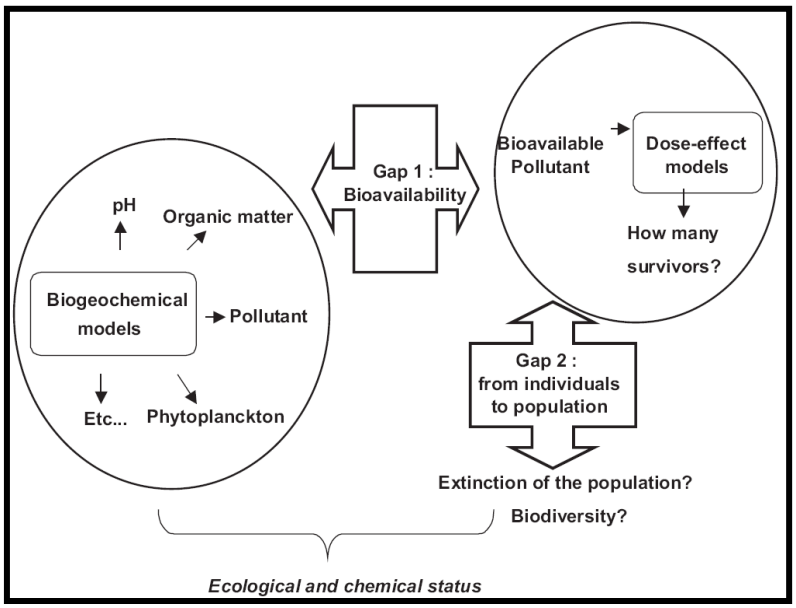


Figure 3.5 : Les manques de connaissances dans les approches mécanistes pour évaluer l'état écologique des cours d'eau au sens de la DCE d'après Wasson *et al.*, 2003b.

CemOA : archive ouverte d'Irstea / Cemagref

Par ailleurs, malgré les efforts effectués pour prendre en compte le rôle épuratoire des zones rivulaires (exemple du modèle SENEQUE, cf. II.2), il apparaît très difficile de préciser et de quantifier à l'échelle du bassin les processus impliqués dans l'élimination et la rétention des nutriments (prélèvement racinaire, dénitrification…), tant ceux-ci sont imbriqués les uns aux autres et sont contrôlés par de multiples paramètres qui régissent le fonctionnement des écosystèmes alluviaux (exemple du modèle REMM).

C'est pour cela que dans la plupart des cas, les modèles cherchant à décrire ces processus ont été appliqués sur des systèmes de petite taille (Inamdar, 2006), de l'ordre de 30 km² (Lowrance *et al.*, 1983) à 3000 km² (Peterjohn and Correll, 1984). Dès que les recherches se sont orientées vers des hydrosystèmes plus vastes, plus de 10 000 km² dans le cas de la Garonne, près de 190 000 km² pour le bassin versant du Rhin, le problème s'est considérablement compliqué (Ruffinoni *et al.*, 2003).

De plus, ces modèles décrivent seulement l'influence des corridors rivulaires sur les processus biogéochimiques. Leur influence sur les processus physiques (diversification de l'habitat) et écologiques (contrôle trophique et thermique) n'est pas prise en compte.

Pour autant ces approches qui s'attachent à comprendre et à décrire finement les phénomènes biophysiques et les interactions entre sous-systèmes sont cruciales pour valider l'ensemble des concepts et guider la construction des modèles plus globaux (approche empirique).

III. Les modèles empiriques

Les modèles physiques butent sur des difficultés importantes lorsqu'on essaie de les appliquer sur de vastes territoires. Il est en effet difficile, à ces échelles, d'appliquer des approches expérimentales, de décrire et modéliser l'ensemble des mécanismes, de caractériser et mesurer l'ensemble des variables, et de gérer la surparamétrisation des modèles. Au-delà de ces aspects combinatoires liés à la dimension du territoire, on risque d'omettre des processus déterminants à une échelle intermédiaire, liés à la complexité des interactions entre échelles. Ceci rend les résultats de la modélisation physique incertains sur de vastes territoires et oblige à rechercher et développer une approche de modélisation alternative et complémentaire : la modélisation empirique.

III.1. Principes et concepts

L'objectif de l'approche de modélisation empirique est de représenter des relations entre les entrées (pressions humaines) et l'état des systèmes (l'état écologique des cours d'eau), pertinentes sur de grands territoires. Son principe, reposant sur l'approche macroécologique (Brown, 1995), est d'identifier ces relations (appelées aussi « patterns » ou « modes d'organisation ») en faisant émerger des relations statistiques de grandes séries de données de pressions et de données écologiques recueillies sur de vastes territoires. Cette approche, alternative aux modèles physiques, permet d'aborder les phénomènes à différentes échelles spatio-temporelles, notamment à l'échelle régionale (ou large échelle au sens de la dimension du territoire), et de faire des déductions pertinentes concernant le fonctionnement de la nature (Lévêque, 2001). L'écologie des paysages, introduite par Troll, 1939, apporte les outils conceptuels et techniques pour étudier ces relations pressions/impacts à travers différentes échelles spatiales en se basant sur la hiérarchie emboitée des écosystèmes d'eau courante (concept de l'habitat).

III.1.1. L'écologie des paysages

Depuis Hynes, 1975 (« in every respect, the valley rules the stream »), les hydro-écologues reconnaissent que les rivières sont influencées non seulement par les éléments du paysage aquatique, notamment au niveau des corridors rivulaires par leur position d'interface (Naiman and Décamps, 1990), mais aussi par l'ensemble de leur bassin versant. L'étude des rivières à travers leur paysage continue d'évoluer (Allan, 2004a), par le besoin

CemOA : archive ouverte d'Irstea / Cemagref

d'étudier leur fonctionnement à des échelles larges permettant de guider les gestionnaires dans leurs décisions, et grâce à l'émergence de l'écologie du paysage comme discipline à part entière (Wiens, 1989 ; Turner *et al.*, 2001 ; Burel and Baudry, 1999).
L'écologie des paysages étudie les relations entre les modes d'organisation spatiale des écosystèmes (patterns) et les processus écologiques. Dans cette approche, on considère que les structures (les formes) spatiales qui composent le paysage aquatique (l'hydrosystème) sont en interaction constante avec les processus fonctionnels des écosystèmes d'eau courante tout en étant pour partie dépendantes des relations entre la société (les activités humaines) et le milieu environnant. L'analyse de la composition, de l'organisation et de la connectivité de ces structures à travers différentes échelles spatiales à partir d'indicateurs du paysage permet de rechercher la gamme d'échelles pour laquelle un ou plusieurs facteurs donnés (naturel ou anthropique) sont pertinents pour expliquer un processus fonctionnel (Burel and Baudry, 1999).

Dès lors, l'enjeu dans les modèles pressions impacts empiriques est de définir des indicateurs du paysage qui traduisent la diversité des mécanismes (voies multiples et variées) par lesquels l'homme modifie et altère l'état écologique des cours d'eau et d'étudier leur relations statistiques avec des indices biologiques (répondant à ces mécanismes) afin d'identifier et de hiérarchiser les causes d'altérations à travers les différentes échelles structurant le paysage aquatique (Figure 3.6).

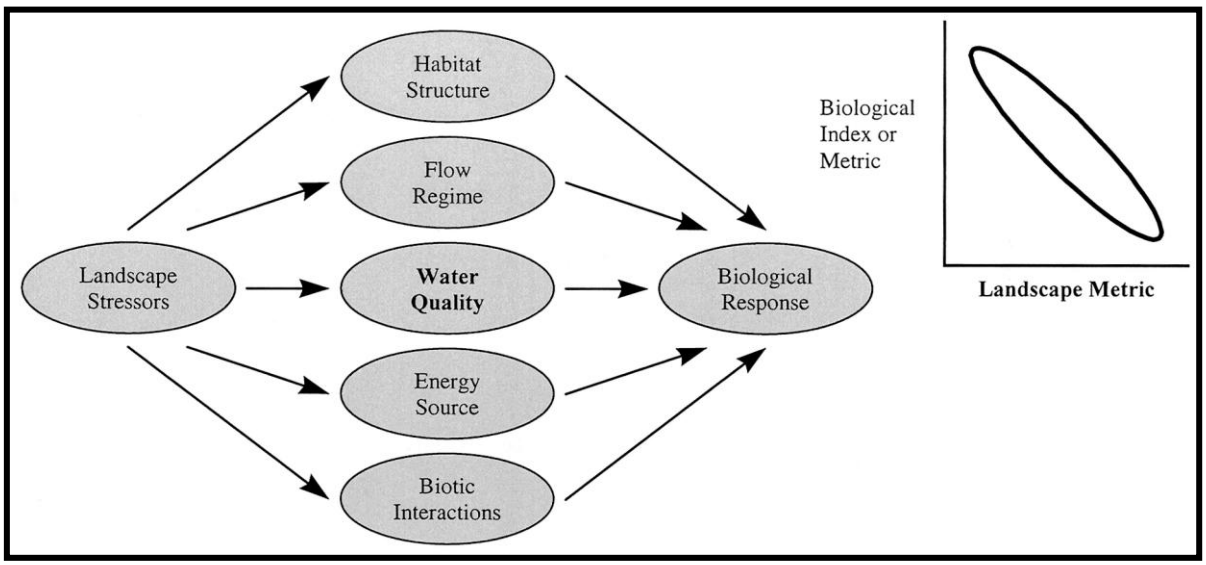


Figure 3.6 : Principe des modèles pressions impacts empiriques. L'ensemble des voies doivent être prises en compte pour établir des relations significatives entre un indicateur du paysage et une réponse biologique d'après Allan, 2004b.

III.1.2. La hiérarchie des systèmes aquatiques (concept de l'habitat)

L'existence des échelles multiples et de leurs interactions n'est plus contestée en écologie des écosystèmes. Leur complexité, et notamment l'imbrication des échelles de fonctionnement, conduit logiquement à considérer plusieurs échelles d'observation pour analyser les structures d'un même système (Lévêque, 2001). C'est pourquoi divers auteurs (Allen and Starr, 1982 ; O'Neill *et al.*, 1986) ont préconisé l'application de la théorie des hiérarchies à l'étude des systèmes écologiques, arguant que tout objet d'étude en écologie est composé d'un ensemble de sous-systèmes et constitue lui-même un élément d'un ensemble plus vaste.

Une hypothèse importante de la théorie des hiérarchies est que les unités constituant les différents niveaux hiérarchiques n'ont pas de dynamiques indépendantes. Les processus qui ont lieu à des niveaux supérieurs exercent des contraintes sur la dynamique des processus intervenant au niveau inférieur. L'organisation hiérarchique des systèmes écologiques peut être effectivement perçue comme un emboîtement d'entités d'ordre inférieur dans des entités plus vastes d'ordre supérieur, à la manière de poupées russes.

CemOA : archive ouverte d'Irstea / Cemagref

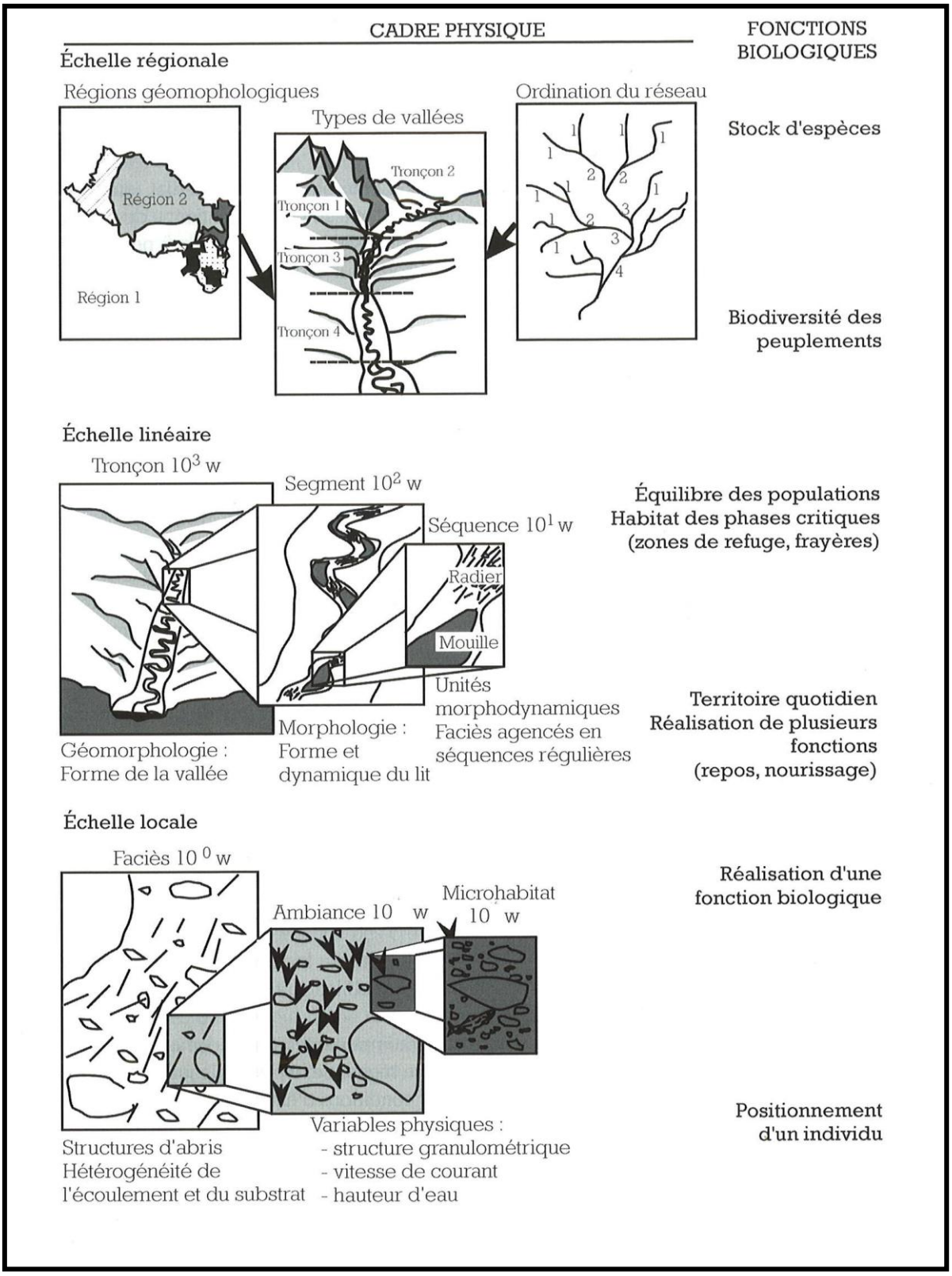


Figure 3.7 : Habitat et morphologie : des échelles emboitées d'après Wasson *et al.*, 1998.

CemOA : archive ouverte d'Irstea / Cemagref

La hiérarchie des écosystèmes aquatiques a été considérablement explorée par la littérature (Allan, 2004a) et bien définie dans le texte Frissell *et al.*, 1986. Elle se base sur le concept d' « habitats » des communautés biologiques. L'habitat est la structure physique perçue par les être vivants et constitue un compartiment « clé » de l'écosystème rivière et « central » (Wasson *et al.*, 1998) car il conditionne aussi pour une large part la disponibilité des ressources trophiques et de l'oxygène (autres facteurs clés de l'écosystème).

La diversité de l'habitat aquatique répond à trois niveaux emboités de la structuration morphologique des rivières : ponctuel, linéaire et régional (Figure 3.7). A l'échelle régionale, la géomorphologie du bassin et la structure fractale du réseau (représenté par le numéro d'ordre) génèrent différents types de vallées. Des discontinuités majeures définissent des tronçons homogènes. A l'échelle linéaire d'un tronçon, les variations locales (pente, substrat, végétation) sous l'action de facteurs naturels et anthropiques définissent des segments, souvent répétitifs, caractérisés par des séquences propres de faciès morphodynamiques. A l'échelle ponctuelle du faciès, l'hétérogénéité longitudinale et transversale de l'écoulement et les singularités physiques génèrent différentes ambiances (hétérogènes) et microhabitats (à la taille des individus). Chacune de ces échelles permet la réalisation de fonctions biologiques nécessaires au maintien de la diversité des peuplements.

Le concept d'habitat est la passerelle la plus utilisée pour étudier les corrélations entre le fonctionnement physique et le fonctionnement écologique des cours d'eau (Souchon, 2002 ; Allan, 2004a; Johnson and Host, 2010).

III.2. Revue des modèles existants

La publication du récent ouvrage « Landscapes influences on Stream Habitats and Biological Assemblages » en 2006 par la « American Fisheries Society » (Hughes *et al.*, 2006) et de la revue bibliographique « Recent developments in landscape approches for the study of aquatic ecosystems » en 2010 par L. Johnson & G. E. Host dans le « Journal of North American Benthologocal Society » montre bien que la modélisation empirique à large échelle des relations pressions/impacts est une question de recherche au niveau international. En Europe, on commence à être attentifs au développement de ces modèles comme en témoigne le travail mené en France, en Estonie, en Slovaquie et au Royaume-Uni publié par Wasson *et al.*, 2010). Cependant, on trouve relativement peu de travaux publiés au niveau européen, comparés aux Etats-Unis, si l'on regarde les références des différentes revues bibliographiques sur ce sujet (e.g., Allan, 2004a, Johnson and Host, 2010). Ceci s'explique en partie par la récente mobilisation des scientifiques des états membres pour mettre en place des méthodes d'évaluation de la qualité écologique compatibles avec la directive, avec en parallèle un travail d'intercalibrage des indices d'évaluation entre chaque état européen (Garcia *et al.*, 2005). Dans cette partie, on s'intéressera tout d'abord aux travaux menés à l'extérieur du territoire national, avant de synthétiser la démarche et les principaux résultats de l'approche développée en France par Wasson *et al.*, 2005).

III.2.1.Synthèse des travaux internationaux

On présente successivement, les données de pressions, les données biologiques, les échelles et zones de travail, les méthodes d'analyse de données employées et une synthèse des résultats des différents travaux analysés.

III.2.1.1 Les données de pressions

Au niveau d'une nation entière, on ne dispose pas à l'heure actuelle de données détaillées sur les diverses pressions renseignant tous les compartiments incriminés, aux différentes échelles appropriées. Lorsque l'information existe, elle n'est pas forcément homogène à l'échelle de travail (grand bassin versant, région, nation) et non spatialisable (i.e. qu'on peut relier topologiquement avec les stations écologiques).

A l'échelle du tronçon, il existe de nombreux protocoles terrain évaluant ces pressions (Maddock, 1999) comme le RHS (River Habitat Survey, Raven *et al.*, 1997) au Royaume Uni, le protocole AusRivAS (Australian River Assessment Scheme, Parsons *et al.*, 2004) ou encore plus récemment le protocole EMAP (Environmental

CemOA : archive ouverte d'Irstea / Cemagref

Monitoring and Assesment Program) aux Etats-Unis (Stoddard *et al.*, 2005). Ils impliquent généralement la surveillance de critères liés aux caractéristiques hydromorphologiques du chenal, aux berges et à l'occupation des sols sur des tronçons de longueur inférieure à 1 km afin de produire des indicateurs de qualité de l'état physique du cours d'eau.

Cependant, ces méthodes, consommatrices de temps, sont difficilement utilisables pour quantifier les données de pressions pour l'ensemble des stations, que ce soit à l'échelle du tronçon ou à l'échelle linéaire (ensemble du corridor rivulaire à l'amont de la station). De plus, la plupart des protocoles ne permettent pas d'obtenir le même niveau d'information sur les différents types de pressions, altérations chimiques et hydromorphologiques, pollutions ponctuelles ou diffuses (Gergel *et al.*, 2007).

Les différentes études pressions/impacts cherchent donc à réduire la complexité en ne mobilisant que les informations homogènes disponibles régionalement ou nationalement au sein de la chaîne de causalité. L'occupation des sols (urbain, industrie, agriculture…), apportant une information spatialisée et homogène sur de grandes surfaces, permet de caractériser la plupart des forces motrices ou grandes structures qui génère les principales pressions s'exerçant sur le milieu aquatique (cf. I.2). Par exemple, la présence de territoires artificialisés peut traduire à la fois des pressions hydromorphologiques (e.g., imperméabilisation des sols, digues, suppression de la ripisylve…) et chimiques (e.g. apport de matière organique, pollution toxique de façon ponctuelle ou diffuse) qui perturbent le fonctionnement et la structure physique du cours d'eau.

La plupart des modèles pressions/impacts à large échelle développés aujourd'hui, utilisent donc l'occupation du sol comme information représentative des pressions. Au sens de la logique DPSIR, ces modèles étudient alors les relations forces-motrices/états (D/S) (on continuera dans la thèse à utiliser le terme générique de pressions/impacts). L'occupation des sols est quantifiée à l'aide des outils de la télédétection et des Systèmes d'Information Géographique (SIG). Ce sont d'ailleurs les progrès grandissants dans ces domaines qui ont permis l'émergence de ces modèles (Johnson and Gage, 1997).

La télédétection est définie comme « l'ensemble des connaissances et techniques utilisées pour déterminer des caractéristiques physiques et biologiques d'objets par des mesures effectuées à distance, sans contact matériel avec ceux-ci » (JO du 11 décembre 1980). Elle mobilise un ensemble d'outils et de méthodes qui permettent d'observer l'environnement terrestre à distance via des capteurs embarqués (satellitaires ou aéroportés). Le principe général de la cartographie par télédétection repose sur l'extraction d'information géographique à partir des données contenues dans les images acquises à un instant donné. L'image numérique de télédétection est composée de pixels (particule élémentaire de l'image) d'une dimension donnée (résolution spatiale). Classer une image de télédétection pour obtenir une carte d'occupation du sol consiste à identifier des objets géographiques à partir des valeurs numériques des pixels qui traduisent l'intensité du signal électromagnétique capté et réfléchi. La partie 2 du rapport détaille les méthodes correspondantes.

Les cartes d'occupation des sols comme National Land Cover Dataset aux Etat-Unis (Vogelmann *et al.*, 2001) ou CORINE Land Cover en Europe (Bossard *et al.*, 2000), réalisées par interprétation d'images satellites à moyenne ou haute résolution (e.g, Landsat Thematic Mapper : 30 m ; Spot 4 : 20 m) ont été extrêmement utilisées pour extraire l'information d'occupation du sol au niveau du bassin versant et du corridor rivulaire (Congalton *et al.*, 2002 ; Goetz, 2006). De façon générale les satellites sont très employés pour inventorier les activités humaines sur de grand territoires (Congalton *et al.*, 2002).

A partir de ces cartes, de nombreux indicateurs spatialisés quantifiant les pressions d'occupation du sol peuvent être construits en mettant en œuvre des techniques implémentées dans les SIG aux différentes échelles d'analyses (Johnson and Gage, 1997 ; Gergel *et al.*, 2002). Ces indicateurs sont invariablement construits par l'agrégation d'un attribut de structure (métrique du paysage) dans une étendue délimitée (emprise spatiale). Les métriques de paysage peuvent être par exemple le nombre de types de postes d'occupation du sol, la proportion de chaque type, la forme des patchs, l'arrangement spatial et la connectivité des patchs (Li and Reynolds, 1995). La partie 3 du rapport détaille les méthodes correspondantes.

Un des intérêts majeur de l'utilisation de l'occupation du sol est de permettre une affectation plus précise des différentes pressions aux bassins versants réels et aux corridors rivulaires des stations biologiques. Cependant, d'une part, elle ne permet pas de quantifier précisément l'intensité des pressions qu'elle génère : par exemple, aux zones de grandes cultures peuvent être associées les pressions de pollutions diffuses (nutriments, pesticides…), mais à l'exception des nitrates, la relation quantitative entre superficie cultivée et rejets polluants est loin d'être simple ; et d'autre part, elle ne permet pas d'identifier réellement les facteurs

CemOA : archive ouverte d'Irstea / Cemagref

incriminants puisqu'elle indique un éventail de pressions possibles (i.e., qu'on « suspecte ») s'exerçant de façon combinée suivant divers mécanismes sur le milieu aquatique.

Son utilisation dans les modèles pressions/impacts ne permet donc pas d'établir un diagnostic très précis de l'état écologique qui permettrait d'identifier les leviers d'actions appropriées pour atteindre les objectifs de la directive cadre. Cependant, elle apporte une première vision en identifiant, hiérarchisant et localisant (aux différentes échelles) les principales structures influençant l'état écologique sur un territoire. Cette vision est essentielle dans le processus de gestion d'un territoire pour cibler les priorités d'action.

III.2.1.2 Les données biologiques

Dans la littérature concernant l'évaluation de la qualité écologique, aucun consensus quant au choix de méthodes d'évaluation d'un site n'est établi.

Bien que les critères diffèrent, les approches (1) multi-métriques américaines (Rapid Bioassessment Protocol, Kerans and Karr, 1994) ou européennes (Assessment system for the ecological Quality of streams and rivers throughout Europe using Macroinvertebrates, AQEM, Hering *et al.*, 2004) et (2) multi-variées telles que RIVPACS (River Invertebrate Prediction and Classification System, Wright *et al.*, 1993) élaborée au Royaume-Uni ou encore AUSRIVAS (Australian River Assessment Scheme, Parsons and Norris, 1996) développée en Australie, permettent toutes de comparer les conditions écologiques actuelles d'un écosystème avec ses conditions de référence (Grenier, 2007).

Qu'elles soient multimétriques ou multivariées, les approches de biosurveillance nécessitent (i) de travailler sur des zones présentant des conditions homogènes, notamment au regard des biotes, de l'utilisation du territoire et des types de perturbations majeurs (Klemm *et al.*, 2003) ; (ii) de calibrer dans chacune de ces zones les indicateurs par rapport à ceux des stations de référence.

Les approches multimétriques et multivariées diffèrent principalement en raison de la classification des sites de référence, a *priori* (multi-métriques) versus a *posteriori* (multi-variées) (Dion, 2009) : la classification des sites a priori découle d'une sélection préliminaire, où les hypothèses et caractéristiques de la structure des communautés sont soulevées et envisagées, tandis que la sélection a posteriori résulte essentiellement de l'analyse de la distribution des variables biologiques suite à un traitement statistique. Brièvement, l'approche multivariée relève les changements dans la structure des communautés en fonction d'un gradient de dégradation, tandis que l'approche multimétrique inclus divers indices simples et variables biologiques, ayant des degrés distincts de sensibilité à la pollution (Pyle, 1999).

Dans la littérature, l'impact de l'occupation du sol est généralement corrélé avec la réponse biologique des poissons ou des macroinvertébrés benthiques, et quelquefois avec les communautés algales. Généralement, on note que les auteurs utilisent préférentiellement les indices multimétriques même si certains travaux utilisent les approches multi-variées (e.g Frimpong et al., 2005). Intuitivement plus simple à assimiler et à produire par les gestionnaires de la qualité de l'eau, l'approche multimétrique permet de fournir, par rapport à l'approche multivariée, des liens plus détaillés entre les causes des perturbations et les effets sur le biote (Boulton, 1999 ; Grenier, 2007).

Par exemple, l'utilisation de l'index d'intégrité biotique (IBI) élaboré par Karr, 1981 permet d'évaluer la santé de l'écosystème de la rivière en fonction des caractéristiques des peuplements de poissons, à partir de différentes caractéristiques des peuplements (ou métriques) sensibles à l'intensité des perturbations anthropiques. Il rend compte notamment de la composition taxonomique, de la structure trophique et de l'abondance des espèces. Les peuplements de macroinvertébrés benthiques sont eux évalués à travers des indices de qualité classiques comme le BMWP (Biological Monitoring Working Party) au Royaume-Uni (Bis *et al.*, 2000), l'indice EPT (nombre total de Ephemeroptera-Plecoptera-Trichoptera) utilisé dans de nombreux pays (Stewart *et al.*, 2001), ou encore des indices multimétriques comme le ICI (Invertebrate Community Index) aux États-Unis (Cormier *et al.*, 2000). Les indices de diversité (Shannon, Hisenhoff) sont parfois utilisés pour caractériser les peuplements biologiques (Wang *et al.*, 2001 ; Wang and Kanehl, 2003). Concernant les diatomées, Pan *et al.*, 2004 ont testé plusieurs métriques auto-écologiques[20] (basé sur Lange-Bertalot, 1979 et van Dam *et al.*, 1994), et le TDI (Trophic Diatom Index).

[20] Métriques relatives à l'environnement de l'espèce étudié (climat, lumière, conditions édaphiques, conditions hydriques)

CemOA : archive ouverte d'Irstea / Cemagref

III.2.1.3 Echelles et zones de travail

Dans ces modèles, le bassin versant amont des stations est généralement l'unité d'étude des relations pressions/impacts. Les approches diffèrent en fonction du mode de spatialisation utilisé pour identifier au sein de ce bassin versant les emprises dans lesquelles l'occupation du sol possède l'influence la plus significative sur l'état écologique. On distingue deux approches.

La première approche, initié par l'USGS (U. S. Geological Surve dans le modèle SPARROW (Smith *et al.*, 1997), prend en compte la distance à la station et au réseau hydrographique d'un type d'occupation du sol en tout point du bassin versant. Elle assume que l'influence d'un type d'occupation du sol à un point donné dans le bassin versant décroit en fonction de sa distance à la station de mesure et au cours d'eau (Van Sickle and Johnson, 2008). Les modèles dérivant de cette approche se différencient selon la méthode pour modéliser les distances à la station et au cours d'eau et le mode de pondération de la fonction décroissante qui cumule l'influence d'une catégorie d'occupation du sol donnée.

Par exemple, King *et al.*, 2005 ont utilisé les distances euclidiennes (Figure 3.8) pour modéliser l'influence d'une parcelle d'occupation du sol sur le bassin par rapport à sa proximité à la station et au cours d'eau. Par contre, Johnson *et al.*, 2007 et Van Sickle and Johnson, 2008 ont utilisé deux modèles de distances distincts pour estimer la distance entre une parcelle d'occupation du sol donnée sur le bassin et la station : (1) la distance dérivant de la direction de flux estimées à l'aide d'un Modèle Numérique de Terrain (MNT) pour estimer la distance entre la parcelle et le point de réseau hydrographique le plus proche ; puis (2) la distance curviligne du réseau hydrographique pour estimer la distance entre ce point d'impact sur le réseau et la station (Figure 3.9). Cette dernière modélisation permet ainsi de prendre en compte les différences de mécanismes entre le transfert de pollution au cours d'eau sur le bassin versant (mécanismes de stockage, de transformation et de transfert) et celui dans le cours d'eau (mécanismes d'autoépuration) en différenciant la pondération des fonctions de distance.

Concernant la pondération (α) de la fonction (f) d'atténuation de la pression en fonction de la distance (d), trois modes ont été utilisés quelque soit la méthode de modélisation des distances : la pondération par l'inverse ($f(d)=d^{-\alpha}$, e.g., Comeleo *et al.*, 1996; Kehmeier, 2000; King *et al.*, 2005), la pondération exponentielle ($f(d)=\exp(-\alpha d)$, e.g., Smith *et al.*, 1997; Weller *et al.*, 1998 ; Van Sickle and Johnson, 2008), et la pondération par seuil : si l'occupation du sol se situe en dessous d'un seuil de distance α (défini par l'analyste) alors elle contribue à l'influence du cours d'eau au-delà elle est nulle (King *et al.*, 2005).

La deuxième approche se base sur la hiérarchie des échelles de l'écosystème aquatique (cf. III.1.2) en délimitant une ou plusieurs emprises spatiales candidates aux niveaux des différentes échelles emboitées. Les modèles dérivant de cette approche utilisent donc généralement trois échelles spatiales (Allan, 2004a, Wang, 2006) (Figure 3.10): le bassin versant (BV) amont (échelle régionale) à la station et deux échelles au niveau du corridor rivulaire : le corridor amont (CRamont) à la station (échelle du linéaire) et le corridor local (échelle ponctuelle).

Les emprises aux échelles du corridor rivulaire correspondent généralement à un buffer (zone tampon) autour du réseau hydrographique qui varie en largeur (distance latérale au cours d'eau) pour le CRamont mais aussi en longueur (distance longitudinale en amont de la station le long du cours d'eau) pour le CRlocal. Ces emprises sont le plus souvent définies par expertise, elles sont donc variables en fonction des études (Van Sickle and Johnson, 2008).

CemOA : archive ouverte d'Irstea / Cemagref

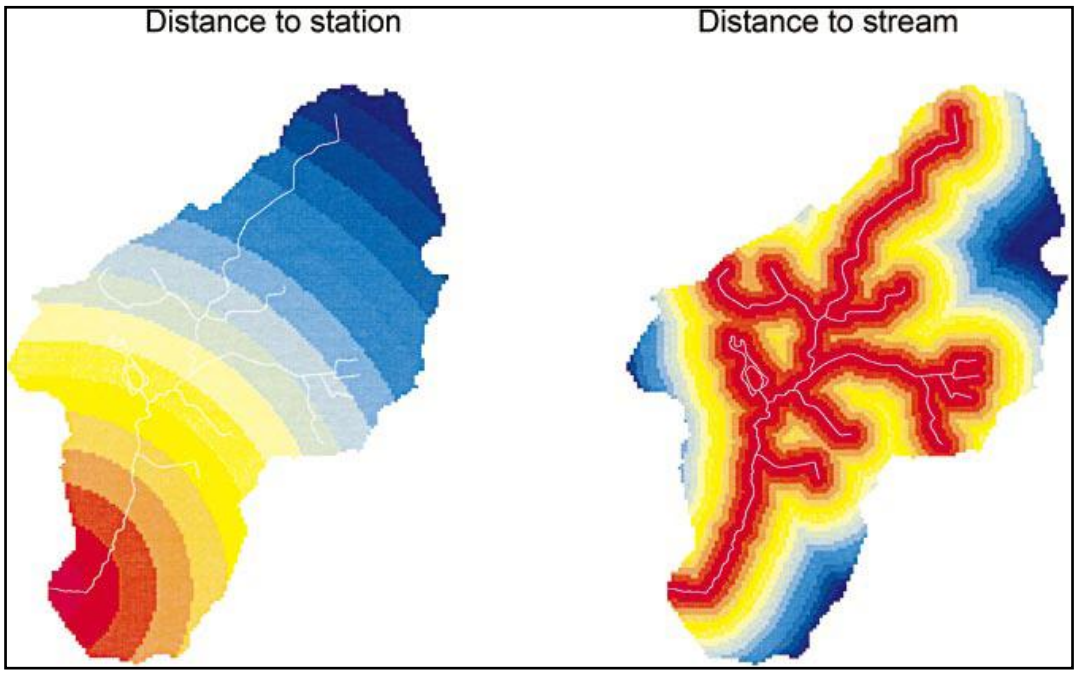


Figure 3.8 : Illustration des distances euclidiennes à la station (à gauche) et au cours d'eau (à droite), (d'après King *et al.*, 2005).

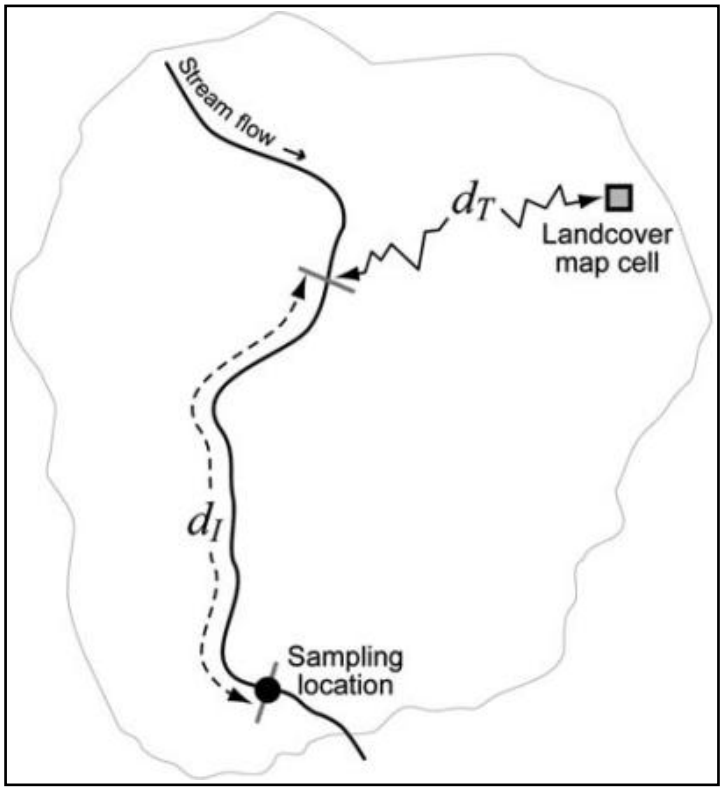


Figure 3.9 : Illustration de la modélisation des distances d'une catégorie d'occupation du sol à un point donné du bassin versant au cours d'eau (direction des flux estimée à partir d'un MNT) et à la station (distance curviligne du réseau) (d'après Van Sickle and Johnson, 2008).

CemOA : archive ouverte d'Irstea / Cemagref

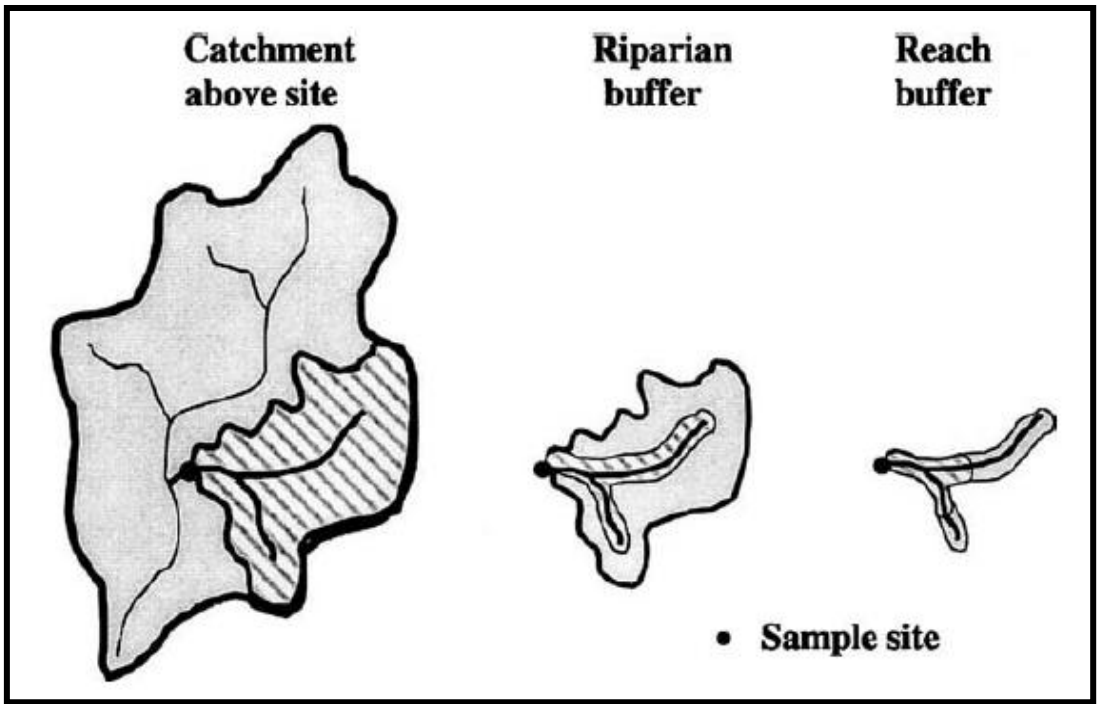


Figure 3.10 : Échelles de travail généralement utilisées dans les modèles pressions/impacts d'après Allan, 2004a.

La première approche, spatialement plus explicite, se focalise principalement sur la pollution diffuse puisqu'elle s'intéresse à pondérer le transfert des polluants en fonction des distances à la station et au cours d'eau. C'est pour cela d'ailleurs que généralement les études utilisant ce type d'approche sont dédiées à l'explication des taux de nutriments (nitrate, phosphate…) au niveau des stations ; on compte toutefois des exceptions comme les études de Van Sickle and Johnson, 2008 qui ont utilisées l'indice multi-métrique pour les communautés de poissons, King *et al.*, 2005 pour les communautés d'invertébrés et Johnson *et al.*, 2007 pour les communautés de poissons, de macroinvertébrés et d'algues. Les modèles dérivant de cette approche ont également la capacité d'estimer les caractéristiques (largeur, étendue, emplacement…) de la végétation rivulaire ayant une influence sur l'interception des flux de polluants (Johnson *et al.*, 2007).

Or, les impacts écologiques ne sont pas seulement déterminés par l'identification et la localisation précise des sources de pollution diffuses (nutriments et sédiments), mais aussi par l'effet combiné des influences anthropiques comme les pollutions ponctuelles, les dysfonctionnements hydrologiques, les modifications hydromorphologiques et les ruptures de connectivité (Borchardt and Richter, 2003) qui peuvent agir sur les processus de fonctionnement de l'écosystème aux différentes échelles. Cette approche ne permet pas prend pas de prendre en compte d'une part l'effet combiné des différentes pressions s'exerçant sur le milieu et d'autre part, l'interaction de ces différentes pressions aux différentes échelles de fonctionnement de l'écosystème rivière.

Par ailleurs, sa mise en place n'est pas forcément évidente puisqu'elle nécessite une bonne connaissance des mécanismes de transfert pour pondérer la fonction de distance et qu'elle est fortement encline aux incertitudes des directions de flux (Van Sickle and Johnson, 2008), notamment en zone de plaine, compte tenu de la résolution spatiale et de l'incertitude des MNT utilisés (généralement > 30 m voire 80 m).

Par contre, la deuxième approche a l'avantage de pouvoir étudier l'effet combiné des différentes pressions en les identifiant et les hiérarchisant aux différentes échelles de fonctionnement de l'écosystème rivière. On s'intéressera donc uniquement, dans la suite de cette partie aux modèles qui suivent cette approche.

Le plus souvent ces modèles sont appliqués au niveau de grands bassins versants. Mais plus ce niveau d'analyse est grand, plus les caractéristiques du milieu, qui déterminent les processus fonctionnels, varient aux différentes échelles d'analyse. Compte tenu de cette variabilité, des auteurs (Richards *et al.*, 1996 ; Dovciak and Perry, 2002) privilégient d'étudier ces relations dans des zones géographiquement homogènes (les écorégions), afin de limiter cette variabilité.

CemOA : archive ouverte d'Irstea / Cemagref

III.2.1.4 Méthodes d'analyses de données

Les approches statistiques que l'on retrouve le plus souvent dans la littérature pour analyser les relations entre les différents indicateurs d'occupation du sol et les indicateurs de l'état écologique sont les techniques de régression simple, de régression multiple, de régressions multi-variées, d'analyse multi-variée des gradients (techniques d'ordination en espace réduit) et de modèles d'équations structurelles (Wang, 2006). On trouve un large spectre de méthodes dans la littérature, le choix de l'outil mathématique se faisant généralement en fonction de l'affinité des laboratoires pour telles ou telles méthodes.

Cependant, l'un des principaux objectifs des méthodes est de contourner la difficulté liée à la colinéarité (corrélation) des variables d'occupation du sol (Johnson *et al.*, 1997; Johnson and Host, 2010), afin de pouvoir correctement évaluer l'effet d'une variable sur la réponse biologique par rapport à une autre. C'est pour cela que l'on recourt souvent à des techniques de réduction et de sélection de variables (e.g, Analyse en Composante Principale, Analyse canonique des correspondances Richards *et al.*, 1996; Stewart *et al.*, 2001; Wang and Kanehl, 2003) afin de limiter cette colinéarité ou bien à des techniques capable de tirer des relations interprétables malgré cette difficulté telle que la « Covariance Structure Analysis » (e.g., Infante *et al.*, 2006; Zorn and Wiley, 2006), la « Stepwise Regression » (e.g., Sponseller *et al.*, 2001 ; Roy *et al.*, 2003b) ou encore la « Partial Least Square regression » (Wasson *et al.*, 2005 ; Wasson *et al.*, 2010).

Les régressions linéaires simples ont été souvent employées pour identifier l'impact relatif d'une variable d'occupation donnée sur les différentes emprises et/ou les différentes échelles spatiales étudiées (e.g., Roth *et al.*, 1996 ; Johnson and Covich, 1997 ; Sponseller *et al.*, 2001). Des méthodes sophistiquées telles que les techniques de partitionnement de la variance (e.g., Richards *et al.*, 1997 ; Johnson *et al.*, 2004) sont utilisées également pour identifier le poids relatif des échelles dans l'étude de ces relations.

III.2.1.5 Synthèse des résultats

De nombreuses études ont été consacrées aux relations statistiques entre l'occupation du sol et les mesures de la qualité des cours d'eau aux différentes échelles de fonctionnement de l'écosystème rivière (Allan, 2004a ; Johnson and Host, 2010). Elles affirment unanimement l'importante influence tant des activités humaines que des conditions naturelles environnantes sur l'état écologique de ces milieux. Parmi les pionniers dans le domaine, Klein, 1979 a mis en évidence l'impact négatif de l'urbanisation à l'échelle du bassin versant sur l'intégrité écologique des cours d'eau. Steedman, 1988, grâce à ce type de modèle, identifie des relations positives entre les communautés de poissons et d'une part le pourcentage de forêt dans le bassin et d'autre part la densité de chenaux dans les corridors rivulaires ; il identifie également un impact fortement négatif dû à l'urbanisation dans le bassin versant. Ces études s'accordent pour désigner l'agriculture et l'urbanisation comme étant les deux principales causes de dégradation (Allan, 2004a) de la qualité écologique des cours d'eau. Mais, elles diffèrent sur l'échelle d'influence de cette occupation du sol (bassin versant versus. corridor rivulaire) (Gergel *et al.*, 2002 ; Allan, 2004a).

III.2.1.5.1 Impact de l'agriculture

L'agriculture est identifiée par l'Agence Européenne de l'Environnement (EEA) comme une force motrice significative dégradant la qualité écologique des cours d'eau, du fait de la variabilité des pressions qu'elle engendre. Elle altère à la fois les processus biogéochimiques via la pollution diffuse (augmentation des apports de sédiments, matières organiques et nutriments, présence de substances dangereuses telle que les métaux et les insecticides), les processus hydromorphologiques (drainage, irrigation, chenalisation des rivières, retenues d'eau, suppression de la ripisylve…), et les processus écologiques (fragmentation des habitats en bordure du cours d'eau). Selon l'agence européenne, «Le bon état écologique des eaux de surfaces (et des eaux souterraines) sera atteint, si l'impact de l'agriculture sur les hydrosystèmes européens est réduit» (EEA, 2003).

Beaucoup d'auteurs désignent l'agriculture intensive comme la cause de la dégradation des habitats des cours d'eau (e.g., Allan *et al.*, 1997, Roth *et al.*, 1996). Par exemple, dans les systèmes agricoles, le travail de la terre augmente le phénomène d'érosion (apports de sédiments) qui est à l'origine du colmatage des substrats avec un impact direct sur la faune (Gayraud *et al.*, 2002). Ce travail modifie également les capacités

CemOA : archive ouverte d'Irstea / Cemagref

d'infiltration de l'eau dans le sol. Il participe donc avec les systèmes de drainage (drains, fossés, raies de curage) et d'irrigation à la diminution des flux liquides, altérant la vie aquatique par la réduction des profondeurs, de la quantité d'habitat disponible et donc de la capacité d'accueil, en aggravant la thermie et en augmentant la concentration en polluants divers.

Ces effets sur la biologie ont été démontrés par plusieurs auteurs. Bis *et al.*, 2000 ont ainsi montré l'impact de l'agriculture sur la diminution tant du nombre d'espèces (richesse taxonomique) que de l'abondance des communautés piscicoles et d'invertébrés, et sur l'évolution des compositions au sein des communautés algales. Cuffney *et al.*, 2000 ont mis aussi en évidence cet impact sur les communautés piscicoles, les invertébrés et les algues en fonction de l'intensité agricole. Des relations négatives entre le pourcentage d'occupation agricole dans le bassin versant et les indices des communautés piscicoles ont été trouvées par Roth *et al.*, 1996 et Walser and Bart, 1999. L'impact de l'agriculture sur les diatomées est plus difficile à mettre en évidence : Pan *et al.*, 2004 n'ont décelé qu'un faible lien entre l'occupation agricole et les communautés lotiques[21] de diatomées durant la saison des basses eaux. Harding *et al.*, 1998 ont relié la diversité actuelle des communautés piscicoles et d'invertébrés avec l'évolution de l'occupation du sol et ont mis en évidence l'impact de l'agriculture depuis 1950 sur l'état écologique des cours d'eau.

III.2.1.5.2 Impact de l'urbanisation

Dans un premier temps les impacts de l'urbanisation ont été principalement associés à des sources de pollution directe ponctuelle, à cause des déversements d'eaux usées traitées ou non (Garcia *et al.*, 2005) modifiant les processus biogéochimiques (apport de matière organique, de matière en suspension, de substances dangereuses…). Cependant, ces processus sont altérés également par la pollution diffuse issue de ces surfaces (ruissellement pluviale non traité, déversoirs d'orages).

Les territoires artificialisés modifient également les processus hydromorphologiques, l'imperméabilisation des sols (habitations, routes et autres infrastructures constituent des surfaces imperméables) altèrent les caractéristiques des évènements d'inondation dans le temps, la fréquence, la durée et l'amplitude (Cormier *et al.*, 2000, Arnold and Gibbons, 1996) avec des conséquences considérables sur la biologie des écosystèmes aquatiques. Par exemple, l'augmentation du débit liquide associé à un maintien, voire une réduction, du débit solide induit des phénomènes importants d'érosion, et souvent une incision du cours d'eau avec des conséquences sur les habitats aquatiques par l'augmentation des contraintes physiques (vitesses, hauteurs d'eau) et par la réduction des alternances naturelles « radiers/mouilles » : ceci est à l'origine d'une banalisation des habitats, de déconnexions des annexes quand elles existent (Chandesris *et al.*, 2008). De plus, beaucoup de travaux de chenalisation sont réalisés pour protéger les zones urbaines des inondations, modifiant ainsi la structure des habitats et le fonctionnement de l'écosystème rivière (Wasson *et al.*, 1998).

Par ailleurs, les obstacles à l'écoulement (barrages, écluses…) altèrent également les processus écologiques en modifiant la capacité de dispersion et de recolonisation des espèces et notamment les espèces migratrices (White and Knights, 1997).

Dans la littérature, beaucoup d'auteurs ont montré l'impact significatif des territoires artificialisés sur l'état écologique. Par exemple, des corrélations négatives entre le taux d'urbanisation et les indices d'invertébrés ont été démontrées par Stepenuck *et al.*, 2002, qui a étudié la réponse des invertébrés en fonction du coefficient d'imperméabilité des sols (indicateur d'urbanisation). Un impact similaire est mis en évidence par Wang and Kanehl, 2003, Snyder *et al.*, 2005. Roy *et al.*, 2003b, ils ont montré des relations négatives entre l'occupation du sol urbain et les indices d'invertébrés. Un seuil situé entre 6 et 15% d'urbanisation revient régulièrement dans les différentes études analysant cet impact (Valette, 2005). Au delà de ce seuil d'urbanisation, il y a une diminution, voire une disparition, des macroinvertébrés polluosensibles au détriment d'espèces plus résistantes Morse *et al.*, 2003 ; Freeman and Schorr, 2004) et une modification du peuplement piscicole (Booth, 1991 ; Luchetti and Fuersteburg, 1993).

[21] Organismes qui vivent dans des hydrosystèmes traversés ou soumis à des courants mesurables

CemOA : archive ouverte d'Irstea / Cemagref

III.2.1.5.3 Bassin versant vs. corridor rivulaire

Les différentes communautés dulçaquicoles répondent donc aux variables d'occupation du sol. Mais à quelle échelle spatiale d'analyse ? La réponse reste confuse dans la littérature car plusieurs travaux se contredisent.

D'une part, on trouve des études ne décelant pas de corrélations à l'échelle du corridor rivulaire. C'est le cas de Roth *et al.*, 1996, pour qui les zones agricoles et forestières prises à l'unité du bassin versant sont les déterminants primordiaux de l'état écologique des cours d'eau alors que la végétation rivulaire se révèle être un faible prédicteur. De même pour Wang *et al.*, 1997, l'impact positif lié à la présence de forêt et l'impact négatif concernant la présence d'agriculture sont démontrés uniquement à l'échelle du bassin versant, pas à l'échelle des corridors rivulaires (l'analyse aux échelles inférieures ne donnant que des faibles corrélations). Les travaux plus récents de Frimpong *et al.*, 2005a montrent que les caractéristiques des communautés piscicoles sont plus fortement corrélées aux caractéristiques naturelles et anthropiques à l'échelle du basin versant qu'elles ne le sont à l'échelle du tronçon. Ils observent également que pour une zone d'étude donnée, le modèle à l'échelle du bassin explique mieux (+15 %) la variabilité de l'indice biologique des cours d'eau que le modèle à l'échelle du corridor rivulaire.

D'autre part, on trouve des études ne détectant pas au contraire de corrélations à l'échelle du bassin. Lammert and Allan, 1999 n'ont décelé aucune relation entre l'occupation du sol du bassin versant et l'état écologique ; seules les caractéristiques naturelles des habitats et de l'occupation du sol à proximité de la station (zone tampon de 100 m) sont corrélées avec la qualité écologique. De même pour Sponseller *et al.*, 2001, qui ont mis en évidence, dans neuf bassins versants en tête de cours d'eau diversifiés du point de vue de l'occupation du sol, que seule la chimie de l'eau était généralement liée aux caractéristiques des bassins versants, alors que la température, le type de substrat et les macroinvertébrés sont eux fortement influencés par l'occupation du sol dans une zone tampon plus proche : 200 m de part et d'autre du cours d'eau. Dans une étude regroupant 172 sites, Meador and Goldstein, 2003 ont conclu que la physico-chimie de l'eau et les conditions ripariennes peuvent être de meilleurs indicateurs de la distribution des communautés de poisson que l'occupation du sol prise à l'échelle du bassin versant. Dans les zones agricoles, l'indice poisson est étroitement relié à l'indice mesurant l'état de la ripisylve, mais cette corrélation est plus faible en zones urbaines. Par ces résultats, ils ont mis en avant l'importance des zones ripariennes pour maintenir et conserver la biodiversité des communautés piscicoles dans les cours d'eau.

Finalement, dans une étude regroupant 38 cours d'eau, Stewart *et al.*, 2001 ont trouvé de fortes relations entre l'occupation du sol et les bioindicateurs (poissons et macroinvertébrés), ceci aussi bien à l'échelle du bassin versant qu'à celle du corridor rivulaire. De même, pour Townsend *et al.*, 2004, les relations pressions d'occupation du sol et impacts sur les macroinvertébrés ne varient pas selon l'échelle d'analyse. Les travaux de Vondracek *et al.*, 2005 confirment ces résultats. Ils ont pour leur part synthétisé neuf études qui analysent l'influence de l'occupation du sol, mesurée aux différentes échelles d'observation des écosystèmes aquatique, sur l'état écologique des cours d'eau. Ces études totalisent 425 cours d'eau sur trois écorégions différentes. Cette synthèse indique que l'ensemble des échelles influence significativement la variabilité de l'état écologique des cours d'eau.

III.2.2. Modèles pressions/impacts en France

Avant le démarrage de la présente thèse, les principaux travaux en France dédiés à l'étude statistique des relations pressions/impacts à large échelle ont été menés principalement par le laboratoire d'hydro-écologie quantitative de Lyon et ont fait l'objet d'un rapport « Modèles pressions/impacts : Approche méthodologique, modèles d'extrapolation spatiale et modèles diagnostics » par Wasson et al. 2005 dans le cadre de l'appui scientifique à la DCE pour le Ministère de l'environnement. On résume succinctement dans cette partie les caractéristiques et les principaux résultats des modèles de diagnostics permettant d'identifier et de hiérarchiser les principales causes d'altérations de l'état écologique.

CemOA : archive ouverte d'Irstea / Cemagref

III.2.2.1 Présentation du modèle pressions/impacts régionalisé

Plusieurs modèles ont été mis en place en France à l'échelle nationale et au niveau des différentes hydro-écorégions (HER) française (cf. Chapitre 1 :II.3), les hydro-écorégions permettant d'identifier des cours d'eaux (qui peuvent être spatialement séparés) souffrant des mêmes « pathologies régionales » sur un territoire et relevant donc d'une même politique de restauration.

La réponse biologique étudiée est un écart à la référence défini pour chaque type de cours d'eau à partir de l'IBGN (cf. Chapitre 1 :III.2). L'EQR détermine ainsi un état écologique indépendant des variations naturelles (typologiques) de l'indice ou, autrement dit, un état « principalement » sous contrôle anthropique.

En France comme pour d'autres pays, on ne dispose pas à l'heure actuelle de données détaillées sur les diverses pressions renseignant tous les compartiments incriminés, aux différentes échelles appropriées. Lorsque l'information existe, elle n'est pas forcément homogène à l'échelle de travail (grand bassin versant, région, nation) et non spatialisable (i.e. qu'on peut relier topologiquement avec les stations écologiques). Par exemple, les sources de données existantes en France ne permettent d'affecter les rejets polluants qu'aux zones hydrographiques pour les rejets urbains (données des Agences de l'eau), et encore avec une incertitude importante, et à l'échelle des cantons pour les pollutions dues aux élevages en raison de la confidentialité des données du recensement agricole (RGA). Concernant les pressions physiques, il n'existe pas de données bancarisées homogènes à l'échelle du territoire métropolitain : on ne trouvait même pas en France, en 2004, un fichier géoréférencé des barrages homogène à l'échelle du territoire national.

Ces modèles utilisent donc l'occupation du sol comme force motrice générant la plupart des pressions s'exerçant sur le milieu aquatique. Cette occupation est calculée à partir de la base de données européenne CORINE Land Cover (Bossard *et al.*, 2000), réalisée par interprétation d'images satellites à moyenne ou haute résolution (e.g, Landsat Thematic Mapper : 30 m ; Spot 4 : 20 m).

L'occupation du sol est analysée à l'échelle des bassins versants amont de chaque station et au niveau du corridor rivulaire local à l'aide d'un buffer délimité sur le drain principal amont/aval de la station en traçant un cercle de 3 km centré sur la station (Pella et al., 2004). Les modèles analysent ensuite l'influence des catégories de chaque type d'occupation du sol aux différentes échelles sur l'ensemble du gradient d'état écologique révélé par l'EQR-IBGN à partir de régressions PLS (Partial Least Square).

III.2.2.2 Principaux résultats

Ces modèles permettent d'identifier et de hiérarchiser les principaux facteurs dégradants ou protecteurs de l'état écologique.

Quelque soit la zone d'étude, ils montrent que les pressions d'origine urbaine sont probablement responsables des situations les plus dégradées, observées sur une fraction non négligeable du réseau.

Les impacts liés aux terres cultivées semblent provoquer des dégradations moins accentuées, mais spatialement très étendues.

Les prairies constituent en général un environnement plutôt favorable à l'échelle nationale, mais cette conclusion doit être fortement nuancée en fonction du contexte régional, lié à l'intensification de l'élevage. Par exemple, au niveau du Massif Armoricain (HER 12), région très typée par sa géographie et ses activités agricoles, les prairies, qui ont généralement un effet positif au niveau du bassin semblent avoir un impact négatif sur l'IBGN lorsqu'elles sont situées au niveau du corridor rivulaire local.

Les espaces « naturels » (ou faiblement anthropisés) ont par contre un effet toujours positif sur l'IBGN sur le bassin et, selon les HER, au niveau du corridor rivulaire local.

CemOA : archive ouverte d'Irstea / Cemagref

IV. Vers des modèles adaptés à la DCE : améliorer les modèles empiriques par une meilleure prise en compte des corridors rivulaires.

Ainsi définis, les modèles empiriques donnent donc un premier niveau de réponse aux attentes de la Directive Cadre sur l'Eau.

Au niveau des grands bassins hydrologiques ou écorégions, ils permettent d'analyser quantitativement et à plusieurs échelles (bassin versant, corridor rivulaire amont et corridor rivulaire local) les relations entre les pressions anthropiques s'exerçant sur le milieu (indicateurs d'occupation du sol) et l'état écologique des cours d'eau (indicateurs de réponse biologique). Ils fournissent ainsi aux acteurs de la gestion de l'eau les moyens de procéder à des évaluations ex-ante (à priori) des effets potentiels de la mise en œuvre de diverses politiques publiques possibles au niveau du bassin (activités humaines) et des corridors rivulaires (préservation/restauration). Compte tenu du délai imparti par la DCE, ils permettent aussi dans un premier temps de cibler les zones prioritaires d'action qui pourront par la suite être investiguées de façon plus exhaustive en utilisant directement des variables de pressions.

Cependant, alors que certains résultats semblent faire l'objet d'un consensus, comme l'impact négatif des zones agricoles et urbaines sur le bassin, d'autres divergent, particulièrement au niveau du poids relatif des différentes échelles d'impact, élément important et attendu par les gestionnaires pour prendre des décisions notamment sur la pertinence de la restauration des espaces rivulaires. Ces résultats contradictoires sont-ils dus aux différents contextes naturels et humains étudiés, ou aux différentes méthodes utilisées ? C'est une question à laquelle il faut répondre avant d'engager des stratégies de préservation et/ou de restauration qui peuvent s'avérer coûteuses.

Trois raisons majeures compliquent le problème:

(1) la difficulté de séparer les facteurs anthropiques et naturels compte tenu de leur covariance (Stauffer *et al.*, 2000 ; Fitzpatrick, 2001 ; Allan, 2004a ; Johnson and Host, 2010). Par exemple, une étude de 25 cours d'eau dans des bassins agricoles du Wisconsin (Fitzpatrick, 2001) démontre la forte influence de l'agriculture sur le milieu mais l'interprétation a été très complexe du fait de la covariance entre les variables anthropiques et naturelles mesurées à plusieurs échelles. Les chercheurs n'ont pas pu séparer de façon significative les effets (étroitement liés) des critères géologiques, hydrologiques et de l'occupation du sol sur le bassin et le long du cours d'eau ;

(2) la disponibilité de données précises sur l'occupation du sol au niveau du corridor rivulaire (Lattin *et al.*, 2004 ; Frimpong *et al.*, 2005b ; Roy *et al.*, 2007). L'information extraite des images à moyenne résolution est en effet en limite d'utilisation, sinon invalide, pour quantifier la composition et la configuration de l'occupation des sols et caractériser l'état des ripisylves (Müller, 1997 ; Congalton *et al.*, 2002), ce qui limite d'emblée la pertinence de l'analyse et l'interprétation de ces relations (Wasson *et al.*, 2005 ; Hollenhorst *et al.*, 2006 ; Wang, 2006 ; Roy *et al.*, 2007) ;

(3) la définition des emprises spatiales candidates aux échelles du corridor rivulaire (Van Sickle and Johnson, 2008) pour étudier l'influence de l'occupation du sol dans ces espaces sur l'état écologique. La plupart des études délimitent seulement une à trois emprises spatiales candidates ce qui limite d'emblée la recherche de l'emprise sur laquelle une source de pressions (i.e. catégorie d'occupation du sol) influence réellement l'état écologique. De plus, on constate dans la littérature une variabilité importante entre les hypothèses de construction de ces différentes emprises aussi bien, sur la dimension latérale (distance au cours d'eau) que longitudinale (distance à la station), ce qui limite la comparaison des résultats entre ces différentes études (Van Sickle and Johnson, 2008).

Seuls Frimpong et al., 2005, ont engagé un travail systématique en définissant plus 100 emprises candidates au niveau du corridor rivulaire et en cherchant à mettre en évidence l'emprise spatiale sur laquelle la végétation arborée au niveau du corridor rivulaire local a le plus d'influence sur l'état écologique. Cependant, ces résultats montrent que la deuxième difficulté (2) est intrinsèquement liée à la troisième (3). La donnée spatiale actuelle, trop grossière pour caractériser l'occupation du sol au niveau du corridor, limite la capacité de localisation précise des différentes sources de pression.

CemOA : archive ouverte d'Irstea / Cemagref

En France (et en Europe), les travaux menés sur l'évaluation de l'état écologique permettent aujourd'hui de construire des modèles robustes sur les relations pressions impacts en diminuant l'effet de covariance entre les facteurs anthropiques et naturels (Wasson *et al.*, 2005 ; Wasson *et al.*, 2007). L'écart à la référence enregistré ne dépend pas des variabilités naturelles (typologiques) de l'indice mais il est principalement sous contrôle anthropique. De plus, les hydro-écorégions (utilisées dans la définition de la typologie physique des cours d'eau) facilitent le transfert des modèles à des contextes socio-écologiques différents (Garcia *et al.*, 2005) et constituent un cadre cohérent pour la définition des stratégies de restauration. Elles permettent d'identifier des cours d'eaux (qui peuvent être spatialement séparés) souffrant des mêmes « pathologies régionales » sur un territoire, et relevant donc d'une même politique de restauration. Néanmoins, ces modèles restent toujours en limite d'utilisation au niveau de l'échelle du linéaire (corridor rivulaire).

D'une part, dans ces modèles, l'occupation du sol dans les espaces rivulaires est issue de la base données CORINE Land Cover construite par photo-interprétation des images satellites Landsat (30 m) avec une unité minimum de cartographie de 25 ha (Bossard *et al.*, 2000). A ce grain d'analyse, la composition et la configuration de l'occupation du sol (générant les pressions sur les milieux) ne sont pas détectées et ne restituent pas les objets de dimension réduite que sont les ripisylves (Figure 3.11).

D'autre part, quelque soit la catégorie d'occupation du sol (ou source de pressions), elles sont étudiées sur une même emprise spatiale au niveau du corridor rivulaire. Cette difficulté limite (i) la compréhension de ces relations car l'effet d'une catégorie d'occupation donnée est plus ou moins perceptible sur l'état écologique selon l'emprise spatiale (distance au cours d'eau et distance à la station) ; (ii) l'information pour les gestionnaires concernant la localisation de ces sources au niveau du corridor.

Dès lors, il est nécessaire de développer des méthodologies efficaces et opérationnelles permettant (i) de quantifier précisément et de façon fiable sur de grands territoires, l'occupation du sol dans les corridors rivulaires, et (ii) de construire des indicateurs spatialisés à partir de cette information afin de mieux localiser les effets dans ces espaces et de mettre en évidence le rôle de la végétation rivulaire dans les modèles pressions/impacts. Cette thèse s'attache à résoudre ces deux difficultés. Ses objectifs sont détaillés dans le Chapitre 4.

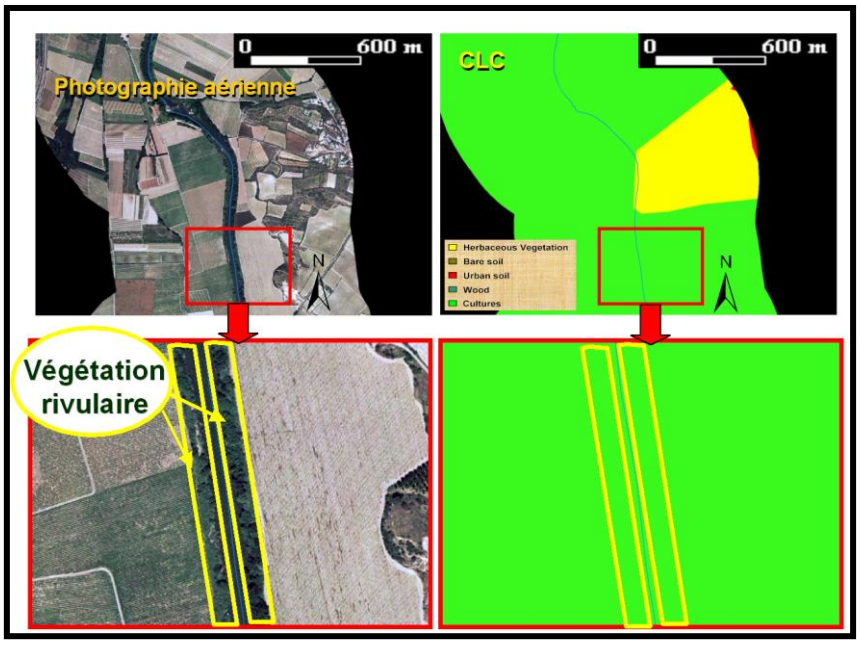


Figure 3.11 : Comparaison entre une photographie aérienne (0.5 m de résolution spatiale) et la donnée CORINE Land Cover (CLC) sur un buffer de 600 m de part et d'autres d'un tronçon de cours d'eau (fleuve Hérault au niveau de sa plaine). L'unité cartographique de CLC (25 ha min) ne permet pas de quantifier précisément la structure de l'occupation du sol dans ces espaces. Les ripisylves (végétation rivulaire) ne sont pas détectées.

CemOA : archive ouverte d'Irstea / Cemagref

Pour finir nous soulignerons que d'autres difficultés, qui ne seront pas étudiées dans cette thèse, viennent complexifier la réponse de l'état écologique des cours d'eau aux perturbations anthropiques :

(1) « l'effet d'héritage » (Allan, 2004a ; Johnson and Host, 2010). Les réponses biologiques n'apparaissent souvent qu'après un temps de latence plus ou moins long, plusieurs années s'avérant nécessaires pour observer, par exemple, les effets d'un dépôt acide, d'une espèce introduite ou d'une pollution des eaux souterraines. La prise en compte de cet effet nécessite un historique de données de pressions conséquent, actuellement inexistant, sachant que ces données sont déjà très difficile à acquérir sur une zone d'étude à un instant donné ;

(2) la « non-linéarité » des réponses biologiques à ces perturbations (Allan, 2004a ; Décamps, 2006 ; Johnson and Host, 2010) : les effets s'ajoutent-ils les uns aux autres de façon linéaire, s'amplifient-ils, introduisent-ils des seuils plus sévères ? La réponse à ces questions n'est pas évidente, elle demande l'utilisation d'outils mathématiques complexes (difficiles à maitriser pour l'analyste) et surtout l'identification précise des pressions s'exerçant sur le milieu, l'occupation du sol informant uniquement sur un panel non exhaustif de pressions potentielles pouvant s'exercer sur le milieu ;

(3) « l'autocorrélation spatiale ». Les observations peuvent avoir des ressemblances liées à leur répartition spatiale dans la zone d'étude. C'est un phénomène bien ancré en écologie du paysage et de nouvelles techniques existent pour répondre à cette difficulté (Legendre, 1993 ; Thompson et al., 1996). Pour l'écosystème rivière, sa prise en compte n'est pas évidente puisque cette dépendance spatiale est liée à un réseau linéaire orienté (amont-aval). C'est pour cela qu'on retrouve peu d'études (King *et al.*, 2005; Maillard and Pinheiro Santos, 2006) traitant cette difficulté dans les modèles pressions/impacts.

CemOA : archive ouverte d'Irstea / Cemagref

Chapitre 4
Objectifs, démarche et zones d'études

SOMMAIRE

Résumé *107*

I. LES OBJECTIFS SPECIFIQUES DE LA THESE 109

I.1. OBJECTIFS DE CLASSIFICATION *109*

I.2. OBJECTIFS D'ANALYSE SPATIALE *110*

I.3. OBJECTIFS DE MODELISATION *110*

II. DEMARCHE ET ZONES D'ETUDES 111

II.1. DEMARCHE DE LA THESE *111*

II.2. PRESENTATION DES ZONES D'ETUDES *111*

II.2.1. La zone pilote : le bassin versant de l'Hérault 111

II.2.2. Zone de modélisation : Nord Ouest de l'HER Tables-Calcaires (HER9-NO) 113

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Résumé

Notre capacité à comprendre empiriquement les relations entre l'homme et l'état écologique des cours d'eau est très fortement dépendante de la disponibilité et du traitement de l'information sur de grands territoires. Alors que la France dispose aujourd'hui d'une base de données spatialisée robuste pour suivre l'état écologique des cours d'eau, elle ne dispose pas encore d'une information précise sur les pressions d'Occupation du Sol dans les Corridors Rivulaires (OSCR).

Les outils de la géomatique peuvent répondre à ces difficultés : d'une part la télédétection à Très Haute Résolution Spatiale (THRS) fournit une information riche et détaillée appropriée à la caractérisation de ces pressions ; et d'autre part les Systèmes d'Information Géographique (SIG) offrent les outils nécessaires pour la construction d'indicateurs spatialisés pertinents permettant d'améliorer la localisation des sources de pressions (i.e. catégorie d'occupation du sol) et leur signification vis-à-vis des mécanismes d'impact de ces sources, et notamment l'influence de la végétation rivulaire, sur l'état écologique des cours d'eau.

Dès lors, l'enjeu de cette thèse est de tester si les potentialités qu'offrent les outils de la géomatique (télédétection à THRS et indicateurs spatialisés) permettent d'améliorer in fine la compréhension des relations pressions impacts notamment au niveau du corridor rivulaire.

Cela implique (1) des objectifs de classification liés à l'extraction de la donnée d'OSCR à partir des images à THRS, (2) des objectifs d'analyse spatiale pour traduire cette donnée en information pertinente (indicateurs spatialisés) et (3) des objectifs de modélisation afin de tester l'apport de ces avancées méthodologiques dans la compréhension des liens pressions impacts puis d'établir un diagnostic de l'état écologique sur la zone étudiée fournissant des éléments de réponses aux gestionnaires sur les stratégies de restauration de cet état.

Pour répondre à ces objectifs des méthodes automatisées (ou semi-automatisés) et opérationnelles devront être construites compte tenu de la quantité de données à manipuler. Deux zones d'études ont été définies pour établir et tester ces méthodologies. Une zone pilote, le bassin versant de l'Hérault (1150 km de réseau hydrographique), qui permet de tester assez rapidement les différentes méthodologies envisagées et de répondre aux différents objectifs de classification et d'analyse spatiale. Une zone de modélisation, le Nord-Ouest de l'hydro-écorégion Tables Calcaires (6000 km de réseau hydrographique) qui possède un nombre de stations (158) suffisants pour mener une démarche statistique robuste et fiable.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

I. Les objectifs spécifiques de la thèse

Selon Wang, 2006, notre capacité à comprendre empiriquement les relations entre l'homme et l'état écologique des cours d'eau est très fortement dépendante de la disponibilité et du traitement de l'information (biologique et de pressions) sur de grands territoires. L'information (et notamment l'information spatialisée) est un élément central de la capacité de l'homme non seulement à connaître et comprendre les structures, dynamiques et évolutions de l'environnement, des sociétés et des territoires mais aussi à échanger entre acteurs de la société (citoyen, État, acteurs socio-économiques) pour définir des stratégies de gestion concertées et les mettre en œuvre (Cemagref, 2009). En France, un investissement important a été consacré à la construction d'une base de données spatialisée nécessaire à l'évaluation de l'état écologique des cours d'eau (délimitation des hydro-écorégions, typologie, station biologiques de référence et de suivi…). On dispose ainsi d'une information robuste sur la biologie des cours d'eau mais on ne dispose pas encore d'une information précise (quantification et localisation) sur les pressions d'Occupation du Sol dans les Corridors Rivulaires (OSCR). Aujourd'hui, grâce à l'augmentation du nombre de satellites en opération et aux avancées technologiques, l'imagerie satellitaire (et aéroportée) ne cesse de devenir plus performante en résolution spatiale comme en résolution spectrale (Royer *et al.*, 2007). La télédétection à Très Haute Résolution Spatiale (THRS à partir de 5 m) fournit une information appropriée à la caractérisation des systèmes rivulaires (Müller, 1997) et de ses pressions correspondantes. Par ailleurs, les progrès technologiques sur les Systèmes d'Information Géographique (SIG), permettent aujourd'hui d'étudier différentes emprises spatiales au niveau du corridor rivulaire afin de mieux localiser les effets de l'occupation du sol dans cet espace et de développer des indicateurs spatialisés potentiellement pertinents vis-à-vis (i) des mécanismes d'impact des différentes sources de pression et (ii) du rôle de la végétation rivulaire. Ainsi, cette thèse s'intéresse aux potentialités de la géomatique (télédétection à THRS et SIG) et à l'analyse et à la modélisation spatiale pour améliorer l'étude des relations pressions/impacts au niveau du corridor rivulaire. Elle a donc des objectifs spécifiques de classification (cf. I.1), d'analyse spatiale (cf. I.2) et de modélisation (cf. I.3). Ces objectifs spécifiques contribuent à l'objectif général de compréhension de l'influence de la ripisylve et de l'occupation du sol dans les corridors rivulaires sur l'état écologique des cours d'eau.

I.1. Objectifs de classification

Le premier enjeu dans la thèse est d'élaborer une méthode opérationnelle de caractérisation fine, sur de grands territoires, de l'Occupation du Sol dans les Corridors Rivulaires (OSCR) à partir des informations contenues dans les images à Très Haute Résolution Spatiale (THRS) et/ou des données exogènes complémentaires. La méthode de caractérisation doit donc répondre à plusieurs objectifs :

- elle doit être reproductible, à coût réduit et applicable sur des types de cours d'eau différents (en montagne, en piémont et en plaine) puisqu'elle a pour but de fournir une information d'OSCR sur l'ensemble du réseau hydrographique français, qui représente un total de près de 500.000 km ;

- elle doit produire une information précise et détaillée pour la construction d'indicateurs spatialisés pertinents. Dans tous les cas, elle doit donner au minimum une information sur les six principales structures (classes) d'occupation du sol :
  - Les **surfaces en eau** (cours d'eau, plans d'eau…) et les **sols nus naturels** (sable, galets, roches nues…) afin de délimiter précisément le lit de la rivière ;
  - Les **sols artificialisés** (tissus urbain continus, discontinus, industriels …) et **sols agricoles** (cultures annuelles, permanentes…) considérés comme les forces motrices générant de forte pressions négatives sur l'état écologique (cf. Chapitre 3 :III.2.1.5) ;
- Les **zones de végétation arborée** (forêt de feuillus, conifères…) et **zones de végétation herbacée** (prairie, lande…) constituant les principaux types de végétation assurant le rôle positif des corridors rivulaires sur l'état écologique des cours d'eau (cf. chapitre 2) ;

CemOA : archive ouverte d'Irstea / Cemagref

- elle doit fournir une information fiable puisque elle sera utilisée in fine dans les modèles pressions/impacts et dans l'orientation des stratégies de restauration ;

- elle doit être facilement portable pour être utilisée dans les autres États européens.

Ces objectifs impliquent de développer un protocole de traitement (système de classification) automatisé (ou semi-automatisé) capable d'intégrer et de combiner l'information issue de données de sources différentes, de s'adapter aux différentes occupations du sol présentes sur le territoire français, et d'évaluer la qualité de la donnée d'OSCR produite sur de grands territoires. Des questions clés se poseront alors sur les méthodes de classification et de validation ainsi que sur le choix des données dans le contexte français. Il est à noter qu'on ne cherche pas à produire une donnée de référence avec une grande précision géométrique (métrique voire submétrique) puisque cette donnée d'occupation du sol est dédiée à la construction d'indicateurs spatialisés agrégeant cette information dans une emprise spatiale donnée. On essayera tout de même d'être le plus précis possible, en fonction des données utilisées, afin de limiter les erreurs des indicateurs s'intéressant aux mécanismes liés aux objets d'occupation du sol au contact du cours d'eau (e.g., les ripisylves).

I.2.Objectifs d'analyse spatiale

Le deuxième enjeu est de construire des indicateurs spatialisés permettant de synthétiser de façon quantifiée l'information de l'OSCR à partir des résultats de classification. Ces indicateurs doivent:

- fournir une information potentiellement explicative de l'état écologique des cours d'eau en traduisant les mécanismes d'impact provenant des activités humaines et de la végétation rivulaire sur le milieu ;

- permettre de mieux quantifier et localiser les effets de cette occupation et de visualiser l'état des corridors rivulaires pour guider les gestionnaires sur les mesures et les zones d'actions prioritaires.

La construction de ces indicateurs nécessite alors le développement de techniques d'analyse spatiale utilisant les Systèmes d'information Géographique (SIG), permettant d'extraire une information sur la composition et la configuration de l'occupation du sol à différents niveaux d'agrégation informatifs et spatiaux et de visualiser cette information sur des cartes didactiques pour les différents acteurs. Des questions se poseront alors sur le choix des indicateurs, sur l'automatisation de leur construction et leur représentation, et sur leur sensibilité à l'incertitude de l'information d'occupation des sols.

I.3.Objectifs de modélisation

Le troisième enjeu est double. Il s'agit tout d'abord d'évaluer l'apport des avancées méthodologiques mise en œuvre dans cette thèse (donnée d'occupation du sol à THRS et indicateurs spatialisés) dans la compréhension des relations pressions/impacts. Il s'agit ensuite d'apporter des éléments de réponse aux différents acteurs sur les poids des différentes causes d'altération (quelles sont les principales structures d'altération ?, où faut-il agir en priorité?, sur quelles emprises ?…) et sur l'influence de la végétation rivulaire (quelles échelles ? dans quelles mesures?, …). Cet enjeu implique donc :

- une analyse comparative afin de quantifier l'apport des avancées méthodologiques. On se basera pour cela sur les modèles pressions/impacts régionalisés développés en France (Wasson *et al.*, 2005, cf. Chapitre 3 :III.2.2).

- la construction d'un modèle pressions/impact pertinent pour établir un diagnostic de l'état écologique sur la zone d'étude afin de fournir des éléments de réponse pour les gestionnaires sur les stratégies de restauration de cet état.

Des questions se poseront sur le choix des indicateurs biologiques et de pressions d'occupation du sol et sur les méthodes statistiques mise en œuvre.

CemOA : archive ouverte d'Irstea / Cemagref

II. Démarche et Zones d'études

II.1.Démarche de la thèse

Pour répondre à ces objectifs, la deuxième partie de la thèse est dédiée au développement de la méthode opérationnelle de caractérisation de l'OSCR, la troisième à la construction des indicateurs spatialisés et la quatrième au développement des modèles pressions/impacts régionalisés.

Deux zones d'études ont été définies pour atteindre ces objectifs. Le bassin versant de l'Hérault servira de zone d'étude pilote pour le développement des méthodologies de la deuxième et troisième partie. Les caractéristiques de ce bassin permettront de tester assez rapidement les différentes méthodologies envisagées et de répondre aux différents objectifs de ces parties. Cependant, compte tenu du faible nombre de stations de mesure de l'état écologique sur ce bassin, une zone d'étude plus large, le Nord-Ouest de l'hydro-écorégion Tables Calcaires (HER 9) a été sélectionnée pour répondre aux objectifs de modélisation. Elle possède un nombre de stations suffisant pour mener une démarche statistique robuste et fiable. Le choix de ces deux zones et leurs caractéristiques sont présentées plus en détail dans la suite de ce chapitre.

II.2.Présentation des zones d'études

II.2.1.La zone pilote : le bassin versant de l'Hérault

Le bassin versant du fleuve Hérault se situe au Sud de la France (à 30 km à l'Ouest de Montpellier) dans la région du Languedoc-Roussillon. Il prend sa source au mont Aigoual (44°7'22.01''N /3°34'29.98''E) dans le sud du massif des Cévennes. Il parcourt le département du Gard puis traverse du nord au sud le département de l'Hérault auquel il donne son nom et se jette dans la mer Méditerranée à Agde (43°16'59.63''N /3°26'38.37''E) (Figure 4.1). Il dispose d'une superficie de 2650 km² environ et comporte 22 sous bassins versants. Il est globalement soumis à un climat méditerranéen avec des pluies au dessus de 2000 mm/an en amont du bassin et en-dessous de 600 mm/an près de la côte.

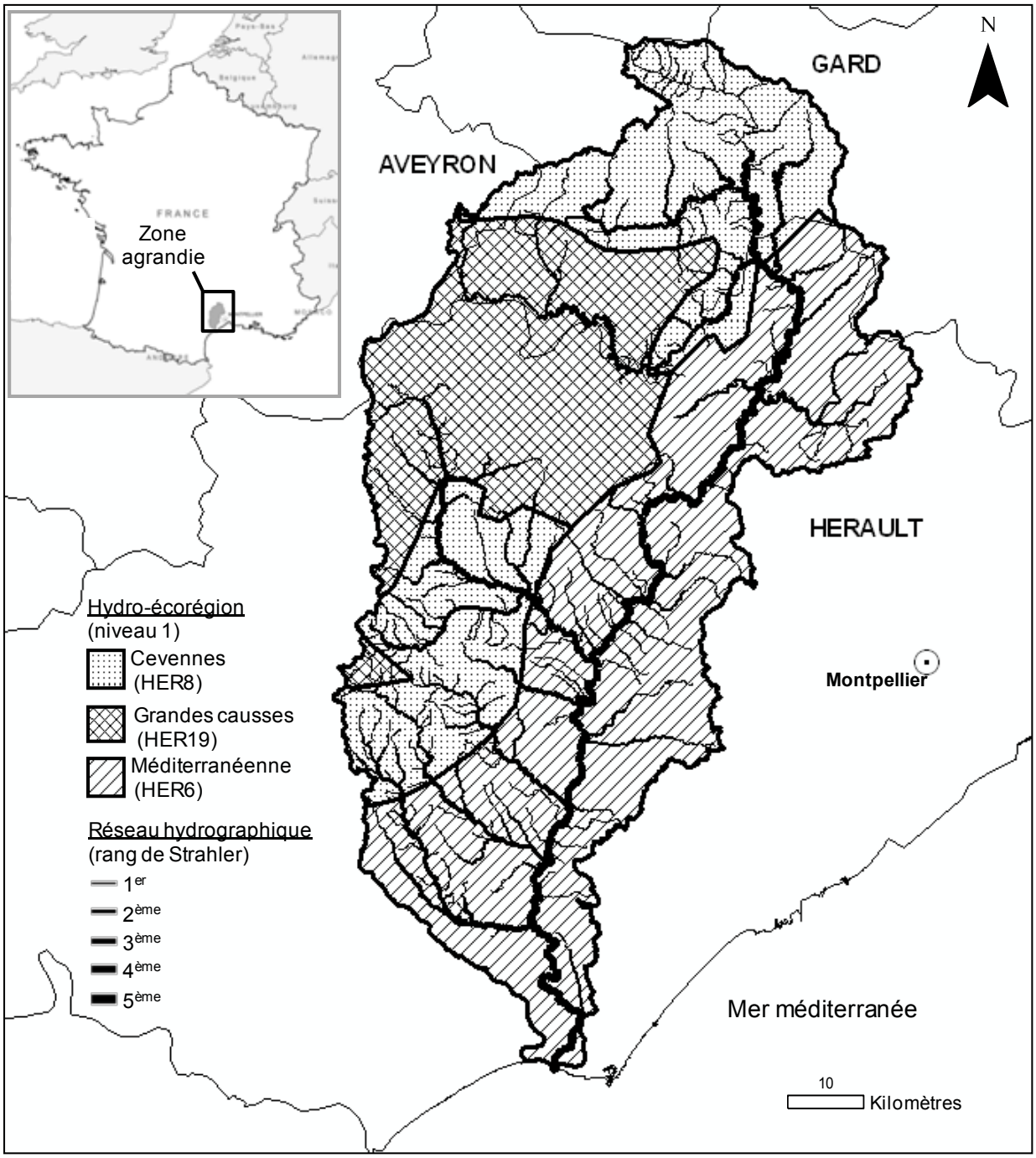
Plusieurs raisons ont suscité le choix de cette zone pilote, liées :

- aux caractéristiques hydrographiques : cette aire est drainée par plus de 1150 km de réseau hydrographique : 48 % sont des cours d'eau de 1er ordre, 19 % du 2ème, 16% du 3ème, 7% du 4ème, 11% du 5ème suivant l'ordination de Strahler, 1957) (Figure 4.1). Ces caractéristiques permettent de tester l'applicabilité et la reproductibilité de la méthode de caractérisation de l'OSCR sur un grand territoire et sur des cours d'eau de tailles différentes (notamment les petits cours d'eau en tête de bassin qui ont un rôle important dans l'effet filtre des zones rivulaires (cf. chapitre2) ;

- au contexte paysager : cette zone possède un paysage contrasté tant sur les caractéristiques naturelles qu'anthropiques. Le bassin parcourt trois hydro-écorégions (HER) différentes (Figure 4.1): l'HER Cévennes (HER8), une chaîne de montagne avec de fortes pentes et un socle cristallin ; l'HER Grandes Causses (HER19), un groupe de grand plateaux calcaires, entaillés par des vallées profondes et au réseau hydrographique superficiel très réduit; et l'HER Méditerranéen (HER6), un relief de plaines, entrecoupé localement de collines calcaires (jusqu'à 500 m) et caractérisé par des pentes faibles, variant localement (Wasson *et al.*, 2002). Ces caractéristiques physiques naturelles influent sur l'utilisation, le type et la structuration de l'occupation du sol. Alors que les couverts naturels (parsemés de zones urbaines) dominent le long des cours d'eau des HER8 et HER19, les activités agricoles (particulièrement la viticulture) et une urbanisation intensive dans la région (Balestrat *et al.*, 2008) s'imposent sur les corridors rivulaires de l'HER6 réduisant les ripisylves à de simples bandes de quelques mètres de part et d'autre du cours d'eau (< à 10m). Cette zone permet ainsi de tester aussi la reproductibilité de la méthode de caractérisation sur un gradient de caractéristiques naturelles et de pressions d'occupation du sol important.

Au-delà de ces deux critères ayant guidé le choix du bassin pilote, l'Hérault présentait l'avantage supplémentaire de la proximité géographique pour la réalisation de la vérité terrain (l'UMR TETIS est basée à Montpellier) et nous disposions déjà des données images aéroportées et satellitaires.

CemOA : archive ouverte d'Irstea / Cemagref



CemOA : archive ouverte d'Irstea / Cemagref

Figure 4.1 : Localisation de la zone pilote : le bassin versant de l'Hérault et présentation de ses caractéristiques hydrographiques et paysagères (hydro-écorégion).

II.2.2.Zone de modélisation : Nord Ouest de l'HER Tables-Calcaires (HER9-NO)

Cette zone d'étude est localisée au Nord Ouest de l'HER Tables-Calcaires (HER9-NO). Elle couvre plusieurs départements situés à l'Ouest de l'agglomération de Paris : La Seine-Maritime (76) et l'Eure qui sont complètement contenus ; puis les Yvelines (78), le Val-d'Oise (95), la Somme (80), l'Eure et Loir (28), l'Orne (61), le Calvados (14) et la Manche (50) qui sont en partie recouverts (Figure 4.2).

Elle contient l'ensemble des bassins versants des affluents de la partie aval de la Seine : l'Eure (5935 km²), l'Epte (1490 km²), l'Iton (1300 km²) et l'Andelle (740 km²) pour les plus grands ; puis plusieurs bassins versants des fleuves tels que la Risle (2310 km²), la Touques (1605 km²), la Dives (1573 km²) et la Béthune (307 km²).

Elle compte en tout 297 sous bassins versants qui forment au total une superficie de 25000 km² environ et sont drainés par plus de 6000 km de réseau hydrographique dont 48 % sont des cours d'eau de 1er ordre, 20 % du 2ème, 18% du 3ème, 6% du 4ème, 2% du 5ème, 1% du 6ème et 5 % du 7ème (la Seine) suivant l'ordination de Strahler.

L'HER Tables-Calcaires (HER9) constitue une région de plaine, d'altitude inférieure à 200 m, couvrant le bassin parisien et le nord du bassin aquitain ; le sous-sol est formé de roches sédimentaires tabulaires en majorité carbonatées.

Cependant, ces roches présentent en superficie des caractéristiques variables de perméabilité induisant des différences dans la densité du réseau de drainage, ou des différences de dureté et de résistance à l'érosion produisant localement des formes de reliefs plus marquées (côtes et collines) (Wasson *et al.*, 2002).

Sur la base de ces différences 13 HER de niveau 2 (HER2) ont été définies. La zone d'étude en couvre 7 : les HER2 Pays de Caen (30) et Cotentin Est (37) qui sont complètement contenues, et les HER2 Pays de Bray (35), Auréole crétacé du Bassin Parisien (38), Nord Loire et Perche (54) et Haute Normandie Picardie (57) recouvertes en partie. Les caractéristiques de chacune de ces HER2 sont présentées dans le détail dans l'***ANNEXE III***.

Plusieurs raisons ont suscité le choix de cette zone de modélisation, liées :

- aux caractéristiques hydro-biologiques : elle contient 158 stations (1038 relevés) mesurant la qualité écologique des cours d'eau à partir de l'Indice Biologique Global Normalisé (IBGN) (Figure 4.2). Ce nombre de stations est suffisant pour assurer la pertinence statistique de la démarche.

- aux contraintes financières et de temps : cette zone présente le meilleur compromis entre le nombre de stations et l'extension spatiale de la zone d'étude qui limite le nombre d'images de données à acquérir et à traiter.

- au contexte anthropique et ripicole : des structures particulières de bandes rivulaires sont présentes sur ce site (alternance de forêt de feuillus et de prairies proches du cours d'eau) dans une zone dominée par une forte occupation agricole, essentiellement portée sur les grandes cultures (oleoproteagineux, polyculture, fourrage) et l'élevage comme le montre la Figure 4.3. Ce paysage agricole est parsemé de tâches urbaines plus ou moins grandes et plus ou moins espacée dans la zone d'étude (Figure 4.4).

CemOA : archive ouverte d'Irstea / Cemagref

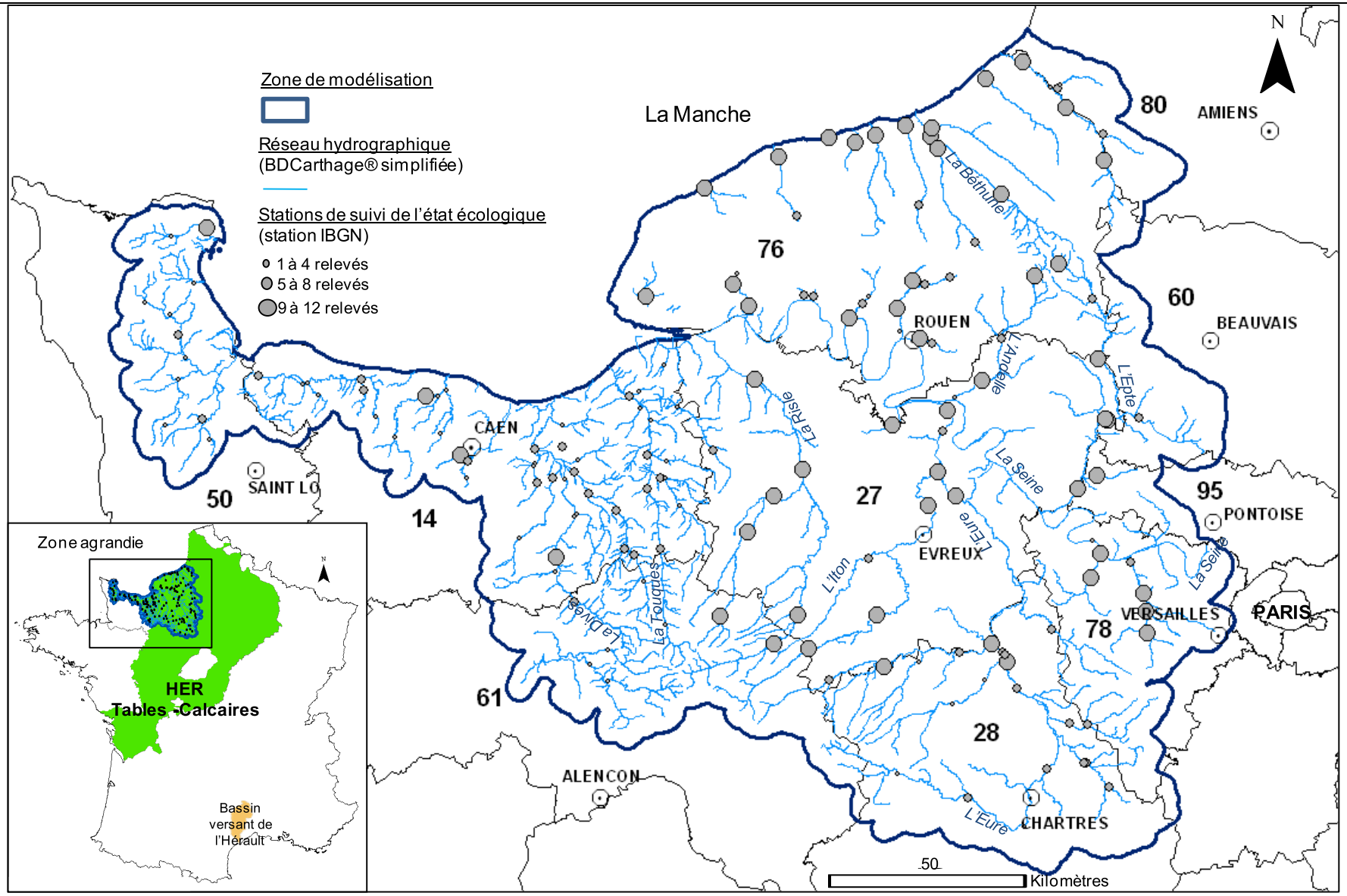


Figure 4.2 : Présentation de la zone de modélisation, Nord Ouest de l'HER Tables-Calcaires (HER9-NO)

CemOA : archive ouverte d'Irstea / Cemagref

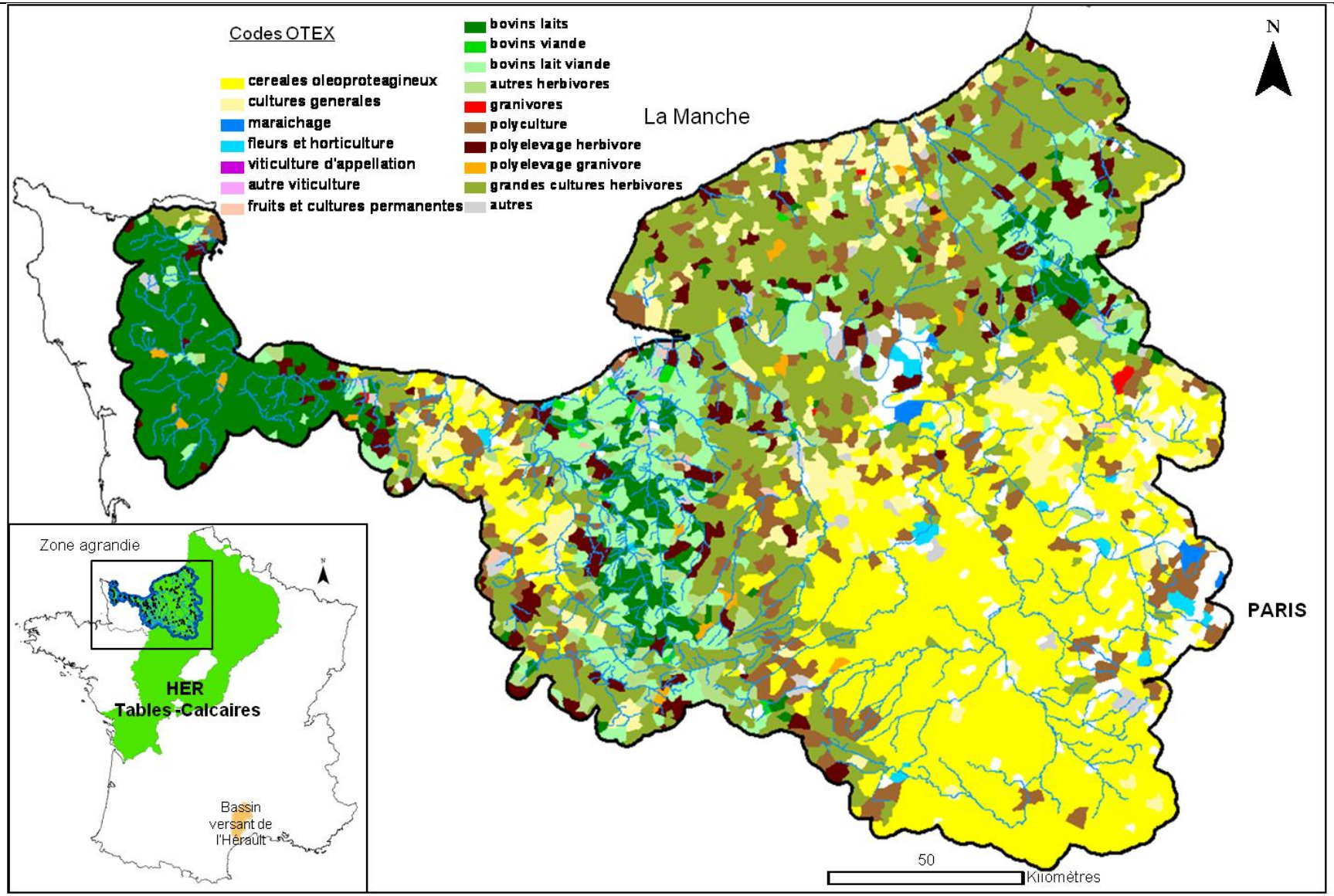


Figure 4.3 : Occupation du sol sur la zone d'étude en fonction de 5. Source : donnée CORINE Land Cover 2000 (IFEN)

CemOA : archive ouverte d'Irstea / Cemagref

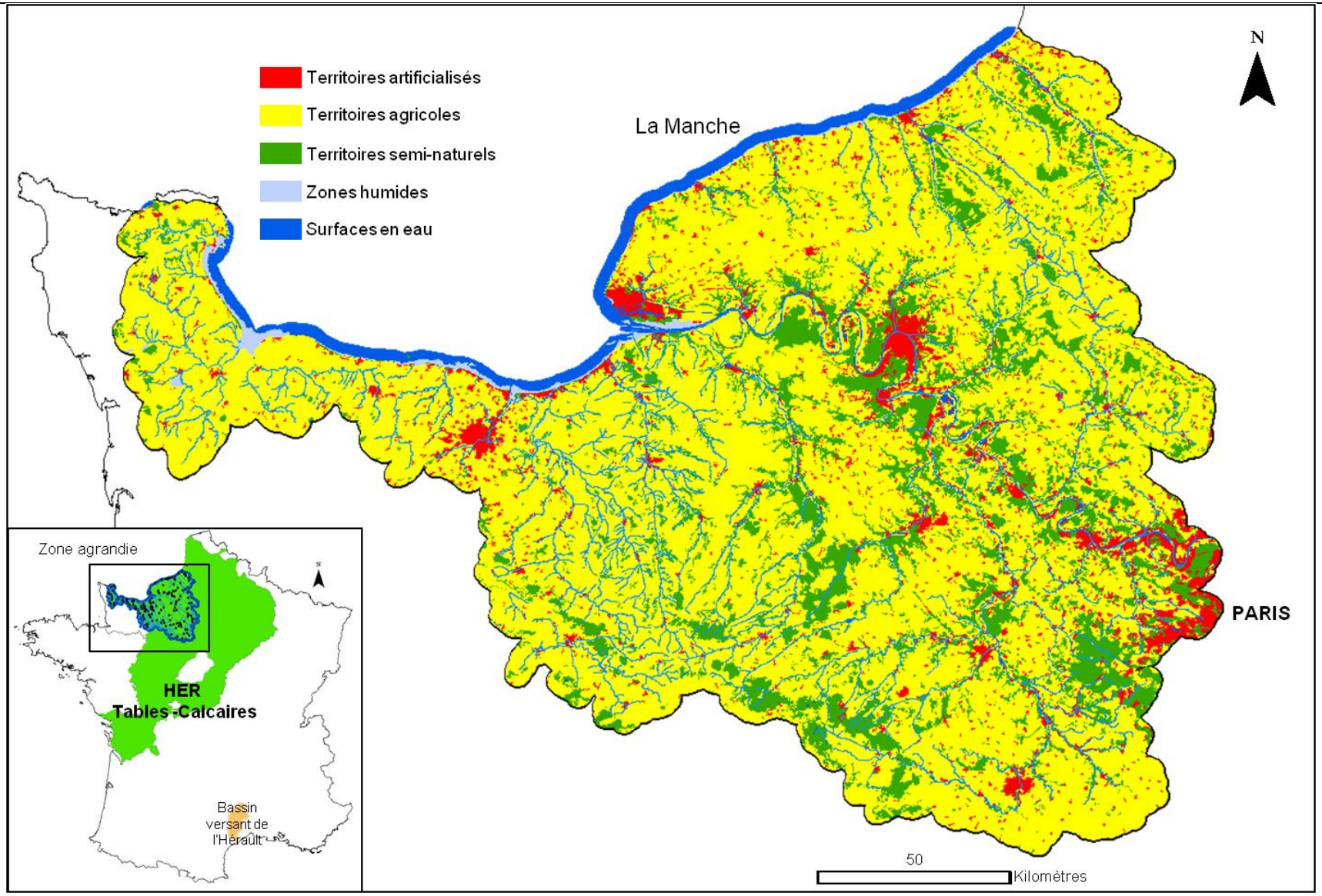


Figure 4.4 : Principales activités agricoles sur la zone de modélisation. Source : donnée OTEX (Orientation Technico-Économique des Exploitations) 18 postes (RGA)

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Deuxième partie :

Extraction de l’occupation du sol le long des cours d’eau par télédétection

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Introduction de la deuxième partie

Cette deuxième partie est consacrée au développement d'une méthode de classification de l'occupation des sols dans les corridors rivulaires à partir de données de télédétection (imagerie aéroportée et satellitaire) et de données auxiliaires existantes. Elle est structurée en deux chapitres.

Le chapitre 5 expose la conception de la procédure de classification. Cette procédure doit permettre d'identifier, dans différents contextes morphologiques, écologiques, et d'activité humaine, les objets de tailles réduites (quelques mètres à dizaines de mètres) localisés le long des cours d'eau. Elle doit être transposable à moindre coût pour l'ensemble du réseau hydrographique national (~500 000 km). Sa conception s'est faite par un travail progressif et itératif sur la zone étude pilote du bassin de l'Hérault. A mesure que nous formalisions et mettions en œuvre les différentes étapes, l'identification des principales difficultés permettait de concevoir et tester des alternatives méthodologiques afin d'optimiser cette procédure.

Le chapitre 6, présente les résultats de classification obtenus par la procédure optimisée sur le bassin de l'Hérault et évalue ses capacités (faisabilité technique et efficacité) pour extraire une donnée d'occupation du sol fine dans le corridor rivulaire sur de grands territoires.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Chapitre 5
Conception de la procédure de classification

SOMMAIRE

Résumé *125*

I. ÉTATS DES CONNAISSANCES 127

I.1. QUELQUES PRINCIPES DE CLASSIFICATION EN TELEDETECTION *127*

I.1.1. Les caractéristiques des images de télédétection 127
I.1.2. Les principes de la classification d'images 129
I.1.3. Les différentes méthodes de traitement 129
I.1.4. L'évaluation de la qualité du résultat de classification 130

I.2. LA TELEDETECTION DES CORRIDORS RIVULAIRES *130*

I.2.1. L'infrarouge proche : une information indispensable 130
I.2.2. La Très Haute Résolution Spatiale (THRS) : une nécessité 131
I.2.3. Revue des travaux existants 132
I.2.4. Conclusion 133

II. CONCEPTION D'UNE PROCEDURE DE CLASSIFICATION REGIONALE DE L'OCCUPATION DES SOLS DANS LES CORRIDORS RIVULAIRES : BESOINS, ARCHITECTURE ET OPTIMISATION 135

II.1. BESOINS ET CONTRAINTES *135*

II.2. ARCHITECTURE : UNE PROCEDURE EN SEPT ETAPES *136*

II.3. TROIS ETAPES A OPTIMISER *136*

II.4. DEMARCHE D'OPTIMISATION DE LA PROCEDURE *137*

III. PRESENTATION DE LA PROCEDURE 137

III.1. SELECTION DES SOURCES DE DONNEES (ETAPE 1) *137*

III.1.1. Présentation des données images 138
III.1.1.1 L'imagerie multispectrale Spot5 XS 138
III.1.1.2 L'imagerie THRS Spot5 PAN 139
III.1.1.3 L'imagerie THRS aéroportée : orthophotos IGN 139
III.1.2. Présentation des données auxiliaires (exogènes) 141
III.1.2.1 CORINE Land Cover (CLC) 141
III.1.2.2 Route 500® et BDCarthage® 141
III.1.2.3 Registre Parcellaire Graphique (RPG) 142
III.1.3. Choix optimisé des données 142
III.1.4. Conclusion de l'étape 1 144

… (suite)

CemOA : archive ouverte d'Irstea / Cemagref

… SOMMAIRE (suite)

III.2. DELIMITATION DES CORRIDORS RIVULAIRES (ETAPE 2) *145*
III.3. STRATIFICATION EN REGIONS HOMOGENES (ETAPE 3) *146*
III.4. APPLICATION DE PRETRAITEMENTS (ETAPE 4) *147*
III.4.1. La correction des images Spot5 XS 148
III.4.2. Le découpage et mosaïquage des données 149
III.4.3. La dégradation de l'orthophoto 149
III.4.4. La création d'un plan thématique 150
III.5. CLASSIFICATION DANS CHAQUE REGION (ETAPE 5) *151*
III.5.1. L'approche orientée objet multi-échelle : principes, méthodes et outils 151
III.5.1.1 Principe des réseaux hiérarchiques de classes 151
III.5.1.2 Choix de la plate-forme logicielle 154
III.5.1.3 Méthodes et algorithmes de segmentation 154
III.5.1.4 Méthodes et algorithmes de classification 156
III.5.1.4.1 Principes de la logique floue 157
III.5.1.4.2 La technique des fonctions d'appartenance (classification « expertisée ») 158
III.5.1.4.3 La technique du plus proche voisin (classification « supervisée ») 158
III.5.2. Optimisation du mode d'échantillonnage des parcelles d'apprentissage 160
III.5.3. Développement de la démarche de classification orientée objet 162
III.5.3.1 Choix du raisonnement de classification 162
III.5.3.2 Une démarche en trois phases 163
III.6. APPLICATION DE POST-TRAITEMENTS (ETAPE 6) *170*
III.6.1. Union des résultats de classification de chaque région 170
III.6.2. Structuration de la nomenclature 171
III.7. VALIDATION DU RESULTAT DE CLASSIFICATION (ETAPE 7) *171*
III.7.1. La matrice de confusion 172
III.7.1.1 Principe et critère d'évaluation de qualité 172
III.7.1.2 Mode d'échantillonnage de validation 173
III.7.2. La fiabilité et la stabilité de la classification 175

IV. SCHEMATISATION DE LA DEMARCHE DE CLASSIFICATION 177

CemOA : archive ouverte d'Irstea / Cemagref

Résumé

L'état des connaissances sur l'utilisation de la télédétection dans les espaces rivulaires montre que peu d'études ont cherché à développer une méthode opérationnelle pour caractériser l'occupation du sol des corridors rivulaires sur de grands territoires. Cependant, cette analyse indique clairement la nécessité (i) d'utiliser des images à Très Haute Résolution Spatiale (10 m à 0.5 m) contenant une information spectrale dans l'infrarouge proche (notamment pour discriminer plus efficacement la végétation dans ces espaces) ; (ii) d'utiliser la classification orientée objet pour extraire efficacement l'information contenue dans ces images ; et (iii) de recourir à l'information existante lorsqu'on travaille sur de grands territoires pour améliorer les résultats de classification et diminuer les temps de traitements.

Une procédure de classification en sept étapes, basée sur ces constats, a été développée afin de fournir une information pertinente pour comprendre les relations pressions / impacts liant l'occupation des sols dans les corridors rivulaires et l'état écologique des milieux aquatiques. Cette procédure a été conçue de manière (i) à identifier et caractériser les objets fins occupant les corridors rivulaires (notamment les ripisylves) selon une typologie la plus détaillée possible (en fonction de l'information apportée par les sources de données) avec pour objectif minimum de retrouver six principales catégories de pressions d'occupation du sol ; (ii) à estimer l'incertitude des résultats de la classification afin qu'elle soit prise en compte dans la construction d'indicateurs et dans l'élaboration des modèles pressions impacts ; et (iii) à être applicable dans différents contextes hydromorphologiques, écologiques et d'activités anthropiques (e.g., de zones de montagnes à des zones de plaine, de région humides à des régions sèches…) avec des moyens humains, financiers et matériels raisonnables.

Pour élaborer une telle procédure, des choix sur les matériels et les méthodes ont été réalisés pour chacune de ces étapes. Ils ont nécessité pour certaines d'entre elles des investigations adaptées afin de les optimiser par rapport aux besoins et aux contraintes de l'étude. La procédure optimisée utilise en entrée des sources de données disponibles sur l'ensemble du territoire français : deux sources de données images (images Spot5 XS - 10 m et orthophotos IGN – 0,5 m) et quatre sources de données auxiliaires (CORINE Land Cover, la Route 500®, les entités surfaciques de la BDCarthage® et le Registre Parcellaire Graphique).

Ces sources sont sélectionnées sur la zone d'étude à l'étape 1 de la procédure. A l'étape 2, l'espace à classer (les corridors rivulaires) est délimité en créant des buffers (de largeurs variables en fonction du rang de Strahler); puis la zone d'étude est stratifiée, à l'étape 3, en régions de types d'occupation du sol homogène (croisement des limites des hydro-écorégions avec les emprises d'images de même date). A l'étape 4, des prétraitements sont appliqués afin de découper les sources de données dans les différentes zones délimitées (corridors rivulaires et régions homogènes) et d'améliorer le contenu informatif des sources de données images et auxiliaires. Ensuite, la classification proprement dite est menée dans chaque région selon une approche orientée objet multi-échelle générique définissant une hiérarchie de quatre niveaux d'objets emboîtés et utilisant un arbre de classification descendant basé sur des règles de décisions booléennes ou floues déterminées par expertise. A l'étape 6, les résultats de classification issus de chaque région sont réunis dans une couche unique structurée selon une nomenclature hiérarchisée. Enfin, à l'étape 7, la qualité du résultat de classification est évaluée à la fois par la construction d'une matrice de confusion basée sur un nombre optimum (robuste et applicable) de 50 parcelles de référence par classe photo-interprétées (et sélectionnées selon un échantillonnage aléatoire stratifié), puis en utilisant les propriétés de la logique floue pour estimer la fiabilité et la stabilité de la classification.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

I. États des connaissances

Avant de présenter un état de l'art de l'application de la télédétection aux corridors rivulaires (I.2.), nous rappelons succinctement (I.1.) quelques principes de base de la télédétection, déterminants pour comprendre les choix que nous avons été amenés à faire pour développer une procédure de classification de l'occupation des sols dans les corridors rivulaires applicable aux échelles régionales.

I.1. Quelques principes de classification en télédétection

L'objectif de la classification d'images de télédétection (aéroportée ou satellitaire) est d'obtenir une répartition (un zonage) thématique à partir de l'information contenue dans ces images. Cette information, qui dépend des caractéristiques de l'image, est de nature spectrale, texturale, morphologique et géométrique. Ces caractéristiques sont présentées, avant d'exposer les principes de la classification et les différentes méthodes mises en œuvre.

I.1.1. Les caractéristiques des images de télédétection

Les données de télédétection, obtenues depuis des plateformes aéroportées ou satellitaires, sont le plus souvent des images numériques structurées sous forme de matrices spatialisées[22]. Elles contiennent une information sur le rayonnement électromagnétique réfléchi ou émis (luminance) par les objets détectés. Il existe une grande variété d'images qui peuvent être décrites selon trois caractéristiques fondamentales : la résolution spectrale, la résolution spatiale et la répétitivité temporelle.

La *résolution spectrale* est liée à l'aptitude d'un système de détection à distinguer des rayonnements électromagnétiques de fréquences (longueurs d'onde) différentes. En télédétection, le rayonnement électromagnétique est le plus souvent mesuré dans les longueurs d'onde du visible (0,38 µm à 0,78 µm)[23], du proche infrarouge (PIR, de 0,78 µm à 1,4 µm) et du moyen infrarouge (MIR, 1,4 à 3 µm), de l'infrarouge thermique (TIR, de 8 à 14 µm) et du radar (domaine des hyperfréquences, ondes qui traversent l'atmosphère (0,75 cm-100 cm) (Robin, 2002). Les images sont selon le cas acquises avec des capteurs optiques, thermiques ou radars. Les capteurs optiques et thermiques sont « passifs » car ils se « contentent » de mesurer la réflexion par les objets à la surface de la Terre des ondes émises par le soleil (optique), ou bien les ondes émises par ces objets eux-mêmes (thermique). À l'inverse, les capteurs radar sont « actifs » car ils émettent des ondes et en recueillent l'écho, ce qui leur donne l'avantage de pouvoir être utilisés de jour comme de nuit et d'être moins sensibles à la couverture nuageuse.

Au sein de ces trois catégories (optique, thermique, radar), l'information peut être mesurée dans une ou plusieurs bandes spectrales (gammes de longueurs d'ondes). Ainsi une image est dite : (1) « panchromatique » si elle mesure le rayonnement dans une seule gamme de longueurs d'ondes ; (2) « multispectrale » si l'information est mesurée dans plusieurs gammes de longueurs d'ondes ; et (3) « hyperspectrale » si l'information est mesurée dans un très grand nombre de bandes spectrales.

En télédétection, on suppose que tout objet ou classe d'objet sur la surface terrestre a, à un instant donné, sa propre « empreinte digitale » dans le spectre électromagnétique (la signature spectrale), puissance en fonction de la longueur d'onde du rayonnement réfléchi ou émis par lui-même (Girard and Girard, 1999).

L'utilisateur doit donc choisir une résolution spectrale adaptée aux comportements des objets géographiques ou phénomènes qu'il veut détecter et différencier. Par exemple, si on s'intéresse à la végétation, l'observation de la Figure 5.1 montre qu'il est intéressant d'acquérir une information dans le rouge et le PIR (bande 2 et 3 du satellite Spot5) puisque elle possède une signature spectrale (dans ces deux bandes) différente de celle de l'eau ou d'une roche (faible réflectance dans le rouge et forte réflectance dans le PIR).

[22] Nous n'abordons pas ici d'autres techniques de télédétection comme le Lidar, l'altimétrie radar, le sondage des gaz atmosphériques...
[23] violet 380–450 nm ; bleu 450–490 nm ; vert 490–570 nm ; jaune 570–585 nm ; orange 585–620 nm ; rouge 620–780 nm

CemOA : archive ouverte d'Irstea / Cemagref

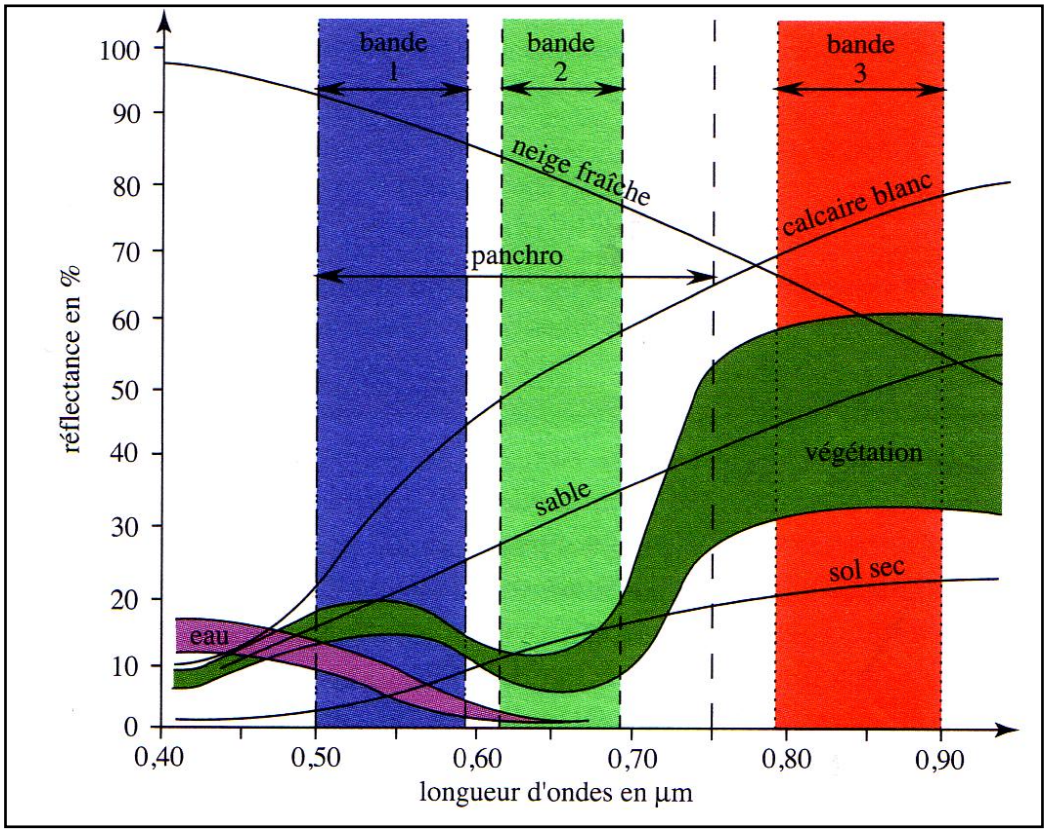


Figure 5.1 : Signatures spectrales typiques (eau, sol, végétation) en réflectance (rapport de la luminance sur l'énergie incidente) et exemples des bandes multi-spectrales retenues pour Spot5 un satellite dédié à l'observation de la terre (d'après un document CNES in Caron *et al.*, 2003)

La *résolution spatiale* d'une image est liée à l'aptitude d'un système de détection à distinguer des objets de positions spatiales différentes. Elle est définie par la dimension spatiale de ses pixels. Une image à 1 mètre de résolution est une image où le pixel correspond à un carré de 1 x 1 mètre au sol. Le pixel est la particule élémentaire de l'image numérique de télédétection. Chaque pixel contient une valeur (un compte) numérique qui traduit l'intensité du signal électromagnétique capté (ou autant de valeurs numériques que de bandes spectrales différentes).

L'utilisateur doit donc choisir une résolution spatiale adaptée à la taille des objets qu'il veut caractériser. Pour les capteurs civils, la taille des pixels peut varier de quelques dizaines de centimètres (ex. pixel de 61 cm de côté pour les images prises par le satellite Quickbird) à plusieurs dizaines de kilomètres (ex. satellites météorologiques géostationnaires). On parle d'imagerie à très haute résolution spatiale (THRS) pour une résolution décimétrique à métrique, de haute résolution (HRS) pour une résolution décamétrique, de moyenne résolution (MRS) pour une résolution hectométrique, de basse résolution pour une résolution kilométrique.

De façon générale, plus la résolution spatiale d'une image est fine (haute), plus petite est la fauchée (largeur de la zone couverte par l'image). Alors que les images des satellites de la NOAA (National oceanic & atmospheric administration) ont une fauchée de 3 000 km pour une résolution de 1100 m, une image Quickbird possède une fauchée de 16,5 km pour une résolution comprise entre 0,6 et 0,8 m.

La *répétitivité temporelle* (ou répétitivité ou cycle orbital) correspond à la fréquence à laquelle un satellite peut acquérir la même scène. Dans le cas d'un suivi régulier, l'utilisateur doit choisir une répétitivité temporelle adaptée à la dynamique des objets et des phénomènes étudiés.

Traditionnellement résolution spatiale et répétitivité temporelle sont inversement liées : un satellite géostationnaire à très haute répétitivité temporelle fournit des images à basse résolution, un satellite à très haute résolution spatiale, qui couvre des bandes de fauchée étroites, ne peut assurer une haute répétitivité temporelle. Toutefois, compte tenu des progrès techniques et de la possibilité de programmation d'images sur des sites géographiques précis, cette incompatibilité entre haute répétitivité temporelle et haute résolution

CemOA : archive ouverte d'Irstea / Cemagref

spatiale tend à s'estomper en termes techniques pour ne subsister qu'en termes économiques : une haute répétitivité temporelle peut techniquement être obtenue aussi bien pour des images à basse résolution spatiale (e.g, les satellites NOAA de 1100 m de résolution spatiale revisitent chaque scène 2 fois par jour), que pour des images à haute résolution spatiale (e.g, le satellite FORMOSAT de 5 mètres de résolution spatiale peut revisiter une scène 1 fois par jour).

I.1.2. Les principes de la classification d'images

La classification consiste à assigner à chaque élément de l'image (pixel ou groupe de pixels appelé « objet ») une classe (catégorie) unique, le plus souvent parmi un ensemble de classes préalablement déterminé (typologie).

Elle se fait sur la base d'attributs (descripteurs ou caractères) définis à partir de l'information contenue dans les images de télédétection. Il existe plusieurs types d'attributs : les attributs spectraux calculés à partir des valeurs spectrales, les attributs texturaux calculés à partir de l'organisation spatiale (configuration et forme) de ces valeurs spectrales (Lake, 1991), les attributs temporels calculés à partir de la dynamique temporelle des valeurs attachées à un pixel ou un objet. De nouveaux descripteurs peuvent être définis en combinant (somme, différence, comparaison, combinaison linéaire…) ces différents attributs (e.g, indices de végétation comme le NDVI, indices d'humidité des sols, …). Enfin d'autres familles d'attributs utiles pour la classification peuvent être définies : attributs morphologiques caractérisant la forme des objets, attributs géométriques caractérisant leurs positions relatives dans l'espace...

Le choix des attributs descriptifs pertinents est un processus clé pour obtenir de bons résultats de classification. Il doit permettre un regroupement de pixels de telle sorte que les pixels qui dans la réalité appartiennent à une même classe, soient le plus proches possible les uns des autres dans l'espace des attributs, et que les pixels ou objets de classes différentes soient les plus éloignés possibles. C'est le postulat de toute classification (Girard and Girard, 1999). Il se fait principalement de deux manières :

(1) de manière expertisée : l'interprète utilise ses connaissances et son expérience sur le comportement des objets et des phénomènes étudiés afin de sélectionner les attributs discriminants. Il s'aide généralement d'outils de visualisation (composition colorée, affichage des valeurs de l'attribut selon une gamme de couleur) ;

(2) de manière automatique en utilisant des algorithmes de sélection et d'extraction de critères discriminants (analyse en composante principales…).

I.1.3. Les différentes méthodes de traitement

Traitements manuels et traitements automatiques

La classification de ces images peut se faire soit par des traitements manuels (la photo-interprétation) soit par des traitements automatiques.

La *photo-interprétation* consiste à analyser l'image visuellement. En fonction de son expérience, l'opérateur reconnaît dans l'image les formes et les couleurs (les descripteurs) qui correspondent à son thème d'étude (chaîne de traitement œil-cerveau). Il délimite alors manuellement les objets détectés et attribue à chacun une classe.

Les *traitements automatiques*[24] consistent remplacer le photointerprète par un algorithme et pour cela à déterminer des règles d'attribution d'un pixel ou d'un objet à une classe en fonction de la valeur de ses attributs, règles qui seront mises en œuvre de façon automatique. Ceci permet de traiter de façon homogène de gros volumes d'images.

Les règles permettent de partitionner automatiquement l'espace défini par les attributs en respectant au mieux les postulats de classification. Elles sont le plus souvent statistiques, arithmétiques ou logiques, et peuvent être hiérarchisées (appliquées successivement selon une arborescence) (Girard and Girard, 1999). Les méthodes de classification statistique par exemple, utilisent dans leur grande majorité pour le partitionnement les notions de distances mathématiques (e.g. distances euclidiennes, de Manhattan ou de Mahalanobis) avec l'objectif de minimiser la distance (variance) intra-groupes et maximiser la distance (variance) inter-groupes.

La classification automatique peut se faire de façon supervisée (assistée) ou non supervisée.

[24] Pour éviter toute méprise en ce qui concerne les traitements dits « automatiques », il y a lieu de rappeler (Girard and Girard, 1999) qu'il y a toujours une intervention de l'opérateur pour paramétrer les traitements.

CemOA : archive ouverte d'Irstea / Cemagref

- Dans le cas d'une *classification non supervisée*, on recherche une partition des données (pixels ou objets) en fonction de leur seule structure, sans présumer de la nature des classes. L'interprète ne donne donc pas d'information a priori sur les classes à définir, mais peut intervenir en choisissant le nombre de classes, les valeurs des seuils….
- Dans le cas d'une *classification supervisée*, le partitionnement s'effectue sur la base d'une typologie de classes préalablement définie. Il peut (1) utiliser des parcelles d'entraînement (de référence) sélectionnées pour chaque classe, chaque pixel ou objet étant alors classé en fonction de son positionnement dans l'espace des attributs, par rapport aux parcelles de référence ; (2) définir, de manière expertisée, des règles logiques d'appartenance pour chacune des classes.

Approche pixel et approche orientée objet

L'application de ces traitements automatiques peut se faire soit par pixel (classification par pixel), soit par objet (classification orientée objet).

La *classification par pixel* consiste à appliquer le même traitement à chaque pixel, indépendamment des pixels voisins, de façon à déterminer la classe à laquelle il sera associé. Elle se base essentiellement sur les valeurs numériques des pixels (signature spectrale) et/ou la configuration dans l'espace de ces valeurs (signature texturale).

La *classification orientée objet* consiste à identifier préalablement des objets composés de plusieurs pixels puis à opérer la classification de ces objets. Ainsi, le pixel n'est plus traité de manière isolée mais dans son contexte. Pour cela on fait appel à un algorithme de segmentation qui divise l'image en plusieurs groupes de pixels montrant des caractéristiques similaires. Il existe de nombreuses méthodes de segmentation d'image rassemblées en trois grandes catégories (Bonn and Rochon, 1992) : les approches par « région » ascendante (agrégations successives de pixels) et/ou descendante (divisions successives de l'image en régions), les approches par « contour » (détection des zones hétérogènes de transition entre deux objets), ou les approches mixtes qui combinent les deux approches précédentes (les informations des éléments de contours sont utilisées pour faciliter le classement et celles sur l'intensité des pixels des régions pour affiner la forme finale des contours). Un groupe de pixels est alors considéré comme un objet (ou région) relevant d'une classe unique, ce qui rend possible une interprétation basée sur les valeurs spectrales, texturales à l'échelle de l'objet mais aussi sur la taille, la forme, le contexte…

I.1.4.L'évaluation de la qualité du résultat de classification

La télédétection est passée d'une phase expérimentale à une phase opérationnelle pour un certain nombre d'applications. Il est donc indispensable de fournir, avec les documents infographiques ou cartographiques qui en sont issus, une évaluation de la qualité de la classification.

On pourrait se contenter d'une appréciation visuelle (analyse qualitative), mais cette méthode reste très subjective. La démarche la plus courante (Foody, 2002) consiste à confronter la classification à des vérités terrain (parcelles de validation) par une analyse quantitative : la matrice de confusion. Cette vérité terrain (la référence) peut être collectée par des observations de terrain, par photo-interprétation, ou bien par une combinaison des deux (Girard and Girard, 1999).

Quatre mesures synthétiques de fiabilité pour évaluer la classification peuvent être calculées (Congalton, 1991): (1) la précision totale; (2) la précision pour l'utilisateur; (3) la précision pour le réalisateur et (4) l'indice de Kappa. Ces différentes mesures de la qualité de la classification seront présentées plus précisément dans la partie III.7 de ce chapitre.

I.2. La télédétection des corridors rivulaires

I.2.1. L'infrarouge proche : une information indispensable

L'analyse des travaux de télédétection réalisés sur les corridors rivulaires montre que quelque soit l'objet d'étude (l'occupation du sol au sens large ou uniquement la ripisylve), l'information spectrale dans l'infrarouge est très souvent utilisée, notamment le Proche Infra Rouge (PIR, de 0,78 µm à 1,4 µm) et le Moyen Infra Rouge (MIR, 1,4 à 3 µm).

CemOA : archive ouverte d'Irstea / Cemagref

Elle est indispensable pour détecter et caractériser plus précisément la végétation (Guyot, 1997; Müller, 1997). En plus d'avoir un comportement spectral dans le visible dû à l'existence de pigments chlorophylliens, la végétation présente un comportement particulier dans le PIR, lié à la structure des tissus, et dans le MIR relatif à sa teneur en eau Figure 5.2).

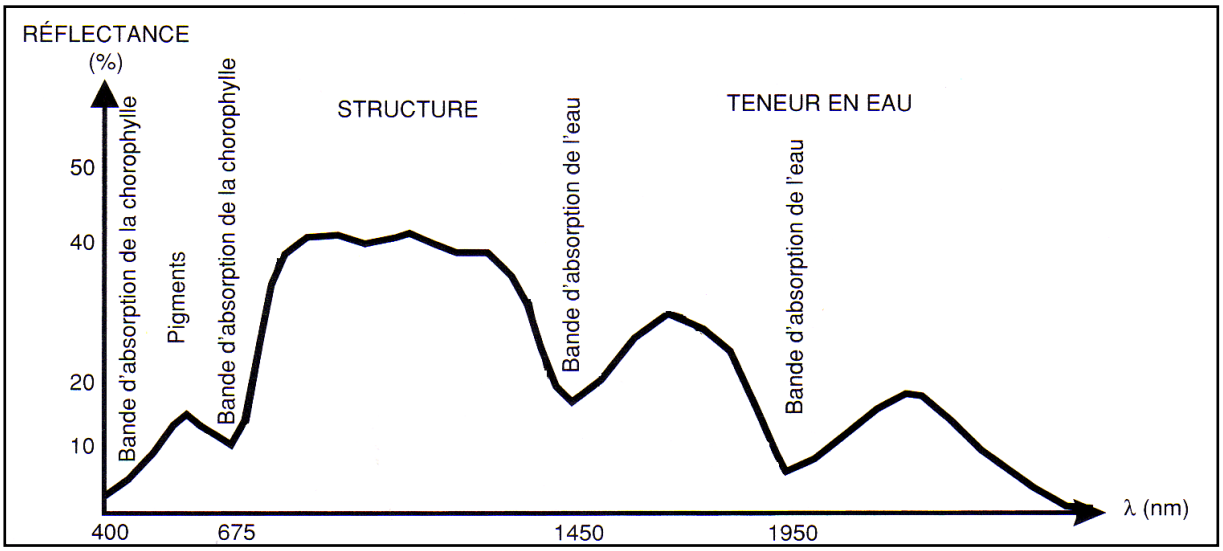


Figure 5.2 : Comportement spectrale de la végétation (modifié de Girard and Girard, 1999)

De nombreux indices de végétation (e.g. NDVI, IRI, SAVI)[25] utilisant cette information ont été employés dans les études dédiées à la caractérisation de la végétation rivulaires (e.g, Nagler *et al.*, 2001 ; Johansen and Phinn, 2006a ; Mollot *et al.*, 2007).

De plus l'information dans le PIR améliore la détection des zones d'eau qui ont une reflectance faible dans cette bande (elle apparaît donc en noir, i.e valeurs de comptes numériques très faibles) (Girard and Girard, 1999).

I.2.2. La Très Haute Résolution Spatiale (THRS) : une nécessité

De nombreux travaux ont montré que la télédétection à Très Haute Résolution Spatiale (THRS[26]) est indispensable pour extraire une information précise dans ces espaces (e.g, Müller, 1997; Congalton *et al.*, 2002).

Les images optiques à haute-moyenne résolution spatiale (Landsat 30 m et Spot4 20 m), traditionnellement utilisées pour inventorier de grands territoires à moindre coût (Congalton *et al.*, 2002), sont en effet le plus souvent inadaptées pour extraire une information précise dans les corridors rivulaires : la forte concentration en voies de communication, en cultures intensives, en ouvrages hydroélectriques, en infrastructures d'urbanisation, y rend l'occupation du sol très fragmentée et hétérogène ; et la végétation rivulaire est souvent réduite à une fine bande le long des cours d'eau. La résolution à 20m ou plus du pixel de ces images n'est donc pas adaptée pour capturer la composition et la configuration de cette occupation (Müller, 1997; Neale, 1997; Shuft *et al.*, 1999). Il en résulte que les caractéristiques des corridors rivulaires ne sont donc pas (ou faiblement) représentées sur les données géographiques issues de l'interprétation de ces images.

Aujourd'hui, avec le progrès des nouvelles technologies d'acquisition, de plus en plus précises, informatives et répétitives, un large choix de capteurs THRS est disponible pour caractériser finement les corridors rivulaires. Ainsi différents types de capteurs ont été utilisés dans les travaux sur le sujet : (1) l'imagerie optique multi-spectrale (e.g. Neale, 1997; Gergel *et al.*, 2007) aussi bien satellitaire (e.g., Quickbird, Ikonos) qu'aéroportée ;

[25] NDVI : Normalized Diffrence Vegetation Index, IR : Infra-Red index, SAVI : Soil-Ajusted Vegetation Index

[26] La THRS sont des images inférieures à 5 m de résolution spatiale. Pour simplifier la lecture, on intègre aussi dans le sigle THRS les images à HRS (Haute Résolution Spatiale) de 10 m de résolution spatiale.

CemOA : archive ouverte d'Irstea / Cemagref

(2) l'imagerie optique hyper-spectrale aéroportée (Bryant and Gilvear, 1999; Leckie *et al.*, 2005; Mollot *et al.*, 2007) ; (3) l'imagerie LIDAR[27] (Witte *et al.*, 2001 ; Dowling and Accad, 2003; Blair *et al.*, 2006).

I.2.3. Revue des travaux existants

Bien que les images LIDAR apportent une information intéressante notamment pour caractériser plus précisément la structure de la végétation rivulaire (Goetz, 2006) et que les images hyper-spectrales permettent de quantifier de façon détaillée l'occupation du sol dans ces corridors (Mollot *et al.*, 2007), leur coût d'acquisition et de traitement reste encore trop élevé aujourd'hui pour les utiliser dans le cadre d'une classification sur de grands territoires.

Nous nous sommes donc principalement intéressés dans cette synthèse aux études utilisant les données plus accessibles de télédétection optiques multi-spectrales (Ashraf *et al.*, 2010).

La technique la plus couramment employée pour le traitement de ce type de données de télédétection est la photo-interprétation. Elle a été très utilisée pour extraire une information d'occupation du sol dans les corridors rivulaires, notamment aux Etats-Unis à partir d'images aériennes (e.g., CBP, 2003). Cependant, bien qu'un photo-interprète puisse extraire relativement facilement des informations utiles sur ces images optiques, le processus de photointerprétation est consommateur de temps, particulièrement lorsque l'unité de cartographie (niveau de détail) est élevée (Leckie *et al.*, 2005), et peut être particulièrement subjectif (Coulter *et al.*, 2000). C'est pour cette raison que la cartographie des corridors rivulaires par photointerprétation n'a pas été systématisée dans la plupart des recherches sur les relations pressions/impacts (Goetz, 2006).

De nombreux auteurs ont donc cherché à caractériser de façon automatique, à partir d'images optiques à THRS, les différentes surfaces présentes dans les corridors rivulaires. Ces travaux se différencient par l'objet d'étude, la taille de la zone d'étude et les techniques de classification utilisées (Tableau 5-1).

La plupart de ces travaux se sont focalisés sur la caractérisation de l'état de la végétation rivulaire afin d'évaluer l'impact des activités humaines sur ces systèmes en cherchant : (1) à délimiter l'étendue des ripisylves (Johansen and Phinn, 2006a ; Hubert-Moy *et al.*, 2006 ; Yang, 2007) ; (2) à quantifier leur structure par des indicateurs comme le pourcentage des différents couverts, la hauteur de canopée (Johansen and Phinn, 2006b; Johansen *et al.*, 2007) ; et (3) à classer les différentes espèces (ou communautés) de végétation (Lonard *et al.*, 2000 ; Weber and Dunno, 2001 ; Davis *et al.*, 2002 ; Johansen and Phinn, 2006a ; Gergel *et al.*, 2007 ; Noonan and Chafer, 2007) ;

D'autres travaux se sont intéressés à quantifier l'occupation du sol dans les corridors rivulaires (y compris la végétation) (Neale, 1997 ; Van der Sande *et al.*, 2003 ; Goetz *et al.*, 2003 ; Volkman *et al.*, 2005 ; Hubert-Moy *et al.*, 2006).

Ces études ont été généralement menées sur des zones géographiques d'extension limitée (tronçon < 100 km et bassin versant < à 400 km²) à l'exception de trois études.

Noonan and Chafer, 2007, ont cartographié l'invasion des sauleraies à l'échelle d'un bassin versant (2158 km²) d'Australie (dans l'état du Nouvelle-Galles du Sud) en utilisant l'imagerie Spot5 XS en bi-date (début-automne et hiver) et les photos aériennes ;

Yang, 2007, a mis en place une méthode pour délimiter les corridors rivulaires sur 593 km de cours d'eau supérieur au rang 4 (selon l'ordination de Strahler) à partir de photos aériennes et d'images Landsat dans la région de Hunter en Australie;

Goetz *et al.*, 2003, ont développé une méthodologie permettant de caractériser les surfaces arborées et imperméables à partir d'images Ikonos sur un bassin versant de 1313 km² dans la région du Mid-Atlantic des Etats-Unis.

On peut remarquer que ces recherches appliquées sur de grandes zones d'études ont fait appel à des données auxiliaires complémentaires, améliorant ainsi les résultats de classification et diminuant les temps de traitements (Goetz *et al.*, 2003 ; Yang, 2007).

Ces travaux ont généralement été réalisés en mettant en œuvre des techniques conventionnelles de classification par pixel, aussi bien « non supervisées » (e.g, algorithme ISODATA[28] Goetz *et al.*, 2003; Johansen

[27] Le LiDAR (Light Detection And Ranging) basé sur l'émission-réception de faisceaux Laser à partir d'une plateforme aéroportée ou satellitaire apporte une information sur la structure verticale des couverts végétaux (Lefsky *et al.*, 2002).

CemOA : archive ouverte d'Irstea / Cemagref

and Phinn, 2006b) que « supervisées » (e.g, maximum de vraisemblance, Weber and Dunno, 2001 ; Volkman *et al.*, 2005 ; Noonan and Chafer, 2007).

Cependant, de nombreux auteurs ont constaté qu'une résolution spatiale plus fine n'améliore pas forcément le résultat des classifications par pixel (Latty and Hoffer, 1981 ; Irons *et al.*, 1985 ; Müller *et al.*, 1993 ; Kressler *et al.*, 2003 ; Lennartz, 2004). L'information spectrale contenue dans l'image à THRS étant très hétérogène (Puissant and Weber, 2004) on constate deux biais : d'une part un pixel ne correspond plus à l'élément (ou objet) que l'on veut classer, et d'autre part, cet élément est composé de plusieurs pixels possédant des comportements spectraux différents qui peuvent appartenir à différentes classes (exemple : une lande composée d'une mosaïque de composantes herbacées et arborées ; une vigne composée d'une alternance de sols nus et de végétation). Ainsi, lorsqu'on classe une image à THRS par pixel, la variabilité intra-classe augmente et la variabilité inter-classe diminue (le postulat de classification n'est plus respecté) et les éléments que l'on cherche à classer sont constitués de pixels de classes différentes (effet « poivre et sel »). La précision du résultat de classification est par conséquent dégradée et les cartes obtenues sont difficilement exploitables à cause de cet effet « poivre et sel » (Ivits and Koch, 2002 ; Durieux *et al.*, 2007).

La classification orientée objet s'avère être une technique plus adaptée pour extraire l'information contenue dans les images à THRS. Cette classification permet de retrouver les éléments que l'on veut classer dans l'image en segmentant préalablement l'image en objets et en classant par la suite ces objets en fonction de leurs attributs spectraux et spatiaux (textures, forme, contexte) plus discriminants que les attributs du pixel. Elle améliore ainsi la discrimination des classes d'occupation du sol par rapport aux classifications par pixel (Repaka, 2004 ; Lennartz, 2004). Cependant peu d'études de télédétection ont exploré les potentialités de cette technique sur les espaces rivulaires (Van der Sande *et al.*, 2003 ; Hubert-Moy *et al.*, 2006 ; Johansen *et al.*, 2007 ; Gergel *et al.*, 2007).

Les résultats de classification ont été validés, pour la plupart de ces études, en construisant des matrices de confusion. Le protocole d'échantillonnage des parcelles de validation utilisées pour la construction de ces matrices dépend des habitudes propres de chaque auteur. On remarque tout de même que pour les recherches appliquées sur de plus grands territoires, les parcelles de validation ont été généralement obtenues par photo-interprétation (e.g, Goetz *et al.*, 2003 ; Yang, 2007). En effet, la collecte de données terrains est difficile à mettre en œuvre compte tenu de la taille de la zone d'étude (Zhu *et al.*, 2000).

I.2.4. Conclusion

L'état des connaissances sur l'utilisation de la télédétection dans les espaces rivulaires montre que peu d'études ont cherché à développer une méthode opérationnelle pour caractériser l'occupation du sol des corridors rivulaires sur de grands territoires.

Cependant, cette analyse indique clairement la nécessité d'utiliser :
- des images comportant une information spectrale dans l'infrarouge proche pour différencier et caractériser plus facilement la végétation dans les corridors rivulaires et extraire les zones d'eau ;
- des images à THRS pour détecter finement l'occupation du sol dans ces espaces ;
- des données géographiques existantes (données auxiliaires) pour améliorer les résultats de classification et diminuer les temps de traitements ;
- la classification orientée objet pour extraire efficacement une information d'occupation du sol à partir de ces données images.

[28] ISODATA : Iterative Self-Organizing Data Analysis Technique

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 5-1 : Études caractérisant les corridors rivulaires à partir d'images optiques multi-spectrales mono-date à THRS (Très Haute Résolution Spatiale) et à l'aide de traitements automatiques. (B=Bleu ; V=Vert ; R=Rouge ; PIR=Proche Infrarouge ; MIR=Moyen Infrarouge ; IRTL=Infrarouge thermique Lointain ; PAN=Panchromatique).

Auteurs	Année	Objet d'étude	Zone d'étude	Capteurs	Résolution spatiale	Résolution spectrale	Techniques de classification	Précision résultats
Gergel et al.	2007	Végétation rivulaire (types)	Corridor rivulaire 32 km²	Satellites Quickbird II	2.8 m	B, R, V, PIR	Objet - supervisée Plus Proche Voisin	70 à 90 %
Johansen et al.	2007	Végétation rivulaire (types et structure)	Corridor rivulaire 50 km²	Satellites Quickbird II	2.8 m 0.7 m	B, R, V, PIR PAN	Objet - supervisée Plus Proche Voisin	73.68 %
Yang	2007	Végétation rivulaire (délimitation)	Bassin versant 1180 km²	Photographie aérienne Landsat ETM+	1-2m 30 m	B, V, R B, V, R, PIR, MIR (2), IRTL	Pixel - supervisée Parallélépipède	81%
Noonan & Chafter	2007	Végétation rivulaire (invasion des sauleraies)	Tronçon 2158 km	Satellite SPOT 5 XS Photographie aérienne	10 m 0.4 m	V, R, PIR, MIR B, V, R	Pixel - supervisée Maximum de vraisemblance Photo-interprétation	77.50%
Johansen & Phinn	2006	Végétation rivulaire (types et délimitation)	Tronçon 2 km	Satellite Ikonos	4 m 1 m	B, V, R, PIR PAN	Pixel –non supervisée ISODATA	54.56 %
Johansen & Phinn	2006	Végétation rivulaire (types et structure)	Tronçon 10 km	Satellites Quickbird II	2.8 m 0.7 m	B, V, R, PIR PAN	Objet - supervisée Plus Proche Voisin	88.54 %
Hubert-Moy	2006	Zone humide riparienne (délimitation et caractérisation)	Non renseigné	Satellite SPOT 5 XS	10 m	V, R, PIR, MIR	Objet - supervisée Plus Proche Voisin Fonction d'appartenance	55 % (délim.) 62 % (caract.)
Volkman et al. 2005	2005	Occupation du sol	Bassin versant 100 km²	Satellites Quickbird II	2.4 m 0.61 m	B, V, R, PIR PAN	Pixel - supervisée Maximum de vraisemblance	84 %
Goetz et al.	2003	Occupation du sol (surface arborée et imperméables)	Bassin versant 1313 km²	Satellite Ikonos	4 m 1 m	B, V, R, PIR PAN	Pixel non supervisée ISODATA Supervisée Maximum de vraisemblance Arbre de décision	83 % 86 % 84 et 97 %
Van der Sande	2003	Occupation du sol	Corridor rivulaire 16 km²	Satellite Ikonos	4 m 1 m	B, V, R, PIR PAN	Objet-supervisée Fonction d'appartenance	74 %
Weber & Dunno	2001	Occupation du sol	Tronçon 2 km	Atlas aéroportée	2.5 m	R, PIR, MIR	Pixel - supervisée Maximum de vraisemblance	40 %
Neale	1997	Occupation du sol	Tronçon 13 km	CCD aéroportée	0.4 m	V, R, PIR	Pixel - supervisée Maximum de vraisemblance	ND

CemOA : archive ouverte d'Irstea / Cemagref

II. Conception d'une procédure de classification régionale de l'occupation des sols dans les corridors rivulaires : besoins, architecture et optimisation

II.1. Besoins et contraintes

A la lumière des travaux menés par différentes équipes (cf. I.2), il est indispensable d'expliciter les objectifs de la procédure de classification de l'occupation des sols dans les corridors rivulaires que nous souhaitons élaborer avant de nous engager dans la conception de son architecture.

Les résultats de la classification doivent fournir une information pertinente pour comprendre les relations pressions/impacts liant l'occupation des sols dans les corridors rivulaires et l'état écologique des milieux aquatiques (objet de la partie 4). Ceci implique plusieurs contraintes :

1. La procédure doit permettre d'identifier et caractériser les objets fins occupant les corridors rivulaires, notamment les ripisylves.
2. Elle doit permettre d'obtenir une information la plus détaillée possible (en fonction des sources de données utilisées) avec pour objectif minimum de discriminer les six principales pressions d'occupation du sol influençant l'état écologique des cours d'eau : surface en eau, sols nus semi-naturels, sols artificialisés, sols agricoles, zones de végétation arborée et zones de végétation herbacée (cf. Chapitre 4 :I.1 pour le détail de ces catégories).
3. Elle doit pouvoir être adaptée à une (ou des) typologie(s) de classes d'occupation des sols, éventuellement hiérarchisée(s), jugée(s) pertinente(s) par les hydrobiologistes et écologues.
4. Les résultats de la classification doivent être accompagnés d'une caractérisation de leur incertitude afin qu'elle soit prise en compte dans la construction d'indicateurs et l'élaboration des modèles pressions/impacts.

La procédure doit pouvoir être mise en œuvre de façon homogène sur l'ensemble du territoire national. Ceci impose trois nouvelles contraintes :

5. Elle doit pouvoir être adaptée à différents contextes hydromorphologiques, écologiques et d'activité anthropique, de zones de montagnes à des zones de plaine, de région humides à des régions sèches, de zones naturelles à des zones fortement anthropisées. Cela implique que la procédure soit conçue à un niveau générique, transposable dans ces différents contextes.
6. Elle doit s'appuyer sur des données disponibles avec une qualité homogène sur l'ensemble du territoire national (données de référence IGN[29], imagerie satellitaire, Référentiel Parcellaire Graphique,…).
7. Elle doit être réalisable avec des moyens humains, financiers et matériels raisonnables.

A noter qu'on ne cherche pas à produire une donnée de référence avec une grande précision géométrique (métrique voire submétrique) puisque cette donnée d'occupation du sol est dédiée à la construction d'indicateurs spatialisés agrégeant cette information dans une emprise spatiale donnée. On essayera tout de même d'être le plus précis possible (de quelques mètres à quelques dizaines de mètres), en fonction des données utilisées, afin de limiter les erreurs des indicateurs s'intéressant aux objets d'occupation du sol au contact du cours d'eau.

[29] IGN : Institut National Géographique

CemOA : archive ouverte d'Irstea / Cemagref

II.2. Architecture : une procédure en sept étapes

Sur la base des besoins et contraintes explicitées ci-dessus, une procédure de classification spécifique a été développée pour extraire l'occupation du sol dans les corridors rivulaires. Elle se compose de sept étapes (Figure 5.3) détaillée dans la partie III. et dont une présentation synoptique détaillée est donnée dans la Figure 5.35:

- A **l'étape 1**, les sources de données images et les sources de données auxiliaires adaptées à la caractérisation de l'occupation du sol dans les corridors rivulaires et disponibles à moindre coût sur la zone d'étude sont sélectionnées.
- A **l'étape 2**, les corridors rivulaires (espace à classer) sont délimités.
- A **l'étape 3,** la zone d'étude est stratifiée (« découpée ») en régions de types d'occupation du sol homogènes afin d'améliorer les résultats de classification.
- A **l'étape 4,** des prétraitements sont appliqués afin de découper les sources de données dans les différentes zones délimitées (corridors rivulaires et régions homogènes) et d'améliorer le contenu des sources de données images et auxiliaires.
- A **l'étape 5,** la classification proprement dite est menée dans chaque région selon une approche orientée objet, de façon à fournir une information d'occupation du sol détaillée.
- A **l'étape 6,** cette information est post-traitée pour obtenir une cartographie sur l'ensemble de la zone d'étude selon une nomenclature structurée.
- A **l'étape 7,** la qualité et la fiabilité de cette carte sont évaluées.

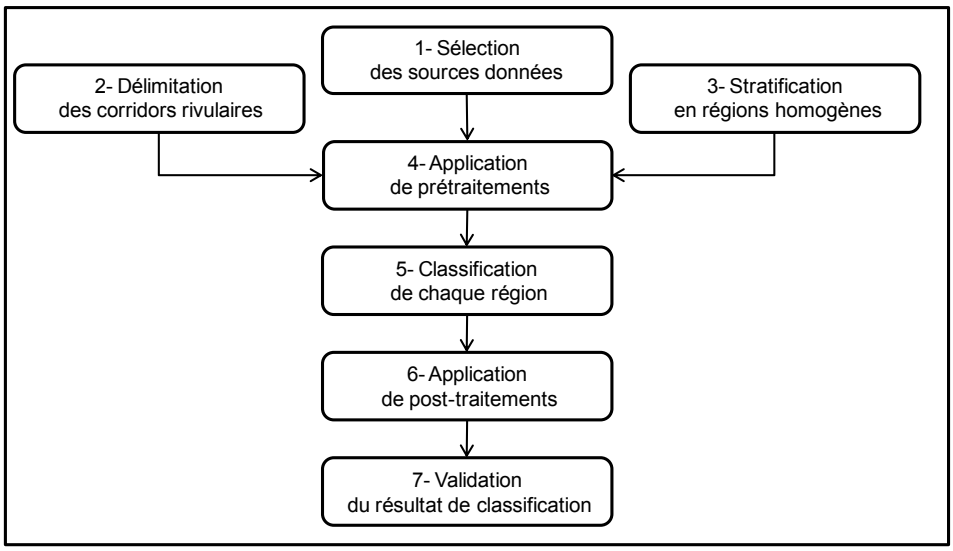


Figure 5.3 : Les sept étapes de la procédure de classification de l'occupation du sol dans les corridors rivulaire.

II.3. Trois étapes à optimiser

Pour chacune de ces étapes, des choix sur les matériels et les méthodes ont été réalisés en lien avec les objectifs de classification.

Pour trois d'entre elles ce choix a nécessité des recherches approfondies afin d'optimiser la procédure :

- Pour l'étape 1 (« Sélection des sources de données ») : trois protocoles de combinaisons différentes de sources de données disponibles sur le territoire français ont été testés afin de sélectionner la combinaison permettant d'obtenir les meilleurs résultats de classification ;

- Pour l'étape 5 (« Classification de chaque région »): deux techniques de classification orientée objet ont été confrontées pour analyser leur aptitude à extraire l'occupation du sol sur de grands territoires (en terme d'applicabilité, de reproductibilité et d'efficacité) ;

CemOA : archive ouverte d'Irstea / Cemagref

- Pour l'étape 7 (« Validation du résultat de classification ») : des analyses statistiques ont été menées pour déterminer le nombre minimal de parcelles de validation par classe nécessaire pour évaluer de façon fiable la qualité du résultat de classification sur l'ensemble de la zone d'étude à partir des matrices de confusion.

II.4. Démarche d'optimisation de la procédure

Compte tenu du lien entre ces trois étapes, il a été nécessaire de procéder de façon itérative pour mener à bien ces investigations.

Nous nous sommes tout d'abord intéressés à choisir la technique de classification orientée objet (optimisation de l'étape 5) en fixant donc arbitrairement les choix sur les sources de données en entrée et le nombre de parcelles de validation par classe pour la construction des matrices de confusion.

Ensuite, en utilisant la technique de classification optimisée, on a testé les potentialités des différentes sources de données images (optimisation de l'étape 1) en fixant arbitrairement le nombre de parcelles de validation par classe.

Enfin, on a évalué le nombre de parcelles de validation par classe optimal (optimisation de l'étape 7) en utilisant les données images et la technique de classification optimisées.

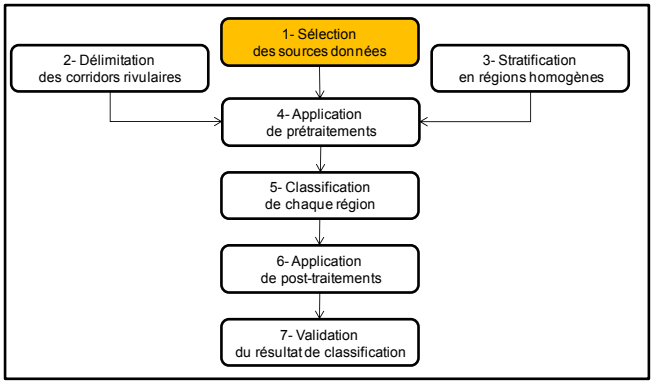
Cette démarche itérative est présentée dans le détail dans l'***ANNEXE VII*** :

- ***ANNEXE VII.1*** : Sélection de la technique de classification orientée objet
- ***ANNEXE VII.2*** : Choix optimisés des sources de données
- ***ANNEXE VII.3*** : Identification de l'échantillonnage optimal de validation

Dans la suite de ce chapitre, on expose les choix sur les matériels et les méthodes qui ont été pris dans les sept étapes de la procédure de classification, en présentant, bien évidement, pour les étapes qui ont nécessité des investigations supplémentaires (étape 1, 5, 7) les principaux résultats issus de la démarche itérative d'optimisation.

III. Présentation de la procédure

III.1. Sélection des sources de données (étape 1)



Plusieurs types de capteurs à THRS on été utilisés pour caractériser l'occupation du sol dans les corridors rivulaires (cf. I.2.2). Compte tenu des objectifs de classification, on s'est cependant intéressé uniquement aux images optiques multi-spectrales couvrant le territoire français. Ces images sont plus facilement disponibles et ont un coût d'acquisition et de traitement beaucoup moins élevé que d'autres types de données de télédétection à THRS (e.g, LIDAR, images optiques hyper-spectrales aéroportées). Disposer d'une information temporelle au cours de l'année (données diachroniques) présenterait un réel intérêt pour améliorer la

CemOA : archive ouverte d'Irstea / Cemagref

discrimination de certaines classes, par exemple entre les cultures et les surfaces herbacées semi-naturelles comme les prairies agricoles et les friches (Van der Sande *et al.*, 2003 ; Khorram, 2003). Toutefois nous avons limité la procédure à l'utilisation d'une seule date d'acquisition, sélectionnée si possible parmi les images d'archives, ceci afin de limiter le coût financier de la procédure comme requis par la contrainte 7 du chapitre II.1.

De plus nous nous sommes également intéressés aux données géographiques françaises disponibles (données auxiliaires), qui peuvent utilement compléter l'imagerie satellitaire et/ou aéroportée. Nous présentons dans cette partie successivement les différentes sources de données présélectionnées puis la combinaison de sources de données optimale pour extraire l'occupation du sol dans les corridors rivulaires.

III.1.1. Présentation des données images

D'après la littérature, l'information de la bande proche infrarouge à THRS est nécessaire pour caractériser les corridors rivulaires. C'est pour cela que les images optiques multi-spectrales issues des satellites Ikonos[30] et Quickbird[31] ont été très utilisées dans ces études (I.2.3). Cependant, le coût d'acquisition de ces images serait exagérément élevé pour couvrir l'ensemble du territoire national. Nous nous sommes donc orientés vers les images multi-spectrales du satellite Spot5 (Spot5 XS) contenant cette information à la résolution de 10m. L'opérateur de ce satellite (Spot Image) propose à l'heure actuelle[32] des tarifs parmi les plus intéressants pour les gestionnaires publics français.

A titre de comparaison le prix 2009 au km² (en prenant en compte le coût des prétraitements géométriques et radiométriques) d'une image Ikonos ou Quickbird multi-spectrale est approximativement de 16 euros/km² contre approximativement 1 euro/km² pour une image Spot5 XS. Cependant la résolution spatiale des données SPOT5 (10 m) semble limitante pour caractériser finement ces espaces.

Deux sources de données supplémentaires, apportant une information à THRS disponible à coût réduit sur le territoire français, ont été envisagées : les orthophotos IGN (0.5 m de résolution spatiale) et les images panchromatiques du satellite Spot5 (Spot5 PAN à 2.5 m de résolution spatiale).

III.1.1.1 L'imagerie multispectrale Spot5 XS

Le capteur de ce satellite dispose d'une fauchée au sol de 60 km et de bandes spectrales à la fois dans le domaine du visible et du proche infrarouge en haute résolution spatiale (10 m) (Figure 5.5). Comme les satellites Spot 1, 2 et 4, Spot5 est placé sur une orbite circulaire pour assurer une résolution constante (l'altitude est de 822 km à l'équateur), quasi polaire pour pouvoir prendre des images de tout point de la terre, et héliosynchrone pour que les conditions d'éclairement solaire varient le moins possible. Cette orbite permet une couverture complète de la terre en 26 jours.

Ces images satellite (Spot5 XS) sont livrées suivant différents niveaux de prétraitement radiométrique et géométrique. Le niveau de base (1A) consiste à corriger les défauts radiométriques provenant des écarts de sensibilité entre les détecteurs élémentaires de l'instrument de prise de vue (égalisation radiométrique) sans aucune correction géométrique. Des corrections géométriques sont réalisées sur les autres niveaux de prétraitement (1B, 2A, 2B et 3) qui se différencient par le niveau de précision géométrique. Plus on augmente de niveau, plus la précision géométrique est élevée (…) et plus le coût de l'image est élevé.

Dans le cas où ces images sont utilisées avec d'autres sources de données (e.g les orthophotos) le niveau 3 (niveau le plus précis) est requis. Il permet de limiter le décalage entre les différentes sources de données qui est souvent à l'origine d'erreurs de classification.

Les corrections du niveau 3 se basent sur la modélisation dynamique du vol du satellite, afin de corriger les distorsions dues aux conditions de prises de vues, et sur des informations géographiques ou cartographiques provenant de points d'appui et d'un MNT (Modèle Numérique de Terrain) pour corriger les erreurs de parallaxe dues au relief.

Un coût d'acquisition supplémentaire, non négligeable, est lié à ce prétraitement. Pour réduire ce coût, des logiciels de traitement d'images (e.g, module LPS d'Erdas) existent pour réaliser les mêmes corrections

[30] Information spectrale dans le Bleu, le Vert, le Rouge et Proche Infrarouge à 4 m de résolution spatiale

[31] Information spectrale dans le Bleu, le Vert, le Rouge et Proche Infrarouge à 2.8 m de résolution spatiale

[32] En France, les gestionnaires auront bientôt accès aux photographies aériennes avec information dans le PIR en cours de production par IFN et IGN. Ils auront le choix avec les données multi-spectrale à THRS des satellites de la constellation Rapid-Eye (multi-spectral 5m R, V, B et Red-Edge) ainsi que, à partir de 2011, avec les images des satellites de la constellation Pléiades (2,8m multispectral avec PIR et 0,7m panchromatique).

CemOA : archive ouverte d'Irstea / Cemagref

géométriques sous-réserve d'avoir un MNT et une source de données géographiques (permettant de sélectionner des points d'appuis) assez précis sur la zone d'étude.

Grâce au programme national ISIS[33] du CNES, qui vise à promouvoir l'utilisation de l'imagerie spatiale auprès de la communauté scientifique française en lui permettant d'acquérir ces données satellitaires à un tarif préférentiel, nous avons pu dans cette étude commander directement ces images au niveau de prétraitement requis (niveau 3).

Choix de la date d'acquisition

Qu'il s'agisse de programmer des acquisitions d'images ou d'exploiter des images d'archive, le choix de la date d'acquisition de ces images est très important, notamment pour distinguer les cultures de la végétation herbacée semi-naturelle (prairies agricoles, friches…). Les couverts de cultures, caractérisés par une succession dans le temps de différents stades phénologiques (e.g germination, levée, montaison, floraison, maturation pour les cultures), peuvent avoir à une date donnée une physionomie identique à celle de la végétation herbacée et donc un comportement spectral similaire. L'objectif sera donc d'acquérir une image à la date où le comportement spectral entre ces deux types de couverts est le plus différent.

Dans la pratique ce choix de date est difficile car cette sélection requiert des informations expertes sur les types de cultures et les pratiques culturales locales qui sont longues à obtenir et difficilement généralisables lorsque l'on travaille sur de grands territoires.

De plus, on doit tenir compte de deux autres critères qui limitent la liste de choix de dates récentes disponibles sur la zone d'étude : (1) un minimum de couverture nuageuse sur l'image (la présence des nuages empêche d'extraire l'information au sol et leur ombre altère la qualité radiométrique de l'image) ; et (2) un taux de couverture maximum sur la zone d'étude (pour réduire le coût d'acquisition de ces images). On évite enfin d'acquérir des images d'hiver (novembre à mars) car l'angle zénithal solaire élevé pendant cette période (soleil bas sur l'horizon) augmente les ombres portées, ce qui bruite l'information captée. Nous privilégions donc une acquisition sur la période mai-août.

III.1.1.2 L'imagerie THRS Spot5 PAN

Spot5 fournit aussi des images panchromatiques (Spot5 PAN) à une résolution spatiale de 2.5 m (Figure 5.5) grâce au procédé Supermode®. A bord du satellite l'acquisition est réalisée par un détecteur spécifique comportant deux barrettes CCD décalées dans le plan focal d'un demi-pixel (2,5 m) en ligne et en colonne ; l'instrument génère alors deux images à 5 m qui sont transmises séparément au sol ; ces images subissent alors une succession de traitements (interpolation, déconvolution, débruitage) pour produire une image à 2.5 m de résolution (Spotimage, 2003).

Ces images satellites constituent une solution économique proche de la THRS plus avantageuse que d'autres sources ou d'autres procédés. Comme pour Spot5 XS, ces images sont livrées selon différents niveaux de prétraitement et sont disponibles à des fins de recherche via le programme ISIS. On a donc pu acquérir ces images au niveau 3 de correction.

Le choix de la date d'acquisition doit se faire aussi en minimisant le taux de couverture nuageuse et en maximisant le taux de recouvrement sur la zone d'étude. Pour les mêmes raisons que dans le cas de l'imagerie Spot5 XS les images d'hiver sont à éviter (angle zenithal solaire trop élevé).

III.1.1.3 L'imagerie THRS aéroportée : orthophotos IGN

Les orthophotos sont des mosaïques numériques d'orthophotographies couleur (5 km par 5 km) issues de la « BD ORTHO® V2 » de l'IGN (Institut Géographique National), renouvelées tous les cinq ans pour chaque département. Les campagnes d'acquisition sont programmées par département et généralement lancées durant les mois d'été pour limiter la couverture nuageuse (IGN, 2009).

Ces mosaïques sont issues de photographies aériennes couleur ortho-rectifiées géographiquement, possédant trois canaux dans le visible (Rouge, Vert, Bleu), et une résolution spatiale de 0.5 m (Figure 5.6). Les déformations dues au relief du terrain, à l'inclinaison de l'axe de prise de vue et à la distorsion de l'objectif sont corrigées.

[33] Incitation à l'utilisation Scientifique des Images SPOT (www.isis-cnes.fr)

CemOA : archive ouverte d'Irstea / Cemagref

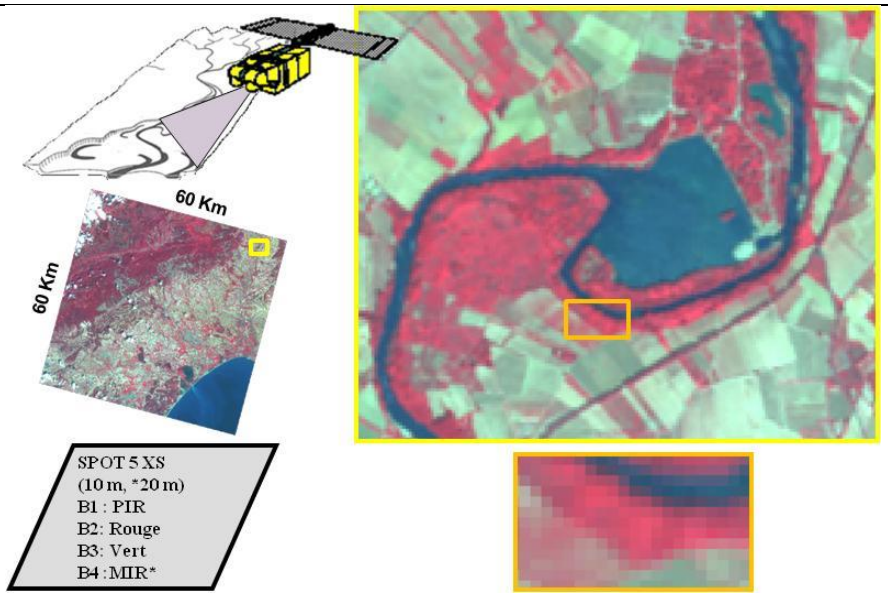


Figure 5.4 : Caractéristiques des images satellite Spot5 XS :

ces images possèdent une information spectrale dans le Vert, le Rouge, le Proche Infrarouge à 10 m et une information dans le Moyen Infrarouge à 20 m. Elles couvrent une surface de 3600 km².

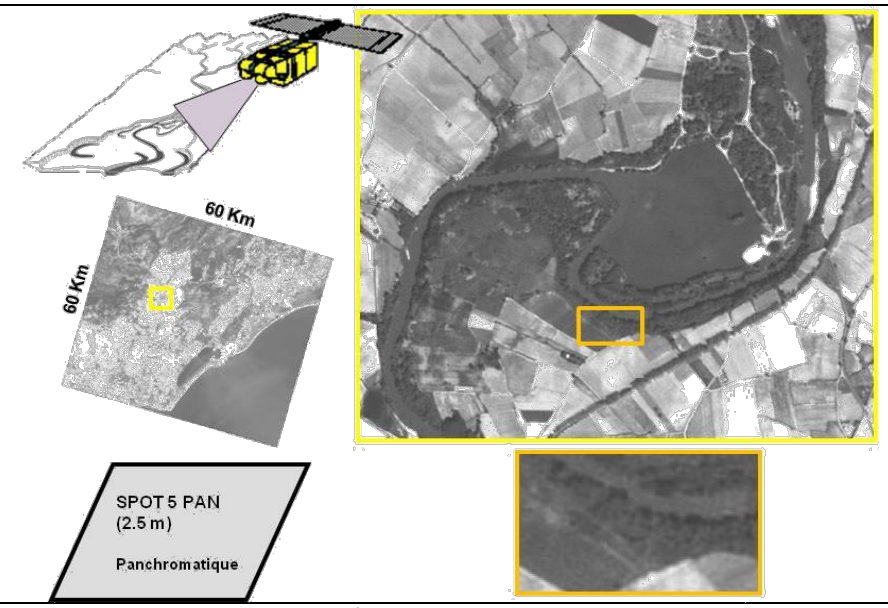


Figure 5.5 : Caractéristiques des images satellite Spot5 PAN:

ces images ont une résolution spatiale de 2.5 m. Elle couvre une surface de 3600 km².

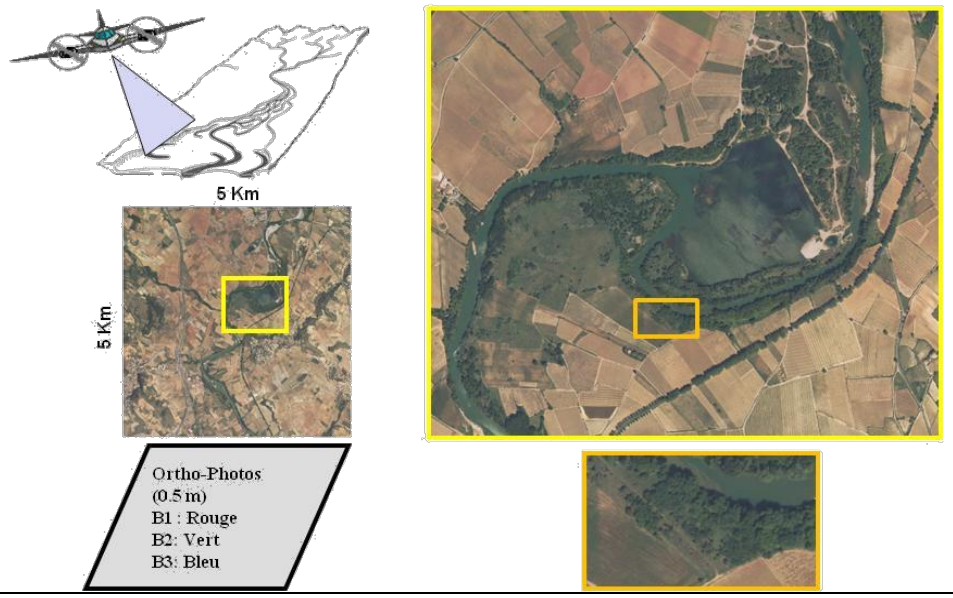


Figure 5.6 : Caractéristiques des orthophotos :

l'orthophoto est une mosaïque de photographies aériennes rectifiées géométriquement d'une résolution spatiale de 0.5 m comportant une information spectrale dans le visible (le Rouge, le Vert, le Bleu). Ces orthophotos sont issus de la BD ORTHO ® IGN et sont livrées par dalles (mosaïque) de 25 km².

CemOA : archive ouverte d'Irstea / Cemagref

Le nombre de photos nécessaire pour couvrir un département est important. Ces clichés présentent des disparités fortes d'ensoleillement, d'apparence colorée, de paysage, de comportement d'objets (carrières, sables, surfaces d'eau, …) et parfois de saison. Pour réduire cette variabilité, des corrections radiométriques sont effectuées par l'IGN afin d'obtenir une mosaïque d'ortho-images (orthophotos) continue et homogène sans surface saturée conséquente à l'échelle du département. Pour cela, chaque cliché est corrigé des effets internes d'éclairement et de couleurs liés à l'ensoleillement. L'ensemble des clichés est ensuite homogénéisé et la mosaïque est globalement rehaussée en couleur et en dynamique afin d'obtenir un rendu le plus naturel possible (Paparoditis *et al.*, 2006).

III.1.2.Présentation des données auxiliaires (thématiques)

Avant de procéder à une classification sur un thème donné, il est recommandé de rechercher s'il existe des données géographiques existantes et fiables permettant de réduire les erreurs et les confusions dans la classification (Girard and Girard, 1999). Disposer de surfaces connues permet d'enrichir la connaissance de l'opérateur sur le site d'étude, de compléter l'information liée aux images et de réduire la quantité de surface à traiter. Quatre sources de données thématiques disponibles à coût réduit à l'échelle du territoire national ont été sélectionnées.

III.1.2.1 CORINE Land Cover (CLC)

La couche d'information géographique CORINE[34] Land Cover 2000 (CLC) est la base de données d'occupation du sol européenne issue de la photo-interprétation assistée par ordinateur d'images Landsat (30 m) et Spot4 (20 m) acquises sur la période de 1999 à 2000 (Bossard *et al.*, 2000). L'unité cartographiée est de 25 ha minimum (500 m × 500 m). Elle repose sur une nomenclature standard hiérarchisée à 3 niveaux et 44 postes (***ANNEXE IV.1***) répartis selon 5 grands types d'occupation du territoire (territoires artificialisés, territoires agricoles, forêts et milieux semi-naturels, zones humides et surfaces en eaux). Elle possède une précision de classification supérieure à 85 %.

L'unité cartographique retenue (25 ha) ne permet clairement pas une caractérisation fine et détaillée de toutes les classes d'occupation du sol dans les corridors rivulaires et notamment de la végétation le long des cours d'eau. Cependant, les grands objets de territoires artificialisés denses, appartenant par exemple aux postes « tissu urbain continu » ou « zone industrielle et commerciale », sont aussi bien détectés avec CLC qu'à partir des images à THRS.

Neuf catégories d'objets appartenant aux territoires artificialisés ont été sélectionnées (***ANNEXE V***). Ils constituent une source de données complémentaire pour l'extraction des sols artificialisés le long des cours d'eau.

III.1.2.2 Route 500® et BDCarthage®

L'extraction automatique des routes est très difficile par télédétection car elles ont un comportement spectral très hétérogène lié aux différents types de revêtements et elles sont souvent masquées par la végétation arborée plantée à proximité. On a donc utilisé la Route 500® (BDR) qui est la base de données routière de l'IGN, décrivant 500 000 km de routes du réseau classé et des éléments d'habillage. La précision de cette donnée est décamétrique. Seules les entités liées au réseau routier ont été utilisées (autoroutes, routes nationales… cf. ***ANNEXE V***).

La BD Carthage® (BDC) est la base de données de référence du réseau hydrographique français. Elle est construite en partie à l'aide des cartes IGN à l'échelle du 1 : 25 000. Elle fournit également une information sur les éléments surfaciques (éléments de surface) en eau : zones couvertes d'eau douce permanente (plans d'eau, bassin, réservoir,…) ou non permanente (zones temporairement recouvertes d'eau, …), les zones couvertes d'eau salée permanente (mer, bassin portuaire,…) ou non permanente (marais salants, zone de vase,…), les glaciers et les névés. Ces informations géographiques sont mobilisées dans le processus de classification orientée objet. Les différentes classes d'éléments surfaciques utilisées sont présentées en ***ANNEXE V***. La précision de cette donnée est décamétrique.

[34] Co-Ordination of Information on the Environnement

CemOA : archive ouverte d'Irstea / Cemagref

III.1.2.3 Registre Parcellaire Graphique (RPG)

Le Registre Parcellaire Graphique (RPG) est le système d'identification géographique informatisée des parcelles agricoles soumises à la Politique Agricole Commune (PAC). Chaque exploitant doit déclarer ses parcelles éligibles en numérisant manuellement sur un fond cartographique (les orthophotos) le contour des îlots de ses parcelles. Les déclarations sont ensuite contrôlées (sur le terrain ou par photo-interprétation d'images optiques multi-spectrales) par les DDEA (Directions Départementales de l'Équipement et de l'Agriculture).

L'ASP (Agence de Service et de Paiement) est chargée de gérer le SIG pour l'identification des îlots et la description de leur occupation culturale et règlemente sa diffusion (site internet ASP). Les établissements publics et les administrations ont accès au RPG anonyme (i.e., privé de toute information personnelle) renseignant uniquement sur la géométrie des îlots et la nature et la surface des couverts déclarés dans chaque îlot (un îlot peut être composé de plusieurs parcelles de natures différentes).

Cette donnée d'une précision métrique est renouvelée tous les ans. Elle apporte donc une information spatialisée annuelle sur l'utilisation des terres agricoles à l'échelle du terroire français. Les types de cultures déclarés sont présentés en ***ANNEXE V***.

Remarque : Au début de la thèse, les établissements publics et les administrations avaient accès uniquement à la géométrie de l'îlot. Par conséquent seule cette information a été utilisée lors du développement de la méthodologie sur le bassin versant de l'Hérault. L'information sur le type de culture a été utilisée par la suite sur la zone de modélisation.

III.1.3. Choix optimisé des données

Un protocole d'analyse a été mis en place sur 350 km de réseau hydrographique d'une région homogène du bassin versant de l'Hérault (zone appartenant à l'HER6) afin de sélectionner, parmi les sources de données images présélectionnées sur le territoire français, la combinaison de sources optimale pour classer l'occupation du sol le long des cours d'eau sur de grands territoires. Trois combinaisons (C1, C2 et C3), contenant au minimum l'information dans l'infrarouge proche du satellite Spot 5 XS et l'information thématique des données auxiliaires, ont été explorées :

- C1 : Spot5 XS + les données auxiliaires
- C2 : Spot5 XS + Spot5 PAN + les données auxiliaires
- C3 : Spot5 XS + orthophotos + les données auxiliaires

A partir de ces trois combinaisons de sources de données, l'occupation du sol a été classée sur la zone d'étude selon une même procédure de classification orientée objet et de validation. Les classifications ont été validées à partir de matrices de confusion en fonction de leur capacité (1) à obtenir la typologie exigée (6 catégories principales d'occupation du sol cf. Chapitre 4 :I.1) et (2) à extraire trois tailles de végétation rivulaire arborée (largeur <15 m, entre 15 et 30 m et > 30 m sur une rive du cours d'eau). Les estimateurs issus de ces matrices sont présentés dans la partie III.7.1.

Pour analyser la qualité géométrique et sémantique des classifications, 50 parcelles de validation ont été choisies pour chaque classe et pour chaque taille de végétation arborée, et ont été numérisées manuellement par un photo-interprète indépendant se basant sur les orthophotos. Cette numérisation a été effectuée à l'échelle du 1 : 1000. Les traitements, les résultats obtenus et leur analyse sont détaillés dans l'***ANNEXE VII.2***.

Le Tableau 5-2 présente les estimateurs globaux (précision globale et indice de Kappa) obtenus pour chacune de ces combinaisons de sources de données.

La combinaison C3 (Spot5 XS + orthophotos + les données auxiliaires) permet d'obtenir les meilleurs résultats de classification comparée aux autres combinaisons (un gain de 10 % environ par rapport à C1 et de 15 % à C2, quelque soit l'estimateur global).

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 5-2 : Estimateurs globaux de la qualité des classifications de l'occupation du sol dans les corridors rivulaires selon les trois combinaisons (C1, C2, C3) de sources de données différentes (DA = Données Auxiliaires).

Données en entrées	Qualité du résultat de classification	
	Précision totale	Indice de Kappa
C1 : Spot 5 XS + DA	81.75 %	75.48 %
C2 : Spot 5 XS + Spot5 PAN + DA	77.08 %	69.25 %
C3 : Spot 5 XS + orthophotos + DA	92.82 %	90.46 %

Le Tableau 5-3 présente la « précision pour l'utilisateur » (pour une classe résultant de la classification, le pourcentage de pixels de cette classe qui sont bien classés vis-vis de la référence) obtenue pour chacune des catégories d'occupation du sol pour les différentes combinaisons de sources de données.

Quelque soit la catégorie d'occupation du sol, la combinaison C3 permet d'obtenir les meilleurs résultats de classification comparée aux autres combinaisons.

Avec cette combinaison de sources de données, les objets classés en « sols artificialisés », « cultures », « zones de végétation arborée », « zones de végétation herbacée » et « surfaces en eau » ont la plus forte fiabilité (plus de 85%). Par contre, les objets classés « sols nus naturels » possèdent une fiabilité plus faible (légèrement inférieure à 70 %, c'est-à-dire que 30% des pixels classés « sols nus naturels » sont en réalité d'une autre nature). Cette classe se confond fortement notamment avec la classe « sols artificialisés » car elle possède un comportement spectral similaire (cf. ***ANNEXE VII.2***).

Tableau 5-3 : Précision pour l'utilisateur par catégorie d'occupation du sol selon les trois combinaisons (C1, C2, C3) de sources de données différentes (DA = Données Auxiliaires).

	C1 : Spot 5 XS + DA	C2 : Spot 5 XS + Spot5 PAN + DA	C3 : Spot 5 XS + orthophotos + DA
sols artificialisés	63,28%	75,80%	95,48%
Cultures	90,22%	88,45%	99,23%
zones de végétation arborée	73,18%	61,49%	93,48%
zones de végétation herbacée et arbustives	70,29%	53,74%	85,76%
sols nus naturels	63,27%	61,54%	67,84%
surfaces en eau	91,24%	93,92%	94,70%

Le Tableau 5-4 présente pour les trois combinaisons de sources de données le pourcentage de végétation arborée rivulaire correctement détecté par la classification en « zone de végétation arborée » pour trois largeurs de végétation.

Pour la végétation arborée rivulaire de largeur supérieure à 15 m, on retrouve des pourcentages de détection similaires quelque soit la combinaison de sources de données utilisées. Par contre, la végétation arborée de largeur inférieure à 15 m est beaucoup mieux détectée en utilisant la combinaison C3 qu'en utilisant les autres combinaisons (un gain d'environ 50 % comparé à C1 et de 10 % comparé à C2).

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 5-4 : Pourcentage de végétation arborée rivulaire correctement détecté pour trois tailles de largeur de végétation selon on les trois combinaisons (C1, C2, C3) de sources de données différentes (DA = Données Auxiliaires).

	C1 : Spot 5 XS + DA	C2 : Spot 5 XS + Spot5 PAN + DA	C3 : Spot 5 XS + orthophotos + DA
< 15 m	23,09%	61,83%	73,10%
entre 15 et 30 m	75,70%	83,19%	82,61%
> 30 m	86,11%	80,24%	83,63%

III.1.4. Conclusion de l'étape 1

Suite à ces analyses, nous avons donc opté pour la combinaison de sources de données C3 (Figure 5.7) pour classer l'occupation du sol dans les corridors rivulaires.

La combinaison C1, utilisant uniquement la donnée image Spot5 XS (10 m), est limitée pour détecter les objets étroits et fragmentés le long des cours d'eau et notamment les bandes fines de végétation arborée rivulaire.

La combinaison C2 apporte en plus l'information Spot5 PAN (2.5 m), ce qui améliore cette détection mais ne permet pas de différencier efficacement les différentes catégories d'occupation du sol, notamment la classe « zones de végétation arborée » de la classe « zone de végétation herbacée et arbustive ». Ces deux classes se confondent car l'utilisation de l'information Spot5 PAN dans la segmentation ne permet pas de séparer correctement les objets contigus appartenant à ces deux classes. Elle est limitée aussi pour différencier les vignes (manque d'information texturale) qui se confondent avec les sols nus.

Par contre, la combinaison C3 qui apporte en plus l'information des orthophotos (0.5 m), plus riche au niveau textural et spectral que Spot5 PAN, améliore la détection des objets étroits et fragmentés dans les corridors rivulaires et la différenciation des différentes catégories d'occupation du sol.

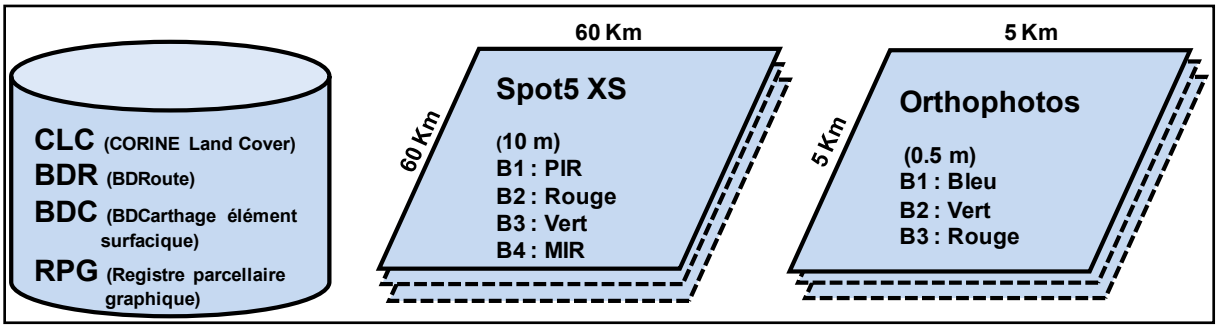
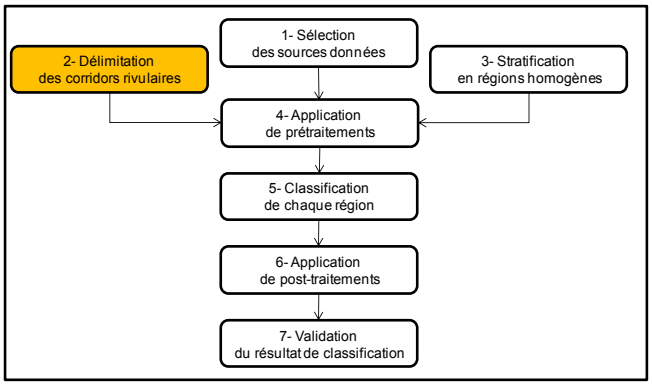


Figure 5.7 : Sources de données optimisées pour la classification de l'occupation dans les corridors rivulaires (étape 1)

CemOA : archive ouverte d'Irstea / Cemagref

III.2. Délimitation des corridors rivulaires (étape 2)



Dans cette étude, on s'intéresse à l'occupation du sol dans les corridors rivulaires donc à une partie seulement du territoire et de l'image. Il est par conséquent souhaitable d'appliquer un masque géographique à partir des limites de ces corridors afin de ne garder que l'espace que l'on veut traiter (et d'éliminer le reste de l'image), ceci pour éviter de traiter l'ensemble de l'image ce qui consommerait temps et ressources de calcul.

La largeur de corridor rivulaire à délimiter dépend naturellement de la largeur du lit majeur du cours d'eau. Les travaux sur la Loire (Souchon *et al.*, 2000) ont montré que les largeurs moyennes du lit majeur peuvent être estimées à partir des rangs de Strahler. On a donc estimé de manière expertisée la largeur de corridor rivulaire en fonction de cette ordination(Tableau 5-5). Ces largeurs ont été surestimées non seulement pour s'assurer d'extraire l'ensemble de l'information présente dans le corridor rivulaire mais aussi pour des besoins méthodologiques. La classification orientée objet nécessite en effet l'information du contexte du pixel et de l'objet : travailler sur un espace plus grand permet d'enrichir l'information contextuelle et d'optimiser la classification.

Tableau 5-5 : Largeur de corridor rivulaire prise en compte pour extraire l'information d'occupation du sol dans ces espaces.

Rangs	1, 2, 3	4, 5	6	7	8
Largeur de corridor rivulaire	400	800	1200	2400	4800

Une fois la largeur fixée, les corridors ont été délimités en construisant des buffers à partir du réseau hydrographique simplifié issu de la BDCarthage® (Figure 5.8). Bien que la BDCarthage® (contenant une longueur totale de 500 000 km d'arcs de cours d'eau environ) permette d'identifier sans ambigüité chaque cours d'eau sur le territoire métropolitain, elle est limitée pour mener des analyses spatiales du réseau naturel (par exemple n'étant pas orientée, elle ne permet pas de suivre un parcours amont - aval). De telles analyses sont rendues encore plus complexes par la présence (Pella *et al.*, 2006) : (1) d'éléments hydrographiques (arcs) artificiels (canaux, réseau de drainage,…) ; (2) de nombreuses erreurs topologiques (discontinuité entre les arc et arcs mal orientés) ; et (3) de plusieurs drains principaux par zone hydrographique (sous bassins-versants).

Dès lors, des traitements ont été effectués sur la base de la BD Carthage® afin d'obtenir un réseau hydrographique simplifié dont les arcs sont parfaitement jointifs et correctement orientés. Deux niveaux de simplification ont été créés : un premier niveau qui contient uniquement les drains principaux de chaque zone hydrographique, soit environ 89 000 km d'arc naturels (Pella *et al.*, 2006) et un deuxième niveau plus riche que le premier qui contient environ 250 000 km d'arcs naturels.

Ce niveau 2 a été choisi comme la donnée hydrographique de référence dans cette étude.

CemOA : archive ouverte d'Irstea / Cemagref

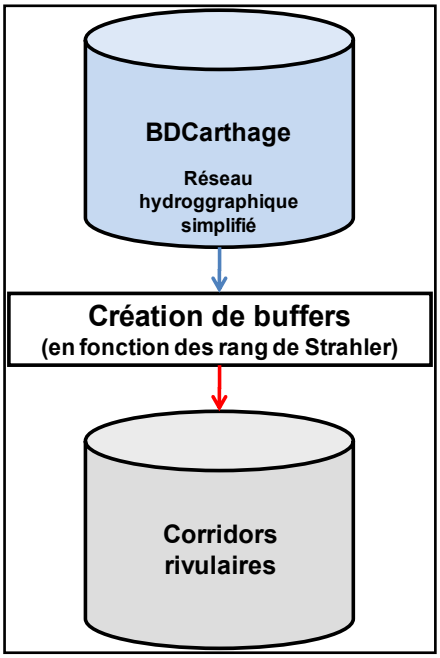
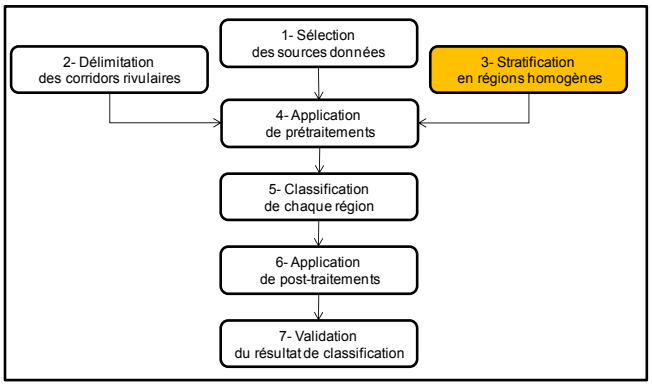


Figure 5.8 : Délimitation des corridors rivulaires (étape 2) en créant des buffers avec une largeur dépendant du rang de Strahler.

III.3. Stratification en régions homogènes (étape 3)



La France possède un contexte géographique très diversifié qui se traduit par la présence de différents types d'occupation du sol, tant naturels qu'anthropiques. En effet, les caractéristiques physiques naturelles des milieux (géologie, climat et relief) influent sur la composition et la configuration de l'occupation du sol dans le paysage. On ne trouvera pas, par exemple, la même végétation naturelle et les mêmes activités humaines (particulièrement agricoles) dans un milieu de montagne que dans un milieu de plaine, sur le versant océanique et sur le versant méditerranéen, sur un socle granitique et sur un socle calcaire.

Cette diversité de types d'occupation du sol augmente le nombre de classes à décrire et complexifie la définition des règles de classification. De plus, elle augmente les risques de confusions (on aura plus de risque de trouver des classes possédant un même comportement spectral et/ou spatial). Il est par conséquent intéressant de stratifier (segmenter, découper) la zone d'étude en régions de même type d'occupation du sol pour réduire le nombre de classes et améliorer les résultats de classification.

Les hydro-écorégions françaises de niveau 1 (HER1), qui sont délimitées en fonction des caractéristiques géologiques, climatiques et topographiques (Chapitre 1 :II.2.1), constituent une première base intéressante pour réaliser cette stratification. Lorsqu'une HER1 ne permet pas de séparer correctement les différents types d'occupation du sol, il est judicieux de travailler au niveau des HER2 (plus précis).

Cependant, pour recouvrir ces zones homogènes il faut le plus souvent acquérir un nombre d'images important, ce qui amplifie le risque d'avoir des dates d'acquisition différentes sur la zone. Cette situation diversifie également l'occupation du sol puisque l'état de certaines surfaces évolue dans le temps (elles

CemOA : archive ouverte d'Irstea / Cemagref

changent de classe). C'est le cas par exemple des parcelles cultivées : selon la date elles pourront être classées en « parcelles labourées», « parcelle en végétation » ou « parcelle à maturité » pour une même culture. Pour limiter cette diversification, on travaillera aussi dans des gammes de dates d'images identiques.

La stratification de la zone d'étude en régions homogènes (de même type d'occupation du sol dans une même gamme de dates) résulte alors du croisement entre les limites des HER (niveau 1 ou 2) et les limites des emprises des images de mêmes gammes de dates. Pour les orthophotos, on utilisera les limites départementales puisqu'elles sont acquises la même année par département (Figure 5.9).

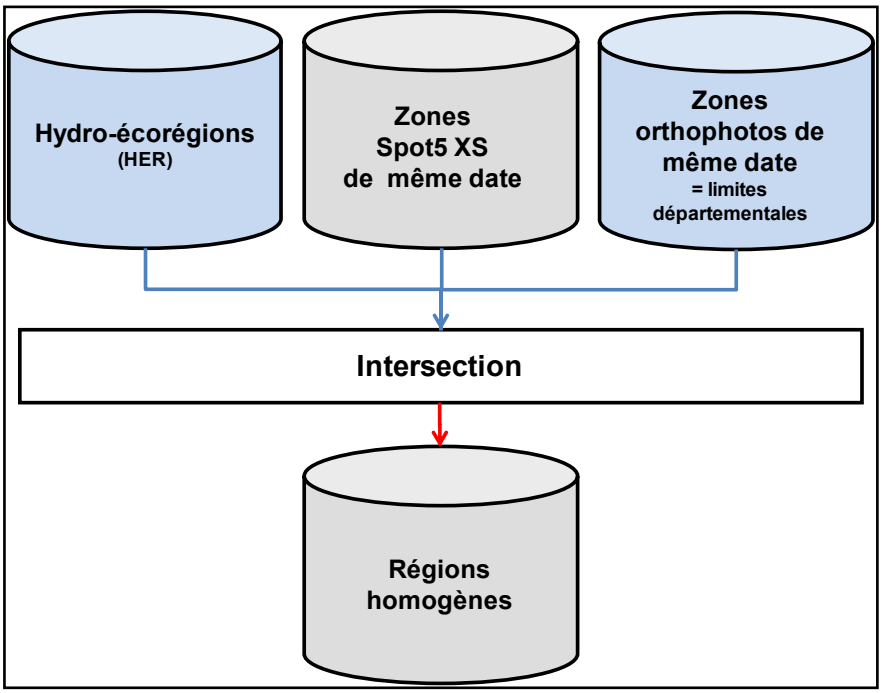
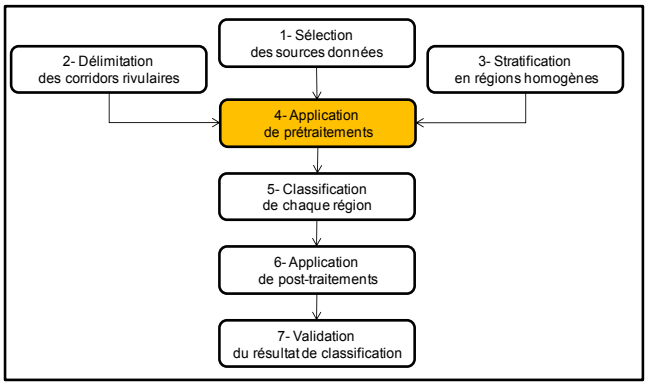


Figure 5.9 : Stratification de la zone d'étude en régions de types d'occupation du sol homogènes (étape 3) à partir des limites des Hydro-écorégions et des emprises d'images de même date.

III.4. Application de prétraitements (étape 4)



Plusieurs prétraitements sont nécessaires sur l'ensemble des données avant d'appliquer le processus de classification orientée objet (Figure 5.10). Ces prétraitements sont appliqués afin de découper les sources de données dans les différentes zones délimitées (corridors rivulaires et régions homogènes) et d'améliorer le contenu des sources de données images.

CemOA : archive ouverte d'Irstea / Cemagref

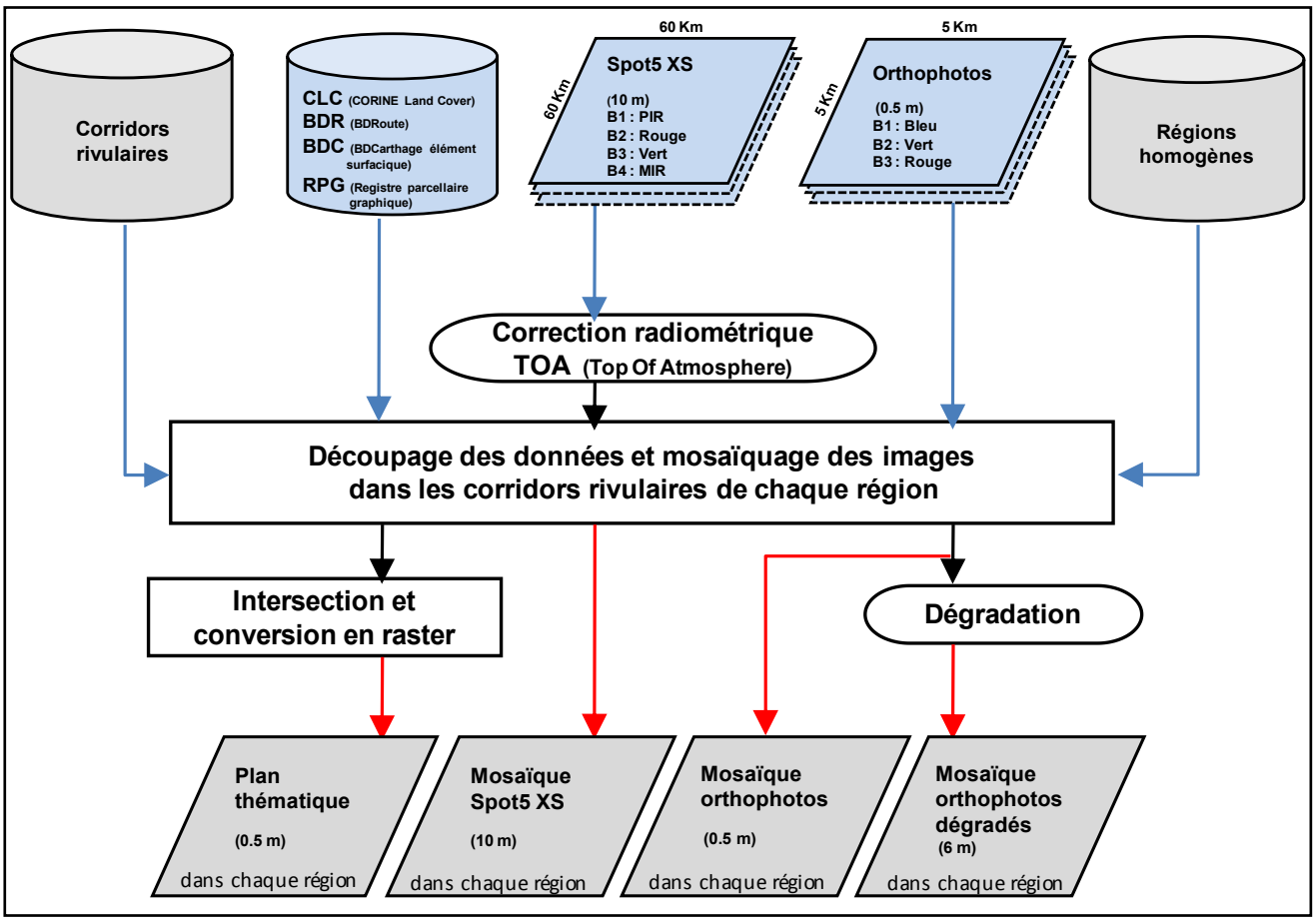


Figure 5.10 : Présentation des prétraitements appliqués (étape 4) sur les sources de données images et thématiques utilisées pour classer l'occupation du sol dans les corridors rivulaires.

III.4.1.La correction des images Spot5 XS

Les images Spot5 XS sont fournies en tant que lignes de comptes numériques qui sont simplement égalisés (corrigés pour le comportement individuel de chaque détecteur de pixel) et sans étalement dynamique numérique. Cependant pour une même cible au sol, le signal électromagnétique capté au niveau du pixel dans un domaine spectral donné dépend notamment:

(1) du gain des capteurs qui évolue en fonction du temps ;
(2) de la configuration angulaire de visée et d'éclairement (position du satellite et du soleil par rapport à la cible) ;
(3) de l'éclairement solaire.

Afin de limiter l'ajustement des règles de classification sur des images acquises à des dates différentes, ceci notamment pour les objets variant peu dans le temps, il est nécessaire de corriger ces perturbations en calculant les réflectances au sommet de l'atmosphère, *Top of Atmosphere* (les effets atmosphériques sont ignorés).

Pour obtenir des mesures physiques indépendantes des caractéristiques du capteur, les comptes numériques doivent être tout d'abord convertis en luminance. Dans une bande spectrale k, la luminance d'un pixel L^k_{TOA} au niveau TOA est liée au compte numérique X^k par la relation suivante :

CemOA : archive ouverte d'Irstea / Cemagref

$$L_{TOA}^{k} = \frac{X^{k}}{A_{k}.G_{m}^{k}} \quad (1)$$

avec :

- A_k le coefficient absolu de calibration pour la bande k, estimé pour la date de l'acquisition d'image. Ce coefficient a été fourni par le CNES (Meygret, 2007) pour chaque image ; il tient compte de la dégradation du capteur avec le temps.

- G^k_m le gain analogique de l'amplificateur à bord pour la bande k (Meygret, 2007).

La luminance L^k_{TOA} de chaque pixel dans chaque bande doit ensuite être normalisée par le flux solaire incident au sommet de l'atmosphère afin d'obtenir une réflectance de surface au niveau TOA, ρ^k_{TOA}. Ceci est calculé par :

$$\rho_{TOA}^{k} = \frac{\pi.L_{TOA}^{k}}{E_{S}^{k}.\cos\theta_{S}.(d_{0}/d)^{2}} \quad (2)$$

avec :

- E^k_S l'éclairement solaire moyen exo-atmosphérique pour chaque bande k. Il peut être récupéré à partir des métadonnées fournies avec l'image ou à partir du code 6S de transfert radiatif (Vermote *et al.*, 1997)

- ϑ_S l'angle zénithal solaire provenant des données auxiliaires.

- d_0/d le rapport entre la distance Terre-Soleil à la date d'acquisition et la distance moyenne Terre-Soleil, calculé pour la date d'acquisition par une équation simple du code 6S (Vermote *et al.*, 1997).

L'angle zénithal solaire varie aussi en fonction de la topographie (pente de la surface). Cet effet n'est pas pris en compte car il nécessite d'avoir accès sur l'ensemble de la France à un MNT aussi précis que les images Spot, ce dont nous ne disposons pas.

III.4.2. Le découpage et mosaïquage des données

Chaque donnée est découpée dans les corridors rivulaires et dans chacune des régions homogènes délimitées. Dans les régions recouvertes par plusieurs images (c'est toujours le cas avec les orthophotos), une mosaïque est réalisée. Compte tenu de la taille des données images (300 Mo pour une orthophoto par exemple) et du nombre d'images pour recouvrir ces régions (60 orthophotos environ sur une région homogène du bassin versant de l'Hérault) des algorithmes de traitements automatisés utilisant un cluster de calcul ont été développés.

III.4.3. La dégradation de l'orthophoto

Les textures sont déterminées par la constitution, l'agencement, la disposition des différentes parties ou éléments d'un objet. Ces textures sont le plus souvent calculées à partir des matrices de cooccurrence selon une direction donnée (Haralick *et al.*, 1973). Cependant ces textures dépendent fortement de la résolution spatiale de la donnée. Dans le cas de l'orthophoto, l'information texturale est parfois trop riche (hétérogène) pour définir une signature texturale spécifique notamment pour discriminer les objets du tissu urbain. Pour réduire cette hétérogénéité, il est nécessaire de dégrader cette information, i.e de ré-échantillonner l'information à un niveau d'information plus global (résolution ou grain plus grossier) en agrégeant les pixels.

CemOA : archive ouverte d'Irstea / Cemagref

Plusieurs niveaux d'agrégation, de résolutions spatiales différentes, ont été créés afin d'identifier (visuellement) celui qui permettait d'obtenir des signatures texturales discriminantes.

Après analyse, une texture calculée sur une dégradation de l'orthophoto à un pixel de 6 m améliore la discrimination des objets du tissu urbain.

Par conséquent, en plus de la mosaïque d'orthophoto d'origine (0.5 m), une mosaïque d'orthophoto dégradée à 6 m est rajoutée en entrée du traitement de classification orientée objet pour chaque région homogène.

III.4.4.La création d'un plan thématique

Afin de faciliter l'intégration des données thématiques dans le processus de classification, les différentes couches vecteur sont unies dans un plan thématique. Ces couches sont superposées en fonction de leur précision géométrique. Lorsque deux ou plusieurs couches se croisent, les limites et le type d'occupation de la couche la plus précise sont conservés (Figure 5.11).

La couche résultante est ensuite rastérisée à une résolution spatiale de 0.5 m (ce passage en raster facilite l'intégration des données dans le logiciel utilisé pour le traitement orienté objet). Chaque pixel de ce plan correspond à un code (une classe) thématique.

Les éléments géographiques linéaires de la Route 500® sont convertis en polygones, avant la création du plan thématique, en créant un buffer. La largeur du buffer dépend du nombre de voies constituant la route (par exemple un buffer de 8 m a été créé pour une route départementale à deux voies).

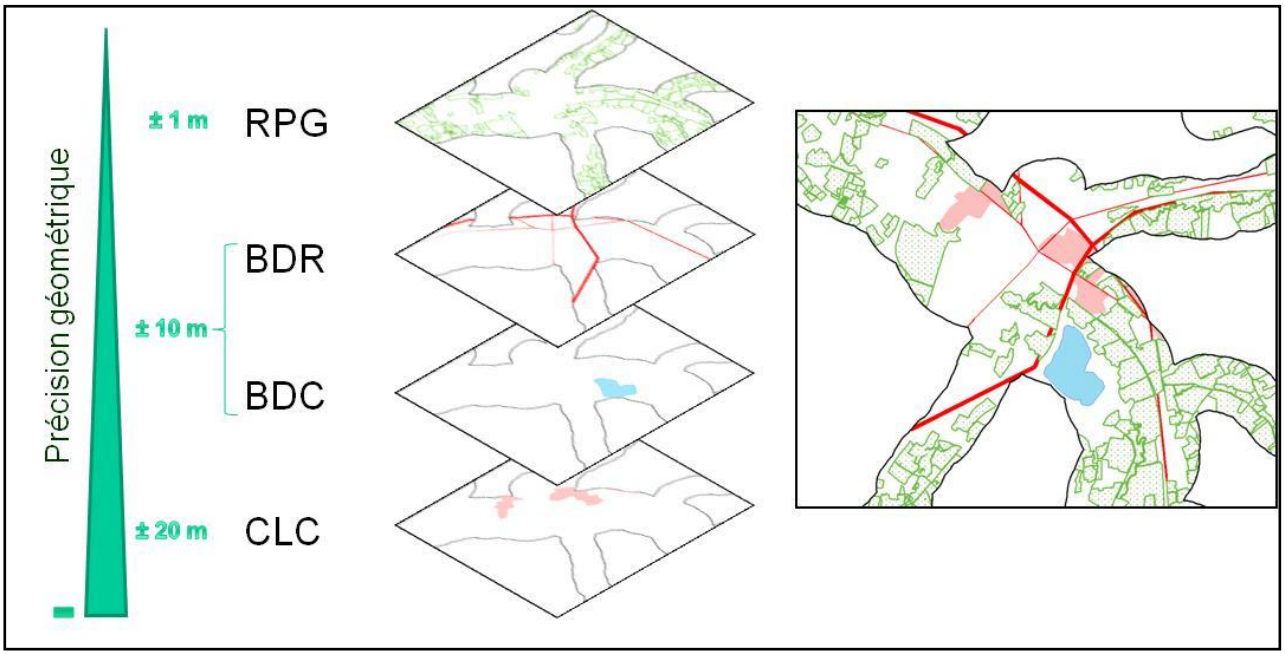
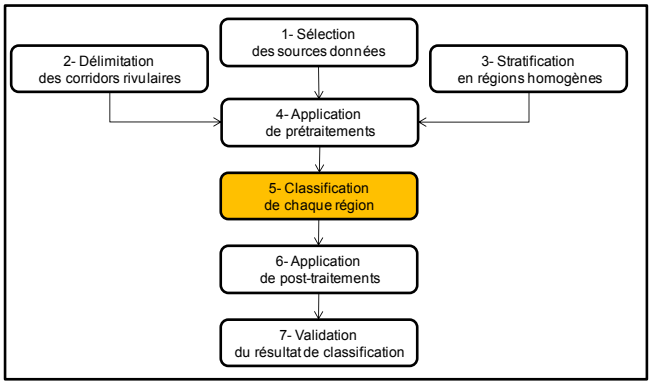


Figure 5.11 : Construction du plan thématique en fonction de la précision de chaque source de données thématiques (CLC=CORINE Land Cover, BDC=BDCarthage®, BDR=Route 500®, RPG=Registre Parcellaire Graphique).

CemOA : archive ouverte d'Irstea / Cemagref

III.5. Classification dans chaque région (étape 5)



Les approches de classification par pixel, strictement spectrales, sont obsolètes pour détecter et extraire les objets des images à THRS (cf. II.2). En effet elles ne tirent pas parti de l'information spatiale (contexte et forme de l'objet à classer) contenue dans l'image, qui est une information clé utilisée par les photo-interprètes (Schneider and Steinwender, 1999). L'approche orientée objet multi-échelle utilise cette information pour améliorer l'interprétation de l'image. Après avoir rappelé les principes de cette approche et présenté les algorithmes de segmentation et de classification envisagés (cf. III.5.1) nous développons des travaux d'optimisation de la stratégie d'échantillonnage des parcelles d'apprentissage (cf. III.5.2) et nous exposons la démarche de classification orientée objet adoptée pour classer l'occupation du sol dans les corridors rivulaires dans chaque région homogène selon les objectifs fixés (cf. III.5.3).

III.5.1. L'approche orientée objet multi-échelle : principes, méthodes et outils

III.5.1.1 Principe des réseaux hiérarchiques de classes

L'approche orientée objet multi-échelle d'une image s'apparente à l'automatisation de l'interprétation visuelle d'une image pratiquée par le photo-interprète (Benz *et al.*, 2004; Devaux, 2005). La démarche intellectuelle inhérente à l'interprétation d'une image par l'œil humain, selon une thématique donnée, passe par un processus mental itératif (Provencher and Dubois, 2007).

Dans ce processus, on commence d'abord par identifier les objets d'intérêt dans l'image et par les décomposer en leurs différents attributs. C'est *la partie analytique* du processus. Ces attributs sont nombreux. Ils sont liés à la fois aux critères d'apparence de l'objet (couleur, texture), de morphologie de l'objet (forme, taille, hauteur), mais surtout à l'environnement (contexte) de l'objet qui peut être analysé à différents niveaux de perception de l'image (échelles) et décrit par ces mêmes attributs. Le contexte spatial est l'un des aspects majeurs expliquant l'aptitude et les facilités de l'homme à identifier des objets complexes dans l'image (Devaux, 2006).

Ensuite débute la *partie interprétative* du processus, où chaque objet d'intérêt de l'image est « interprété » en se basant sur la description de ses différents attributs. Ceci amène à émettre, au fur à mesure de la description, des hypothèses sur la nature des objets observés et se termine par une interprétation de cette nature et l'attribution d'une classe à l'objet. Ce processus se répète jusqu'à ce que l'identification de l'ensemble des objets d'intérêt de l'image soit réalisée.

Cette démarche visuelle est fortement liée à la connaissance du photo-interprète. Il adopte dans la plupart des cas une démarche descendante: il commence par décrire les éléments les plus simples du paysage à une échelle globale et affine ensuite chacun de ces éléments à une échelle plus locale (Girard and Girard, 1999).

L'approche orientée objet multi-échelle reproduit cette démarche. Elle permet d'extraire l'information attributaire nécessaire à différents niveaux de perception de l'image en créant un réseau hiérarchique d'objets (partie analytique de l'interprétation visuelle) et formalise la connaissance du photo-interprète (*partie interprétative* de l'interprétation visuelle de l'image).

CemOA : archive ouverte d'Irstea / Cemagref

Ce formalisme repose sur une construction ontologique[35] géographique qui est un moyen simple de représenter une connaissance complexe (Devaux, 2006). Cette représentation regroupe les différentes classes décrivant complètement les objets d'intérêt. Ces classes sont liées les unes aux autres par des relations hiérarchiques (hiérarchie de classes : une classe peut être une sous-classe d'une classe mère, cf. Figure 5.12), Elles sont décrites par des règles de classification qui traduisent la connaissance experte de l'observateur.

La hiérarchie de classes peut être vue comme un arbre de classification (de décision) lorsque la connaissance du photo-interprète peut être formalisée de manière descendante. La racine de l'arbre correspond à l'ensemble de l'information de l'image apportée en entrée du processus de classification, les nœuds (les segments intermédiaires) constituent les classes intermédiaires et les feuilles (les segments terminaux) constituent les classes recherchées. La racine et les différents segments sont reliés par des branches qui symbolisent les règles de classification permettant de différencier les différentes classes (intermédiaires ou finales). Tous les chemins menant de la racine aux feuilles n'ont pas forcément le même nombre de nœuds et de branches. L'arbre procède par division successives des classes. Un nœud peut donner alors deux ou plusieurs classes-enfants qui héritent des propriétés (des règles de classification) de leur classe-mère (Figure 5.12).

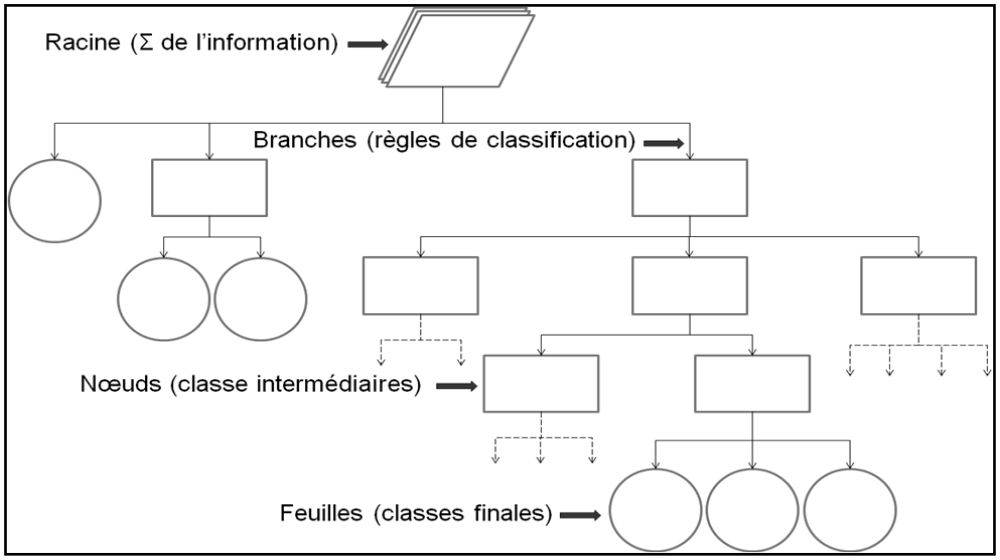


Figure 5.12 : Arbre de classification descendant. Les classes filles héritent des propriétés de classe mère.

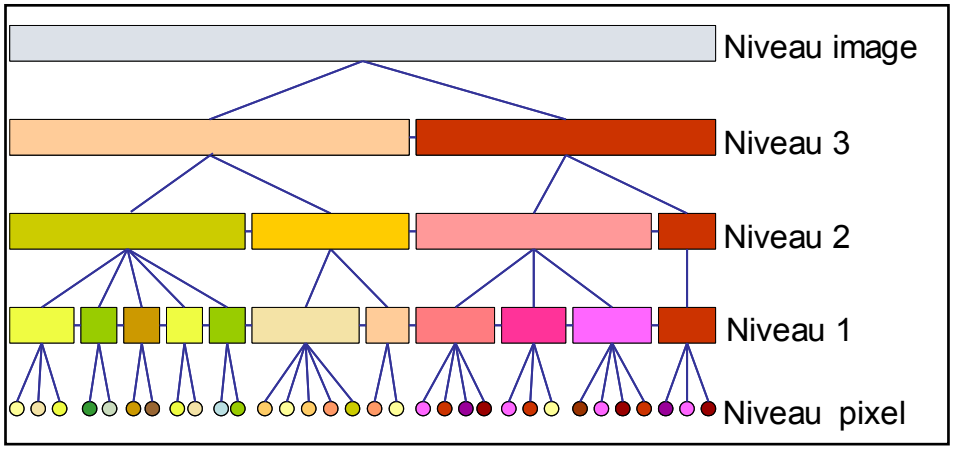


Figure 5.13 : Représentation abstraite d'une image d'après un réseau hiérarchique d'objets possédant trois niveaux de segmentation modifiés de Benz *et al.*, 2004

[35] Une ontologie est la spécification d'une conceptualisation d'un domaine de connaissance (Grüber, 1993)

CemOA : archive ouverte d'Irstea / Cemagref

Le formalisme de la hiérarchie de classe se base sur un réseau topologique d'objets (i.e. la frontière d'un super-objet est identique aux frontières de l'union de ses sous-objets, Figure 5.13). Dans ce contexte chaque objet est décrit par des attributs liés : (1) à son information radiométrique (critères spectraux et texturaux des pixels constitutifs de l'objet) ; (2) à sa géométrie (issue de calculs décrivant la forme de l'objet) ; (3) à son contexte au même niveau d'échelle (relations aux objets voisins), à l'échelle inférieure (relations aux sous-objets) et à l'échelle supérieure (relations aux super-objets) ; et (4) à sa hiérarchie (localisation de l'objet dans le réseau hiérarchique). L'appartenance d'un objet à une classe, définie dans la hiérarchie de classes, est décrite à partir de règles de classification utilisant ces différents attributs.

Par exemple, lorsqu'un photo-interprète doit caractériser les types de forêt (feuillus, conifères, mixte) dans un milieu naturel sur une image (Figure 5.14-A), il identifie tout d'abord ces objets dans leur ensemble. Il commence par distinguer le milieu naturel dans l'image (analyse à l'échelle globale) par rapport aux autres grands groupes d'occupation du sol (champs cultivés et sols artificialisés par exemple) à la condition que les attributs du milieu naturel soient suffisamment distincts de ceux de ces grands groupes. L'observateur se focalise ensuite sur le milieu naturel. Dans ce milieu, il détecte les massifs forestiers (analyse à l'échelle intermédiaire) par rapport aux autres couverts (végétation herbacée et sol naturel par exemple). Enfin, dans chacun de ces massifs, il détecte les différents types de forêts (analyse à l'échelle locale).

Si l'on applique une analyse orientée objet pour cette caractérisation (Figure 5.14-B), l'arbre de classification (hiérarchie de classe) mis en place se basera alors sur trois niveaux de segmentation de tailles d'objets différentes (grossier, moyen, fin) qui correspondent aux différentes échelles de perception du contexte spatial des objets à extraire utilisées par l'observateur (globale, intermédiaire, locale). Il décrit d'abord, au niveau grossier, les classes « milieu naturel » et « non milieu naturel ». Ensuite, au niveau moyen, la classe « milieu naturel » est subdivisée en deux classes « zone de forêt » et « non zone de forêt ». Enfin au niveau fin, la zone de forêt est subdivisée en trois classes de types de forêts : « feuillus », « conifères » et « mixtes ».

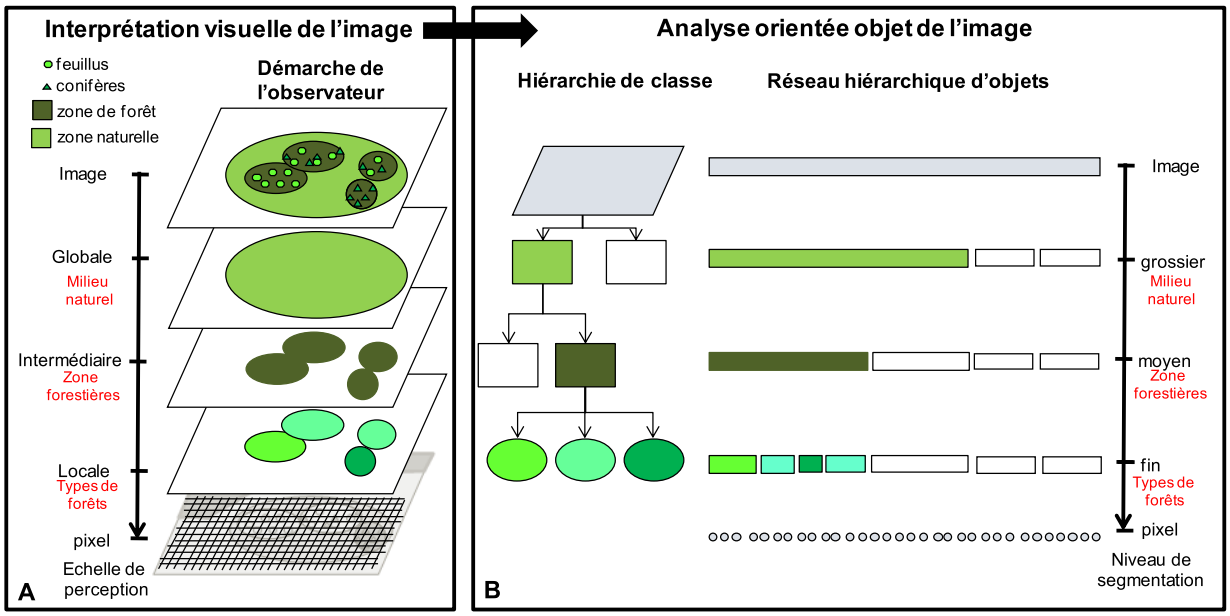


Figure 5.14 : Automatisation (ou formalisation de la connaissance) de l'interprétation visuelle d'une image (A) par l'analyse orientée objet (B). L'exemple présentée dans cette illustration concerne la classification des différents types de forêts (feuillus, conifères, mixtes).

Ainsi, grâce à sa capacité notamment à utiliser le contexte spatial à différents niveaux de perception, cette approche est une méthode très efficace pour extraire les objets d'intérêt dans une image.

Elle facilite aussi la transposition des connaissances accumulées à de nouvelles zones géographiques qui présentent des contextes similaires en minimisant les adaptations nécessaires (Devaux, 2006). En effet, les

CemOA : archive ouverte d'Irstea / Cemagref

classes des objets manipulés dans un contexte géographique similaire (urbain ou rural par exemple) sont sensiblement les mêmes d'une région géographique à l'autre. Par conséquent, le transfert des représentations des connaissances se trouve facilité par des modifications limitées aux seules spécificités des contextes locaux étudiés et non pas à l'ensemble de la connaissance.

III.5.1.2 Choix de la plate-forme logicielle

De nombreux logiciels (généralement commerciaux) proposent des fonctions de classification orientée objet utilisant des techniques de segmentation et de classification différentes (e.g., BerkeleyImgseg 0.54, ENVI Feature Extraction 5.4). Mais peu ont implémenté des solutions pour mener une approche orientée objet multi-échelles. La plupart ne permettent pas (ou difficilement) l'exploitation de l'information entre les différents niveaux de segmentation créés.

Les récents développements sur l'analyse orientée objet, basée sur l'approche « Fractal Net Evolution » (FNEA, Blaschke and Hay, 2001), fournissent les techniques de segmentation et de classification nécessaires pour extraire les objets d'intérêt en créant un réseau hiérarchique d'objets (analyse multi-échelle de l'image) et en affectant à chaque objet de ce réseau une classe à partir de règles basées sur la logique floue.

Cette approche, qui a révolutionné le traitement des images à haute et à très haute résolution spatiale, donne des résultats proches de la qualité des classifications par photo-interprétation, tout en étant plus rapide, moins cher et plus reproductible (Benz *et al.*, 2004; Durieux *et al.*, 2007 ; Aplin and Smith, 2008). Elle est implémentée dans le logiciel commercial Definiens Developper 7.0® (auparavant appelé eCognition et encore connu sous ce nom).

Nous avons choisi dans cette thèse d'utiliser ce logiciel pour développer la méthode de classification orientée objet de l'occupation du sol dans les corridors rivulaires pour quatre raisons principales:

- il permet d'implémenter une chaîne de traitement automatisable selon une analyse orientée objet de l'image multi-échelles ;

- il permet de définir des règles de classification basées sur la logique floue. Cette logique est intéressante puisque (1) elle facilite la gestion de l'information hétérogène (dans l'espace et dans le temps) des données multi-sources dans la définition des règles de classification (El hajj, 2008), (2) elle permet de prendre en compte l'incertitude de ces données, l'approximation du raisonnement mis en œuvre et le chevauchement entre les classes (Benz *et al.*, 2004) ; et (3) elle associe des facteurs de confiance (appartenance) aux objets classés : alors qu'une classification standard attribuerait à l'objet uniquement la classe de valeur d'appartenance la plus forte, la logique floue permet de décrire la stabilité de la classification, sa fiabilité et la manière dont la probabilité d'appartenance à différentes classes se combine pour un objet (la façon dont les classes « se mélangent » au sein de cet objet).

- Il fournit une série complète d'attributs pour chaque objet et offre la possibilité d'en calculer de nouveaux. Les différents types d'attributs sont présentés en ***ANNEXE VI***.

- Il propose enfin des solutions pour distribuer les différents traitements sur un serveur de calcul (cluster) ce qui permet de traiter des données volumineuses plus rapidement.

III.5.1.3 Méthodes et algorithmes de segmentation

La base des traitements Definiens Developper 7.0® est la segmentation multi-résolution, une technique brevetée pour l'extraction des objets d'une image. Elle a été développée pour extraire, à un haut niveau de qualité et à différentes résolutions selon le choix de l'utilisateur (structures fines ou grossières), les objets d'une ou plusieurs images. Elle permet donc la construction d'un réseau hiérarchique d'objets (Figure 5.13). Cette technique est adaptée pour détecter les objets d'une image même dans le cas de données très texturées, comme les images radars, les images satellites à très haute résolution, ou les photos aériennes.

La segmentation multi-résolution est une technique ascendante de croissance par paires de régions qui débute par un (ou plusieurs) objet(s) constitué(s) d'un seul pixel de l'image. Les plus petits objets sont ensuite

CemOA : archive ouverte d'Irstea / Cemagref

fusionnés de façon itérative dans de plus grands. Le procédé fondamental d'optimisation qui guide ce processus de fusion par paires d'objets, consiste à réduire au minimum l'hétérogénéité pondérée intra-objets. A chaque itération, sont fusionnés les deux objets adjacents de l'image qui, une fois réunis ont la plus faible croissance d'hétérogénéité. Si la plus faible croissance d'hétérogénéité excède le seuil maximal de croissance d'hétérogénéité, le processus s'arrête. Une description complète de cet algorithme peut être trouvée dans Baatz and Schäpe, 2000, Benz *et al.*, 2004, et Definiens, 2008.

Pour paramétrer la segmentation multi-résolution, l'opérateur définit l'hétérogénéité f sur laquelle l'algorithme se basera pour tester la fusion par paires d'objets et le seuil $f_{max,}$ la valeur maximale d'hétérogénéité (paramètre d'échelle) au-delà de laquelle le processus de fusion est arrêté.

L'hétérogénéité f résulte d'un compromis (**poids *W***) entre des critères d'hétérogénéité liés d'une part à la couleur $h_{spectral}$ et d'autre part à la forme de l'objet h_{shape} :

$$\mathrm{f} = (1-w).h_{spectral} + w.h_{shape} \quad (1)$$

- le critère d'hétérogénéité spectrale $h_{spectral}$ est obtenu par la somme des écarts-type σ_t des valeurs spectrales de chaque bande t, pondérés par w_t :

$$\mathrm{h}_{spectral} = \sum w_t \sigma_t \quad (2)$$

- le critère d'hétérogénéité morphologique h_{shape} résulte aussi d'un compromis (**poids W_s**) entre deux critères : un critère de « lissé » h_{smooth} et un critère de « compacité » h_{cmpct} :

$$\mathrm{h}_{shape} = w_s h_{cmpct} + (1-w_s) h_{smooth} \quad (3)$$

  - le critère de « lissé » h_{smooth} est défini par le rapport entre la longueur de coté l et la longueur b du plus court côté donné par la fenêtre englobante de l'objet :

$$\mathrm{h}_{smooth} = \frac{l}{b} \quad (4)$$

  - le critère de compacité h_{cmpct} est défini par le rapport entre la longueur de la bordure l et la racine carrée du nombre de pixel n formant cet objet d'image.

$$\mathrm{h}_{cmpct} = \frac{l}{\sqrt{n}} \quad (4)$$

Pour définir l'hétérogénéité f, l'opérateur peut agir ainsi sur les différents poids qui interviennent dans sa formule (Figure 5.15):

- W : qui donne plus ou moins d'importance à la forme (contre la couleur). Ce critère est nommé paramètre de forme (**shape**) dans le logiciel.
- W_t : qui donne plus ou moins d'importance aux bandes spectrales
- W_s : qui donne plus ou moins d'importance à la compacité (contre le lissé) dans le critère de forme. Ce critère est nommé paramètre de compacité (**compactness**) dans le logiciel.

CemOA : archive ouverte d'Irstea / Cemagref

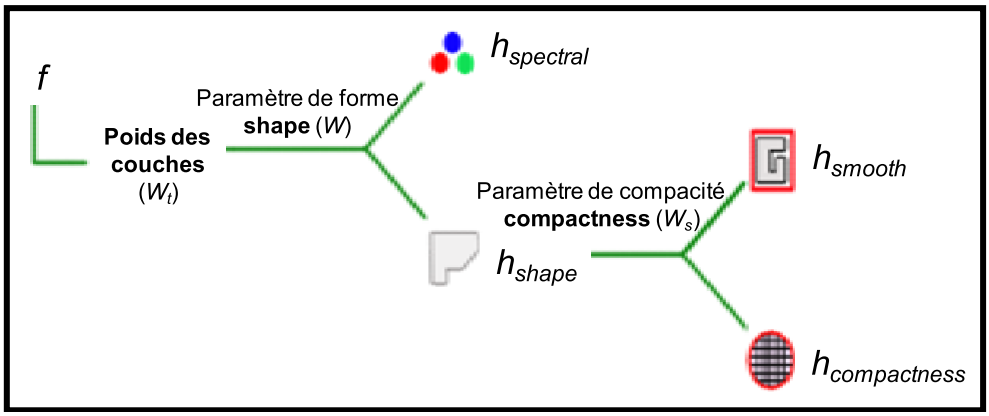


Figure 5.15 : Définition du paramètre d'hétérogénéité f de l'algorithme de segmentation multi-résolution

L'hétérogénéité maximale f_{max} est définie par le paramètre d'échelle (**scale**) dans le logiciel. Il dépend bien évidemment de la définition de l'hétérogénéité f. Ce paramètre d'échelle conditionne au final la taille des objets obtenus dans la segmentation.

La modification de la pondération respective du paramètre d'échelle (**scale**), de forme (**shape**) et de compacité (**compactness**) génère des résultats de segmentation différents (avec les mêmes poids de bandes en entrée), mais qui n'ont pas de relations linaires explicites avec les valeurs des paramètres. Par conséquent, la phase de segmentation reste une étape empirique où l'optimisation des résultats (formes et nombre des objets obtenus) s'obtient par itérations successives (Carleer *et al.*, 2005). C'est l'utilisateur qui estime visuellement la qualité des résultats, étant entendu qu'il recherche la meilleure adéquation entre les contours des objets-images (issus de la segmentation) et ceux des objets qu'il veut extraire.

A noter que cet algorithme peut se baser sur des données thématiques soit intégralement soit de façon complémentaire aux données images. Dans ce cas, les objets délimités intègrent obligatoirement les limites des objets décrits dans les données thématiques.

Ce logiciel propose aussi d'autres algorithmes de segmentation et notamment un algorithme qui permet, après une première segmentation puis classification, de refaçonner les objets dans un même niveau de segmentation en fusionnant les objets contigus appartenant à la même classe (algorithme de fusion).

III.5.1.4 Méthodes et algorithmes de classification

Après chaque segmentation, le logiciel calcule pour chacun des segments une série d'attributs (***ANNEXE VI***). Ces attributs sont utilisés pour construire les règles de classification. Une règle de classification utilise un seul critère (attribut) de l'objet, une classe peut être décrite par une combinaison de plusieurs règles de classification, et un critère peut être utilisé pour des classes différentes.

Deux techniques sont proposées dans Definiens Developper 7.0® pour définir ces règles:

(1) la ***technique des fonctions d'appartenance***, qui repose sur le développement par l'opérateur de règles générées de manière cognitive à partir de sa connaissance et de son observation de l'image (classification dite « expertisée ») ;

(2) la ***technique du plus proche voisin***, qui génère automatiquement les règles de classification sur la base d'un apprentissage à partir d'objets de référence (classification dites « supervisée »).

Ces deux algorithmes sont basés sur la logique floue qui permet de rendre compte des incertitudes inhérentes aux images de télédétection et des risques de confusion rencontrés lors de l'interprétation de l'image (par exemple, en géographie, la différenciation entre un habitat dispersé ou groupé est très subjective). Après avoir présenté synthétiquement les principes de la logique floue, nous présentons les deux techniques de classification.

CemOA : archive ouverte d'Irstea / Cemagref

III.5.1.4.1 Principes de la logique floue

La logique floue est une logique multicritère prenant en compte l'incertitude des connaissances. L'idée fondamentale est de remplacer les deux affirmations logiques booléennes "vrai" et "faux" par une gamme continue de valeurs [0… 1], où 0 signifie "faux" et 1 signifie "vrai" et où toutes les valeurs entre 0 et 1 représentent une transition du faux vers le vrai. En évitant de fixer des seuils arbitraires, la logique floue est capable de se rapprocher de la complexité du monde réel bien mieux que les systèmes de classification booléens. La logique floue peut tenir compte de l'incertitude de la connaissance humaine et mettre en application des règles linguistiques (Zadeh, 1965, Zadeh, 1978). Par conséquent, les systèmes de classification floue conviennent particulièrement bien à la gestion des sources d'imprécision lors de l'extraction des informations de télédétection (Benz *et al.*, 2004). Les systèmes flous se composent de trois notions principales, « flouification », base de connaissance floue et « déflouification », qui sont brièvement décrites dans la suite de cette partie.

La ***flouification*** décrit le passage d'un système booléen à un système flou. C'est la conversion des données numériques (par exemple la valeur d'un critère) en valeurs floues, sous la forme d'une fonction d'appartenance (règle de classification floue). Cette fonction d'appartenance décrit un ensemble flou, i.e. toutes les valeurs d'un attribut propres à une classe spécifique (Figure 5.16). En général, plus l'ensemble flou est large, plus l'ambigüité est forte sur la définition de la classe par ce critère.

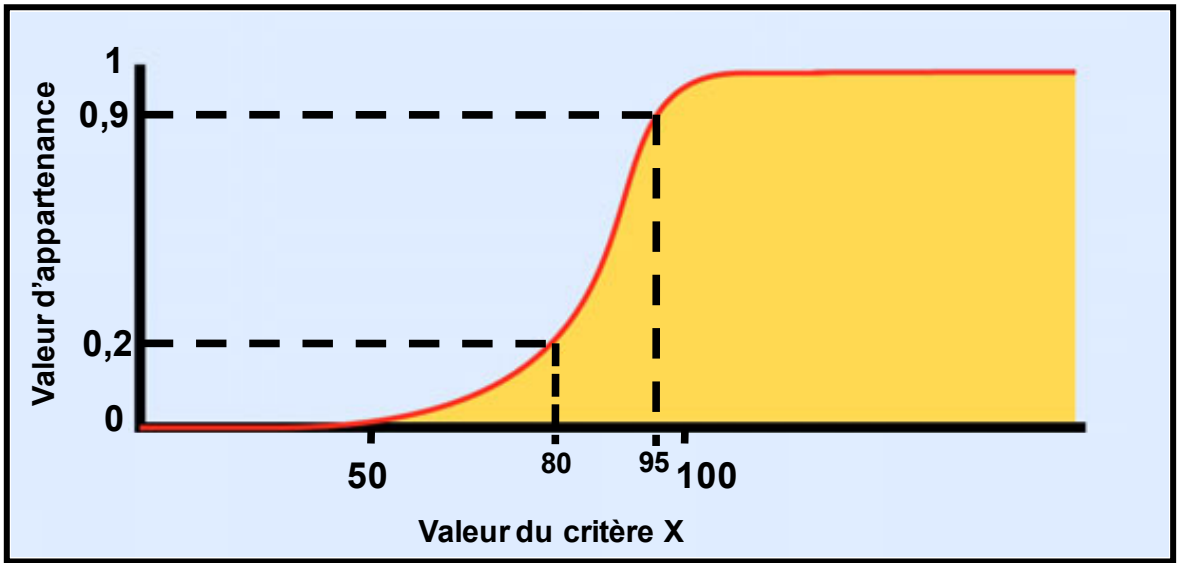


Figure 5.16 : Exemple de fonction d'appartenance sigmoïde pour décrire un ensemble flou : la valeur d'appartenance à une classe en fonction du critère X entre la valeur 50 et 100. Si la valeur X de l'objet = 95 alors sa valeur d'appartenance = 0.9 pour la classe. Par contre si X = 80 alors sa valeur d'appartenance = 0.2.

Des règles floues avancées peuvent être crées lorsque la classe est décrite par plusieurs critères. Chaque règle définie pour chacun des critères participe au processus de décision. Dès lors, une méthode d'inférence doit être définie, à l'aide d'opérateurs logiques, pour déterminer la manière dont les degrés de confiance de toutes les conclusions des règles sont agrégés afin d'attribuer aux objets une valeur d'appartenance à la classe.

La ***base de connaissance floue*** regroupe pour chacun des objets les valeurs d'appartenance discrètes pour chacune des classes définies. Elle se compose d'un n-uplet $f_{class,obj}$ de degrés d'appartenance μ, qui décrit les possibilités d'appartenance de l'objet obj aux n classes *class* considérées :

$$f_{class,obj} = \left(\mu_{class_1}(obj), \mu_{class_2}(obj), \ldots, \mu_{class_n}(obj) \right)$$

CemOA : archive ouverte d'Irstea / Cemagref

Ce n-uplet contient toutes les informations sur la stabilité de la classification, sa fiabilité. Cette propriété sera utilisée dans l'évaluation de la qualité du résultat (cf. III.7.2).

La ***déflouification*** est le processus inverse de la flouification. Pour produire des résultats tels que des cartes classiques pour l'occupation et l'utilisation des sols, les résultats flous doivent être traduits en valeurs booléennes pour assigner une classe unique à l'objet. Pour la déflouification des résultats de classification, la classe avec le degré d'appartenance le plus élevé est choisie (maximum du n-uplet).

A noter que si le degré d'appartenance d'une classe est inférieur à une certaine valeur (définie par l'opérateur), aucune classification n'est appliquée, ceci afin d'assurer un minimum de fiabilité.

III.5.1.4.2 La technique des fonctions d'appartenance (classification « expertisée »)

Cette technique permet à l'interprète de générer des fonctions d'appartenance (floues ou non floues) pour décrire les règles de classification. La définition de ces fonctions fait appel aux connaissances et à l'expérience de l'opérateur sur les comportements des objets étudiés. On parle alors de classification « expertisée ».

Ces fonctions utilisent bien sûr la couleur mais aussi la taille, la forme, le contexte des objets. Lorsque la classe est décrite par plusieurs critères, ces fonctions sont combinées à l'aide d'opérateurs logiques. Ils permettent ainsi la formulation de la connaissance et des concepts en créant une base de règles complexe.

Le logiciel propose plusieurs formes de fonctions prédéfinies, floues ou non floues. Il permet aussi la création de nouvelles fonctions (manuellement) et donne la possibilité de décrire les seuils non flous directement de manière linguistique.

Pour aider l'interprète dans la définition de ces règles, un outil de visualisation des attributs est intégré dans Definiens (« Feature View »). Il permet d'afficher les valeurs de l'attribut en niveau de gris, ou sur un dégradé du bleu (valeurs basses) au vert (valeurs élevées). A l'aide de cette visualisation l'interprète peut identifier plus facilement par photo-interprétation le (ou les) critère(s) discriminant(s) pour une classe et les valeurs seuils de la (ou des) fonction(s) floue(s) ou non floue(s).

La définition de la forme de la fonction et des valeurs seuils est la partie la plus délicate dans l'élaboration d'une règle de classification. C'est à l'opérateur de restituer et implémenter ses connaissances expertes associées à la problématique traitée. De ce fait, sa responsabilité dans les traitements augmente de plus en plus (Devaux, 2006).

III.5.1.4.3 La technique du plus proche voisin (classification « supervisée »)

Cette technique génère les règles de classification de façon automatique à partir d'un apprentissage. C'est une technique de classification « supervisée » qui, comme toute méthode supervisée, requiert des échantillons de référence (parcelles d'apprentissage) pour chaque classe ainsi qu'un espace d'attributs discriminant.

Cet espace peut être sélectionné directement par l'interprète (en fonction de sa connaissance) ou généré automatiquement par le logiciel à l'aide de l'outil « Feature Space Optimization ». A partir de l'apprentissage, cet outil permet d'extraire les critères qui construisent l'espace le plus discriminant pour la classification.

Dans cet espace, l'objet à classer appartiendra à la classe de la parcelle d'apprentissage la plus proche (en distance). C'est le principe de la classification par plus proche voisin (Figure 5.17).

La distance à un échantillon est traduite automatiquement par le logiciel en fonction d'appartenance de forme gaussienne. La classification par plus proche voisin renvoie ainsi une valeur d'appartenance à chaque classe entre 0 et 1. Plus l'objet à classer sera proche d'un échantillon d'une classe, plus sa valeur d'appartenance à la classe de l'échantillon sera forte (Figure 5.18).

CemOA : archive ouverte d'Irstea / Cemagref

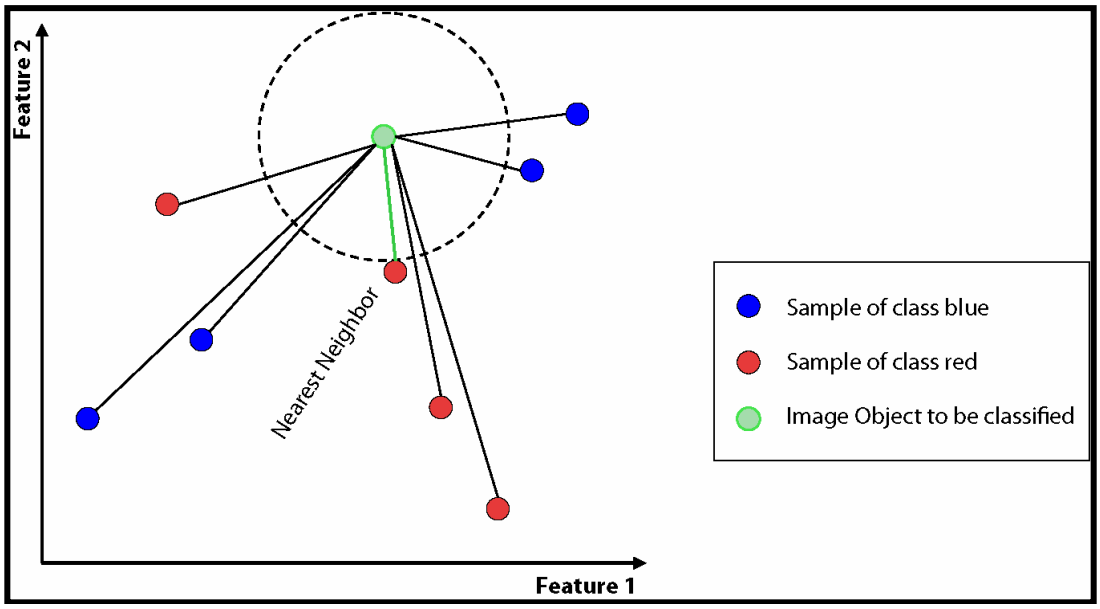


Figure 5.17 : Principe de la classification par plus proche voisin (Definiens, 2008)

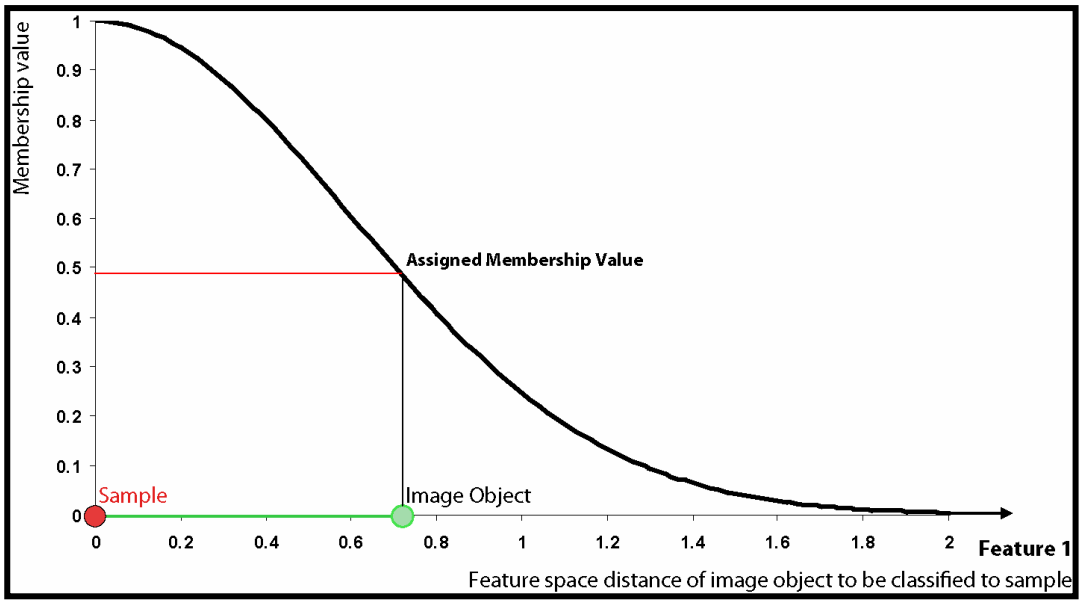


Figure 5.18 : Fonction d'appartenance générée automatiquement par le logiciel (Definiens, 2008)

La fonction d'appartenance effective pour une classe (à chaque point de l'espace d'attribut défini) est une combinaison (une fusion) des fonctions floues gaussiennes définies à partir de tous les échantillons d'une même classe. La Figure 5.19 montre cette combinaison pour un espace à une dimension (un critère), la Figure 5.20 pour un espace à deux dimensions.

L'utilisation de la technique du plus proche voisin (comme toute méthode supervisée) est recommandée lorsque la discrimination entre plusieurs classes de même niveau hiérarchique nécessite plusieurs attributs.

Les méthodes supervisées sont très souvent employées en classification orientée objet pour extraire l'information d'occupation du sol. Généralement, cette technique est appliquée sur un seul niveau de segmentation et définie sur un seul niveau hiérarchique de classes.

C'est une technique efficace qui demande moins de connaissance experte que la technique par fonctions d'appartenance pour définir les règles de classification. Cependant son efficacité dépend du mode d'échantillonnage (Girard and Girard, 1999) ce qui nous a amenés à développer un travail particulier d'optimisation présenté ci-dessous.

CemOA : archive ouverte d'Irstea / Cemagref

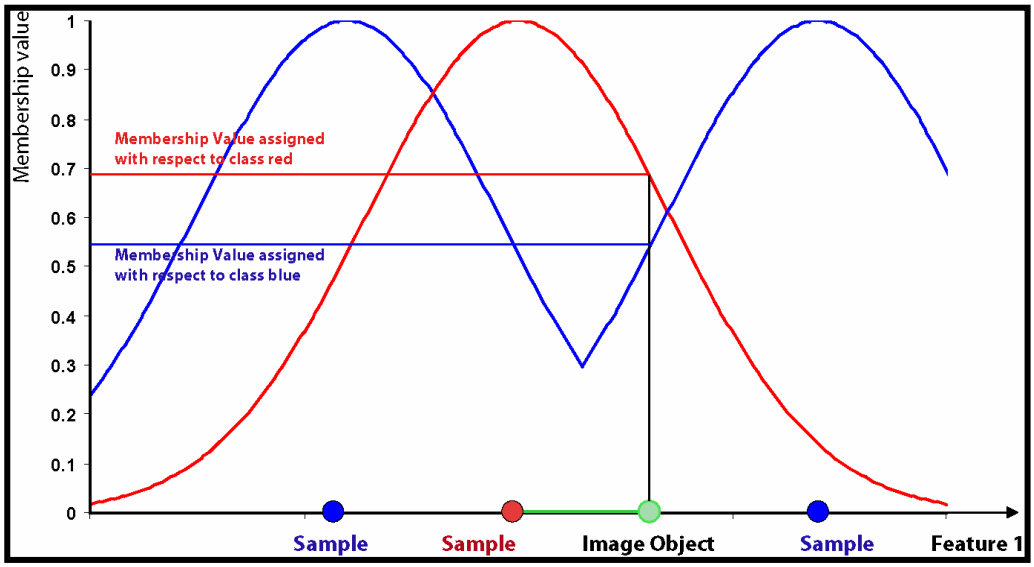


Figure 5.19 : Représentation de la combinaison des fonctions d'appartenance générée automatiquement par le logiciel dans un espace à une dimension (Definiens, 2008)

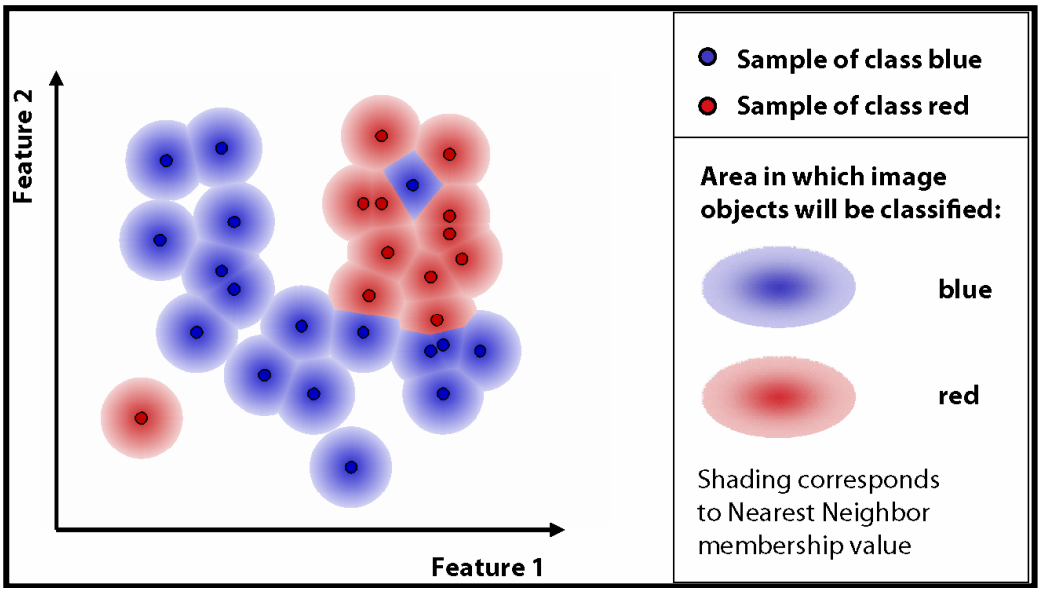


Figure 5.20 : Représentation de la combinaison des fonctions d'appartenance générée automatiquement par le logiciel dans un espace à deux dimensions (Definiens, 2008)

III.5.2.Optimisation du mode d'échantillonnage des parcelles d'apprentissage

L'objectif de la procédure d'échantillonnage des parcelles d'apprentissage est de choisir, pour chaque classe, des parcelles qui représenteront au mieux la variabilité interne de cette classe (variabilité intra-classe). Tout l'apprentissage reposera sur ces parcelles. Le mode d'échantillonnage des parcelles d'apprentissage constitue donc une étape clef de la technique de classification par plus proche voisin.

De façon générale, l'utilisateur dispose de peu d'informations sur la façon de déterminer le jeu d'apprentissage (comment le sélectionner ? quelle est sa taille optimale ?…). La plupart des études en télédétection utilisent un raisonnent en termes de nombre « x » de parcelles par classe, mais ce nombre « x » est le plus souvent choisi empiriquement et est rarement justifié (Congalton, 2004).

CemOA : archive ouverte d'Irstea / Cemagref

Lorsqu'on travaille sur de grands territoires (ce qui implique un très grand nombre d'objets), cette étape peut devenir très coûteuse en temps car elle implique d'identifier la classe de chacune de ces parcelles soit sur le terrain, soit par photo-interprétation.

Pour savoir si une technique de classification supervisée orientée objet peut être mise en œuvre de façon réaliste à cette échelle de travail, on s'est donc intéressé à définir la taille d'échantillonnage optimale permettant d'obtenir les meilleurs résultats de classification. Pour cela, deux analyses ont été menées sur des zones d'étude de tailles différentes.

La première analyse (Golden, 2006) a été réalisée sur un corridor rivulaire de l'Ain (12 km de cours d'eau) en utilisant l'information d'une orthophoto. 1846 objets ont été obtenus après une segmentation délimitant les différents couverts d'occupation du sol sur la zone (Perrez Correa, 2004). Pour chacun de ces objets une catégorie d'occupation du sol a été attribuée par photo-interprétation (selon une typologie de 8 classes) et 42 attributs ont été extraits. A partir de cette base de données d'objets, plusieurs techniques de classification supervisée (utilisant ou non des techniques d'optimisation de l'espace des variables) ont été testées pour identifier la taille d'échantillon optimale, en nombre « x » d'objets par classe sélectionnés au hasard.

Ces résultats ont été valorisés dans un article (communication) « Influence of training sampling protocol and of feature space optimization methods on supervised classification results » par S. Durrieu *et al.*, 2007 pour les IGARSS[36] de 2007. L'article est présenté dans son intégralité dans l'***ANNEXE VIII.1***. Les résultats montrent un optimum de 50 parcelles par classe quelque soit la technique de classification supervisée (Figure 5.21). On note par ailleurs que l'analyse discriminante linéaire (ALD) donne les meilleurs résultats de classification (80 % pour 50 parcelles/classe).

Comme cette étude a été menée sur un petit territoire, une deuxième analyse (basée sur les résultats et les méthodes de la première analyse) a été réalisée sur une zone d'étude plus importante (une région homogène du bassin versant de l'Hérault – 350 km de cours d'eau) en utilisant l'information des orthophotos et de Spot5 XS. La base de données objets est cette fois-ci constituée de 31 334 objets. Pour chacun de ces objets une catégorie d'occupation du sol a été obtenue lors de la mise en place de la technique de classification expertisée (selon une typologie de 9 classes) et 22 attributs ont été extraits (attributs utilisés par la technique expertisée). Cette analyse est présentée en détail à l'***ANNEXE VIII.2*** et comporte deux protocoles d'étude qui utilisent uniquement l'analyse discriminante linéaire (ADL), technique permettant d'obtenir les meilleurs résultats selon la première analyse. Pour limiter le biais lié aux erreurs de classification, les protocoles d'étude ont été itérés un grand nombre de fois, par tirage aléatoire des parcelles d'apprentissage.

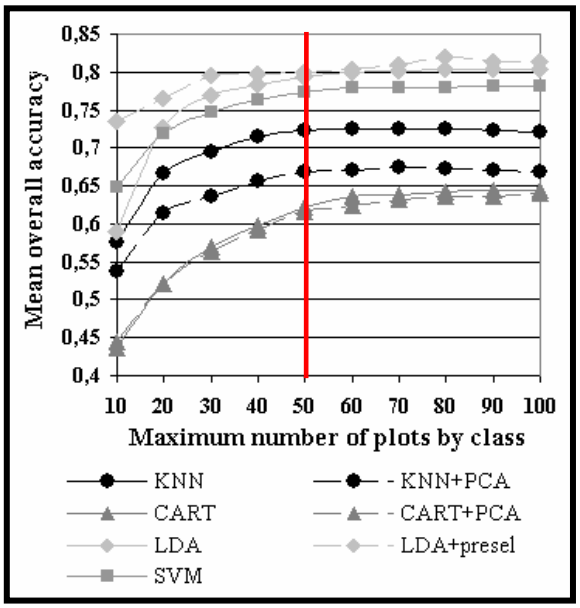


Figure 5.21 : Evolution des moyennes d'objets bien classés (précision globale) en fonction de l'augmentation de la taille de l'échantillon d'apprentissage (en nombre d'objets par classe) pour plusieurs techniques de classification supervisée (avec ou sans méthode d'optimisation de variable) appliquées sur un corridor rivulaire de l'Ain (12 km, 1846 objets) d'après Durrieu *et al.*, 2007.

Dans le premier protocole l'évolution des moyennes d'objets bien classés (précision globale) en fonction de l'augmentation de la taille de l'échantillon en nombre d'objets par classe a été analysée sur la moitié de la zone

[36] International Geoscience And Remote Sensing Symposium

CemOA : archive ouverte d'Irstea / Cemagref

(soit 15667 objets). On retrouve le même résultat (50 parcelles/classe) que dans la première analyse (Figure 5.22).

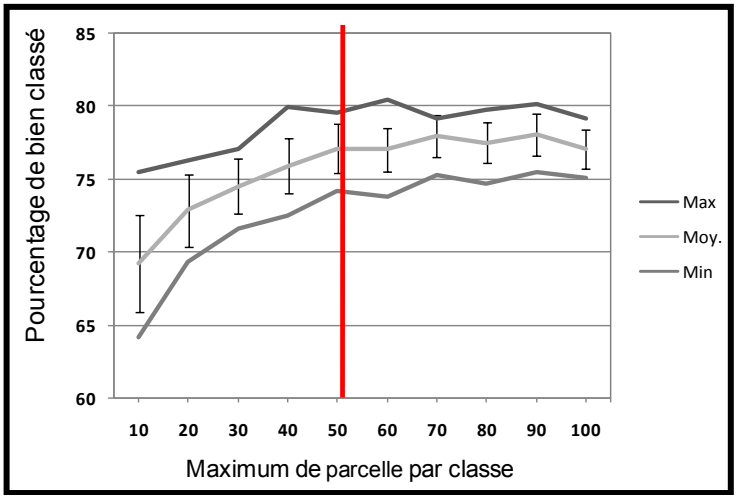


Figure 5.22 : Cette figure présente l'évolution des moyennes d'objets bien classés (précision globale) en fonction de l'augmentation de la taille de l'échantillon en nombre d'objets par classe pour l'analyse discriminante linéaire (sans méthode d'optimisation de variables) appliqué sur un corridor rivulaire de l'Hérault (15667 objets).

Dans le deuxième protocole nous avons analysé l'évolution des moyennes d'objets bien classés (précision globale) selon la taille de la zone d'étude (nombre d'objets), en utilisant une taille d'échantillon fixe de 50 parcelles par classe. La taille de la zone d'étude a été testée depuis 7500 objets jusqu'à 31 000 objets (zone complète). Les résultats montrent que quelque soit la taille de la zone d'étude, un jeu d'apprentissage de 50 parcelles par classe suffit pour entrainer un algorithme de classification supervisée dédié à la caractérisation de l'occupation du sol dans les corridors rivulaires sur une zone présentant un contexte géographique homogène et des types d'images de même date (Figure 5.23).

Au vu des résultats convergents de ces trois types de tests indépendants, ce nombre de parcelles d'apprentissage par classe nous semble un nombre raisonnable pour déployer cette technique sur de grands territoires.

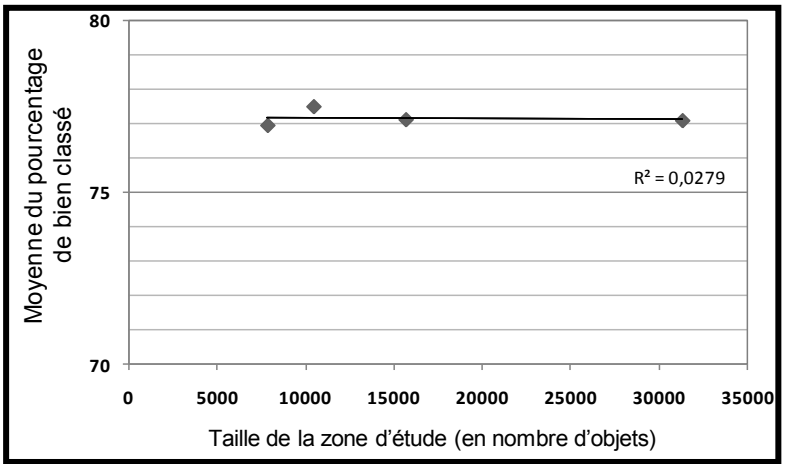


Figure 5.23 : Cette figure présente l'évolution des moyennes d'objets bien classés (précision globale) en utilisant l'analyse discriminante linéaire (sans méthode d'optimisation de variables) avec 50 parcelles/classe en fonction de l'augmentation de la taille de la zone d'étude (nombres d'objets).

III.5.3. Développement de la démarche de classification orientée objet

III.5.3.1 Choix du raisonnement de classification

La démarche de classification orientée objet doit être aussi « générique » que possible. Elle doit pouvoir être transposée (i.e. facilement adaptée) sur des régions de contextes géographiques différents (par exemple en HER[37] méditerranéenne comme en HER Tables calcaires) sans impact significatif sur son efficacité à extraire

[37] HER = Hydro-écorégion

CemOA : archive ouverte d'Irstea / Cemagref

l'information d'occupation des sols (i.e. de façon fiable, détaillée, et économique en termes de coûts de traitements).

L'enjeu est donc de développer une démarche :

- utilisant des règles de classification robustes et efficaces pour éviter d'en définir de nouvelles à chaque nouvelle région d'étude.
- assez souple pour permettre d'intégrer de nouvelles classes spécifiques à une nouvelle zone d'étude (ex. « sols nus rouges » en HER « Cévennes », « forêts de résineux » en HER « Landes ») sans modifications importantes du raisonnement ;
- compréhensible pour qu'elle puisse être assimilée simplement par un nouvel opérateur.

A la lumière de ces besoins, nous nous sommes orientés vers un raisonnement de classification fondé sur une hiérarchie sémantique de classes descendante (arbre de classification descendant).

Cette hiérarchie est construite pas à pas en décrivant, au fur et à mesure de l'interprétation du paysage, les classes les plus simples à discriminer (par exemple eau/terre ; puis au sein de la classe terre : sol nu/végétation ; puis au sein de la classe végétation : forêt/culture…).

Cette construction se fait tout en gardant à l'esprit l'objectif d'extraire au minimum les six catégories d'occupation du sol requises pour l'étude des relations pressions impacts (cf. Chapitre 4 :I.1). Elle s'arrête lorsque la définition de nouvelles classes n'est plus possible, i.e lorsqu'il n'est plus possible d'extraire efficacement de façon automatique de nouvelles classes. Elle dépend ainsi à la fois de l'expérience de l'opérateur et des sources de données utilisées.

Cette approche de classification descendante facilite la transposition des connaissances sur une nouvelle zone :

- Elle réduit fortement la complexité relative à l'extraction de certaines classes d'occupation du sol, ce qui favorise l'utilisation de règles simples et réutilisables.
- Elle se rapproche du raisonnement du photo-interprète (du plus facile au plus complexe), ce qui favorise son appropriation par un nouvel opérateur.
- Elle permet d'intégrer une classe d'occupation du sol spécifique à une région donnée en rajoutant simplement une nouvelle branche dans l'arbre de classification (exemple : pour la classe « vignes », une branche de classification supplémentaire peut être rajoutée à partir de la classe « terres agricoles nues »).

III.5.3.2 Une démarche en trois phases

Nous avons donc développé une démarche orientée objet, basée sur un arbre de classification descendant (cf. III.5.3.1), utilisant l'information du plan thématique (données auxiliaires), de l'orthophoto (0.5 m), de l'orthophoto dégradée (6 m) et de l'imagerie Spot5 XS (10m).

Dans un premier temps, on avait envisagé spontanément une démarche en deux phases : la première consistant à segmenter et classer l'image en fonction du plan thématique de manière à intégrer l'information connue sur certains objets de l'image, puis la seconde consistant à classer le reste de l'image sur un seul niveau de segmentation.

Cette démarche en deux phases, que nous avons testée, bute sur un obstacle : la forte hétérogénéité des tailles d'objets que les méthodes de segmentation ne permettent pas de gérer correctement en une seule fois (cf. III.5.1.3.). Les objets possèdent en effet des tailles différentes : des objets plutôt étroits et fragmentés, comme les bancs de sable le long des cours d'eau, la végétation arborée isolée ou encore le bâti diffus, et des objets plus gros comme le parcellaire agricole ou les tissus urbain denses.

La solution permettant de contourner cette difficulté, liée à la configuration du paysage, est de compléter le niveau calculé à partir du plan thématique par deux niveaux d'organisation emboîtés : le premier considérant des objets « micro » et le second des objets « macro ».

Ainsi, pour extraire efficacement l'occupation du sol dans ces espaces, nous avons développé une démarche en trois phases successives de segmentation-classification (Figure 5.24). Chacune des phases utilise le résultat de segmentation et de classification obtenu à la phase précédente. L'arbre de classification est construit progressivement au cours de ces trois phases.

CemOA : archive ouverte d'Irstea / Cemagref

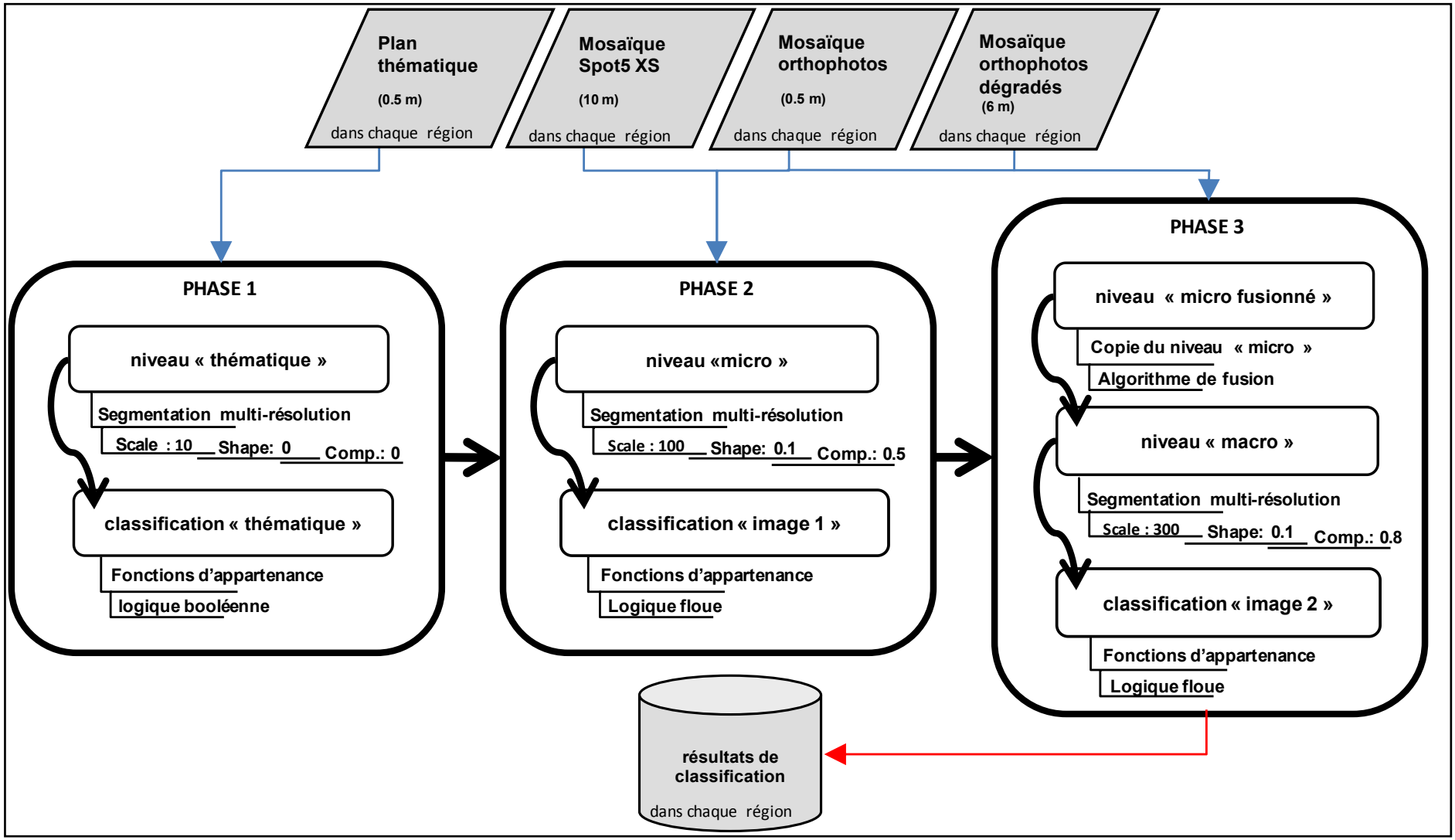


Figure 5.24 : Les 3 phases de classification orientée objet par région (étape 5)

CemOA : archive ouverte d'Irstea / Cemagref

La **phase 1** exploite l'information du plan thématique construit à partir des sources de données auxiliaires. Un niveau de segmentation (niveau « thématique » Figure 5.25-A) est créé selon les limites de ce plan en utilisant l'algorithme de segmentation multi-résolution. Les objets segmentés à ce niveau sont ensuite classés en fonction de l'information du plan (classification « thématique » Figure 5.25-B).

L'arbre de classification commence par définir deux classes intermédiaires : « thématique » (objets contenant une information dans le plan thématique) et « non thématique » (objet ne contenant pas d'information dans le plan thématique). La classe « thématique » est ensuite subdivisée en plusieurs classes finales correspondant aux différentes catégories fournies par cette couche d'information (Figure 5.25-C).

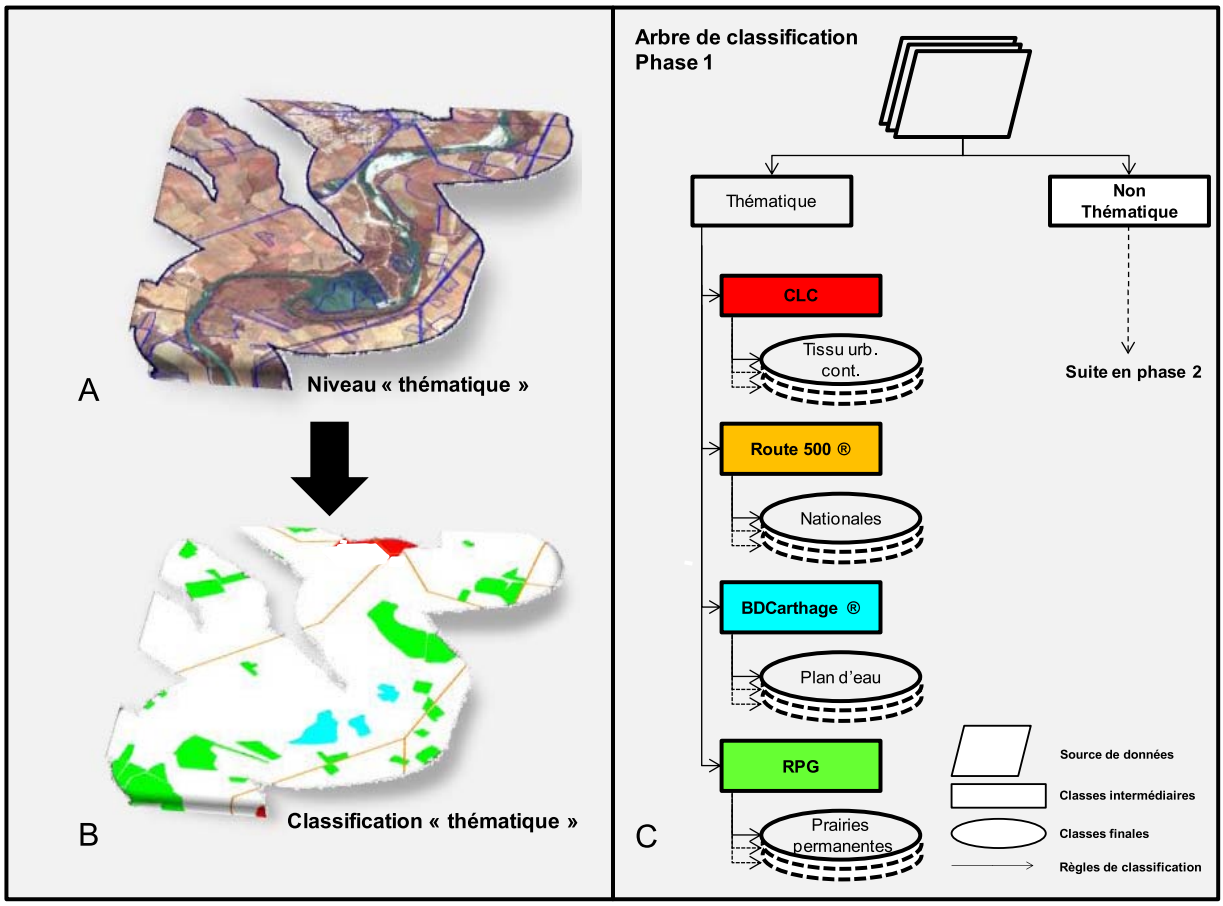


Figure 5.25 : Illustrations de la segmentation (A), de la classification (B) et présentation du développement de l'arbre de classification (C) à la phase 1

Ces classes sont décrites en utilisant des fonctions d'appartenances booléennes. On accorde ainsi un degré de confiance maximum (de 1) à l'information provenant des sources de données auxiliaires.

Cette première phase est spécifique, elle permet d'intégrer l'information thématique connue sur certains objets (par exemple les objets « routes » fournis par la BD routes) et de construire une première base solide et précise d'objets de classes connues, qui pourront être utilisés par la suite dans les processus de segmentation-classification des objets non connus (classés « non thématique »).

Par ailleurs, elle permet de réduire en partie le temps de traitement, puisque les processus de segmentation et de classification suivants seront appliqués uniquement sur les objets non connus.

La **phase 2** est consacrée à l'identification et la caractérisation des objets de petite taille présents dans le paysage rivulaire. Un nouveau niveau de segmentation, plus fin (niveau de segmentation « micro » Figure 5.26-A), fondé sur les limites du niveau « thématique », est créé en utilisant également l'algorithme de segmentation multi-résolution. Cette segmentation est réalisée uniquement sur les objets classés « non thématique ». Elle se base à la fois sur le canal PIR de Spot5 XS et de façon accrue (pondération deux fois plus grande du coefficient wt correspondant cf. III.5.1.3) sur l'information spectrale plus précise des orthophotos.

Cette segmentation est réalisée uniquement sur les objets classés « non thématique ».

CemOA : archive ouverte d'Irstea / Cemagref

Les objets fins segmentés sont classés à partir de l'information des sources de données images selon un raisonnement de classification descendant. L'arbre de classification subdivise la classe « non thématique » en deux classes : classe «cours d'eau » et la classe « terre » ; puis au sein de la classe terre : la classe « «sol sans ou avec peu de végétation» et la classe « sol avec végétation » ; ainsi de suite jusqu'à ce qu'il ne soit plus possible de séparer de façon fiable et simple de nouvelles classes à ce niveau de segmentation (Figure 5.26-C).

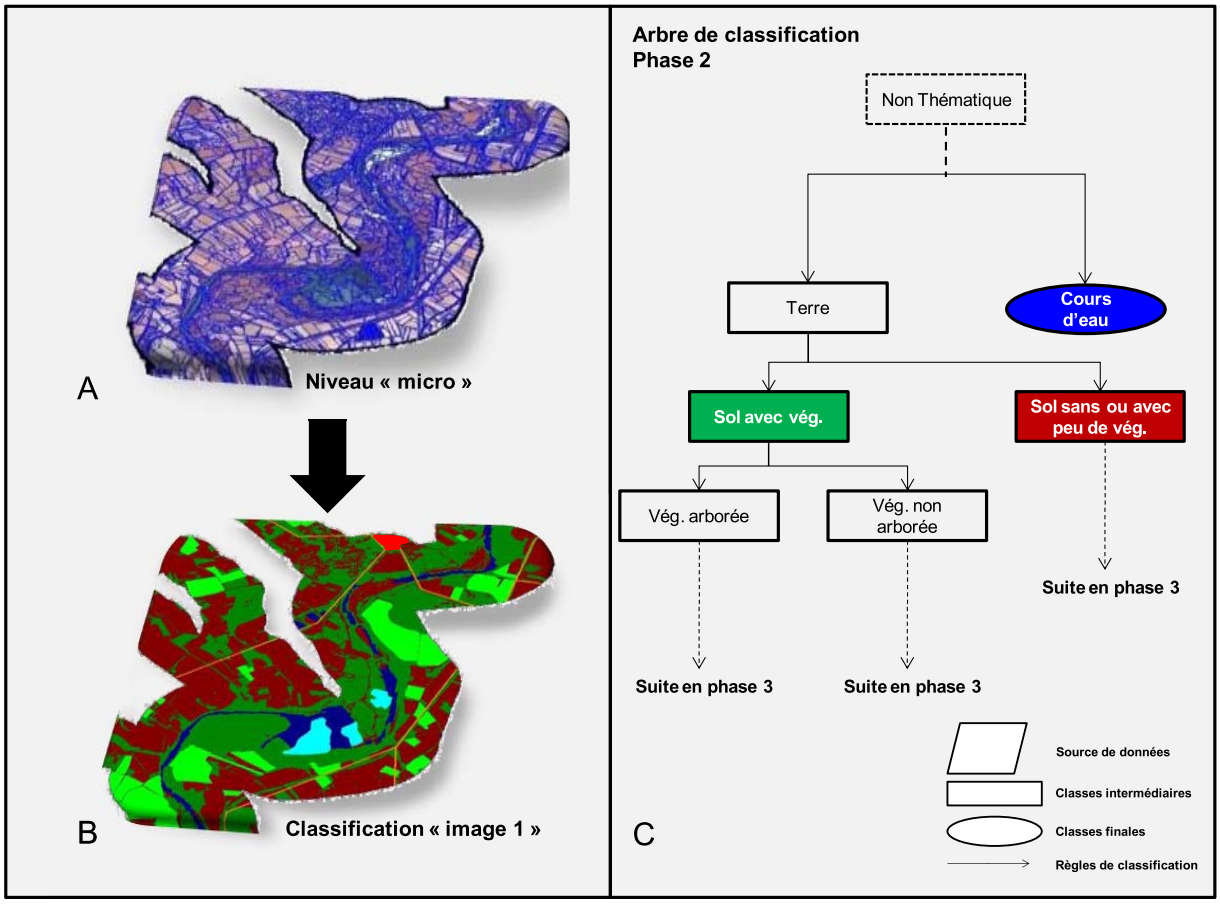


Figure 5.26 : Illustrations de la segmentation (A), de la classification (B) et présentation du développement de l'arbre de classification (C) à la phase 2

Généralement cette première classification (classification « image 1 » Figure 5.26-B) aboutit à quatre classes : une classe finale («cours d'eau») et trois classes intermédiaires («sol sans ou avec peu de végétation», «zone de végétation arborée» et «zone de végétation non arborée»).

Ces classes sont décrites à partir d'une seule ou bien d'une combinaison (à l'aide d'opérateur logique) de fonctions d'appartenances floues ou non floues basées sur des critères issus de l'information des sources de données images.

La définition d'une fonction d'appartenance est réalisée en quatre étapes (un exemple est présenté Figure 5.27) :

- étape 0 : l'opérateur a sélectionné en fonction de son expérience ou par photo-interprétation le (ou un des) critère(s) permettant de discriminer la classe ;
- étape 1 : il visualise tout d'abord ce critère à l'aide de l'outil « Feature View » fourni par le logiciel de traitement d'images;

CemOA : archive ouverte d'Irstea / Cemagref

- étape 2 : il cherche grâce à cet outil le seuil permettant de discriminer la classe. Lorsque qu'un seuil fixe ne peut pas être déterminé, il définit alors un intervalle autour de ce seuil correspondant à une marge d'ambiguité liée au phénomène étudié et à son incertitude ;

- étape 3 : il traduit ensuite cette recherche dans une fonction d'appartenance non floue si le seuil est fixe, et floue si la définition du seuil est ambigüe en se basant sur les bornes de l'intervalle identifié à l'étape 2 ;

- étape 4 : l'opérateur applique enfin la fonction et vérifie visuellement le résultat de la classification.

Lorsqu'on utilise un critère issu de l'information fournie par des sources de données images, il n'est pas toujours facile de définir un seuil strict (logique du « tout ou rien »).

On ne passe pas, par exemple, brutalement d'objets « eau » à objets « terre » en utilisant l'information du canal PIR. Il existe toute une gradualité de situations (eau plus ou moins profondes) entre ces deux classes d'objets correspondant à un intervalle de valeurs de PIR plus ou moins grands. La logique floue permet de prendre en compte cette spécificité et dépasse ainsi cette logique du « tout ou rien ». Elle permet aussi à l'opérateur d'introduire l'incertitude liée à sa connaissance du phénomène étudiée.

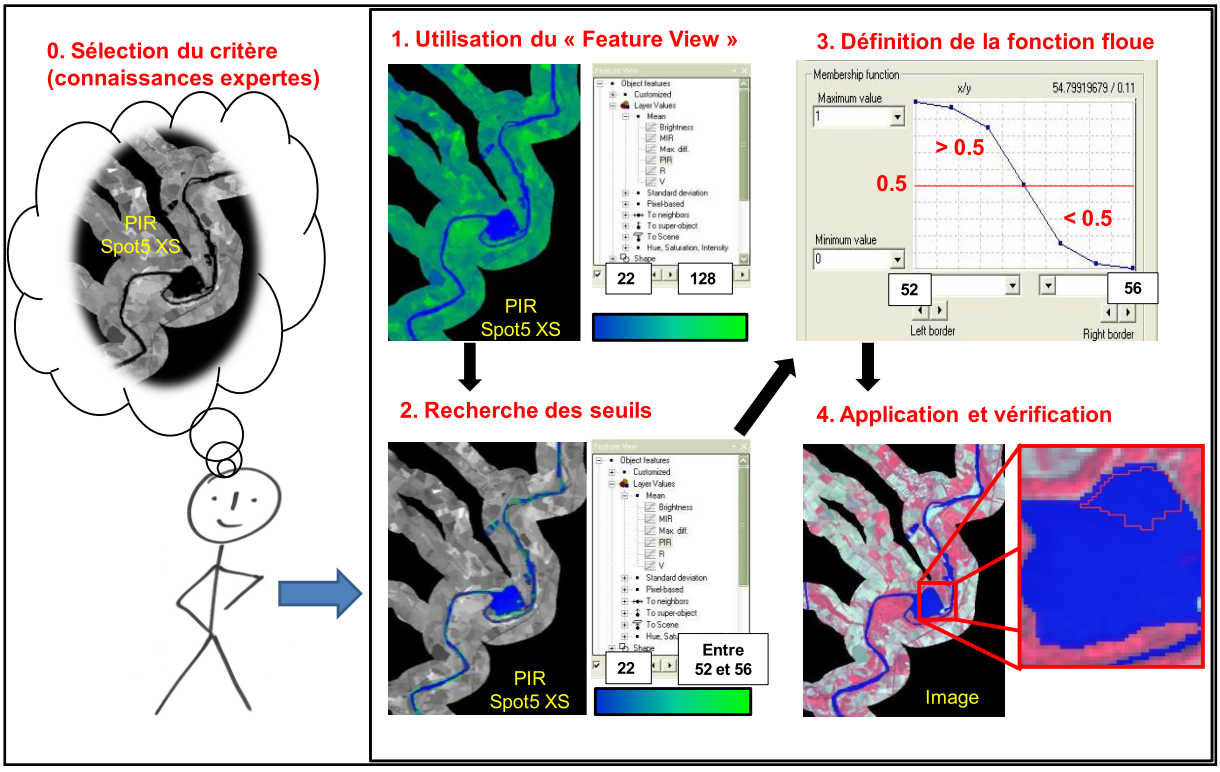


Figure 5.27 : exemple de la définition d'une fonction floue pour extraire les objets de la classe « cours d'eau » sur une région homogène donnée. Cette fonction utilise l'information du PIR du satellite Spot5 XS puisque l'eau possède une faible réflectance dans cette bande spectrale.

La **phase 3** continue à détailler les classes obtenues à la phase 2 en se basant sur un nouveau niveau de segmentation plus large (niveau « macro », Figure 5.29-A) qui permet à la fois de délimiter les objets de taille moyenne mais aussi de garder les objets isolés délimités au niveau « micro ». Ce niveau résulte d'une segmentation multi-résolution utilisant uniquement l'information de Spot5 XS en se basant sur les limites du niveau « micro » et d'un niveau « micro fusionné » (Figure 5.28).

CemOA : archive ouverte d'Irstea / Cemagref

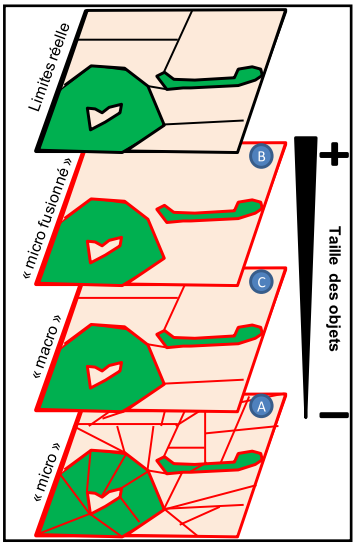


Figure 5.28 : création du niveau de segmentation « macro » (C) : ce niveau se base à la fois sur les objets fins créés au niveau « micro » (A) et le niveau « micro fusionné » (B) résultant d'une copie du niveau « micro » et d'une fusion des objets contigus de même classe. Ce système permet de garder les petits objets isolés dans la zone étudiée au niveau « macro ».

En utilisant toujours le même raisonnement de classification, les classes intermédiaires de la phase 2 sont détaillées dans la phase 3. Un exemple est présenté Figure 5.29-C pour la classe « sol sans ou avec peu de végétation ». A cette phase on aboutit à une classification finale (classification « image 2 », Figure 5.29-B), la plus détaillée possible en fonction de l'information des sources de données et des connaissances de l'opérateur.

Comme pour la phase 2, les classes sont décrites en utilisant des fonctions d'appartenances floues ou non floues basées sur des critères issus de l'information des sources de données images.

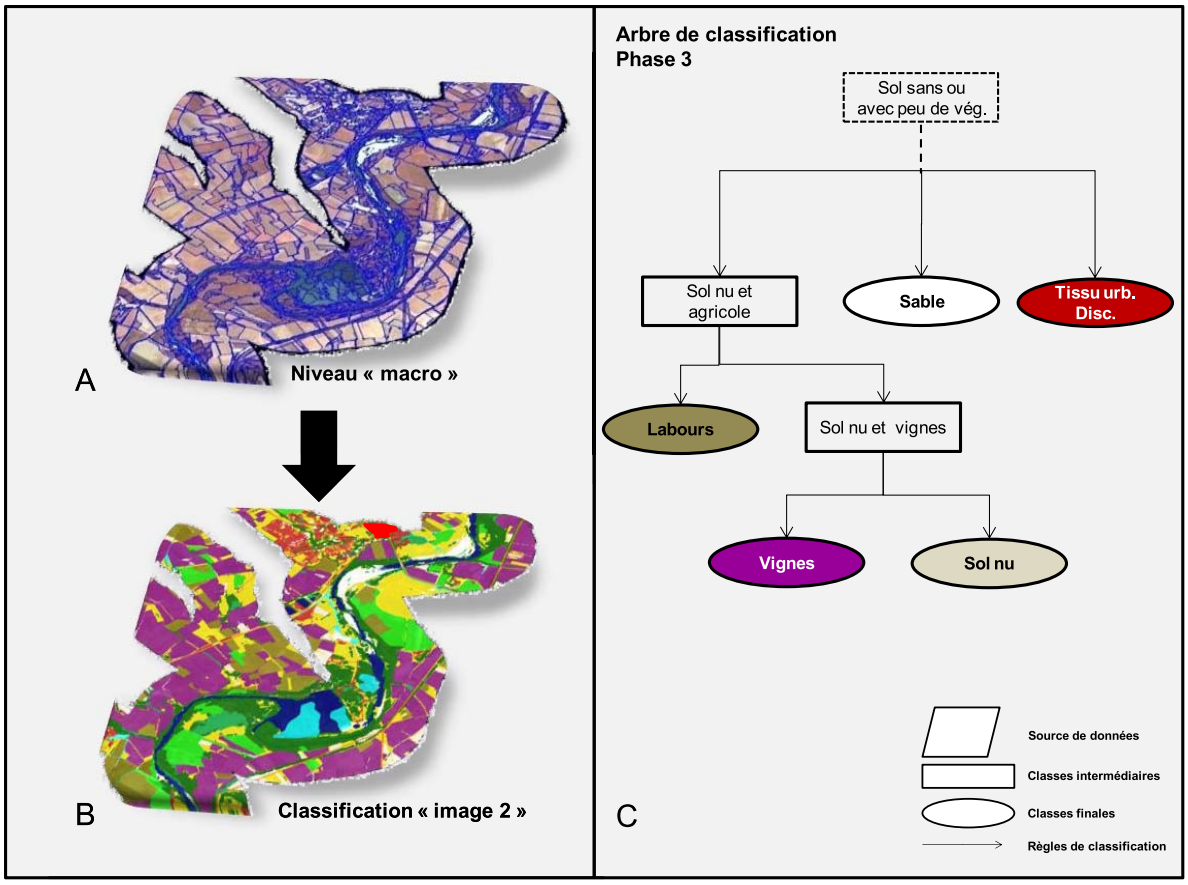


Figure 5.29 : Illustrations de la segmentation (A), de la classification (B) et présentation du développement de l'arbre de classification pour la classe « sol sans ou avec peu de végétation » (C) à la phase 3

CemOA : archive ouverte d'Irstea / Cemagref

Ainsi, cette démarche d'analyse orientée objet multi-échelle se base sur un réseau hiérarchique d'objets composé de quatre niveaux de segmentation emboîtés (Figure 5.30) et utilise un arbre de classification construit progressivement selon un raisonnement descendant attribuant des classes aux objets appartenant à trois niveaux de ce réseau (niveau « thématique », niveau « mico » et niveau «macro »).

Ces classes sont décrites à partir de règles de classification classiques, lorsqu'elles utilisent l'information issue des sources de données auxiliaires, et floues lorsqu'elles utilisent des critères basés sur l'information des sources de données images. Ces critères sont sélectionnés parmi les attributs spectraux (moyenne, ratio…), texturaux (textures haralick : entropie, homogénéité…), de forme (superficie, rapport longueur sur largeur…) et de voisinage (distance, frontière commune avec une classe donnée…) fournis par le logiciel. Ils peuvent aussi être sélectionnés parmi des attributs personnalisés comme les indices de végétation (e.g., NDVI, IRI) ou d'autres indices spectraux (indice de brillance BI, indice d'humidité du sol NDWI…).

Dans la pratique, cette démarche est réalisée sur une première région homogène et est adaptée par la suite aux autres régions. L'article présenté dans le chapitre 6 (cf. Chapitre 6 :II.2) illustre la mise en œuvre et teste la robustesse de cette démarche sur le bassin de l'Hérault. On trouvera dans cet article l'arbre de classification complet et les critères utilisés pour définir les règles de classification.

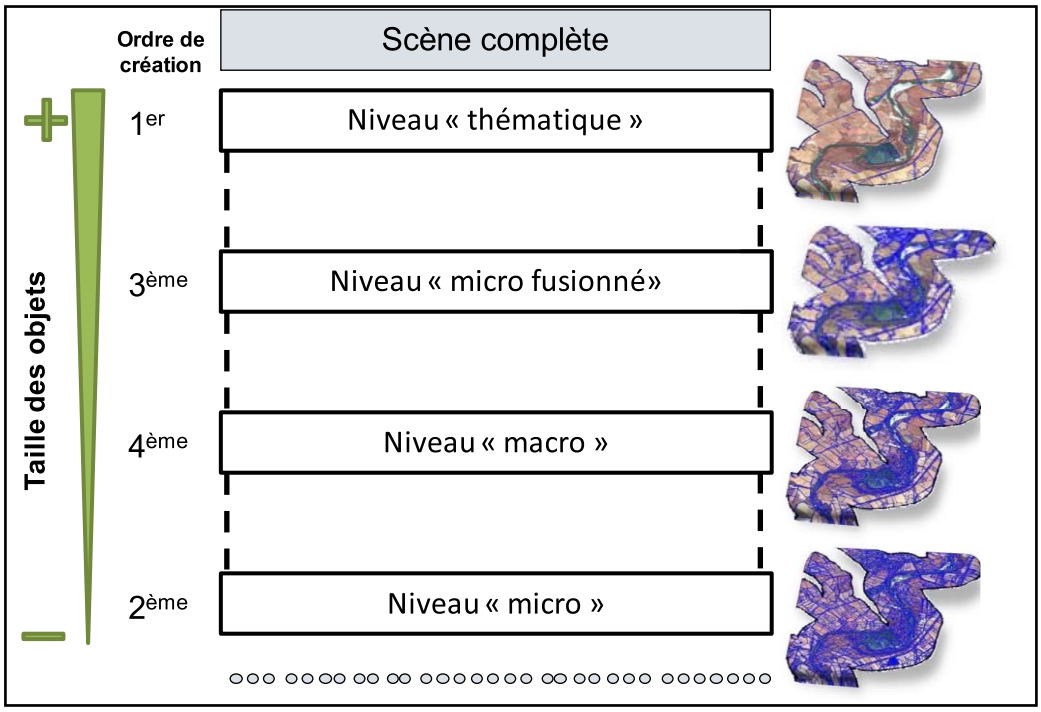


Figure 5.30 : réseau hiérarchique d'objets de la démarche de classification. Il se compose de quatre niveaux de segmentation emboîtés

Par ailleurs, afin d'être certain que l'utilisation de la méthode de classification par fonctions d'appartenance était le choix le plus judicieux, nous avons voulu tester l'autre méthode de classification, utilisant la technique du plus proche voisin, disponible dans le logiciel.

En effet, la sélection de 50 parcelles d'apprentissage par classe (échantillonnage optimal cf. III.5.2) nous semblant raisonnable à réaliser dans chacune des régions, cette technique de classification supervisée a été donc testée sur le dernier niveau de segmentation. Elle a été appliquée pour classer directement les classes finales définies à ce niveau (9 classes) (***ANNEXE VII.1***).

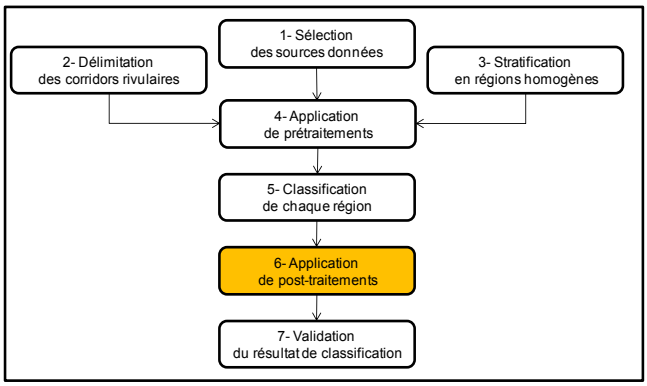
A l'issue de ce test nous pouvons dire que cette technique s'est avérée d'une part moins fiable (une perte de 10 % dans la précision globale du résultat par rapport à la technique « expertisée ») et d'autre part très difficile à mettre en œuvre : alors que la méthode par fonctions d'appartenance peut être distribuée (sur un serveur de calcul), la méthode par plus proche voisin impose de travailler sur l'ensemble de la zone étudiée ce

CemOA : archive ouverte d'Irstea / Cemagref

qui augmente considérablement les temps de traitements (10 fois plus de temps), et la rend rédhibitoire lorsqu'on veut l'appliquer à l'ensemble d'un bassin versant.

De plus, la technique par fonctions d'appartenance offre plus de modularité que la technique supervisée. En effet, il est possible d'ajouter des règles ou d'en enlever simplement selon les besoins de l'utilisateur sans remettre en question l'ensemble de la classification.

III.6. Application de post-traitements (étape 6)



Deux types de post-traitements (Figure 5.31) sont appliqués dans cette procédure de classification afin d'obtenir une carte d'occupation du sol continue sur la zone d'étude selon une nomenclature compréhensible, adaptable et utilisable pour la construction des indicateurs spatialisés.

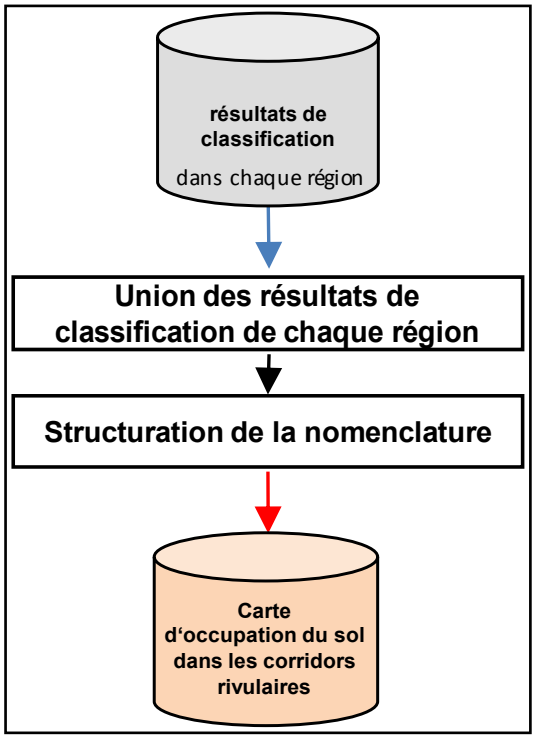


Figure 5.31 : Présentation des post-traitements appliqués (étape 6) sur les résultats de classification obtenus dans chaque région.

III.6.1.Union des résultats de classification de chaque région

CemOA : archive ouverte d'Irstea / Cemagref

Un premier post-traitement est nécessaire pour obtenir une donnée d'occupation du sol continue sur l'ensemble de la zone d'étude. Un algorithme issu des SIG est utilisé pour unir les résultats de classification (couches vecteur) obtenus pour chaque région homogène. Les objets de même classe en limite de chaque région sont fusionnés manuellement.

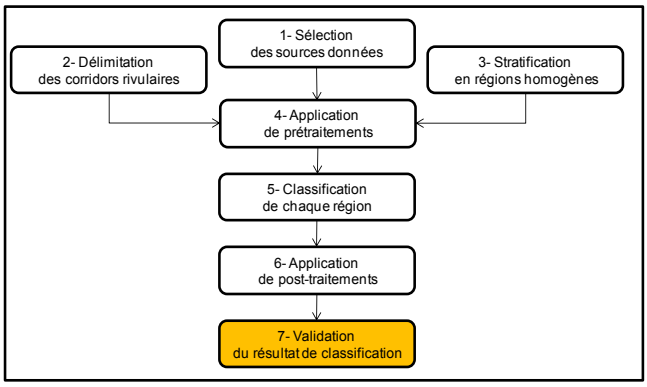
III.6.2.Structuration de la nomenclature

Le résultat de classification sur l'ensemble de la zone d'étude ainsi obtenu, est structuré dans une nomenclature compréhensible par différents utilisateurs qui permet d'étudier l'occupation du sol à différents niveaux d'organisation.

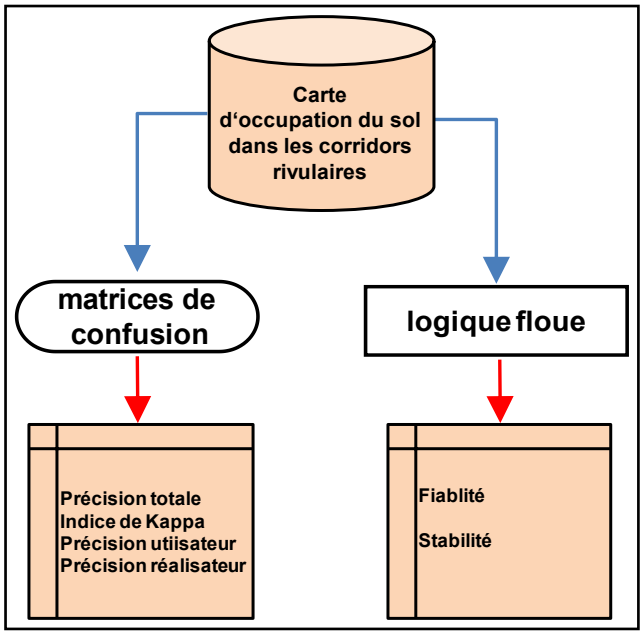
La nomenclature de classification de CORINE Land Cover a été prise comme modèle. Elle comprend trois niveaux hiérarchiques (Bossard *et al.*, 2000, ***ANNEXE IV.1***). Cette structure hiérarchique est devenue de plus en plus courante car elle est efficace dans la définition des classes (Girard and Girard, 1999).

Cette structure permet aussi d'intégrer aisément de nouvelles classes d'occupation ou de nouveau niveaux hiérarchiques ce qui facilite son utilisation sur des régions très diverses.

III.7. Validation du résultat de classification (étape 7)



Deux techniques sont utilisées pour valider le résultat de classification obtenu sur l'ensemble de la zone d'étude (Figure 5.32): (1) en réalisant une matrice de confusion à partir d'un échantillon de validation, et (2) en étudiant, grâce aux propriétés de la logique floue, la fiabilité et la stabilité de la classification.



CemOA : archive ouverte d'Irstea / Cemagref

Figure 5.32 : Validation du résultat de classification (étape 7).

III.7.1.La matrice de confusion

III.7.1.1 Principe et critère d'évaluation de qualité

La comparaison statistique entre la vérité terrain (échantillon de validation) et le résultat de la classification se présente comme une matrice de confusion (Tableau 5-6) avec plusieurs indicateurs synthétiques de fiabilité pour évaluer la classification. Dans la matrice de confusion la « vérité terrain » figure en ligne et la « classification » en colonne. Les valeurs de la matrice peuvent représenter des nombres de pixels ou des nombres d'objets.

Ainsi, la diagonale de la matrice correspond en général aux individus (pixels ou objets) bien classés (en jaune sur le Tableau 5-6). Si l'on considère la ligne i (individus qui dans la réalité appartiennent à la classe i) on trouvera dans la colonne j (case i,j) le nombre d'individus qui sur le terrain sont de la classe i mais ont été attribués à la classe j par la procédure de classification. Si l'on considère la colonne j (individus qui ont été attribués à la classe j par la procédure de classification) on trouvera dans la ligne i (case i,j) le nombre d'individus qui ont été attribués à la classe j par la procédure de classification mais appartiennent en réalité à la classe i.

Tableau 5-6 : Notation sur une matrice de confusion

		Classification			
		Unité 1	Unité i	Unité n	Total
Référence	Unité 1	x11	x1i	x1n	N1
	Unité i	xi+	xii	x+i	Ni
	Unité n	xn1	xni	xnn	Nn
	Total	M1	Mi	Mn	N

On distingue plusieurs mesures ou indicateurs de fiabilité (Girard and Girard, 1999), comme les estimateurs par classe exprimant la qualité de la classification pour chacune des classes :

- la *précision pour l'utilisateur* est le pourcentage d'individus (pixels ou objets) d'un groupe issu de la classification qui sont correctement classé vis-à-vis de la référence. Elle peut se lire selon les colonnes. D'après le Tableau 5-6, la précision de la classe i pour l'utilisateur (en %), s'exprime de la manière suivante :

$$PUi = \frac{xii}{Mi}$$

Cette précision pour l'utilisateur indique la probabilité pour qu'un objet qui a été classé dans la classe i soit effectivement, dans la réalité, un objet de classe i.

- la *précision pour le réalisateur* (producteur) correspond au pourcentage d'individus (objets ou pixels) d'une classe de référence qui sont correctement classés par la procédure de classification ; elle se lit selon les lignes. D'après le Tableau 5-6, la précision de la classe i pour le réalisateur (en %) s'exprime de la manière suivante :

$$PRi = \frac{xii}{Ni}$$

Cette précision pour le réalisateur indique le taux de réussite avec lequel il a réussi à classer les objets appartenant dans la réalité à la classe i (l'échantillon de validation doit bien évidemment être strictement indépendant de tout échantillon d'apprentissage).

CemOA : archive ouverte d'Irstea / Cemagref

De cette matrice, on peut également calculer des estimateurs globaux qui donnent une appréciation globale de la qualité des résultats de classification :

- La précision totale (PT) désigne le nombre d'individus (objets ou pixels) bien classés, (correspondance entre les groupes et les classes) rapporté au nombre total d'individus (objets ou pixels). Elle correspond donc à la proportion d'individus bien classés mais ne donne pas d'information sur les classes possédant une meilleure fiabilité. D'après le Tableau 5-6, la précision totale s'exprime de la manière suivante :

$$PT = \frac{\sum_{i=1}^{l} x_{ii}}{N}$$

- L'indice Kappa (KIA) d'une classification tient compte des erreurs en lignes et en colonnes (erreurs de déficits et d'excédents). Il exprime la réduction de l'erreur par rapport à celle obtenue par une classification qui serait réalisée de façon aléatoire. Il varie de 0 à 1. Par exemple, la valeur de 0.75 exprime que la classification utilisée évite 75 % des erreurs obtenues par une procédure travaillant complètement au hasard. D'après le Tableau 5-6, l'indice de Kappa (KIA) s'exprime de la manière suivante :

$$KIA = \frac{\left| N\sum_{i=1}^{l} x_{ii} - \sum_{i=1}^{l} (x_{i+} x_{+i}) \right|}{N^2 - \sum_{i=1}^{l} (x_{i+} x_{+i})}$$

III.7.1.2 Mode d'échantillonnage de validation

L'objectif n'est pas d'évaluer la précision géométrique de la donnée d'occupation du sol, mais d'évaluer uniquement la qualité sémantique du résutat (qui est d'ailleurs intrinsèquement liée à la qualité géométrique). L'objet résultant de la classification sera donc utilisé comme unité d'échantillonnage.

On réduit ainsi le temps du processus de validation car la prise en compte de la qualité géométrique dans la validation nécessite une numérisation des objets sur l'image par photo-interprétation ce qui augmente considérablement le temps de ce processus. On a pu le constater dans l'étude comparative des résultats de classification issus de combinaisons de sources de donnée différentes. Dans cette étude, la qualité géométrique avait été analysée afin d'apprécier les capacités de trois combinaisons contenant des informations de résolutions spatiales différentes (Spot 5 XS dans la première combinaison, Spot5 Pan dans la deuxième, et orthophotos dans la troisième) pour extraire les objets d'occupation du sol dans le corridor rivulaire.

L'objet ou le polygone (ou segment, i.e. résultat de la segmentation) a souvent été utilisé comme unité d'échantillonnage (Zhan *et al.*, 2005; Tiede *et al.*, 2006 ; Schöpfer and Lang, 2006; Grenier *et al.*, 2008), compte tenu du fait que les critères utilisés pour évaluer la qualité de la classification sont calculés au sein de chaque segment. Cependant, les tailles des objets issus de la segmentation-classification ne sont pas homogènes puisqu'ils ont été définis à partir de différents niveaux de segmentation (différents paramètres d'échelle). Par conséquent, pour tenir compte de cette hétérogénéité, nous privilégions la construction des matrices de confusion en fonction de la superficie de ces objets (en pixels).

L'échantillonnage de validation doit être représentatif de l'ensemble de la zone d'étude et sélectionné sans biais pour assurer la robustesse statistique de l'analyse (Congalton, 2004). Un algorithme a donc été développé (sous Python 2.4) pour sélectionner aléatoirement les parcelles de validation avec l'objectif d'obtenir un échantillonnage bien-distribué sur la zone d'étude (réparti de façon homogène).

CemOA : archive ouverte d'Irstea / Cemagref

Cet algorithme divise le plan d'échantillonnage (emprise de la zone d'étude) en X cellules de grille de la même taille (avec X le nombre de parcelles à sélectionner par classe). Chaque cellule qui présente une surface d'intersection avec le corridor rivulaire de la zone d'étude sert de strate géographique pour l'échantillonnage. L'algorithme se déplace de strate en strate en sélectionnant à chaque passage un objet appartenant à chaque classe. Cette sélection commence par une strate choisie au hasard (Figure 5.33). Le processus s'arrête lorsqu'on obtient le nombre de parcelles de validation requis pour chaque classe.

Une fois les parcelles de validation sélectionnées, la classe de chacune d'entre elles est vérifiée par photo-interprétation en utilisant les orthophotos comme donnée de référence. Une vérification sur le terrain serait un processus plus long et plus coûteux compte tenu de la taille de la zone d'étude (Zhu *et al.*, 2000).

Pour réaliser une matrice de confusion, il faut disposer d'un nombre suffisant de parcelles afin d'obtenir des résultats de validation statistiquement robustes. Cependant, comme pour l'échantillonnage d'apprentissage, l'utilisateur dispose de peu d'informations sur le nombre de parcelles de validation minimum à sélectionner par classe. Ce choix est lié à l'expérience de l'utilisateur : alors que Congalton, 1991, conseille un nombre de 50 parcelles/classe, Girard and Girard, 1999, en proposent 30.

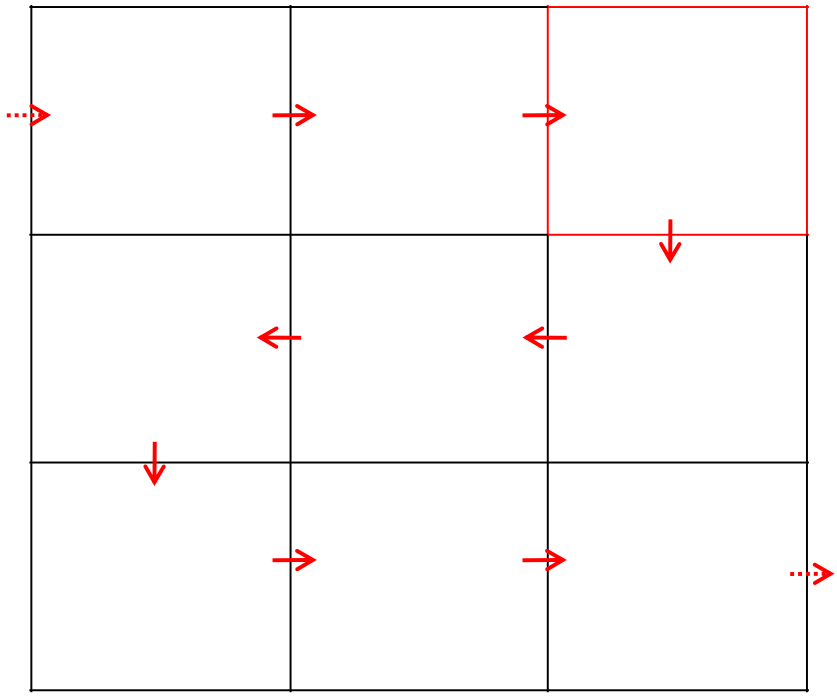


Figure 5.33 : Sens de déplacement de l'algorithme d'échantillonnage des parcelles de validation. L'algorithme commence la sélection par la strate rouge sélectionnée aléatoirement.

Nous nous sommes donc intéressés à rechercher la taille d'échantillon (nombre de parcelles de validation par classe) permettant d'obtenir des résultats de validation robustes. Cette analyse est présentée dans le détail en ***ANNEXE VII.3***. Pour cela, à partir d'un jeu de parcelles comportant 300 parcelles de contrôle par classe d'occupation du sol (selon une typologie à 9 classes), sélectionnées à l'aide de l'algorithme présenté précédemment, plusieurs matrices de tailles d'échantillon de validation croissantes ont été construites en piochant aléatoirement dans ce jeu de parcelles. Ce processus a été itéré un grand nombre de fois pour tester la robustesse de chaque échantillon de taille différente.

L'analyse montre tout d'abord qu'en moyenne (des itérations de chaque échantillon de taille donnée) on obtient un résultat de pourcentage de bien classé (précision globale) similaire quelque soit le nombre de parcelles de validation (84.22 % ± 0.34 pour une typologie à 9 classes). Ceci est logique puisque le grand nombre d'itérations ramène à la moyenne générale de la classification.

Par contre, on observe que plus on augmente le nombre de parcelles de validation, plus le résultat de classification est robuste (faible écart-type) c'est-à-dire peu variable en fonction de l'échantillon. D'après la

CemOA : archive ouverte d'Irstea / Cemagref

Figure 5.34, on constate qu'à partir de 50 parcelles l'évaluation de ce pourcentage se stabilise. Nous retenons donc pour les travaux de validation une norme de 50 parcelles de validation par classe.

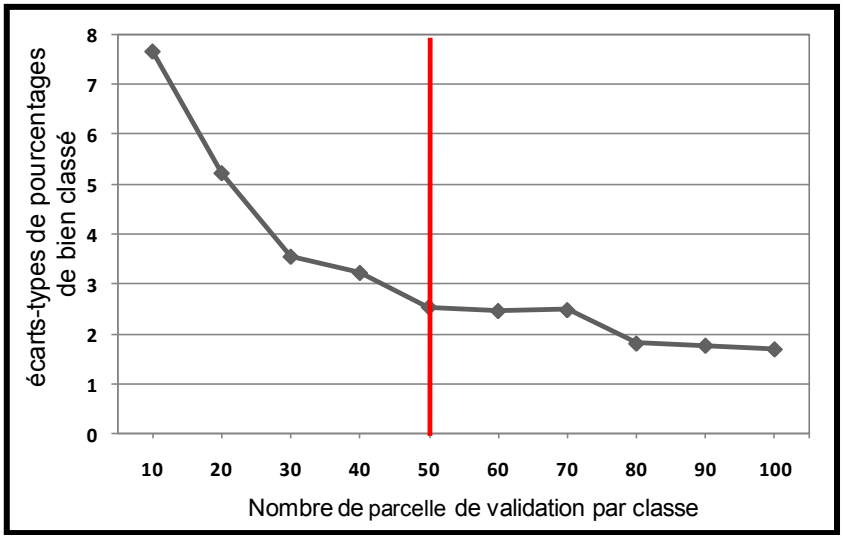


Figure 5.34 : Cette figure présente l'évolution des écarts types de pourcentage d'objets bien classé (précision globale) en fonction de l'augmentation de la taille de l'échantillon de validation.

III.7.2.La fiabilité et la stabilité de la classification

Grâce à la logique floue utilisée dans le processus de classification, la fiabilité et la stabilité de la classification peuvent être évaluées (Benz *et al.*, 2004).

En sortie du résultat de classification, chaque objet possède une valeur (ou degré) d'appartenance à sa classe. Plus ce degré sera élevé (proche de 1), plus la probabilité pour l'objet d'appartenir à cette classe sera forte et plus la classification sera fiable. On peut alors analyser statistiquement la fiabilité par classe (moyenne, écart type, minimum et maximum des valeurs d'appartenance des objets appartenant à la même classe) et de la même façon la fiabilité de classification (moyenne des différents estimateurs statistiques).

On connaît aussi la manière dont les classes se mélangent au sein d'un objet (cf. III.5.4.1). Ce mélange permet de mesurer la stabilité de l'objet vis-à-vis de sa classe finale (classe pour laquelle l'objet à la plus forte valeur d'appartenance). La stabilité correspond à la différence entre les deux meilleurs degrés au sein de l'objet. Plus cette différence sera forte, plus cet objet sera stable. On peut alors analyser statistiquement la stabilité par classe (moyenne, écart type, minimum et maximum des valeurs de stabilité des objets appartenant à la même classe) et de la même façon la stabilité de la classification (moyenne des différentes estimateurs statistiques).

Par rapport à la technique par matrice de confusion, la validation à l'aide de la logique floue permet de spatialiser les incertitudes de la donnée d'occupation du sol puisqu'elle évalue la fiabilité et la stabilité pour chaque objet. Cette propriété permet donc de propager de façon plus précise l'incertitude de cette donnée sur celle des indicateurs spatialisés.

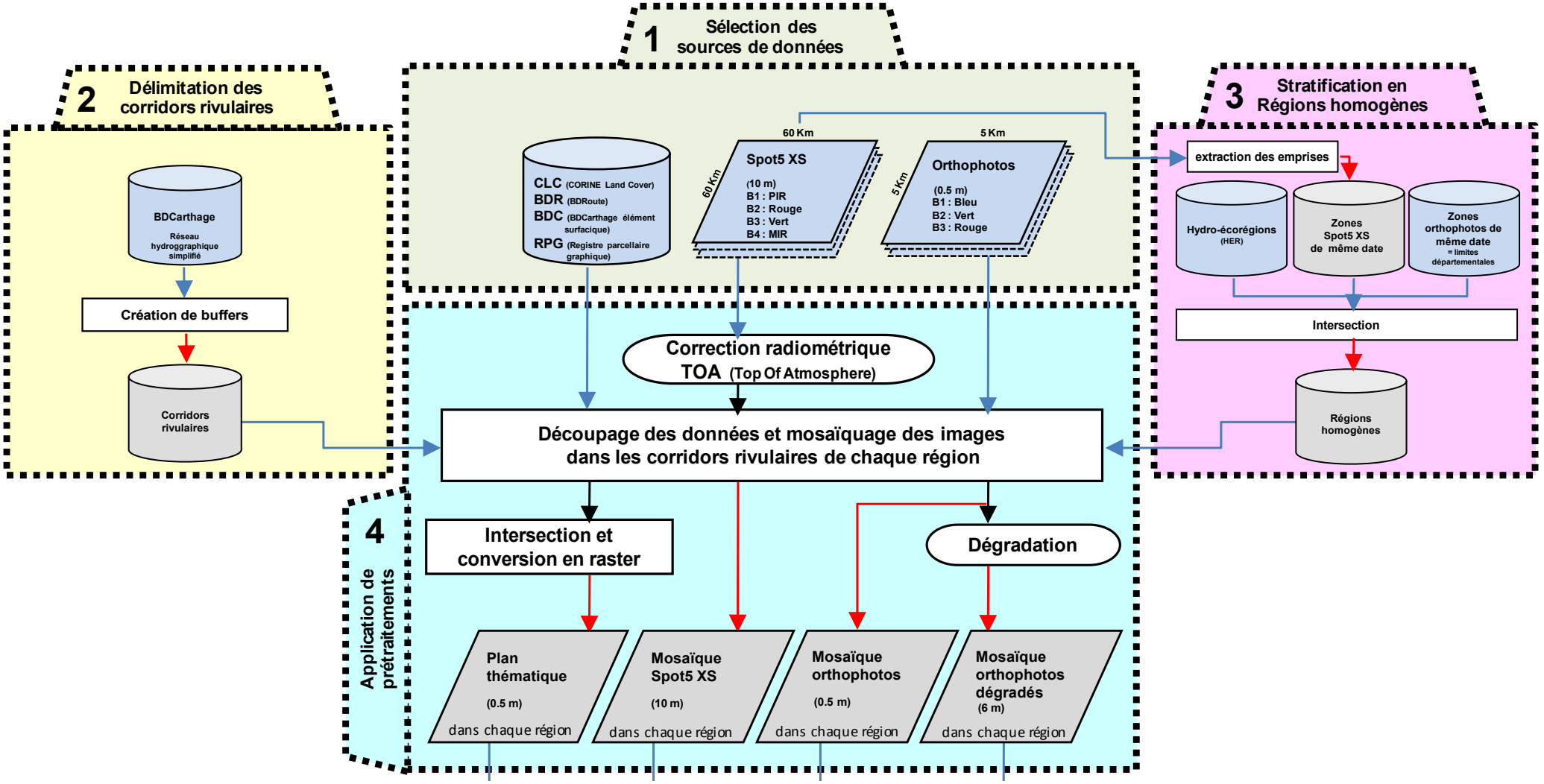
Il est toutefois important de noter deux limites à cette approche de la fiabilité et de la stabilité :

- cette évaluation est largement dépendante de l'opérateur puisque c'est lui qui définit par ces connaissances et son interprétation les ensembles flous décrits par les fonctions d'appartenances.
- elle ne constitue pas une approche de la précision puisqu'elle est menée indépendamment de toute vérité terrain : ce n'est donc pas une évaluation « absolue » de la qualité de la classification mais une évaluation « relative » de son niveau de fiabilité et stabilité dans le système de connaissance de l'opérateur.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

IV. Schématisation de la démarche de classification



CemOA : archive ouverte d'Irstea / Cemagref

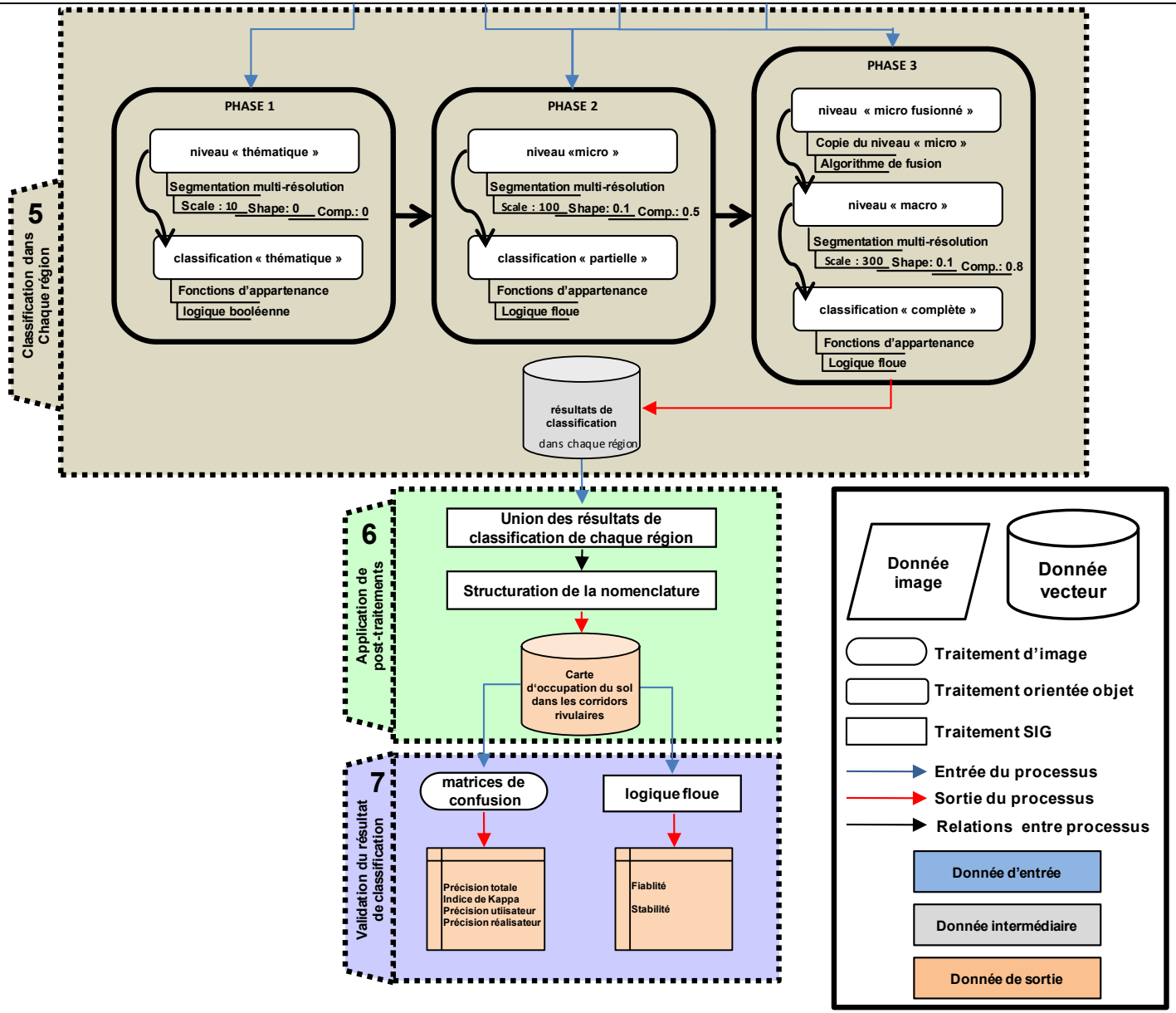


Figure 5.35 : Schématisation de la démarche de classification

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Chapitre 6
Application sur le bassin versant de l'Hérault

SOMMAIRE

Résumé *182*

I. OBJET DU CHAPITRE 184

II. EFFICACITE ET APPLICABILITE ET DE LA PROCEDURE DE CLASSIFICATION 184

II.1. RESUME ETENDU DE L'ARTICLE *184*

II.1.1. Présentation de la zone d'étude 184
II.1.2. Développement de la procédure 184
II.1.3. Principaux résultats 186
II.1.3.1 Qualité des résultats de classification 186
II.1.3.1 Contribution des sources de données et autres critères 187
II.1.3.2 Faisabilité technique de la procédure de classification 188
II.1.3.3 Apport incontournable de la THRS 189

II.2. ARTICLE SOUMIS A INTERNATIONAL JOURNAL OF REMOTE SENSING (IJRS) *190*

III. L'INCERTITUDE DE LA CARTE D'OCCUPATION DU SOL DANS LE CORRIDOR RIVULAIRE 224

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Résumé

La procédure de classification (procédure en 7 étapes développée au chapitre précédent) a été mise en œuvre sur l'ensemble du bassin versant de l'Hérault. Ce travail a fait l'objet d'une publication soumise à International Journal of Remote Sensing (IJRS) qui est en cours de révision. Un résumé étendu de l'article ainsi que le texte de la publication (en anglais) dans son intégralité sont présentés dans ce chapitre.

Cette mise en œuvre a permis d'évaluer l'efficacité et l'applicabilité de la procédure pour extraire de façon fine et détaillée l'occupation du sol dans les corridors rivulaires, sur de grands territoires. Dans l'étape 1, 138 orthophotos et 2 images Spot5 XS d'archives datant des mois d'avril à mai ont été acquises sur la zone d'étude, en plus des données auxiliaires. Dans l'étape 2, 553 km² environ de corridors rivulaires ont été délimité. Dans l'étape 3, six régions homogènes (de même date et de même type d'occupation du sol) ont été définies sur le bassin versant de l'Hérault à partir des emprises des images Spot5 XS et des limites des différentes hydro-écorégions. Une fois les données découpées (et mosaïquées pour les orthophotos) dans ces espaces et prétraités dans l'étape 4, la démarche de classification basée sur une approche orientée objet descendante a été appliquée dans chaque région homogène (étape 5). Les résultats de classification dans chaque région ont été par la suite réunis afin d'obtenir une carte d'occupation du sol continue sur la zone d'étude selon une nomenclature emboîtée à trois niveaux (étape 6), avant d'être validés à partir des matrices de confusion et des propriétés de la logique floue (fiabilité et stabilité) (étape 7).

Au final, en combinant de façon judicieuse l'information de ces données multi-source, une donnée détaillée (22 classes) et fine d'occupation du sol dans les corridors rivulaires a été extraite sur l'ensemble du bassin versant de l'Hérault avec une bonne qualité (89 % de précision totale) selon la typologie requise à minima pour étudier les relations pressions/impacts (surface en eau, sols semi-naturels, sols artificialisés, sols agricoles, zones de végétation arborée et zones de végétation herbacée).

En plus d'être efficace et fiable cette procédure s'est avérée (i) robuste puisque l'occupation du sol a été extraite dans ces espaces de façon fiable quelque soit la région (précision totale variant de 83 à 95 % selon les régions) ; (ii) facilement transférable puisque le même arbre de classification (excepté pour classer les parcelles issues du registre parcellaire graphique) a été utilisé et qu'aucune nouvelle règle n'a dû être définie, seuls les seuils de ces règles étant ajustés en passant d'une région à une autre ; (iii) rapidement applicable sur de grands territoires puisque trois jours de traitement orienté objet ont suffit pour caractériser l'occupation du sol le long des 1150 km du réseau hydrographique du bassin versant de l'Hérault ; et (iv) portable et évolutive puisqu'elle permet d'intégrer facilement de nouvelles sources de données, notamment en gérant les conflits entre les différentes sources grâce à la logique floue. De plus, cette procédure limite les coûts pour les gestionnaires français car elle utilise en partie des données déjà en leur possession ou accessibles à coût réduit sur le territoire français, et emploie des règles simples qui facilitent l'appropriation de la méthodologie par un nouvel opérateur.

Cependant des efforts restent à fournir (i) au niveau de la classification pour discriminer plus efficacement les classes de « sols nus naturels », de « sols artificialisés » et de « sols nus agricoles » ; (ii) au niveau de la validation à partir des propriétés de la logique floue car elle s'est avérée imprécise compte tenu de son absence de confrontation à une vérité terrain, et de la difficulté pour l'opérateur à définir des ensembles flous recouvrant a priori l'ensemble des erreurs possibles ; (iii) au niveau de la spatialisation de l'incertitude associée à la carte d'occupation des sols.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

I. Objet du chapitre

La conception de la procédure de classification présentée au chapitre 5 s'est faite par un travail progressif et itératif sur la zone d'étude pilote du bassin de l'Hérault.

La présentation de l'application sur le bassin de l'Hérault, objet du présent chapitre, n'est donc pas la simple mise en œuvre *à posteriori* d'une procédure conçue « sur le papier » mais une synthèse et une illustration sur le site pilote de la démarche.

Ce travail a fait l'objet d'une publication soumise à International Journal of Remote Sensing (IJRS) qui est en cours de révision. Le texte de la publication (en anglais) est intégralement présenté dans ce chapitre. Il inclut des éléments de présentation de la procédure que nous avons détaillée dans le chapitre 5 et illustre les résultats obtenus à chaque étape sur le site pilote. Il approfondit particulièrement l'applicabilité de la procédure (i.e. transposabilité des règles et arbre de classification orientée objet) et l'analyse de la qualité des résultats de classification à l'aide des matrices de confusion (l'évaluation de la qualité de la classification à l'aide des propriétés de la logique floue n'est pas traitée dans cet article). Par ailleurs, il montre l'intérêt de la THRS pour le développement d'indicateurs pertinents pour la gestion des corridors rivulaires (objet de la partie 3).

Dans la deuxième partie de ce chapitre, un résumé étendu de l'article est proposé au lecteur (II.1.) pour limiter la redondance d'information mais il peut bien évidemment s'il le désire se référer à l'article directement (II.2.).

Ensuite, dans la troisième partie du chapitre (cf. III.), une discussion dédiée à l'évaluation de l'incertitude des cartes d'occupation du sol dans le corridor rivulaire est proposée notamment en présentant les résultats obtenus à l'aide des propriétés de la logique floue.

II. Efficacité et applicabilité et de la procédure de classification

II.1. Résumé étendu de l'article

II.1.1. Présentation de la zone d'étude

On rappelle ici que le bassin versant du fleuve Hérault (cf. Chapitre 4 :II.2.1) a été choisi comme site pilote pour cette étude car il présente (i) une taille intermédiaire (2650 km² de superficie, 1150 km de réseau hydrographique) suffisante pour juger de la capacité de mettre en œuvre la méthode sur de grands territoires ; et (ii) une diversité de contextes paysagers à la fois en termes d'hydro-écorégions (3 représentées sur le bassin), en termes de morphologie de cours d'eau (de cours d'eau de montagne jusqu'à un fleuve de plaine) et d'occupation des sols dans les corridors rivulaires (de naturelle à urbanisée).

Au-delà de ces deux critères ayant guidé le choix du bassin pilote, l'Hérault présentait l'avantage supplémentaire de la proximité géographique, donc d'une connaissance sur le type d'occupation du sol *à priori* plus importante que sur une autre région et la possibilité de vérités terrain. Nous disposions par ailleurs d'une base de données images et auxiliaires qui a dû être complétée au cours de la thèse pour appliquer la procédure de classification sur l'ensemble du bassin versant (Tableau 6-1 et Tableau 6-2, dernière colonne).

II.1.2. Développement de la procédure

Nous avons acquis tout d'abord les données images et auxiliaires identifiées dans l'étape 1 de la procédure (cf. Chapitre 5 :III.1). Ces données sont présentées dans le Tableau 6-1 pour les données images et dans le Tableau 6-2 pour les données auxiliaires. Parmi les images d'archive de Spot5 XS sur la zone, on a opté pour des

CemOA : archive ouverte d'Irstea / Cemagref

images de printemps (avril – mai) afin de mieux différencier les cultures annuelles (e.g., blé, orge…) en forte activité végétative durant cette période comparée à la végétation herbacée semi-naturelle. Mis à part les postes d'occupation du sol artificialisé de CORINE Land Cover (datant de 2000), les données apportent une information entre 2005 et 2006.

Tableau 6-1 : Présentation des données images utilisées sur le bassin versant de l'Hérault

	Nombres d'images	Date d'acquisition	Localisation	Données collectées durant la thèse
orthophotos	138	été 2005	dép. de l'Hérault	oui
		été 2006	dép. du Gard	oui
Spot5 XS (images d'archive)	2	27 avril 2005	Nord du BV	oui
		14 mai 2005	Sud BV	non

Tableau 6-2 : Présentation des données auxiliaires utilisées sur le bassin versant de l'Hérault (MAJ = Mise A Jour)

	Date de production	Données collectées durant la thèse
CORINE Land cover	2000	Non
Route 500 ®	MAJ de 2006	Oui
BDCarthage ®	MAJ de 2006	Non
Registre Parcellaire Graphique	2006	Oui

553 km² environ de corridor rivulaire ont été délimités (étape 2, cf. Chapitre 5 :III.2) et six régions homogènes (de même date et de même type d'occupation du sol) ont été définies sur le bassin versant de l'Hérault à partir des emprises des images Spot 5 XS et des limites des différentes hydro-écorégions (étape 3 de la procédure cf. Chapitre 5 :III.3). Les limites des emprises des orthophotos n'ont pas été prises en compte car la variabilité radiométrique des images acquises en 2005 et 2006 est très faible.

Des prétraitements ont été appliqués sur ces données (étape 4 cf. Chapitre 5 :III.4) afin de découper les sources de données dans les différentes zones délimitées (corridors rivulaires et régions homogènes) et d'améliorer le contenu des sources de données images.

Ensuite, la démarche de classification basée sur une approche orientée objet descendante utilisant la technique de classification par fonctions d'appartenance (cf. Chapitre 5 :III.5) à partir de l'information combinée des orthophotos, de Spot5 XS et des données auxiliaires a été mise en œuvre dans chaque région homogène (chaque strate) avec pour objectif d'obtenir la typologie à 6 catégories requise *à minima* pour étudier les relations pressions/impacts (surface en eau, sols nus semi-naturels, sols artificialisés, sols agricoles, zones de végétation arborée et zones de végétation herbacée, cf. Chapitre 4 :I.1). On retrouve dans l'article l'arbre de classification (la hiérarchie de classe, Figure 5) et les paramètres de segmentation et de classification utilisés à chacune des phases de l'analyse orientée objet (Tableau 4).

CemOA : archive ouverte d'Irstea / Cemagref

Les résultats de classification dans chaque strate ont été réunis afin d'obtenir une carte d'occupation du sol continue sur la zone d'étude (étape 6, Chapitre 5 :III.6) selon une nomenclature emboîtée à trois niveaux.

Enfin, ces résultats ont été validés (étape 7, cf. Chapitre 5 :III.7) selon la typologie à 6 catégories uniquement à partir des matrices de confusion (construites en sélectionnant 50 parcelles de validation par classe).

II.1.3. Principaux résultats

On trouvera dans l'article les détails des différents résultats, nous reprenons ici les quatre constats qui nous semblent les plus instructifs et significatifs : la qualité des résultats de classification (cf. II.1.3.1), la contribution des différentes sources de données (cf. Chapitre 6 :II.1.3.1) ; la faisabilité technique de la procédure (cf. II.1.3.2) et l'apport incontournable de la THRS (cf. II.1.3.3) pour quantifier les indicateurs de pressions d'occupation du sol dans ces espaces.

II.1.3.1 Qualité des résultats de classification

A partir de l'information combinée des images Spot5 XS, des orthophotos et des données auxiliaires, une information détaillée (22 classes) et fine d'occupation du sol dans les corridors rivulaires, structurée dans une nomenclature emboitée à trois niveaux, a été extraite sur l'ensemble du bassin versant de l'Hérault (cf. Tableau 6 dans l'article). Cette donnée d'occupation du sol a été agrégée selon la typologie requise *à minima* pour étudier les relations pressions/impacts (6 catégories : surface en eau, sols semi-naturelles, sols artificialisés, sols agricoles, zones de végétation arborée et zones de végétation herbacée).

D'après les résultats de la matrice de confusion (cf. Table 7 dans l'article), construite selon la typologie à 6 classes, 89 % des pixels des objets vérifiés sur le corridor rivulaire du bassin versant de l'Hérault (553 km² environ) ont été correctement classés. Cette qualité de précision totale de la classification est confirmée par l'indice de Kappa puisque seulement 83 % des pixels auraient été bien classés selon une procédure de validation complètement au hasard.

Cependant on constate, que les catégories « sols artificialisés » et « sols nu semi-naturels » ont une plus faible précision aussi bien pour le réalisateur que l'utilisateur (la précision pour le réalisateur est de 60 % et celle de l'utilisateur de 65 % pour la catégorie « sols artificialisés » et de 35 et 51 % pour la catégorie « sols nu semi-naturels ») à cause des fortes confusions existantes entre ces deux classes. Ces erreurs sont dues aux comportements spectraux très similaires entre ces deux classes (e.g., toits des maisons, parkings/ roches nues, sables semi-naturels …).

On note toutefois que ces erreurs concernent essentiellement les objets isolés dans le paysage. Pour séparer les objets « sols artificialisés » d'autres types de sols sans ou avec peu de végétation (comme les sols nu semi-naturels) on se base sur un critère de texture Harralick (« GLDV entropy all directions ») à partir de l'information de la bande bleue dans l'orthophoto dégradée à 6 m. Plus l'information spectrale contenu dans l'objet est hétérogène et spatialement désordonnée (c'est le cas de la plupart des objets urbains), plus ce critère de texture est fort. Cependant, dans le cas d'objet isolés ce critère est moins efficace. En effet, les objets « sols artificialisés » isolés, généralement plus petit et homogène spectralement (e.g., parking, immeubles…), ne peuvent pas être discriminés par ce critère de texture. A l'inverse, certains objets « sols nu-semi-naturels » isolés contiennent une information spectrale hétérogène et spatialement désordonnée car ils sont recouverts d'une végétation parsemée, et sont donc considérés par le critère de texture comme des « sols artificialisés ».

Par ailleurs, la catégorie « sols nus semi-naturels » se confond également avec la catégorie « sols agricoles » car certains objets de ces catégories ont des comportements spectraux fortement similaires (e.g., terre labourée/sols nus semi-naturels ; vignes/ sols nus semi-naturels …).

D'après ces résultats, on peut donc affirmer que cette méthode de classification fournit une carte d'occupation du sol relativement détaillée (22 classes) qui possède globalement une information précise sur les six principales pressions d'occupation du sol. Cependant des progrès sont nécessaires pour les classes « sols artificialisés » et « sols nus semi-naturel » car confondre une catégorie susceptible d'altérer l'état écologique avec une catégorie non dégradante peut avoir des conséquences importantes dans l'étude de ces relations.

De plus, cette classification ne tient pas directement compte des ombres portées contenues dans les images, notamment dans les orthophotos. Ces ombres sont essentiellement liées aux zones à fort relief, au bâti

CemOA : archive ouverte d'Irstea / Cemagref

et à la végétation arborée. Dans le processus de classification les objets « ombre », quelque soit leur origine, sont d'abord classés au niveau micro en « végétation arborée », sur la base de l'indice de brillance calculé à partir des bandes spectrales de l'orthophoto. Au niveau macro, les objets « végétation arborée » (comprenant donc certains objets « ombre ») proches de zones artificialisés sont classés en « tissu urbain discontinu » (de manière à obtenir une tâche de tissu urbain discontinu). Ce raisonnement de classification a pour effet de classer les objets « ombre » :

(i) en « végétation arborée » lorsqu'ils sont liés aux zones à fort relief, ce qui est cohérent dans les fonds de vallée étroits et encaissés généralement caractérisés par une épaisse ripisylve ;

(ii) en « tissu urbain discontinu » lorsqu'ils sont liés au bâti et à la végétation arborée proche du bâti, ce qui est cohérent pour les ombres contenues dans la tâche de tissu urbain discontinu mais moins pour celles qui sont en périphérie de cette tâche ;

(iii) en « végétation arborée» lorsqu'ils sont liés à la végétation arborée, ce qui cohérent pour les ombres contenues dans les zones de forêts denses, mais moins pour celles qui sont en périphérie de zones de forêts denses.

Ainsi, même si l'objectif n'est pas d'obtenir une cartographie très précise géométriquement, des progrès sont à réaliser pour identifier les ombres et les gérer en périphérie de groupes d'objets (tâche urbaine, zones de forêts denses,…) de manière à ne pas surestimer les indicateurs spatialisés liés aux catégories d'occupation des sols auxquelles sont attribuées ces ombres. Lors de la création des mosaïques orthophotos (opération qui consiste à produire une seule image à partir de l'ensemble des clichés aériens ortho-rectifiés) les clichés pris avec une inclinaison proche du nadir[38] sont favorisés, ce qui limite les ombres portées dans les images (Paparoditis *et al.*, 2006).

II.1.3.1 Contribution des sources de données et autres critères

La majorité des objets (77 % environ) dans les corridors rivulaires de ce bassin ont été classés uniquement à partir de l'information des orthophotos et des images satellites Spot5 XS. Ces objets recouvrent 65 % du corridor rivulaire étudié.

L'information spectrale des orthophotos est apparue indispensable dans la phase de segmentation pour détecter (délimiter) les objets de taille réduite et fragmentés dans ces espaces. Dans la phase de classification, aussi bien leur information spectrale que texturale a été nécessaire. Par exemple, l'information spectrale a permis de discriminer précisément la végétation arborée à partir de l'indice de brillance (calculé sur toutes les bandes de l'orthophoto) car les objets de cette classe possèdent sur cette zone d'étude une valeur dans cet indice plus faible que les autres objets de végétation ; puis l'information texturale a servi notamment dans la discrimination de la classe vigne (par rapport aux terres labourées et aux sols nus semi-naturels) en utilisant l'indice de texture Harralick d'entropie calculé dans toutes les directions à partir de la bande rouge au sein de l'objet (« Red-ortho GLDV entropy all directions »).

Concernant les images Spot5 XS, l'information spectrale dans le Proche Infra Rouge (PIR) a été essentielle dans la phase de segmentation fine pour améliorer la délimitation des objets « eau » (appartenant essentiellement au cours d'eau) de la végétation bordant ces objets ; puis dans la phase de classification elle a permis via notamment les indices de végétation, tel que le NDVI (« Normalized Difference Vegetation Index ») ou l'IRI (InfraRed Index), de différencier les différents couverts de végétation. Par exemple, un seuil flou dans le NDVI a été utilisé pour différencier les objets recouverts par une végétation parsemée des objets avec taux de couverture plus dense ; puis les cultures en activité ont pu être discriminées plus précisément à l'aide de l'IRI car elles sont caractérisées sur la zone d'étude par une forte valeur de cet indice.

Ces sources d'information ont été également combinées avec la donnée RPG (cf. Chapitre 5 :III.1.2.3) afin de discriminer les grandes occupations agricoles au sein de ces parcelles (labours, vignes, cultures, prairies). La donnée RPG s'est avérée primordiale pour traiter les objets agricoles sur la zone d'étude. Elle concerne 15 % environ des objets de la carte d'occupation du sol et recouvre 32 % de la superficie du corridor rivulaire du bassin versant de l'Hérault. Ces objets sont très difficiles à identifier en classification mono-date uniquement à partir de l'information des images : ils sont caractérisés par différents stades de développement selon leur utilisation agricole, qui se traduisent par différents comportements spectraux augmentant ainsi les risques de

[38]Le nadir est le point de la sphère céleste représentatif de la direction verticale descendante, en un lieu donné (par opposition à zénith, le point du ciel situé à la verticale montante d'un lieu donné).

CemOA : archive ouverte d'Irstea / Cemagref

confusion avec d'autres classes (exemple : cultures de blé dans les premiers stades de développement et végétation herbacée semi-naturelle). Avec la donnée RPG, la majorité des cultures (ou prairies), exceptées les vignes (culture non éligible par la PAC), ont pu être détectées facilement sur la zone d'étude. De plus ces parcelles, qui ont été délimitées par photo-interprétation sur l'orthophoto par les exploitants, constituent une base solide pour la segmentation, ce qui a pour effet d'améliorer *in fine* la classification.

Cependant, une fois ces parcelles détectées (malgré le choix de la date des images Spot5 XS), la distinction entre cultures et prairies uniquement à partir de l'information des images n'était pas aisée. Dans la zone d'étude, peu de terres agricoles sont dédiées aux prairies, car l'élevage n'est pas fortement développé dans cette zone, excepté dans les régions de montagne. On a donc adapté les règles et l'arbre de classification de ces parcelles en fonction de l'hydro-écorégion étudiée.

Le reste des objets de la carte d'occupation du sol (7 % environ) a été classé à partir de l'information des autres données auxiliaires définissant en tout 11 classes. Ces classes représentent sur cette zone d'étude un faible pourcentage (2.98 %) de la superficie étudiée.

Ainsi, chaque source d'information a été utilisée seule ou de façon combinée (avec d'autres sources d'information) pour classer l'occupation du sol dans le corridor rivulaire. Sur cette zone d'étude les données images sont la source majoritaire d'information. Mais selon la zone d'étude ces proportions peuvent relativement changer. C'est le cas notamment de la zone de modélisation, qui fortement occupée par les grandes cultures et des zones dédiées à l'élevage, est beaucoup plus recouvert par la donnée RPG (cf. Chapitre 9 :II.2.1.3.2).

En plus des critères liés directement aux informations issues de ces différentes sources de données, des critères de contexte (reliées aux classes des objets voisins) ont été utilisés notamment pour obtenir des tâches de tissu urbain discontinu : tous les objets de végétation semi-naturelle (herbacée ou arborée) à proximité de zones de bâti (une zone périphérique de 10m) ont été classés en végétation urbaine.

II.1.3.2 Faisabilité technique de la procédure de classification

Cette procédure s'est avérée :

(1) Reproductible puisque, quelque soit les différentes régions étudiées, la qualité (précision totale) des classifications obtenues est bonne à très bonne (de 83 % à 95 % environ) ;

(2) Facilement transférable sur un autre territoire car le même arbre de classification (excepté pour les parcelles du RPG) a été utilisé quelque soit la région traitée, ceci malgré la diversité paysagère de la zone d'étude, et aucune nouvelle règle (utilisation d'un autre critère) n'a dû être définie à chaque adaptation. Seuls les seuils des fonctions d'appartenance utilisés pour décrire les classes de l'arbre ont été ajustés par l'opérateur ;

(3) Rapidement applicable sur de grands territoires puisque les temps de traitement ont été relativement faibles compte tenu des coûts d'adaptation réduits de la démarche orientée objet et de la possibilité de distribuer sur un serveur de calculs les différents processus de cette procédure. A titre indicatif, la durée de traitement orientée objet, en intégrant le temps d'ajustement des règles, a été de 3 jours grâce à l'exploitation d'un cluster de 8 coeurs. Le temps d'ajustement dépend bien évidemment de la connaissance de la zone d'étude par l'opérateur et de la diversité de l'occupation qu'elle présente.

(4) Portable et évolutive (i.e. apte à intégrer de nouvelles sources de données) puisqu'elle permet de gérer facilement l'information issue de données multi-source, en réglant notamment les conflits entre ces sources grâce à logique floue. Cette propriété a été précieuse notamment pour discriminer la classe intermédiaire « sol avec de la végétation » de la classe intermédiaire « sol sans ou avec peu de végétation ». Suivant la source de donnée, Spot 5 XS ou orthophotos, l'information pour attribuer l'une des deux classes à un objet peut être contradictoire car elles n'ont pas été acquises à la même saison (avril-mai pour les images Spot5 XS et été pour les orthophotos). Ce cas de figure se présente particulièrement pour les terres agricoles dont le comportement spectral varie au cours de la saison (« parcelles labourées», « parcelle en végétation » ou « parcelle à maturité » suivant la saison pour une même culture).

De plus, cette procédure limite les coûts pour les gestionnaires français car elle utilise des données numérisée déjà en leur possession (par exemple les produits IGN BD Ortho® , BD Carthage®…) ou des données

CemOA : archive ouverte d'Irstea / Cemagref

images à coût réduit sur le territoire français (Spot5 XS) ; et emploie des règles simples qui facilitent l'appropriation de la méthodologie par de nouveaux opérateurs.

II.1.3.3 Apport incontournable de la THRS

La comparaison d'un indicateur spatialisé, correspondant au pourcentage de zones semi-naturelles et d'activités humaines proches du cours d'eau (buffer de 5 m de part et d'autre du cours d'eau), calculé d'une part à partir de l'information d'occupation du sol issue de la Très Haute Résolution Spatiale (THRS) et d'autre part à partir de la base de donnée CORINE Land cover (CLC) démontre l'apport incontournable de la THRS pour quantifier de façon précise et fiable l'occupation du sol dans les corridors rivulaires, et notamment la végétation rivulaire (ripisylve). Alors qu'avec CLC, les activités humaines dominent sur cette bande de 5m de part et d'autre du cours, ceci quelque soit le tronçon étudié, la THRS révèle que la majorité des tronçons sont dominés par la présence de végétation semi-naturelle dans cet espace proche des berges. Ce travail sur les indicateurs spatialisés est détaillé dans la troisième partie de ce mémoire.

CemOA : archive ouverte d'Irstea / Cemagref

II.2.Article soumis à International Journal of Remote Sensing (IJRS)

Object oriented approach for operational fine-scale regional mapping of land cover within river corridors from optical imagery and thematic data

T. TORMOS*†, P. KOSUTH†, S. DURRIEU†, B. VILLENEUVE‡, J.G. WASSON†‡ and S. DUPUY†
† Earth Observation & Geo-information for Environment and Land Management, Cemagref, 500 rue Jean François Breton, 34093 Montpellier, France
‡ Laboratory of Quantitative Hydroecology, Cemagref, 3 bis quai Chauveau, 69336 Lyon, France

Accurate mapping, at regional scale, of land cover diversity within riparian areas is a major challenge for better understanding the influence of riparian landscapes and related natural and anthropogenic pressures on river ecological status. As the structure (composition and spatial organization) of Land Cover within River Area (RALC) is generally not accessible using moderate-scale satellite imagery, finer spatial resolution imagery and specific mapping techniques are needed. For this purpose, we developed a classification procedure based on a specific multi-scale Object Based Image Analysis (OBIA) scheme dedicated to produce fine-scale and reliable RALC maps in different geographic context (relief, climate and geology) by exploiting information (available for managers and stakeholders) from very high spatial resolution imagery (satellite or airborne) and/or metric or decametric spatial thematic data, using fuzzy expert knowledge classification rules. This OBIA scheme was tested over the Hérault river watershed (southern of France) presenting a contrasted landscape (different ecoregions) and a total stream length of 1150 km, using the combination of multispectral and high spatial resolution imagery (10m Spot5 XS and 0.5m aerial photography) and several metric and decametric spatial thematic data. A RALC map was produced (22 classes) with an overall accuracy of 89% and a Kappa index of 83% according to a land cover pressures typology (six categories). Results of this experimentation demonstrate that the application of OBIA to multi-source spatial data provides an efficient approach for the mapping and monitoring of RALC that can be implemented operationally at regional or national scale. We further analyzed the influence of map resolution on the quantification of riparian spatial indicators to highlight the interest of such data for studying landscape influence at riparian scale.

1. Introduction

Riparian vegetation, being an interface between terrestrial and aquatic systems, influences the biodiversity and water quality of streams and river ecosystems in many specific ways (Naiman et al. 2005). Maintaining and restoring good riparian buffer area conditions could therefore constitute a major action to improve the ecological status of these ecosystems and reach by 2015 the targets of the European Water Framework Directive (WFD). Prior to the definition of efficient stream management and restoration strategies and actions, an improved understanding, at regional scale, of the influence of multiple human pressures along rivers and riparian vegetation on river ecological status is required (Allan 2004).

* Corresponding author. Email : thierry.tormos@onema.fr

CemOA : archive ouverte d'Irstea / Cemagref

For this purpose, Riparian Area Land Cover (RALC) maps serve as a basis to calculate spatial indicators related to human pressures along rivers and riparian vegetation properties (e.g., composition, continuity, strip width…). These indicators will help in understanding and predicting river ecological status (Shuft et al. 1999, Tormos et al. (In press) 2010).

Traditional RALC photo-interpretation mapping methods, based on aerial photography, are time-consuming and expensive to accurately characterize RALC over large study areas (Coulter et al. 2000). As a result, most studies analysing the relations between RALC spatial indicators and water bodies ecological status, have used maps derived from optical moderate spatial resolution satellites such as *Landsat* (30 m) or SPOT 4 (20 m). These satellites are commonly used to investigate earth surface over large territories ((Congalton et al. 2002). However, considering the spatial extension of riparian ecosystems and the diversity of land cover types in these systems, moderately high spatial resolution images appear insufficient for a fine and detailed characterization of RALC (Müller 1997, Apan 2002, Davis et al. 2002). Therefore, as observed by several authors (e.g., Frimpong et al. 2005, Hollenhorst et al. 2006, Roy et al. 2007), studying correctly influence of human pressures along rivers and riparian vegetation on river ecological status requires a reliably and fine-scale RALC map over large areas (at watershed, regional or national scale).

On the one hand, with progresses of optical HSR (High Spatial Resolution, 5 to 10 m) and VHSR (Very High Spatial Resolution, < 5 m) sensor technologies, a variety of remotely sensed images are currently available for mapping RALC at different levels of spatial detail (i.e. spatial resolution) and spectral detail (i.e. number of spectral bands). Goetz et al. (2003) have been the first to carry out a detailed map over a 1313km² tree cover and impervious surface watershed in central Maryland, using 11 Ikonos image data (4 m pixels). Since then, Yang (2007) has implemented a method to delineate and quantify the extent of riparian vegetation from digital aerial photographs along a 593 km river linear of 1st to 4th stream order. Noonan and Chafer (2007) have computed an invasive willow map from SPOT 5 XS images (10 m pixels) over the 2158 km long Cox river drainage system. Johansen et al. (2008) have developed a method to quantify riparian conditions (percentage of canopy cover, canopy continuity and organic litter) from Quickbird images (2.4 pixels) at scales from 1 km to hundreds of km of stream length. Nevertheless, these studies at broad scale did not characterize both the main human pressures (agricultural and urban areas) and riparian vegetation conditions.

On the other hand, with emergence of Geographical Information Systems (GIS) technologies in the last decades, stakeholders and/or managers currently have numerous thematic data containing spatial information on land cover and land use produced, for different management objectives on their territory. Such data sources could provide significant information for improving characterization of RALC.

Therefore, developing cost-effective, accurate, time-efficient and spatially extensive mapping techniques able to combine information from HSR or VHSR optical remotely-sensed data and existing spatial thematic data, and using semi-automatic or automatic processing is a major challenge.

With increased spatial resolution (smaller size of pixels), spectral information within an image is spatially heterogeneous, and conventional pixel-based classification techniques are no more suitable (Latty and Hoffer 1981, Irons et al. 1985, Müller et al. 1993, Kressler et al. 2003, Lennartz 2004, Yan et al. 2006). The pixel-based classification accuracy is relatively low and the resulting maps suffer from a well-

CemOA : archive ouverte d'Irstea / Cemagref

known "salt and pepper" effect (Ivits et al. 2005, Durieux et al. 2007b). Object-Based Image Analysis (OBIA) is an efficient solution: in a first step a segmentation process identifies and builds up homogeneous regions (segments or image objects) and in a second step a classification process is applied to these objects using spectral as well as spatial information such as texture, shape and context features, to increase the discrimination between spectrally similar land cover types.

Recent developments in OBIA, based on a Fractal Net Evolution Approach (FNEA), have revolutionized the processing of high to very high remote sensing data by providing effective computer-assisted classification techniques whose results come close to the quality of manual photo-interpretation, while being much faster and cheaper and much more reproducible (Blaschke and Hay 2001, Benz et al. 2004, Carleer and Wolff 2006, Durieux et al. 2007a, Aplin and Smith 2008, Zhou and Troy 2008, Blaschke 2009).

Moreover, the FNEA incorporates an object-oriented framework and various image classification techniques (Blaschke and Hay 2001). In particular, it uses fuzzy classification that improves class description with understandable linguistic concepts built from expert knowledge and facilitates fusion of multi-source heterogeneous information fusion (Benz et al. 2004).

In consequences, the present research was devoted to develop a reproducible classification procedure using an OBIA scheme based on FNEA approach, able to combine multi-source high spatial resolution data (remote sensing images and thematic data) with affordable cost for manager and stakeholders, to produce a reliable detailed and finely resolved RALC map over large and geographically contrasted territories. The method has been designed to map and monitor the RALC over the overall French river network (500 000 km) in order to obtain an information about the main land cover categories in the river corridor generating pressures on river ecological status. We therefore selected a large pilot study region, in order to test the classification procedure feasibility using available cost-effective high resolution multi-spectral images and existing thematic data on French territory. We also examined the impact of map resolution on the quantification of spatial indicators at riparian scale, by comparing sets of indicators derived either from the resulting maps or from the European CORINE Land Cover map, considered here as a reference since it is the only land cover map available at present on the whole territory.

2. Methods

2.1 Overview

A classification procedure, composed of 7 stages, was designed in order (i) to produce a detailed and finely resolved RALC map over large riparian territories in different geographical contexts (i.e, relief, climate and geology), based on information from multi-source spatial data affordable for managers or stakeholders on a given studied zone, and (ii) to validate its accuracy. The flowchart showing the different stages of this procedure is presented in figure 1.

The first three stages are dedicated to input data selection (st.1) and delineation, in these data, of the region under study (st.2) and homogeneous subregions in terms of both geographical contexts and data acquisition dates (st.3). The three following stages are dedicated to object based image analysis and classification (st.5) with a pre-

CemOA : archive ouverte d'Irstea / Cemagref

processing (st.4) and a post-processing (st.6) stages. A final stage is dedicated to the validation of resulting RALC (riparian area land cover map). It should be underlined that the choice of the classification typology is defined by final users, upstream of the methodology implementation (prior to stage 1).

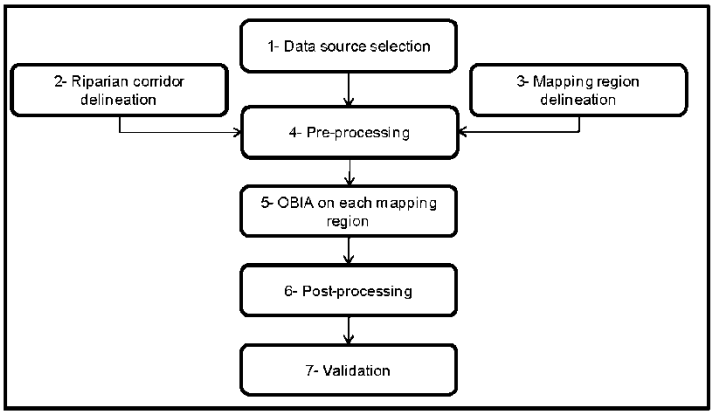


Figure 1. Flowchart of the land cover within riparian areas classification procedure

2.2 Details of the seven stages

2.2.1 Stage 1: Data source selection. In this stage all sources of spatial data available and suitable for detailed and fine characterization of RALC, optical remotely sensed images and/or thematic data, are acquired.

***Optical images*:** there are numerous sensors currently on-board satellites or planes that observe the earth and produce images that vary in terms of spatial, spectral, and temporal resolutions. The choice of the theses image parameters critically impacts on the feasibility and accuracy of the regional RALC mapping (Ashraf et al. 2010).

Spatial resolution, which refers to the earth's surface area covered by a single pixel of the image, is the most significant parameter that influences the accuracy of RALC classifications because of the small size and heterogeneous nature of land cover objects in riparian areas (Müller 1997). Moderate spatial resolution satellite data (≥ 20 m pixels such as the early SPOT 4 and *Landsat* images) have yielded degraded classifications compared to HSR and/or VHSR data (Congalton et al. 2002, Tormos et al. (In press) 2010). As a result, for mapping RALC HSR and/or VHSR are indispensable.

Spectral resolution, which is the measure of the sensitivity of a sensor to record information across the electromagnetic spectrum, critically influences what can be extracted from images. As a result, acquiring images with more spectral bands means more capacity to identify different types of land cover and produce a detailed land cover map. Therefore, (Becker et al. 2007) suggested that narrow and continuous hyperspectral data are required to get information about the composition and physical properties of different materials observed in riparian vegetation areas. Moreover, among spectral bands, in all riparian mapping studies, Near InfraRed (NIR) bands appear indispensable to discriminate vegetation classes. Many vegetation indexes derived from

CemOA : archive ouverte d'Irstea / Cemagref

the NIR bands improve riparian vegetation classification in these studies (e.g., Davis et al. 2002, Johansen and Phinn 2006, Akasheh et al. 2008).

Temporal resolution, which refers to the frequency of image acquisition for a given area, critically influences accuracy of a land cover map. Several studies have shown that the use of information on temporal (seasonal) changes in the spectral signature of land covers, which highly change over time such as crops, improves the map accuracy by reducing spectral confusion and increasing the number of land cover classes (e.g., Panigrahy and Sharma 1997, Lucas et al. 2007). The number of images and the acquisition dates required depend on temporal changes in the spectral signature of crops as a function of crops phenological stages and on the different land cover classes. Consequently, higher is the temporal resolution, easier it is to acquire images of a given area with suitable acquisition dates. Moreover a high temporal resolution image becomes critical when an area is often cloudy, as several passes may be required before a cloud-free image is obtained.

However, the choice of remote sensing data generally results from a trade-off between cost and accuracy (Phinn et al. 2003). In most cases for a national or regional stakeholder and/or manager, no such data are available with affordable cost on their territory.

Currently, only aerial spectrometry images are able to providing hyperspectral information at HRS or VHRS. Hyperspectral satellite sensors (e.g., *Hyperion* 30 m pixels and *CHRIS* 25 to 50 m pixels) are heavily compromised in terms of spatial resolution. Not only costly, aerial spectrometry images require more resource and time to store and analyze collected data if captured over a large region at maximum spatial and spectral resolution (Ashraf et al. 2010). As a result, aerial or satellite HRS or VHSR multispectral data (i.e., discrete and broad spectral bands) are generally the most available and appropriate data for cost-effective large scale land cover mapping.

Several aerial or satellite HRS or VHSR multispectral data contain information in NIR band (indispensable for the discrimination of class vegetation). Some have very high temporal resolution ability (such as RapidEye satellite 5 m pixels or Formasat-2 8 m pixels with 1 day temporal resolution), but generally time series of images of a given area are not always available with suitable acquisition dates and the cost generated by the purchase of multiple images is too expensive. Moreover high temporal data have generally a lower spatial extent, which increases the number of images for covering the studied zone.

Therefore, aerial or satellite HSR or VHSR multispectral data (with NIR information) acquired in single summer dates appears as the most optimized solution to obtain relevant results to support appropriate management solutions. Obviously, this solution can be changed depending on the economic issues for the territory lead by the stakeholders and/or managers in their territory. But, whatever the solution, two requirements are indisputable (i) the HSR and/or VHSR, and (ii) spectral information from visible to NIR bands.

Thematic data: Information from existing spatial thematic data available at an affordable cost for managers and/or stakeholders can be used in complement to image information. Several studies have shown that spatial thematic data (such as agricultural parcels boundaries or hydrography entities) can significantly improve image classification results by increasing accuracy levels (e.g., Raclot et al. 2005, Yang 2007). For a fine characterization of RALC, spatial thematic data with metric or decametric geometric precision are required.

CemOA : archive ouverte d'Irstea / Cemagref

2.2.2 Stage 2: Riparian area delineation. As development of riparian vegetation is widely influenced by hydrological regime, riparian area corresponds to the limits of the floodplain. Floodplain width can be estimated and extrapolated using the Strahler stream order (Souchon et al. 2000). Consequently, we consider the riparian area as a buffer around the hydrographic network data whose width depends on the Strahler stream order. The different widths according to the stream order are based on the results of Souchon et al., 2000 and were willingly overestimated in order to ensure to extract all relevant information in riparian areas (see table 1). Buffer is constructed using common GIS techniques integrated in ArcGis® 9.2 software.

Table 1. Riparian area widths according to the Strahler stream order

Strahler stream order	*Riparian area width*
1,2,3	400 m
4,5	800 m
6	1200 m
7	2400 m
8	4800 m

2.2.3 Stage 3: Delineating homogeneous mapping regions. In order to achieve satisfactory results, land cover classification over large territories requires to divide these territories in homogeneous regions.

Large territories (hundreds of km) can cover different geographical contexts resulting in different types of landscape (natural and/or anthropic). Geographical characteristics of the territory (geology, climate and relief) impact on the composition and the configuration of the landscape: for instance the natural vegetation and human activities are significantly different between mountainous regions and plain regions. In consequences, more geographically diversified is the territory, higher is the number of land cover classes to be described. Additionally, a given homogeneous region covered by two images at significantly different dates (end of spring, end of summer) will present distinct landscape features.

The number of homogeneous regions within the territory may even increase taking into account the different dates of data acquisition necessary to cover the studied area.

As a result, when classifying land cover over a large territory, the risk of confusion between classes increases as both probability of having similar spectral behaviours between distinct classes and probability of having heterogeneous spectral behaviour within the same class increase. Therefore, to confront these difficulties, it is necessary to stratify (divide) the study area in regions homogenous in terms of geographic characteristics and image acquisition date.

Eco-region boundaries are often used to stratify a study area with intent to map it, as an eco-region often correspond to specific patterns of land cover, urban settlement, agricultural variables, and resource-based industries (Gallant et al. 2004). Moreover eco-region data, that define homogeneous geographic regions with respect to geology, climate and relief characteristics, exists and is available in most of countries. Many studies have used ecoregion boundaries, for example Wasson et al. (2010) used

CemOA : archive ouverte d'Irstea / Cemagref

European ecoregion boundaries for analysing large-scale relationships between land cover indicators and river ecological status indicators in several countries in Europe.

As a result, the territory under study is divided in homogeneous mapping regions resulting from the intersection between eco-region boundaries and footprint boundaries of images having different acquisition dates or periods.

2.2.4 Stage 4: Pre-processing. This stage is dedicated to the preparation of the selected data sources (remotely-sensed and spatial thematic data) for mapping RALC on each mapping region (stage 5). Five pre-processing tasks are required.

First, geometric corrections of remotely-sensed data are needed to produce orthorectified images in order to remove image distorsions due to topographical variation in the surface of the earth and the elevation angle with which the image was acquired. Orthorectificaton method using a digital elevation model and ground control points are well-known in remote sensing analysis and implemented in the different image processing softwares.

Second, all data sources have to be projected in the same projected coordinate system in order to reduce classification errors linked to shift errors between the different data sources.

Third, image radiometric corrections are necessary to calibrate radiometric information between images acquired at different dates with the same sensor. This facilitates the transposition of classification rules between the different mapping regions. With relatively good weather conditions, Top Of Atmosphere (TOA) reflectance corrections based on pre-launch calibration coefficients, the solar zenith angle and the normalized solar irradiance reflectance (El Hajj et al. 2008) can be sufficient. With relative cloudy weather conditions, at-surface reflectance corrections based additionally on atmospheric parameters, including water vapour generally derived from MODIS sensor are required.

Fourth, to facilitate the integration of information from spatial thematic data in vector format in the OBIA, the production of raster thematic data is required. For that purpose, thematic vector data are overlapped as regard to their geometric precision in decreasing order of precision. Thus, when two or several vector data are intersected, the information with the most precise boundaries is conserved. The resulting intersected data vector is then converted to a raster layer with the same resolution as the image with the highest spatial resolution.

Finally, to prepare data for OBIA over each mapping region, pre-processed data are clipped according to the limits of the riparian corridor and each mapping region. A mosaic image is created when the mapping region is covered by several images.

These five pre-processing tasks are indispensable to map RALC. Obviously, the number of processing required depends on the geometric and radiometric precision quality with which remotely-sensed images are acquired. In addition to these five pre-processing tasks, the operator can implement other ones that he deems interesting for enriching or enhancing the information contained within the image.

2.2.5 Stage 5: OBIA on each mapping region. In this stage, an OBIA scheme, based on the FNEA embedded in Definiens Developer Earth® 7 (previously called eCognition) is implemented in order to extract RALC for each mapping region from information of pre-processed data (stage 4). The OBAI is first designed on one mapping

CemOA : archive ouverte d'Irstea / Cemagref

region before being implemented on other mapping regions with adjustments for taking into accounts the region specificities.

With this software fuzzy classification can be used to extract the objects of interest, at the scale of interest. For this purpose, a hierarchical image object network, able to segment images simultaneously at fine and coarse scales, is created from the multi-resolution algorithm, and a hierarchical classification scheme (class hierarchy) allowing the semantic organization of classes is built. A complete description of the multi-resolution algorithm (region-growing algorithm) can be found in (Baatz and Schäpe 2000, Benz et al. 2004, Definiens 2008). Through the process of classification, each image object is assigned, or not assigned at all, to a given class with a confidence degree and thus connected with the class hierarchy. The result of the classification is a network of classified image objects with concrete features, concrete relations to each other and concrete relations to the classes in the class hierarchy (Definiens 2008).

In this study, two specific steps were developed for designing the RALC OBIA scheme: (1) a method to build the class hierarchy and (2) a method to build the hierarchical image object network.

(1) The ***class hierarchy*** is constructed following a step by step top-down approach which divides iteratively the easiest to extract classes, according to available data sources (for example "water surface"/"land surface", then within "land surface" "soils with high vegetation" / "soils with low vegetation", and so on). Such an approach divides the feature space into finer and finer units that improve the extraction of interest classes (Safavian and Landgrebe 1991). Additionally, it promotes the use of simple rules that are easier to transpose to other mapping regions, and specifics classes of a given mapping region can be integrated in the class hierarchy without questioning its construction (Devaux 2005).

For each decision in the class hierarchy, rules are developed using fuzzy or crisp membership function based on one or several relevant features. Expert knowledge is crucial for the selection of relevant features, as well as the determination of threshold values of the chosen features for creation and classification rules (Benz et al. 2004). The relevant features are selected and their threshold values are determined by expert knowledge both from the knowledge already acquired by the operator and the knowledge acquired using the "Feature View" tool of Definiens. "Feature View" is a display method to visualize features over the entire project view. It helps the user to visualize the properties of image objects in a graphical way and therefore provides intuitive access to the peculiarity of a given feature over all image objects in a project (Definiens 2008). The software provides a huge number of radiometric, textural, contextual features for the classification and allows to calculate new features, such as vegetation indexes.

The number of final classes in the class hierarchy depends on both information made available by data sources, the knowledge of the operator and the minimum typology objective which defined for the planned application.

(2) The ***hierarchical image object network*** is composed of four segmentation levels in order to extract the objects of different sizes that are present in the riparian landscapes : from narrow and fragmented objects, such as sand banks along rivers, isolated tree vegetation or small impervious surfaces, to large objects such as agricultural plots or continuous urban areas. The hierarchical image object network is presented in figure 2.

CemOA : archive ouverte d'Irstea / Cemagref

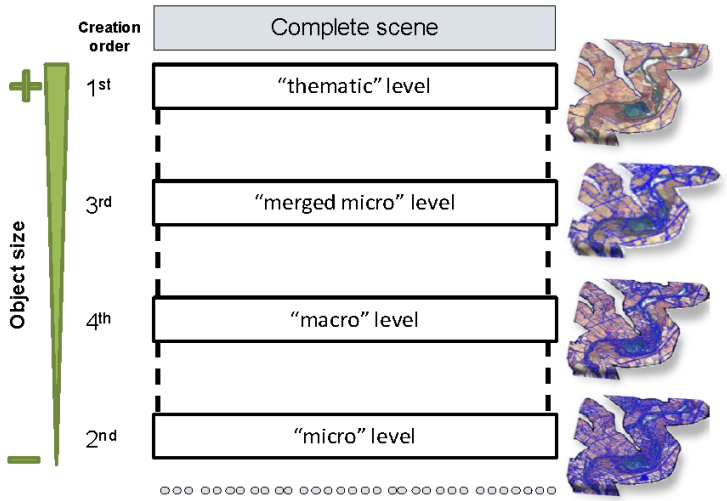


Figure 2. The hierarchical image object network designed for mapping land cover within riparian area.

The construction of this OBIA scheme is organized in three successive phases, each of them involving specific segmentation and classification processes. This structure has proved to be the most efficient to both integrate spatial thematic data information (*phase 1*) and identify and classify objects of heterogeneous shapes and sizes that are present in river corridor landscapes (*phases 2 and 3*). These phases and processes are summarized in the workflow figure 3.

In *Phase 1*, a segmentation level (called "thematic" level) is first created implementing the multi-resolution segmentation algorithm using only the thematic raster layer (considered as a thematic source by the software) that contains reliable information from the spatial thematic data. As a result, the studied zone is segmented according to the different limits resulting of the intersection between thematic data entities. Then, the resulting image objects are assigned –or not- to a thematic class using boolean rules (the "thematic" classification). *Phase 1* is a particular phase that allows to integrate external information on a number of objects of the mapping region and to construct a robust image object basis that will be used for the next phases.

In *Phase 2*, a new segmentation level (called "micro" level) is created. This level is devoted to detect narrow and fragmented objects in the riparian landscapes, using the multi-resolution segmentation algorithm. If different image sources are used, the construction of this level, based on the limits of the "thematic" level, relies in priority on the information from images with the highest spatial resolution. Each object of this level inherits classes of the "thematic classification", using class-related features at super-object level. Then a primary classification is applied to "No thematic" objects which are classified following the top-down approach developed for the construction of the class hierarchy. Generally the "No thematic class" is divided in two classes: the «water surface » class and the « land surface » class; then within the « land surface class »: the "soils with high vegetation" class and the "soils with low vegetation" class,

CemOA : archive ouverte d'Irstea / Cemagref

and so on until it is no longer possible to reliably separate and define simply new classes at this level of segmentation. Fuzzy rules defined by expert knowledge are used for this classification process (called "first image" classification).

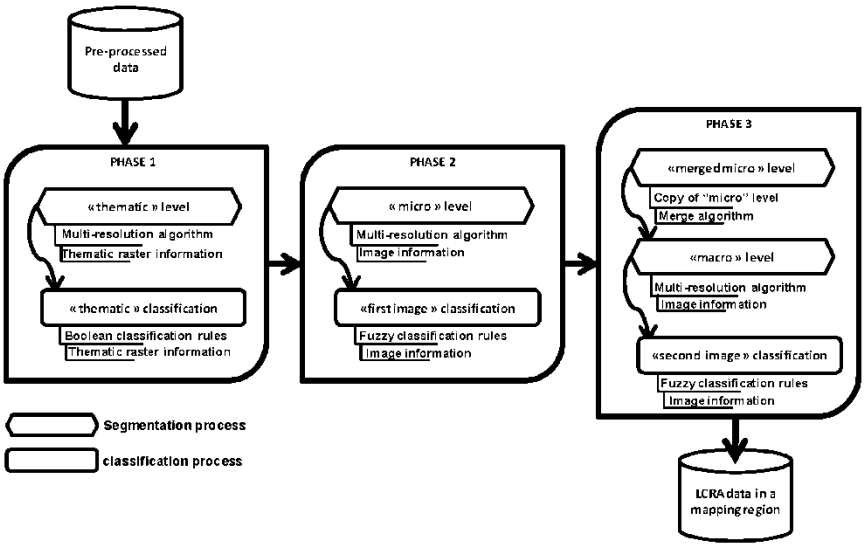


Figure 3. Workflow of the River Corridor Land Cover Object Based Image Analysis scheme.

In Phase 3, a new segmentation level (called "macro" level) is created, using the multi-resolution segmentation algorithm. This level aims at defining medium size objects while keeping isolated objects that have been delineated at the "micro" level. For this purpose a "merged micro" level is first created from two processes: a copy of the "micro" level and the merging of neighbour objects of a same class (according to the final typology obtained during the "first image" classification) into a single object of this class (Definiens, 2008). Then the "macro segmentation" is implemented between the limits of the "micro" level and the "merged micro level" (see figure 2), using spectral information from suitable images. Each object of this level first inherits classes of the "first image" classification contained in the "merged micro" level, using class-related features at super-object level. Then a second classification is applied to all classes that can be divided in more detailed classes following the same top-down approach developed for the construction of the class hierarchy. Fuzzy rules defined by expert knowledge are used for this classification process (called "second image" classification). It stops when no more classes can be reliably separated and simply defined at this level of segmentation.

CemOA : archive ouverte d'Irstea / Cemagref

2.2.6 Stage 6: Post-processing. In this stage, RALC classification results produced on the different mapping regions are first collected (according to the most detailed typology obtained) in vector format and combined on a unique vector layer.

Then, the typology is structured into a hierarchical classification nomenclature. Hierarchical structures have been demonstrated to be efficient for class definition and well-adapted to analyse land cover at different scales and resolutions (Anderson et al. 1976). This last point is of major importance to study the influence of land cover on river ecological status. Additionally to the advantage of providing significant information at different spatial levels, this structure is flexible to integrate additional information as it becomes available (for example a new type of land cover in a new region) and is extendable to higher hierarchical levels at a later stage (Gurnell et al. 1994). The land cover classification nomenclature is designed according to CORINE Land Cover (CLC) comprising three hierarchical levels (Bossard et al. 2000). Level 1 and 2 are the same as CLC nomenclature and Level 3 contains the detailed typology obtain in the OBIA process.

2.2.7 Stage 7: Map validation. One of the final goals of the river corridor land cover mapping is to allow the calculation of suitable river corridor pressure indicators. The reliability of the classification results is therefore of major importance and has to be assessed and quantified.

In this stage, accuracy assessment of land cover pressure map was performed. However, because the objective of the classification is not to obtain a geographically precise map as metric topographic map, only a semantic quality of the classification (i.e., assessment of objects nature) is performed using confusion matrix (Foody 2002). The confusion matrix is computed after grouping classes according to a pressure typology appropriate for the planned application.

Given that the features used for classification are calculated at the object scale, objects or polygons have been chosen as sampling units for the selection of control data (Zhan et al. 2005, Schöpfer and Lang 2006, Tiede et al. 2006, Grenier et al. 2008). However, land cover map results from the implementation of the multi-level OBIA scheme and contains different object sizes, from fine to large ones. Consequently, confusion matrix is computed using area of the selected control objects (expressed as a number of 0.5 m pixels).

To select control (reference) data with the objective of a spatially and thematically well-distributed sample in the studied zone (as suggested by (Congalton 2004), the entire sampling frame is divided into N equal grid cells. Next, a stratified random sampling is performed using grid cells as geographic strata (equal area for all strata). The number of cells (N) is equal to the desired sample size, in order to have at least one object of each class per grid cell. As suggested by (Congalton 1991) and confirmed by specific comparative analysis implemented during this study, 50 samples are collected for each class in order to build the confusion matrix.

Considering the size of the study area (and mapping regions), collecting field data for the control sample would be extremely labour intensive and time consuming. As suggested by (Zhu et al. 2000), selected control objects are photo-interpreted according to the pressure typology and using the image with highest spatial resolution as reference data. To maintain objectivity of photo-interpretation, classified map was not viewed

CemOA : archive ouverte d'Irstea / Cemagref

during the process and was realized by a "neutral" and experimented photo-interpreter that was not involved in the method development.

Four statistics are extracted from the confusion matrix to access land cover pressure map accuracy (Congalton 1991): (1) the overall accuracy that indicates the proportion of pixels that are correctly classified; (2) the user's accuracy that indicates to the map user which percentage of a mapped class actually corresponds to this class in the reality, (3) the producer's accuracy that indicates for each class which percentage of the reality is correctly classified in the map and (4) the Kappa index which is a statistical measure of the agreement, beyond chance, between map and reality.

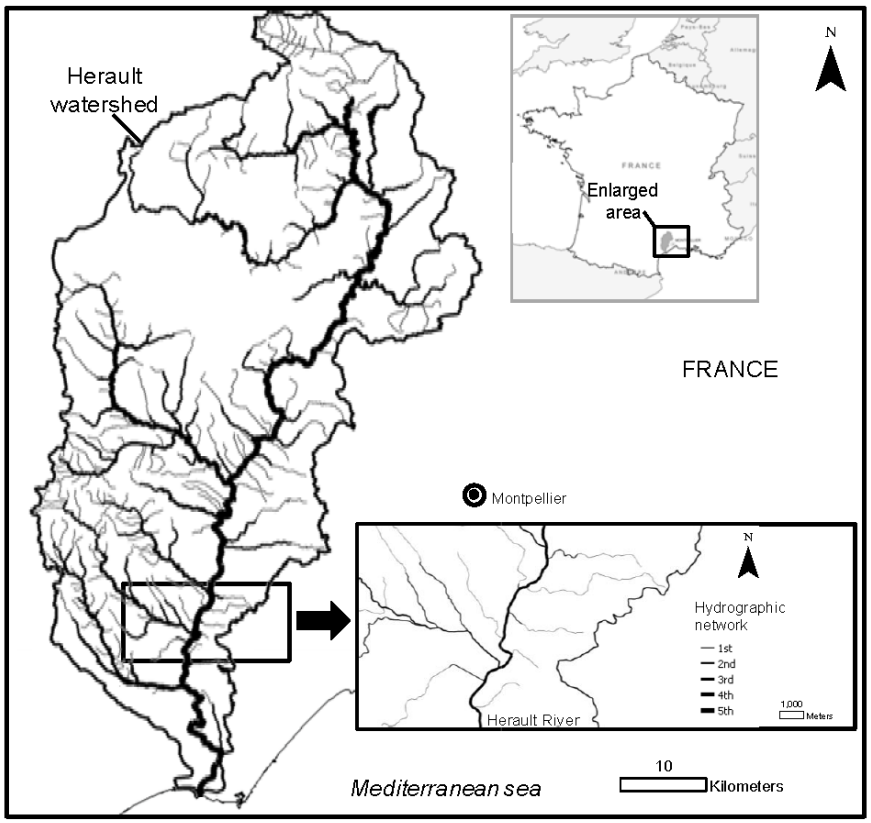


Figure 4. Study area: the Hérault river Basin, Languedoc-Roussillon, southern France. The river mouth is located at long. 3.443°W lat. 43.283°N, the basin area is 2650 km²; hydrographic network length is 1150 km

3. Test area

This study focuses on the riparian corridor of the Hérault River basin, located in Gard and Hérault Departments (Languedoc-Roussillon region, southern France), 30 km

CemOA : archive ouverte d'Irstea / Cemagref

west of the city of Montpellier (see figure 4). This basin includes 22 watersheds with a total area of about 2650 mm². It is submitted to Mediterranean climate, with precipitation rates above 2000 mm/year upstream and below 600 mm/year near the coast. Hérault River basin was chosen as a pilot study for the French territory because it presents a particularly contrasted landscape with a pressure gradient from upstream to downstream. While the upstream part of the basin is mainly composed of forests and semi-natural areas, the downstream part is strongly dominated by agriculture, especially vineyards, and by intensive urbanization originating from both the fast economic growth of the neighbouring city of Montpellier and the development of tourism (Balestrat et al. 2008). Along the downstream part of the basin the remaining riparian vegetation consists in narrow vegetated strips (< 30 m).

Stream orders over the study area range from the 1st to the 5th order according to the Straher's stream system ordination (Strahler, 1952) and the total stream length is 1150 km in which the 1st order accounts for about 48 %, 2nd order 19 %, 3rd order 16%, 4th order 7% and 5th order 11% according to the French hydrographic network database (BDCarthage®).

4. Implementation of classification procedure

The 7 stage classification procedure defined in the "method" part (see part 2) was implemented in order to produce a detailed and finely resoluted RALC map over the Hérault watershed. A specific land cover typology has been designed for studying pressures / impacts relationships. It is composed of six classes, representing the main human structures and riparian characteristics influencing stream ecological status: "C1-water surfaces" and "C6-semi-natural bare soils"; "C2-agricultural areas" and "C3-urban areas" considered as the main causes of stream ecological status alteration (Allan 2004); and "C4-tree vegetation", "C5-semi-natural herbaceous and shrub vegetation" constituting the main natural elements of the river corridor landscape that maintain biodiversity and regulate non-point source pollution (Naiman et al. 2005). The classification must allow, as a minimum requirement, to produce a land cover map according this 6 class typology.

4.1 Selection of data sources (stage 1)

Data sources were chosen according to their cost-effectiveness and their availability on French territories for managers and stakeholders. The comparative economic analysis of the different data is not presented in this paper : it has led to collect two types of HRS and VHRS multispectral remotely-sensed data in a single date and four spatial thematic data. Characteristics of the selected image data and thematic data are presented respectively in part 4.1.1 and 4.1.2. The main characteristics are summarized in table 2.

4.1.1 Characteristics of remotely-sensed data. Digital colour aerial photographs (orthophotos) and Spot5 multispectral satellite images were acquired on the Hérault watershed.

Aerial photographs provide the textural information required to detect narrow and fragmented cover types along rivers (Müller 1997, Neale 1997). In France, colour aerial

CemOA : archive ouverte d'Irstea / Cemagref

photographs with three visible bands (Blue, Green, and Red) and a 0.5 m spatial resolution are regularly acquired by IGN over the national territory using CDD digital cameras (Paparoditis et al. 2006). Acquisitions are generally realized in spring or summer to avoid large cloud cover. They are used for producing the digital orthophotos (5 × 5 km²) available from BDOrtho®. This database is updated every five years for each French Department. Radiometric equalizations are computed on raw images, before orthorectification, to minimize the radiometric heterogeneities between images that are due, among other causes, to surface bidirectional reflectance effects, to variations in atmospheric conditions and to changes in day of acquisition. Radiometric equalizations are computed between orthophotos of the same department (Paparoditis et al. 2006).

In addition to orthophotos, Spot5 XS (60 × 60 km²) scenes were acquired in order to get an information in the NIR band, which is required for the discrimination between vegetated covers using various vegetation indices (Müller 1997, Davis et al. 2002). Spot5 XS (60 × 60 km²) scenes have four spectral bands (B1 to B4) : bands B1 (green: 0.50–0.59 µm), B2 (red: 0.61–0.68 µm) and B3 (near infrared: 0.78–0.89 µm) have a spatial resolution of 10 m and band B4 (mid-infrared, 1.58–1.75 µm) of 20 m.

A total of 138 orthophotos and 2 Spot5 XS archive scenes was needed to cover the study area. Orthophotos were acquired during spring 2005 and 2006 (corresponding to the latest update for Hérault and Gard departments) and Spot5 XS archive images on the 14th of May 2004 and the 27th of April 2005. Among unclouded archive images, these dates were the most adapted in this study zone for the identification of several vegetation classes, particularly to discriminate grassland from other cultures. Spot5 XS images were acquired at standard level 3 correction. The level 3 products are corrected for radiometric and geometric distortions and orthorectified using ground control points and a digital elevation model (with a 25 m spatial resolution) to correct residual parallax errors due to topography.

4.1.2 Characteristics of spatial thematic data. Four existing spatial thematic data were collected that provide relevant information for the RALC classification.

First, some land cover classes were selected from the 2000 CORINE Land Cover (CLC) map produced by photo-interpretation of Landsat images (30 m). CLC is based on a standard 44 classes nomenclature organized in a 3 level hierarchy. Delimited CLC polygons, which are map unit of 25 ha minimum, either are homogenous zones or comprise a combination of zones with a recognizable structure (Bossard et al. 2000). Although CLC is not accurate enough for characterizing the narrow and fragmented objects along rivers, it is appropriate for the extraction of large objects within river corridor, especially large artificial surfaces. As a result, several artificial surface categories (continuous urban zones, industrial zones, artificial non-agricultural vegetated areas…) were selected as data source for primary classification of some objects within the river corridor.

Second, as roads are particularly difficult to extract automatically from remote-sensing images (because of their heterogeneous spectral behaviour), we collected the road network geodatabase (BDRoute®) produced by IGN. It provides information on the major roads in France (freeway and main roads).

Third, we used the surface hydrography entities produced by the hydrography database (BDCarthage®) in France (rivers, lakes and reservoirs).

CemOA : archive ouverte d'Irstea / Cemagref

Finally, as suggested by (Raclot et al. 2005) we incorporated thematic data for improving the discrimination between grassland and crops using single date remote sensing images. The Digital Parcel Register (DPR) from the Ministry in charge of Agriculture was used. This data set includes the boundaries of blocks of adjacent plots as declared by farmers benefiting from European subsidies for at least one crop in the frame of the European common agricultural policy. The block boundaries are annually photo interpreted and digitized by farmers on orthophotos. In order to ensure a temporal consistency with orthophoto, we used the 2006 DPR that only provides information on the limits of the agricultural plots (more recent DPR also provides information on the percentage of the various types of crops inside the block).

Table 2: The main characteristics of the data sources selected for mapping RALC in the French context.

	Data	*Spatial resolution or geometric precision*	*Spectral resolution or thematic information*	*acquisition date or production year*	*Producers*
Image data	Orthophotos	0.5 m	Visible bands (B, G, R)	Summer 2006	IGN®
	SPOT 5 XS	10 m (20 m for MIR)	G, R, NIR, MIR	14 May 2004 27 April 2005	Spotimage©
Spatial Thematic data	CLC (Corine Land Cover)	Map unit of 25 ha minimum	Artificial continuous areas	2000	European land cover program
	BDRoutes®	decametric	Roads	2006	IGN®
	BDCartharge®	decametric	Hydrographic surfaces	2006	IGN®
	DPR (Digital Parcel Register)	metric	block boundaries of adjacent parcels of farmers	2006	Ministry of agriculture

4.2 Delineation of riparian area and mapping regions (stages 2 & 3)

Riparian buffer limits were computed from the French hydrographic network database (BDCarthage®). Riparian widths were defined according to the Strahler stream order (see table 1). The resulting riparian buffer along the Herault hydrographic network has an overall 553.40 km² area.

Mapping regions were built according to French Hydro-EcoRegion (HER) and Spot5 XS footprint boundaries (see figure 5). The HER concept was developed to provide the first level typology of running waters, for assessment in accordance with the European WFD (European commission 2000, Wasson et al. 2004). HER are considered as regions with homogeneous landscape and a limited variability of environmental variables (climate, topology, geology). A detailed description of how French HER were delineated can be found in (Snelder et al. 2008).

Hérault River Basin limits intersect 3 HER: namely, from upstream to downstream, the Cévennes HER (HER8), a range of mountains with crystalline basement and high slope; the Great Causses HER (HER19) a group of median limestone plateaus; the

CemOA : archive ouverte d'Irstea / Cemagref

Mediterranean HER (HER6), an alluvial and littoral plain with a very low morphology and sparse rainfalls. The two SPOT5 XS images cover respectively the north (N) and the south (S) of the Herault river basin. We finally obtained 6 mapping regions: the HER8-N, HER6-N and HER6-S ones comprising the major part of the Hérault river corridor (see table 3).

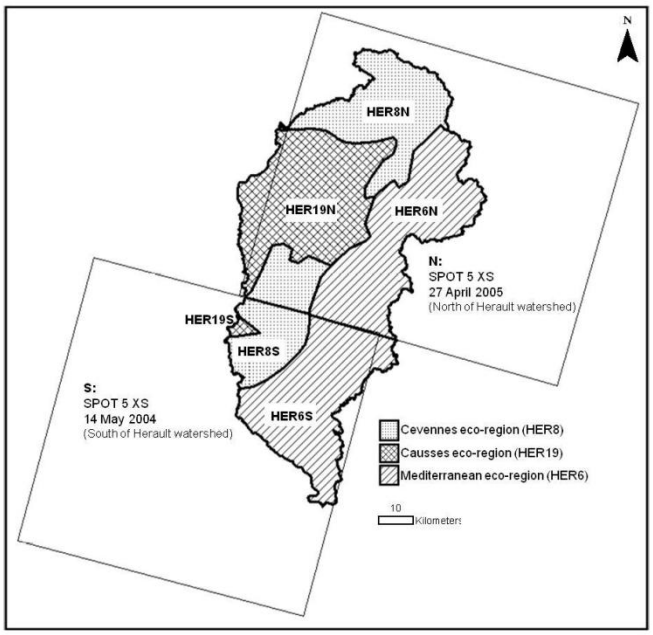


Figure 5. Stratification of the study area for mapping: delimitation of mapping regions.

Table 3. Stratification of the study area in mapping regions: river corridor buffer area per mapping region.

Mapping regions	Eco-region	Spot5 XS acquisition date	River Corridor surface (km²)	River Corridor surface (%)
HER8-N	**HER8**	25th of April 2005	140.28	25.35
HER6-S	**HER6**	14th of May 2004	137.27	24.80
HER6-N	**HER6**	25th of April 2005	136.04	24.58
HER8-S	**HER8**	14th of May 2004	71.41	12.90
HER19-N	**HER19**	25th of April 2005	62.70	11.33
HER19-S	**HER19**	14th of May 2004	5.70	1.04
	Total	2004-2005	553.40	100

16

CemOA : archive ouverte d'Irstea / Cemagref

4.3 Application of pre-processing (stage 4)

Due to the fact that remotely-sensed data were acquired with orthorectification corrections (see 4.1.1), no additional geometric corrections were needed.

First, all data sources were projected in the same projection coordinate system (extended Lambert 2).

Second, raw numerical counts of Spot5 XS scenes were converted to Top Of Atmosphere (TOA) in order to use radiometric information comparable between the 2 SPOT5 XS scenes. This took into account a) the calibration parameters for the acquisition date, which are absolute coefficients and the analog gain values, b) the solar zenith angle and c) the normalized solar irradiance reflectance (a complete description of the calculation of TOA reflectance can be found in (El Hajj et al. 2008)).

Third, the different vector data (selected artificial surface polygons from CLC, roads polylines from BDRoute®, surface entities from BDCarthage® and blocks of parcels from Digital Parcels Register DPR), were overlapped taking into account the geometric precision of each vector data source (in decreasing order of precision: CLC, BDCarthage® , BDRoute® and DPR). The intersected vector data was converted to a raster layer with the same resolution as orthophotos (0.5 m).

Fourth, orthophoto mosaics were created on each mapping region (several orthophotos were needed for each mapping region). The TOA converted Spot5 XS raster data, the thematic raster layer and orthophotos were then clipped according to the riparian buffer area and each mapping region boundaries..

Additionally, a new raster layer with a 6-m spatial resolution was produced on each mapping region from the riparian orthophoto mosaic. This re-sampling operation changes the texture of the objects: radiometric heterogeneity is reduced for some classes and specific textural characteristics can be identified for other classes that improve their discrimination (e.g. urban continuous or discontinuous areas).

These pre-processing tasks resulted in four pre-processed data to be used as inputs to the OBIA on each mapping region. They are illustrated in figure 6.

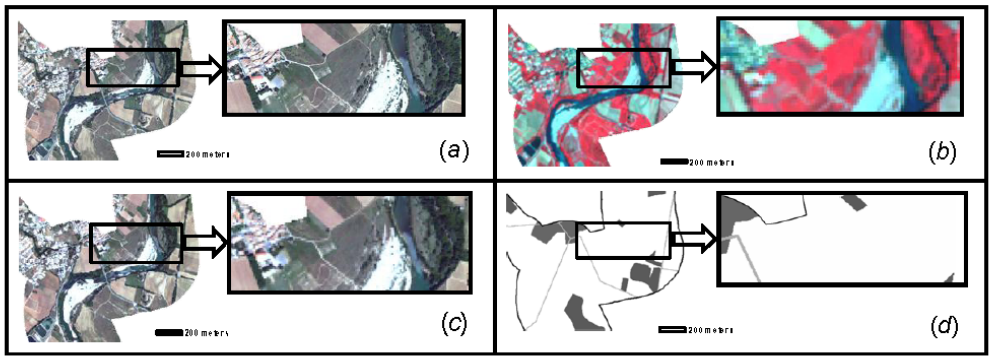


Figure 6. Illustration of pre-processed data incorporated in the OBIA: in (*a*) the orthophoto (0.5m), in (*b*) Spot5 XS (10 m), in (*c*) the degraded orthophoto (6 m) and in (*d*) the ancillary raster layer (0.5 m).

CemOA : archive ouverte d'Irstea / Cemagref

4.4 Application of OBIA on each mapping region (stage 5)

The OBIA was designed on the HER6-S mapping region. The complete class hierarchy is presented in figure 5. Multi-resolution segmentation parameters and features used for classication in the 3 phases of OBIA are presented in table 4. Features used in this study are radiometric, textural, contextual features provided by the software and specific spectral indexes (see table 5). For example, we used Normalized Difference of Vegetation Index (NDVI), which was derived from the red and near-infrared Spot 5 XS wavebands, because NDVI is a very useful feature for the differentiation of vegetation and non vegetation in urban areas (Carleer and Wolff 2006). Using the "Feature view" tool, we investigated suitable threshold values of NDVI feature for separating vegetation areas. The same approach is used for discriminating all classes.

The OBIA designed on the HER6-S were implemented on other mapping regions without defining new classes, simply by adjusting the threshold of classification parameters, except on DPR objects (in Phase 3 of the OBIA) which are divided into different agricultural classes. Because the agricultural practices can change according to the geographical context (HER), this part of the class hierarchy is different from one mapping region to another. For example, a plot from DPR characterized by an herbaceous vegetation spectral behaviour in a mountain area with high slope, such as HER8-N, is more likely to be "natural or semi-natural vegetation" for grazing than "crop" areas, as it would be elsewhere in the study zone.

Table 5. Description of the customized features used in classification procedure in the OBIA (G: Green, R: Red, NIR: Near InfraRed, MIR: Medium InfraRed from Spot5 XS

Name of Index	*Equation*	*Source*
Normalized difference Vegetation Index (NDVI)	$NDVI = \frac{(NIR - R)}{(NIR + R)}$	standard indice
Infra-Red Index (IRI)	$IRI \quad \frac{(NIR - MIR)}{(NIR + MIR)}$	standard indice
Normalized Difference Water Index (NDWI)	$NDWI = \frac{(R - MIR)}{(R + MIR)}$	(Gao 1996)
Brightness Index (BI)	$BI = \sqrt{\sum X_i^2}$	standard indice
Water Index (WI)	$WI = \frac{(ratioG - ratioNIR)}{(ratioG + ratioNIR)} \times MaxDiff$	This study

CemOA : archive ouverte d'Irstea / Cemagref

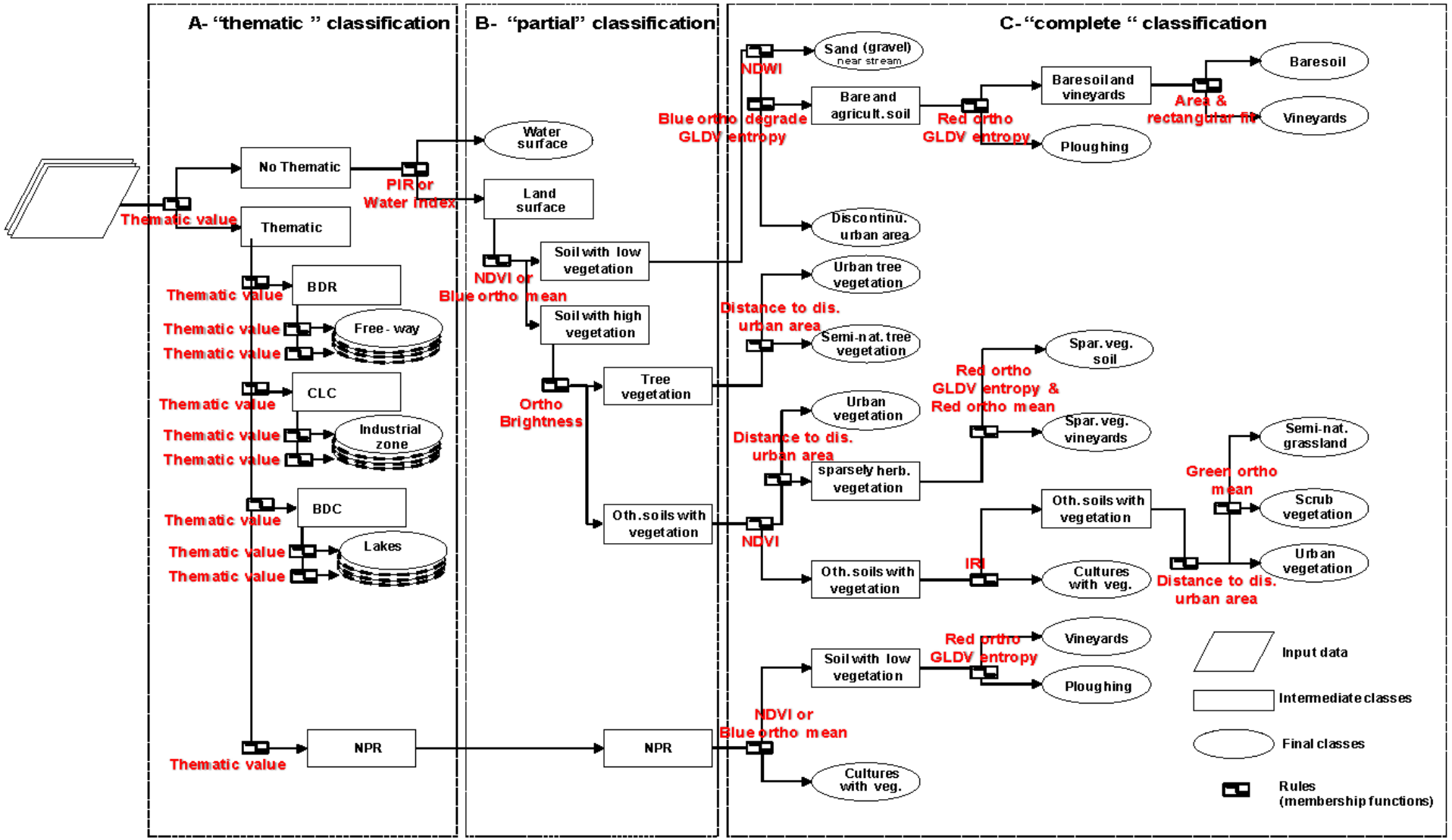
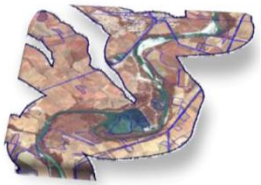
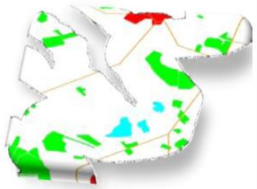
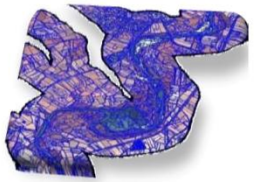
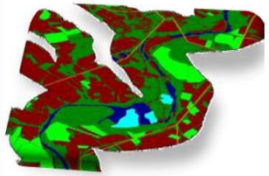
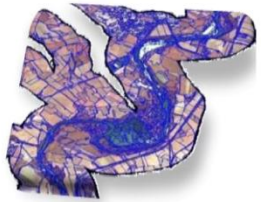
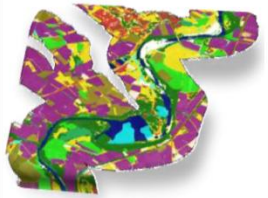


Figure 5. Class hierarchy developed in the OBIA scheme for mapping land cover in river corridors with Digital Parcel Register class hierarchy for HER6-S mapping region

CemOA : archive ouverte d'Irstea / Cemagref

Table 4. Object features used to classify objects from each phase during the OBIA (SXS for Spot5 XS; ortho for orthophotos; degr. ortho for degraded orthophotos; TL for thematic layer; and B for blue band, G for green band, R for red band, NIR for near infra red, GLDV for grey level difference vector)

	Segmentation					**Classification**						
	Parameters				Illustration	Featured used						Illustration
	Bands	Scale	Form	Compactness		ortho	degr. orhto	SXS	TL	Form	Context	
5.1 Thematic level	TL	100	0	0	**a)** 	not used	not used	not used	Mean	not used	-Similarity to classes	**b)** 
5.2 Structure level	ortho: $R_{(2)}$, $G_{(2)}$, $B_{(2)}$ SXS : $NIR_{(1)}$	10	0.1	0.5	**c)** 	-B mean -R GLDV entropy -BI	not used	-NDVI -WI -G ratio -NIR -mean	not used	not used	-Existence of super-object -Similarity to classes	**d)** 
5.3 LCRC level	SXS: MIR, NIR,G, V	300	0.1	0.8	**e)** 	-R mean -G mean -R GLDV entropy -B GLDV entropy -BI	-B GLDV entropy	-BI -NDVI -NDWI -IRI -G ratio	-Mean	-Area -Rectang. fit	-Existence of super-object -Real border to -Similarity to classes	**f)** 

CemOA : archive ouverte d'Irstea / Cemagref

4.5 Application of post-processing and validation of RALC map (stage 6& 7)

RALC classification results produced on the different mapping regions (according to the most detailed typology obtained) were collected in vector format and combined on a unique vector layer.

Then, the typology were structured into a hierarchical classification nomenclature designed according to CORINE Land Cover (CLC) following three hierarchical levels (Bossard et al. 2000). Level 1 and 2 are the same as CLC nomenclature. Level 3 does not contain the full detailed typology obtained in the OBIA process: some land cover classes in the class hierarchy are gathered in order to obtain a level 3 with more structured and comprehensible semantic information. For example, the "urban vegetation" class was grouped with the "discontinuous urban area" class and the "soil with sparsely vegetation" class was grouped with the "bare soil" class.

Finally, confusion matrixes were built according to the pressures typology (six classes) following the methodology detailed in part 2.2.7, for each mapping region as well as for the whole Hérault River watershed.

5. Results & Discussions

The method developed to produce a reliable, detailed and finely resolved Riparian Area Land Cover map over large and geographically contrasted territories, presented in (2), was implemented (4) on the study area (3). Resulting RALC maps are of major importance to understand the way pressures, either natural or anthropogenic, exerted within the river corridor influence the stream ecological status. We analyse and discuss successively the quality of classification results (5.1.), the specific contribution of each data source to these results (5.2.), the operational characteristics and performances of the method (5.3.), the key role of fine scale mapping for the quantification of relevant riparian spatial indicators (5.4.).

5.1 Quality of classification results

A relatively fine and detailed RALC map (22 classes in total, see level 3 of the hierarchical nomenclature scheme in table 6) has been obtained through OBIA of orthophotos and Spot5 XS images acquired each in single-date combined with spatial thematic data. Reliable distinction between different crop classes being hard to achieve, crop classes were aggregated in two main classes: "ploughing" and "cultures with vegetation". Similarly, because all grass pasture cannot be separated from natural areas covered with grass/forb, a "semi-natural grassland" class was created. An illustration of the resulting map can be seen on figure 6(*a*) according to level 3 of the hierarchical nomenclature scheme.

Table 7 shows the error matrix built according to the pressures typology (six categories) designed to study pressures / impacts relationships. The correspondence between level 3 land cover classes of the hierarchical nomenclature and the land cover pressure categories is presented in table 8. According to this matrix, 89.01% of pixels over the riparian area of Hérault watershed (553.40 km²) were correctly classified. This global accuracy is confirmed by the Kappa index value as 83.05% of pixels could be well classified according to a completely random validation process.

Highly accurate results were obtained for "C1-water surfaces", "C2-agricultural areas" and "C4-forested areas" (producer's accuracies are respectively 99.80%, 96.21%, 88.42%, and user's accuracy are respectively 98.09%, 89.33%, 97.53%). "C5-semi-natural herbaceous

CemOA : archive ouverte d'Irstea / Cemagref

vegetation" class presents a lower producer's accuracy (82.44%) especially because of confusions with "C2-agricultural areas" (5.50%).

The poorest accuracy was for both "C3-urban areas" and "C6-semi-natural bare soils" categories (user's accuracies are respectively 65.15% and 50.65%, producer's accuracies are respectively 59.25% and 34.53%) because of confusion between these two categories due to their similar spectral behavior (eg, rooftops, parkings / bare rocks, semi-natural sand ...). These errors mainly concern objects that are isolated in the landscape. The discrimination between "C3-urban areas" objects and other soil types with low or no vegetation (such as semi-natural bare soils) is based on a texture Harralick feature ("entropy GLDV all directions") extracted from the the degraded orthophoto blue spectral band (see figure 5). The more spatially heterogeneous and disordered is the spectral information contained within the object (common for urban objects), the higher is its texture feature. However, in the case of isolated objects this feature is less discriminating. Some "C6-semi-natural bare soils" isolated objects (for instance bare limestone common in mountainous region), are characterized by heterogeneous and spatially disorganized spectral information as they are sparsely covered by vegetation : they can be classified, according to the textural feature, as "C3-urban areas".

Some confusion also exist between the "C6-semi-natural bare soils" category and the "C2-agricultural areas" category because some objects of these categories have similar spectral behavior (e.g., ploughing/ semi-natural bare soils; vineyards/semi-natural bare soils...).

From these results it can be argued that this classification method provides a detailed and fine RALC map with globally accurate information about the six main land cover pressures. But progress is needed for "C3-urban areas" and "C6-semi-natural bare soils" categories: confusing a category likely to affect the river ecological status with a non-degrading category may have important implications in the study of pressures/impacts relationships. For "C6-semi-natural bare soils" isolated objects in mountainous region, with the information available in this study, contextual features as distance to roads might improve this discrimination and solve the difficulty.

Moreover, this classification does not account directly for shadows contained within images, especially within VHSR orthophotos. These shadows are mainly related to areas with high relief, the buildings and tree vegetation. "Shadow" objects are classified as "tree vegetation" at the micro level using the brightness index calculated from the orthophoto bands and then "tree vegetation" (thus including "shadow" objects) close to buildings areas are classified at the macro level into "urban vegetation" (with the objective of delineating discontinuous urban areas). This classification approach has several secondary effects : (i) it classifies "shadow" objects associated with high relief zones in "tree vegetation" which is rather consistent in narrow valley bottoms generally characterized by a dense riparian forest; (ii) it finally classifies "shadow" objects linked to buildings and to tree vegetation close to the buildings in "discontinuous urban area" which is rather consistent for shadows contained within the discontinuous urban area but not for those on the periphery, and (iii) it classifies "shadow" objects associated with tree vegetation in "tree vegetation" which is rather consistent for shadows contained in dense forested areas but not for those in the periphery. So even if the goal is not to obtain a geometrically precise map, progress is to be done to manage the shadows on the periphery of discontinuous urban area and dense tree vegetation objects so as not to overestimate the spatial indicators related to land cover categories affiliated with these shadows. However as generally orthophotos are acquired during the nadir period, "shadows" within images remain very limited.

CemOA : archive ouverte d'Irstea / Cemagref

Table 6. The 3-level land cover classification nomenclature for the River Corridor Land Cover mapping, and the corresponding data sources used to classify categories of the level 3 (rs: remote sensing images; th: thematic data; and rs+th: both data sources).

Level 1 (4 classes)	Level 2 (12 classes)	Level 3 (22 classes)	Data sources
1. Artificial surfaces	1.1. Urban fabric	1.1.1. Continuous urban fabric	th
		1.1.2. Artificial soil	rs
	1.2. Ind., com. and transport units	1.2.1. Industrial or commercial units	th
		1.2.2. Road and rail networks and associated land	th
	1.3. Mine, dump and construction sites	1.3.1. Mineral extraction sites	th
	1.4. Artificial non-agricultural vegetated areas	1.4.1. Sport and leisure facilities	th
	1.5. Road network	1.5.1. Local road	th
		1.5.2. Department road	th
		1.5.3. National road	th
		1.5.4. Freeway	th
2. Agricultural areas	2.1. Arable land	2.1.1. Ploughing	rs + th
		2.1.2. Cultures with vegetation	rs + th
	2.2. Permanent crops	2.2.1. Vineyards	rs + th
3. Forests and semi-natural areas	3.1. Forests	3.1.1. Tree vegetation	rs
		3.1.2. Tree and shrub vegetation	rs
	3.2. Shrub and/or herbaceous vegetation association	3.2.1. Semi-natural grassland	rs
		3.2.2. Pasture or brushwood	rs + th
	3.3. Open spaces with little or no vegetation	3.3.1. Bare soil	rs
		3.3.2. Sand	rs
5. Water bodies	5.1. Inland waters	5.1.1. Water courses	rs
		5.1.2. Lakes and reservoirs	th
	5.2 Marine waters	5.2.1. Sea and ocean	th

Table 7: Error matrix of the RCLC map on Herault River basin, expressed in number of 0.5m pixels.

	Classification						
Reference data	C1-water surf.	C2-agr. areas	C3-urb. areas	C4-For. areas	C5-herb. veg.	C6- bare soils	Total
C1-water surf.	**696510**	0	1405	0	0	0	697915
C2-agr. areas	0	**2991112**	7032	0	62292	48556	3108992
C3-urb. areas	0	30070	**131457**	0	16549	43776	221852
C4-For. areas	13110	2908	0	**392188**	35335	0	443541
C5-herb. veg.	425	206343	0	9938	**1017691**	0	1234397
C6-bare soils	0	117828	6188 9	0	0	**94770**	274487
Total	710045	3348261	201783	402126	1131867	187102	5981184
User's accuracy	98.09%	89.33%	65.15%	97.53%	89.91%	50.65%	
Producer's accuracy	99.80%	96.21%	59.25%	88.42%	82.44%	34.53%	
Overall accuracy	89.01%						
Kappa index	83.05%						

23

CemOA : archive ouverte d'Irstea / Cemagref

Table 8. The correspondence between land cover classes of level 3 hierarchical nomenclature (right) and land cover pressure categories (left)

Land cover pressure category	*Land cover classes of level 3*
C1-water surfaces	511, 512, 521
C2-agricultural areas	211, 212, 221
C3-urban areas	111, 112, 121, 122, 131, 141, 151, 152, 153, 154
C4-tree vegetation	311, 312
C5- herbaceous / shrub vegetation	321, 322
C6-semi-natural bare soil	331, 332

5.2 Contribution of the different data sources and other features

Each data source, either remote sensing imagery or thematic information, contributes in a specific way to the land cover classification. The last column in table 6 indicates for each class of the detailed typology (level 3 of the hierarchical nomenclature: 22 classes) the data sources used for its classification. Table 9 summarizes statistics about the contribution of each data source.

77 % of the objects, covering 65 % of the riparian area, have been classified based on information from orthophotos and Spot 5 XS satellite images. Spectral information from orthophotos (0.5 m) was necessary to detect narrow and fragmented objects in the riparian landscape and particularly to classify accurately "tree vegetation" objects (using brightness indices) and "vineyards" object (using Grey-Level Divergence Vector (GLDV) and Harralick texture in all directions for the red band information). Spectral information from Spot5 XS proved to be essential to improve the delineation of water bodies versus riparian tree vegetation, as well as to improve the classification of vegetation covers using indices based on near infra-red and medium infra-red information (such as NDVI Normalized Difference Vegetation Index or IRI InfraRed Index).

Some objects were classified using both information from DPR data and remote sensing data. The DPR data was paticularly used to treat agricultural objects on the study area. This data relates to about 15% of RALC objects and covers 32% of the Herault watershed riparian area. Generally, these objects are very difficult to identify using information from single-date images (Raclot et al. 2005). Depending on this date, on agricultural use and practices, these objects are at different stages of development, which result in different spectral behaviour and an increased risk of confusion with other classes (e.g., wheat crops in the early stages of development can be confused with semi-natural herbaceous vegetation). With the DPR data, the majority of agricultural areas (or pasture areas) are detected except those containing only vineyards, as this culture does not benefit from European subsidies. Once these areas are detected, information from the remotely-sensed data sources are used to discriminate the main agricultural use within plots (ploughing, vineyards, cultures with vegetation and grasslands). However the distinction between crops and grassland from the image information alone, was not easy (despite the choice of Spot5 XS images acquisition date). In the study area, few farmlands are dedicated to grasslands, because breeding is not highly developed, except in the

CemOA : archive ouverte d'Irstea / Cemagref

mountain regions. We therefore adapted the rules and the classification tree of these areas according to the hydro-ecoregion studied. Thus, DPR data was indispensable for improving classification results of agricultural areas and particularly for improving the discrimination between semi-natural herbaceous vegetation (grass, for, pasture) and the different cultures. Proper identification of grassy or herbaceous vegetation in riparian corridors is of major importance, as this cover can provide beneficial impact on stream status and its development may constitute appropriate restoration action, especially in agricultural regions (Lyons et al. 2000). Moreover, as the boundaries of blocks of plots are in perfect agreement with orthophotos, this thematic data significantly improved the segmentation results and so the classification results too. DPR data will be an important source of information in France for mapping agricultural areas, especially since the information on the different types of cultures is available (which was not the case during the achievement of this study).

The remaining objects (7%) were classified using information from other spatial thematic data defining a total of 11 classes. These classes concern a low percentage (2.98%) of the studied zone.

Thus, each information source was used alone or in combination with other information sources, to classify RALC. In this study, image data are the main source of information. But according to the study area, relative rates of contribution of the different sources of information may change. A study, not presented here, was recently carried out over 6000 km of hydrographic network located in the Normandie region in France : as this territory is strongly occupied by crops and areas dedicated to livestock, a larger proportion of objects were classified using DPR data compared to this study.

In addition to features directly related to the information borne by these different sources of data, context features (relating to classes of neighbors objects) have been used, particularly for better delineation of discontinuous urban areas: all objects of semi-natural vegetation (grass or trees) near areas of housing (a peripheral zone of 10m) were classified into urban vegetation.

Table 9. Contribution of data sources in the definition of classes contained in the level 3 of the hierarchical nomenclature

Data sources	Number of classes	Object Percentage	Area percentage
Remotely-sensed images	7	77.77%	65.43%
Ancillary data	11	6.53%	2.98%
Both data	4	15.70%	31.59%

5.3 Technical feasibility of the classification procedure

The implementation of this classification procedure on the Hérault watershed has proved some of its qualities. This procedure is:

(1) ***Reproducible*** ; Whatever the mapping region, and therefore whatever the geographic context, river size and image acquisition date, the overall quality (overall accuracy) of the classification was good to very good. Figure 7 presents the global accuracy measures for each of the 6 mapping regions within the study area : overall accuracy is greater than 89 % and Kappa index greater than 83 % except for HER19-N mapping region. Most errors on this

CemOA : archive ouverte d'Irstea / Cemagref

region are due to the frequent confusions between "C3-urban areas" and "C6-semi-natural bare soils" categories, as bare limestone objects sparsely covered by vegetation objects are widely present in this region (see 5.1).

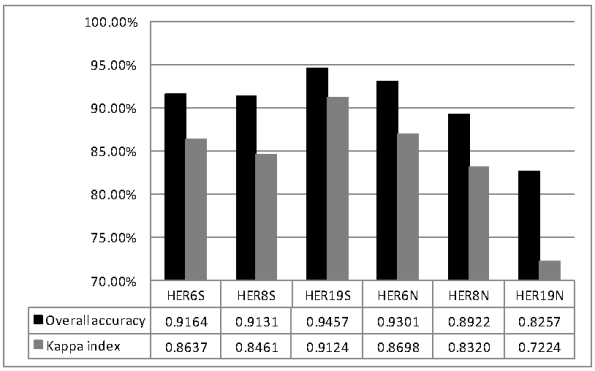
(2) ***Easy transferable*** : Transfering the method to another area has proved to be easy. Despite the diversity of landscapes in the study area, it has been possible to use the same class hierarchy (except for plot blocks of DPR data) on all mapping regions, without having to introduce new features at each adaptation. The operator only had to ajust the threshold values of the different features used to define classes. Class hierarchy modifications appeared necessary only for DPR children classes because agricultural practices of DPR are distinct according to the mapping region. However, as this class hierarchy was constructed following a top-down approach the integration of new specific classes (for instance new crops) for a given mapping region was facilitated. Thus. "Knowledge-based rules" classification technique ("expert" classification technique) revealed to be a robust, efficient and easily transferable approach for broad scale mapping of RALC, compared to the "supervised" technique commonly used in OBIA. The supervised Nearest Neighbour classification technique, that was implemented on the Herault River Reach in a previous study (Tormos et al. (In press) 2010), appeared to be highly time-consuming for the definition and in situ acquisition of training sample objects, and was hardly transferable to other areas. Moreover, the use of top-down approach promotes the use of simple rules that facilitate the appropriation of the methodology by new operators.

(3) ***Quickly applicable over large areas*** : The duration of OBIA processing were relatively low. The processing time, including the time necessary for adjustment of numerical rules, was 3 days with the use of a cluster machine containing 8 cores. The adjustment time obviously depends on the operator knowledge of the study area and on the land cover diversity. As observed by Lucas et al. 2007, "expert" classification technique is more flexible : all rules could be refined with the full control of the user, at any time in the classification process and, in most cases, without changing the class allocation of other objects (which is generally not the case with "supervised" classification technique). Moreover, while the "expertised" classification technique can be distributed on a cluster machine, thus drastically reducing the computation time, "supervised" classification technique cannot be distributed because it requires to work on the whole mapping region in order to use representative training samples.

(4) ***Portable and scalable*** (i.e., able to integrate new data sources) : The method can easily manage information from multiple data sources by resolving conflicts between these sources using fuzzy logic. This property has been particularly valuable for discriminating between the intermediate "soil with high vegetation" and "soil with low vegetation". Depending on the source data, SPOT 5 XS or orthophotos, information for assigning one of two classes in an object can be contradictory because they were not acquired in the same season (April-May for images Spot5 XS and Summer for orthophotos). This is particularly the case for agricultural land whose spectral behaviour varies during the season. Thus, fuzzy logic appears indispensable to treat multi-source information in this OBIA classification. Thanks to fuzzy logic, new data sources could easily be integrated and combined with the initial data sources. HSR Light Detection and Ranging (LIDAR) data, for example, that is becoming more and more available on large territories with reduced costs (Johansen et al. 2010), can provide additional information on the cover vertical dimension, highly relevant for accurate RALC mapping, especially for discriminating vegetation classes using vegetation heights (Whited et al. 2002, Goetz 2006, Johansen et al. 2010) and for discriminating artificial classes (C3) from

CemOA : archive ouverte d'Irstea / Cemagref

semi-natural bare soils (C6) using building heights. Moreover LIDAR information could improve the processing of "shadow effects", as LIDAR data are not affected by these effects.



	HER6S	HER8S	HER19S	HER6N	HER8N	HER19N
Overall accuracy	0.9164	0.9131	0.9457	0.9301	0.8922	0.8257
Kappa index	0.8637	0.8461	0.9124	0.8698	0.8320	0.7224

Figure 7. Overall accuracy and Kappa index according to a given mapping region

5.4 Interest of fine-scale RALC map for computing accurate riparian spatial indicators

In order to assess the capacity of the resulting RALC map to provide relevant spatial indicators of the riparian land cover, we compared the Linear Spatial Indicator (LSI) defined in (Tormos et al. (In press) 2010) computed from this RALC map (called here HSRI-map) with the same indicator computed from CORINE Land Cover (CLC) map produced from moderate remotely-sensed images with a minimum cartographic unit of 25 ha. The comparison was realized on 10 sections (A to J) of a 80 km Hérault River mainstream reach located in the basin plain (see figure 6(*b*)). The LSI corresponds to the proportion of area occupied by a given land cover category on a buffer close to the river (10 m width). It was computed according to 2 pressure categories: "semi-natural pressure" (aggregation of C4, C5 and C6 pressure categories) and "human pressure" (aggregation of C2 and C3 pressure categories).

Figure 8 shows the Linear Spatial Indicator computed from HSRI-map (right) and from CLC map (left) according to these two categories: "semi-natural pressure" in green and "human pressure" in red. LSI indicators are significantly different depending on the land cover map that is used to calculate them: while RALC HSRI-map shows that "semi- natural pressure" at the contact with the river is predominant for all sections except the most downstream (due to the presence of Agde city at the River mouth), CLC map would indicate that "human pressure" is dominant on all sections.

Not surprisingly, spatial resolution of land cover maps has a strong influence on our ability to quantify landscape indicators (Lattin et al. 2004, Tormos et al. (In press) 2010). Such result highlights the key need of HSRI-maps to quantify in a relevant way land cover pressures and riparian vegetation properties near the stream. Spatial indicators derived from HSRI maps provide accurate and original information on the presence and intensity of pressures close to the stream and on forested riparian strip attributes (uniformity, mean width and continuity) (Tormos et al. (In press) 2010).

CemOA : archive ouverte d'Irstea / Cemagref

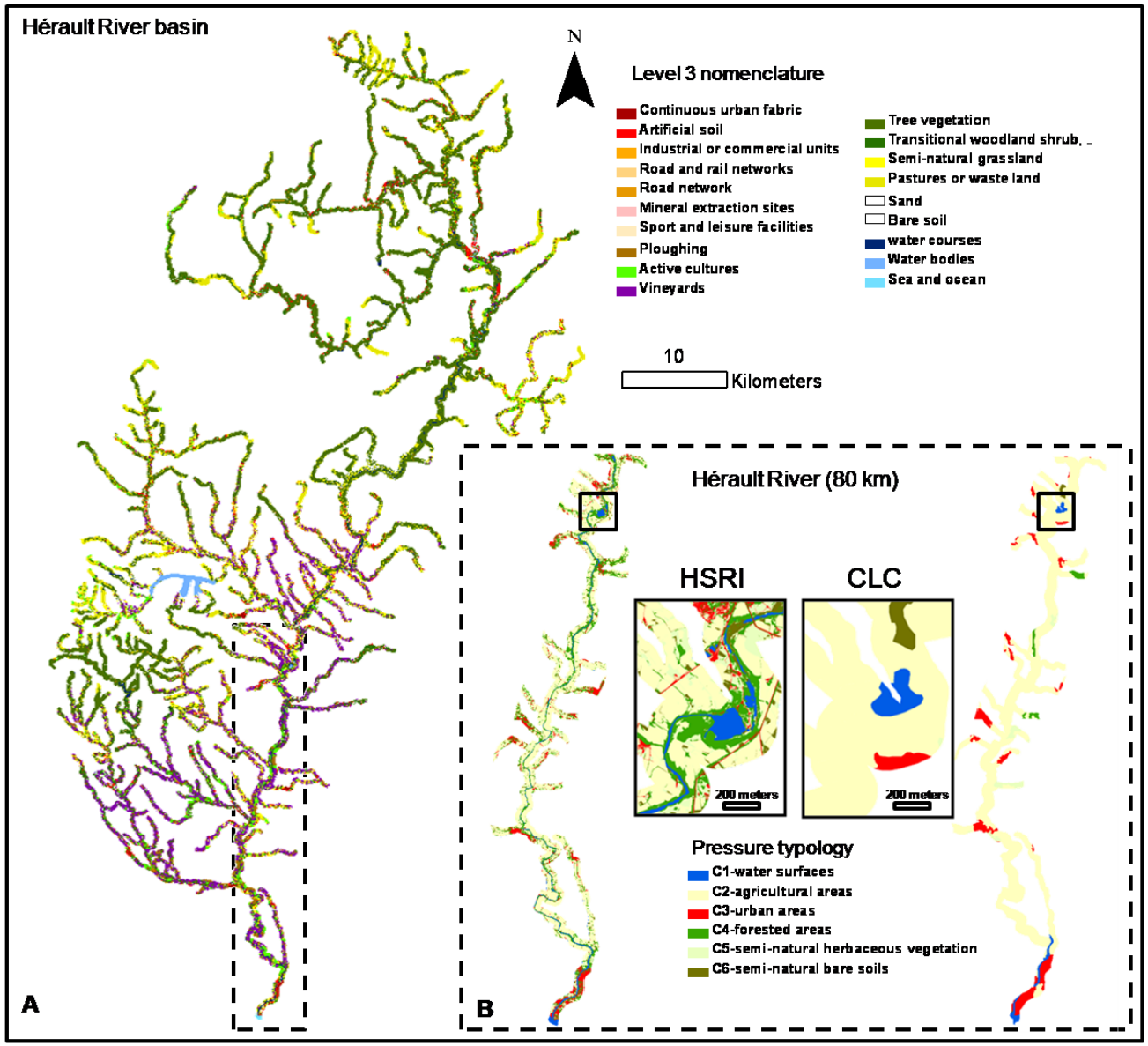


Figure 6. In (*a*) the River Corridor Land Cover (RALC) map on Hérault river basin according to the level 3 hierarchical classification nomenclature. In (*b*) the RALC map on 80 km of the Hérault River main streamaccording to the pressure typology derived from High Spatial Resolution Images (HSRI) on the left and CORINE Land Cover (CLC) on the right.

CemOA : archive ouverte d'Irstea / Cemagref

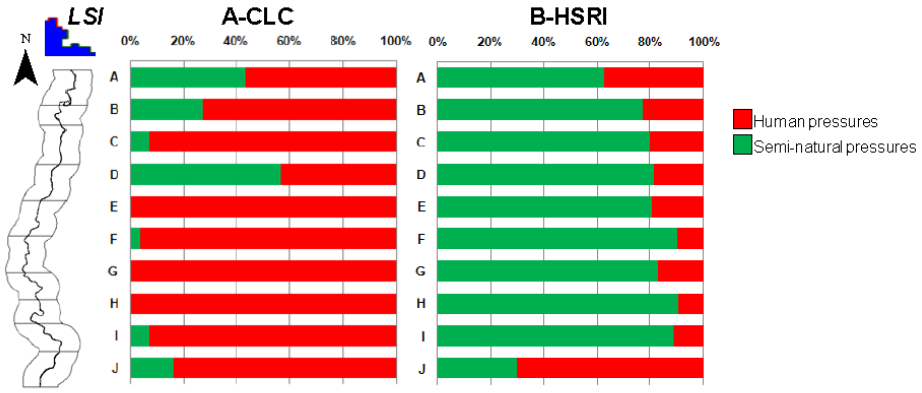


Figure 8. Linear Spatial Indicator (LSI): the surface area percentage of a given land cover category extracted on a 10-meter buffer around a river reach. In A from CORINE Land Cover (CLC) map, in B from HSRI-map produced by our classification procedure using orthophotos, SPOT 5 XS and thematic spatial data.

6. Conclusions and future works

In this research we designed an accurate, spatially extensive, time-efficient and cost-effective classification procedure, able to combine information from HSR or VHSR optical remotely-sensed data and existing spatial thematic data, for producing fine scale and detailed Riparian Area Land Cover maps. Such maps are relevant to study large scale response of river ecological status indicators to human pressures and riparian vegetation indicators.

This classification procedure consists in seven stages from the selection of data sources to the accuracy assessment of the resulting map. Its implementation on the Hérault watershed (1150 km of hydrographical network and 553 km² of riparian area) produced a detailed (22 classes) and fine scale RALC map providing accurate information (overall accuracy higher than 89 %) on the main land cover categories in the river corridor that generate pressures on river ecological status ("water surfaces", "semi-natural bare soil", "agricultural areas", "urban areas", "tree vegetation" and "semi-natural herbaceous and shrub vegetation").

Multi-level Object Based Image Analysis (OBIA) following a top-down approach for the class hierarchy construction and expert knowledge fuzzy numerical rules for classification, were clearly required to ensure a reliable and rapid transferability of the classification procedure over different mapping regions defined by an homogeneou geographic context and homogeneous image acquisition period for each remote sensing sensor.

Fuzzy logic was indispensable to combine the information from multiple sources data. It allows the classification procedure to be portable and scalable by facilitating the integration of new data sources, an important quality in the context of rapidly changing remote-sensing imagery market and of acquisition of new thematic data sources by managers or stakeholders.

This fine scale RALC map, processed with a GIS, ensures a huge flexibility for accurate calculation of landscape metrics to characterize human pressures and riparian influences on aquatic ecosystems. It proved to be incomparably more accurate than CORINE Land Cover map for the quantification of riparian spatial indicators. It can be used to study the mechanisms through which land cover along river corridors impacts stream ecosystems, and could improve the prediction of river ecological status.

CemOA : archive ouverte d'Irstea / Cemagref

Given these results, two challenges emerge for a better understanding of the relationships between land cover pressures and stream ecological status.

The first challenge focuses on spatial riparian indicators. It lies in (i) the definition of a set of complementary spatial indicators related to the riparian functions and to the mechanisms of land cover pressures on stream conditions (chemical, physical or biotic) and (ii) the development of GIS techniques to automatically build these indicators from RALC maps. Sensitivity of these spatial indicators to the uncertainty in land cover classification will have to be analysed in detail and the uncertainty values will have to be associated to these indicators. The inherent properties of fuzzy logic, that indicates how the classes are confounded within the object (Benz et al. 2004) and thus provides a spatialized information on uncertainties that confusion matrix do not provide, may be a high value tool for assessing uncertainty of spatial indicators .

The second challenge consists in the spatialized modelling of the stream ecological (or chemical) response to riparian pressure indicators derived from the fine RALC maps and to other pressure variables, in order to identify the width and composition of riparian vegetation that should be recommended at regional or national scale to maintain river corridor functions for stream integrity and ecological status.

Acknowledgments

This study has been made possible by a PhD grant from the French Ministry of Research. The work has been partly realised in the framework of a research contract between ONEMA and Cemagref. Acquisition of remote sensing data benefited from the support of the ISIS programme of the French national space agency CNES (Spot5 images, distribution Spot Image), while IGN airborne orthophotos were made available by the French Ministry of Agriculture.

References

AKASHEH, O. Z., NEALE, C. M. U., and JAYANTHI, H., 2008, Detailed mapping of riparian vegetation in the middle Rio Grande River using high resolution multi-spectral airborne remote sensing. J. Arid Environ., 72, 1734-1744.

ALLAN, J. D., 2004, Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. Annual Review of Ecology, Evolution, and Systematics, 35, 257-284.

ANDERSON, J. R., HARDY, E. E., ROACH, J. T., and WITMER, R. E., 1976, A Land Use And Land Cover Classification System For Use With Remote Sensor Data, Geological Survey Professional.

APAN, A. A., 2002, Mapping and analysis of changes in the riparian landscape structure of the Lockyer Valley catchment, Queensland, Australia. Landscape and Urban Planning, 59, 43-57.

APLIN, P., and SMITH, G. M., 2008, Advances in Object-Based Image Classification: ISPRS Congress, Proceedings of Commission VII.

ASHRAF, S., BRABYN, L., HICKS, B. J., and COLLIER, K., 2010, Satellite remote sensing for mapping vegetation in New Zealand freshwater environments: A review. New Zealand Geographer, 66, 33-43.

BAATZ, M., and SCHÄPE, A., 2000, Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. Angewandte Geographische Informationsverarbeitung, 12, 12-23.

BALESTRAT, M., CHÉRY, J.-P., and VALETTE, E., 2008, Suivi des changements d'occupation et d'utilisation du sol pour la compréhension des dynamiques périurbaines: XLV colloque de l'Association Régionale de Langue Française.

CemOA : archive ouverte d'Irstea / Cemagref

BECKER, B. L., LUSCH, D. P., and QI, J., 2007, A classification-based assessment of the optimal spectral and spatial resolutions for Great Lakes coastal wetland imagery. Remote Sensing of Environment, 108, 111-120.

BENZ, U. C., HOFMANN, P., WILLHAUCK, G., LINGENFELDER, I., and HEYNEN, M., 2004, Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing, 58, 239-258.

BLASCHKE, T., 2009, Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, Available online 28 August 2009.

BLASCHKE, T., and HAY, G. J., 2001, Object-oriented Image Analysis and Scale-space: Theory and Methods for Modeling and Evaluating Multiscale Landscape Structures. International Archives of Photogrammetry and Remote Sensing, 34, 22-29.

BOSSARD, M., FERANEC, J., and OTAHEL, J., 2000, CORINE land cover technical guide-Addeendum 2000, European Environment Agency.

CARLEER, A. P., and WOLFF, E., 2006, Urban land cover multi-level region-based classification of VHR data by selecting relevant features. International Journal of Remote Sensing, 27, 1035 - 1051.

CONGALTON, R., 2004, Putting the map back in map accuracy assessment. A peer-reviewed chapter IN: Lunetta, R.S., and J.G. Lyon (Eds.), Remote Sensing and GIS Accuracy Assessment, CRC Press, Boca Raton, FL 304p.

CONGALTON, R. G., 1991, A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37, 35-46.

CONGALTON, R. G., BIRCH, K., JONES, R., and SCHRIEVER, J., 2002, Evaluating remotely sensed techniques for mapping riparian vegetation. Computers and Electronics in Agriculture, 37, 113-126.

COULTER, L., STOW, D., HOPE, A., O'LEARY, J., TURNER, D., LONGMIRE, P., PETERSON, S., and KAISER, J., 2000, Comparison of high spatial resolution imagery for efficient generation of GIS vegetation layers. Photogrammetric Engineering and Remote Sensing 66, 1329-1335.

DAVIS, P. A., STAID, M. I., PLESCIA, J. B., and JOHNSON, J. R., 2002, Evaluation of airborne image data for mapping riparian vegetation within the Grand Canyon, U.S. Department of the Interior, U.S. Geological Survey.

DEFINIENS, 2008, Definiens Developper 7.0® User Guide, Definiens AG.

DEVAUX, N., 2005, Détection automatique et cartographie de l'habitat rural dans les pays en développement par télédétection et analyse spatiale: 7èmes Rencontres de Théo Quant., p. 26-28.

DURIEUX, L., KROPACEK, J., DE GRANDI, G. D., and ACHARD, F., 2007a, Object-oriented and textural image classification of the Siberia GBFM radar mosaic combined with MERIS imagery for continental scale land cover mapping. International Journal of Remote Sensing, 28, 4175 - 4182.

DURIEUX, L., LAGABRIELLE, E., and NELSON, A., 2007b, A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. ISPRS Journal of Photogrammetry and Remote Sensing, 63, 399-408.

EL HAJJ, M., BÉGUÉ, A., LAFRANCE, B., HAGOLLE, O., DEDIEU, G., and RUMEAU, M., 2008, Relative Radiometric Normalization and Atmospheric Correction of a SPOT 5 Time Series. Sensors, 8, 2774-2791.

EUROPEAN COMMISSION, 2000, Directive 2000/60/EC, Establishing a framework for community action in the field of water policy, European Commission PE-CONS.

FOODY, G. M., 2002, Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80, 185-201.

FRIMPONG, E. A., SUTTON, T. M., LIM, K. J., HRODEY, P. J., ENGEL, B. A., SIMON, T. P., LEE, J. G., and LE MASTER, D. C., 2005, Determination of optimal riparian forest buffer dimensions for stream biota- landscape association models using multimetric and multivariate responses. Canadian Journal of Fisheries and Aquatic Sciences, 62, 1-6.

CemOA : archive ouverte d'Irstea / Cemagref

GALLANT, A. L., LOVELAND, T. R., SOHL, T. L., and NAPTON, D. E., 2004, Using an Ecoregion Framework to Analyze Land-Cover and Land-Use Dynamics. Environmental Management, 34, S89-S110.

GAO, B. C., 1996, NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment, 58, 257-266.

GOETZ, S. J., 2006, Remote sensing of riparian buffers: Past progress and future prospects. Journal of the American Water Resources Association, 42, 133-143.

GOETZ, S. J., WRIGHT, R., SMITH, A. J., ZINECKER, E., and SCHAUB, E., 2003, Ikonos imagery for resource management: Tree cover, impervious surfaces and riparian buffer analyses in the mid-Atlantic region. Remote Sens. Environ, 88, 195-208.

GRENIER, M., LABRECQUE, S., BENOIT, M., and ALLARD, M., 2008, Accuracy assement method for wetland object-based classification Geographic Object Based Image Analysis (GEOBIA 2008), ISPRS Commision IV.

GURNELL, A. M., ANGOLD, P., and GREGORY, K. J., 1994, Classificatin of river corridors: Issues to be addressed in developing in operational methodology. Aquatic Conservation: Marine and Freshwater Ecosystems, 4, 219-231.

HOLLENHORST, T., HOST, G., and JOHNSON, L., 2006, Scaling issues in mapping riparian zones with remote sensing data: quantifying errors and sources of uncertainty. . In Scaling and uncertainty analysis in ecology, edited by J. Wu, B. Jones, H. Li, and O. Loucks (Dordrecht, The Netherlands: Springer), pp. 275-295.

IRONS, J. R., MARKHAM, B. L., NELSON, R. F., TOLL, D. L., WILLIAMS, D. L., LATTY, R. S., and STAUFFER, M. L., 1985, The effects of spatial resolution on the classication of Thematic Mapper data. International Journal of Remote Sensing, 6, 1385-1403.

IVITS, E., KOCH, B., BLASCHKE, T., JOCHUM, M., and ADLER, P., 2005, Landscape structure assessment with image grey-values and object-based classification at three spatial resolutions. International Journal of Remote Sensing, 19.

JOHANSEN, K., and PHINN, S., 2006, Mapping Structural Parameters and Species Composition of Riparian Vegetation Using IKONOS and Landsat ETM+ Data in Australian Tropical Savannahs. Photogram. Eng. Remote Sensing, 72, 71-79.

JOHANSEN, K., PHINN, S., LOWRY, J., and DOUGLAS, M., 2008, Quantifying indicators of riparian condition in Australian tropical savannas: integrating high spatial resolution imagery and field survey data. International Journal of Remote Sensing, 29, 7003 - 7028.

JOHANSEN, K., PHINN, S., and WITTE, C., 2010, Mapping of riparian zone attributes using discrete return LiDAR, QuickBird and SPOT-5 imagery: Assessing accuracy and costs. Remote Sensing of Environment, 114, 2679-2691.

KRESSLER, F. P., KIM, Y. S., and STEINNOCHER, K. T., 2003, Object-oriented land cover classification of panchromatic KOMPSAT-1 and SPOT-5 data. Geoscience and remote sensing, 19, 263-269.

LATTIN, P., WIGINTON, P., MOSER, T. J., PENISTON, B. E., LINDEMAN, D. R., and OETTER, D. R., 2004, Influence of remote sensing imagery source on quantification of riparian land cover/land use. Journal of the American Water Resources Association, 40, 215-227.

LATTY, R. S., and HOFFER, R. M., 1981, Computer-based Classification Accuracy Due to the Spatial Resolution Using Per-point Versus Per-field Classification Technique: Symposium of Machine Processing of Remotely Sensed Data,, p. 384-392.

LENNARTZ, S. P., 2004, Classifying and mapping forest cover types using ikonos imagery in the northeastern United States: Proceedings of the ASP R S.

LUCAS, R., ROWLANDS, A., BROWN, A., KEYWORTH, S., and BUNTING, P., 2007, Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. ISPRS Journal of Photogrammetry and Remote Sensing, 62, 165-185.

LYONS, J., TRIMBLE, S. W., and PAINE, L. K., 2000, Grass versus trees: Managing riparian areas to benefit streams of central North America. Journal of the American Water Resources Association, 36, 919-930.

CemOA : archive ouverte d'Irstea / Cemagref

MÜLLER, E., 1997, Mapping riparian vegetation along rivers: old concepts and new methods. Aquatic Botany, - 58, - 437.

MÜLLER, E., DÉCAMPS, H., and DOBSON, M., 1993, Contribution of space remote sensing to river studies. Freshwater Biology, 29, 301-312.

NAIMAN, R. J., DÉCAMPS, H., and MCCLAIN, M. E., 2005, Riparia : ecology, conservation, and management of streamside communities (Amsterdam ; Boston: Elsevier Academic, ©2005).

NEALE, C. M. U., 1997, Classification and mapping of riparian systems using airborne multispectral videography. Restoration Ecology, 5, 103-112.

NOONAN, M., and CHAFER, C., 2007, A method for mapping the distribution of willow at a catchment scale using bi-seasonal SPOT5 imagery. Weed Research, 47, 173-181.

PANIGRAHY, S., and SHARMA, S. A., 1997, Mapping of crop rotation using multidate Indian Remote Sensing Satellite digital data. ISPRS Journal of Photogrammetry and Remote Sensing, 52, 85-91.

PAPARODITIS, N., SOUCHON, J. P., MARTINOTY, G., and PIERROT-DESEILLIGNY, M., 2006, High-end aerial digital cameras and their impact on the automation and quality of the production workflow. ISPRS Journal of Photogrammetry and Remote Sensing, 60, 400-412.

PHINN, S. R., STOW, D. A., FRANKLIN, J., MERTES, L. A. K., and MICHAELSEN, J., 2003, Remotely sensed data for ecosystem analyses: Combining hierarchy theory and scene models. Environmental Management, 31, 429-441.

RACLOT, D., COLIN, F., and PUECH, C., 2005, Updating land cover classification using a rule-based decision system. International Journal of Remote Sensing, 26, 1309 - 1321.

ROY, A. H., FREEMAN, B. J., and FREEMAN, M. C., 2007, Riparian influences on stream fish assemblage structure in urbanizing streams. Landscape Ecology, 22, 385-402.

SAFAVIAN, S. R., and LANDGREBE, D., 1991, A survey of decision tree classifier methodology. IEEE Trans. Systems Man Cybernet, 21, 660-674.

SCHÖPFER, E., and LANG, S., 2006, Object fate analysis - A virtual overlay method for the categorisation of object transition and object-based accuracy assessment: Proceedings of the 1st International Conference on Object Based Image Analysis (OBIA 2006).

SHUFT, M. J., MOSER, T. J., WIGINGTON JR, P. J., STEVENS JR, D. L., MCALLISTER, L. S., CHAPMAN, S., and ERNST, T. L., 1999, Development of landscape metrics for characterizing riparian-stream networks. PE&RS - Photogrammetric Engineering and Remote Sensing, 65, 1157-1167.

SNELDER, T., PELLA, H., WASSON, J.-G., and LAMOUROUX, N., 2008, Definition Procedures Have Little Effect on Performance of Environmental Classifications of Streams and Rivers. Environmental Management, 42, 771-788.

SOUCHON, Y., ANDRIAMAHEFA, H., COHEN, P., BREIL, P., PELLA, H., LAMOUROUX, N., MALAVOI, J.-R., and WASSON, J.-G., 2000, Régionalisation de l'habitat aquatique dans la bassin de la Loire, Cemagref Lyon BEA/LHQ, Agence de l'eau Loire Bretagne.

TIEDE, D., LANG, S., and HOFFMANN, C., 2006, Supervised and forest type-specific multi-scale segmentation for a one-level-representation of single trees: 1st International Conference on Object Based Image Analysis (OBIA 2006).

TORMOS, T., KOSUTH, P., DURRIEU, S., VILLENEUVE, B., and WASSON, J.-G., (In press) 2010, Improving the quantification of land cover pressures on stream ecological status at riparian scale using High Spatial Resolution Imagery. Physics and Chemistry of the Earth.

WASSON, J.-G., CHANDESRIS, A., PELLA, H., and BLANC, L., 2004, Les hydro-écorégions : une approche fonctionnelle de la typologie des rivières pour la Directive cadre européenne sur l'eau. Ingénieries-EAT, 40, 3-10.

WASSON, J. G., VILLENEUVE, B., IITAL, A., DOBIASOVA, M., BACIKOVA, S., PELLA, H., MENGIN, N., MURRAY-BLIGH, J., TIMM, H., and CHANDESRIS, A., 2010, Large scale relationships between basin and riparian land cover and ecological status of European rivers: examples with invertebrate indices from France, Estonia, Slovakia and United Kingdom. Freshwater Biology, 55, 1465 - 1482.

CemOA : archive ouverte d'Irstea / Cemagref

WHITED, D., STANFORD, J. A., and KIMBALL, J. S., 2002, Application of airborne multispectral digital imagery to quantify riverine habitats at different base flows. River Research and Applications, 18, 583-594.

YAN, G., MAS, J. F., MAATHUIS, B. H. P., XIANGMIN, Z., and VAN DIJK, P. M., 2006, Comparison of pixel-based and object-oriented image classification approaches - a case study in a coal fire area, Wuda, Inner Mongolia, China. International Journal of Remote Sensing, 27, 4039 - 4055.

YANG, X., 2007, Integrated use of remote sensing and geographic information systems in riparian vegetation delineation and mapping. International Journal of Remote Sensing, 28, 353 - 370.

ZHAN, Q., MOLENAAR, M., TEMPFLI, K., and SHI, W., 2005, Quality assessment for geo-spatial objects derived from remotely sensed data. International Journal of Remote Sensing, 26, 2953-2974.

ZHOU, W., and TROY, A., 2008, An object-oriented approach for analysing and characterizing urban landscape at the parcel level. International Journal of Remote Sensing, 29, 3119 - 3135.

ZHU, Z., YANG, L., STEHMAN, S. V., and CZAPLEWSKI, R. L., 2000, Accuracy Assessment for the U.S. Geological Survey Regional Land-Cover Mapping Program: New York and New Jersey Region. Photogrammetric Engineering and Remote Sensing, 66, 1425-1435.

CemOA : archive ouverte d'Irstea / Cemagref

III. L'incertitude de la carte d'occupation du sol dans le corridor rivulaire

Dans cette thèse deux techniques ont été mises en place pour quantifier la qualité sémantique (i.e., nature des objets) de la carte d'occupation du sol dans le corridor rivulaire (cf. Chapitre 5 :III.7) :

- la matrice de confusion, qui confronte les résultats de classification à une vérité terrain, apporte sur l'ensemble de la zone d'étude analysée une information synthétique sur la manière dont certaines confusions entre classes affectent les résultats de la classification,
- la logique floue, qui met en oeuvre des ensembles flous de règles de classification définis par l'opérateur indépendamment de toute vérité terrain, apporte pour chaque objet une information différenciée sur la manière dont les classes se mélangent au sein de cet objet.

Selon les résultats de la matrice de confusion, la classification mise en œuvre sur le bassin versant de l'Hérault procure globalement une information précise sur les six principales pressions d'occupation du sol (89 % de pixels bien classés par rapport à la vérité terrain collectée par photo-interprétation des images ortho-photos). Alors que les classes «surface en eau », « sols agricoles », « zones de végétation arborée semi-naturelles» et « zones de végétation herbacée semi-naturelles » ont été extraites de façon précise (99 %, 96 %, 88 %, et 82 % respectivement de pixels bien classés pour chacune de ces catégories), les classes « sols nus semi-naturels » et « sols artificialisés » ont était extraites de façon plus incertaine (65 % et 51 % respectivement de pixels bien classés du point de vue de l'utilisateur pour chacune de ces catégories) avec de fortes confusions entre elles, ainsi qu'avec la classe « sols agricoles » pour la classe « sols naturels ».

Selon les résultats de validation utilisant les propriétés de logique floue la classification est fiable et stable quelque soit la catégorie (une fiabilité et une stabilité proches de 1 ont été obtenues pour l'ensemble des six catégories d'occupation du sol, cf. Tableau 6-3).

Ces résultats en apparence contradictoires ne le sont pas, mais permettent de mieux cerner les limites et l'incertitude de la classification. Le système de classification en logique floue considère le résultat comme fiable, c'est-à-dire de son point de vue sans ambiguïté de mise en œuvre de ses règles de classification. Ceci ne l'empêche pas dans les cas cités plus haut d'être erroné (classes « sols nus semi-naturels », « sols artificialisés »). Dans ces cas les règles floues de classification sont discriminantes (pas d'ambiguïté) mais peu adaptées (des erreurs).

L'opérateur a délimité les ensembles flous des différents critères définissant les classes. Par excès de confiance ou par manque d'expertise, il a défini des ensembles flous trop étroits sur certaines classes, sous-estimant l'ambigüité existante entre les classes (par exemple celles se confondant avec les sols nus semi-naturels). Cet exemple illustre parfaitement la difficulté pour un opérateur de définir *à priori* les fonctions floues des règles de classification.

Tableau 6-3 : Statistiques de la fiabilité de la stabilité des six catégories d'occupation du sol extraites dans le corridor rivulaire de l'Hérault

	Fiabilité				Stabilité			
Catégories	Moyenne	Ecart-type	Minimum	Maximum	Moyenne	Ecart-type	Minimum	Maximum
C1	1	0	1	1	1	0	1	1
C2	0.9799	0.0697	0.5	1	0.97	0.12	0	1
C3	0.9998	0.0015	0.56	1	0.97	0.14	0	1
C4	0.9991	0.0145	0.5	1	0.99	0.02	0.01	1
C5	0.9931	0.0381	0.11	1	0.98	0.08	0.01	1
C6	0.9982	0.0174	0.5	1	0.99	0.06	0.01	1
Moyenne	0.995	0.0235	0.5283	1	0.9833	0.07	0.1717	1

CemOA : archive ouverte d'Irstea / Cemagref

Ceci exige dès lors à une certaine prudence concernant l'interprétation des résultats d'incertitude extraits à partir des seules propriétés de la logique floue. Pourtant, cette validation (on devrait plutôt parler d'évaluation) n'en demeure pas moins très intéressante et complémentaire de celle issue de la matrice de confusion. Elle permet de spatialiser les erreurs de classification, puisqu'on connait la manière dont les classes se mélangent au sein de chaque objet, et ainsi de déduire plus précisément l'incertitude des indicateurs spatialisés liée à ces erreurs (ce point est discuté dans le Chapitre 8 :IV. de ce mémoire).

Des progrès doivent donc encore être réalisés dans l'évaluation de l'incertitude d'une carte d'occupation des sols et notamment la spatialisation de ces erreurs. Les informations apportées par les deux techniques de validation sont complémentaires : la première, issue de la matrice de confusion, est primordiale car elle est construite à partir d'une vérité terrain ; la deuxième, issue des propriétés de la logique floue, apporte une information locale et est indispensable pour correctement propager les erreurs de classification au niveau des indicateurs spatialisés car elle permet de spatialiser l'incertitude sur la carte d''occupation du sol. L'enjeu est donc de trouver des méthodes permettant de combiner ces informations complémentaires afin d'avoir une évaluation spatialisée de l'incertitude de la carte d'occupation liée à une vérité terrain.

On soulignera enfin que le problème de l'incertitude géométrique (localisation, contour, géométrie, topologie) des objets n'a été abordé qu'indirectement à travers les études comparatives sur la résolution des données source (cf. Chapitre 5 :III.1.3), et l'analyse de l'effet de leur précision planimétrique (e.g., BD Carthage cf. Chapitre 7 :III.2.1.1). Ce problème est toutefois moins sensible que le problème de l'incertitude sémantique, la plupart des indicateurs spatialisés intégrant cette information sur une emprise géographique de dimensions significatives.

CemOA : archive ouverte d'Irstea / Cemagref

Synthèse de la deuxième partie

L'état des connaissances sur la télédétection des espaces rivulaires révèle ***qu'il n'existe pas aujourd'hui de méthode permettant d'extraire finement et efficacement l'occupation des sols des corridors rivulaires sur de grands territoires de façon opérationnelle***. Cependant, il indique clairement la nécessité d'utiliser la ***classification orientée objet*** à partir de ***l'imagerie à THRS*** contenant une information dans ***l'infrarouge proche*** (PIR et MIR) pour extraire efficacement l'occupation du sol dans ces espaces.

Une procédure de classification, s'appuyant sur cet état des connaissances, a donc été conçue et mise au point sur le bassin versant de l'Hérault. Elle est composée ***de sept étapes***. Son optimisation a nécessité des analyses supplémentaires qui montrent que :

- ***L'orthophoto, les images satellites Spot5 XS et les données auxiliaires (BD Carthage, CLC, RPG,…)*** constituent la ***combinaison*** des sources de données (disponibles sur le territoire français) ***la plus efficace*** aussi bien pour extraire les catégories d'occupation du sol d'intérêt que pour caractériser les ripisylves (notamment la végétation arborée rivulaire).
- La ***classification supervisée*** peut être menée efficacement quelque soit la taille de la zone homogène (contexte géographique homogène et images de même date) avec un jeu d'apprentissage de ***50 parcelles/classe***.
- La ***classification expertisée*** (par fonctions d'appartenance) est cependant ***plus efficace et applicable sur de grands territoires*** que la classification supervisée. Elle offre plus de modularité, peut être distribuée sur un serveur de calcul, et est facilement transférable sur de nouvelles zones.
- ***Les estimateurs de la qualité de la classification*** évalués à partir d'une matrice de confusion ***sont stables*** (robustes) lorsque la ***matrice est construite*** à partir d'un échantillon constitué de ***50 parcelles/classe*** sélectionnées de façon aléatoire et réparties sur la zone d'étude.

La mise en œuvre de la procédure de classification optimisée sur l'ensemble des régions homogènes du ***bassin versant de l'Hérault*** montre que cette ***procédure est*** :

- ***Efficace*** : 89 % des objets d'occupation du sol vérifiés ont été correctement classés selon la typologie de pressions d'occupation du sol à 6 classes.
- ***Générique*** : le raisonnement de classification orientée objet descendant, favorisant la description de règles simples, robustes et facilitant l'intégration de nouvelles classes, a été appliqué sur l'ensemble du bassin sans avoir à définir de nouvelles règles de classification mais simplement en ajustant les seuils des règles existantes.
- ***Applicable*** : trois jours de traitement orienté objet ont suffit pour caractériser l'occupation du sol le long des 1150 km du réseau hydrographique du bassin versant de l'Hérault (560 km² de corridor rivulaire) à partir de données déjà disponibles ou à coûts réduits pour les gestionnaires.
- ***Portable et évolutive*** : de nouvelles sources de données peuvent être facilement intégrées grâce aux propriétés de la logique floue.

Elle montre également :

- ***L'intérêt des données auxiliaires*** pour améliorer (enrichir) le résultat de classification et réduire les temps de traitements.
- ***Les difficultés pour extraire correctement les classes « sols nus semi-naturels » et « sols artificialisés »***
- ***La limite d'utilisation des estimations de la fiabilité et de la stabilité basées sur les propriétés de la logique floue*** compte tenu de la difficulté pour l'opérateur à définir l'ambigüité existante entre les classes et de leur déconnexion d'une vérité terrain.
- ***L'enjeu en suspens de spatialisation de l'incertitude associée à la cartographie de l'occupation des sols.***

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Troisième partie :

Les indicateurs spatialisés du corridor rivulaire

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Introduction de la troisième partie

Cette partie est structurée en deux chapitres.

Un premier chapitre (chapitre 7) présente la définition et la méthode de construction des différents indicateurs spatialisés du corridor rivulaire utilisés dans cette étude pour quantifier les indicateurs d'occupation du sol et de la végétation rivulaire dans le cadre des études des relations pressions/impacts.

Un deuxième chapitre (Chapitre 8) met en œuvre certains de ces indicateurs sur le bassin versant de l'Hérault en analysant leur sensibilité vis-à-vis de la résolution spatiale, et propose, à titre d'exemple, plusieurs représentations cartographiques d'indicateurs susceptibles d'intéresser les gestionnaires et les décideurs.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Chapitre 7
Conception des indicateurs spatialisés du corridor rivulaire

SOMMAIRE

Résumé *234*

I. OBJECTIFS, CONCEPTS DE BASE ET ETAT DES CONNAISSANCES 236

I.1. OBJECTIFS DE LA CONSTRUCTION D'INDICATEURS *236*

I.2. L'ANALYSE DES STRUCTURES SPATIALES *236*

I.2.1. Principes de base 236
I.2.2. Utilisation de l'outil SIG 238

I.3. LES INDICATEURS D'OCCUPATION DU SOL RIVULAIRES DANS LES MODELES PRESSIONS/IMPACTS *239*

I.3.1. Quels niveaux d'organisation ? 239
I.3.2. Quelles métriques (ou attributs) du paysage ? 240
I.3.3. Quelles emprises spatiales ? 240
I.3.4. Conclusion 241

II. PRESENTATION DES INDICATEURS SPATIALISES DU CORRIDOR RIVULAIRE 241

II.1. DEFINITION GENERIQUE *241*

II.2. LES INDICATEURS D'OCCUPATION DU SOL *243*

II.2.1. Choix des métriques du paysage 243
II.2.1.1 Le pourcentage surfacique 243
II.2.1.2 La fragmentation 243
II.2.2. Délimitation des emprises spatiales 244
II.2.3. Trois familles d'indicateurs particulières : linéaire, contact et lit majeur 245
II.2.3.1 Le linéaire (ICR-regfixe-Lin) 246
II.2.3.2 Le lit majeur (ICR-regvar-Maj) 246
II.2.3.3 Le contact (ICR-ireg-Con) 246

…SOMMAIRE (suite)

CemOA : archive ouverte d'Irstea / Cemagref

…SOMMAIRE (suite)

II.3. LES INDICATEURS SPECIFIQUES A LA RIPISYLVE *247*
II.3.1. Description générale 247
II.3.2. Six types d'indicateurs spécifiques de la ripisylve 248
II.3.2.1 La longueur 248
II.3.2.2 La fragmentation 248
II.3.2.3 La superficie 248
II.3.2.4 La densité de ruptures 248
II.3.2.5 La largeur moyenne 248
II.3.2.6 L'uniformité 249

III. METHODE DE CONSTRUCTION 249

III.1. LA PLATE-FORME LOGICIEL SIG *249*
III.2. DELIMITATION DES BUFFERS *249*
III.2.1. Prétraitements 250
III.2.1.1 Construction de la donnée hydrographique « polygone eau » 250
III.2.1.2 Création d'un système de repérage curviligne 252
III.2.2. Procédure de construction 253
III.2.2.1 Création d'un évènement linéaire 253
III.2.2.2 Découpage de la donnée réseau « polygone eau » 253
III.2.2.3 Construction des buffers 253
III.2.2.3.1 Le buffer régulier à distance fixe (« regfixe ») 254
III.2.2.3.2 Le buffer régulier à distance variable (« regvar ») 254
III.2.2.3.3 Le buffer irrégulier 254
III.3. EXTRACTION DES ATTRIBUTS DU PAYSAGE *255*
III.3.1. En mode vecteur 255
III.3.2. En mode raster 255

IV. TABLEAU RECAPITULATIF DES FAMILLES ET TYPES SPECIFIQUES D'INDICATEURS SPATIALISES DU CORRIDOR RIVULAIRE 256

CemOA : archive ouverte d'Irstea / Cemagref

Résumé

Un indicateur spatialisé du corridor rivulaire au niveau d'une station écologique est défini, de façon générique, par un attribut de structure ou métrique du paysage (composition et/ou configuration) calculé pour une catégorie d'occupation du sol (selon une certaine typologie) dans une emprise du corridor rivulaire de dimensions données (représentée par un buffer). Ce buffer est le plus souvent déterminé par une distance latérale au cours d'eau (largeur de part et d'autre du cours d'eau) et une distance amont et/ou aval à la station.

Une première famille d'indicateurs est définie pour quantifier les pressions générées sur le milieu aquatique par l'occupation du sol environnante. Ils renseignent sur la présence, l'intensité et la configuration spatiale des pressions générées par une catégorie d'occupation du sol donnée au travers du pourcentage surfacique (m^2/m^2, métrique la plus utilisée pour quantifier les pressions d'occupation du sol dans les modèles pressions/impacts) ou de la fragmentation traduisant l'hétérogénéité spatiale de cette catégorie.

Ces métriques peuvent être extraites dans plusieurs emprises spatiales traduisant ainsi le rayon d'influence de ces pressions, plus ou moins proche au cours d'eau et à la station. Un grand nombre d'indicateurs peut donc être construit en fonction de la taille de ces emprises. L'utilisation de plusieurs emprises candidates (mode de spatialisation « multiple ») améliorera l'identification, dans les modèles pressions/impacts, de l'emprise au sein de laquelle une catégorie d'OS a une influence significative sur la variabilité de la réponse biologique du cours d'eau.

Parmi ces possibilités d'indicateurs, trois familles traduisant des mécanismes d'impact spécifiques et/ou des localisations particulières dans le corridor rivulaire, ont été détaillées. Ces indicateurs diffèrent en fonction de la méthode et de la distance latérale au cours d'eau prise en compte pour délimiter le buffer d'analyse.
La première famille, les indicateurs du linéaire construits à partir d'un buffer aux limites régulières (i.e. qui suivent la linéarité du cours d'eau) d'une distance fixe de 5 m de part et d'autre du cours d'eau, s'intéresse aux pressions d'une catégorie d'occupation du sol donnée qui s'exercent au plus proche du cours d'eau et apportent une information sur l'état des berges rivulaires.
La deuxième famille, les indicateurs du lit majeur définis par un buffer régulier à distance variable (en fonction du rang du cours d'eau), s'intéresse aux pressions exercées par une catégorie d'occupation du sol donnée présente au sein du lit majeur, espace de liberté théorique du cours d'eau. Ils caractérisent ainsi l'état global de l'espace rivulaire.
La troisième famille, les indicateurs de contact construits à partir d'un buffer irrégulier (i.e. qui ne suivent pas la linéarité du cours d'eau) constitué par la fusion des objets au contact du cours d'eau à partir d'un buffer irrégulier, s'intéresse aux pressions exercées par les objets d'une catégorie d'occupation du sol donnée directement au contact du cours d'eau. Elle permet également de mettre en évidence la présence de ruptures dans la bande de végétation rivulaire et de préciser ainsi l'intensité de ces pressions au niveau de ces ruptures.

Six types d'indicateurs spécifiques au rôle des ripisylves ont été également définis : (1) la longueur et (2) la fragmentation de la ripisylve qui dérivent de la famille des indicateurs du linéaire ; (3) la superficie de la ripisylve qui dérive de la famille des indicateurs de « contact » ; (4) la densité de ruptures (nombre de patchs distincts par km de cours d'eau), (5) la largeur moyenne (en m) et (6) l'uniformité (en fonction de la variation de la largeur des patchs) de la ripisylve définies à partir du buffer irrégulier de contact. Ces indicateurs permettent de quantifier de façon générale la présence et l'intensité de cette végétation le long du cours (longueur et superficie), sa connectivité (fragmentation et densité de ruptures) caractéristique essentielle de son fonctionnement, mais aussi plus spécifiquement sa capacité à jouer son rôle de barrière vis-vis des polluants diffus (fragmentation, densité de ruptures, largeur moyenne et uniformité).

La plupart de ces indicateurs sont construits en mode vecteur, plus adapté aux opérateurs topologiques linéaires et générant moins de volume de données qu'en mode raster. La délimitation automatisée des différents buffers a nécessité le développement d'algorithmes SIG spécifiques notamment pour définir la distance amont et ou/aval à la station. La définition de cette distance se base sur les outils de la segmentation dynamique utilisant les systèmes de repérage curviligne et les informations topologiques du réseau hydrographique.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

I. Objectifs, concepts de base et état des connaissances

I.1. Objectifs de la construction d'indicateurs

L'enjeu de cette partie est de définir, à partir de la carte d'occupation du sol issue du traitement de l'imagerie THRS (Très Haute Résolution Spatiale), des indicateurs spatialisés pertinents dans le cadre de l'étude des relations pressions/impacts. Cette recherche de pertinence suppose plusieurs qualités de la part des indicateurs :

(1) ils doivent potentiellement expliquer une partie de la variabilité de l'état écologique des cours d'eau. Ils doivent donc être reliés aux principaux mécanismes d'impact des activités humaines ou de la végétation rivulaire sur l'état écologique des cours d'eau ;

(2) ils doivent permettre de localiser précisément au sein du corridor rivulaire l'emprise des sources de pressions ayant un impact significatif sur l'état écologique d'une station, et être suffisamment caractéristiques de l'état de ces corridors pour informer et orienter les gestionnaires sur les mesures et les zones d'actions prioritaires dans ces espaces ;

(3) leur signification doit être à la fois synthétique et interprétable en termes d'aide à la décision et à la communication pour les gestionnaires.

La qualité de ces indicateurs dépend aussi de la facilité à les construire (i.e., l'automatisation de leur construction) et du coût de traitement associé (i.e., en termes de temps) puisqu'ils doivent pouvoir être appliqués sur de grands territoires.

Avant de présenter les indicateurs développés pour répondre à ces objectifs, on expose succinctement les concepts de base de l'analyse des structures spatiales et une vue générale des indicateurs de l'occupation du sol rivulaire que l'on retrouve généralement dans les modèles pressions/impacts.

I.2. L'analyse des structures spatiales

La construction des indicateurs spatialisés repose sur les concepts de l'analyse des structures spatiales dans un paysage, qui utilise les techniques issues des Systèmes d'Information Géographique (SIG). Cette partie présente les principes de base de cette analyse et quelques rudiments sur les techniques SIG.

I.2.1.Principes de base

Pour caractériser des paysages, les comparer, tester des liens entre structures paysagères et caractéristiques écologiques, l'utilisation d'indicateurs spatialisés (ou indicateurs du paysage) est indispensable (Burel et Baudry, 1999). Ces indicateurs spatialisés résument une structure spatiale à un nombre, ou plus exactement à un champ scalaire. Il sont invariablement définis par l'agrégation d'un ou plusieurs attributs de structure du paysage dans une zone délimitée (emprise ou échelle spatiale de l'indicateur) (Turner et al., 1991).

Dans un des textes fondateurs de l'écologie du paysage, Forman et Gordon (1981) ont proposé une nomenclature, fournissant le cadre nécessaire à la description des différents attributs de structure du paysage. Cette nomenclature, qui distingue les divers éléments que l'on peut repérer dans un paysage, a été définie dans l'optique de rechercher un « effet du paysage » sur les processus écologiques.

Elle définit dans un premier temps la matrice, l'élément dominant, englobant, et en son sein des « patchs » qui peuvent être soit des taches (forêts, habitations) soit des corridors (éléments linéaires comme les routes). L'ensemble des taches constitue une mosaïque et l'ensemble des corridors un réseau. Au sein des patchs, on

CemOA : archive ouverte d'Irstea / Cemagref

peut distinguer une lisière qui a de très fortes interactions avec la matrice ou les patchs voisins, et un milieu intérieur dans lequel les interactions avec l'extérieur sont très faibles à nulles. Plus les patchs sont allongés, plus le ratio lisière/intérieur est élevé (Figure 7.1).

La notion de matrice est obsolète (Burel et Baudry, 1999) lorsque l'on travaille sur un milieu terrestre. En effet, ce milieu est caractérisé non pas par un espace englobant (la matrice) tous les types d'occupations du sol présents dans le paysage mais plutôt par un espace composé d'un gradient de situations d'occupation du sol. On préfère alors parler de mosaïques de patchs (Urban et al., 1987), i.e. un ensemble de patch contigus de nature différente.

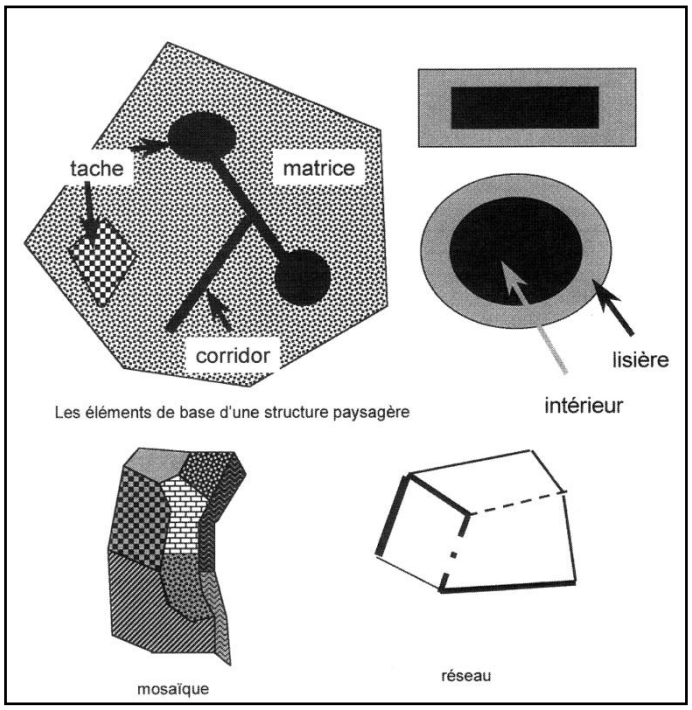


Figure 7.1 : Les catégories d'éléments du paysage d'après Burel et Baudry, 1999

La structure du paysage ainsi décrite peut être quantifiée à la fois par sa composition et par sa configuration (Dunning et al., 1982 ; Turner, 1989), ces deux aspects pouvant, indépendamment ou non, influencer les processus écologiques et les organismes.

La composition est décrite par des attributs (ou critères ou mesures) relatifs à la présence et à la quantité de chaque type de patch dans le paysage mais sans information spatiale explicite. Il existe de nombreuses mesures de la composition du paysage, incluant la proportion de chaque type, la diversité des types de patchs (richesse spécifique, diversité de Shannon, régularité…).

La configuration se réfère, par contre, aux attributs relatifs à la distribution physique ou au caractère spatial des patchs dans le paysage. Certaines mesures de la configuration, comme l'isolement ou la connectivité des patchs, sont calculées en fonction de l'emplacement relatif d'un type de patchs vis-à-vis d'un autre type, d'une frontière du paysage ou d'un élément d'intérêt. D'autres mesures, comme la forme et la taille, sont relatives aux caractéristiques morphologiques propres aux patchs.

Notons qu'il n'est pas toujours évident de catégoriser un attribut tel que la mesure de la composition ou de la configuration du paysage. Certains attributs peuvent apporter une information couplant ces deux types de mesures. Par exemple, la densité moyenne d'un type de patch reflète à la fois la quantité du type de patch présent (composition) et sa distribution spatiale (configuration).

Certains attributs de composition et/ou de configuration peuvent être formulés à différents niveaux d'organisation : au niveau du patch individuel, au niveau de la classe de patchs (patchs de même type) ou au niveau du paysage (tous les patchs). La dimension fractale par exemple, qui est une mesure de la complexité de la forme (Mandelbrot, 1982), peut être calculée pour chaque patch (niveau du patch), moyennée pour des

CemOA : archive ouverte d'Irstea / Cemagref

patchs de même type (niveau de la classe) ou pour tous les patchs du paysage (niveau du paysage). D'autres, en revanche, ne peuvent être calculés qu'à certains niveaux d'organisation : on ne peut considérer l'attribut de densité que sur une population de patchs, et non sur un seul patch.

L'analyse des structures spatiales d'un paysage, via les indicateurs spatialisés, se fait dans une perspective de lien avec les phénomènes et les processus d'intérêt. Dans cette analyse, des questions se posent alors aussi bien sur les attributs de paysage à extraire (composition et/ou configuration), que sur le niveau d'organisation de cette extraction (patch, classe, ou paysage) et sur l'étendue spatiale où seront extraits ces attributs (emprise). En plus de ces questions, il faut décider aussi de la résolution spatiale (ou échelle cartographique) (Burel et Baudry, 1999) et de la typologie d'occupation du sol à prendre en compte pour décrire ces processus (Suarez Seoane, 1998).

I.2.2. Utilisation de l'outil SIG

Le Système d'Information Géographique est un outil parfaitement adapté au calcul des indicateurs spatialisés d'un paysage. Un SIG est un système informatique qui permet de stocker, de gérer, de représenter et de manipuler des informations hétérogènes localisées (i.e. géoréférencées dans le même système de coordonnées). Chaque thème d'information est représenté par une couche qui contient la représentation cartographique des objets spatiaux (objets élémentaires de même nature) et la table de données descriptives (table attributaire) qui leur est associée.

Il existe deux types de représentations (ou modèles) géographiques pour modéliser le monde réel (Figure 7.2): (1) le modèle vecteur et (2) le modèle raster.

Chacun de ces deux modèles de données géographiques, présente des avantages spécifiques pour construire des indicateurs spatialisés à l'aide d'opérateurs spatiaux.

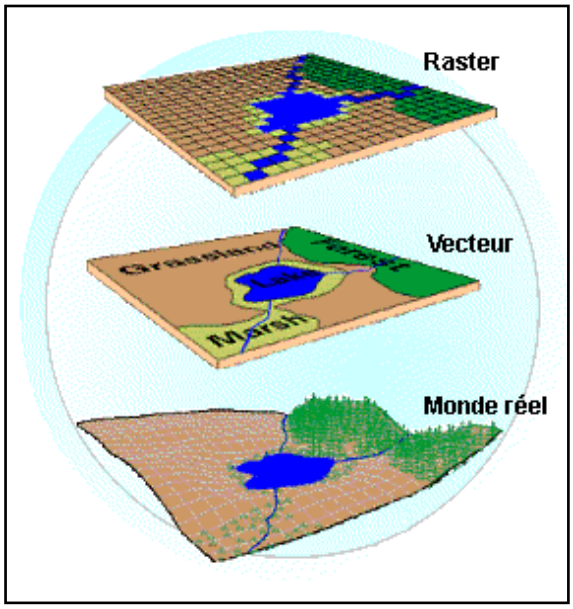


Figure 7.2 : Les deux types de modèles exploités par les Systèmes d'Information Géographique pour représenter le monde réel (Esri France).

Dans la représentation cartographique vecteur, les informations sont regroupées sous la forme de coordonnées x, y. Les objets de type ponctuel sont dans ce cas représentés par un simple point (un couple de coordonnées (x,y)). Les objets linéaires (routes, fleuves…) sont eux représentés par une succession de coordonnées $(x,y)_i$ (un arc). Les objets polygonaux (territoire géographique, parcelle…) sont, quant à eux, représentés par une succession de coordonnées délimitant une surface fermée (un arc fermé). Ce type de représentation est particulièrement utilisé pour représenter des données discrètes.

CemOA : archive ouverte d'Irstea / Cemagref

Cette représentation permet l'utilisation de plusieurs opérateurs spatiaux mobilisant des informations alphanumériques:

- Les *opérateurs dimensionnels* (spatiaux, métriques) qui permettent d'extraire des informations sur la géométrie des différentes entités géographiques (longueur, orientation d'un arc, périmètre, surface d'un polygone…).
- Les *opérateurs topologiques* qui permettent de sélectionner des objets en fonction de leur emplacement par rapport à des objets de la même couche ou d'autres couches (objets voisins, disjoints, inclus, selon une certaine distance…) ; d'agréger, de généraliser les objets dans une couche ; de créer des zones tampons autour des objets ; d'unir, de combiner, d'intersecter (…) les objets de couches différentes ; et de parcourir et interroger un réseau linéaire (chemin le plus court, accumulation d'un poids…).
- Les *opérateurs statistiques* qui permettent d'analyser l'information sémantique liée aux objets et d'établir des représentations thématiques cartographiques (selon différents modes de discrétisation).
- Les opérateurs *stochastiques* qui permettent d'étudier les relations entre les objets (autocorrélation spatiale, semi-variogramme).

Le modèle raster, quant à lui, est constitué d'une matrice de points (grille de cellules) pouvant tous être différents les uns des autres. Il s'adapte parfaitement à la représentation de données variables continues telles que la nature d'un sol, l'altitude, la pluviométrie… Il permet l'utilisation de plusieurs opérateurs spatiaux regroupés en quatre types de fonctions :

- Les *fonctions locales* qui utilisent des opérateurs mathématiques, statistiques (…) sur chaque cellule individuellement (sans influence du voisinage sur le résultat). Elles permettent la reclassification de l'information numérique d'une grille et le croisement de l'information entre plusieurs grilles (combinaisons d'informations).
- Les *fonctions focales* agissent sur une seule grille en tenant compte des valeurs des cellules voisines à l'aide d'un élément structurant (fenêtre) en général un carré qui définit le voisinage de chaque pixel (taille minimale de la fenêtre 3x3). Elles permettent de calculer par exemple les pentes ou les azimuts à partir d'une grille contenant l'information de l'altitude (MNT)
- Les *fonctions incrémentales* agissent sur l'ensemble des pixels. Elles affectent (dans une grille de sortie) une valeur à chaque pixel par rapport à un point de départ (calcul de la distance euclidienne, de la distance pondéré par un cout, calculs d'accumulation…).
- Les *fonctions zonales* qui permettent les calculs statistiques, entre deux grilles ou une grille et une couche vectorielle, en fonction de zones (cellules ou formes de valeurs identiques).

Les études en écologie du paysage utilisent les opérateurs spatiaux aussi bien en modèle raster qu'en modèle vecteur, notamment à partir du logiciel FRAGSTAT (McGarigal and Marks, 1995).

I.3. Les indicateurs d'occupation du sol rivulaires dans les modèles pressions/impacts

I.3.1. Quels niveaux d'organisation ?

Les indicateurs spatialisés utilisés dans la littérature pour traduire les pressions exercées par l'occupation du sol aux échelles du corridor rivulaire (corridor local et corridor amont, cf. Chapitre 3 :III.2.1.3) sur l'état écologique des cours d'eau (chimique, physique ou biologique) sont construits au niveau de la classe, i.e. séparément pour chaque catégorie d'occupation du sol (type de patch). En effet l'objectif de ces modèles est d'identifier l'impact relatif des différentes catégories d'occupation du sol pour établir un diagnostic de l'état écologique.

L'occupation du sol est donc généralement catégorisée selon une typologie de classes significative pour une zone d'étude et une échelle donnée, en se basant pour cela sur une structure de nomenclature hiérarchisée (e.g., Anderson et al., 1976 ; Bossard et al., 2000) qui facilite l'agrégation des classes d'occupation du sol afin de définir des indicateurs conceptuellement pertinents (Schuft et al, 1999).

On retrouve également dans les études pressions/impacts des indicateurs calculés au niveau du paysage mais à l'échelle du bassin versant. Par exemple, Snyder *et al.*, 2003 ont utilisé l'indice de contagion (mesure la

CemOA : archive ouverte d'Irstea / Cemagref

connectivité au niveau du paysage entre les patchs des différentes classes, McGarigal and Marks, 1995) qui s'est avéré non explicatif de l'état écologique évalué à partir des communautés de poissons et de macroinvertébrés benthiques. Par contre, les métriques au niveau du patch individuel ne sont pas utilisées dans ces études.

I.3.2. Quelles métriques (ou attributs) du paysage ?

Toutes les études ont utilisé le pourcentage surfacique comme métrique pour étudier l'impact d'une catégorie d'occupation du sol sur le paysage, aussi bien à l'échelle du bassin versant qu'aux échelles du corridor rivulaire. Le pourcentage surfacique est un descripteur de base, qui permet de quantifier la présence et l'intensité d'une catégorie d'occupation du sol le long des cours d'eau, susceptible d'exercer diverses pressions sur le milieu aquatique. C'est la métrique la plus pertinente pour expliquer l'état écologique des cours d'eau (Gergel et al., 2002 ; Allan et al., 2004).

Par contre elle ne traduit pas les modalités spatiales d'un phénomène (connectivité, distance au cours d'eau et à la station…). Peu d'auteurs ont utilisé aux échelles du corridor rivulaire des métriques relatives à la configuration spatiale des patchs de même type dans le paysage. Ces métriques sont le plus souvent étudiées à l'échelle du bassin versant notamment en prenant en compte la distance (euclidienne, direction des flux…) au cours d'eau et à la station afin de traduire l'impact des flux d'eau (pics de crues) et de polluants sur le milieu (e.g., Snyder *et al.*, 2003 ; King *et al.*, 2005). Cependant, Snyder *et al.*, 2003 montrent sur leur zone d'étude que les variables prenant en compte la configuration spatiale sont moins significatives que les variables de composition (pourcentage) pour expliquer la variabilité de l'état écologique des cours d'eau (poissons et macroinvertébrés benthiques).

Concernant la végétation rivulaire, peu d'études ont utilisé des métriques relatives à son rôle sur l'état écologique, probablement à cause de la résolution spatiale de la donnée d'occupation du sol trop grossière pour caractériser précisément la végétation dans ces espaces.

On trouve néanmoins l'étude de Stewart *et al.*, 2001 qui se base sur les travaux de Shuft *et al.*, 1999. Dans cette étude, des métriques évaluant la continuité (ou fragmentation) de la bande rivulaire, calculées à partir d'une information d'occupation du sol issue de la photo-interprétation des images orthophotos (entre 0.6 m et 1 m) et de données auxiliaires, ont été utilisés pour expliquer la réponse biologique des communautés de poissons. Ils montrent que la présence de taxon polluosensible (par rapport à l'oxygène et la température) et l'indice d'intégrité biologique sont corrélés avec le nombre de ruptures par km et la moyenne de ces ruptures dans un buffer de 10 à 20 m de cette végétation rivulaire. Ces résultats illustrent notamment le rôle d'ombrage (qui maintient les eaux froides et bien oxygénées) et d'abri (zone de refuge) de la ripisylve pour les communautés de poissons.

D'autres métriques en lien avec le rôle de zone tampon de cette végétation ont été également développées. Par exemple, dans une approche heuristique Weller *et al.*, 1998 utilisent, en plus de la fréquence de ruptures, des métriques liées à la forme de la bande rivulaire. Ils démontrent que plus cette bande est large (en moyenne) et uniforme, plus la capacité de rétention des flux de polluants diffus (nitrates, sédiments) est efficace.

I.3.3. Quelles emprises spatiales ?

L'emprise spatiale au niveau du corridor rivulaire, quelque soit l'échelle (CRamont ou CRlocal), correspond le plus souvent à un buffer (zone tampon) autour du cours d'eau (réseau hydrographique) (Gergel *et al.*, 2002) qui varie en largeur (distance latérale au cours d'eau) pour le CRamont mais aussi en longueur (distance longitudinale en amont et en aval de la station) pour le CRlocal.

Dans la plupart des études une seule emprise spatiale candidate (voire deux ou trois) est utilisée pour analyser l'impact de l'occupation du sol sur l'état écologique. Seuls Frimpong *et al.*, 2005b ont délimités plusieurs emprises candidates afin de localiser plus précisément la zone d'influence de la végétation rivulaire sur les communautés de poisson.

Ces emprises sont le plus souvent définies par expertise, elles sont donc variables en fonction des études (Van Sickle and Johnson, 2008) aussi bien au niveau de leur dimension latérale que longitudinale. Par exemple, alors que Stewart *et al.*, 2001 s'intéressent à des buffers avec des largeurs de 10 et 30 m de part et d'autre du cours d'eau, Allan *et al.*, 1997 définissent des buffers de 100 et 250 de large. Concernant la dimension

CemOA : archive ouverte d'Irstea / Cemagref

longitudinale, Sponseller *et al.*, 2001 utilise par exemple des buffers avec une distance en amont de la station de 200, 1000 et 2000 m à l'échelle du CRlocal, Wasson *et al.*, 2005 considère uniquement un buffer avec une distance en amont de 2.5 km.

Ces emprises diffèrent aussi en fonction de leur méthode de construction. Par exemple, pour la définition de la distance en amont (et quelque fois aval) de la station, certains auteurs considèrent la distance curviligne (plus précise) du réseau amont de la station (e.g., Sponseller *et al.*, 2001 ; Harding *et al.*, 1998), alors que d'autres utilisent la distance radiale centrée sur la station (e.g., Wang *et al.*, 2001 ; Van Sickle *et al.*, 2004 ; Wasson *et al.*, 2005).

Pour la distance latérale au cours d'eau, quelques soient les études, c'est la distance euclidienne qui est utilisée définissant un buffer aux limites régulières (i.e qui suivent la linéarité du cours d'eau). Cette distance peut être fixe (e.g., Sponseller *et al.*, 2001 ; Harding *et al.*, 1998) ou variable en fonction de la taille du cours d'eau (e.g. Wasson *et al.*, 2005 ; Wasson *et al.*, 2010).

I.3.4. Conclusion

La plupart des indicateurs d'occupation du sol aux échelles du corridor rivulaire utilisés dans les modèles pressions/impacts sont calculés au niveau des patchs de même type (i.e. catégorie d'occupation du sol). Ils s'intéressent principalement au pourcentage surfacique (métrique de composition) d'une catégorie d'occupation du sol donnée. Ils traduisent ainsi la présence et l'intensité des pressions potentielles s'exerçant sur l'état écologique générées par une catégorie d'occupation du sol. Les métriques concernant la configuration spatiale de cette occupation sont peu utilisées et s'avèrent dans certaines études beaucoup moins efficaces que les métriques de composition (pourcentage surfacique) pour expliquer la variabilité de l'état écologique. Ces métriques à ces échelles sont extraites dans un ou plusieurs buffers (emprises) autour du cours d'eau (réseau hydrographique) qui varient en largeur (distance euclidienne au cours d'eau) pour le CRamont mais aussi en longueur (distance curviligne ou circulaire en amont et/ou en aval à la station) pour le CRlocal. La définition de ces emprises est fortement variable en fonction des études.

II. Présentation des indicateurs spatialisés du corridor rivulaire

II.1. Définition générique

Dans cette étude, on définit de façon générique « l'indicateur spatialisé» par un attribut de structure de paysage (composition et/ou configuration) calculé pour une catégorie d'occupation du sol (selon une certaine typologie) dans une emprise spatiale d'une dimension donnée.

L'emprise spatiale au niveau du corridor rivulaire correspond à un buffer d'analyse (une zone tampon) autour d'un tronçon de cours d'eau. Ce buffer est déterminé au niveau d'une station écologique en fonction (i) d'une distance latérale au cours d'eau et (ii) d'une distance le long du réseau hydrographique amont et/ou aval à la station (Figure 7.3). Ces indicateurs quantifient les pressions susceptibles d'être générées par un type d'occupation du sol donné, présent au sein de différentes emprises spatiales dans l'environnement de la station. Ils sont utilisés dans les modèles pressions/impacts aux échelles du corridor rivulaire (CRamont et CRlocal, cf. Chapitre 3 :III.2.1.3). La construction de ce buffer nécessite une localisation précise de la station écologique sur un réseau hydrographique référencé et correctement orienté (cf. III.2).

Dans l'illustration précédente (Figure 7.3), on représente un buffer « régulier » (i.e. dont les limites suivent la linéarité du cours d'eau) défini selon une distance euclidienne fixe du cours d'eau. Cependant, on verra par la suite (cf. II.2.2) que ce buffer aux limites régulières peut être également défini à partir d'une distance euclidienne variable au cours d'eau dépendant par exemple de la taille du cours d'eau, ou qu'un buffer aux limites irrégulières (i.e., qui suivent les contours des objets d'occupation du sol étudiés) pourra également être utilisé pour quantifier l'impact de l'occupation du sol qui s'exerce sur une station écologique.

CemOA : archive ouverte d'Irstea / Cemagref

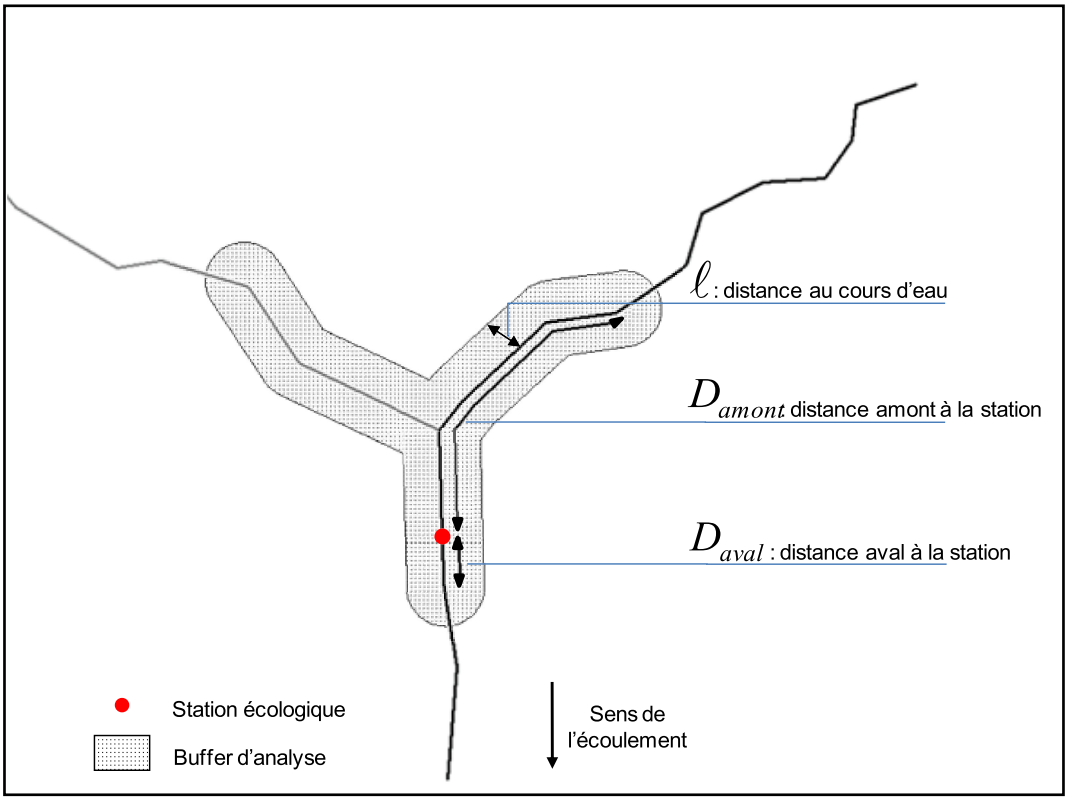


Figure 7.3 : Buffer d'analyse pour le calcul d'un indicateur spatialisé au niveau d'une station écologique selon une largeur l fixe sur cette illustration et une longueur L déterminée par une distance amont Damont et/ou une distance aval Daval à la station.

Dans cette étude, on a mis en place une terminologie pour dénommer un indicateur spatialisé du corridor rivulaire. Elle est construite selon la notation suivante :

ICR- Buf- ℓ – D_{amont} – D_{aval} – Nat – OS

avec ICR pour indicateur du corridor rivulaire, Buf pour le type de buffer construit, ℓ pour la largeur pris en compte de part et d'autre du cours d'eau pour construire le buffer, D_{amont} pour la distance amont de la station, D_{aval} pour la distance aval à la station, Nat pour le type d'attribut de structure du paysage (ou métrique) et OS pour la catégorie d'occupation du sol analysée.

Par exemple, l'indicateur du pourcentage surfacique (%) de forêt sur un buffer (régulier à distance fixe) de 30 m de part et d'autre du cours d'eau sur 1 km en amont et 100 m en aval de la station sera nommé : ICR-regfixe-30m-1km-100m-%-forêt. Dans ce même exemple, si la distance aval à la station n'est pas prise en compte l'indicateur sera alors nommé : ICR-regfixe-30m-1kkm-%-forêt. Par conte, s'il s'intéresse à l'ensemble du réseau amont à la station (i.e échelle du CRamont) il aura comme nom : ICR-regfixe-30m-∑amont-100m-%-forêt.

A noter que dans cette étude, on considère que les indicateurs construits sur une emprise spatiale de même distance au cours d'eau (de même largeur de part et d'autre du cours d'eau) sont de la même famille. Les trois exemples d'indicateurs présentés au-dessus appartiennent donc à la même famille.

Cette terminologie est un moyen simple pour identifier les différents indicateurs d'occupation du sol au niveau du corridor rivulaire mais comme on le verra par la suite des mots clés distinctifs s'insèreront afin de nommer des indicateurs traduisant des mécanismes d'impact spécifiques ou bien des localisations particulières dans le corridor rivulaire.

CemOA : archive ouverte d'Irstea / Cemagref

Ainsi, à partir de cette définition, un grand nombre d'indicateurs spatialisés peut être construit au niveau d'une station écologique. Ces indicateurs se différencient en fonction de la dimension du buffer d'analyse (en largeur et/ou en longueur), de la nature de l'attribut de structure du paysage ou de la catégorie d'occupation du sol considérée.

On présente, dans la suite de cette partie les caractéristiques des indicateurs spatialisés développés pour quantifier à la fois les pressions générées par une catégorie d'occupation du sol (i.e indicateurs de pressions d'occupation du sol cf. II.2) et l'influence spécifique liée aux rôles de la ripisylve (cf. II.3) qui s'exercent sur l'état écologique des cours d'eau.

II.2. Les indicateurs d'occupation du sol

II.2.1. Choix des métriques du paysage

Deux métriques, calculées pour des patchs de même type, ont été choisies dans cette étude pour étudier l'influence d'une catégorie d'occupation du sol donnée sur l'état écologique : le pourcentage surfacique (cf. II.2.1.1) et la fragmentation (cf. II.2.1.2).

II.2.1.1 Le pourcentage surfacique

Le pourcentage surfacique est l'attribut de structure du paysage le plus utilisé (aussi bien à l'échelle du corridor rivulaire qu'à l'échelle du bassin versant) et considéré comme le plus explicatif de l'état écologique des cours d'eau (cf. I.3.2).

C'est un descripteur de base qui ne traduit pas les modalités spatiales d'un phénomène, mais permet de quantifier la présence et l'intensité d'une catégorie d'occupation donnée du sol exerçant potentiellement des pressions, sur l'emprise dans laquelle elle a été analysée, qui modifient un ou plusieurs processus de fonctionnement (hydromorphologique, biogéochimique et écologique) de l'écosystème rivière.

Par exemple, le pourcentage de terres arables sur une bande de 100 m de part et d'autre du cours d'eau cumulé sur 2 km en amont de la station traduit l'intensité et la présence des pressions potentielles générées par cette catégorie d'occupation du sol sur cette emprise (apport de polluants et de substances toxiques, apports de sédiments…) qui modifient les différents processus de cet écosystème.

Cet attribut de composition peut être mesuré sur n'importe quelle taille de buffer. Cette quantité, exprimée en pourcentage (m²/m²), est calculée pour chaque catégorie d'occupation du sol selon la formule suivante :

$$\sigma_{Si} = \frac{S_i}{S_{tot}} \times 100 \quad (1)$$

avec S_i la surface cumulée de la catégorie i d'occupation du sol dans le buffer d'analyse et S_{tot} la surface totale de ce buffer.

Lorsque cet attribut est utilisé, on parle d'indicateur de pourcentage (noté « % » dans le nom de l'indicateur).

II.2.1.2 La fragmentation

CemOA : archive ouverte d'Irstea / Cemagref

La fragmentation permet de caractériser la structure de la surface occupée par un type d'occupation du sol : en un seul patch (homogène) ou bien découpée en une multitude de patchs (hétérogène).

Dans cette étude, on a choisi d'utiliser cette métrique de configuration car elle peut apporter dans certains cas une information complémentaire au pourcentage surfacique.

Par exemple, Gergel (2005) montre à l'aide d'une approche heuristique que l'hétérogénéité spatiale des sources de polluants diffus générées par les terres agricoles explique une part importante des taux de nutriments retrouvés dans le cours d'eau.

La fragmentation de la végétation au bord du cours d'eau permettra de traduire également les capacités de rétention de ces polluants par les ripisylves (cf. II.3.2.2).

Cet attribut de configuration peut mathématiquement être calculé sur n'importe quelle taille de buffers. Cette mesure de fragmentation varie entre 0 (non fragmenté) et 1 (très fragmenté). Elle est calculée pour chaque catégorie d'occupation du sol selon la formule suivante :

$$\sigma_{Frag_i} = 1 - \frac{1}{N_i} \quad (2)$$

avec N_i (>=1) le nombre de patchs de la catégorie *i* d'occupation du sol dans le buffer d'analyse.

Lorsque cet attribut est utilisé, on parle d'indicateur de fragmentation (noté « frag » dans le nom de l'indicateur).

II.2.2. Délimitation des emprises spatiales

Dans la littérature (cf. I.3.3) on a vu que la définition de ces emprises est très variable selon les travaux étudiant les relations pressions/impacts. La plupart des études utilisent une seule emprise spatiale candidate pour étudier l'influence combinée (ou non) des catégories d'occupation du sol sur l'état écologique des cours d'eau. Ce mode de spatialisation « simple » limite la compréhension des relations pressions/impacts puisque l'effet d'une catégorie d'occupation du sol donnée (combinée avec d'autres) peut être plus ou moins perceptible sur l'état écologique en fonction de l'emprise spatiale considérée au niveau du corridor rivulaire.

Dès lors, on comprend qu'il est intéressant d'analyser sur différentes emprises spatiales candidates l'influence de l'occupation du sol au niveau du corridor rivulaire sur l'état écologique des cours d'eau. Ce mode de spatialisation « multiple » permettra d'identifier plus précisément l'emprise spatiale sur laquelle une catégorie d'occupation du sol aura l'influence la plus significative sur la variabilité de la réponse biologique. Ce mode sera utilisé dans la partie 4 de ce mémoire dédiée à l'étude statistique de ces relations.

Ainsi, dans l'absolu, plusieurs indicateurs d'occupation du sol peuvent être considérés. Ils étudient le pourcentage ou la fragmentation d'une catégorie d'occupation du sol donnée dans diverses emprises spatiales, i.e., dans des buffers construits selon une distance latérale au cours d'eau et une distance amont (et/ou aval) à la station différente.

De plus, ces buffers peuvent être délimités de différentes façons. Dans la littérature (cf. I.3.3), on a vu que quelque soit les études, la distance latérale au cours d'eau correspond à la distance euclidienne définissant un buffer aux limites régulières (i.e., qui suivent la linéarité du cours d'eau). Cette distance peut être fixe ou variable.

Le buffer régulier à distance fixe (« regfixe », cf. III.2.2.3.1) permet d'étudier l'influence d'une catégorie d'occupation du sol au sein d'une bande précise de part et d'autre du cours d'eau quelque soient les caractéristiques de ce cours d'eau. Par contre, le buffer régulier à distance variable (« regvar », cf. III.2.2.3.2) permet d'étudier cette influence au sein d'une zone particulière dont les limites sont estimées en fonction des caractéristiques du cours d'eau. Par exemple, si on s'intéresse à des indicateurs étudiant l'occupation du sol dans la zone riparienne, la largeur de bande de part et d'autre du cours d'eau pourra être estimée en fonction de la taille du cours d'eau (son rang de Strahler).

CemOA : archive ouverte d'Irstea / Cemagref

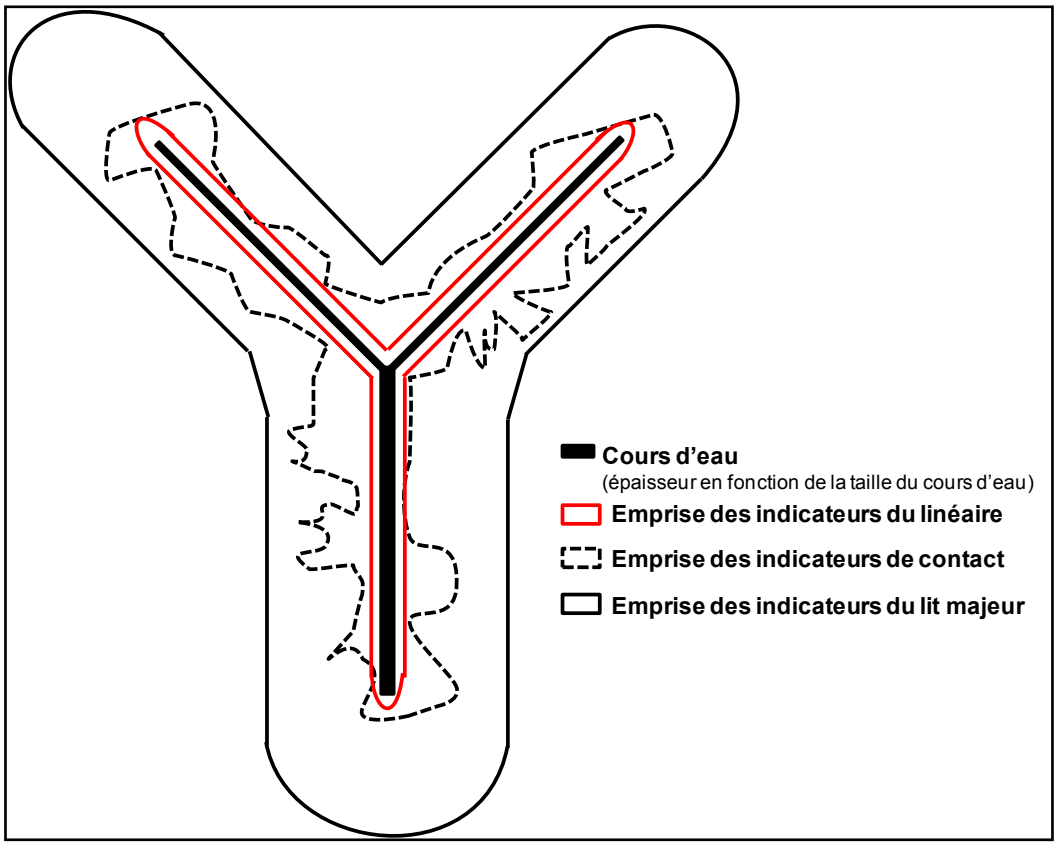
En plus de ces deux types de buffer (« regfixe » et « regvar »), un troisième type, le buffer « irrégulier » (« ireg ») peut également être utilisé pour quantifier l'influence de l'occupation du sol. Il est défini par des limites irrégulières qui ne suivent pas la linéarité du cours d'eau mais les contours des « objets » d'occupation du sol étudiés dont les limites se situent à une distance euclidienne donnée du cours d'eau (cf. III.2.2.3.3). Alors que l'utilisation des buffers « réguliers » ne prend en compte que la partie des objets d'occupation du sol contenue à l'intérieur de ces buffers, oubliant de ce fait la partie de ces objets hors limite des buffers, l'utilisation des buffers « irréguliers » permet de prendre en compte les caractéristiques complètes de ces objets.

L'utilisation de ces trois types de délimitation de buffers (régulier à distance fixe ; régulier à distance variable ; irrégulier) multiplie d'autant le nombre d'indicateurs possibles pour étudier l'influence d'une catégorie d'occupation du sol s'exerçant au niveau d'une station écologique.

Parmi ces possibilités, trois familles (i.e. une famille regroupe des indicateurs calculés sur une emprise spatiale de même distance au cours d'eau), traduisant des mécanismes d'impact spécifiques et/ou des localisations particulières dans le corridor rivulaire, ont été détaillées (cf. II.2.3).

II.2.3. Trois familles d'indicateurs particulières : linéaire, contact et lit majeur

Les trois familles d'indicateurs que nous retenons sont particulièrement différentes par la méthode et la distance latérale au cours d'eau prises en compte pour délimiter le buffer d'analyse (Figure 7.4). Alors que la famille « indicateurs du linéaire » (ICR- regfixe-Lin, cf. II.2.3.1) est définie à partir d'un buffer régulier fixe (i.e., la même largeur pour n'importe quelle taille de cours d'eau) au plus proche du cours d'eau, la famille « indicateurs du lit majeur » (ICR--regvar-Maj, cf. II.2.3.2) est définie par un buffer régulier variable (i.e., une largeur différente en fonction de la taille du cours d'eau) correspondant aux limites du lit majeur. La famille « indicateurs de contact » (ICR-ireg-Con, cf. II.2.3.3), quant à elle, est caractérisée par un buffer irrégulier «de contact » (i.e, défini par la fusion des objets en contact avec le cours d'eau).



CemOA : archive ouverte d'Irstea / Cemagref

Figure 7.4 : Schématisation des emprises spatiales des trois familles d'indicateurs spécifiques d'indicateurs d'occupation du sol

II.2.3.1 Le linéaire (ICR-regfixe-Lin)

Ces indicateurs correspondent à un buffer étroit d'une largeur fixe de 5 m de part et d'autre du cours d'eau. Sur ce buffer le pourcentage surfacique est sensiblement équivalent au pourcentage de longueur (m/m) compte tenu de la faible largeur du buffer d'où le terme de « linéaire » pour dénommer ces indicateurs.

Ils s'intéressent aux pressions d'une catégorie d'occupation du sol donnée qui s'exercent au plus proche du cours d'eau et apportent une information sur l'état des berges rivulaires.

Dans la terminologie des indicateurs mise en place dans cette étude, le ℓ est remplacé par le mot clé « Lin », ces indicateurs débutent donc par « ICR-regfixe-Lin ».

II.2.3.2 Le lit majeur (ICR-regvar-Maj)

Ces indicateurs sont construits dans un buffer supposé délimiter le lit majeur du cours d'eau. Comme pour définir l'espace des corridors rivulaires (cf. Chapitre 5 :III.2) la largeur du lit majeur, en absence d'une couche plus précise sur les fonds de vallée des rivières en France, est estimée à partir du rang de Strahler.

Tableau 7-1 : Largeur de lit majeur pris en compte en fonction du rang de Strahler pour calculer les indicateurs du lit majeur.

Rangs	1, 2, 3	4	5	6	7	8
Largeur de corridor rivulaire	100	140	250	600	1200	2400

Ces indicateurs s'intéressent aux pressions exercées par une catégorie d'occupation du sol donnée présente au sein du lit majeur, espace de liberté théorique du cours d'eau. Ils caractérisent ainsi l'état global de l'espace rivulaire.

Dans la terminologie des indicateurs mise en place dans cette étude, le ℓ est remplacé par le mot clé « Maj », ces indicateurs débutent donc par « ICR-regvar-Maj ».

II.2.3.3 Le contact (ICR-ireg-Con)

Ces indicateurs considèrent uniquement les objets directement en contact avec le cours d'eau. Ces objets fusionnés constituent un buffer aux limites irrégulières.

Ils s'intéressent aux pressions exercées par les objets d'une catégorie d'occupation du sol donnée, directement au contact du cours d'eau (ces objets en contact étant pris dans leur globalité, par opposition aux indicateurs du linéaire qui ne prennent en compte que la partie de ces objets située dans une bande de 5m). Ils permettent également, lorsqu'ils sont construits à partir des catégories anthropiques d'occupation du sol (i.e., territoires agricoles ou urbains), de mettre en évidence la présence de ruptures dans la bande rivulaire semi-naturelle (végétation arborée et/ou herbacée et/ou sol nu naturel) et précisent ainsi l'intensité de cette pression au niveau de ces ruptures. La Figure 7.5 montre un exemple d'indicateur de contact dans le cas d'une ripisylve arborée interrompue par des surfaces agricoles.

CemOA : archive ouverte d'Irstea / Cemagref

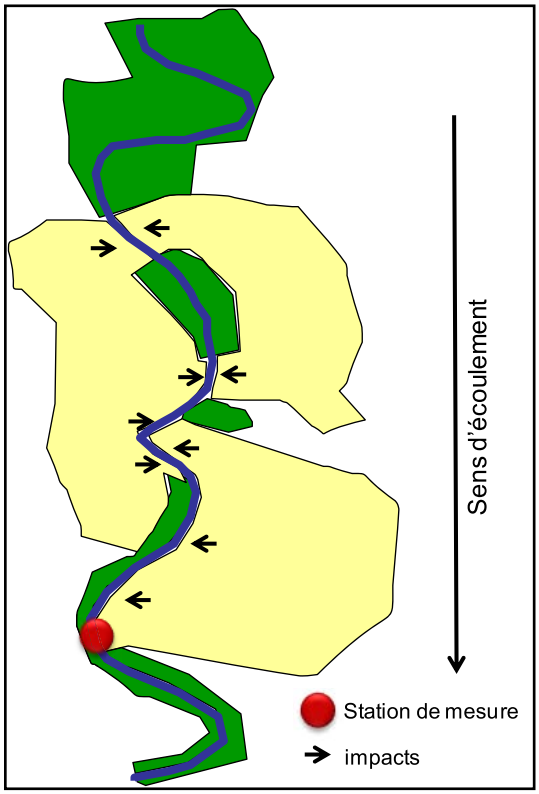


Figure 7.5 : Schématisation de l'indicateur de contact agricole. Cet indicateur traduit la présence et l'intensité au niveau des ruptures de la bande arborée, donc les pressions agricoles directement en contact avec le cours d'eau. En jaune, les territoires agricoles, en vert la végétation arborée rivulaire, en bleu l'eau.

Dans la terminologie des indicateurs mise en place dans cette étude, le ℓ est remplacé par le mot clé « Con » (pour Contact), ces indicateurs débutent donc par « ICR-ireg-Con».

II.3. Les indicateurs spécifiques à la ripisylve

II.3.1. Description générale

Les indicateurs présentés dans la partie II.2 sont applicables à différentes catégories d'occupation des sols, et notamment à la végétation semi-naturelle. Ici, nous nous intéressons plus particulièrement à la ripisylve, c'est-à-dire aux patchs de végétation arborée et/ou arbustifs et/ou herbacées très proches du cours d'eau.

Six types d'indicateurs pertinents pour caractériser ces patchs de ripisylve, dont les trois premiers appartiennent aux familles d'indicateurs d'occupation du sol présentées plus haut, ont été définis :

- Deux types appartenant à la famille « indicateurs du linéaire » (ICR-regfixe-Lin) : la «longueur de ripisylve » (cf. II.3.2.1) et « la fragmentation de la ripisylve » (cf. II.3.2.2) ;
- Un type appartenant à la famille « indicateurs de contact » (ICR-ireg-Con) : « la superficie de ripisylve » (cf. II.3.2.3) ;
- Trois nouveaux types qui sont construits à partir du buffer « contact » (i.e., constitué par la fusion des objets au contact du cours d'eau) : « la densité de ruptures » (cf. II.3.2.4), « la largeur moyenne » (cf. II.3.2.5) et « l'uniformité de la ripisylve » (cf. II.3.2.6).

On emploie le terme « type » car ces indicateurs peuvent être calculés à différents niveaux d'agrégation des classes de végétation semi-naturelle.

CemOA : archive ouverte d'Irstea / Cemagref

II.3.2. Six types d'indicateurs spécifiques de la ripisylve

II.3.2.1 La longueur

Ces indicateurs appartiennent à la famille ICR-regfixe-Lin caractérisée par un buffer étroit de 5 m de part et d'autre du cours d'eau (cf. II.2.3.1). Ils mesurent le pourcentage surfacique de végétation semi-naturelle en amont et/ou an aval de la station et indiquent par conséquent la présence de ripisylve sur les berges du cours d'eau qui joue un rôle important sur l'état écologique des cours d'eau (cf. Chapitre 2).

II.3.2.2 La fragmentation

Ces indicateurs appartiennent également à la famille ICR-regfixe-Lin. Ils mesurent la fragmentation (le nombre de patch distincts, cf. II.2.1.2) de la ripisylve traduisant ainsi la connectivité (ou encore le nombre de ruptures) de cette végétation qui est une caractéristique importante dans les multiples fonctions de la végétation rivulaire (Naiman et Decamps, 1997).

La connectivité est essentielle aussi bien (i) pour son rôle de barrière contre les pollutions diffuses (e.g., Weller et al., 1998) ; (ii) pour maintenir la fonction de connexion favorisant la biodiversité terrestre (Forman et Godron, 1997 ; Burel et Baudry, 1999) ; et (iii) pour son rôle de facteur de contrôle du fonctionnement écologique des cours d'eau afin de garantir la biodiversité aquatique (Barton et al., 1985 ; Feld, 2004).

II.3.2.3 La superficie

Ces indicateurs appartiennent à la famille ICR-ireg-Con caractérisé par un buffer constitué par la fusion des objets au contact du cours d'eau. Ils mesurent le pourcentage surfacique (cf. II.2.1.1) des patchs de végétation semi-naturelle par rapport à l'ensemble des patchs en contact avec le cours d'eau, traduisant ainsi la présence et l'intensité (plus ou moins de surface) de cette végétation qui joue un rôle important sur l'état écologique des cours d'eau.

II.3.2.4 La densité de ruptures

Le type « fragmentation de la ripisylve » ne permet pas de comparer directement des tronçons de longueurs différentes, aussi introduisons nous une nouvelle métrique du paysage qui mesure la « densité de ruptures de la ripisylve », indicateur construit à partir du buffer irrégulier de contact. Elle s'exprime en nombre de ruptures par kilomètre de cours d'eau et est calculée selon la formule suivante :

(4) $$\frac{(N_{ryp})}{L}$$

avec N_{ryp} le nombre de ruptures de la ripisylve, qui est sensiblement égal au nombre de patchs de la catégorie de la végétation arborée rivulaire au contact du cours d'eau et L la longueur du tronçon du cours d'eau analysé.

L'indicateur de densité de ruptures (drup) de la végétation arborée à 1 km de la station écologique par exemple sera nommé ICR-ireg-Con-1km-drup-forêt.

II.3.2.5 La largeur moyenne

Ce type, défini à partir du buffer irrégulier de contact, correspond à la quantification de la largeur moyenne de ripisylve (sur les deux rives du cours d'eau). Cette largeur moyenne, exprimée en m, est mesurée selon la formule suivante :

(5) $$\ell_{moy} = \frac{1}{N}\sum_{i}^{N} \ell_i$$

avec ℓ_i la largeur moyenne d'un patch de végétation arborée au contact du cours d'eau (cf III.3.2 pour l'extraction de cet attribut) et N le nombre total de patchs.

CemOA : archive ouverte d'Irstea / Cemagref

La largeur moyenne de la végétation rivulaire arborée est un attribut significatif de l'efficacité du rôle de barrière que joue cette bande, vis-vis des pollutions diffuses (Petersen, 1992 ; Weller et al, 1998). En effet, plus cette bande est large, plus elle est efficace pour réguler les flux latéraux provenant des pressions agricoles et/ou urbaines.

L'indicateur de largeur moyenne (largm) de la ripisylve à 1 km de la station écologique par exemple sera nommé ICR-ireg-Con-1km-largm-forêt.

II.3.2.6 L'uniformité

L'uniformité de la ripisylve caractérise la variation moyenne de la largeur de la ripisylve (sur les deux rives du cours d'eau). Cette mesure d'uniformité (σ_{Uryp}) varie linéairement entre 0 (variation ≥ ± 100 m) et 1 (variation = 0 m). Elle est calculée selon la formule suivante :

(6) $$\sigma_{Uryp} = -0.01 \times \frac{\sum_{i=1}^{N} \sigma\ell_i}{N} + 1$$

avec $\sigma\ell_i$ la variation de la largeur d'un patch de végétation arborée au contact du cours d'eau (cf III.3.2 pour l'extraction de cet attribut) et N le nombre total de patchs.

L'uniformité (régularité) de la végétation rivulaire arborée est également un attribut significatif de l'efficacité du rôle de barrière que joue cette bande vis-vis des pollutions diffuses (Weller et al, 1998) puisque cette régulation sera d'autant moins efficace si la largeur de cette bande est irrégulière.

L'indicateur d'uniformité (ufor) de la ripisylve à 1 km de la station écologique par exemple sera nommé ICR-ireg-Con-1km-ufor-forêt.

III. Méthode de construction

III.1. La plate-forme logiciel SIG

Pour mener à bien l'analyse spatiale permettant de construire les indicateurs, nous avons utilisé le logiciel ArcGis 9.2® de la société ESRI. Deux raisons principales nous ont conduits à choisir ce logiciel :

- c'est un SIG complet qui permet d'exploiter simultanément les modes de représentation raster et vecteur ;
- les différentes fonctions proposées par le logiciel peuvent être intégrées dans le langage de programmation *python* permettant l'automatisation des traitements (développement d'algorithmes). De plus, *python* possède une grande bibliothèque standard fournissant des outils complémentaires pour le calcul de ces indicateurs (notamment la librairie *NumPy* pour le calcul scientifique).

On privilégiera le modèle vecteur dans cette étude car il est plus adapté à l'analyse des formes linéaires (notamment pour étudier un réseau) et l'information stockée dans cette représentation est beaucoup moins volumineuse qu'en modèle raster ce qui est un point important compte tenu du volume de données à traiter.

III.2. Délimitation des buffers

Pour construire les indicateurs spatialisés du corridor rivulaire, on est amené à délimiter plusieurs géométries de buffers: le buffer régulier fixe (« regfixe »), le buffer régulier variable (« regvar ») et le buffer irrégulier (« ireg »). La délimitation de ces buffers nécessite tout d'abord des prétraitements (cf. III.2.1) avant de mettre en œuvre la procédure de construction (cf. III.2.2).

CemOA : archive ouverte d'Irstea / Cemagref

III.2.1. Prétraitements

III.2.1.1 Construction de la donnée hydrographique « polygone eau »

La précision planimétrique de la donnée hydrographique est déterminante car c'est à partir de cette donnée que sont délimités les buffers d'analyse et de calcul des indicateurs. Elle est d'autant plus déterminante que la largeur du buffer considéré est faible. Par exemple, la famille d'indicateurs du linéaire (ICR-regfixe-Lin, cf. II.2.3.1) sera plus sensible aux erreurs liées à la position du cours d'eau que la famille d'indicateurs du lit majeur (ICR-regvar-Maj, cf. II.2.3.2).

Le référentiel hydrographique utilisé dans cette étude est une version simplifiée et corrigée de la BDCarthage® niveau 2 (cf. Chapitre 5 :III.2). Il présente deux inconvénients : il est linéaire alors que le lit mineur des cours d'eau possède une certaine largeur ; il présente parfois des décalages spatiaux significatifs par rapport à la réalité constatée sur des orthophotos ou des images satellitaires à haute résolution.

Des procédures ont été définies et mises en œuvre pour résoudre ou contenir ces inconvénients.

(1) Une limite inhérente à la BDCarthage est qu'elle est « linéaire » et ne permet pas de représenter le lit mineur du cours d'eau qui est une entité surfacique (i.e., un polygone). Si l'on calcule par exemple un indicateur de la famille du linéaire (buffer étroit de 5 m de part et d'autres du cours d'eau) pour un cours d'eau de grande taille (grande largeur de lit mineur), l'indicateur renseignera uniquement la présence de l'eau dans ce buffer.

Une donnée hydrographique « polygone eau » a donc été créée pour représenter l'objet surfacique « lit mineur du cours d'eau ». Elle est construite à partir de la fusion des objets classés « cours d'eau » par la cartographie d'occupation du sol à THRS. Toutefois cette fusion peut présenter des discontinuités, du fait que certaines parties du cours d'eau sont soit recouvertes par les ripisylves (notamment pour les cours d'eau de petite taille) ou des infrastructures (ponts, barrages…), soit sans eau au moment de la date d'acquisition des images (cours d'eau temporaire), soit simplement non détectables sur les images (cours d'eau étroits, recouvert par la végétation rivulaire). Elle est donc combinée et unie à un buffer très fin (largeur 0.5 m) autour du référentiel hydrographique BD Carthage (Figure 7.6). On obtient ainsi un polygone continu sur l'ensemble du réseau hydrographique.

(2) La BDCarthage® issue de la numérisation des entités hydrographiques sur les cartes au 1/25 000ème de l'IGN[39] peut s'avérer être décalée par rapport aux orthophotos plus précises (Figure 7.7). Ces décalages peuvent engendrer des erreurs dans la quantification des indicateurs, notamment ceux qui sont calculés à proximité immédiate du cours d'eau (e.g., indicateurs linéaires ou indicateurs de contact).

Comme il n'est pas envisageable dans le cadre de ce travail de « corriger » la BDCarthage® sur l'ensemble du réseau hydrographique, une analyse a été menée sur un sous bassin versant de l'Hérault (16 cours d'eau) pour étudier la sensibilité des indicateurs aux erreurs de décalage de cette donnée.

Pour cela, nous avons dans un premier temps construit un réseau hydrographique corrigé par photo-interprétation sur la base des orthophotos. Nous avons ensuite construit deux versions des indicateurs sur chacun des cours d'eau du sous-bassin versant : l'une à partir de la donnée hydrographique « polygones eau » complétée d'un buffer très fin autour de la BDCarthage® (appelé dans l'analyse « réseau non corrigé »), l'autre à partir de la donnée hydrographique « polygones eau» complétée d'un buffer très fin autour du réseau hydrographique corrigé (appelé dans l'analyse « réseau corrigé »).

Les méthodes et les résultats obtenus dans cette analyse sont détaillés en ***ANNEXE IX***.

L'analyse, sur ce sous bassin versant, montre que la valeur des indicateurs linéaires (buffer étroit, cf. II.2.3.1) varient faiblement lorsque des erreurs de décalages existent et que cette variation ne s'accentue pas en fonction de l'importance de ces erreurs sur un tronçon de cours d'eau analysé.

[39] Institut Géographique National

CemOA : archive ouverte d'Irstea / Cemagref

Par exemple pour l'indicateur de longueur de ripisylve (cf. II.3.2.1), on observe pour la majorité des cours d'eau une différence (en valeur absolue) inférieure à 5 % entre les valeurs de cet indicateur construit à partir du « réseau non corrigé » et construit à partir du « réseau corrigé ».

De plus, on observe que cette différence est faiblement sensible quelque soit le décalage moyen (Figure 7.8) ou le décalage maximum (Figure 7.9) entre ces deux données hydrographiques.

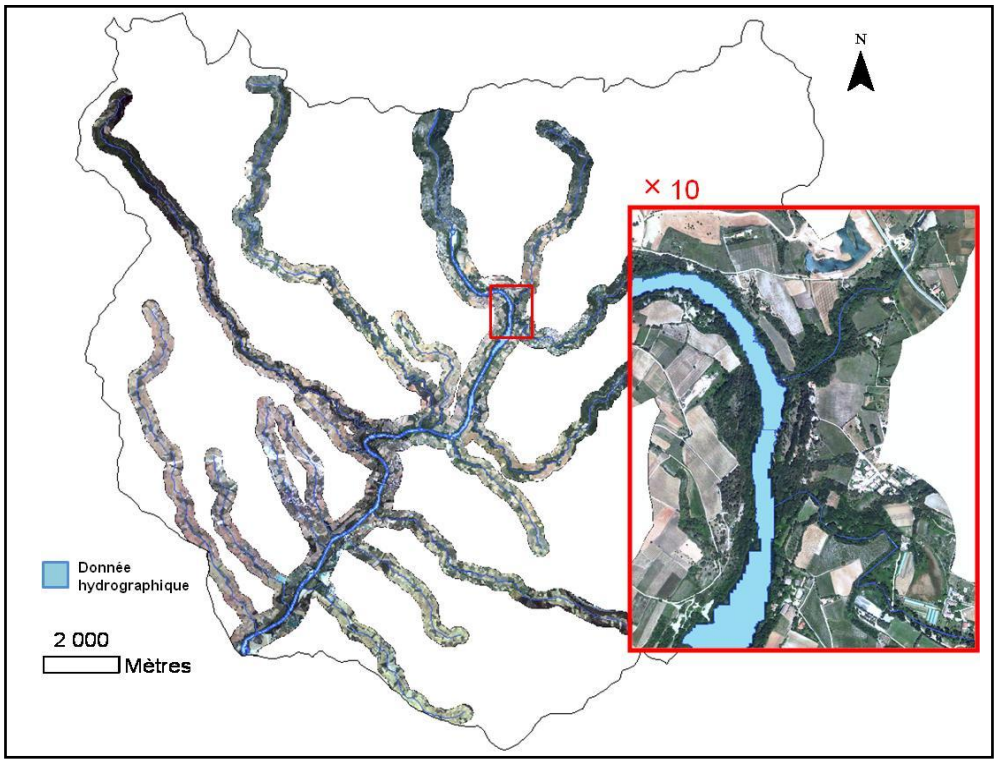
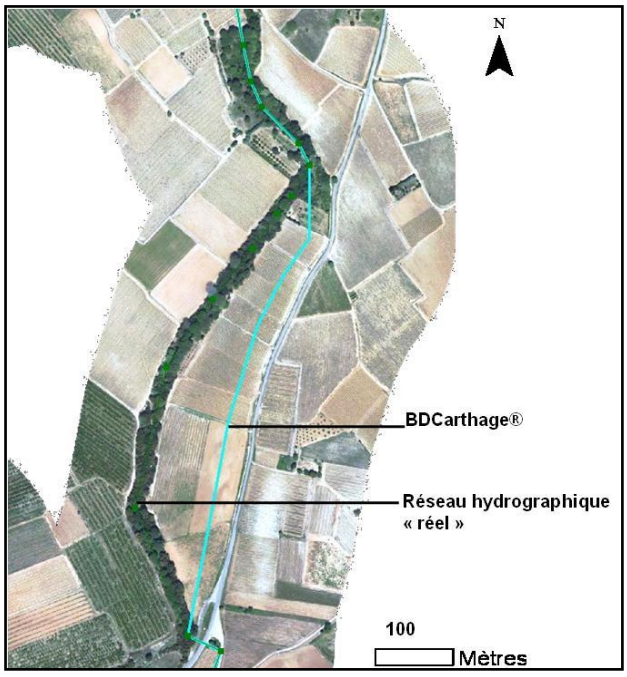


Figure 7.6 : Illustration de la donnée hydrographique obtenue par la combinaison des objets classés « cours d'eau » de la cartographie de l'occupation du sol à THRS et d'un buffer fin (1 m de largeur) autour de la BDCarthage® niverau 2.



CemOA : archive ouverte d'Irstea / Cemagref

Figure 7.7: Illustration du décalage entre la BDCarthage° et le réseau hydrographique « réel » photo-interprétés sur une orthophoto.

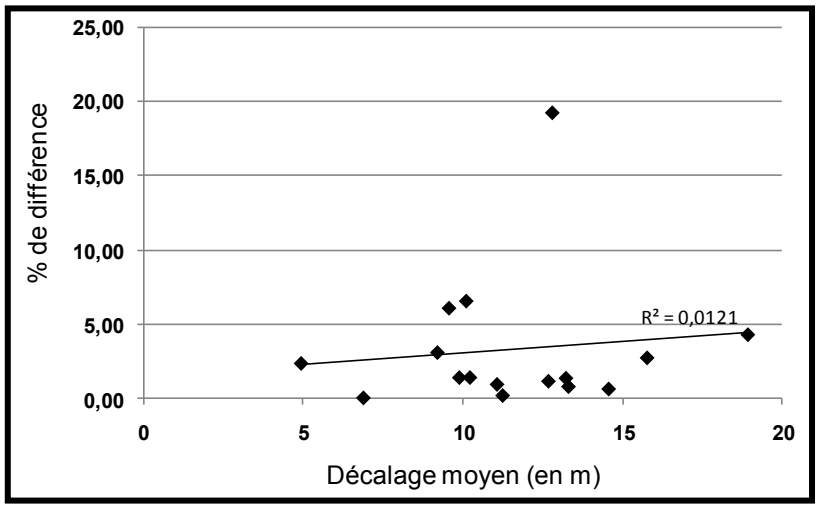


Figure 7.8 : Différence (en valeur absolue) entre l'indicateur de longueur de ripisylve construit à partir du « réseau non corrigé » et le même indicateur construit à partir du « réseau corrigé » en fonction du décalage moyen entre ces deux réseaux pour chaque cours d'eau (16) du sous-bassin versant de l'Hérault.

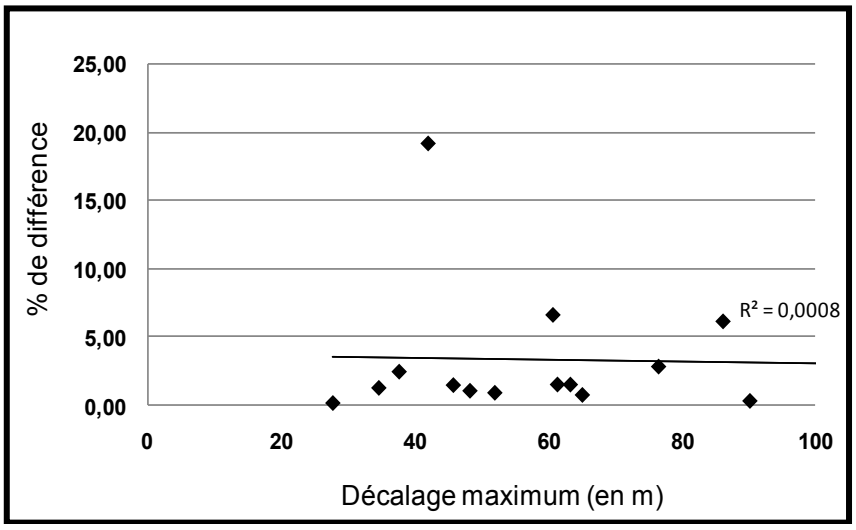


Figure 7.9 : Différence (en valeur absolue) entre l'indicateur de longueur de ripisylve construit à partir du « réseau non corrigé » et le même indicateur construit à partir du « réseau corrigé » en fonction du décalage maximum entre ces deux réseaux pour chaque cours d'eau (16) du sous-bassin versant de l'Hérault.

La donnée hydrographique « polygones eau», pour l'analyse et le développement des indicateurs spatialisés du corridor rivulaire, sera donc bien construite en combinant la fusion des objets, classés « cours d'eau » par la cartographie d'occupation du sol à THRS, et un buffer très fin (0.5 m) autour de la BDCarthage®niveau 2. Cette donnée plus précise améliorera la quantification de ces indicateurs.

III.2.1.2 Création d'un système de repérage curviligne

Les buffers d'analyse dépendent à la fois de la distance latérale au cours d'eau (largeur de buffer) et d'une longueur déterminée par une distance amont et/ou aval à la station (cf. Figure 7.3).

CemOA : archive ouverte d'Irstea / Cemagref

Dès lors, afin de définir une distance précise amont et/ou aval à la station, il est nécessaire d'une part de connaître la position exacte de la station et d'autre part de pouvoir définir des distances curvilignes (suivant le réseau) à partir de ce positionnement (i.e. de parcourir le réseau hydrographique).

Pour cela, un système de repérage linéaire (curviligne) a été mis en place. C'est un système de mesure à une dimension, permettant d'identifier un emplacement par sa position relative le long d'une entité linéaire comme les autoroutes par exemple (sortie pour le village d'Ollioules: au 5ème km de l'Autoroute A50). Ce système est lié dynamiquement aux entités géométriques sous-jacentes.

Il permet ainsi de localiser des phénomènes (évènements) géographiques sous forme de mesure linéaire sans modifier la géométrie des entités sous-jacentes. Ces évènements peuvent être ponctuels (station écologique localisée au 20ème km du tronçon de cours d'eau n°2) ou linéaires (portion de linéaire polluée localisée entre le 10ème et le 30ème km du tronçon de cours d'eau n°4).

Notre système de repérage curviligne est construit sur la géométrie du référentiel hydrographique (BDCarthage® niveau 2). Cette géométrie doit au préalable être correctement connectée et orientée. Les erreurs de connections ont pu être traitées à l'aide des outils topologiques semi-automatiques fournies par le logiciel SIG, celles d'orientations ont demandé le développement d'un algorithme spécifique qui permet de façon automatique de modifier le sens de digitalisation d'un arc mal orienté en se basant sur les outils de réseau proposés par le logiciel.

III.2.2. Procédure de construction

Cette procédure est réalisée en trois étapes : (1) la définition d'un évènement linéaire sur le système de repérage curviligne au voisinage de la station pour définir la distance amont et/ou aval à la station (cf. III.2.2.1) ; (2) le découpage de la donnée réseau « polygone eau» à partir de cet évènement linaire (cf. III.2.2.2); et (3) la construction du buffer à partir de la donnée réseau « polygone eau» découpée pour définir la distance latérale au cours d'eau (cf. III.2.2.3).

III.2.2.1 Création d'un évènement linéaire

Pour créer un évènement linéaire, on utilise les propriétés de la segmentation dynamique qui s'appuie sur un système de référencement curviligne. On parle de segmentation dynamique car ces outils permettent de créer dynamiquement des évènements ponctuels ou linéaires sur le réseau sans modifier sa géométrie.

Un algorithme utilisant les outils de segmentation dynamique proposés par le logiciel SIG et la topologie du réseau a été développé pour créer de façon automatique un évènement linéaire définissant un tronçon de réseau hydrographique en fonction d'une distance donnée en amont et/ou en aval de la station.

III.2.2.2 Découpage de la donnée réseau « polygone eau »

Une fois cet évènement linéaire délimité, un algorithme a été développé pour découper automatiquement la donnée hydrographique « polygone eau» en fonction des limites de l'évènement linéaire, si nécessaire (dans certains cas les buffers seront construits directement à partir de la donnée linaire BDCarthage® utilisée dans cette étude donc directement à partir de l'évènement linaire, cf. III.2.2.3.2).

Pour cela un buffer est créé autour de l'évènement linaire de manière assez large (fonction de la taille du cours d'eau étudié) et à angle droit par rapport au cours d'eau au niveau des extrémités de l'entité linaire, afin de découper correctement la donnée hydrographique « polygone eau» en fonction de la distance amont et/ou aval à la station spécifiée.

III.2.2.3 Construction des buffers

Une fois la donnée hydrographique « polygone eau » découpée selon une distance amont et/ou aval à la station, un des types de buffer peut être construit en fonction d'une distance latérale au cours d'eau spécifiée. On présente successivement les méthodes utilisées pour construire le buffer régulier à distance fixe

CemOA : archive ouverte d'Irstea / Cemagref

(« regfixe », cf. III.2.2.3.1), le buffer régulier à distance variable (« regvar », cf. III.2.2.3.2), le buffer irrégulier (« ireg », cf. III.2.2.3.3).

III.2.2.3.1 Le buffer régulier à distance fixe (« regfixe »)

Le buffer « fixe » est défini par une largeur invariable quelque soit la taille du cours d'eau à partir de la donnée hydrographique « polygone eau » (Figure 7.10-A). Il est utilisé pour construire les indicateurs appartenant notamment à la famille ICR-regfixe-Lin (cf. II.2.3.1).

Ces buffers sont construits à partir de l'outil « classique » de création d'une zone tampon fourni par le logiciel SIG qui se base sur la distance euclidienne.

III.2.2.3.2 Le buffer régulier à distance variable (« regvar »)

Le buffer «variable » est défini par une largeur variable en fonction de la taille du cours d'eau, taille fournie par la BDCarthage® qui contient l'information sur le rang de Strahler comparée à la donnée hydrographique « polygone eau». L'utilisation de cette donnée linéaire (moins précise que la donnée hydrographique « polygone eau») a peu d'incidence sur le résultat de l'indicateur compte tenu de la taille des buffers (> 200 m de part et d'autre du cours d'eau). Ce buffer est utilisé uniquement pour construire les indicateurs de la famille ICR-regvar-Maj (cf. II.2.3.2). Ce sont les mêmes types de buffers qui ont été construits pour délimiter les corridors rivulaires et extraire l'occupation du sol le long des cours d'eau (cf. Chapitre 5 :III.2).

Ils sont également construits à partir de l'outil « classique » de création d'une zone tampon du logiciel SIG, qui se base sur la distance euclidienne.

III.2.2.3.3 Le buffer irrégulier

Le buffer irrégulier est défini par la fusion de tous les objets d'occupation du sol situés à une distance euclidienne, défini par l'analyste, de la donnée hydrographique « polygone eau ». Il est notamment utilisé pour les indicateurs de la famille ICR-ireg-Con (cf. II.2.3.3) et pour les types d'indicateurs spécifiques à la ripisylve : les indicateurs de densité de ruptures (cf. II.3.2.4), de largeur moyenne (cf. II.3.2.5) et d'uniformité (cf. II.3.2.6). La Figure 7.10-B illustre le buffer irrégulier délimité pour construire ces indicateurs spécifiques (buffer irrégulier de contact).

Ces buffers sont construits, à l'aide d'opérateurs topologiques proposés par le logiciel SIG, (i) en sélectionnant au préalable l'ensemble des objets d'occupation sol à une distance latérale au cours d'eau inférieure à une valeur, ou bien au contact de cette donnée hydrographique (des objets partageant une même limite avec cette donnée) puis (ii) en fusionnant ces objets en un seul polygone.

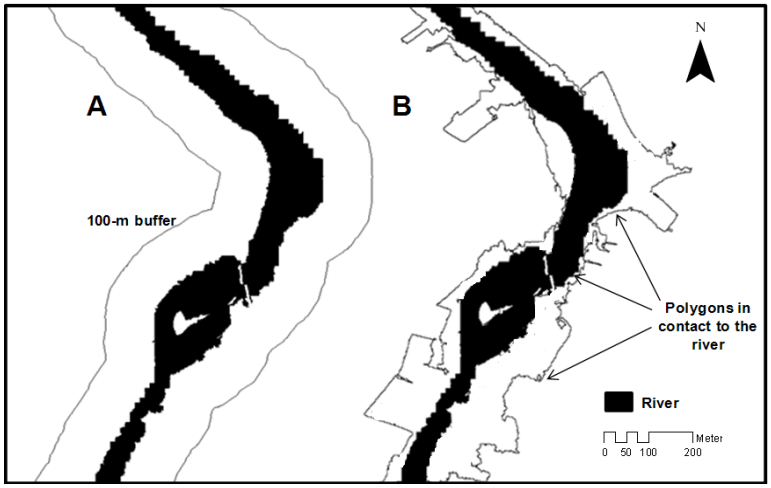


Figure 7.10 : Exemple de la délimitation d'un buffer régulier « fixe » en A et d'un buffer irrégulier s'intéressant aux objets au contact du cours d'eau en B, d'après Tormos et al., 2010.

CemOA : archive ouverte d'Irstea / Cemagref

III.3. Extraction des attributs du paysage

Pour extraire les différents attributs du paysage utilisés par les indicateurs spatialisés du corridor rivulaire, il est indispensable, avant d'engager le calcul des indicateurs, de procéder à la fusion des patchs de même nature (selon une typologie donnée) présentant une limite commune (patchs voisins) car certains indicateurs présentés précédemment prennent en compte le nombre de patchs distincts d'un type d'occupation du sol donné. L'extraction de ces attributs a été effectuée en partie en mode vecteur mais aussi en mode raster.

III.3.1. En mode vecteur

La surface, nécessaire au calcul du pourcentage surfacique, ainsi que le nombre de patch nécessaire au calcul de la fragmentation et de la densité de rupture de la ripisylve, ont été mesurés en mode vecteur à l'aide des opérateurs dimensionnels et attributaires proposés par le logiciel SIG.

III.3.2. En mode raster

Par contre, l'extraction de la largeur moyenne et de la variation de la largeur d'un patch ont été calculés en mode raster. Une matrice de distance euclidienne par rapport à la donnée hydrographique est au préalable créée à l'intérieur du buffer « contact » (fonction incrémentale). La largeur moyenne d'un patch au contact du cours d'eau correspond alors à la moyenne des distances euclidiennes dans ce patch, et la variation de la largeur à l'écart-type (fonctions de statistiques zonales).

CemOA : archive ouverte d'Irstea / Cemagref

IV. Tableau récapitulatif des familles et types spécifiques d'indicateurs spatialisés du corridor rivulaire

Tableau 7-2 : Récapitulatif des familes et types spécifiques d'indicateurs spatialisés du corridor rivulaire

	Famille/Type d'indicateurs	Type de buffer	Attribut de structure	Catégorie OS	Traitement SIG	Mesure de l'indicateur
Occupation du sol	**Le linéaire (l-Lin)**	regfixe	% surfacique (pour) ou fragmentation (frag)	pour chaque catégorie	vecteur	présence et intensité des pressions d'occupation du sol ou répartition spatiale de ces pressions (+ ou – homogène) au plus proche du cours d'eau (sur les berges)
	Le lit majeur (l-Maj)	regvar	% surfacique (pour) ou fragmentation (frag)	pour chaque catégorie	vecteur	présence et intensité des pressions d'occupation du sol ou répartition spatiale de ces pressions (+ ou – homogène) sur le lit majeur
	Le contact (l-Con)	ireg	% surfacique (pour) ou fragmentation (frag)	pour chaque catégorie	vecteur	présence et intensité des pressions d'occupation du sol ou répartition spatiale de ces pressions (+ ou – homogène) des polygones au contact du cours d'eau
Ripisylves	**La longueur**	regfixe	% surfacique (pour)	végétation semi-naturelle	vecteur	Longueur de ripisylve en amont et/ou aval de la station
	La fragmentation	regfixe	fragmentation (frag)	végétation semi-naturelle	vecteur	connectivité de la ripisylve (+ ou – fragmentée)
	La superficie	ireg	% surfacique (pour)	végétation semi-naturelle	vecteur	Superficie des polygones de ripisylve au contact du cours d'eau
	La densité de ruptures	ireg	nombre de ruptures par km (drup)	végétation semi-naturelle	vecteur	connectivité de la ripisylve (+ ou – fragmentée) par km de cours d'eau
	La largeur moyenne	ireg	largeur moyenne d'un patch (largm)	végétation semi-naturelle	raster	largeur de bande effective contre les pollutions diffuses sur de part et d'autre du cours d'eau
	L'uniformité	ireg	la variation de la largeur d'un patch (ufor)	végétation semi-naturelle	raster	régularité de la bande effective contre les pollutions diffuses de part et d'autre du cours d'eau

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Chapitre 8
Analyse et mise en œuvre des indicateurs sur le bassin versant de l'Hérault

-

SOMMAIRE

Résumé *260*

I. OBJECTIF DU CHAPITRE 262

II. INFLUENCE DE LA RESOLUTION SPATIALE SUR LES INDICATEURS 262

II.1. PRESENTATION DE L'ARTICLE *262*

II.1.1. Site d'étude et objectifs 262
II.1.2. Principaux résultats 264

II.2. ARTICLE ACCEPTÉ À PHYSICS AND CHEMISTRY OF THE EARTH (SPECIAL ISSUE: MAN AND RIVER SYSTEMS II) *264*

III. REPRESENTATIONS CARTOGRAPHIQUES DE L'ETAT DU CORRIDOR RIVULAIRE 276

III.1. SUR LA PARTIE AVAL DE L'HERAULT *276*

III.1.1. Présentation du site d'étude 276
III.1.2. Choix des indicateurs et représentations cartographiques 276

III.2. SUR LES MASSES D'EAU DU BASSIN VERSANT DE L'HERAULT *280*

III.2.1. Présentation des masses d'eau 280
III.2.2. Choix des indicateurs et représentations cartographiques 280

IV. L'INCERTITUDE DES INDICATEURS SPATIALISES 285

V. CONCLUSION 286

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Résumé

Ce chapitre met en œuvre les différents types d'indicateurs décrit dans le chapitre 7 sur le bassin versant de l'Hérault en analysant : (i) l'influence de la résolution spatiale de la donnée d'occupation du sol sur leurs valeurs; et (ii) leur capacité à fournir une information cartographique explicite pour guider les gestionnaires dans leurs stratégies de restauration sur le corridor rivulaire. Ces indicateurs ont été calculés non pas au niveau de stations écologiques mais au niveau d'entités hydrographiques linéaires particulières (tronçons d'étude ou masses d'eau par exemple), car le nombre de stations écologiques (3 stations) sur le bassin versant de l'Hérault n'est pas suffisamment important pour répondre aux objectifs de l'étude.

L'influence de la résolution spatiale sur la valeur des indicateurs a été évaluée sur plusieurs indicateurs calculés sur 10 tronçons de 8 km sur la partie aval du fleuve de l'Hérault. Sur chacun de ces tronçons ces indicateurs ont été construits à partir de deux sources de données d'occupation du sol de résolutions spatiales différentes : la donnée d'occupation du sol CORINE Land Cover (CLC) issue de la photo-interprétation d'images à moyenne résolution (20 m et 30 m) avec une unité cartographique minimum de 25 ha et la cartographie d'occupation du sol à Très Haute Résolution Spatiale (THRS) obtenue sur les 80 km du corridor rivulaire de l'Hérault à partir d'une classification orientée objet supervisée qui utilise l'information combinée des orthophotos et de l'image SPOT5 XS.

Cette étude, qui a fait l'objet d'un article dans Physics and Chemisty of the Earth, montre clairement la supériorité de l'information d'occupation du sol à THRS pour quantifier plus précisément la composition et la configuration du paysage rivulaire, notamment pour les indicateurs qui s 'intéressent aux espaces situés à proximité du cours d'eau (dans une bande de largeur inférieure à 150 m). Elle montre également que la donnée CLC ne permet pas de détecter les bandes fines arborées le long du cours d'eau et que lorsqu'une bande est détectée par CLC sa largeur est considérablement surestimée. La donnée à THRS apparait donc incontournable pour étudier les relations pressions/impact au niveau du corridor rivulaire.

Cette donnée permet également de produire des représentations cartographiques didactiques pour orienter les gestionnaires dans la définition des mesures de restauration. Les représentations cartographiques des différents indicateurs, mises en œuvre sur des tronçons de la partie aval de l'Hérault et au niveau de ces masses d'eau, illustrent cette capacité. Elles mettent en évidence l'intensité des pressions susceptibles de s'exercer sur ces différents tronçons et caractérisent l'état de la végétation rivulaire (largeur, uniformité, continuité de la végétation arborée). Au-delà de leur utilité pour la Directive Cadre sur l'Eau, le développement de tels indicateurs prépare un autre enjeu relatif à la mise en œuvre de la politique « trame verte et bleue » concernant les corridors de biodiversité à l'échelle nationale.

Par ailleurs, ce chapitre aborde également, de façon partielle, le problème de la sensibilité de ces indicateurs liée aux erreurs des cartes d'occupation du sol et à l'incertitude qui en résulte. Si les méthodologies pour étudier cette influence sont clairement identifiées, leur mise en œuvre n'a pas été menée en raison des difficultés rencontrées pour spatialiser les erreurs de la donnée d'occupation du sol, étape indispensable pour correctement quantifier cette source d'incertitude.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

I. Objectif du chapitre

L'objectif de ce chapitre est de mettre en œuvre sur le bassin versant de l'Hérault, zone test, la construction des différents types d'indicateurs décrits dans le chapitre 7, tout en analysant :

(1) l'influence de la résolution spatiale de la donnée d'occupation du sol sur leurs valeurs ;

(2) leur capacité à fournir une information cartographique explicite pour guider les gestionnaires dans leurs stratégies de restauration sur le corridor rivulaire.

Ces indicateurs ont été calculés non pas au niveau de stations écologiques mais au niveau d'entités hydrographiques linéaire particulières (tronçons d'étude ou masses d'eau par exemple) car le nombre de stations écologiques (3 stations) sur le bassin versant de l'Hérault n'est pas suffisamment important pour répondre aux objectifs de l'étude.

Dès lors, le buffer d'analyse est déterminé uniquement en fonction d'une distance latérale à cette entité hydrographique (i.e. au cours d'eau cf. Figure 8.1), la longueur de tronçon prise en compte étant la longueur de l'entité hydrographique. Ces indicateurs caractérisent l'état du corridor rivulaire en considérant un environnement plus ou moins proche de l'entité hydrographique, selon cette dimension spatiale. Au niveau terminologique (cf. Chapitre 7 :II.1), les indicateurs ne prendront donc pas de valeur pour l'attribut D_{amont} et D_{aval}.

La première analyse présentée dans ce chapitre s'est intéressée à étudier l'influence de la résolution spatiale sur des indicateurs calculés sur des tronçons de 8 km sur la partie aval du fleuve Hérault. Elle a fait l'objet d'une publication acceptée par la revue Physics and Chemistry of the Earth (PCOE) dans un numéro spécial faisant suite au colloque international « Man and River Systems II » de 2006. Le texte de la publication (en anglais) est intégralement présenté dans la seconde partie de ce chapitre.

La deuxième analyse, présentée en troisième partie de ce chapitre, propose à titre d'exemple des représentations cartographiques de certains indicateurs au niveau des masses d'eau du bassin versant de l'Hérault définies par le SANDRE (Le Service d'Administration Nationale des Données et Référentiels sur l'Eau) et sur différents cours d'eau d'un sous-bassin versant de l'Hérault.

Enfin cette étude a également analysé la sensibilité de ces indicateurs aux erreurs des cartes d'occupation du sol. Si les méthodologies pour étudier cette influence *sont clairement identifiées, leur mise en œuvre n'a pas été menée en raison des* difficultés rencontrées pour spatialiser les erreurs de la donnée d'occupation du sol, étape indispensable pour correctement quantifier cette source d'incertitude (cf. Chapitre 6 :III.).

II. Influence de la résolution spatiale sur les indicateurs

II.1. Présentation de l'article

II.1.1.Site d'étude et objectifs

Cette étude a été menée sur le corridor rivulaire (600 m de part et d'autre du cours d'eau) de la partie aval du fleuve Hérault (80 km). Ce corridor est caractérisé par de fines bandes de végétation rivulaire (< 10 m) interrompues par endroits par les territoires agricoles (notamment de la vigne) et par l'urbanisation à l'embouchure de l'Hérault (ville d'Agde).

CemOA : archive ouverte d'Irstea / Cemagref

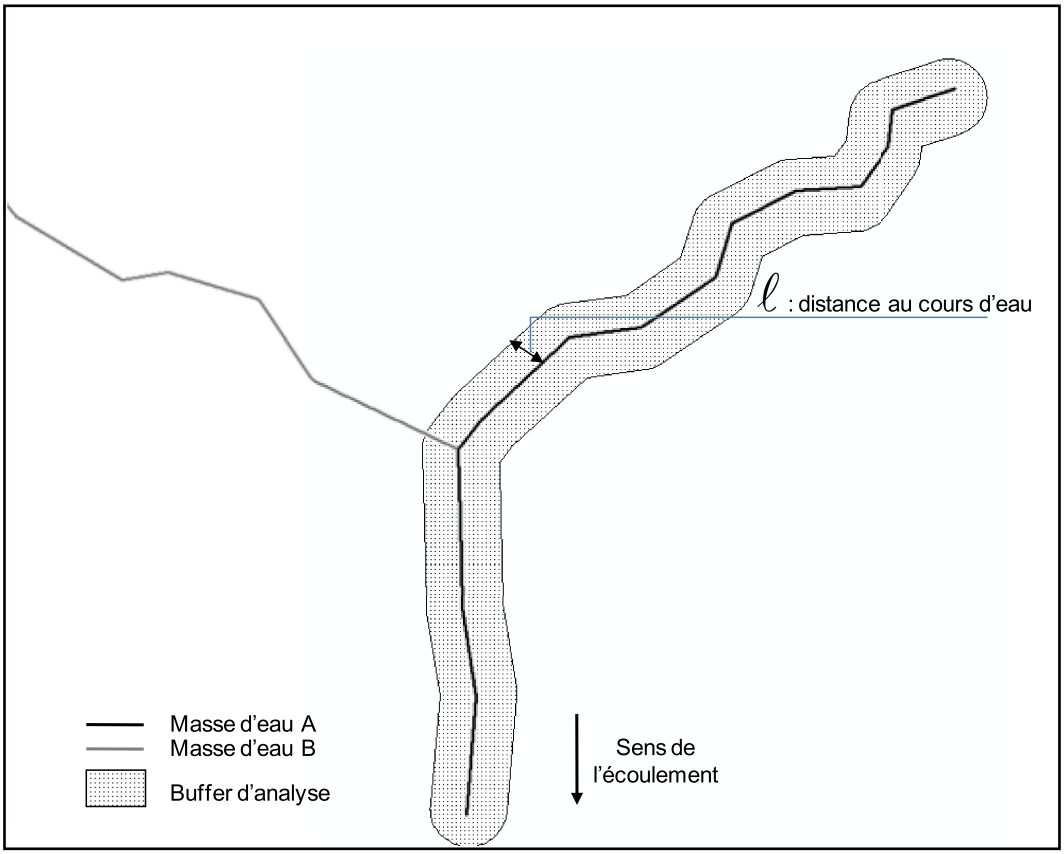


Figure 8.1 : Buffer d'analyse pour le calcul d'un indicateur spatialisé au niveau d'un tronçon de cours d'eau (pour la masse d'eau A dans cet exemple) selon une largeur l fixe sur cette illustration.

L'objectif principal de l'article est de mettre en œuvre les indicateurs du corridor rivulaire développés au chapitre 7 et d'étudier l'influence de la résolution spatiale sur l'information fournie par ces indicateurs.

Pour cela, la zone d'étude a été découpée en 10 tronçons (de A à J) de 8 km de long. Sur chacun des tronçons six types d'indicateurs spatialisés ont été construits à partir de deux sources de données d'occupation du sol de résolutions spatiales différentes :

- la donnée d'occupation du sol CORINE Land Cover (CLC) issue de la photo-interprétation d'images à moyenne résolution (20 m et 30 m) avec une unité cartographique minimum de 25 ha (Bossard *et al.*, 2000) ;

- la cartographie d'occupation du sol dans les corridors rivulaires à Très Haute Résolution Spatiale (THRS). Cette carte à THRS a été obtenue sur les 80 km du corridor rivulaire de l'Hérault (600 m de part et d'autre du fleuve) à partir d'une classification orientée objet supervisée qui utilise l'information combinée des orthophotos et de l'image SPOT5 XS (Tormos, 2005)[40].

L'article présente tout d'abord la méthodologie de classification développée pour extraire l'occupation du sol dans les corridors rivulaires à THRS (Partie 2 de cette thèse), la méthodologie de construction des indicateurs (Chapitre 7) et la méthodologie de comparaison de l'information entre CLC et la carte à THRS. Puis, il expose et discute les différents résultats obtenus.

[40] Ce travail a été réalisé lors de mon DEA Sciences de l'Eau dans l'Environnement Continental à l'UMR TETIS

CemOA : archive ouverte d'Irstea / Cemagref

II.1.2. Principaux résultats

Cet article démontre l'intérêt de la THRS et les limites des données d'occupation du sol à moyenne résolution comme CLC, pour quantifier précisément des indicateurs spatialisés aptes à caractériser l'état des corridors rivulaires et potentiellement explicatifs de l'état écologique des cours d'eau.

Premièrement, du fait de sa résolution spatiale, la carte d'occupation du sol à THRS apporte beaucoup plus d'information que CLC (par exemple alors que CLC contient uniquement 15 objets de forêt sur la zone d'étude, la carte à THRS en dénombre 2463). De plus, cette information plus précise est également plus fiable (70 % de précision globale pour la carte à THRS contre 54 % pour CLC) pour la construction des indicateurs spatialisés.

Deuxièmement, lorsqu'on étudie l'évolution de l'occupation du sol selon une distance latérale au cours d'eau sur l'ensemble de la zone d'étude, on s'aperçoit clairement de la différence d'information apportée par ces deux sources de données d'occupation du sol. Alors que l'information fournie par CLC est sensiblement la même quelque soit la distance au cours d'eau, la carte d'occupation du sol à THRS apporte une information variée en fonction de cette distance et notamment à proximité du cours d'eau où elle révèle la présence d'une végétation arborée (ripisylves). Un tel résultat montre, sans surprise, que CLC est moins précise que la carte à THRS, mais que cette différence diminue lorsqu'on s'éloigne du cours d'eau (à partir de 150 m) comme l'ont observé Schuft et al. (1999) et Lattin et al. (2004). Cette analyse préliminaire met clairement en évidence l'intérêt de l'information d'occupation du sol à THRS pour quantifier la composition et la configuration du paysage rivulaire.

Troisièmement, la comparaison de six indicateurs spatialisés construits à partir de ces deux sources de données confirme la limite de CLC pour quantifier ces indicateurs notamment ceux construits à proximité du cours d'eau. Cette donnée ne permet pas de détecter les bandes fines arborées le long du cours d'eau et lorsqu'une bande est détectée sa largeur est considérablement surestimée. Par contre, les indicateurs construits à partir de l'information à THRS révèlent la présence de ripisylves sur chacun des tronçons avec une continuité, une largeur moyenne et une uniformité variable selon les tronçons (excepté sur le tronçon J à l'embouchure de l'Hérault dominé par l'urbanisation). Cependant, une forte pression agricole est mise en évidence sur certains tronçons par l'indicateur de contact. Ces résultats montrent que le corridor rivulaire de la partie aval du fleuve Hérault est caractérisé par une ripisylve le long du cours d'eau mais qui peut être fragilisée (discontinu et irrégulière) par endroit soumettant ainsi le cours d'eau à des pressions agricoles et urbaines plus fortes.

Cet article montre clairement que l'information de l'occupation du sol à THRS est incontournable pour étudier les relations pressions-impacts au niveau du corridor rivulaire. La partie IV de cette thèse (chapitres 9 et 10) indiquera si cette information permet effectivement d'améliorer la prédiction de l'état écologique et de mettre en évidence le rôle attribué aux ripisylves à l'échelle locale à des échelles plus larges.

II.2. Article accepté à Physics and Chemistry of the Earth (special issue: Man and River Systems II)

CemOA : archive ouverte d'Irstea / Cemagref

ARTICLE IN PRESS

Physics and Chemistry of the Earth xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Physics and Chemistry of the Earth

journal homepage: www.elsevier.com/locate/pce

Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery

T. Tormos [a,*], P. Kosuth [a], S. Durrieu [a], B. Villeneuve [b], J.G. Wasson [b]

[a] Remote Sensing & Geo-information for Environment and Land Management, cemagref, 500 rue Jean François Breton, 34093 Montpellier, France
[b] Laboratory of Quantitative Hydroecology, cemagref, 3 bis quai Chauveau, CggP 220, 69336 Cedex 09, Lyon, France

ARTICLE INFO

Article history:
Received 12 September 2007
Received in revised form 5 February 2010
Accepted 17 July 2010
Available online xxxx

Keywords:
Stream ecological status
Riparian corridor
River corridor
High Spatial Resolution Imagery
Spatial indicators
Water Framework Directive

ABSTRACT

The aim of this paper is to demonstrate the interest of High Spatial Resolution Imagery (HSRI) and the limits of coarse land cover data such as CORINE Land Cover (CLC), for the accurate characterization of land cover structure along river corridors and of its functional links with freshwater ecological status on a large scale. For this purpose, we compared several spatial indicators built from two land cover maps of the Herault River corridor (southern France): one derived from the CLC database, the other derived from HSRI. The HSRI-derived map was obtained using a supervised object-based classification of multi-source remotely-sensed images (SPOT 5 XS-10 m and aerial photography-0.5 m) and presents an overall accuracy of 70%. The comparison between the two sets of spatial indicators highlights that the HSRI-derived map allows more accuracy in the quantification of land cover pressures near the stream: the spatial structure of the river landscape is finely resolved and the main attributes of riparian vegetation can be quantified in a reliable way. The next challenge will consist in developing an operational methodology using HSRI for large-scale mapping of river corridor land cover, for spatial indicator computation and for the development of related pressure/impact models, in order to improve the prediction of stream ecological status.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Preservation and restoration of the ecological quality of river ecosystems is a major social issue. It is the aim of several European Community actions such as the Water Framework Directive (WFD 2000) that provides a new legislative framework to manage, protect and restore surface waters in Europe. Prior to the definition of efficient management and restoration strategies and actions, an improved understanding of the mechanisms through which land use impacts stream ecosystems is needed (Allan, 2004).

Landscape ecology emphasizes the interaction between spatial patterns and ecological processes (Turner, 1989) and provides relevant conceptual and technical tools to establish relationships between land use and stream ecological status. Simple spatial (or landscape) indicators that describe the amount and arrangement of human-altered land in a watershed provide a direct way of quantifying man-induced pressure. They can be correlated with many stream indicators currently used to characterize river ecological status (Gergel et al., 2002), such as water chemistry and biotic variables.

A spatial indicator is invariably defined by aggregating a landscape structure attribute over a delimited area (spatial scale). Some examples of structural attributes are the number of different cover types, the proportion of each cover type, the shape of patches, and the spatial arrangement and connectivity of patches (Li and Reynolds, 1995). They are built using land cover databases and GIS tools and provide valuable environmental information in addition to traditional physical measurements (field samples and in situ data, e.g., River Habitat Survey). They are particularly useful in the context of regional monitoring schemes whose objective is to characterize the status of a large number of aquatic systems, and the pressure on them, over broad geographic regions (Jones et al., 2001).

Riparian buffer zones, or river corridors, are located at a key interface position between land and river, and provide multiple ecological goods and services (Gregory et al., 1991; Naiman and Decamps, 1997). They play a major role in river ecological status and can therefore constitute keystone units for actions towards preservation and restoration of stream quality and ecology. However, although many studies have demonstrated that upstream land use influences stream ecological status via numerous and complex pathways (Allan et al., 1997; Stewart et al., 2001; Strayer et al., 2003; Townsend et al., 2003), findings diverge concerning the relative influence of watershed vs. riparian zones on biotic status (Gergel et al., 2002). Consequently, understanding the links between stream condition and land cover at the riparian scale is a key issue for stream restoration management.

* Corresponding author.
E-mail address: thierry.tormos@onema.fr (T. Tormos).

1474-7065/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.pce.2010.07.012

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

CemOA : archive ouverte d'Irstea / Cemagref

ARTICLE IN PRESS

Frimpong et al. (2005a) tried to determine optimal riparian buffer dimensions (length and width). Efforts have been made to optimise buffer dimensions incorporated into models, but none has explicitly determined a single optimum based on both longitudinal and lateral buffer dimensions. The longitudinal dimension was conclusively determined, but the lateral dimension was optimal only with respect to the resolution of the land cover data used. Land cover maps derived from Landsat images (30 m resolution, e.g., CORINE Land Cover database) or from other moderately high resolution satellite images have been widely used to estimate the surface area of the different cover types within the riparian buffer (Goetz, 2006). However, such spatial resolutions may not be fine enough for the accurate quantification of the surface area and arrangement of land cover types along the riparian corridor, and of their impact on river ecology (Müller, 1997). For instance narrow strips (less than 10 m large) of grass or tree vegetation do not appear on these maps, although they are thought to play a significant protective role.

Progress in High Spatial Resolution Imagery (HSRI) acquisition, both from satellites and from airborne platforms with digital cameras or scanner systems, and recent developments in image classification techniques, such as object-based image analysis, offer the capacity to characterize riparian land cover in greater detail (Gergel et al., 2007; Johansen et al., 2008). Given this potential, we focused in this paper on: (i) the ability of HSRI to map land cover accurately along the river corridor; (ii) the development of synthetic spatial indicators and (iii) the influence of the spatial resolution of land cover data on these indicators. For this purpose, we developed an object-oriented classification approach to obtain land cover information at the riparian scale from HSRI, and tested it on a reach of the Herault River (France). Then, we compared the spatial information from the HSRI-derived map with CORINE Land Cover (CLC) data to highlight the interest of HRSI for precise land cover mapping at the riparian scale. Finally, after having built six synthetic spatial indicators, we analyzed the impact of differences in land cover maps (HSRI and CLC) on indicator values.

2. Data and methods

2.1. Study area

Our research was implemented on the riparian corridor of the downstream reach of the Herault River, located in southern France (Fig. 1). It is a 5th order river according to the Strahler stream system ordination (Strahler, 1952) and the total stream length of the studied reach is 80 km. This part of the Herault watershed comprises an alluvial and a littoral plain. Agriculture, especially vineyards, is the main land use type in this region. Recently, simultaneously with the development of tourism, the fast economic growth of cities has lead to the intensive urbanization of the Herault plain (Balestrat et al., 2008). The riparian corridor width is usually correlated to the width of the river bed. Studies on the Loire watershed (in France) have

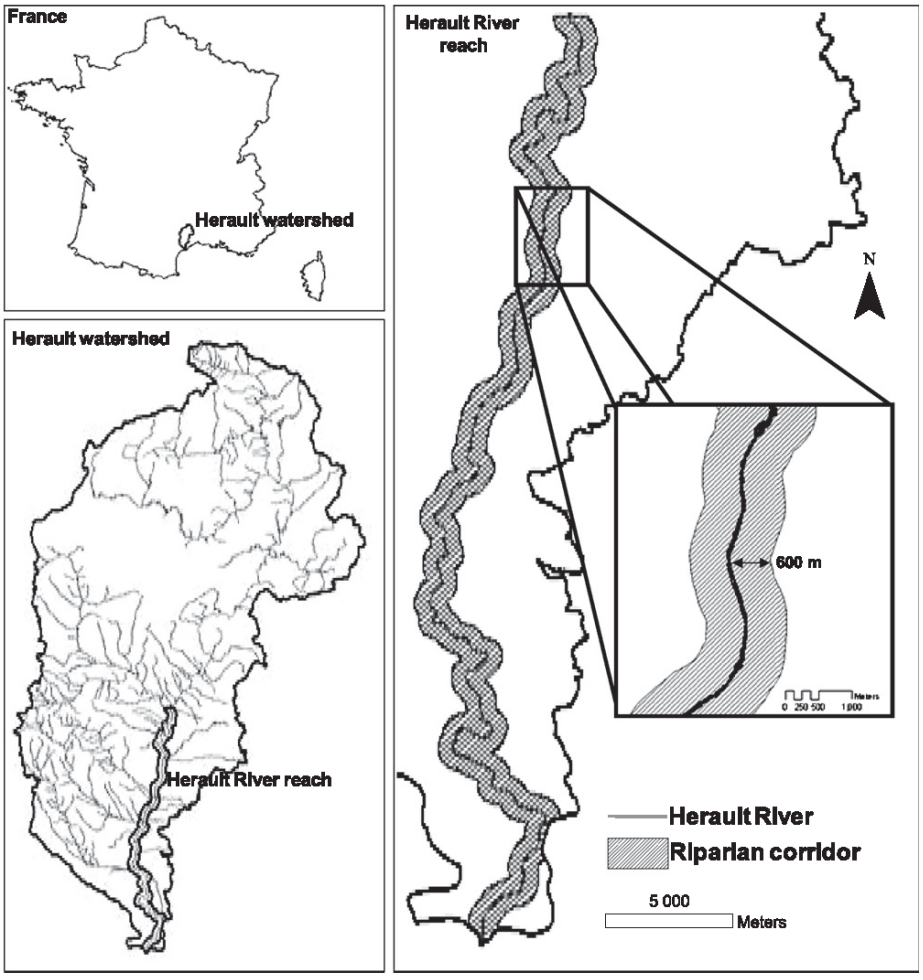


Fig. 1. Study area: the riparian corridor (right) of the downstream reach of the Herault River (bottom left), located in the South of France (top left).

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

CemOA : archive ouverte d'Irstea / Cemagref

ARTICLE IN PRESS

indicated that it is possible to estimate the average width of river beds from the Strahler stream order (Souchon et al., 2000). Following these results, riparian buffer distances along both sides of the rivers were determined according using a distance of 600 m for the 5th stream order. Data was obtained from the French hydrographic network database, the BDCarthage® produced by the French National Geographic Institute (IGN) from topographic maps (1:50 000) and Spot imagery.

2.2. Land cover pressure typology

A typology of land cover pressure was designed based on a literature review of mechanisms degrading or maintaining stream ecological status (Naiman et al., 2005) and on the recent advances in large scale analysis of the relationships between land cover and stream conditions (Allan, 2004; Hughes et al., 2006). As a result, six thematic classes of interest were defined: "C1-water surfaces", "C2-agricultural areas", "C3-urban areas", "C4-forested areas", "C5-semi-natural herbaceous vegetation" (meadow and pasture land) and "C6-natural bare soil".

C1 and C6 categories were defined in order to delineate river water bodies (stream bed and stream banks). C2 and C3 are considered as the two categories causing the main alteration of stream ecological status according to large scale statistical analyses (Allan, 2004). Agricultural practices on land adjacent to streams can lead to soil erosion and subsequent runoff of fine sediments, nutrients, and pesticides (e.g., Cuffney et al., 2000; Schulz and Liess, 1999). Urbanization leads to enhanced runoff, channel erosion, and reduced water quality due to inputs of metals, oils, and road salts (e.g., Booth and Jackson, 1997; Hammer, 1972; Paul and Meyer, 2001). Both land uses degrade the composition and abundance of riparian vegetation.

C4 and C5 are the main types of semi-natural vegetation (wooded and grassy) of the corridor that maintain biodiversity and regulate non-point source pollution (Lyons et al., 2000; Naiman and Decamps, 1997). Wooded vegetation provides stream shading, large woody debris and fine organic matter and both regulates the flux of upland derived sediments, nutrients and other chemicals and stabilizes stream banks. The forested riparian buffer was widely analysed in large-scale relationship studies and was demonstrated to play an important role on fish and macroinvertebrate communities (Stewart et al., 2001).

2.3. Data

2.3.1. High spatial resolution remotely-sensed data

Considering the spatial extent of riparian areas and the diversity of land cover types within these areas, their study requires multisource High Spatial Resolution Imagery (HRSI) data (Müller, 1997). Two HSRI data sets available on the whole French territory were chosen.

First, Aerial photographs (5×5 km^2) with 0.5 m spatial resolution and spectral information in the visible bands were collected for fine detection of riparian land cover objects. These aerial photographs are orthorectified images (orthophotos) distributed by the French national geographical agency IGN® (Paparoditis et al., 2006). Fifteen orthophotos acquired in summer 2001 were used to cover the study area.

An orthorectified SPOT 5 XS (60×60 km^2) scene with four spectral bands (B1–B4) was used in addition to the orthophotos for a more accurate discrimination of vegetation classes and extraction of water body objects (Tormos et al., 2006). Bands B1 (green: 0.50–0.59 μm), B2 (red: 0.61–0.68 μm) and B3 (near infrared: 0.78–0.89 μm) have a spatial resolution of 10 m and band B4 (mid-infrared, 1.58–1.75 μm) has a spatial resolution of 20 m. SPOT 5 XS scenes are orthorectified satellite images, available from SPOT Image©. To cover large geographic areas or specific locations, 10 m resolution SPOT 5 XS images are often the most cost effective and efficient solution. For this study we used a SPOT 5 XS archive image acquired on 2004 May 14th.

2.3.2. CORINE Land Cover data

The harmonised European land cover database, CORINE Land Cover (CLC), was built by visual interpretation of both Landsat and SPOT satellite images. Interpretation of the images is based on transparencies overlaid on 1/100.000 hard copy prints of satellites images (Bossard et al., 2000). It is based on a standard nomenclature organized into a 3-level hierarchy containing 44 classes. CLC features, characterized by a 25 ha minimum area, are either homogenous areas or combinations of land cover types with a certain recognizable structure. For this study, CLC database was aggregated according to the 6-class land cover pressure typology (see Section 2.2).

2.4. Land cover classification from HSRI

2.4.1. Image processing

When spatial resolution is increased, image spectral and spatial information becomes highly heterogeneous and conventional pixel-based classification techniques are no longer suitable to classify land cover (Durieux et al., 2008; Ivits and Koch, 2002). Blaschke et al. (2000) suggest as an efficient solution the use of object-based image analysis where, in a first step, homogeneous regions are built up through a segmentation process (segments are also called image objects) and, in a second step, a classification process is applied to these image objects using spectral as well as spatial information such as texture, shape and context features. Such methods improve the discrimination level between spectrally similar land cover types in riparian areas (Goetz et al., 2003).

The software used to implement the method in this study was eCognition4® from Definiens. It provides a complete set of tools for object-based image analysis with a multi-resolution segmentation algorithm (Baatz et al., 2000). The multi-resolution approach, based on a region-growing procedure, was used for land cover classification along the river corridor. A complete description of the region-growing algorithm can be found in (Baatz and Schäpe, 2000) or (Benz et al., 2004). The size and shape of the resulting objects can be empirically determined by the user (Blaschke and Hay, 2001).

Various classification methods (supervised or not) can be used. Their application to image objects offers several advantages compared to pixel-based methods: spectral, textural, contextual, shape and scale information associated to each object can be integrated into the classification hierarchical set of rules, or into the classification feature space for supervised classifications, to improve the quality of classification results (Benz et al., 2004).

In a first step, riparian land cover objects were delineated at a given segmentation level using both aerial photographs and PIR information from the SPOT 5 XS image.

Then, the resulting image objects were classified using a nearest neighbour (NN) supervised algorithm in order to discriminate thematic classes according to the land cover pressure typology. The optimal feature space for the nearest neighbour supervised classification method was identified using the Feature Space Optimization (FSO), a classification support tool in the eCognition software. The FSO enabled us to select an optimal set of variables from the range of available spectral, textural, contextual and shape variables, according to patterns expressed in the data set of training objects. The training data set, representing 5% of all image objects, was visually selected on images. Due to the spectral heterogeneity among training objects of a given thematic class, we had to create new "children classes" (see Table 1) for each thematic class (except for

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

CemOA : archive ouverte d'Irstea / Cemagref

ARTICLE IN PRESS

Table 1
Hierarchical structure of classes between the six thematic classes and the 12 children classes.

Thematic classes	Children classes
C1-water surfaces	Water surfaces
C2-agricultural areas	Active cultures vineyards ploughing
C3-urban areas	Building areas roads
C4-forested areas	Tree vegetation Shrub vegetation
C5-semi-natural herbaceous vegetation	Sparsely vegetated area meadow
C6-semi-natural bare soil	Dark bare soils light bare soils

C1) to avoid spectral overlapping. For example, we distinguished two types of "semi-natural bare soil" (C6) because bare soils with two distinct spectral responses exist in the Herault River plain. The class of each training object was identified in the field (May 2005) using the children class nomenclature. Finally, each image object was classified by allocating the children class of the nearest training sample object in the optimal feature space, and the resulting classification was exported in vector format after merging objects according to their land cover class (see Section 2.2).

2.4.2. Map validation

The assessment of the accuracy of the resulting land cover map was based on the analysis of a confusion matrix, computed using a set of sampled reference objects (representing 2% of all image objects) which were visually selected on images and identified on the field (May 2005) according to the land cover pressure typology. The reference objects were different from the training objects and were not used in the classification process. The confusion matrix was built according to pixels contained in the reference objects using Erdas Imagine® software. The accuracy assessment reports three values (Congalton, 1991): (1) the overall accuracy which indicates the proportion of pixels that are correctly classified by the method; (2) the user's accuracy which is a statistic that indicates to the map user which percentage of a map class actually corresponds to this class and (3) the producer's accuracy which is a statistic that indicates for each class which percentage of the reality is correctly classified in the map.

2.5. GIS analysis

Two GIS analyses were conducted. A preliminary analysis was performed in order to compare the relevance of information from the CLC- and from the HSRI-derived map (see Section 2.5.2), before developing the spatial land cover pressure indicators from HSRI-derived map (see Section 2.5.3). These analyses required specific GIS processing presented in Section 2.5.1.

2.5.1. GIS processing

The domain over which spatial indicators are computed at the riparian scale is generally defined by combining a lateral distance to the river with longitudinal distances upstream and downstream from the ecological station where stream ecological status is measured. These distances are often chosen arbitrarily (Frimpong et al., 2005b). However, as no ecological station exists on this part of the Herault River, spatial indicators were computed on 10 sections (A–J) of equal length (8 km). As a result, the GIS analysis focussed on the lateral distance to the river. It was conducted following a three step approach, using ArcGis® software.

First, the river bed was delineated for a precise delimitation of the lateral distance. For this purpose we selected from the HSRI-derived map all objects classified as "water surfaces" (C6 class) and located near the Herault River hydrographic network (i.e., water object intersected by a 30-m buffer along the hydrographic network). This selection was visually checked on orthophotos and cor-

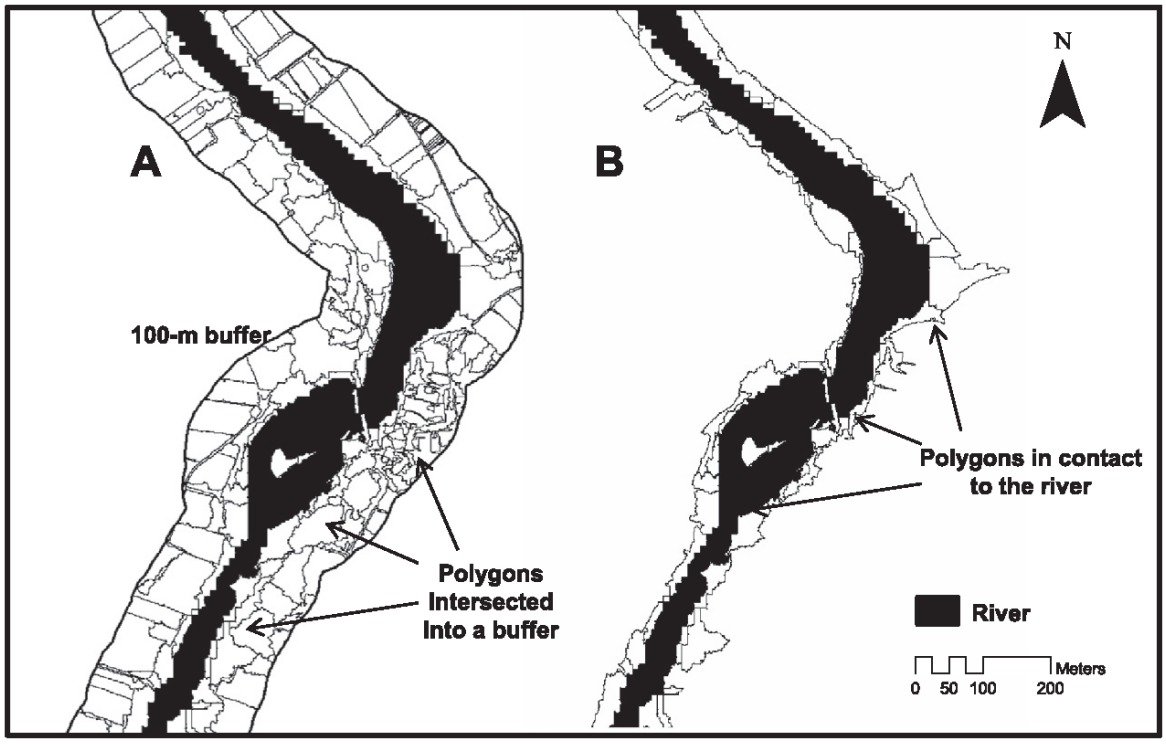


Fig. 2. Presentation of the two GIS techniques for the extraction of relevant polygons according to a lateral distance to the river bed. In A, the "buffer technique": first a buffer is computed along the river bed according to a given distance and then land cover polygons are clipped into the buffer limits. In B, the "contact technique": all polygons of a given land cover type in contact (having a common border) with the river bed are selected.

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

CemOA : archive ouverte d'Irstea / Cemagref

Table 2
Characteristics of the six spatial indicators developed in this study.

Name	Landscape metric	Riparian scale	Land cover categories	Pressures variables meaning
Linear Spatial Indicator (LSI)	Surface area (in percent)	10-m buffer	For each land cover category	Presence and intensity of land cover pressures near the stream
Floodplain Spatial Indicator (FSI)	Surface area (in percent)	600-m buffer	For each land cover category	Presence and intensity on the floodplain
Contact Spatial Indicator (CSI)	Surface area (in percent)	Polygons in contact to river	For each land cover category	Presence and intensity of patterns directly in contact to the river.
Forested Riparian zone Continuity (FRC)	Fragmentation (numbers of patches/km)	Polygons in contact to river	Forested areas (C4)	Discontinuity (number of gaps) of the forested riparian strip
Average Forested Riparian zone Width (AFRW)	Mean width (in meters)	Polygons in contact to river	Forested areas (C4)	Average width of the riparian buffer strip
Forested Riparian zone Uniformity (FRU)	Standard deviation of width (in meters)	Polygons in contact to river	Forested areas (C4)	Uniformity of the riparian buffer strip

rected when necessary before merging all polygons in order to construct a uniform polygon representing the Herault River bed.

Then, relevant land cover polygons were selected according to their lateral distance to the river bed, using two techniques: (i) by clipping land cover data according to the limits of a buffer with a given distance from the river bed polygon (the "buffer technique" – Fig. 2A) and (ii) by selecting a given type of land cover polygons in contact (having a common border) with the river bed ("the contact technique" – Fig. 2 B). While the first technique defines a fixed lateral distance, the second defines a variable one. Depending on the technique employed and the lateral distance chosen for the first technique, the meaning of the spatial indicator is different (see Section 2.5.3).

Finally, structural attributes of these polygons (e.g., area, perimeter...) were extracted using traditional GIS techniques and specific processing to quantify the width distribution (transversal to the river) of polygons in contact with the river.

2.5.2. Comparison of land cover maps

To compare information from CLC and the HSRI-derived map on the study area, we first computed the confusion matrix for the CLC map, using the same sample of reference objects as for the HSRI map (see Section 2.4.2) (comparison of accuracy between the two maps). Then, we analysed the land cover composition (area percentage of each land cover class) according to different buffer widths (lateral distances) along the whole Herault River reach (80 km). Two ranges of lateral distances were defined. The first range aims to analyse pressure characteristics in the immediate proximity of the river (5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 m). The second range concerns the vicinity of the river (100, 150, 200, 250, 300, 350, 400, 450, and 500 m). The lateral distance was delimited using the buffer technique (see Fig. 2A).

2.5.3. Spatial indicators of land cover pressure

Six spatial indicators characterizing the river corridor land cover where used as potential measurements of pressure on the stream ecological status. Spatial indicators were developed for each section of Herault River main reach (10 section from A to J of equal length). Table 2 summarizes indicator characteristics and the GIS processing used to construct them (see 3.3.1).

The first three spatial indicators are of the "area percentage" type. They focus on the proportion of a given land cover type along the river corridor in the delimited area: the Linear Spatial Indicator (LSI), the Floodplain Spatial Indicator (FSI), and the Contact Spatial Indicator (CSI). Area percentages, indicating the presence and intensity of land cover pressure are the main landscape structure attributes used at the riparian scale in studies of large-scale relationships and remain the main explanatory land cover variable of stream condition. While the LSI represents the presence and intensity of pressure on a buffer close to the stream (10 m width), the FSI quantifies it on the whole river corridor (600 m width). In contrast the CSI is not restrained to a buffer and deals with pressure polygons directly in contact with the river.

Three additional indicators were designed to characterize the forested component of the riparian buffer more specifically. In many streams, the presence of forested riparian buffer strips can efficiently reduce groundwater nitrogen loads and surface runoff phosphorus loads (Naiman and Decamps, 1997). Furthermore, the spatial patterns of forested riparian zones may influence their ability to act as nutrient sinks. Thus indicators that characterize the spatial pattern of forested riparian areas can be useful. A simple model of an upland contributing area and a forested riparian buffer by Weller et al. (1998) explored the relationship between the spatial configuration and the nutrient retention capacity of riparian zones. According to the results of this heuristic model, we defined: (i) the Average Forested Riparian zone Width (AFRW) which quantifies approximately the average buffer width of forested riparian strips, and is the best predictor of lateral runoff for non-retentive buffers; (ii) the Forested Riparian zone Uniformity (FRU) which indicates the variability (the uniformity) of this forested buffer width. Irregular buffer widths are less efficient than uniform buffer widths because transfer through gaps dominates lateral runoff and (iii) the Forested Riparian zone Continuity (FRC) which indicates the frequency of gaps (the fragmentation or discontinuity) of the forested riparian strip and is the best predictor of lateral runoff for narrow, retentive buffers.

3. Results and discussion

3.1. Land cover classification from HSRI

Table 3 shows the error matrix, overall accuracy and both user's and producer's accuracies of the HSRI derived land cover classification. The error matrix is expressed in number of pixels. Lines represent classes in reality; while columns represent classes on the map: each pixel thus appears in one cell. Overall accuracy was 70%. All "C1-water surfaces" objects were correctly classified. Accurate results were obtained for "C2-agricultural areas" and "C4-forested areas" too (user's and producer's accuracies are 88% and 73%, and 74% and 64%, respectively). "C3-urban areas" presents a lower accuracy (user's and producer's accuracy are 46% and 57% respectively). The accuracy of the map obtained with the object-based method is poor for both "C5-semi-natural herbaceous vegetation" and "C6-semi-natural bare soils" classes (user's and producer's accuracies are 27% and 37%, and 9% and 40%, respectively).

In the case of the C5 class, both commission and omission errors were encountered especially with "C2-agricultural areas". Confu-

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

CemOA : archive ouverte d'Irstea / Cemagref

Table 3
Error matrix and accuracy values of riparian land cover classification from HSRI (in number of pixels, 0.5 m pixel size).

Reality/map	C1-water surf.	C2-agr. areas	C3-urb. areas	C4-For. areas	C5-herb. veg.	C6-bare soils	Total
C1-water surf.	**432,463**	0		0	0	0	432,463
C2-agr. areas	0	**1,042,961**	12,338	75,628	167,680	129,309	1,427,916
C3-urb. areas	0	21,426	**76,363**	30,851	3967	2288	134,895
C4-For. areas	0	11,335	30,152	**370,075**	143,726	20,189	575,477
C5-herb. veg.	0	102,162	35,654	14,619	**114,776**	41,689	308,900
C6-bare soils	0	10,301	13,072	6 342	0	**19,663**	49,378
Total	432,463	1,188,185	167,579	497,515	430,149	213,138	2,929,029
User's accuracy (%)	100.00	87.78	45.57	74.38	26.68	9.23	
Producer's accuracy (%)	100.00	73.04	56.61	64.31	37.16	39.82	
Overall accuracy (%)	70.20						

Table 4
Error matrix and accuracy values of the CORINE Land Cover map (in number of pixels, 0.5 m pixel size).

Reality/map	C1-water surf.	C2-agr. areas	C3-urb. areas	C4-For. areas	C5-herb. veg.	C6-bare soils	Total
C1-water surf.	**0**	169,782	0	262,681	0	0	432,463
C2-agr. areas	0	**13,71,148**	0	56,768	0	0	14,27,916
C3-urb. areas	9008	78,045	**34,608**	13,234	0	0	134,895
C4-For. areas	49,339	335,548	0	**190,590**	0	0	575,477
C5-herb. veg.	0	303,287	0	5613	**0**	0	308,900
C6-bare soils	0	38,827	10,301	250	0	**0**	49,378
Total	58,347	22,96,637	44,909	529,136	0	0	29,29,029
User's accuracy (%)	0.00	59.70	77.06	36.02	0.00	0.00	
Producer's accuracy (%)	0.00	96.02	25.66	33.12	0.00	0.00	
Overall accuracy (%)	54.50						

sion between these two classes is probably due to the fact that the SPOT 5 image was acquired during the spring season, at a time of year when some crops (e.g., wheat, barley) show a spectral behaviour similar to semi-natural grass cover. Using satellite images acquired during the summer season may help to avoid spectral confusion between these classes. Grassy or herbaceous vegetation in riparian corridors can have a beneficial impact on stream status and may be appropriate restoration options in some situations, especially in agricultural regions (Lyons et al., 2000). Therefore, proper identification of this class is of major importance. For the C6 class, the low accuracy originates from strong confusions with "C2-agricultural areas" and "C3-urban areas". Spectral behaviour of bare soils can be similar to that of ploughings, vineyards and house roofs (tile cover). The use of specific textural attributes could improve the discrimination between these classes, thus leading to an increased accuracy. Furthermore, the lagtime between remotely-sensed image acquisition (May 2004 for SPOT 5 XS image and summer 2001 for orthophotos) and reference data set acquisition (May 2005) can also explain part of the misclassifications.

In the perspective of large-scale mapping of land cover along river corridors, classification methods have to be cost-effective and easy to reproduce in distinct areas. However, implementation of the nearest neighbour (NN) technique on the Herault River reach was very time-consuming for the definition and the in situ acquisition of training sample objects and the classification rules automatically inferred by the FSO (see Section 2.4.1.) were hard to understand and to refine.

Consequently, the NN method seems to be unsuitable for large-scale mapping. On the contrary, rule-based classification appears more adapted. In this approach, a logical decision tree is built based on feature space and physical features of each class that can be applied in other places. It allows easy integration of remotely-sensed data with other sources of geo-referenced information, such as land use data and spatial texture, to obtain higher classification accuracy, especially for herbaceous vegetation. All rules can be refined with full user control, at any time in the classification process and, in most cases, without changing the class allocation of other objects (Lucas et al., 2007). This was implemented on the study area and on the overall Herault River hydrographic network (1500 km long), but is not presented here.

Table 5
Comparison of the number of objects for each land cover type for both CORINE Land Cover (CLC) and High Spatial Resolution Imagery (HSRI)-derived maps.

	Object number		Object percentage	
	CLC	HRSI	CLC (%)	HRSI (5)
C1-water surfaces	5	320	8.47	3.00
C2-agricultural areas	6	2947	10.17	27.59
C3-urban areas	28	2792	47.46	26.14
C4-forested areas	15	2463	25.42	23.06
C5-semi-natural herbaceous vegetation	2	1821	3.39	17.05
C6-semi-natural bare soil	3	338	5.08	3.16
Total	59	10,681		

3.2. Comparison of land cover maps

When comparing the number of polygons (objects) for each land cover type between CORINE Land Cover (CLC) and the HSRI-derived map (see Table 5 and Fig. 8), it clearly appears that, due to its finer spatial resolution, the HSRI map provides much more information than the CLC database. For example, only 15 polygons of "C4-forested areas" are detected with the CLC database against 2463 with the HSRI-derived map, which means that the CLC resolution is much too low to detect the heterogeneity of forested areas in riparian areas. Furthermore the different classes in the CLC map appear not to be aggregated in the same way: C2-agricultural areas appears to be strongly aggregated (10% of objects in CLC against 28% in HSRI) while C3-urban areas is under aggregated (47% in CLC against 26% in HSRI). This will have an impact on spatial indicators that depends on the relative number of objects.

Table 4 shows the error matrix, overall accuracy and both user's and producer's accuracies of the CLC map. Overall accuracy was 54%. "C5-semi-natural herbaceous vegetation" and "C6-semi-natural bare soils" reference objects were not detected by CLC (no pixel

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

CemOA : archive ouverte d'Irstea / Cemagref

ARTICLE IN PRESS

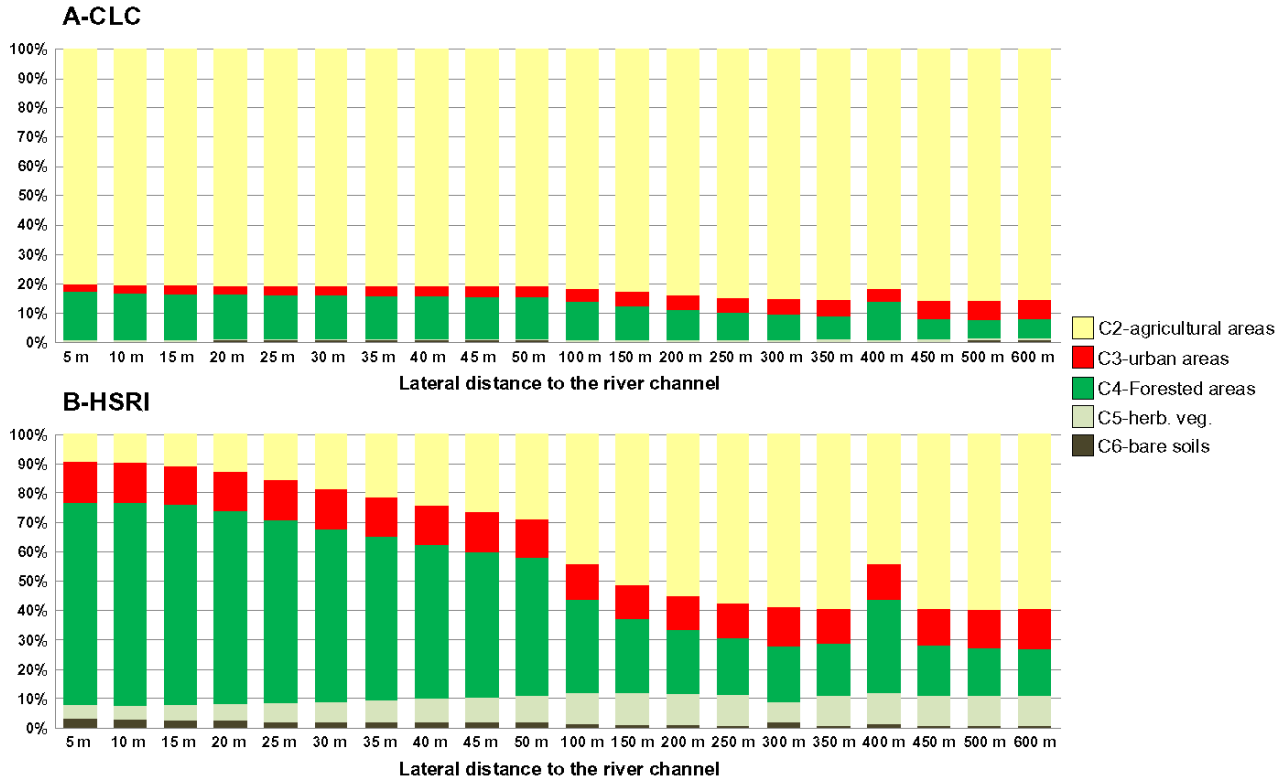


Fig. 3. Variation of the area percentage of a given land cover type according to the lateral distance to the river channel on the Herault River reach (A) from CORINE Land Cover (CLC) database, (B) from the High Spatial Resolution Imagery (HSRI)-derived map.

in the confusion matrix) and all "C1-water surfaces" objects were incorrectly classified (both user's and producer's accuracies were 0%). Highest accuracy results were obtained for "C2-agricultural areas" and "C3-urban areas" (user's and producer's accuracies were 59% and 96%, and 77% and 26%, respectively). "C4-forested areas" presented a lower accuracy (user's and producer's accuracy were 36% and 33% respectively).

Accuracy results of the CLC map are lower than those of the HSRI-derived map and confirm the limits of coarse land cover data for the accurate mapping of land cover along a river corridor. The "C4-forested areas" class is poorly characterized by the CLC map due to the strong confusion with "C1-water surfaces" and "C2-agricultural areas". Although a high percentage of the "C3-urban areas" class from CLC corresponds to the reality, a low percentage of C3 reference objects is detected by CLC. On the contrary, because of the predominance of the "C2-agricultural areas" class along the riparian corridor of the Herault River, the majority of C2 reference objects are correctly detected (96%) by CLC, while only 59% of CLC C2 objects actually correspond to the reality.

In Fig. 3, the evolution of the land cover composition (area percentage indicator) according to the lateral distance from the Herault River (buffer width from 5 m to 600 m) shows that major landscape patterns inside the riparian area which are revealed by HSRI maps, are totally smoothed in CLC maps. In the case of CLC (Fig. 2A), spatial information does not significantly change with the buffer width whatever the land cover class. The "C2-agricultural areas" class is predominant both at the contact and in the vicinity of the river. In contrast, the HSRI-derived map (Fig. 2B) highlights a spatial change in the land cover structure according to the lateral distance from the river channel, especially for "C4-forested areas" and "C2-agricultural areas" classes. Near the river (between 5 and 50 m on both sides of the river) C4 is predominant, while C2 is dominant for riparian buffer zones larger than 100 m. Such results prove that the CLC database is, as expected, less accurate than the HSRI-derived map, and that the differences between the maps decrease as buffer width increases (up to 150 m). This is of major importance regarding linear corridors whose structure cannot be correctly described by CLC data. Similar results have been presented in the literature (Lattin et al., 2004; Shuft et al., 1999). It is most likely that, in many situations within river corridors, the generalisation level of the CLC database will lead to under- or overestimation of the effect of cover types which impact on stream chemical and biotic metrics (Goetz, 2006). This work clearly demonstrates the interest of HSRI information for improving the quantification of land cover amount and arrangement patterns at the riparian scale.

3.3. Spatial indicators of land cover pressures

Figs. 4–7 show results of synthetic spatial indicators computed from both land cover maps (CLC and HSRI) for the 10 sections of the Herault River reach.

First, information given by each spatial indicator differs according to the land cover data used (CLC or HSRI) except for the Floodplain Spatial Indicator (FSI). For instance, the Linear Spatial Indicator (LSI) is significantly different for the two land cover maps: while "C4-forested areas" is predominant with the HSRI land cover map, "C2-agricultural areas" is dominant with the CLC database on most of the sections. In contrast, the FSI shows a predominance of agricultural areas (C2) for both sets of land cover data. These results confirm the occurrence of large differences from both land cover data, which get blurred with higher buffer widths. Furthermore, when forest areas (C4) are detected in contact with the river using the CLC database, we observe that cover areas are

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

CemOA : archive ouverte d'Irstea / Cemagref

ARTICLE IN PRESS

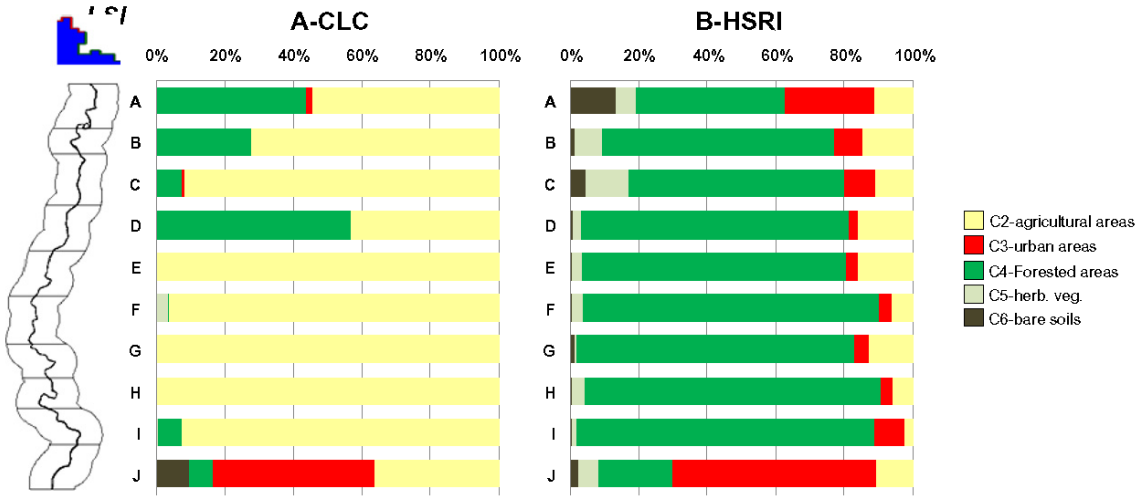


Fig. 4. Linear Spatial Indicator (LSI): the surface area percentage of a given land cover category extracted on a 10-m buffer. In A from CORINE Land Cover (CLC) database, in B from the High Spatial Resolution Imagery (HSRI)-derived map.

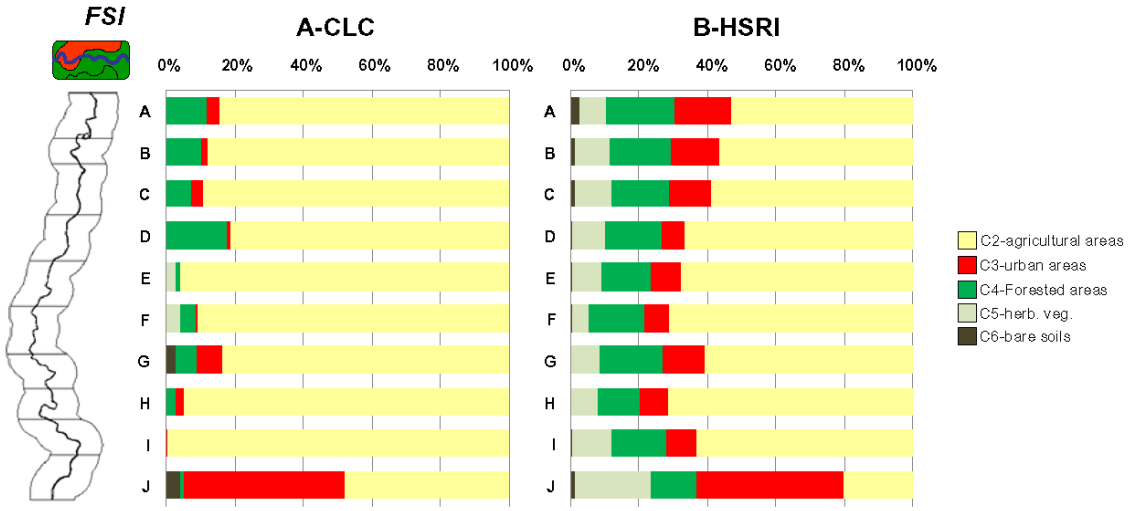


Fig. 5. Floodplain Spatial Indicator (FSI): the surface area percentage of a given land cover category extracted on a 600-m buffer. In A from CORINE Land Cover (CLC) database, in B from the High Spatial Resolution Imagery (HSRI)-derived map.

consistently overestimated: the value of the Average of Forest Riparian Width (AFRW) indicator extracted from the CLC database is always overvalued compared to the HSRI map. Fig. 8 compares maps from both land cover data sets on a fraction of section D. The overestimation of the forested riparian strip by CLC is obvious, while the HSRI map better agrees with reality.

The analysis of HSRI-derived indicators on river sections gives interesting and complementary information on the presence and the intensity of pressure and on the characteristics of the forested riparian strip along the Herault River reach.

The LSI (Fig. 4B) indicates a high presence of forested areas (C4) on the riparian strip close to the stream for all sections except for section J which is dominated by urban areas due to the presence of Agde city at the river mouth.

In addition, the CSI (Fig. 6B) specifies the intensity of pressure bordering the stream. Although the CSI confirms the impact of urban land cover on section J and the presence of a forested riparian strip on the other ones, it also reveals and quantifies the existence of gaps along this forested band. Sections D, E, F, G, and I are potentially affected by nutrient, sediment, and pollutant fluxes from agricultural plots in direct contact with the river. Fig. 7B shows indicators related to attributes of the forested riparian strip and the occurrence of gaps in each of the sections is highlighted by the Forested Riparian zone Continuity (FRC) indicator.

The nitrogen retention capacity of the riparian buffer zone could be assessed using both the Average Forested Riparian zone Width (AFRW) and the Forested Riparian zone Uniformity (FRU). A, G and I sections are characterized by a wide forested riparian strip along the river (AFRW is 148 m, 125 m and 147 m, respectively) but this strip is more uniform on G (FRU is 63) than on A and I (FRU is 102 m and 89 m, respectively).

Although many other factors (e.g., climate, topography, land use, stream type and structure, slope, soil type, and drainage characteristics) influence (i) the volume of sediment and agricultural chemicals reaching the stream network and (ii) the riparian functions (Naiman et al., 2005), this set of indicators is liable to provide explanatory variables for assessing stream ecological status. The detailed land cover map provides considerable flexibility for the

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

CemOA : archive ouverte d'Irstea / Cemagref

ARTICLE IN PRESS

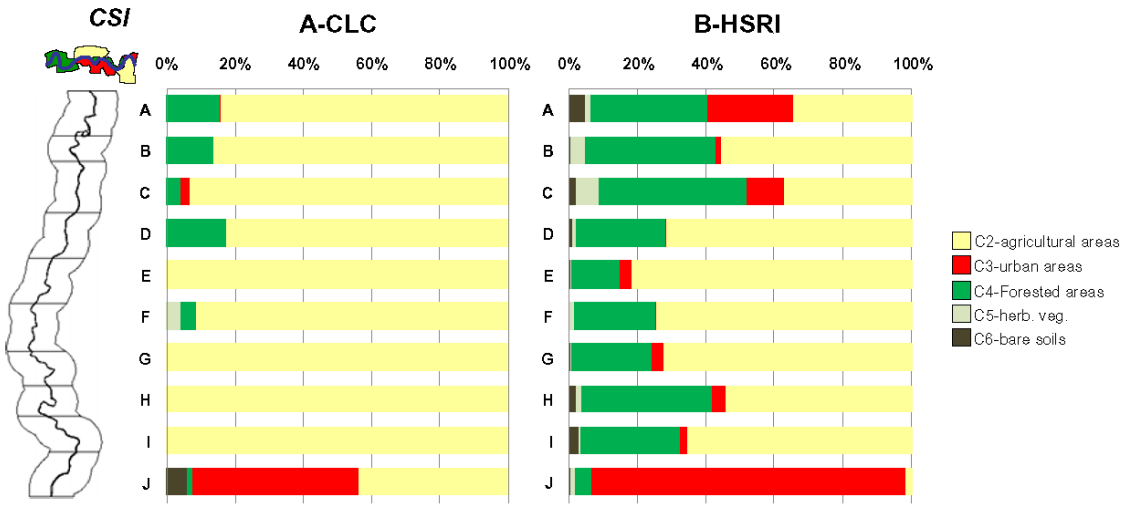


Fig. 6. Contact Spatial Indicator (CSI): the surface area percentage of a given land cover category in contact with the river channel. In A from CORINE Land Cover (CLC) database, in B from the High Spatial Resolution Imagery (HSRI)-derived map.

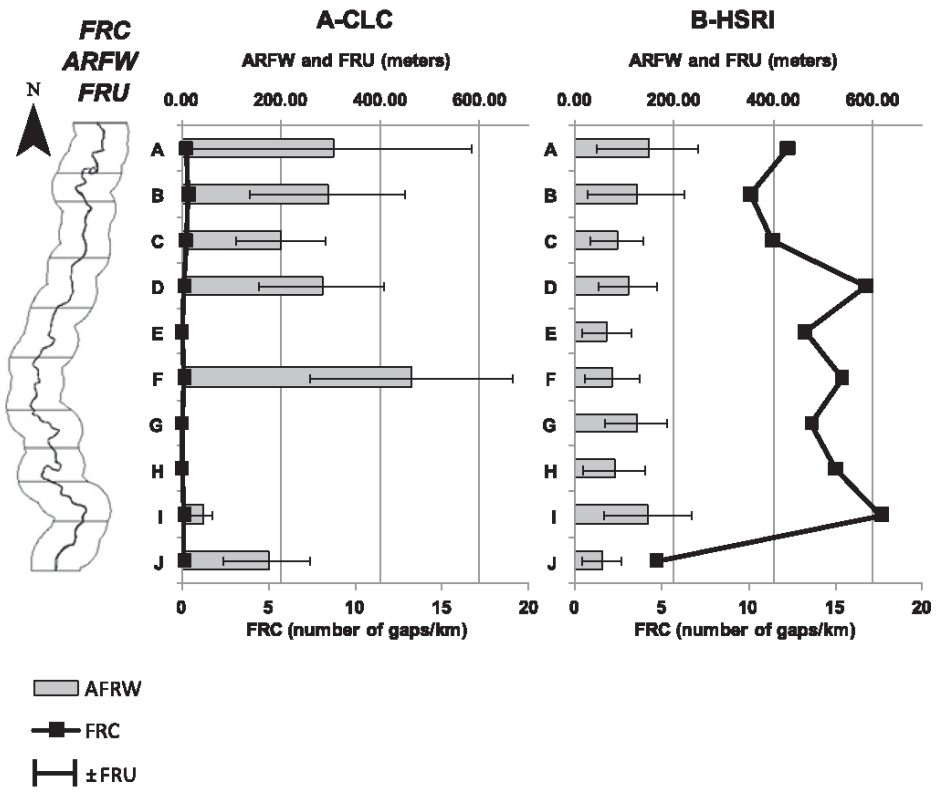


Fig. 7. Forested Riparian zone indicators: the Forested Riparian zone Continuity (FRC): the number of forested areas patches/km in contact with the river; the Average Forested Riparian zone Width (AFRW): the average of the forested riparian strip width (in meters); and the Forested Riparian zone Uniformity (FRU): the variability (standard deviation) of the forested riparian strip width (in meters). In A from CORINE Land Cover (CLC) database, in B from the High Spatial Resolution Imagery (HSRI)-derived map.

calculation of various structure attributes of land cover pressure and riparian vegetation along rivers. However, the error in land cover classification will have to be analysed in detail to associate uncertainty values to these indicators (Gergel et al., 2007).

Based on this approach some properties of the river corridor can be examined more precisely, like for instance the continuity (connectivity) of the forested riparian strip, considered as a key attribute in assessing riparian functions (Naiman and Decamps, 1997). Additionally to the number of gaps (FRC), the longitudinal extent, mean length, and maximum length of gaps could be assessed. Such complementary information highlights properties of riparian connectivity. Indeed, while a number of short gaps in the forested riparian strip may have little effect on stream temperature, they may be significant channels for direct transport to the stream of upland nutrients, sediments and pollutants (Naiman and Decamps, 1997). On the other hand, large gaps in forested veg-

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

CemOA : archive ouverte d'Irstea / Cemagref

ARTICLE IN PRESS

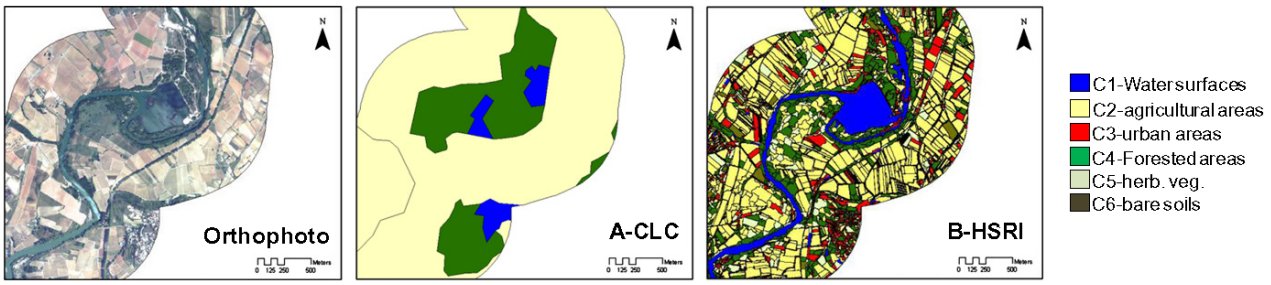


Fig. 8. Comparison with the orthophoto on a part of the section D of the results in A from CORINE Land Cover (CLC) and in B from the High Spatial Resolution Imagery (HSRI)-derived map. Note the overestimation of forested areas that affects riparian vegetation strip indicators.

etation, approaching 1 km in length, have been demonstrated to have significant effects on stream temperature and trout distribution (Barton et al., 1985).

4. Summary and conclusion

This study, based on an implementation on the Herault River, has demonstrated that High Spatial Resolution Imagery (HSRI) using object-based image analysis allows the extraction of land cover information along the river corridor with an acceptable degree of confidence (70% of pixels well classified according to ground reference data), showing far more accuracy and detail than CLC data (54% of pixels well classified according to ground reference data). The comparison of land cover information and spatial indicators computed from the CORINE Land Cover (CLC) database and from the HSRI-derived map attests that CLC is too coarse to quantify land cover pressure and riparian vegetation properties near the stream (along a 150 m wide strip) in a relevant way. Spatial indicators derived from the HSRI map provide accurate and original information on the presence and intensity of pressure close to the stream and on forested riparian strip attributes (uniformity, mean width and continuity). Detailed land cover data introduced in GIS allows large flexibility and reliability for the calculation of spatial indicators, which aim to characterize the mechanisms through which land cover along river corridors impacts stream ecosystems.

Given these results, new challenges have emerged in view of gaining a better understanding of the relationships between land cover pressure and stream ecological status.

The first challenge consists in the development of a rule-based object-oriented classification method for large-scale mapping of land cover along river corridors using HSRI. Because supervised classification techniques (such as nearest neighbour used in this study) are strongly dependent on the training sample which is time-consuming to define and collect, we argue that rule-based classification would be more a more efficient technique to map land cover on large territories.

The second challenge consists in: (i) the definition of a set of complementary spatial indicators related to the riparian functions and to the mechanisms of land cover pressure on a given stream condition (chemical, physical or biotic) and (ii) the development of GIS techniques to define these indicators automatically. Sensitivity of these spatial indicators to the error in land cover classification will have to be analysed in detail and resulting uncertainty values will be associated to these indicators.

Finally, the last challenge consists in the study and modelling of the ecological (or chemical) response to spatial indicators derived from the HSRI map and other pressure variables, in order to identify the minimum width and composition of riparian vegetation that should be recommended on a regional or national scale to maintain river corridor functions for stream integrity and ecological status.

Acknowledgments

This study has been made possible by a PhD Grant from the French Ministry of Research. The work has been partly realised in the framework of a research contract between ONEMA and Cemagref. Acquisition of remote sensing data benefited from the support of the ISIS programme of the French national space agency CNES (Spot 5 images), while IGN airborne orthophotos were made available by the French Ministry of Agriculture.

References

Allan, J.D., 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annual Review of Ecology, Evolution, and Systematics 35, 257–284.
Allan, J.D., Erickson, D., Fay, J., 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater Biology 37, 149–161.
Baatz, M., Schäpe, A., 2000. Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. Angewandte Geographische Informationsverarbeitung 12, 12–23.
Baatz, M., Heynen, M., Hofmann, P., Lingenfelder, I., Mimler, M., Schäpe, A., Weber, M., Willhauck, G., 2000. ECognition User Guide. Definiens AG, München. p. 483.
Balestrat, M., Chéry, J.-P., Valette, E., 2008. Suivi des changements d'occupation et d'utilisation du sol pour la compréhension des dynamiques périurbaines, XLV colloque de l'Association Régionale de Langue Française, Rimouski (Québec).
Barton, D.R., Taylor, W.D., Biette, R.M., 1985. Dimension of riparian buffer strips required to maintain trout habitat in southern Ontario streams. North American Journal of Fisheries Management 5, 364–378.
Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing 58, 239–258.
Blaschke, T., Hay, G.J., 2001. Object-oriented image analysis and scale-space. Theory and methods for modeling and evaluating multiscale landscape structures. International Archives of Photogrammetry and Remote Sensing 34, 22–29.
Blaschke, T., Lang, S., Lorup, E., Strobl, J., Zeil, P., 2000. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. Umweltinformation Für Planung, Politik und Öffentlichkeit/environmental information for planning. Politics and the Public 2, 555–570.
Booth, D.B., Jackson, C.R., 1997. Urbanization of aquatic systems: degradation thresholds, stormwater detection, and the limits of mitigation. Journal of the American Water Resources Association 33, 1077–1090.
Bossard, M., Feranec, J., Otahel, J., 2000. CORINE Land Cover Technical Guide-Addeendum 2000, Technical Report No. 40. European Environment Agency.
Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37, 35–46.
Cuffney, T.F., Meador, M.R., Porter, S.D., Gurtz, M.E., 2000. Responses of physical, chemical, and biological indicators of water quality to a gradient of agricultural land use in the Yakima River Basin, Washington. Environmental Monitoring and Assessment 64, 259–270.
Durieux, L., Lagabrielle, E., Nelson, A., 2008. A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. ISPRS Journal of Photogrammetry and Remote Sensing 63, 399–408.
Frimpong, E.A., Sutton, T.M., Lim, K.J., Hrodey, P.J., Engel, B.A., Simon, T.P., Lee, J.G., Le Master, D.C., 2005a. Determination of optimal riparian forest buffer dimensions for stream biota- landscape association models using multimetric and multivariate responses. Canadian Journal of Fisheries and Aquatic Sciences 62, 1–6.

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

CemOA : archive ouverte d'Irstea / Cemagref

ARTICLE IN PRESS

Frimpong, E.A., Sutton, T.M., Lim, K.J., Hrodey, P.J., Engel, B.A., Simon, T.P., Lee, J.G., Le Master, D.C., 2005b. Determination of optimal riparian forest buffer dimensions for stream biota landscape association models using multimetric and multivariate responses. Canadian Journal of Fisheries and Aquatic Sciences 62, 1–6.

Gergel, S.E., Turner, M.G., Miller, J.R., Melack, J.M., Stanley, E.H., 2002. Landscape indicators of human impacts to riverine systems. Aquatic Sciences 34, 118–128.

Gergel, S.E., Stange, Y., Coops, N.C., Johansen, K., Kirby, K.R., 2007. What is the value of a good map? An example using high spatial resolution imagery to aid riparian restoration. Ecosystems 10, 688–702.

Goetz, S.J., 2006. Remote sensing of riparian buffers: past progress and future prospects. Journal of the American Water Resources Association 42, 133–143.

Goetz, S.J., Wright, R., Smith, A.J., Zinecker, E., Schaub, E., 2003. Ikonos imagery for resource management: tree cover, impervious surfaces and riparian buffer analyses in the mid-Atlantic region. Remote Sensing of Environment 88, 195–208.

Gregory, S.V., Swanson, F.J., McKee, W.A., Cummins, K.W., 1991. An ecosystem perspective of riparian zones. BioScience 41, 540–551.

Hammer, T.R., 1972. Stream channel enlargement due to urbanization. Water Resource Research 8, 1530–1540.

Hughes, R.M., Wang, L., Seelbach, P.W., 2006. Landscapes Influences on Stream Habitats and Biological Assemblages. American Fisheries Society, Bethesda, Maryland.

Ivits, E., Koch, B., 2002. Object-oriented remote sensing tools for biodiversity assessment: a European approach. In: Geoinformation for European-wide Integration. Proceedings of the 22nd EARSeL Symposium, June 4–6, 2002, Prague, Czech Republic. Millpress Science, Netherlands.

Johansen, K., Phinn, S., Lowry, J., Douglas, M., 2008. Quantifying indicators of riparian condition in Australian tropical savannas: integrating high spatial resolution imagery and field survey data. International Journal of Remote Sensing 29, 7003–7028.

Jones, K.B., Neale, A.C., Nash, M.S., Van Remortel, R.D., Wickham, J.D., Riitters, K.H., O'Neill, R.V., 2001. Predicting nutrient and sediment loadings to streams from landscape metrics: a multiple watershed study from the United States Mid-Atlantic Region. Landscape Ecology 16, 301–312.

Lattin, P., Wiginton, P., Moser, T.J., Peniston, B.E., Lindeman, D.R., Oetter, D.R., 2004. Influence of remote sensing imagery source on quantification of riparian land cover/land use. Journal of the American Water Resources Association 40, 215–227.

Li, H., Reynolds, J.F., 1995. On definition and quantification of heterogeneity. Oikos 73, 280–284.

Lucas, R., Rowlands, A., Brown, A., Keyworth, S., Bunting, P., 2007. Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. IJRS International Journal of Photogrammetry and Remote Sensing 62, 165–185.

Lyons, J., Trimble, S.W., Paine, L.K., 2000. Grass versus trees: managing riparian areas to benefit streams of central North America. Journal of the American Water Resources Association 36, 919–930.

Müller, E., 1997. Mapping riparian vegetation along rivers: old concepts and new methods. Aquatic Botany 58, 437.

Naiman, R.J., Decamps, H., 1997. The ecology of interfaces: Riparian Zones. Annual Review of Ecology and Systematics 28, 621–658.

Naiman, R.J., Décamps, H., McClain, M.E., 2005. Riparia: Ecology, Conservation, and Management of Streamside Communities. Elsevier Academic, ©2005, Boston, Amsterdam.

Paparoditis, N., Souchon, J.P., Martinoty, G., Pierrot-Deseilligny, M., 2006. High-end aerial digital cameras and their impact on the automation and quality of the production workflow. ISPRS Journal of Photogrammetry and Remote Sensing 60, 400–412.

Paul, M.J., Meyer, J.L., 2001. Streams in the urban landscape. Annual Review of Ecology and Systematics 32, 333–365.

Schulz, R., Liess, M., 1999. A field study of the effects of agriculturally derived insecticide input on stream macroinvertebrate dynamics. Aquatic Toxicology 46, 155–176.

Shuft, M.J., Moser, T.J., Wigington Jr., P.J., Stevens Jr., D.L., McAllister, L.S., Chapman, S., Ernst, T.L., 1999. Development of landscape metrics for characterizing riparian-stream networks. PE&RS – Photogrammetric Engineering and Remote Sensing 65, 1157–1167.

Souchon, Y., Andriamahefa, H., Cohen, P., Breil, P., Pella, H., Lamouroux, N., Malavoi, J.-R., Wasson, J.-G., 2000. Régionalisation de l'habitat aquatique dans le bassin de la Loire, Synthèse. Cemagref Lyon BEA/LHQ, Agence de l'eau Loire Bretagne, p. 291.

Stewart, J.S., Wang, L., Lyons, J., Horwatich, J.A., Bannerman, R., 2001. Influences of watershed, riparian-corridor, and reach-scale characteristics on aquatic biota in agricultural watersheds. Journal of the American Water Resources Association 37, 1475–1487.

Strahler, A.N., 1952. Hypsometric (area-altitude) analysis of erosional topography. Bulletin Geological Society of America 63, 1117–1142.

Strayer, D.L., Beighley, R.E., Thompson, L.C., Brooks, S., Nilsson, C., Pinay, G., Naiman, R.J., 2003. Effects of land cover on stream ecosystems: roles of empiricals models and scaling issues. Ecosystems 6, 407–423.

Tormos, T., Kosuth, P., Durrieu, S., Wasson, J.G., Pella, H., Villeneuve, B., Perez Correa, M., 2006. Use of multi-platform-sensing for characterisation of land use in river corridors. In: International Society for Photogrammetric and Remote Sensing-Commission 1 Symposium, Paris.

Townsend, C.R., Dolédec, S., Norris, R., Peacock, K., Arbuckle, C., 2003. The influence of scale and geography on relationships between stream community composition and landscape variables: description and prediction. Freshwater Biology 48, 768–785.

Turner, M.G., 1989. Landscape ecology: the effect of pattern on process. Annual Review of Ecology and Systematics 20, 171–197.

Weller, D.E., Jordan, T.E., Correll, D.L., 1998. Heuristic models for material discharge from landscapes with riparian buffers. Ecological Applications 8, 1156–1169.

CemOA : archive ouverte d'Irstea / Cemagref

Please cite this article in press as: Tormos, T., et al. Improving the quantification of land cover pressure on stream ecological status at the riparian scale using High Spatial Resolution Imagery. J. Phys. Chem. Earth (2010), doi:10.1016/j.pce.2010.07.012

III. Représentations cartographiques de l'état du corridor rivulaire

Cette partie illustre, à titre d'exemple, l'intérêt pour les gestionnaires et les décideurs de disposer de cartographies d'indicateurs spatialisés caractérisant l'état du corridor rivulaire. Les indicateurs qui semblaient a priori les plus «didactiques » ont été sélectionnés pour réaliser ces cartes mais il est bien évident que ces cartographies auront d'autant plus d'intérêt qu'elles seront construites à partir des indicateurs (1) les plus explicatifs de l'état écologique des cours d'eau (résultats des modèles pressions/impacts de la partie 4) et (2) jugés pertinents par les gestionnaires. Cette illustration a été menée sur deux aires d'étude : sur la partie aval du fleuve Hérault (80 km) constituée de 8 tronçons de cours d'eau de 10 km de long et sur les masses d'eau de l'ensemble de ce bassin (380 km). Tous les indicateurs ont été construits à partir de la donnée d'occupation du sol à THRS dont nous venons de démontrer la supériorité par rapport à ceux construits à partir de la donnée CLC.

III.1. Sur la partie aval de l'Hérault

III.1.1. Présentation du site d'étude

Cette étude a été menée sur le corridor rivulaire (600 m de part et d'autre du cours d'eau) de la partie aval du fleuve Hérault (80 km). Ce corridor est caractérisé par de fines bandes de végétation rivulaire (<10m) interompues par endroits par les territoires agricoles (notamment de la vigne) et par l'urbanisation à l'embouchure de l'Hérault (ville d'Agde). Pour la cartographie des indicateurs du corridor rivulaire, cette partie du fleuve a été découpée en 10 tronçons (de A à J) de 8 km de long (Figure 8.2).

III.1.2.Choix des indicateurs et représentations cartographiques

Cinq indicateurs ont été choisis pour représenter l'état du corridor rivulaire sur chacun des tronçons de la partie aval du fleuve Hérault :

- la largeur moyenne (ICR-ireg-Con-largm-forêt) et l'uniformité (ICR-ireg-Con-ufor-forêt) de la ripisylve (Figure 8.3).

- Trois indicateurs de contact: la superficie de ripisylve (ICR-ireg-Con-%-forêt), la pression agricole (ICR-ireg-Con-%-agri) et la pression urbaine (ICR-ireg-Con-%-urb) au contact du fleuve Héraut (Figure 8.4).

Tous ces indicateurs sont calculés sur l'emprise des objets en contact avec le cours d'eau.

CemOA : archive ouverte d'Irstea / Cemagref

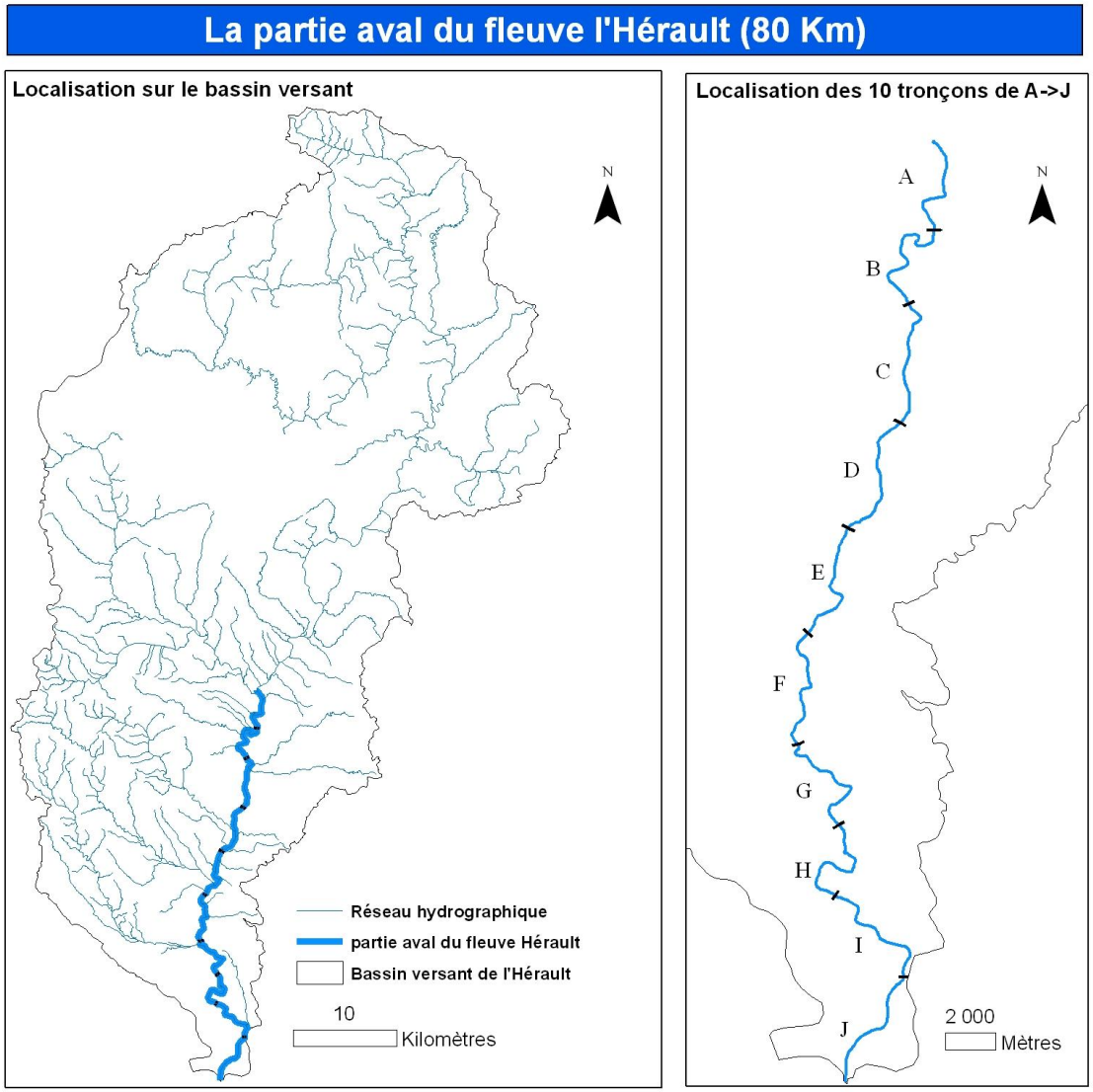


Figure 8.2 : Présentation de la partie aval du fleuve Hérault (80 km)

CemOA : archive ouverte d'Irstea / Cemagref

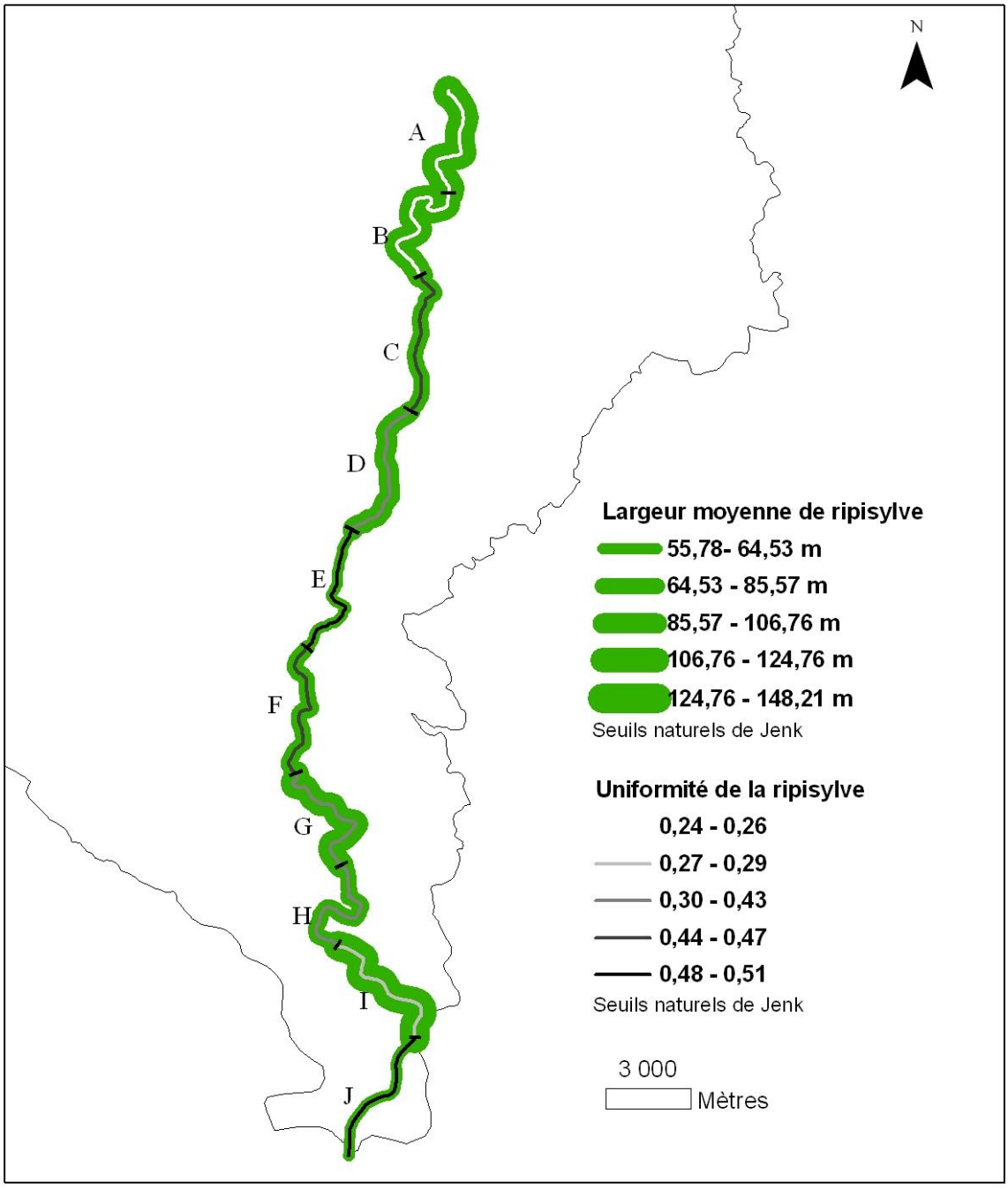


Figure 8.3 : Cartographie de la largeur moyenne et de l'uniformité de la ripisylve sur chaque tronçon (A->J) de la partie aval de l'Hérault.

CemOA : archive ouverte d'Irstea / Cemagref

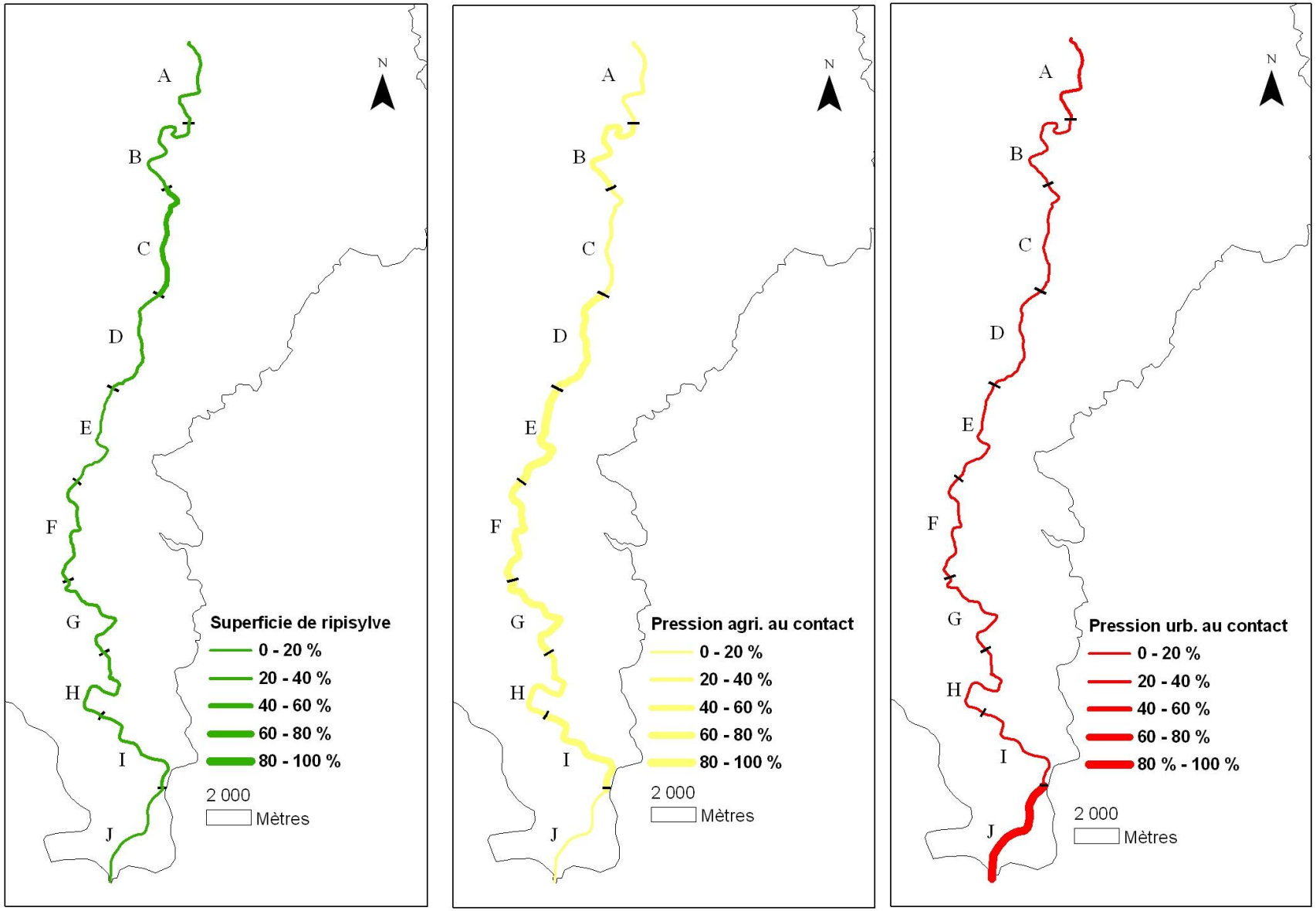


Figure 8.4 : Cartographie de trois indicateurs de contact : superficie de la ripisylve, pressions agricole (agri.) et urbaine (urb.) de contact sur chaque tronçon de la partie aval de l'Hérault.

CemOA : archive ouverte d'Irstea / Cemagref

La cartographie de la largeur moyenne et de l'uniformité de la ripisylve (Figure 8.3) indique la présence d'une végétation rivulaire arborée sur cette partie du fleuve de part et d'autre du cours d'eau avec des largeurs moyennes conséquentes (> 100 m pour les tronçons A, G et I). Cette ripisylve est globalement irrégulière (indicateur d'uniformité inférieur à 0.5, ce qui correspond à une variation de la largeur de plus de 50 m). Par ailleurs, on remarque que les ripisylves les moins larges sont plus uniformes (tronçons E et J) ce qui s'explique par le fait que sur ces tronçons la bande de ripisylve est étroite et quasi-continue.

La cartographie des indicateurs de contact (Figure 8.4) confirme la présence de ripisylve le long du fleuve mais indique que cette végétation rivulaire présente des interruptions (puisque le pourcentage de végétation arborée au contact du cours d'eau est inférieur à 50 % pour l'ensemble des tronçons). Le fleuve est donc soumis, dans la plupart des tronçons, à une pression agricole au niveau de ces interruptions, et également à la pression urbaine notamment pour le tronçon J fortement artificialisé.

III.2. Sur les masses d'eau du bassin versant de l'Hérault

III.2.1. Présentation des masses d'eau

Treize masses d'eau ont été définies par le SANDRE sur le bassin versant de l'Hérault (Figure 8.5). Ces masses d'eau appartenant à une seule et même hydro-écorégion ont été délimitées en fonction de deux critères : (1) le peuplement piscicole dominant basé sur le contexte piscicole (Salmonicole, Cyprinicole, Intermédiaire) et (2) la classe de taille (rang de Strahler) (Sandre).

III.2.2.Choix des indicateurs et représentations cartographiques

Quatre indicateurs ont été choisis pour représenter l'état du corridor rivulaire sur chacune des masses d'eau du bassin versant de l'Hérault :

- La longueur (ICR-regfixe-Lin-%-forêt) et la densité de ruptures (ICR-ireg-Con-drup-forêt) de la ripisylve (Figure 8.6).

- Deux indicateurs du lit majeur : la pression agricole (ICR-regvar-Maj-%-agri) et la pression urbaine (ICR-regvar-Maj-%-urb) sur le lit majeur (Figure 8.7 et Figure 8.8).

Le premier indicateur est calculé sur une largeur fixe de 5 m, le second est calculé sur l'emprise des objets en contact avec la masse d'eau, les deux autres sont calculés sur une largeur variable correspondant au lit majeur et estimée à partir du rang de Strahler selon la règle définie à la partie Chapitre 7 :II.2.3.2 .

La cartographie de la longueur et de la densité de ruptures de la ripisylve (Figure 8.6) indique la présence généralisée (> 50 %) d'une végétation rivulaire arborée au niveau des berges. Cette ripisylve est plus fortement présente en tête du bassin versant (jusqu'à 95 %). Elle est caractérisée en revanche par de nombreuses ruptures, excepté sur les masses d'eau 4 et 10.

La cartographie des indicateurs du lit majeur (Figure 8.7 et Figure 8.8) indique la présence de pressions anthropiques (agricoles et urbaines) susceptibles d'altérer l'état écologique des masses d'eau. Une pression agricole qui peut atteindre plus de 50 % notamment sur les masses d'eau 8 et 13, et une pression urbaine supérieure à 50 % pour les masses d'eau 11 et 12.

CemOA : archive ouverte d'Irstea / Cemagref

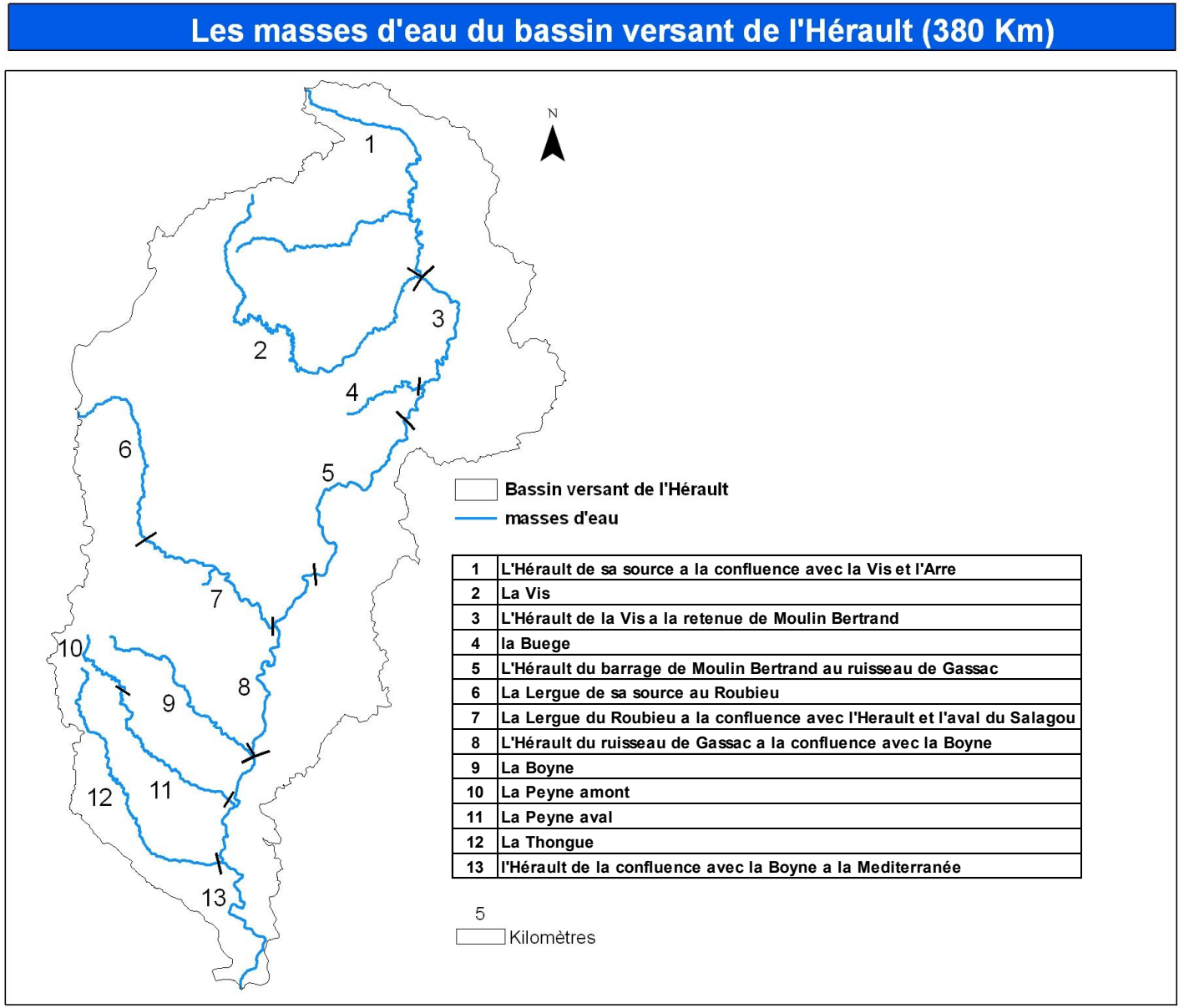


Figure 8.5 : présentation des treize masses d'eau du bassin versant de l'Hérault

CemOA : archive ouverte d'Irstea / Cemagref

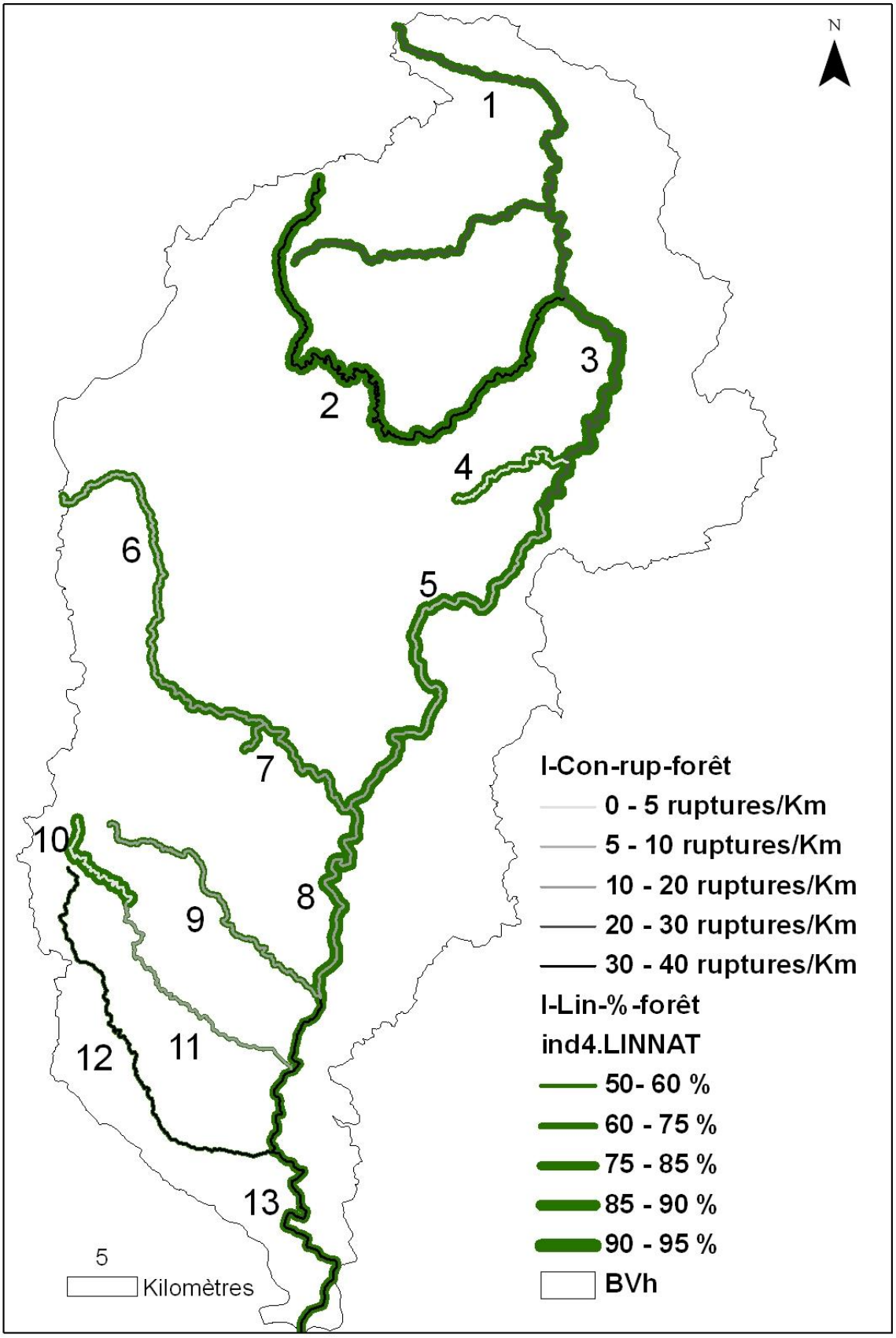


Figure 8.6 : Cartographie de la longueur (ICR-regfixe-Lin-%-forêt) et de la densité de ruptures (ICR-ireg-Con-drup-Forêt) de la ripisylve de chaque masse d'eau du bassin versant de l'Hérault

CemOA : archive ouverte d'Irstea / Cemagref

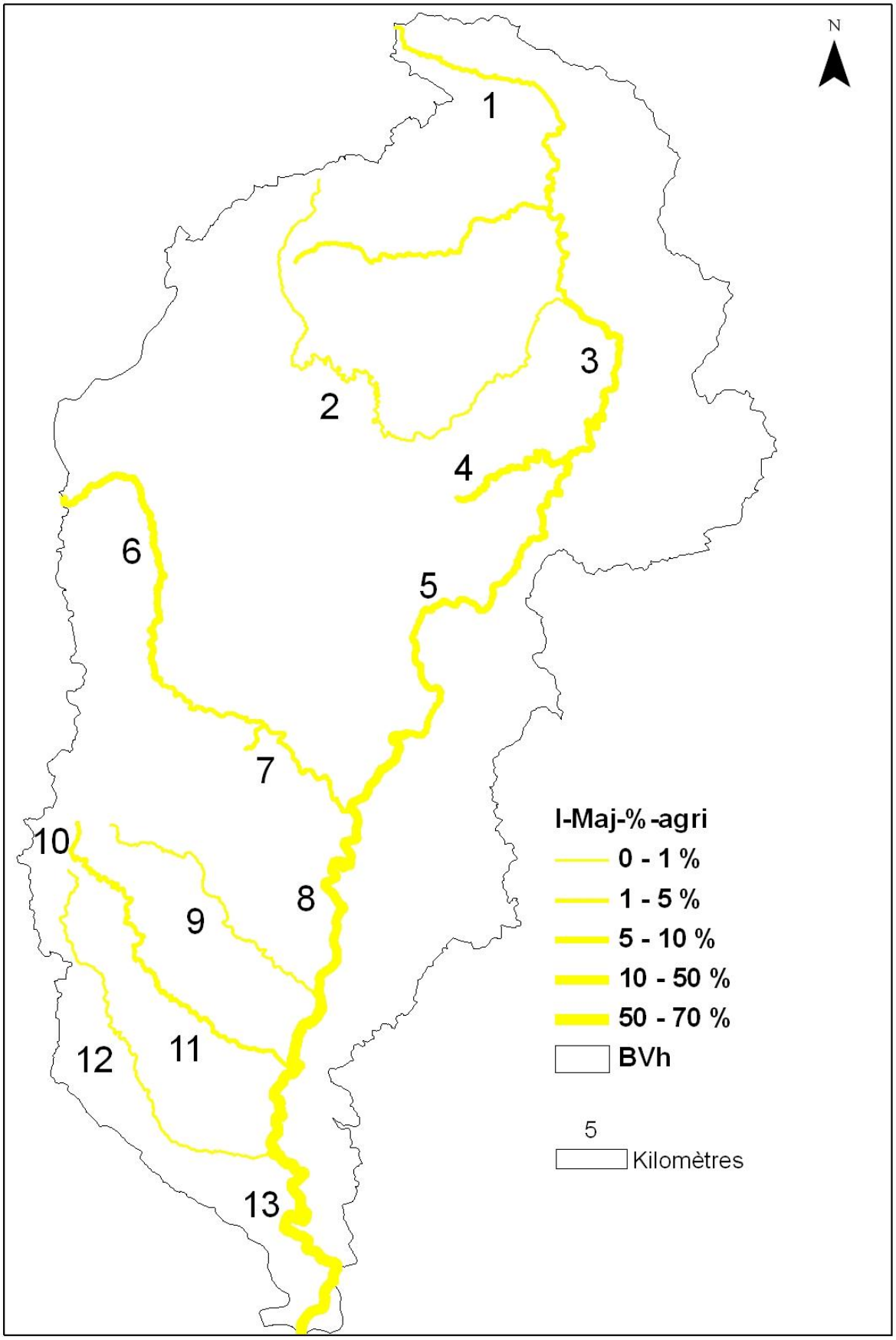


Figure 8.7 : Cartographie de la pression agricole sur le lit majeur (ICR-regvar-Maj -%-agri) de chaque masse d'eau du bassin versant de l'Hérault

CemOA : archive ouverte d'Irstea / Cemagref

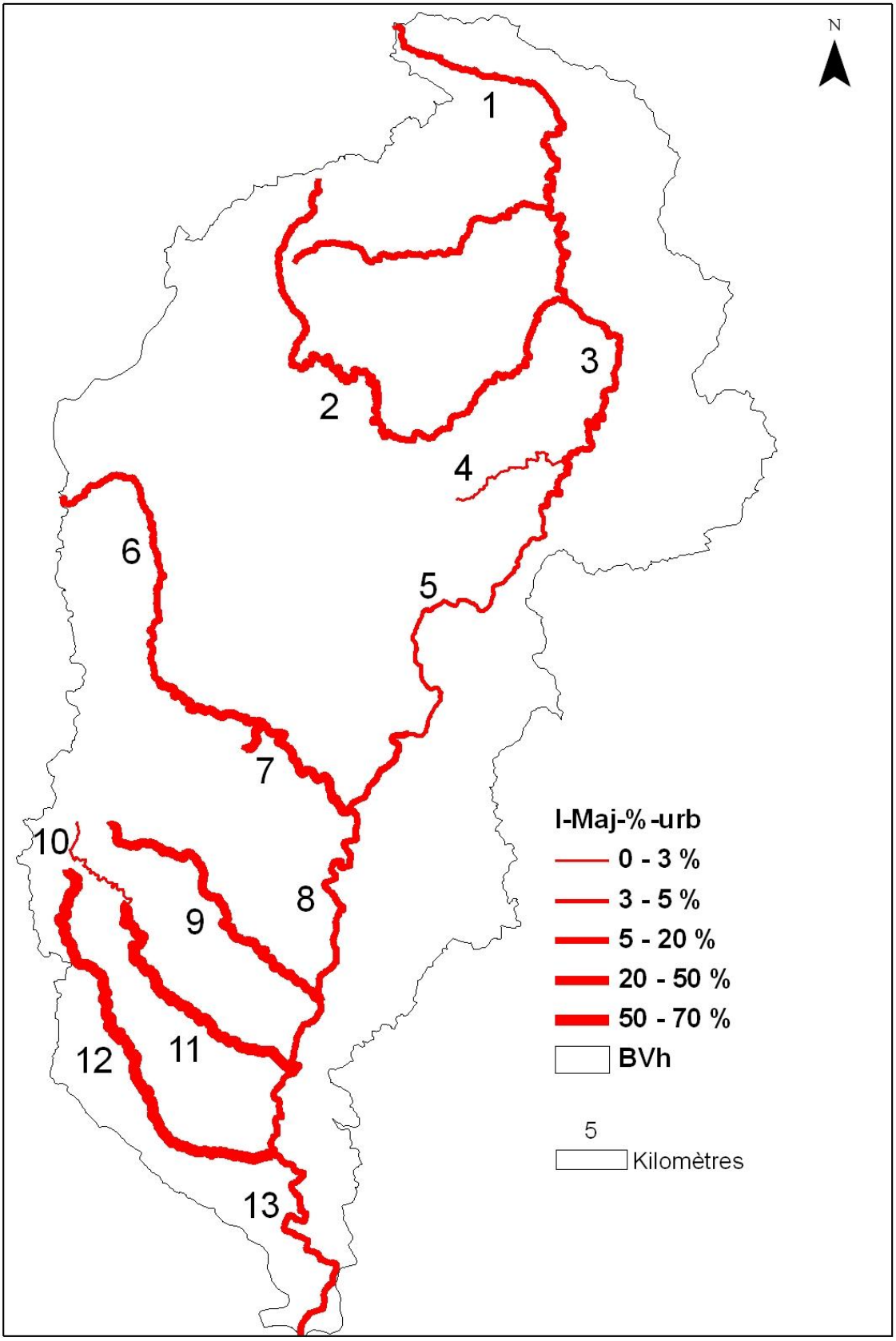


Figure 8.8 : Cartographie de la pression urbaine sur le lit majeur (ICR-regvar-Maj -%-urb) de chaque masse d'eau du bassin versant de l'Hérault

CemOA : archive ouverte d'Irstea / Cemagref

IV. L'incertitude des indicateurs spatialisés

Il est indispensable, en quantifiant des indicateurs spatialisés à des fins de représentation, de modélisation ou de gestion (indicateurs pour les gestionnaires), de leur associer la valeur de leur incertitude. Elle permet d'évaluer la pertinence de la représentation, de la modélisation, ou –plus critique- de la prise de décision qui peut en dépendre. Dans le cadre de cette thèse, l'estimation de l'incertitude des indicateurs permet par la suite de mesurer plus précisément l'incertitude des résultats des modèles pressions / impacts régionalisés. Deux sources d'incertitudes peuvent affecter le résultat d'un indicateur spatialisé.

La première source est relative aux erreurs de décalage planimétrique de la donnée hydrographique utilisée pour construire l'emprise spatiale dans laquelle l'attribut du paysage est quantifié. Dans la partie III.2.1.1 et ***ANNEXE IX***, un premier travail a été réalisé sur un sous-bassin versant de l'Hérault qui compare le résultat d'un indicateur spatialisé avec des emprises délimitées suivant la donnée hydrographique à disposition sur le territoire français, la BDCarthage®, puis en fonction d'un tracé corrigé par photo-interprétation.

Ce travail montre que la valeur des indicateurs spatialisés étudiés (indicateurs linéaires construits à partir d'un buffer étroit à priori plus sensible aux erreurs de décalage de la donnée hydrographique) ne varie que faiblement lorsque des erreurs de décalages existent et que cette variation ne s'accentue pas en fonction de l'importance de ces erreurs sur un tronçon de cours d'eau analysé.

A partir de ces résultats on peut conclure que les indicateurs spatialisés sont faiblement sensible aux erreurs de décalage de la BDCarthage®. Bien évidemment, ce travail, réalisé sur une zone d'étude localisée, devra être reproduit à plus grande échelle afin de confirmer cette observation.

La deuxième source d'incertitude est relative aux erreurs de la carte d'occupation du sol. Les erreurs de cette carte sont de deux types : (i) de type géométrique, i.e., la précision des limites des objets extraits par rapport à la réalité sur le terrain ; et (2) de type sémantique, i.e., la nature des objets par rapport à cette même réalité.

En connaissant, l'incertitude attachée aux erreurs géométriques et à la nature des objets il est possible de générer une multitude de cartes d'occupation des sols réalistes, en bruitant la carte d'OS initiale avec cette incertitude. On peut alors recalculer le même indicateur sur cette multitude de cartes d'OS "bruitées" et faire une analyse statistique de l'échantillon de valeurs de l'indicateur ainsi obtenu (approche de type Monte Carlo) pour obtenir son incertitude. Cette méthode est d'autant plus efficace lorsqu'on on a une bonne connaissance des erreurs géométriques et sémantiques de la carte d'occupation du sol. Toutefois elle nécessite une importante capacité de calcul.

Dans la procédure de classification mise en œuvre dans la partie 2 de la thèse, on a choisi de ne pas quantifier la qualité géométrique de la carte d'occupation car l'objectif de l'exercice n'est pas d'obtenir une donnée de référence cartographique. Cependant, La qualité géométrique est sensiblement liée à la qualité sémantique : si la segmentation est de mauvaise qualité, i.e., une mauvaise délimitation des objets que l'on veut extraire dans l'image, on aura de forte chance que la qualité sémantique de la carte d'occupation soit également de mauvaise qualité car l'information radiométrique récupérée à l'échelle des objets mal segmentés ne correspondra alors pas correctement à la signature de l'objet qu'on cherche à extraire.

Deux techniques ont été mises en place dans la thèse pour quantifier la qualité sémantique de la carte d'occupation du sol dans le corridor rivulaire (cf. Chapitre 5 :III.7) : la matrice de confusion et l'utilisation des propriétés de la logique floue. Alors que la matrice de confusion permet de confronter le résultat de classification à une vérité terrain en fournissant une information sur la manière dont les classes se confondent sur l'ensemble de la zone d'étude analysée, la logique floue apporte une information au niveau de chaque objet sur la manière dont les classes se mélangent en son sein, ce mélange dépendant des ensembles flous des règles de classification définis par l'opérateur.

Pour quantifier l'incertitude attachée à la nature des objets résultant de la carte d'occupation du sol du corridor rivulaire deux approches sont donc possibles :

- soit on se base sur l'information apportée par la matrice de confusion : on peut générer facilement des cartes d'occupation du sol bruitées (« fausses » cartes) en faisant pour chaque objet un tirage aléatoire selon son risque de confusion avec d'autres classes. Cette solution est cependant limitée pour deux raisons : on a de fortes chances de générer beaucoup de fausses « cartes » irréalistes et de risquer ainsi de surestimer fortement l'incertitude d'un indicateur spatialisé et surtout on ne prend pas en compte le fait que ces erreurs peuvent être spatialisées, c'est-à-dire hétérogènes sur le territoire, et donc impacter différemment un indicateur suivant l'emprise dans laquelle il analyse la structure de l'occupation du sol.

CemOA : archive ouverte d'Irstea / Cemagref

- soit on se base sur l'information apportée par la logique floue : on peut également facilement générer de « fausses » cartes d'occupation des sols en faisant pour chaque objet un tirage aléatoire selon les possibilités de confusion entre les différentes classes spécifiques à cet objet, telles que celles fournies par les résultats de la classification floue. Cette deuxième approche permet donc, comparée à la première, de prendre en compte la spatialisation des erreurs de classification. Cependant, comme le montre l'application de la procédure de classification sur le bassin versant de l'Hérault (cf. Chapitre 6 :III.), l'information d'incertitude apportée par la logique floue est relative et non absolue : elle représente le risque d'ambiguïté de classification tel qu'évalué par l'opérateur sur la base des ensembles flous qu'il a définis ; elle se fait indépendamment de la vérité terrain et n'apporte pas d'information sur l'incertitude absolue.

Ainsi, à cause des difficultés liées à la spatialisation des erreurs de classification (caractère glabal des matrices de confusion, subjectivité de l'analyse de fiabilité de la logique floue) et des moyens et temps de calcul qui seraient nécessaires pour mener à bien cette étude, la sensibilité des indicateurs d'occupation du sol aux erreurs de classification du sol n'a pas été menée dans cette thèse.

V. Conclusion

La construction d'indicateurs représentatifs de la composition et de la configuration de l'occupation des sols dans les corridors rivulaires, et notamment de la végétation arborée, nécessite d'utiliser une donnée d'occupation des sols à Très Haute Résolution Spatiale. La donnée CLC s'avère insuffisamment résolue pour rendre compte du type d'occupation des sols en contact avec les masses d'eau.

Les représentations cartographiques des différents indicateurs d'une part sur des tronçons de la partie aval de l'Hérault et d'autre part au niveau des 13 masses d'eau de l'Hérault illustrent la capacité de ces indicateurs à fournir une information cartographique explicite pour guider les gestionnaires dans leurs stratégies de restauration sur le corridor rivulaire.

La quantification de l'incertitude associée aux indicateurs spatialisés d'occupation des sols dans les corridors rivulaires, notamment aux indicateurs de ripisylves, peut être réalisée par une approche de type Monte Carlo en calculant le même indicateur pour un grand nombre de cartes d'occupation des sols générées à partir de la carte « mère » bruitée selon la matrice de confusion de la classification ou bien selon des paramètres d'incertitude liés à chaque objet (issus de la classification floue). La spatialisation des erreurs de classification de l'occupation des sols reste toutefois un point clef non résolu pour mener à bien ce travail.

Au-delà de la Directive Cadre sur l'Eau, cette recherche prépare un autre enjeu relatif à la mise en place de la politique « trame verte et bleue » concernant les corridors de biodiversité à l'échelle nationale. La construction d'indicateurs d'évaluation de l'état des corridors rivulaires pertinents à large échelle est un enjeu majeur pour la mise œuvre de cette politique. De tels indicateurs sont nécessaires à la fois pour dresser un « état des lieux » national et suivre les effets des actions entreprises.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

- L'indicateur spatialisé du corridor rivulaire est défini de façon générique par ***un attribut de structure*** (composition et/ou configuration) calculé pour ***une catégorie d'occupation du sol*** (selon une certaine typologie) au sein d'***une emprise spatiale rivulaire de dimension donnée*** (représentée par un buffer délimité autour d'un tronçon de cours d'eau selon une certaine distance en amont et/ou aval de la station et selon une certaine distance latérale à ce tronçon, largeur de part et d'autre du cours d'eau).

- ***Deux attributs*** ont été sélectionnés: (i) ***le pourcentage surfacique*** (attribut le plus explicatif de l'état écologique) qui traduit la présence et l'intensité des pressions susceptibles d'être générées par une catégorie d'occupation du sol (OS); et (ii) ***la fragmentation*** qui nuance l'intensité des pressions générées pour certaines de ces catégories dont l'influence est susceptible de varier en fonction de leur emplacement (hétérogénéité spatiale) le long du cours d'eau.

- Ces métriques peuvent être extraites dans ***plusieurs emprises spatiales*** traduisant ainsi le rayon d'influence de ces pressions, plus ou moins proche au cours d'eau et à la station. ***Un grand nombre d'indicateurs*** peut donc être construit en fonction de la taille de ces emprises. L'utilisation de plusieurs emprises candidates (***mode de spatialisation « multiple »***) améliorera l'identification, dans les modèles pressions/impacts, de l'emprise au sein de laquelle une catégorie d'OS a une influence significative sur la variabilité de la réponse biologique du cours d'eau.

- Parmi ces possibilités d'indicateurs, ***trois familles***, qui diffèrent en fonction de la méthode et de la distance latérale au cours d'eau prise en compte pour délimiter le buffer d'analyse, traduisent ***des mécanismes d'impact spécifiques et/ou des localisations particulières dans le corridor rivulaire*** : (1) ***les indicateurs du linéaire*** qui, construits à partir d'un buffer aux limites régulières (i.e. qui suivent la linéarité du cours d'eau) d'une distance fixe de 5 m de part et d'autre du cours d'eau, s'intéressent ***aux pressions*** d'une catégorie d'occupation du sol donnée qui s'exercent ***au plus proche du cours d'eau*** et apportent ***une information*** sur ***l'état des berges rivulaires*** ; (2) les ***indicateurs du lit majeur***, qui définis par un buffer régulier à distance variable (en fonction du rang du cours d'eau), s'intéressent ***aux pressions*** exercées par une catégorie d'occupation du sol donnée présente ***au sein du lit majeur***, espace de liberté théorique du cours d'eau. Ils caractérisent ainsi l'état global de l'espace rivulaire ; et (3) les indicateurs de contact qui, construits à partir d'un buffer irrégulier (i.e. qui ne suivent pas la linéarité du cours d'eau) constitué par la fusion des objets au contact du cours d'eau à partir d'un buffer irrégulier, s'intéressent ***aux pressions exercées par les objets*** d'une catégorie d'occupation du sol donnée ***directement au contact du cours d'eau*** ; ils permettent également de mettre en évidence ***la présence de ruptures dans la bande de végétation rivulaire*** et de ***préciser ainsi l'intensité de ces pressions au niveau de ces ruptures***.

- Six types d'indicateurs spécifiques ***au rôle des ripisylves*** ont été également définis : (1) ***la longueur*** et (2) ***la fragmentation*** de la ripisylve qui dérivent de la famille des indicateurs du linéaire ; (3) ***la superficie*** de la ripisylve qui dérive de la famille des indicateurs de « contact » ; (4) la ***densité de ruptures*** (nombre de patchs distincts par km de cours d'eau), (5) ***la largeur moyenne*** (en m) et (6) ***l'uniformité*** (en fonction de la variation de la largeur des patchs) de la ripisylve définies à partir du buffer irrégulier de contact. Ces indicateurs permettent de quantifier de façon générale ***la présence et l'intensité*** de cette végétation le long du cours d'eau (longueur et superficie), ***sa connectivité*** (fragmentation et densité de ruptures) caractéristique essentielle à son fonctionnement, mais aussi plus spécifiquement ***sa capacité à jouer son rôle de barrière*** vis-vis des polluants diffus (fragmentation, densité de ruptures, largeur moyenne et uniformité).

- ***La délimitation automatisée des différents buffers*** a nécessité ***le développement d'algorithmes SIG (Systèmes d'Information Géographique) spécifiques*** notamment pour définir la distance amont et ou/aval à la station. La définition de cette distance ***se base sur*** les outils de la ***segmentation dynamique*** utilisant **les systèmes de repérage curviligne** et les ***informations topologiques du réseau hydrographique***.

- ***La mise en œuvre*** de certains de ces indicateurs ***sur le bassin de l'Hérault*** démontre : (i) ***l'intérêt déterminant de la donnée à THRS*** pour ***quantifier correctement ces indicateurs*** au niveau du corridor rivulaire ; et (ii) ***leur capacité*** à fournir ***une information cartographique explicite pour guider les gestionnaires*** dans leurs stratégies de restauration sur le corridor rivulaire.

- ***L'étude de la sensibilité*** de ces indicateurs ***aux erreurs des cartes d'OS*** restée ***inachevée*** : les méthodologies pour étudier cette influence sont clairement identifiées (tirages de Monte-Carlo) mais les ***difficultés rencontrées pour spatialiser les erreurs de la donnée d'OS***, au-delà de la seule matrice de confusion qui n'est pas spatialisée, demeurent un point de blocage à lever.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Quatrième partie :

Modélisation pressions/impacts régionalisée - Prise en compte des corridors rivulaires

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Introduction de la quatrième partie

Cette partie est structurée en deux chapitres.

Un premier chapitre (Chapitre 9) qui présente les données disponibles sur la zone d'étude, les méthodes et la démarche de modélisation mise en place pour d'une part tester les avancées méthodologiques développées dans les parties précédentes dans l'étude des relations pressions/impacts à large échelle au niveau du corridor rivulaire et d'autre part établir un diagnostic de l'état écologique sur la zone d'étude.

Un deuxième chapitre (Chapitre 10) qui applique cette démarche afin de répondre aux objectifs, dresse un bilan critique des méthodes mises en œuvre et discute de l'implication de ces modèles dans l'élaboration des stratégies de restauration pour atteindre les objectifs de la directive cadre.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Chapitre 9
Modèles pressions/impacts : matériels, méthode et démarche

SOMMAIRE

Résumé *296*

I. ENJEUX, OBJECTIFS ET ZONE D'ETUDE (RAPPELS) 298

I.1. ENJEUX DE LA MODELISATION PRESSIONS/IMPACTS *298*

I.2. PRINCIPES ET LIMITES DES MODELES EMPIRIQUES *298*

I.3. OBJECTIFS DE L'ETUDE *299*

I.4. PRESENTATION DE LA ZONE D'ETUDE *300*

II. DESCRIPTION DES DONNEES BIOLOGIQUES ET DE PRESSIONS 300

II.1. LES INDICATEURS D'ETAT ECOLOGIQUE DES STATIONS *300*

II.1.1. Méthode d'évaluation : l'EQR-IBGN$_{moyen}$ 300

II.1.2. Limites et Incertitudes de l'IBGN 302

II.2. LES INDICATEURS DE PRESSIONS D'OCCUPATION DU SOL *303*

II.2.1. Méthode d'évaluation des indicateurs de pressions 303

II.2.1.1 Définitions (rappels) 303

II.2.1.2 Choix de l'attribut du paysage 303

II.2.1.3 Données d'occupation du sol 304

II.2.1.3.1 CORINE Land Cover (CLC) 304

II.2.1.3.2 La donnée à Très Haute Résolution Spatiale (THRS) 304

II.2.1.3.3 Représentations cartographiques 305

II.2.1.4 Choix de la typologie 305

…SOMMAIRE (suite)

CemOA : archive ouverte d'Irstea / Cemagref

…SOMMAIRE (suite)

II.2.1.5 Choix des emprises spatiales 308
II.2.1.5.1 Emprise spatiale au niveau du bassin versant 308
II.2.1.5.2 Emprise spatiales au niveau du corridor rivulaire 308
II.2.1.5.2.1 Mode de spatialisation des sources de pressions (fixe ou multiple) 308
II.2.1.5.2.2 Prise en compte de l'aval de la station 309
II.2.1.5.2.3 Emprises à l'échelle du CRamont (mode multiple) 309
II.2.1.5.2.4 Emprises à l'échelle du CRlocal 310
II.2.1.5.2.4.1 Emprise du mode fixe 310
II.2.1.5.2.4.2 Emprises du mode multiple 311
II.2.1.6 Récapitulatif et dénomination des indicateurs 312
II.2.2. Limites et incertitudes des indicateurs d'occupation du sol 313

III. METHODE DE CONSTRUCTION D'UN MODELE PRESSIONS/IMPACTS 314

III.1. UNE PROCEDURE EN DEUX ETAPES 314
III.2. METHODE DE SELECTION DES EMPRISES (1ERE ETAPE) 314
III.3. METHODE D'ANALYSE DES RELATIONS PRESSIONS/IMPACTS (2EME ETAPE) 316
III.3.1. Choix la méthode statistique 316
III.3.2. Présentation de l'algorithme de régression PLS 319
III.3.3. Construction d'un modèle PLS 321

IV. DEMARCHE DE MODELISATION 322

IV.1. UNE DEMARCHE EN DEUX PHASES 322
IV.2. PHASE 1 : CONSTRUCTION DES DIFFERENTS MODELES PRESSIONS/IMPACTS 324
IV.3. PHASE 2 : ANALYSES COMPARATIVES DES MODELES 324
IV.3.1. Critères de comparaison 324
IV.3.2. Évaluation des apports respectifs des différentes voies d'améliorations (analyse 1) 325
IV.3.1. Sélection du modèle optimal (analyse 2) 325

CemOA : archive ouverte d'Irstea / Cemagref

Résumé

Ce chapitre présente les données disponibles sur la zone d'étude, les méthodes de modélisation et la démarche d'ensemble élaborée pour répondre aux objectifs de modélisation de cette étude. Ces objectifs sont doubles. Le premier, lié à des enjeux méthodologiques, a pour but de quantifier l'apport respectif, dans l'étude des relations pressions/impacts à large échelle, des voies d'amélioration développées dans cette thèse au niveau du corridor rivulaire : la classification de l'occupation du sol à partir de l'imagerie à Très Haute Résolution Spatiale (THRS) et le mode de spatialisation multiple (définition de plusieurs emprises candidates pour les indicateurs d'occupation du sol). Le deuxième, lié à des enjeux opérationnels, a pour but de donner des éléments de réponse aux gestionnaires sur l'identification et la localisation des facteurs dégradant ou protégeant l'état écologique.

Une démarche spécifique a donc été mise en place pour répondre à ces objectifs. Elle construit dans une première phase plusieurs modèles pressions/impacts (21 au total) qui diffèrent soit en fonction de l'échelle ou de la combinaison d'échelles d'analyse (bassin versant et/ou corridor rivulaire amont et/ou corridor rivulaire local), soit en fonction de la source de données d'occupation du sol au niveau des deux échelles du corridor rivulaire (CORINE Land Cover ou données à THRS) ou soit en fonction du mode de spatialisation au niveau de l'échelle du corridor local (fixe ou multiple). Puis, dans une deuxième phase, elle compare les résultats de ces modèles afin de quantifier les apports respectifs des différentes voies d'amélioration et de sélectionner le modèle le plus pertinent pour établir un diagnostic de l'état écologique sur la zone d'étude (partie Nord-Ouest de l'hydro-écorégion Tables Calcaires).

Tous ces modèles cherchent à expliquer, par les indicateurs d'occupation des sols en se basant sur une population de 155 stations écologiques, la variabilité de l'écart à la référence (EQR) estimé à partir de la moyenne de l'Indice Biologique Global Normalisé (l'IBGN basé sur la polluosensibilité et la diversité des macroinvertébrés benthiques) sur la période de 1992 à 2004.

Ils sont construits à partir d'une procédure en deux étapes. La première étape sélectionne, dans le cas où on se retrouve devant divers choix d'emprises spatiales pour une source de pression donnée (i.e., catégorie d'occupation du sol), l'emprise spatiale sur laquelle cette source de pression aura l'influence la plus significative sur l'état écologique. Cette sélection se base sur la comparaison des corrélations bi-variées entre la réponse biologique et la source de pression mesurée sur les différentes emprises candidates. La deuxième étape, à partir des sources de pressions s'exprimant sur leur emprise spatiale la plus significative (sélectionnés dans la 1ère étape), identifie la combinaison des différentes sources de pressions aux différentes échelles expliquant au mieux la variabilité de la réponse biologique. Cette combinaison est déterminée de façon automatique à l'aide d'une méthode de construction itérative et progressive (en terme de significativité) basée sur l'algorithme de régression Partial Least Square qui permet d'obtenir des coefficients de régression interprétables même en cas de corrélation forte entre les prédicteurs (ce qui est le cas avec les indicateurs d'occupation du sol).

CemOA : archive ouverte d'Irstea / Cemagref

I. Enjeux, objectifs et zone d'étude (rappels)

I.1. Enjeux de la modélisation pressions/impacts

Afin d'orienter efficacement les gestionnaires dans l'élaboration des mesures de préservation et/ou de restauration nécessaires pour atteindre l'objectif de « bon état » fixé par la Directive Cadre Eau (DCE), il est indispensable d'établir un diagnostic de l'état écologique des cours d'eau en identifiant et hiérarchisant, à large échelle (nationale, régionale, grands bassin versants), les causes d'altération ou de protection de cet état.

L'état écologique n'est pas seulement influencé par les sources de pollution ponctuelles, mais aussi par l'ensemble des pressions anthropiques plus complexes qui modifient les processus fonctionnels de l'écosystème rivière (processus physiques, biogéochimiques et écologiques), en générant des pollutions diffuses, des dysfonctionnements hydrologiques, des variations du flux de sédiments, des modifications hydromorphologiques et des ruptures de connectivité (Borchardt and Richter, 2003). Ces différentes pressions agissent, dans la plupart des cas, de façon simultanée (pressions combinées), à différentes échelles spatiales de fonctionnement de l'écosystème rivière (de l'échelle locale jusqu'à l'échelle du bassin versant, cf. Chapitre 3 :III.1.2) et sont réparties irrégulièrement sur le territoire (Allan, 2004a).

Des études au cas par cas n'étant donc pas envisageables, la recherche des causes d'altération de l'état écologique implique clairement une modélisation régionale des relations entre ces pressions multiples aux différentes échelles et l'état écologique. De tels modèles sont appelés modèles pressions/impacts ou modèle pressions/états selon la logique DPSIR (cf. Chapitre 3 :I.2). Ils doivent permettre d'identifier et localiser l'impact relatif des différentes sources de pressions (i.e., catégories d'occupation du sol) dégradantes ou protectrices aux différentes échelles afin d'indiquer aux gestionnaires sur quelles causes agir en priorité, à quelles échelles et sur quelles emprises spatiales.

I.2. Principes et limites des modèles empiriques

les approches mécanistes sont limitées comme nous l'avons vu dans le chapitre 3 pour répondre à ces enjeux : d'une part, ils ne permettent pas de prédire la présence, la structure et le fonctionnement des biocénoses, paramètres clés de l'évaluation de l'état écologique au sens de la DCE ; et d'autre part, ils restent encore difficilement applicables sur de vastes territoires (difficultés à décrire et modéliser l'ensemble des mécanismes, à caractériser et mesurer l'ensemble des variables, et à gérer la surparamétrisation des modèles).

Par contre, les approches empiriques permettent d'établir un diagnostic à large échelle de l'état écologique au sens de la DCE par l'étude des relations statistiques entre les diverses pressions calculées aux différentes échelles de fonctionnement de l'écosystème rivière (de l'échelle locale jusqu'à l'échelle du bassin versant) et les indicateurs de cet état écologique.

Cependant, dans la plupart des cas, on ne dispose pas, à cette échelle de travail, d'une information homogène et spatialisée (i.e. qu'on peut relier topologiquement avec les stations écologiques) sur les diverses pressions. La plupart des modèles pressions/impact à large échelle développés aujourd'hui, utilisent donc l'occupation du sol comme une des forces motrices (au sens de la logique DPSIR cf. Chapitre 3 :I.2) à l'origine de la plupart des pressions s'exerçant sur le milieu aquatique. Cette information spatialisée et homogène sur de grandes surfaces est issue généralement du traitement d'images de télédétection. Elle est ensuite analysée au sein d'un Système d'Information Géographique (SIG) afin de quantifier et de localiser (aux différentes échelles), à l'aide d'indicateurs d'occupation du sol, les sources de pressions s'exerçant à chaque station.

Ces modèles pressions/impacts (ou plus exactement forces-motrice/états par rapport à la logique DPSIR) dérivent des approches de l'écologie du paysage (Wiens, 2002 ; Allan, 2004a). Ils étudient la réponse biologique en fonction de la structure (composition et/ou la configuration) de l'occupation du sol à travers différentes échelles de fonctionnement du paysage étudié.

Leur application apporte un premier cadre descriptif et directif pour les gestionnaires sur l'identification et la localisation (échelles et emprises spatiales) des principales pressions générées par l'occupation du sol,

CemOA : archive ouverte d'Irstea / Cemagref

susceptibles d'influencer significativement l'état écologique. Par la suite, dans ce cadre, des analyses plus précises, qui prendront en compte l'ensemble des éléments du système selon la logique DPSIR, pourront être développées de manière à identifier et quantifier des leviers d'action concrets et réalisables par la société humaine.

Cependant, alors que certains résultats semblent faire l'objet d'un consensus, comme par exemple l'impact négatif des zones agricoles et urbaines sur le bassin, d'autres divergent, particulièrement au niveau du poids relatif des différentes échelles d'impact (cf. Chapitre 3 :III.2.1.5.3), élément important et attendu par les gestionnaires pour prendre des décisions notamment sur la restauration des espaces rivulaires. Deux difficultés relatives à l'analyse des relations pressions/impacts au niveau du corridor rivulaire compliquent le problème.

La première difficulté réside dans le manque d'information adaptée à cette échelle : la résolution de la donnée spatiale est généralement trop grossière pour quantifier précisément la composition et la configuration de l'occupation des sols et caractériser l'état des ripisylves, ce qui limite d'emblée l'analyse et l'interprétation de ces relations (Hollenhorst *et al.*, 2006 ; Wasson *et al.*, 2005 ; Wang, 2006 ; Roy *et al.*, 2007).

La deuxième difficulté réside dans la spatialisation de ces sources de pressions (Frimpong *et al.*, 2005b ; Van Sickle and Johnson, 2008) : généralement, quelque soit la source, elles sont étudiées sur une même emprise spatiale (la délimitation de cette emprise se base sur des hypothèses définies préalablement par l'analyste) ceci au niveau des deux échelles du linéaire : corridor local et corridor amont. Cette difficulté limite (i) la description de ces relations car l'effet d'une catégorie d'occupation du sol donnée est plus ou moins perceptible sur la réponse biologique en fonction de l'emprise spatiale sur laquelle elle est étudiée; puis (ii) l'information pour les gestionnaires concernant la localisation de ces sources de pressions au niveau du corridor.

Bien évidemment, la deuxième difficulté est intrinsèquement liée à la première. La donnée spatiale étant trop grossière pour caractériser l'occupation du sol au niveau du corridor, elle limite la spatialisation des différentes sources de pressions.

I.3. Objectifs de l'étude

Pour répondre à ces difficultés (précision de la donnée d'occupation du sol puis quantification et spatialisation des sources de pressions au niveau du corridor rivulaire), on a vu précédemment que (i) le traitement combiné de l'imagerie à Très Haute Résolution Spatiale (THRS) et de données auxiliaires, de façon semi-automatique, permet d'apporter une information spatiale précise et fiable sur la composition et la configuration de l'occupation des sols au niveau du corridor rivulaire (cf. deuxième partie) ; et (ii) que l'utilisation des outils des SIG permet aujourd'hui de définir pour chaque catégorie d'occupation du sol plusieurs emprises candidates au niveau du corridor rivulaire (mode de spatialisation « multiple ») pour étudier leur état écologique (cf. troisième partie).

Les objectifs de cette étude des modèles pressions/impacts sont doubles :

(1) Le premier, lié à des enjeux méthodologiques, a pour but de quantifier l'apport respectif de ces voies d'amélioration (cartographie à THRS de l'occupation des sols, indicateurs d'une catégorie d'occupation du sol sur différentes emprises) dans l'étude des relations pressions/impacts et la prise en compte du rôle du corridor rivulaire ;

(2) Le deuxième, lié à des enjeux opérationnels, a pour but de donner des éléments de réponse aux gestionnaires sur l'identification et la localisation des facteurs dégradant ou protégeant l'état écologique.

Une démarche de modélisation spécifique a été mise en place pour répondre à ces objectifs. Elle consiste à identifier pour chaque type de catégorie d'occupation du sol (source de pression) les indicateurs dont l'emprise spatiale présente statistiquement l'influence la plus significative sur l'état écologique des cours d'eau, puis à construire différents modèles, à différentes échelles, à partir de ces indicateurs, et enfin à sélectionner le modèle le plus pertinent du point de vue des gestionnaires. Les modèles pressions/impacts développés dérivent de l'approche méthodologique mise en place dans les travaux de Wasson et al., 2005 (cf. Chapitre 3 :III.2.2) et Wasson et al., 2010. Cette approche permet (i) de diminuer l'effet de covariance entre les facteurs

CemOA : archive ouverte d'Irstea / Cemagref

anthropiques et naturels en utilisant l'écart à la référence (EQR) comme variable biologique qui ne dépend pas des variabilités naturelles (typologiques) de l'indice biologique mais principalement des facteurs anthropiques ; et (ii) d'identifier des cours d'eau (qui peuvent être spatialement séparés) souffrant des mêmes « pathologies régionales » sur un territoire et relevant donc d'une même politique de restauration puisqu'ils sont construits à l'échelle des hydro-écorégions.

I.4. Présentation de la zone d'étude

Pour cette étude, nous avons choisi de travailler sur la partie Nord-Ouest de l'HER Tables Calcaires (HER 9-NO). Cette zone est présentée dans le détail dans le Chapitre 4 :II.2.2. On rappelle ici que le choix de cette zone est lié :

- aux caractéristiques hydro-biologiques : elle contient 158 stations (1038 relevés) mesurant la qualité écologique des cours d'eau à partir de l'Indice Biologique Global Normalisé (IBGN) (Figure 4.2). Ce nombre de stations est jugé suffisamment grand pour assurer la pertinence statistique de la démarche.
- aux contraintes financières et de temps : cette zone présente le meilleur compromis entre le nombre de stations et l'extension spatiale de la zone d'étude qui limite le nombre d'images à acquérir et à traiter.
- au contexte anthropique et ripicole : des structures particulières de bandes rivulaires sont présentes sur ce site (alternance de forêts de feuillus et de prairies proches du cours d'eau) dans une zone dominée par une forte occupation agricole et par la présence répétée de tâches urbaines (milieux « rurbains »).

II. Description des données biologiques et de pressions

II.1.Les indicateurs d'état écologique des stations

II.1.1. Méthode d'évaluation : l'EQR-IBGN$_{moyen}$

La méthode utilisée pour évaluer l'état écologique des stations est la méthode mise en place en France au sens de la DCE à partir de l'Indice Biologique Global Normalisé (IBGN). Cette méthode est détaillée dans le chapitre 1 de ce mémoire.

On rappelle ici que l'IBGN (A.F.N.O.R. 1992) est basé sur les peuplements de macro-invertébrés benthiques. 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 favorisant la recherche sur les habitats marginaux (Archaimbault and Dumont, 2010). 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 données disponibles pour cet indice proviennent de la base de données « GIRAFE » du Cemagref. Cette base lors du démarrage de la thèse contenait 3640 stations, pour lesquelles on disposait de 12318 relevés (ou notes IBGN) couvrant la période 1992-2004.

Parmi les stations GIRAFE, des sites de référence pour chaque type de cours d'eau (HER X Rang) ont été sélectionnés selon des critères de « très faible pression anthropique », en utilisant à la fois l'avis d'experts des DIREN, selon un questionnaire précis, et des variables d'occupation du sol calculées sous SIG. Les valeurs de référence de l'IBGN ont été validées par comparaison avec des données indépendantes du Cemagref.

CemOA : archive ouverte d'Irstea / Cemagref

Les notes IBGN sont alors transformées, selon les préconisations de la DCE, en EQR (pour Ecological Quality Ratio), selon la formule : EQR= (IBGN-1)/(REF-1) avec REF correspondant à la valeur IBGN de référence (valeur médiane des IBGN des stations de référence).

Cet EQR-IBGN, correspondant à une mesure de « l'écart à la référence », est utilisé dans toutes les analyses suivantes pour déterminer un état écologique normalisé (entre 0 et 1) indépendant des variations naturelles (typologiques) de l'indice ou, autrement dit, un état « principalement » sous contrôle anthropique.

De manière à classer ces stations selon les cinq classes de qualité (Très bon, Bon, Moyen, Médiocre et Mauvais), pour chaque type de cours d'eau, la limite inférieure du « très bon état » est fixée au 25ème percentile de la distribution des valeurs observées sur les sites de référence.

Ensuite, en première hypothèse, la limite inférieure du « bon état » est fixée en divisant en 4 classes égales l'étendue des valeurs entre la limite inférieure du très bon état et la valeur minimum « réaliste » de l'indice, fixée à 1 pour l'IBGN. Chaque limite inférieure de classe est exprimée en EQR-IBGN.

La zone d'étude est recouverte par 158 stations hydro-biologiques (IBGN). 3 stations (2 stations sur l'Orne et 1 station sur la Seyne) ont été supprimées car il n'était pas possible d'obtenir une information d'occupation du sol à THRS au niveau de leur corridor rivulaire amont (les bassins versants amont de ces 3 stations n'étant pas inclus dans la zone d'étude). La population statistique (ou jeu de données) se compose donc de 155 stations IBGN avec un total de 1034 relevés (campagnes de prélèvement) sur la période de 1992 à 2004.

On observe sur ce jeu de données que le nombre de campagnes de prélèvement (de visites) sur cette période est très variable en fonction du site (Figure 9.1). Les sites ont été visités en moyenne 6.7 fois (soit environ une fois tous les deux ans), les extrêmes étant 1 fois (4 stations) et 11 fois (17 stations). On note aussi que les sites n'ont été visités qu'une seule fois par année de prélèvement.

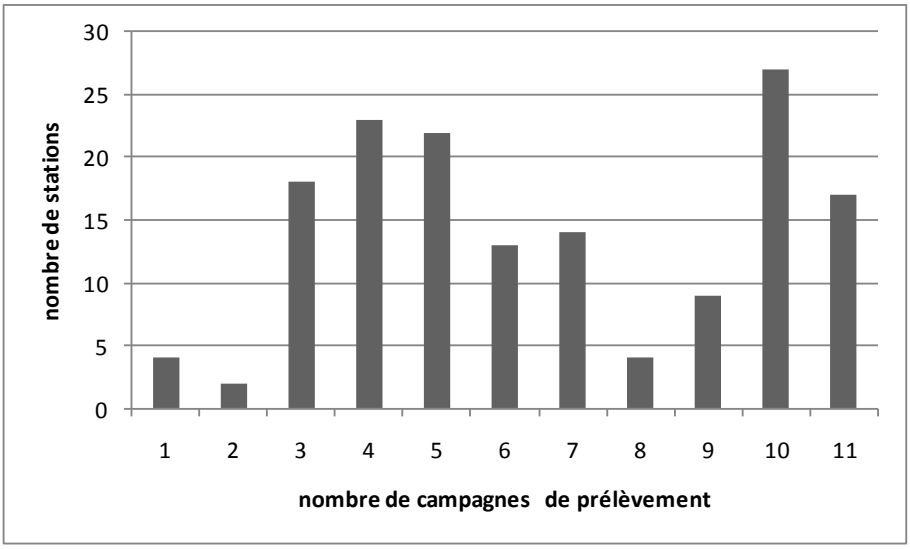


Figure 9.1 : Nombre de stations par nombre de campagnes de prélèvement sur la période de 1992 à 2004.

La moyenne des valeurs mesurées à une même station est l'outil d'inférence le plus simple pour synthétiser au mieux les mesures effectuées sur un site mais seulement lorsque la qualité du site est stable (que se soit avec une variabilité cyclique ou non).

Une analyse de la stabilité des chroniques IBGN a donc été réalisée afin de s'assurer de la validité de l'utilisation de la moyenne comme estimateur de l'état écologique d'un site. Cette analyse dérive de la méthodologie mise en place par Ferreol, 2010 (in press). Dans cette méthode, le critère minimum retenu est une absence de tendance linéaire[41]. La mise en évidence de ce critère dépend de la p-value[42] de la pente du

[41] Ce critère est suffisant pour détecter des évolutions sur des échelles de temps suffisamment longues dans la mesure où toute autre tendance non linéaire mais monotone aura forcement un impact sur la valeur estimée de la pente de la régression linéaire simple. Par contre toute évolution non-monotone est alors assumée comme étant un artefact dû aux variations saisonnières.

[42] La p-value permet de prendre une décision quant au rejet ou non de la nullité du coefficient directeur dans un modèle linéaire simple. En effet, si la p-value est inférieure au niveau de signification α spécifié par l'expérimentateur, le test est dit statistiquement significatif à ce niveau de signification. L'hypothèse nulle H0 est ainsi rejetée au niveau de signification α.

CemOA : archive ouverte d'Irstea / Cemagref

modèle linéaire simple de chaque site ; lorsque cette valeur est supérieure à 0.1, on considère que la pente n'est pas significative et que, par conséquent, la chronique considérée ne montre pas de tendance significative. Afin d'éviter les écueils habituels des modèles linéaires simples, un premier test, distance de cook (Cook, 1977), est appliqué afin d'éliminer les points influents isolés (i.e., une observation qui se trouve « loin » des autres observations) ; une fois enlevés la normalité des résidus, à l'aide du test de Shapiro-Wilk (Shapiro and Wilk, 1965), est vérifiée : dans le cas où la normalité n'est pas (suffisamment) conservée, donc que la régression linéaire est un (trop) mauvais modèle et donc que la valeur de la pente n'est pas représentative, la chronique IBGN du site est directement considérée comme instable ; dans le cas inverse la stabilité de la chronique d'un site IBGN est alors jugée à partir de la valeur de p-value de la pente du modèle.

Comme l'existence d'une tendance linéaire ne peut être mise en évidence qu'à partir de chroniques pour lesquelles il existe un nombre de données suffisant, cette méthode requière 10 notes au minimum. Or, compte tenu de la moyenne de prélèvement des sites de notre échantillon, nous avons appliqué cette méthode uniquement sur des sites ayant été prélevée 6 fois au minimum afin d'avoir une idée de la stabilité des chroniques IBGN sur au moins la moitié de notre échantillon. Ce qui impose d'emblée une certaine prudence sur l'analyse des résultats.

L'analyse de la stabilité des chroniques IBGN a donc reposé au total sur un échantillon de 86 sites (un peu plus de la moitié des sites de la population). Les résultats issus de cette analyse sont exposés dans l'***ANNEXE X***. Ils montrent que 60 % des sites étudiés possèdent des chroniques significativement stables sur cette période.

La moyenne apparaît donc être globalement le meilleur estimateur de l'état écologique d'un site sur notre jeu de donnée de 155 stations (au vu des 86 sites étudiés).

On notera tout de même, que les tendances observées sur les sites restant de l'échantillon sont relativement faible, 86 % des tendances ont une pente (en valeur absolue) inférieure à 0.4. Ce qui n'est pas négligeable puisque cette pente de 0.4 correspond par exemple sur une période de 12 ans à un gain ou une perte de 5 points environ sur la note IBGN.

Ces résultats sont bien évidement à prendre avec précautions compte tenu du nombre de prélèvements pris en compte pour analyser ces tendances, mais Il faudra bien garder à l'esprit lors de l'interprétation des résultats que pour certaines stations la moyenne n'est pas forcément l'inférence statistique la plus fiable pour synthétiser les mesures observées.

Par ailleurs, cette analyse révèle également que sur les sites évolutifs (possédant une tendance) de l'échantillon, 86 % tendent vers une dégradation de l'état écologique au regard des macroinvertébrés benthiques.

II.1.2. Limites et Incertitudes de l'IBGN

L'IBGN repose uniquement sur deux métriques (1) la sensibilité des taxons et (2) la diversité (variété taxonomique).

(1) Les taxons indicateurs de l'IBGN sont sélectionnés plutôt en fonction de leur réponse à la pollution organique (Tachet *et al.*, 2000). Ils sont donc vraisemblablement particulièrement plus sensibles par exemple à la proximité de territoires urbanisés générant ce type de pollution (rejets directs par les stations d'épuration notamment). Cependant, ils s'avèrent aussi sensibles à différents types de contaminants chimiques et toxiques (Camargo *et al.*, 2004), ainsi qu'à des dégradations de l'habitat (e.g., colmatage) (Gayraud *et al.*, 2002). En revanche, l'accélération de l'écoulement, liée à un ouvrage hydraulique, se traduira par une raréfaction des habitats lentiques, mais aussi par une meilleure disponibilité en oxygène donc un milieu favorable aux espèces rhéophiles polluosensibles (Wasson *et al.*, 1998).

(2) De façon complémentaire, le nombre de taxons est sensible aux altérations de toutes natures (physique, chimique, toxique) à partir du moment où elles sont suffisamment marquées pour éliminer plusieurs familles d'invertébrés (Wasson *et al.*, 2005). En revanche, les premiers stades de l'eutrophisation (par apport de nutriments) ou des modifications légères de l'habitat (e.g., éclairement) peuvent provoquer une augmentation de la richesse faunistique (Wasson *et al.*, 2006).

Malgré certaines limites, on peut considérer que l'IBGN permet de mesurer une dégradation générale des peuplements en réponse à un large spectre de perturbations anthropiques. Il a été comparé à d'autres méthodes européennes dans le cadre de l'exercice d'inter-étalonnage requis par la DCE (Buffagni *et al.*, 2007) et est actuellement utilisé pour évaluer l'état du compartiment « invertébrés benthiques des cours d'eau » dans les réseaux nationaux (réseau de référence, réseau de surveillance, réseau de contrôle opérationnel et

CemOA : archive ouverte d'Irstea / Cemagref

réseaux complémentaires). De plus, le calcul de l'IBGN est assez robuste ceci malgré l'existence de sources de variabilité liées à la subjectivité du protocole et aux circonstances particulières lors du prélèvement comme la présence d'un stade dormant chez l'espèce au moment de la récolte (Berryman, 1990).

On gardera donc à l'esprit lors de l'interprétation des résultats que (i) l'IBGN est actuellement l'indicateur le plus robuste pour juger de l'état écologique d'une station mais que les métriques qu'il synthétise ne rendent pas compte à priori de toute la diversité des pressions; et (ii) que la moyenne des IBGN, observée sur la période de prélèvement de 1992 à 2004, n'est qu'un estimateur grossier, pour synthétiser les mesures observées et donc estimer l'état réel des stations (cf. II.1.1).

II.2. Les indicateurs de pressions d'occupation du sol

II.2.1. Méthode d'évaluation des indicateurs de pressions

II.2.1.1 Définitions (rappels)

Pour évaluer le degré d'anthropisation des cours d'eau de façon cohérente et homogène à l'échelle du territoire national, on rappelle que la seule possibilité au vu des données existantes est d'analyser, à l'aide d'indicateurs spatialisés (indicateurs du paysage), l'occupation du sol qui constitue l'une des « forces motrices » (selon la logique DPSIR cf. Chapitre 3 :I.2) à l'origine de la plupart des pressions sur les milieux aquatiques.

Un indicateur spatialisé est défini par un attribut de structure de paysage (composition et/ou configuration) calculé pour une catégorie d'occupation du sol (un type d'occupation du sol selon une certaine typologie) à l'échelle d'une emprise spatiale donnée. Il est donc défini par le quadruplet type d'occupation du sol / structure de paysage /emprise spatiale / source de données.

Les indicateurs spatialisés utilisés dans la littérature pour traduire les pressions exercées par l'occupation du sol sur l'état chimique, physique ou biologique des cours d'eau sont construits au niveau de la classe, i.e. séparément pour chaque type d'occupation du sol. Dans ces études, l'occupation du sol est catégorisée selon une typologie de classes significative (conceptuellement pertinente) pour une zone d'étude et une échelle donnée.

L'emprise spatiale sur laquelle est extrait l'attribut de structure du paysage est généralement localisée en amont de chaque station écologique selon trois échelles (cf. Chapitre 3 :III.2.1.3) : l'échelle du Bassin Versant (BV) et deux échelles au niveau du corridor rivulaire : le corridor rivulaire amont (CRamont) et le corridor rivulaire local (CRlocal). Ces échelles correspondent respectivement aux échelles régionale, linéaire et ponctuelle dans la hiérarchie des échelles de fonctionnement des écosystèmes d'eau courante (cf. Chapitre 3 :III.1.2). Ces trois échelles seront explorées dans cette étude.

On présente successivement, à la suite de cette partie, l'attribut du paysage, les sources de données d'occupation du sol, la typologie et les emprises spatiales (aux différentes échelles) utilisés dans cette étude.

II.2.1.2 Choix de l'attribut du paysage

Plusieurs attributs de paysage peuvent être utilisés pour quantifier les pressions d'occupation du sol (cf. Chapitre 7 :II.). C'est le cas par exemple des attributs spécifiques pour quantifier l'effet protecteur de la végétation rivulaire (largeur et uniformité de la bande, continuité…).

Pour ne pas complexifier l'étude par l'augmentation du nombre d'indicateurs à interpréter, et ainsi répondre plus facilement aux objectifs fixés, on travaillera uniquement avec les indicateurs utilisant le pourcentage surfacique. Cet attribut, facilement interprétable, traduit la présence et l'intensité (plus ou moins de surface) d'une catégorie d'occupation du sol susceptible de générer des pressions sur le milieu aquatique.

Par exemple, un indicateur relatif au tissu urbain continu peut traduire tout à la fois les pressions hydromorphologiques (e.g., rectification du tracé, digues, suppression de la ripisylves…) et chimiques (e.g.,

CemOA : archive ouverte d'Irstea / Cemagref

apport de matière organique, pollution toxique…) exercées par la zone urbaine qui perturbent le fonctionnement et la structure physique du cours d'eau (Borchardt and Ritcher, 2003).

II.2.1.3 Données d'occupation du sol

Deux sources de données d'occupation du sol de résolutions différentes sont disponibles sur la zone d'étude pour quantifier les indicateurs de source de pressions. La donnée CORINE[43] Land Cover (CLC, cf. II.2.1.3.1) de résolution 25ha est plus gros grossière que la donnée métrique issue de l'information à Très Haute Résolution Spatiale (THRS cf. II.2.1.3.2). A chaque source de données correspond une typologie arborescente organisée en plusieurs niveaux.

II.2.1.3.1 CORINE Land Cover (CLC)

La couche d'information géographique CORINE Land Cover 2000 (CLC) est la base de données d'occupation du sol européenne issue de la photo-interprétation assistée par ordinateur d'images Landsat (30 m) et Spot4 (20 m) acquises entre la période de 1999 à 2000 (Bossard et al., 2000). L'unité cartographiée est de 25 ha minimum. Elle repose sur une nomenclature standard hiérarchisée à 3 niveaux et 44 postes (***ANNEXE IV.1***) répartis selon 5 grands types d'occupation du territoire (territoires artificialisés, territoires agricoles, forêts et milieux semi-naturels, zones humides et surfaces en eaux). Elle possède une précision de classification supérieure à 85 %.

C'est le type de base de données couramment utilisé dans les modèles pressions/impacts à large échelle pour quantifier les indicateurs d'occupation du sol quelque soit l'échelle d'analyse (bassin versant ou corridor rivulaire).

II.2.1.3.2 La donnée à Très Haute Résolution Spatiale (THRS)

La donnée d'occupation du sol à Très Haute Résolution Spatiale (THRS) au niveau du corridor rivulaire a été obtenue sur la zone d'étude à partir de la procédure de classification développée dans la deuxième partie de ce mémoire.

On rappelle que cette procédure utilise en entrée des sources de données disponibles gratuitement ou à moindre coûts pour les gestionnaires sur l'ensemble du territoire français : deux sources de données images (images Spot5 XS - 10 m et orthophotos IGN– 0,5 m) et quatre sources de données auxiliaires (CORINE Land Cover, la Route 500®, les entités surfaciques de la BDCarthage® et le Registre Parcellaire Graphique).

La classification est menée selon une approche orientée objet multi-échelles définissant une hiérarchie de quatre niveaux d'objets emboîtés et utilisant un arbre de classification descendant mettant en œuvre des règles de décision booléennes ou floues. Ces règles sont déterminées par expertise et combinent l'information des différentes sources de données d'entrée.

L'application de cette procédure sur le corridor rivulaire de la zone d'étude (5600 km²) est présentée en ***ANNEXE XI***. Elle a nécessité la mobilisation de 11 images SPOT 5 XS[44] et 2455 orthophotos[45]. En plus des données mobilisées pour l'application sur le bassin versant de l'Hérault (cf. Chapitre 6 :II.1.1), on disposait sur cette zone d'étude de la donnée du RPG fournissant une information sur les types de cultures par îlots déclarés par les agriculteurs., Cette donnée s'est avérée très intéressante car elle a permis d'une part de détailler les classes d'agriculture (en attribuant pour chaque îlot le type de culture occupant majoritairement sa superficie : 80 % des îlots sur la zone d'étude possèdent un seul type de culture) et d'autre part de réduire le temps de traitement des images (les données du RPG couvrent plus de 46 % de la superficie des corridors rivulaires analysés).

[43] Co-Ordination of Information on the Environnement
[44] Acquises avec l'appui du programme ISIS du CNES, facilitant l'accès de la communauté scientifique à l'imagerie satellitaire Spot.
[45] Mises à disposition par le Ministère de l'Agriculture dans le cadre de sa convention avec IGN

CemOA : archive ouverte d'Irstea / Cemagref

La donnée d'occupation du sol résultant de la procédure de classification des données multi-sources présente plusieurs intérêts.

Du point de vue sémantique, elle fournit une information détaillée comportant 62 postes d'occupation du sol structurés selon une nomenclature hiérarchisée à quatre niveaux (***ANNEXE IV.2***) calquée sur la nomenclature de CLC. La qualité de cette donnée n'a pas été évaluée car l'étape de validation par construction de matrices de confusion (étape 7 de la procédure, cf. Chapitre 5 :III.7.1) n'a pas pu être menée dans le cadre de la thèse.

Du point de vu géométrique, elle apporte une précision quasi-métrique car la majorité des objets (plus de 90 %) ont été délimités à partir de l'information des orthophotos à 0.5 m de résolution spatiale et de la donnée RPG qui se base également sur les orthophotos à 0.5 m.

Du point de vue temporel, cette donnée apporte une information hétérogène, puisque les images Spot5 XS (images d'archives) ont été acquises sur la période de 2003 à 2007, les orthophotos sur la période de 2002 à 2006 et les données auxiliaires en 2006, excepté CLC qui date de 2000. Néanmoins, l'identification d'une date précise est possible pour certains postes d'occupation des sols obtenus par des règles de classification s'appliquant à une seule et même source de donnée. C'est le cas par exemple de certains postes artificialisés dérivant uniquement de la donnée CLC ou encore de ceux obtenus uniquement à partir de la donnée RPG. Cette date est spécifiée pour ces postes dans la présentation de la nomenclature (cf. ***ANNEXE IV.2***).

II.2.1.3.3 Représentations cartographiques

Des représentations cartographiques pour chaque donnée sont proposées au niveau du corridor rivulaire, Figure 9.2 pour la donnée CLC, et Figure 9.3 pour la donnée THRS.

De manière à rendre ces cartes plus facilement lisibles et comparables, ces deux représentations ont été réalisées à partir d'une même typologie simplifiée en 8 catégories d'occupation du sol, correspondant à une agrégation de un ou plusieurs postes d'occupation du sol (cf. Chapitre 10 :I.2.1).

En comparant visuellement ces deux représentations, on observe sans surprise que l'information apportée par la donnée THRS est beaucoup plus détaillée que celle fournie par CLC. A titre indicatif, la donnée THRS comporte environ 100 000 fois plus de polygones que la donnée CLC (selon la typologie de niveau 3).

II.2.1.4 Choix de la typologie

Étant donné que la description les relations pressions/impacts sur la zone d'étude se fait de manière totalement exploratoire, i.e. sans connaissance préalable étayée sur les types de sources de pressions existant dans cette zone, nous avons décidé de développer l'ensemble des modèles à partir d'un niveau détaillé d'occupation du sol. Cette typologie « détaillée » correspond au troisième niveau de la nomenclature des données d'occupation du sol. Elle est composée de 44 postes pour CLC (***ANNEXE IV.1***) et de 52 postes pour la donnée à THRS (***ANNEXE IV.2***).

Nous avons volontairement laissé de côté (en raison de contraintes temporelles de travail) la question du niveau de nomenclature optimal pour la construction des modèles pressions/impacts.

CemOA : archive ouverte d'Irstea / Cemagref

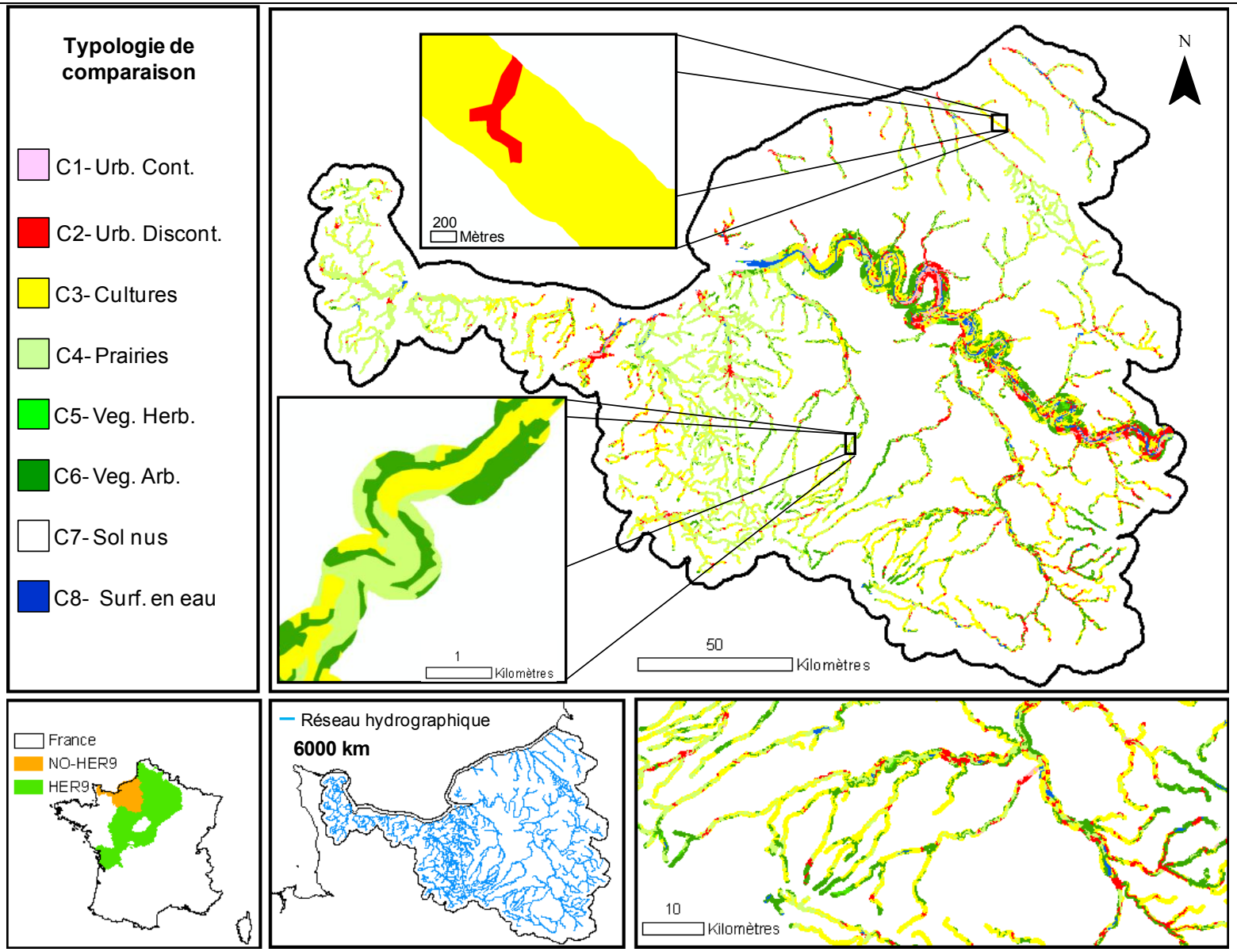


Figure 9.2 : Cartographie de l'occupation du sol dans l'HER9-NO au niveau du corridor rivulaire selon la donnée à CLC agrégée suivant la typologie de comparaison en 8 catégories

CemOA : archive ouverte d'Irstea / Cemagref

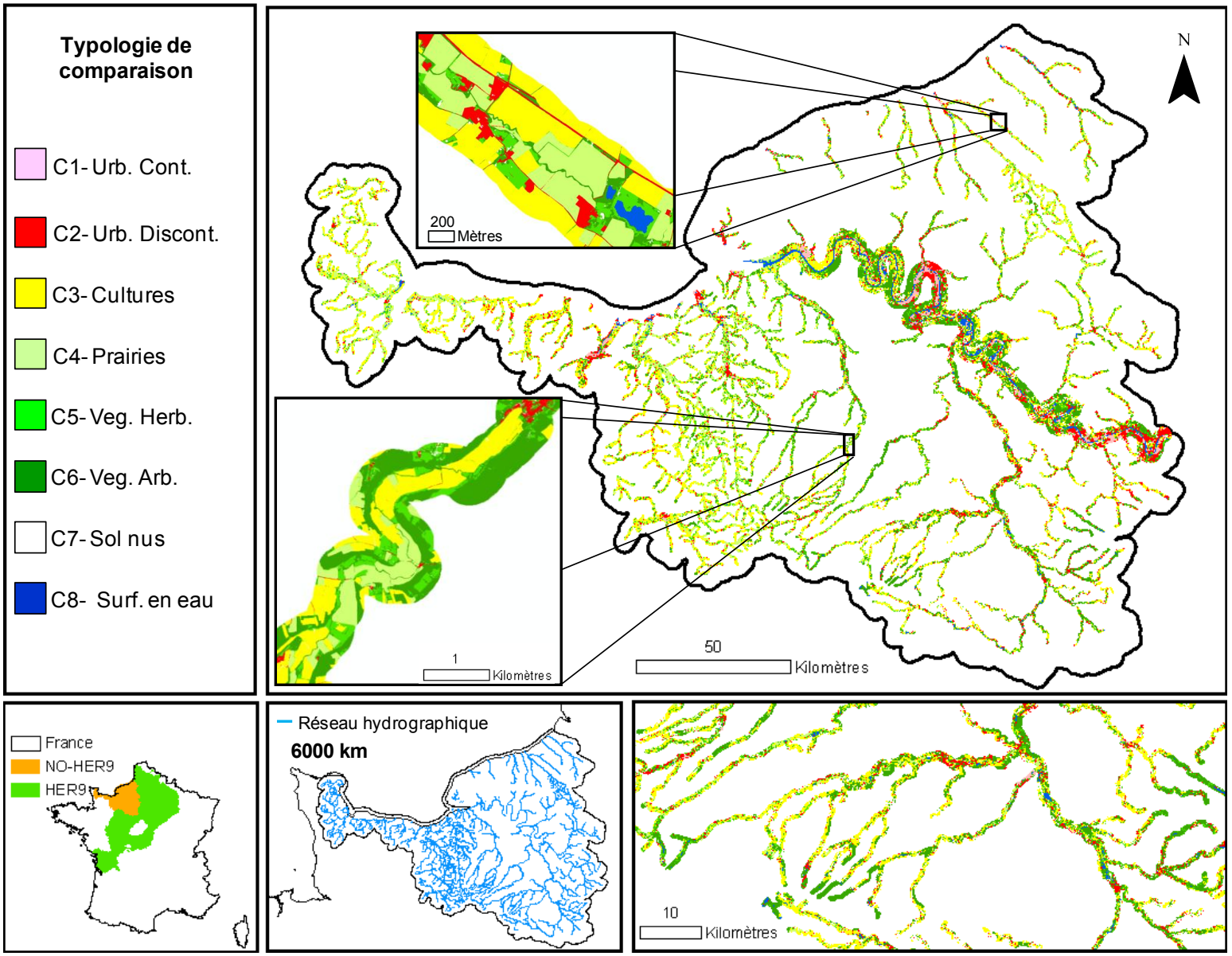


Figure 9.3 : Cartographie de l'occupation du sol dans l'HER9-NO au niveau du corridor rivulaire selon la donnée à THRS agrégée suivant la typologie de comparaison en 8 catégories

CemOA : archive ouverte d'Irstea / Cemagref

II.2.1.5 Choix des emprises spatiales

II.2.1.5.1Emprise spatiale au niveau du bassin versant

L'emprise spatiale utilisée dans ces modèles à l'échelle du bassin versant (BV) correspond au bassin versant amont de chaque station de suivi de l'état écologique.

Compte tenu du nombre de stations ponctuelles sur lesquelles se trouvent les données biologiques observées, un outil spécifique de délimitation et caractérisation du bassin versant a été utilisé. Cet outil, développé par Pella *et al.*, 2004 est basé sur l'analyse du modèle numérique de terrain dérivé de la BDALTI® de l'Institut Géographique National (IGN) au pas de 250 m, sur le découpage en zones hydrographiques et le tracé du réseau hydrographique de la BD Carthage® niveau 2 (cf. Chapitre 5 :III.2). Etant donné que ce calcul de bassin versant est basé sur le MNT de l'IGN au pas de 250 m, la surface minimale de détection est de quelques kilomètres carrés. En zone de relief peu marqué, l'outil utilisé touche ses limites et le recours à l'expertise manuelle est parfois indispensable.

Ensuite, à partir de l'enveloppe du bassin versant ainsi délimitée, il est possible de calculer le pourcentage de surface des différentes catégories d'occupation du sol (cf. exemple Figure 9.4).

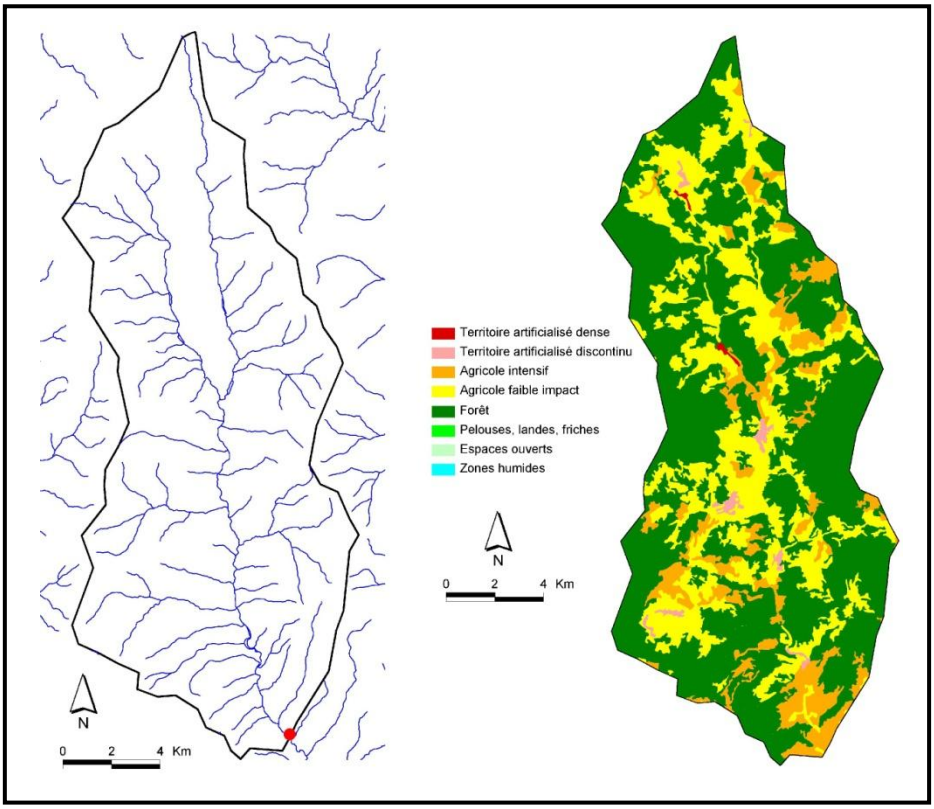


Figure 9.4 : Emprise spatiale à l'échelle du bassin versant. Exemple de calcul à partir d'une station (en rouge) et des pressions anthropiques dérivées de l'analyse de CORINE Land Cover (d'après Pella *et al.*, 2004)

II.2.1.5.2Emprise spatiales au niveau du corridor rivulaire

II.2.1.5.2.1 Mode de spatialisation des sources de pressions (fixe ou multiple)

Les emprises spatiales définies au niveau du corridor rivulaire pour étudier l'effet d'une source de pressions donnée sont délimitées généralement à l'aide d'un buffer autour du réseau hydrographique. Deux modes de spatialisation des sources de pressions sont possibles au niveau du corridor rivulaire selon le type de buffer utilisé :

CemOA : archive ouverte d'Irstea / Cemagref

- Le mode de spatialisation « fixe » : ce mode utilise une seule dimension d'emprise (un seul buffer). Par conséquent toutes les sources de pressions influençant l'état écologique des cours d'eau, quelles qu'elles soient, sont étudiées dans un même espace au niveau du corridor rivulaire. La délimitation de ce buffer se base sur des hypothèses définies préalablement par l'analyste. C'est le mode de spatialisation couramment utilisé dans les modèles pressions/impacts à large échelle.

- Le mode de spatialisation « multiple » (cf. Chapitre 7 :II.2.2) : ce mode consiste à construire plusieurs buffers candidats de dimensions différentes. Par conséquent, l'influence d'une source de pressions donnée sur l'état écologique des cours d'eau peut-être étudiée en considérant plusieurs espaces.

Alors qu'on utilisera dans cette étude ces deux modes de spatialisation à l'échelle du CRlocal (cf. II.2.1.5.2.4), seul le mode « multiple » sera utilisé à l'échelle du CRamont (cf. II.2.1.5.2.3)

II.2.1.5.2.2 Prise en compte de l'aval de la station

Dans cette étude, on considérera une portion de 500 m au maximum à l'aval de la station IBGN pour définir les emprises au niveau du corridor rivulaire. Cette portion aval, prise en compte dans l'étude de modélisation pressions/impacts de Wasson et al. (2005), a été considérée pour deux principales raisons :

- premièrement, l'occupation du sol à proximité en aval de la station peut avoir une influence sur la note IBGN. Elle peut avoir aussi bien une influence négative (e.g., la présence d'un tissu urbain s'accompagne généralement de la présence d'ouvrages de protection en amont de ce tissu impactant l'état physique de la station) que positive (e.g., la présence d'une végétation naturelle abondante à proximité peut servir de zone refuge aux macroinvertébrés qui pourront à la suite recoloniser le milieux, bien évidement en aval de cette zone, phénomène de dérive des invertébrés, mais aussi sur une distance en amont assez proche car certains d'entre eux se déplacent via les airs pendant leur âge adulte).

- deuxièmement, les stations, même si elles sont représentées géométriquement par des points sur le réseau, correspondent en réalité à un tronçon de rivière d'environ 100 m de long. Les points de mesure ne sont pas forcément positionnés en aval des tronçons étudiés. Le fait donc de prendre une partie aval permet de compenser l'éventuelle erreur de positionnement des stations sur le réseau.

II.2.1.5.2.3 Emprises à l'échelle du CRamont (mode multiple)

Seul le mode de spatialisation « multiple » est utilisé à cette échelle. L'échelle du CRamont correspond à l'espace le long (à proximité) de l'ensemble du réseau amont de la station écologique étudiée. Les buffers diffèrent donc uniquement par leur largeur (distance latérale au cours d'eau).

Dans cette étude, 15 buffers de largeurs différentes (par commodité on appelle « largeur » la demi-largeur de part et d'autre du cours d'eau) ont été définis selon un pas régulier (cf. Chapitre 7 :III.2.2.3.1) : 10 buffers entre 0 et 50 m selon un pas de 5 m et 5 buffers entre 50 et 300 m d'un pas de 50 m. Ces buffers sont construits sur la totalité du réseau hydrographique amont ainsi que sur un tronçon de 500 m du réseau hydrographique aval de chaque station écologique (Figure 9.5).

Les emprises spatiales définies à partir de ces buffers seront candidates pour étudier l'effet d'une source de pressions donnée sur l'état écologique. Le pas entre les largeurs des buffers est plus réduit (pas de 5 m) dans la bande de 50 m car c'est un espace potentiellement privilégié pour la mise en œuvre de stratégies de préservation et/ou de restauration.

Les buffers sont délimités suivant la méthodologie décrite dans le Chapitre 7 :III. et construits à partir de la donnée hydrographique « polygone » (cf. Chapitre 7 :III.2.1.1). Une fois l'enveloppe d'un buffer délimitée, il est facile de calculer à l'aide d'un SIG le pourcentage de surface des différentes catégories d'occupation du sol sur son emprise.

CemOA : archive ouverte d'Irstea / Cemagref

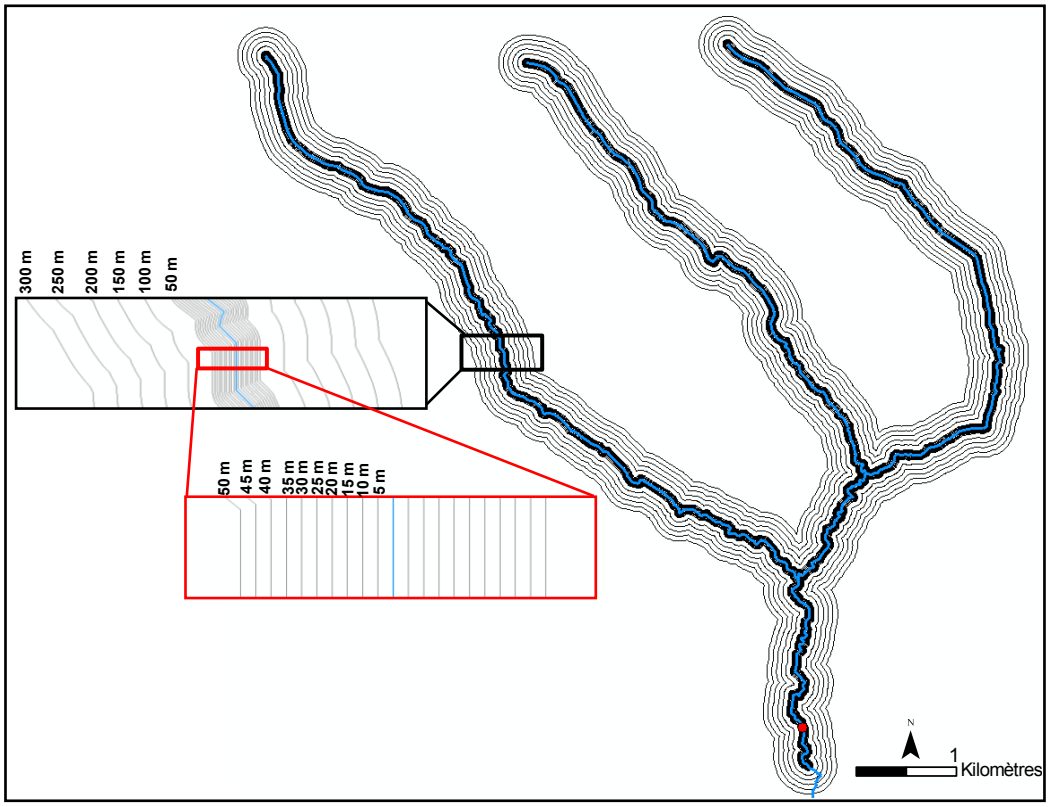


Figure 9.5 : Emprises spatiales mises en place dans le cadre du mode de spatialisation « multiple » à l'échelle du corridor rivulaire amont. Pour chaque station écologique (en rouge) ces buffers sont construits sur la totalité du réseau amont ainsi que sur une longueur de réseau aval de 500 m.

II.2.1.5.2.4 Emprises à l'échelle du CRlocal

Les deux modes de spatialisation, « fixe » et « multiple », sont utilisés à cette échelle. L'emprise définie pour le mode de spatialisation « fixe » est présentée dans le § II.2.1.5.2.4.1, celle définie pour le mode « multiple » dans le § II.2.1.5.2.4.2.

II.2.1.5.2.4.1 Emprise du mode fixe

Les hypothèses de construction de cette emprise sont celles utilisées dans les modèles pressions/impacts de Wasson et al. (2005) à l'échelle du corridor rivulaire local (CRlocal). Cette emprise permet d'étudier l'influence des sources de pression sur le lit majeur amont et aval de la station. Les hypothèses sont les suivantes (Figure 9.6) :

- Le tronçon étudié mesure environ 3 km de long soit 2.5 km à l'amont et 0.5 km à l'aval de la station considérée.
- La largeur de la zone à prendre en compte de part et d'autre du cours d'eau dépend de la largeur du lit majeur. Elle est estimée à partir des rangs de Strahler : 100 m pour les cours d'eau de rang 1, 2 et 3 ; 140 m pour ceux de rang 4 ; 250 m pour ceux de rang 5 ; 600 m pour ceux de rang 6 ; 1200 m pour ceux de rang 7 et enfin 2400 m pour ceux de rang 8.

Seuls les arcs de même rang sont retenus pour le calcul de la zone tampon. Il est donc possible, lorsque la station est proche d'une confluence qui change le rang en aval, de ne considérer que la portion du cours d'eau de même rang que celle sur laquelle est située la station.

Une fois l'enveloppe du buffer « fixe » délimitée pour chaque station écologique, il est possible de calculer le pourcentage de surface des différentes catégories d'occupation du sol sur son emprise.

CemOA : archive ouverte d'Irstea / Cemagref

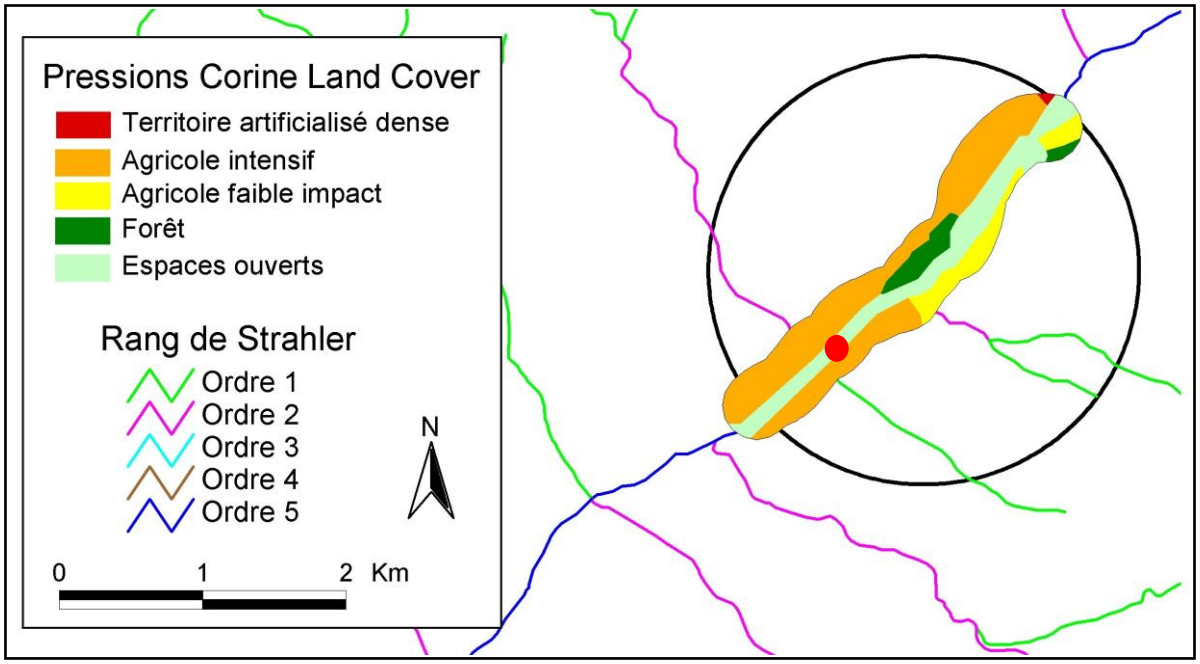


Figure 9.6 : Emprise spatiale définie par le mode de spatialisation « fixe » à l'échelle du CRlocal : exemple de délimitation pour une station écologique (en rouge) de l'emprise sur laquelle seront calculés les indicateurs de pressions anthropiques dérivées de l'analyse de CORINE Land Cover (Pella *et al.*, 2004)

II.2.1.5.2.4.2 Emprises du mode multiple

Dans ce mode, plusieurs buffers ont été définis à l'échelle du CRlocal selon des largeurs (distance latérale au cours d'eau) et des longueurs (distance longitudinale à la station $Dist_{amont}$) croissantes. La distance avale à la station ($Dist_{aval}$) dépend de la $Dist_{amont}$: lorsque celle-ci est inférieure ou égale à 1500 m alors $Dist_{aval}$ est égale à 1/3 de $Dist_{amont}$, et dans le cas contraire $Dist_{aval}$ est fixée à 500 m.

En tout, 150 buffers de couples largeurs, longueurs différents ont été définis. 10 longueurs $Dist_{amont}$ ont été établies selon un pas de 100 m de 0 à 500 m et selon un pas de 500 m de 500 à 3000 m (Figure 9.7-A). Puis pour chaque longueur, 15 largeurs différentes ont été mises en place selon un pas latéral de 5 m entre 0 et 50 m et d'un pas de 50 m entre 50 et 300 m (comme pour le CRamont). Une illustration des différentes largeurs est proposée Figure 9.7-B pour la longueur 2500 m.

Ces buffers seront candidats pour étudier l'effet d'une source de pressions donnée sur l'état écologique. Pour les mêmes raisons qu'à l'échelle du CRamont, les distances entre les buffers sont plus réduites (pas de 5 m) dans la bande de 50 m (de part et d'autre du cours).

Les buffers sont délimités suivant la méthodologie décrite dans le Chapitre 7 :III. et construits à partir de la donnée hydrographique « polygone » (cf. Chapitre 7 :III.2.1.1). Une fois l'enveloppe d'un buffer délimitée, il est facile de calculer à l'aide d'un SIG le pourcentage de surface des différentes catégories d'occupation du sol sur son emprise.

CemOA : archive ouverte d'Irstea / Cemagref

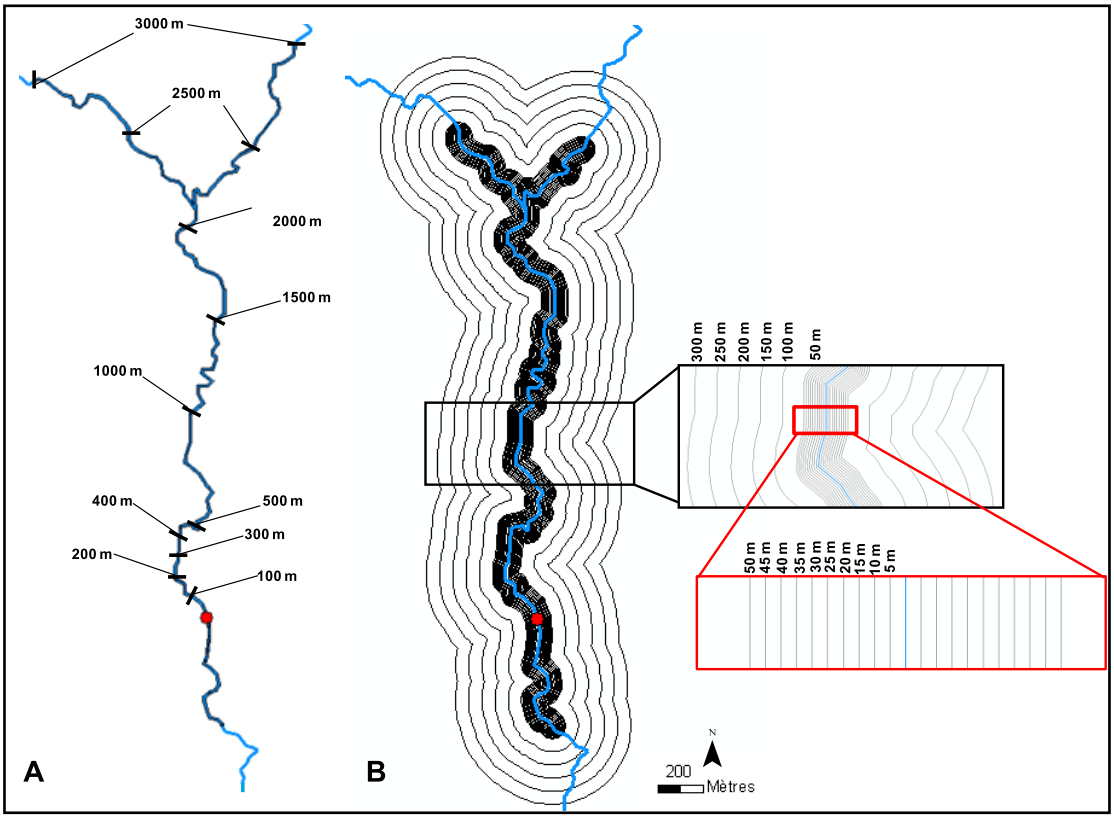


Figure 9.7 : Emprises spatiales mises en place dans le cadre du mode de spatialisation « multiple » à l'échelle du CRlocal. Ces buffers, de longueurs et de largeurs différentes, sont construits autour du réseau hydrographique en amont et en aval de la station écologique (en rouge). En A : Les différentes longueurs étudiées ; en B : Les différentes largeurs sur un exemple de longueur (ici 2500 m)

II.2.1.6 Récapitulatif et dénomination des indicateurs

Les indicateurs utilisés dans le cadre de cet étude se différencient par leur triplet « type d'occupation du sol » / « emprise spatiale » / « source de données ». Nous rappelons en effet que nous nous sommes limités à un seul attribut de structure du paysage : le pourcentage de superficie d'un type d'occupation du sol au sein d'une emprise (par exemple 28% de couvert arboré).

Le Tableau 9-1 récapitule le nombre d'indicateurs construits. Un indicateur a été généré pour :

- chaque source de données d'occupation du sol (2)
  - chaque poste d'occupation du sol de cette source de donnée (44 pour CLC ; 52 pour THRS)
    - chacune des échelles (BV, CRamont, CRlocal).
      - Chacune des emprises spatiales de ces échelles

Au final, 15980 indicateurs, traduisant le pourcentage d'un type d'occupation du sol dans une emprise spatiale donnée, pour une source de données particulière, ont été construits pour chacune des 155 stations IBGN de la zone d'étude. Ces indicateurs ont été stockés et structurés dans une base de données de manière à faciliter les traitements de modélisation.

Les indicateurs spatialisés sont nommés suivant la terminologie mise en place dans la troisième partie de ce mémoire (cf. Chapitre 7 :II.1) avec « IBV » pour les indicateurs construits à l'échelle du bassin versant et « ICR » pour ceux construits au niveau du corridor rivulaire (CRamont et CRlocal). A noter qu'un nouvel argument s'ajoute à la fin des noms des indicateurs afin d'identifier la source de donnée utilisée pour les construire (CLC ou THRS).

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 9-1 : Récapitulatif des indicateurs d'occupations du sol générés pour chaque station écologique de la zone d'étude

Données d'OS	Nombre de postes d'OS	Echelles d'analyses	Nombre d'emprises candidates	Nombre d'indicateurs
CLC	44	BV	1	44
		CRamont	15	660
		CRlocal	151	6644
			Total	**7348**
TRHS	52	CRamont	15	780
		CRlocal	151	7852
			Total	**8632**
			Total	**15980**

On remarquera que les indicateurs utilisant l'emprise spatiale de 5 m, correspondent à la famille des indicateurs du linéaire, et les indicateurs construits avec l'emprise spatiale du mode fixe correspondent à la famille des indicateurs du lit majeur (cf. respectivement Chapitre 7 :II.2.3.1 et Chapitre 7 :II.2.3.2).

II.2.2.Limites et incertitudes des indicateurs d'occupation du sol

L'occupation du sol est l'une des forces motrices (source de pressions), générant potentiellement des pressions dégradantes ou améliorantes sur l'état écologique des cours d'eau. Un indicateur d'occupation du sol traduit donc un éventail de pressions possibles s'exerçant de façon combinées suivant divers mécanismes sur le milieu aquatique.

Pour des raisons évidentes développées ci-dessous, l'occupation du sol ne permet pas de quantifier de façon précise une pression donnée et encore moins le ou les mécanismes mis en œuvre.

D'une part, la relation quantitative entre la superficie d'un type d'occupation du sol et les rejets polluants ponctuels ou diffus est loin d'être simple. Par exemple, l'occupation du sol « tissu urbain continu » ou « tissu urbain discontinu » ne tient compte ni de la densité de population des zones urbanisées, ni du taux de traitement des rejets urbains, et représente assez mal les sources ponctuelles de rejets toxiques. Aux zones de grandes cultures peuvent être associées les sources de pollutions diffuses (nutriments, pesticides…), mais à l'exception des nitrates, la relation quantitative entre superficie cultivée et rejets polluants reste complexe. Elle dépend aussi d'autres paramètres liés aux systèmes de production, aux types de culture, aux pratiques culturales, qui peuvent être très différents suivant les régions. De la même façon, les pollutions dues aux élevages ne sauraient être directement reliées aux seules superficies de prairies car l'intensification des pratiques d'élevage est aussi très différente selon les régions.

D'autre part, l'occupation du sol ne révèle qu'indirectement les pressions hydromorphologiques. Certaines altérations physiques comme la chenalisation peuvent être systématiquement suspectées dans les zones urbanisées où de terres cultivées, mais nombre de pressions importantes comme les barrages, les dérivations, la navigation, l'endiguement, ne sont pas du tout prises en compte par les variables d'occupation du sol (Wasson et al., 2005).

Par ailleurs, par leur mode de construction, les indicateurs d'occupation du sol présentent une certaine incertitude liée pour partie aux erreurs de décalage de la donnée hydrographique (de façon très limitée cf. Chapitre 7 :III.2.1.1) et pour partie aux erreurs de classification de l'occupation du sol, cette incertitude n'étant pas quantifiée compte tenu des difficultés rencontrées pour spatialiser les erreurs de classification (ce point est discuté dans le Chapitre 6 :III. et Chapitre 8 :IV.).

On devra donc garder à l'esprit les limitations et les incertitudes de ces indicateurs pour l'interprétation des résultats.

CemOA : archive ouverte d'Irstea / Cemagref

III. Méthode de construction d'un modèle pressions/impacts

III.1. Une procédure en deux étapes

Le développement d'un modèle pressions/impacts suit une procédure en deux étapes :

- la **1ère étape** (cf. III.2) permet, pour chaque échelle étudiée et chaque source de pression, de sélectionner l'emprise spatiale sur laquelle cette source de pression a l'influence la plus significative sur l'état écologique observé aux stations. Cette étape est indispensable car on ne dispose pas des moyens de mener une analyse multi-variable sur 15980 variables dont certaines sont très fortement corrélées. Par exemple au niveau du corridor rivulaire, l'occupation du sol estimée dans les différents buffers candidats est fortement similaire particulièrement pour les buffers de dimensions proches (Van Sickle, 2003 ; Allan, 2004a; King *et al.*, 2005). Cette étape nous permet donc de réduire la dimension du problème en sélectionnant pour une échelle donnée et pour chaque source de pressions les indicateurs qui seront intégrées dans l'analyse des relations pressions/impacts (2ème étape).

- la **2ème étape** (cf. III.3) permet, parmi les indicateurs sélectionnés dans la 1re étape, d'identifier la combinaison d'indicateurs de pressions aux différentes échelles sur leurs emprises spatiales significatives, exerçant l'influence la plus marquée sur la réponse biologique.

III.2. Méthode de sélection des emprises (1ère étape)

L'objectif est, pour une échelle donnée (Corridor rivulaire amont ou corridor rivulaire local), de sélectionner l'emprise spatiale sur laquelle une source de pression donnée a l'influence la plus significative sur l'état écologique des stations. Cette emprise « influente » est identifiée en comparant les corrélations bi-variées entre les indicateurs de pression calculés sur les différentes emprises candidates et la réponse biologique aux stations, et en choisissant l'emprise assurant la plus forte corrélation (e.g., Johnson and Covich, 1997 ; Sponseller *et al.*, 2001)

Autrement dit, l'objectif est de rechercher statistiquement, parmi des indicateurs de même échelle et de même source de pression (i.e. catégorie d'occupation du sol), l'indicateur possédant le lien le plus fort avec l'EQR-IBGN$_{moyen}$.

Le lien entre un indicateur de pressions d'occupation du sol donné et l'EQR-IBGN$_{moyen}$ est mesuré à l'aide du coefficient de détermination R^2 issu de la Régression Linaire Simple (RLS).

La RLS consiste à rechercher, à partir d'observations (y_i, x_i), $i = 1, \ldots, n$, une équation de la forme :

$$(1)\quad y_i = ax_i + b + \varepsilon_i$$

où y est la variable dépendante (ici l'état écologique aux stations mesuré par l'EQR-IBGN), x la variable explicative (ici l'indicateur de pressions d'occupation du sol) et ε le résidu du modèle suivant une loi normale N $(0, \sigma 2)$. Le paramètre a (pente) et b (ordonnée à l'origine) sont estimés par le principe des moindres carrés qui consiste à choisir ces paramètres de façon à minimiser la somme des carrés des erreurs, selon l'équation :

CemOA : archive ouverte d'Irstea / Cemagref

(2) $$SC_{res} = \sum_i \left(y_i - (ax_i + b) \right)^2$$

les estimateurs de ***a*** et ***b*** par le critère des moindres carrés sont :

(3) $$\hat{a} = \frac{\sum_i (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}$$ et (4) $$\hat{b} = \hat{y} - \hat{a}\bar{x}$$

où $\bar{x} = (x_1 + x_2 + ... + x_n)/n$ et $\bar{y} = (y_1 + y_2 + ... + y_n)/n$; les valeurs $\hat{y}_i = \hat{a}x_i + \hat{b}$ (*i = 1, …, n)* sont appelées valeurs ajustées et les différences $r_i = y_i - \hat{y}_i$ (*i = 1, …, n)* sont dites résidus.

Le coefficient de détermination R², aussi appelé coefficient de corrélation simple est ensuite estimé. Il correspond à la proportion de la variation de *y* expliquée par la variation de *x*, autrement dit au degré de liaison entre ces deux variables qui se traduit mathématiquement par l'équation suivante :

(4) $$R^2 = 1 - \frac{SCR}{SCT}$$

où SCR représente la somme des carrés des résidus totaux qui s'écrit $\sum_i (r_i)^2$; et SCT la somme des carrés totaux qui s'écrit $\sum_i (\bar{y} - y_i)^2$.

L'indicateur qui, parmi les indicateurs d'une même source de pressions et de même échelle mais d'emprises différentes, obtient le plus fort R² avec l'EQR-IBGN$_{moyen}$ sera retenu comme l'indicateur « influent » (calculé sur l'emprise « influente ») pour cette source de pression à cette échelle. C'est lui qui sera utilisé dans les modèles pressions/impacts régionalisés.

Lorsque plusieurs indicateurs ont un R² équivalent, l'indicateur calculé sur l'emprise la plus proche de la station est sélectionné. A l'échelle du CRIocal par exemple, l'emprise la plus proche de la station correspond à l'emprise avec la plus petite longueur et largeur de buffer.

On privilégie ainsi, à capacité explicative équivalente, la sélection d'emprises de taille minimum ce qui pertinent dans une optique d'aménagement du territoire (pour un type d'action donné, quelle est l'emprise spatiale minimum d'action permettant d'améliorer de façon significative l'état écologique des cours d'eau ?).

CemOA : archive ouverte d'Irstea / Cemagref

III.3. Méthode d'analyse des relations pressions/impacts (2ème étape)

III.3.1. Choix de la méthode statistique

L'objectif de cette méthode est (i) de quantifier la part de la variabilité de l'état écologique observé aux stations qui peut être expliquée par une combinaison des sources de pressions aux différentes échelles, (ii) d'identifier cette combinaison et (iii) de l'interpréter en terme de hiérarchisation de l'impact relatif de chacune des sources de pressions, afin de donner des conseils clairs aux gestionnaires sur les stratégies de restauration envisageables.

De façon à ne pas complexifier la méthode statistique, on fait dans cette étude l'hypothèse que les sources de pression agissent et se combinent de façon linéaire. On sait pertinemment que la réponse biologique ne répond pas de façon linéaire en fonction de l'intensité de la source de pressions (Allan, 2004a) et que les effets induits par les différentes sources de pression ne s'ajoutent pas linéairement (Décamps, 2006), toutefois cette linéarisation du modèle constitue une première approximation des relations pressions/impacts autour d'une situation moyenne. On recherche donc la combinaison « linéaire » des sources de pressions aux différentes échelles possédant une influence significative sur (une corrélation avec) la réponse biologique.

La Régression Linéaire Multiple (RLM) est la méthode d'analyse des relations la plus fréquemment utilisée en écologie (Graham, 2003) pour identifier l'effet combiné des facteurs abiotiques sur une variable biologique. Elle rend compte des relations linéaires simples existant entre une variable dépendante ou à expliquer (dans notre cas la réponse biologique) et un ensemble de prédicteurs ou variables explicatives (dans notre cas les indicateurs de pressions d'occupation du sol aux différentes échelles).

Supposons que l'on a *p* variables explicatives: X_1, X_2,…, X_p et une variable dépendante Y :

- On note *n* le nombre d'individus statistiques considérés, y^i la $i^{ème}$ observation de la variable Y et x^i_j celle de la variable X_j. Pour simplifier, nous supposons dans toute la suite que ces variables sont centrées et réduites. Le modèle linéaire est défini alors par l'équation matricielle ci-dessous :

$$Y = X\beta + \varepsilon \quad (5)$$

dans laquelle :

- **Y** est le vecteur $(y_1, y_2, ..., y_n)^t$ des *n* valeurs observées de la variable expliquée Y ;
- **X** est la matrice des données à *n* lignes et *p* colonnes, la colonne j (de 1 à *p*) étant définie par le vecteur $(x^j_1, x^j_2, ..., x^j_p)^t$;
- $\beta = (\beta_1, \beta_2, ..., \beta_p)^t$ est le vecteur des coefficients de régression ;
- ε est le vecteur résiduel $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_p)^t$ défini par un échantillon indépendant de la variable résiduelle ε de variance σ 2. on suppose fréquemment que la variable ε suit la loi normale $N(0,\sigma^2)$.

- On note **R** la matrice de corrélation entre les *p* variables X_1, …, X_p. On la suppose inversible (de rang *p*): elle possède donc *p* valeurs propres strictement positives. La méthode classique d'estimation des paramètres est fondée sur le critère des moindres carrés ordinaires (MCO).

L'estimateur **B** = $(B_1, B_2, ..., Bp)^t$ de β est alors donné par la formule ci-dessous :

$$B = R^{-1} X^T Y \quad (6)$$

CemOA : archive ouverte d'Irstea / Cemagref

Et admet comme variance :

$$(7) \qquad V_B = \frac{\sigma^2}{n} R^{-1}$$

où $\sigma^2 = \frac{\sum_i (y_i - \hat{y}_i)^2}{n - p}^{-1}$

De manière très générale, chaque coefficient de régression β_j est interprétable comme la variation relative de Y rapportée à celle de la variable explicative x_j, si et seulement si toutes les autres variables restent constantes.

Par contre, si les variables explicatives présentent des liaisons ou sont linéairement dépendantes (« colinéaires »), les coefficients de régression ne sont plus interprétables indépendamment les uns des autres, puisqu'on ne peut plus concevoir que l'une varie sans que les autres varient.

Or, dans notre étude on est dans une situation où les variables explicatives sont colinéaires. Cette colinéarité est liée (i) à la nature des indicateurs et (ii) le fait qu'on travaille à l'échelle régionale.

(i) les pourcentages d'occupation du sol sur une emprise spatiale donnée (e.g., bassin versant amont) ne sont pas des prédicteurs « indépendants » puisque leur somme pour les différents types d'occupation du sol vaut 1 : l'augmentation du pourcentage d'une classe entraîne nécessairement la diminution du pourcentage d'une ou de plusieurs autres classes (Van Sickle, 2003).

(ii) l'occupation du sol, soumis aux mêmes usages et aux mêmes aménagements régionaux, possède une structure uniforme et la présence de certains types d'occupation est liée à la présence d'autres. Les classes d'occupation du sol sont spatialement auto-corrélées (King *et al.*, 2005).

De plus cette colinéarité est accentuée par le fait que le modèle est conçu de façon non explicite au niveau spatial. La position des stations de mesure écologique sur le réseau hydrographique n'est pas prise en compte par le modèle. Ainsi des points de mesures assez proches appartenant au même réseau hydrographique auront des variables d'occupation du sol à l'échelle du BV et du CRamont très similaires.

Nous pouvons illustrer, à titre d'exemple, la colinéarité existante dans le jeu de prédicteurs par l'analyse de la corrélation (au sens de *Pearson*) entre deux couples d'indicateurs :

- I-BV-111-CLC / I-BV-112-CLC => r =0.71

Cet exemple illustre les corrélations qui peuvent exister entre certains postes d'occupation du sol dans le paysage. Pour ce couple, on observe une forte corrélation à l'échelle du bassin versant des pressions urbaines continues et discontinues quantifiées avec Corine Land Cover. Cette corrélation peut s'expliquer par le développement systématique des tissus urbains discontinus aux abords des grandes villes. Le même raisonnement est valable aux autres échelles (CRamont et CRlocal).

- I-BV-112-CLC / I-CR-20-2000-112-THRS => r =0.38

Cette corrélation montre le lien relativement significatif même lorsqu'on change d'échelle. L'occupation du sol du bassin (ici le tissu urbain discontinu) influe sur celle du corridor rivulaire local (ici une bande de 2000 m de long et de 40 m de large).

La colinéarité des variables explicatives (ou « quasi-colinéarité ») est un phénomène très bien reconnu aujourd'hui dans le domaine de la statistique (e.g., Cramer, 1985 ; Foucart, 2006) et intégré dans le domaine de l'écologie (Graham, 2003).

D'un point de vue mathématique, la défaillance de la RLM (basée sur les MCO) est purement analytique (Foucart, 2006) : lorsque qu'on se retrouve dans le cas de colinéarité entre les variables explicatives, la matrice de corrélation R est presque non inversible (det R ≈ 0) et donc la variance des estimateurs des coefficients de régression, V_B, devient très grande (cf. équation(7)). Les signes des coefficients estimés peuvent alors être contraires à ceux des vraies valeurs et leur amplitude erronée. Le coefficient de détermination R^2 peut aussi

CemOA : archive ouverte d'Irstea / Cemagref

devenir très instable (Foucart, 2000). Même une très faible colinéarité) entre les variables explicatives provoque des résultats peu interprétables (Graham, 2003).

Ainsi, pour le praticien, l'interprétation des coefficients issus de la RLM utilisant les MCO devient aussi délicate que dangereuse. Par exemple, le fait que le signe d'un coefficient de régression puisse être changé par la colinéarité peut être particulièrement gênant pour étudier l'effet propre d'une variable xj sur Y.

Cependant, on trouve dans la littérature, plusieurs approches permettant de contourner ce problème de colinéarité. De nombreux ouvrages, articles et mémoire d'études présentent ces approches. On peut classer ces approches en trois catégories : (1) les régressions pas à pas ; (2) les régressions Ridge ; et (3) les régressions basée sur les méthodes factorielles.

Les régressions « pas à pas » (Hocking, 1976) :

L'idée de ces méthodes est de limiter le nombre de variables explicatives en les sélectionnant ou les éliminant suivant leurs coefficients de corrélation partielle[46] avec la variable expliquée. Elles permettent ainsi de réduire les colinéarités éventuelles.

Cependant, dans ces approches « pas à pas », seules les variables les plus influentes (selon une hypothèse de linéraité) seront sélectionnées. L'absence d'une estimation de la contribution des prédicteurs moins influents pour expliquer la variable Y peut être non satisfaisante pour le praticien. Dans notre cas, il est important d'estimer l'influence relative de l'ensemble des facteurs significatifs : l'identification de facteurs de dégradation ou d'amélioration est primordiale même s'ils ont une influence relativement faible, surtout si ces facteurs concernent l'espace à proximité du cours d'eau (e.g., présence de ripisylves). Ces approches ne répondent donc pas entièrement aux objectifs de modélisation.

Les régressions Ridge (Hoerl and Kennard, 1970) :

L'idée de cette méthode est de contourner l'obstacle de la colinéarité entre variables explicatives en renonçant à la méthode des MCO pour estimer les paramètres du modèle. Elle modifie la matrice de covariance (elle introduit un biais contrôlé) pour rendre possible le calcul de son déterminant et de son inversion à l'aide du paramètre « ridge ». La valeur affectée à ce paramètre est définie par l'analyste qui cherche le meilleur compromis « biais-variance ». Si cette valeur est trop petite, la régression Ridge est trop semblable à la régression par les MCO, et ne peut donc lutter efficacement contre la colinéarité (forte variance et perte d'interprétabilité des coefficients de régressions). Au contraire, si elle est trop grande, la variance des paramètres du modèle sera fortement diminuée (donc interprétable) mais le biais des prédictions du modèle sera trop important. Le choix du paramètre « ridge » est donc très important. Cette méthode permet donc de régresser (d'obtenir par régression) la variable dépendante en prenant en compte l'ensemble des prédicteurs même en cas de colinéarité.

Mais, cette approche possède deux inconvénients. Le premier réside dans le choix de la valeur optimale du paramètre « ridge », que la théorie seule ne peut déterminer efficacement (Foucart, 2000). Malgré, plusieurs propositions de méthode automatique pour sa détermination, il reste généralement déterminé empiriquement à l'aide de la représentation graphique «ridge trace » (coefficients de régression en fonction du paramètre « ridge »). Le second inconvénient résulte de la complexité de l'implémentation de l'algorithme et de la lourdeur des calculs (diagonalisation et inversion des matrices).

Les méthodes factorielles :

L'idée de ces méthodes est de réduire le nombre de prédicteurs. On construit un petit nombre de variables synthétiques (variables latentes ou composantes principales) non corrélées entre elles qui sont des combinaisons linéaires de tout les prédicteurs. On élimine ainsi le problème de colinéarité dans le jeu de données des prédicteurs. On distingue deux méthodes factorielles :

- La PCR (Principal Component Regression) (Massy, 1965) :

Cette méthode consiste à régresser de façon ascendante la variable dépendante sur les composantes principales résultant d'une ACP de tous les prédicteurs.

[46] Le coefficient de corrélation partielle permet d'évaluer directement la proportion de la variation non expliquée de Y qui devient expliquée grâce à l'ajout de la variable X_j

CemOA : archive ouverte d'Irstea / Cemagref

Elle commence par la réalisation d'une ACP d'ordre *k* des p prédicteurs X qui va donner C_k composantes principales non corrélées (orthogonales). On choisit comme variables explicatives les composantes principales. On élimine ensuite les composantes principales dont (1) la corrélation avec la variable dépendante Y est faible (pour réduire le nombre de variables explicatives comme dans une approche « pas à pas ») et (2) la variance est faible (composantes principales considérés comme instable).

Cependant, la PCR possède deux inconvénients majeurs. Premièrement, le choix de *k* n'est pas évident. Deuxièmement, les composantes principales issues de l'ACP des prédicteurs ne sont pas forcément les plus corrélées avec la variable dépendante. Cette méthode n'est pas optimisée pour trouver la meilleure combinaison linéaire des X pour expliquer la distribution des valeurs de Y.

- La PLS (Partial Least Square) (Wold, 1966 ; Geladi and Kowalski, 1986 ; Tenenhaus, 1998):

Comme la PCR, la PLS est une méthode de régression ascendante. Mais au lieu de décomposer d'abord la matrice X en un ensemble de composantes principales, et de les régresser contre Y dans une étape séparée, la régression PLS utilise l'information de Y en même temps que le processus de décomposition.

On commence d'abord par extraire la première variable latente (t_1) qui résulte d'un compromis entre maximiser la corrélation entre t_1 et Y (régression ordinaire) et maximiser la variance de t_1 (ACP des prédicteurs). Ensuite, on régresse Y en fonction de t_1. Le résidu obtenu de cette première régression est utilisé par la suite comme une nouvelle variable dépendante qu'on régresse par une variable latente t_2 construite de la même façon que t_1. Cette procédure itérative se poursuit jusqu'à la régression de Y sur t_1, t_2, …, t_p où *p* est égal au nombre de prédicteurs.

L'idée principale de PLS est de donner le plus d'information possible sur Y dans les premières variables latentes construites. Elle demande moins de composantes que la PCR pour donner une bonne description et prédiction de Y. Le nombre de variables latentes retenu dans le modèle PLS est déterminé généralement par validation croisée.

La PLS est donc une méthode de régression efficace même dans le cas de colinéarités entre les variables explicatives. Le signe et la valeur des coefficients de régression du modèle peuvent être interprétés de façon fiable pour quantifier l'impact relatif des différents prédicteurs sur la variable indépendante, objectif de modélisation.

Un des avantages notables de cette méthode est la simplicité de son algorithme (comparé aux régressions « Ridge »). Ni inversion, ni diagonalisation de matrices, mais seulement une succession de régressions simples, autrement dit des calculs de produits scalaires.

La facilité d'implémentation de son algorithme a rendu cette méthode très populaire en chimie (e.g., Wold *et al.*, 2001), et elle commence à faire son apparition dans le domaine de l'écologie (e.g., Wasson *et al.*, 2005 ; Carrascal and Seoane, 2009)

Ainsi on utilisera la méthode de régression PLS pour étudier l'impact relatif des pressions d'occupation du sol sur l'état écologique des cours d'eau (EQR-$IBGN_{moyen}$), compte tenu de la colinéarité et de surparamétrisation existantes dans nos jeux de données explicatifs résultant des différentes hypothèses de construction de modèle.

On présente à la suite de cette partie l'algorithme de régression PLS utilisé dans cette étude (cf. III.3.2) puis la méthode mise en œuvre pour obtenir à partir de cette technique un modèle pressions/impacts significatif (cf. III.3.3).

III.3.2. Présentation de l'algorithme de régression PLS

Comme on l'a vu précédemment, la régression PLS est une méthode spécifiquement construite pour gérer les problèmes de colinéarité et de surparamétrisation (nombre de prédicteurs candidats largement supérieur au nombre d'observations). Elle a été proposée initialement par Wold, 1966 et Wold, 1982. Il existe de nombreuses versions de l'algorithme de régression PLS univarié. Elles découlent toutes de l'algorithme NIPALS[47] mais diffèrent au niveau des normalisations et des calculs intermédiaires. Elles aboutissent toutes à la même régression (Tenenhaus, 1998).

[47] Les principales de l'algorithme NIPALS (Nonlinear estimation by Iterative Partial Lesat Squares), présenté par Wold (1966), sont à la base de la régression PLS.

CemOA : archive ouverte d'Irstea / Cemagref

Dans cette étude, nous avons utilisé la version proposée par le logiciel XLSTAT® (2007). L'algorithme est présenté ci-dessous :

On cherche à réaliser une régression d'une variable à expliquer y sur des prédicteurs x_1, x_2, .., x_p qui peuvent être fortement corrélés entre eux ou être plus nombreux que le nombre d'observations. Par ailleurs, les coefficients doivent être interprétables dans le sens où l'on doit pouvoir mesurer la contribution de la variable x_j à la construction de la variable y à l'aide du coefficient de régression.

- *On construit d'abord une composante* $t_1 = w_{11}x_1 + \ldots + w_{1p}x_p$ *où* $w_{1j} = \frac{\mathrm{cov}(x_j, y)}{\sqrt{\sum_{j=1}^{p} \mathrm{cov}^2(x_j, y)}}$, *ce qui revient à chercher une* ***combinaison linéaire des prédicteurs maximisant la covariance entre les prédicteurs et la variable dépendante***.

- *On effectue ensuite* ***une régression simple de y sur t_1*** $y = c_1 t_1 + y_1$ *où c1 est le coefficient de régression et y_1 le vecteur des résidus d'où une première équation de régression* $y = c_1 w_{11} x_1 + \ldots + c_1 w_{1p} x_p + y_1$ *dont les coefficients sont très faciles à interpréter.*

- *On construit ensuite de la même façon une* ***composante t_2****, combinaison linéaire des x_p non corrélée à t_1 et expliquant bien le vecteur des résidus y_1 :* $t_2 = w_{21}x_{11} + \ldots + w_{2p}x_{1p}$ *où* $w_{2j} = \frac{\mathrm{cov}(x_{1j}, y_1)}{\sqrt{\sum_{j=1}^{p} \mathrm{cov}^2(x_{1j}, y_1)}}$

- *On effectue ensuite* ***une régression de y sur t_1 et t_2:*** $y = c_1 t_1 + c_2 t_2 + y_2$ *ce qui permet d'obtenir une équation de régression plus précise que la première.*

- *Cette procédure itérative se poursuit jusqu'à la* ***régression de y sur t_1, t_2, …, t_H*** *où H est égal au nombre de prédicteurs.*

- ***Le nombre de composantes à retenir est ensuite déterminé par validation croisée*** *selon un critère de minimisation de l'erreur, le* ***PRESS*** *(PRédiction Error Sum of Squares) qui correspond à la somme de tous les carrés des erreurs de prévision calculées sur les jeux tests.*

On obtient ainsi une équation de régression comparable à l'équation du modèle de régression linéaire classique. Cependant, le processus de sélection des composantes t implique quelques différences :

- les coefficients des prédicteurs sont interprétables même en cas de corrélation forte entre les prédicteurs,
- Ils sont comparables (signe et amplitude) avec le coefficient de corrélation simple entre un prédicteur et la variable à expliquer,
- une partie seulement de la variabilité des prédicteurs est utilisée pour construire cette équation de régression.

L'équation finale s'interprète enfin comme une équation de régression linéaire, les coefficients reflètent bien l'effet d'un prédicteur relativement aux autres et le R^2 est un bon estimateur de l'efficacité du modèle.

CemOA : archive ouverte d'Irstea / Cemagref

III.3.3. Construction d'un modèle PLS

Les estimations des coefficients de régression fournies par l'étude de l'échantillon ne sont interprétables qu'accompagnées d'indications quantitatives fixant le degré de confiance que l'on peut leur accorder.

Par convention, un résultat est dit « significatif » lorsqu'il n'y a pas plus de 5 chances sur 100 qu'il ait été produit par les fluctuations du hasard. Un coefficient est donc jugé significatif lorsque sa valeur ne risque pas d'être nulle ou de changer de signe selon un seuil de significativité de 5 %, autrement dit, lorsque son intervalle de confiance (IC) à 95 % ne recouvre pas la valeur 0.

Une méthode de construction a donc été mise en place pour obtenir un modèle significatif, i.e., un modèle constitué uniquement des prédicteurs possédant des coefficients de régression significatifs.

On utilise généralement la méthode de ré-échantillonnage jackknife pour estimer ces IC. Cette méthode a été recommandée par Wold, 1982 dans ses travaux initiaux sur la PLS et par Martens and Martens, 2000. Elle a été introduite par Quenouille, 1949, Tukey, 1958, et consiste à écarter chaque observation tour à tour (échantillonnage sans remise) et de recalculer la statistique (ici les coefficients de régression) afin d'estimer l'intervalle de confiance des coefficients. Ce processus est répété n fois, avec n le nombre d'individus dans l'échantillon étudié.

L'obtention d'un modèle PLS avec des coefficients de régression significatifs (avec un seuil de significativité de 5 %) se fait nécessairement selon une méthode de construction itérative. On construit dans un premier temps un modèle (Mod_1) à partir de l'ensemble des prédicteurs. Puis, dans un deuxième temps, un modèle (Mod_2) est construit en intégrant uniquement les variables significatives du Mod_1. Le résultat du Mod_2 est bien évidemment différent du résultat du Mod1 car ses composantes (variables latentes) ont été déterminées avec un jeu de prédicteurs différents (en nombre) de celui du Mod1. Dans ce nouveau modèle, certain prédicteurs peuvent alors s'avérer non significatifs. Ce processus est donc répété jusqu'à ce que l'ensemble des variables significatives du Mod_{i-1}, soient toutes significatives dans le Mod_i (modèle final).

Cependant, si on construit dès le départ un modèle avec un seuil de significativité à 5 %, on risque d'éliminer un grand nombre de variables qui sont non significatives dans ce modèle construit à partir de variables latentes se basant sur l'ensemble des prédicteurs, mais qui peuvent s'avérer significatives dans le modèle final (construit selon des variables latentes se basant sur moins de prédicteurs).

Une méthode de construction itérative et progressive (en terme de significativité) a donc été mise en place. Le fait de diminuer progressivement le seuil de significativité permet de diminuer le risque de perdre des variables s'avérant significatives dans le modèle final.

Cette méthode a été automatisée, elle comporte au minimum 5 itérations :

- **1ère itération** : un modèle PLS (Mod_1) est construit avec l'ensemble des prédicteurs avec une estimation des IC des coefficients de régression selon un seuil de significativité de 50 % (1 chance sur 2 que le résultat ait été produit par les fluctuations du hasard) ;
- **2ème itération** : un modèle PLS (Mod_2) est construit avec les prédicteurs jugés significatifs dans le Mod_1 avec une estimation des IC des coefficients de régression selon un seuil de significativité de 33 % (1 chance sur 3 que le résultat ait été produit par les fluctuations du hasard) **;**
- **3ème itération** : un modèle PLS (Mod_3) est construit avec les prédicteurs jugés significatifs dans le Mod_2 avec une estimation des IC des coefficients de régression selon un seuil de significativité de 10 % (1 chance sur 10 que le résultat ait été produit par les fluctuations du hasard) ;
- **4ème itération** : un modèle PLS (Mod_4) est construit avec les prédicteurs jugés significatifs dans le Mod_3 avec une estimation des IC des coefficients de régression selon un seuil de significativité de 5 % (1 chance sur 5 que le résultat ait été produit par les fluctuations du hasard).
- **5ème itération** : un modèle PLS (Mod_5) est construit avec les prédicteurs jugés significatifs dans le Mod_4 avec une estimation des IC des coefficients de régression selon un seuil de significativité de 5 % (1 chance sur 5 que le résultat ait été produit par les fluctuations du hasard). Si toutes les variables significatives du Mod_4 restent significatives dans le Mod_5 alors le processus s'arrête. Sinon une ou plusieurs $n^{ème}$ itération supplémentaires, avec une estimation des IC des coefficients de régression selon un seuil de significativité de 5 %, sont reproduites jusqu'à obtenir le modèle finale.

CemOA : archive ouverte d'Irstea / Cemagref

IV. Démarche de modélisation

IV.1. Une démarche en deux phases

La démarche de modélisation est conçue pour permettre de répondre aux deux objectifs définis dans cette étude (cf. I.3).

Le premier objectif vise à confirmer, et quantifier, l'apport respectif des deux voies d'amélioration mises en place dans cette thèse :

- (i) Est-ce que le mode de spatialisation « multiple » des sources de pressions au niveau du corridor rivulaire permet d'améliorer (comparé au mode « fixe ») l'analyse des relations pressions/impacts ?
- (ii) Est-ce que l'utilisation de données d'occupation du sol issues de l'imagerie THRS pour calculer les indicateurs de pression au niveau du corridor rivulaire permet d'améliorer (comparé aux indicateurs issus de CLC) l'analyse des relations pressions/impacts ?

Le deuxième objectif est d'apporter aux gestionnaires des éléments permettant d'établir des priorités d'action sur la zone d'étude en identifiant, hiérarchisant et localisant les principales sources de pressions notamment au niveau du corridor rivulaire.

Le premier objectif va nous amener à comparer entre eux des modèles construits selon différents modes de spatialisation et avec différentes sources de données d'occupation des sols. Une fois identifié le modèle le plus explicatif, nous pourrons apporter les éléments de réponse nécessaires à l'objectif 2.

A partir d'un même réseau de stations écologiques, plusieurs modèles peuvent être construits en fonction :

(1) de l'échelle ou de la combinaison d'échelles considérée ;

Dans cette étude, on a choisi d'explorer l'utilisation des trois échelles couramment employées dans l'étude des relations pressions/impacts (cf. Chapitre 3 :III.2.1.3): le bassin versant (BV) ; le corridor rivulaire amont (CRamont) et le corridor rivulaire local (CRlocal). Par conséquent, 7 modèles d'échelle et de combinaison d'échelle sont possibles pour étudier ces relations : 3 modèles « mono-échelle » (BV, CRamont, CRlocal), 3 modèles « bi-échelle » (BV/CRamont, BV/CRlocal, CRamont/CRlocal), et 1 modèle « tri-échelle » (BV/CRamont/CRlocal).

(2) de la donnée d'occupation du sol utilisée pour quantifier les indicateurs de pressions;

Deux sources de données d'occupation du sol dans le corridor rivulaire sont disponibles sur la zone d'étude pour quantifier les indicateurs de pressions liés à l'occupation du sol (THRS ou CLC cf. II.2.1.3). Par conséquent, 2 voies sont possibles pour construire les modèles utilisant l'échelle du CRamont et/ou du CRlocal.

(3) du mode de spatialisation des sources de pressions au niveau du corridor rivulaire.

Deux modes de spatialisation ont été mis en place à l'échelle du CRlocal (le mode « fixe » et le mode « multiple » cf. II.2.1.5.2.4). Par conséquent, 2 voies sont possibles pour construire les modèles utilisant l'échelle du CRlocal.

Au final, 21 modèles sont donc envisageables (Tableau 9-2) pour étudier les relations pressions/impact sur la zone d'étude.

De manière à répondre aux deux objectifs cités plus haut, une démarche en deux phases a été mise en place :

- La phase 1 (cf. IV.2) ou seront construits les 21 modèles correspondant aux différentes échelles, différentes données source d'occupation des sols ; différents modes de spatialisation des emprises de pressions;

- La phase 2 (cf. IV.3) qui comparera les résultats de ces modèles afin de répondre aux différents objectifs de l'étude, tant méthodologiques (objectif 1) que thématiques (objectif 2).

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 9-2 : Présentation des hypothèses de construction des 21 modèles pressions/impacts. Ces hypothèses diffèrent en fonction de l'échelle(s) d'analyse ; de la source de donnée d'occupation du sol (CLC ou THRS) et du mode de spatialisation au niveau du corridor rivulaire. Ce tableau indique au final le nombre de prédicteurs (d'indicateurs d'occupation du sol) qui seront utilisés par chacun des modèles.

	Echelle(s) d'analyse	Identifiant du modèle	Mode de spatialisation dans le CR	Sources d'OS dans le CR	Nombres de prédicteurs
1. Modèle mono-échelle	BV	Mod-1.1.1			44
	CRamont	Mod-1.2.1	multiple	CLC	660
		Mod-1.2.2	multiple	THRS	780
	CRlocal	Mod-1.3.1	fixe	CLC	44
		Mod-1.3.2	fixe	THRS	52
		Mod-1.3.3	multiple	CLC	6600
		Mod-1.3.4	multiple	THRS	7800
2. Modèle bi-échelle	BV CRamont	Mod-2.1.1	multiple	CLC	704
		Mod-2.1.2	multiple	THRS	824
	BV CRlocal	Mod-2.2.1	fixe	CLC	88
		Mod-2.2.2	fixe	THRS	96
		Mod-2.2.3	multiple	CLC	6644
		Mod-2.2.4	multiple	THRS	7852
	CRamont CRlocal	Mod-2.3.1	multiple (CRamont) fixe (CRlocal)	CLC	704
		Mod-2.3.2	multiple (CRamont) fixe (CRlocal)	THRS	832
		Mod-2.3.3	multiple	CLC	7260
		Mod-2.3.4	multiple	THRS	8580
3. Modèle tri-échelle	BV CRamont CRlocal	Mod-3.1.1	multiple (CRamont) fixe (CRlocal)	CLC	748
		Mod-3.1.2	multiple (CRamont) fixe (CRlocal)	THRS	876
		Mod-3.1.3	multiple	CLC	7304
		Mod-3.1.4	multiple	THRS	8624

CemOA : archive ouverte d'Irstea / Cemagref

IV.2.Phase 1 : Construction des différents modèles pressions/impacts

Dans cette phase, les 21 modèles pressions/impacts qui diffèrent par l'échelle(s) d'analyse, et/ou par la source de donnée d'occupation du sol au niveau du corridor rivulaire (CLC ou THRS) et/ou par le mode de spatialisation au niveau du CRlocal (fixe ou multiple), seront construits à partir de la méthode présentée dans la partie III. de ce chapitre. En fonction de leur hypothèse le nombre d'indicateurs (de prédicteurs) considéré n'est pas le même (Tableau 9-2, colonne 5).

L'étape 1 de cette méthode (sélection pour chaque type d'occupation des sols de l'emprise la plus influente) sera bien évidemment effective uniquement pour les modèles présentant plusieurs emprises pour une échelle donnée, c'est-à-dire pour les modèles utilisant le mode spatialisation « multiple » à l'échelle du corridor rivulaire.

IV.3.Phase 2 : Analyses comparatives des modèles

Cette phase compare les résultats des 21 modèles afin de répondre en deux temps aux objectifs de l'étude :
* objectif méthodologique : évaluation des apports respectifs des différentes voies d'amélioration (données THRS d'occupation du sol ; mode de spatialisation multiple au niveau du corridor rivulaire) (IV.3.2) ;
* objectif thématique : sélection du modèle optimal, quantification de l'impact relatif des différentes pressions et éléments de réponse aux objectifs opérationnels de restauration (Chapitre 9 :IV.3.1).

Au préalable on présente les critères sur lesquels ces modèles seront comparés (cf. IV.3.1).

IV.3.1. Critères de comparaison

La finalité d'un modèle pressions/impacts dans notre étude est d'établir un diagnostic de l'état écologique afin de cibler et localiser les priorités d'action. Ce diagnostic sera d'autant plus pertinent pour les gestionnaires si :

(1) le modèle explique une grande partie de la variabilité de l'état écologique ;

(2) le modèle identifie un large spectre de facteurs significatifs ;

(3) ces facteurs significatifs sont de nature gérable sur des emprises spatiales identifiées.

On utilisera donc, quelque soit l'objectif de l'analyse comparative des modèles, trois critères de comparaison :

- Le ***critère 1*** correspond à la capacité explicative du modèle. Cette capacité est évaluée à l'aide du coefficient de détermination R^2 qui traduit le pourcentage de variabilité de la réponse biologique expliqué par les indicateurs de pressions significatifs.

- Le ***critère 2*** correspond à la capacité descriptive du modèle. Cette capacité est évaluée en comptabilisant le nombre d'indicateurs significatifs identifiés par le modèle.

- Le ***critère 3*** correspond à la pertinence des facteurs pour les finalités de gestion. Cette pertinence dépend à la fois de la nature du facteur et de son emprise spatiale (e.g. les ripisylves).

CemOA : archive ouverte d'Irstea / Cemagref

IV.3.2. Évaluation des apports respectifs des différentes voies d'améliorations (analyse 1)

Cette première analyse a pour but de répondre aux objectifs méthodologiques de l'étude. Elle doit permettre de tester l'apport des deux voies d'améliorations qui ont été proposées dans les premières parties de la thèse au niveau du corridor rivulaire : (i) une donnée d'occupation du sol plus fine (donnée à THRS) et (ii) un mode de spatialisation des sources de pressions plus explicite (mode « multiple »).

Dans cette analyse, on testera d'abord l'apport du mode de spatialisation « multiple » avant de tester celui de la donnée d'occupation du sol à THRS en se basant sur les trois critères de comparaison définis précédemment (cf. IV.3.1).

Pour étudier l'apport du mode de spatialisation « multiple » (qui consiste à construire plusieurs buffers de largeur et/ou de longueur différente de façon à mieux appréhender les distances influentes des sources de pressions au cours d'eau et/ou à la station écologique), on comparera les résultats entre les mêmes types de modèles utilisant ce mode de spatialisation et le mode « fixe » (qui étudie lui les sources de pressions sur un seul et même buffer d'une largeur et/ou d'une longueur définies par l'expertise de l'analyste).

Ces deux modes de spatialisation dans le corridor rivulaire ont été mis en place uniquement à l'échelle du CRlocal. On s'intéressera donc uniquement aux modèles utilisant cette échelle (16 modèles au total) pour évaluer l'apport du mode de spatialisation « multiple » par rapport au mode « fixe ». 8 couples de modèles de même échelle ou combinaison d'échelles et de même source de donnée d'occupation du sol seront donc comparés dans cette étude (Tableau 9-3).

Pour étudier ensuite l'apport de la donnée d'occupation du sol à THRS, on comparera les résultats entre les mêmes types de modèle utilisant les indicateurs issus de cette donnée et ceux utilisant CLC (donnée plus grossière que la THRS).

Ces deux sources de donnée d'occupation du sol ont été utilisées aux échelles du corridor rivulaire. On s'intéressera uniquement aux modèles utilisant au moins une de ces échelles (20 modèles au total) pour évaluer l'apport de la donnée THRS par rapport à CLC. 10 couples de modèles de même échelle ou combinaison d'échelles et de même mode de spatialisation seront donc comparés dans cette étude (Tableau 9-4)

IV.3.1.Sélection du modèle optimal (analyse 2)

Cette deuxième analyse a pour but de sélectionner le modèle optimal afin de répondre aux objectifs opérationnels de l'étude.

Pour cela, les 21 modèles seront comparés en fonction des trois critères définis précédemment (cf. IV.3.1). Cette sélection se fera par élimination en appliquant dans l'ordre les critères de comparaison.

Le modèle sélectionné sera ensuite utilisé pour établir un diagnostic de l'état écologique sur la zone d'étude.

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 9-3 : Présentation des 8 couples de modèles dont les résultats seront comparés pour étudier l'apport du mode de spatialisation « multiple » par rapport au mode « fixe ». En gris les couples de modèles utilisant la donnée à THRS.

Echelle(s) d'analyse	Identifiants des modèles comparées	Source OS dans le CR
CRlocal	Mod-1.3.1/Mod-1.3.3	CLC
	Mod-1.3.2/Mod-1.3.4	THRS
BV CRlocal	Mod-2.2.1/Mod-2.2.3	CLC
	Mod-2.2.2/Mod-2.2.4	THRS
CRamont CRlocal	Mod-2.3.1/Mod-2.3.3	CLC
	Mod-2.3.2/Mod-2.3.4	THRS
BV CRamont CRlocal	Mod-3.1.1/Mod-3.1.3	CLC
	Mod-3.1.2/Mod-3.1.4	THRS

Tableau 9-4 : Présentation des 10 couples de modèles dont les résultats seront comparés pour étudier l'apport de la donnée à THRS par rapport à CLC. En gris les couples de modèles utilisant le mode de spatialisation « multiple » à l'échelle du CRlocal.

Echelle(s) d'analyse	Identifiants des modèles comparées	Mode de spatialisation dans le CR
CRamont	Mod-1.2.1/Mod-1.2.2	Multiple
CRlocal	Mod-1.3.1/Mod-1.3.2	Fixe
	Mod-1.3.3/Mod-1.3.4	Multiple
BV CRamont	Mod-2.1.1/Mod-2.1.2	Multiple
BV CRlocal	Mod-2.2.1/Mod-2.2.2	Fixe
	Mod-2.2.3/Mod-2.2.4	Multiple
CRamont CRlocal	Mod-2.3.1/Mod-2.3.2	multiple (CRamont) fixe (Crlocal)
	Mod-2.3.3/Mod-2.3.4	Multiple
BV CRamont CRlocal	Mod-3.1.1/Mod-3.1.2	multiple (CRamont) fixe (Crlocal)
	Mod-3.1.3/Mod-3.1.4	Multiple

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

Chapitre 10
Modèles pressions/impacts : Application sur la zone d'étude (HER9-NO)

SOMMAIRE

Résumé *330*

I. CARTOGRAPHIE DE L'ETAT ECOLOGIQUE ET DES PRESSIONS 332

I.1. ETAT ECOLOGIQUE DES STATIONS *332*

I.2. ETAT DES PRESSIONS *335*

I.2.1. Choix de la typologie d'occupation du sol et des emprises spatiales 335
I.2.1.1 Une typologie commune entre la donnée CLC et la donnée à THRS 335
I.2.1.2 Présentation des emprises spatiales analysées 335
I.2.2. Distribution spatiale de l'occupation du sol 336
I.2.2.1 En fonction de l'emprise spatiale 336
I.2.2.1.1 Avec la donnée CLC 336
I.2.2.1.2 Avec la donnée à THRS 339
I.2.2.1.3 Comparaison entre la donnée CLC et la donnée THRS 339
I.2.2.2 En fonction du positionnement des stations sur le réseau 340
I.2.2.3 Distribution géographique sur la zone d'étude 342

I.3. PREMIERES OBSERVATIONS *347*

II. RESULTATS DE LA DEMARCHE DE MODELISATION 348

II.1. PRESENTATION DES RESULTATS *348*

II.1.1. Construction des modèles (phase 1) 348
II.1.1.1 Illustration de la construction d'un modèle (Mod-3.1.4) 348
II.1.1.1.1 Etape 1 : Sélection des emprises spatiales 348
II.1.1.1.1.1 Sélection de l'emprise la plus significative pour la source de pression « prairies permanentes » 348
II.1.1.1.1.2 Emprises sélectionnées aux échelles du corridor rivulaire 349
II.1.1.1.2 Etape 2 : Construction du modèle Régression PLS 351
II.1.1.2 Tableau récapitulatif des différents modèles 351
II.1.2. Comparaison des modèles (phase 2) 353
II.1.2.1 Évaluation de l'apport respectif des différentes voies d'amélioration 353
II.1.2.1.1 Apport du mode de spatialisation « multiple » 353
II.1.2.1.2 Apport de la donnée d'occupation du sol à THRS 354
II.1.2.2 Sélection du modèle optimal 357

…SOMMAIRE (suite)

CemOA : archive ouverte d'Irstea / Cemagref

…SOMMAIRE (suite)

II.2. BILAN ET FIABILITE DES RESULTATS *359*
II.2.1. Apport des voies d'améliorations 359
II.2.2. Choix du modèle optimal 362
II.2.3. Approche de la fiabilité 363

III. ANALYSE CRITIQUE DES METHODES 365

III.1. METHODE DE CONSTRUCTION DES MODELES PRESSIONS/IMPACTS *365*
III.1.1. Sélection de l'emprise spatiale significative 365
III.1.2. Évaluation des incertitudes des résultats de modélisation 366
III.2. METHODE DE COMPARAISON DES MODELES *366*

IV. PREMIER DIAGNOSTIC DE L'ETAT ECOLOGIQUE 369

IV.1. PRESENTATION DU MODELE « OPTIMAL » *369*
IV.2. ANALYSE DES SOURCES DE PRESSIONS INFLUENÇANT L'ETAT ECOLOGIQUE *372*
IV.2.1. Les sources de pressions dégradantes 372
IV.2.1.1 L'impact de l'urbain discontinu et des zones industrielles & commerciales 372
IV.2.1.2 L'impact des routes dans le corridor rivulaire 374
IV.2.1.3 L'impact des terres arables 374
I.1.1. Les sources de pressions protectrices 377
I.1.1.1 Le rôle des prairies et de la végétation herbacée semi-naturelle 377
I.1.1.2 Le rôle des ripisylves et de la végétation arborée 378
I.1.1.3 Les effets « contradictoires » 379
I.2. IMPLICATIONS POUR LA GESTION *381*

CemOA : archive ouverte d'Irstea / Cemagref

Résumé

Ce chapitre met en œuvre la démarche présentée au chapitre précédent pour répondre aux objectifs de modélisation de l'étude. La cartographie générale de l'état écologique et des grandes sources de pressions sur la zone d'étude précise que le diagnostic portera principalement sur des cours d'eau de petites tailles et que 60 % des stations sont en bon état. Elle confirme également que la donnée CLC est limitée, contrairement à la donnée à THRS, pour caractériser précisément l'occupation du sol au niveau du corridor rivulaire. Elle permet de mettre en évidence de façon visuelle certaines relations entre l'occupation du sol et l'état écologique (EQR-IBGN$_{moyen}$).

Les résultats de la démarche de modélisation approfondissent ces premières observations. Ils confirment que l'utilisation des voies d'amélioration augmente significativement le pouvoir explicatif des modèles ainsi que le nombre de sources de pressions identifiées comme influençant la réponse biologique. Ils précisent la localisation de l'emprise significative de ces sources de pressions. Ils démontrent par exemple, l'impact limitant des routes à l'échelle du corridor local dans une bande de 100 m de large de part et d'autre du cours d'eau, sur 200 à 400 m en amont de la station, l'effet protecteur des prairies à l'échelle du corridor amont et l'effet protecteur des ripisylves (végétation arborée) à l'échelle du corridor local sur une bande de 20 m de large de part et d'autre du cours d'eau. Ces avancées fiabilisent donc la compréhension des relations pressions/impacts et apportent de nouvelles informations aux gestionnaires qui peuvent être pertinentes pour la gestion et la restauration.

Le modèle retenu pour établir un diagnostic de l'état écologique utilise ces avancées méthodologiques et étudie les impacts de l'occupation du sol sur l'ensemble des trois échelles (BV, CRamont, CRlocal). Ce modèle explique environ 38 % de la variabilité de la réponse biologique (variabilité maximale qu'on a pu expliquer avec les modèles les plus explicatifs) tout en identifiant un plus grand nombre de sources de pressions significatives que les autres modèles (23 au total).

Son interprétation confirme l'impact marqué des terres artificialisées (urbains discontinus, zones industrielles et commerciales puis routes) aussi bien à l'échelle du bassin versant qu'aux échelles du corridor rivulaire ainsi que l'impact des terres arables à l'échelle du bassin versant qui modifient les processus de fonctionnement (biogéochimiques, hydromorphologiques et écologiques) de l'écosystème rivière. L'objectif de « bon état » ne semble pas irréaliste dans ces régions très agricoles puisque le modèle met en évidence les effets protecteurs de la végétation arborée et de la végétation herbacée permanente dans une bande rivulaire gérable (20 m de part et d'autre du cours d'eau), les deux ayant une contribution positive cumulée sur l'EQR-IBGN$_{moyen}$ de même amplitude que l'effet négatif des terres arables sur le BV. Toutefois, la possibilité de gagner quelques points d'IBGN ne devra pas s'appuyer uniquement sur la restauration/préservation des espaces rivulaires (même s'ils conditionnent en partie les processus de fonctionnement). Elle devra aussi porter sur la réduction de la pollution résiduelle et toxique (notamment au niveau des déversoirs d'orages), sur la modification des pratiques agricoles à l'échelle du bassin, et sur la restauration/préservation de l'état hydromorphologique des rivières compte tenu de la nature des impacts mis en évidence (impacts des routes notamment).

L'ensemble de ces résultats doit cependant être interprété avec prudence compte tenu (i) que leur robustesse vis-à-vis de l'incertitude sur les indicateurs n'a pas été évaluée ; et (ii) que les dépendances d'échelles, permettant de vérifier l'effet propre d'un facteur à une échelle donnée sur la variabilité de la réponse biologique, et de confronter l'effet des facteurs à une échelle par rapport à une autre, n'ont pas été analysées.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

I. Cartographie de l'état écologique et des pressions

On présente tout d'abord dans cette partie une cartographie de l'état écologique et des grandes sources de pressions sur la zone d'étude. Cette cartographie générale permettra d'améliorer la compréhension des résultats de modélisation et du diagnostic de l'état écologique sur la zone d'étude.

I.1. Etat écologique des stations

Pour chaque station, on a calculé l'IBGN moyen ($IBGN_{moyen}$) gommant ainsi les sources de variabilité saisonnière de l'indice IBGN pour les sites ayant fait l'objet d'un grand nombre de mesures. Ces $IBGN_{moyen}$ ont été transformés en EQR ($EQR\text{-}IBGN_{moyen}$, cf. Chapitre 9 :II.1.1).

Chaque station a été ensuite classée à partir de l'$EQR\text{-}IBGN_{moyen}$ en se référant aux hypothèses de limites de classes d'états[48] liées aux types de cours d'eau de chaque station. D'après ces limites, 60 % des stations de la zone d'étude sont en bon état. Une cartographie de l'état des stations selon la valeur de l'$EQR\text{-}IBGN_{moyen}$ a été établie (Figure 10.1). Elle permet de visualiser des répartitions régionales de l'état écologique. On observe que les stations du centre (département du Calvados) sont majoritairement au dessus de la limite du bon état alors qu'on constate l'inverse pour les stations à l'Est (département de la Manche) et à l'Ouest (département proche de Paris) de la zone d'étude.

La majorité des stations concernent des cours de très petites (87 %) et petites tailles (11 %), seules 3 stations concernent les cours d'eau de rang 5 (Tableau 10-1). Le diagnostic de l'état écologique sur l'HER9-NO rendra compte uniquement des causes d'altération et des facteurs protecteurs pour des systèmes de petite et moyenne taille (qui représentent 94 % du réseau de drainage de la zone d'étude).

Tableau 10-1 : Répartition des stations de l'HER9-NO en fonction de la dimension du cours d'eau (rang de Strahler)

Classe de cours d'eau	Rang de Strahler	Nombre de stations	% de stations
Très petit	**1**	29	**18,71%**
	2	50	**32,26%**
	3	56	**36,13%**
Petit	**4**	17	**10,97%**
Moyen	**5**	3	**1,94%**

[48] Les hypothèses de limites de classes ainsi que les valeurs de référence sont consultables dans le guide technique actualisant les règles d'évaluation de l'état des eaux douces de surface de métropole, publié en mars 2009 par le Ministère chargé de l'environnement.

CemOA : archive ouverte d'Irstea / Cemagref

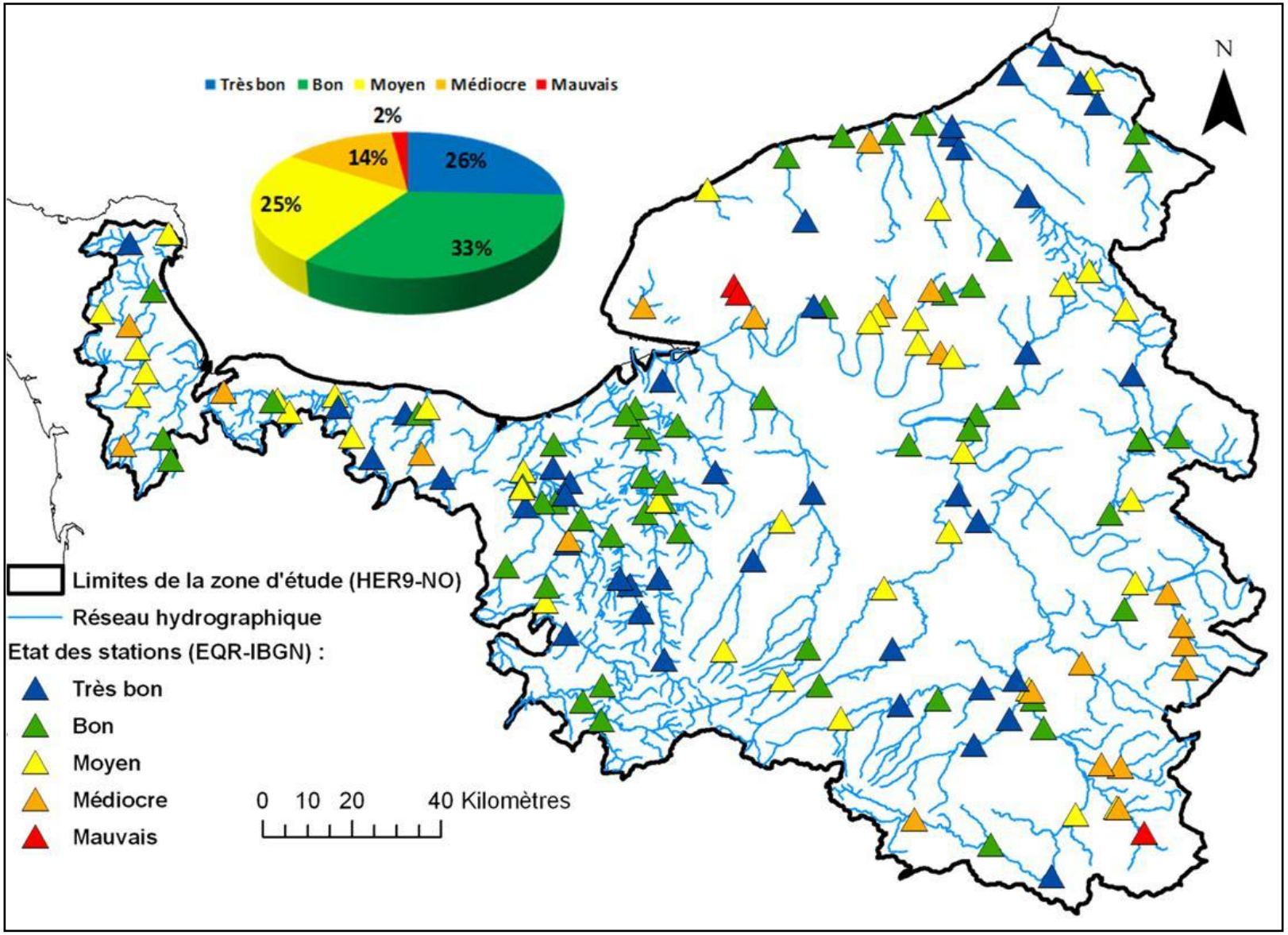


Figure 10.1 : Cartographie de l'état écologique des stations de l'HER9-NO selon l'EQR-IBGN moyenné sur la période de 1992 à 2004 en fonction des 5 modalités d'état.

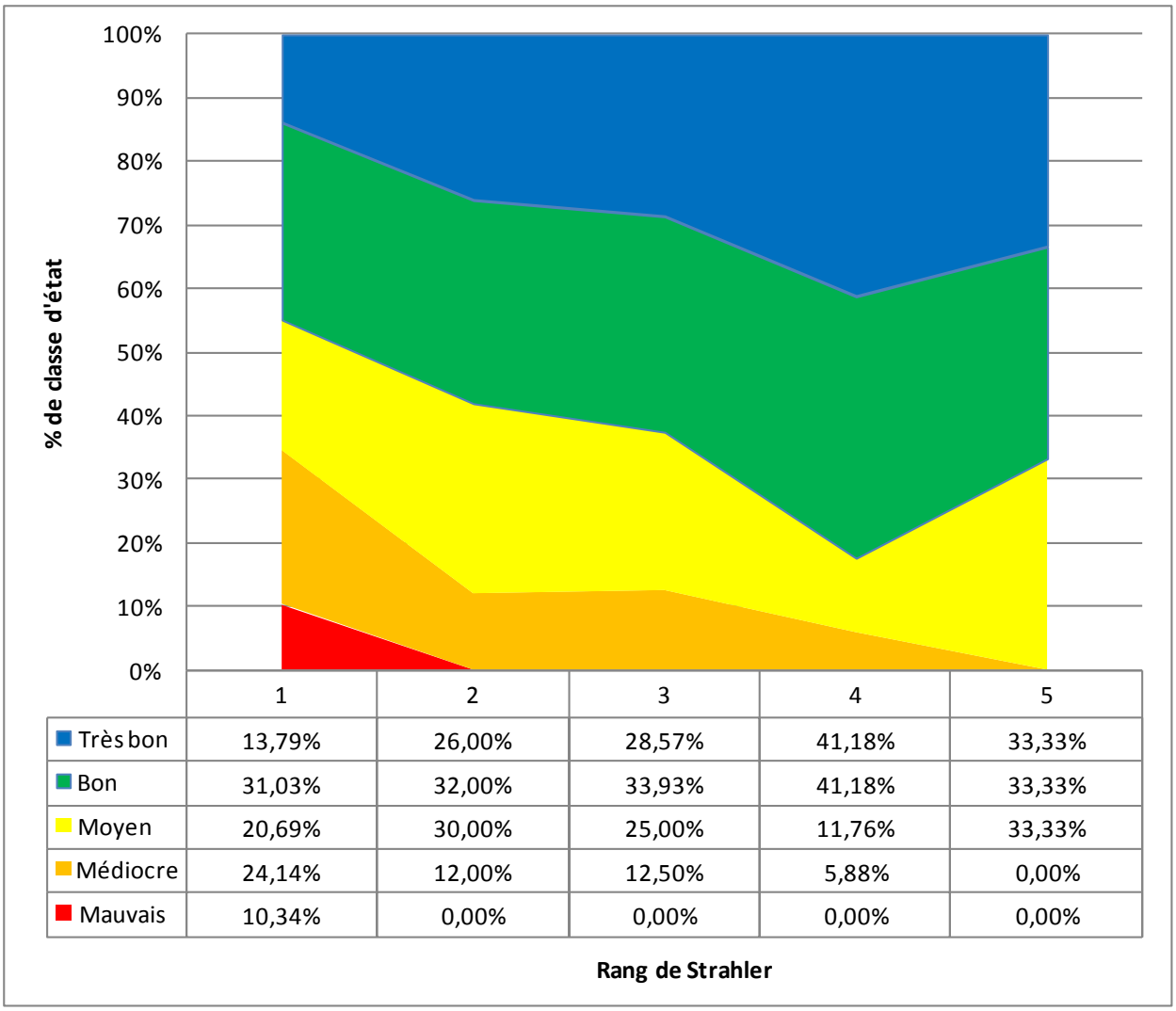
CemOA : archive ouverte d'Irstea / Cemagref

La Figure 10.2 qui illustre la variation de la qualité des stations en fonction du rang de Strahler. On remarque de manière générale que la qualité de l'état écologique (EQR-IBGN$_{moyen}$) sur la zone d'étude augmente en fonction de la taille des cours d'eau du rang 1 jusqu'au rang 4. L'évolution au rang 5 est peu fiable compte tenu du faible nombre de stations représentant cette taille de cours d'eau.

Cette observation montre :

(1) que les têtes de bassins (cours d'eau de rang 1) de la zone d'étude sont fortement impactées : 55 % des stations de cours d'eau de rang 1 sont sous la barre du bon état selon l'EQR-IBGN$_{moyen}$. 10 % d'entre eux sont même en Mauvais état (classe la plus dégradante).

(2) qu'il existe certainement, à mesure que l'on va vers l'aval et vers des tronçons d'ordre de Strahler plus élevé, des mécanismes améliorant la qualité de l'état écologique tels que, par exemple, la présence de zone tampon interceptant la pollution diffuse de ces têtes de bassin ou la présence d'espaces naturels abritant des réservoirs biologiques.



	1	2	3	4	5
Très bon	13,79%	26,00%	28,57%	41,18%	33,33%
Bon	31,03%	32,00%	33,93%	41,18%	33,33%
Moyen	20,69%	30,00%	25,00%	11,76%	33,33%
Médiocre	24,14%	12,00%	12,50%	5,88%	0,00%
Mauvais	10,34%	0,00%	0,00%	0,00%	0,00%

Figure 10.2 : Évolution de la qualité des stations, évaluée à partir de l'EQR-IBGN$_{moyen}$ (sur la période de 1992 à 2004) selon les 5 modalités d'état (Très bon, Bon, Moyen, Médiocre, Mauvais), en fonction du rang de Strahler

CemOA : archive ouverte d'Irstea / Cemagref

I.2. Etat des pressions

Au niveau du bassin versant le calcul des indicateurs de pression se fait sur la base de la carte d'occupation des sols Corine Land Cover.
Au niveau du corridor rivulaire, plusieurs possibilités, en termes de carte d'occupation des sols et d'emprises spatiales, sont envisageables pour étudier l'état des pressions sur la zone d'étude. On présente donc tout d'abord la typologie d'occupation des sols et les emprises spatiales choisies au niveau du corridor rivulaire (cf. I.2.1) avant d'analyser la distribution spatiale des différentes sources de pressions (cf. I.2.2).

I.2.1. Choix de la typologie d'occupation du sol et des emprises spatiales

I.2.1.1 Une typologie commune entre la donnée CLC et la donnée à THRS

Une typologie d'occupation du sol commune à la donnée CLC et la donnée à THRS, a été conçue de manière à pouvoir comparer l'information entre ces deux données. Cette typologie se compose de 8 catégories correspondant à une agrégation de un ou plusieurs postes d'occupation du sol (Tableau 10-2). Ces catégories ont été définies conceptuellement en fonction des pressions qu'elles sont susceptibles de générer sur l'état écologique.

Tableau 10-2 : Règles d'agrégation (pour chacune des sources de données) pour obtenir la typologie commune composée de 8 catégories.

Catégories	*Code OS (CLC)*	*Code OS (THRS)*
C1 : Espace artificialisé continu	**111 ; 121 ; 122 ; 123 ; 124 ; 131 ; 132 ; 133**	**111 ; 121 ; 122 ; 123 ; 124 ; 131 ; 133**
C2 : Espace artificialisé discontinu	**112 ; 141 ; 142**	**112 ; 113 ; 141 ; 142 ; 151 ; 152 ; 153 ; 154 ; 155**
C3 : Terres arables	**211 ; 212 ; 213 ; 221 ; 222 ; 223 ; 241 ; 242 ; 243 ; 244**	**211 ; 212 ; 213 ; 214 ; 221 ; 222 ; 223 ; 241 ; 242 ; 251 ; 252 ; 253**
C4 : Prairies agricoles	**231**	**231 ; 232 ; 233 ; 234**
C5 : Végétation herbacée et arbustive (semi-nat.)	**321 ; 322 ; 323**	**321**
C6: Végétation arborée (semi-nat.)	**311 ; 312 ; 313 ; 324**	**311**
C7 : Sol sans ou avec peu de végétation (semi-nat.)	**331 ; 332 ; 333 ; 334 ; 335**	**331**
C8 : Surface en eau	**411 ; 412 ; 421 ; 422 ; 423 ; 511 ; 512 ; 521 ; 522 ; 523**	**411 ; 412 ; 413 ; 414 ; 415 ; 416; 417 ; 418 ; 421 ; 422 ; 423 ; 424; 425 ; 426 ; 427 ; 428 ; 429**

I.2.1.2 Présentation des emprises spatiales analysées

Les sources de pressions sont analysées au niveau des trois échelles de fonctionnement de l'écosystème rivière (le BV, le CRamont et le CRlocal).

Au niveau du BV, l'emprise correspond au bassin versant amont de chaque station (cf. Chapitre 9 :II.2.1.5.1). Au niveau du corridor rivulaire, plusieurs emprises (buffers) sont envisageables pour étudier les pressions aussi bien à l'échelle du CRamont que du CRlocal. On rappelle que ces emprises sont définies par un buffer qui peut varier en largeur (distance au cours d'eau) pour le CRamont mais aussi en longueur (distance amont et/ou aval à la station) pour le CRlocal.

CemOA : archive ouverte d'Irstea / Cemagref

Trois emprises ont été sélectionnées parmi les buffers construits dans le mode de spatialisation « multiple » pour chacune des échelles (cf. Chapitre 9 :II.2.1.5.2.3 pour le CRamont et cf. Chapitre 9 :II.2.1.5.2.4.2 pour le CRlocal) : l'emprise la plus éloignée et la plus proche du cours d'eau (et/ou de la station pour le CRlocal) puis une emprise intermédiaire. Les caractéristiques de ces buffers sont présentées dans le Tableau 10-3.

Tableau 10-3 : Caractéristiques des emprises sélectionnées au niveau du corridor rivulaire pour l'analyse de l'état des pressions.

	Buffer CRamont (largeur)	**Buffer CRlocal (largeur/longueur)**
Emprise la plus éloignée	300 m	300/3000 m
Emprise intermédiaire	100 m	100/1000 m
Emprise la plus proche	5 m	5/100 m

I.2.2. Distribution spatiale de l'occupation du sol

Dans cette partie, on analyse tout d'abord la distribution de la composition de l'occupation du sol en fonction de l'emprise spatiale utilisée aux différentes échelles (cf. I.2.2.1). Cette analyse sera menée à la fois avec la donnée CLC (cf. I.2.2.1.1) et la donnée à THRS (cf. I.2.2.1.2) de façon à mettre en évidence la différence d'information entre ces deux données (cf. I.2.2.1.3). Ensuite, on étudiera cette distribution en fonction du positionnement des stations (cf. I.2.2.2) sur le réseau et de façon générale sur la zone d'étude (cf. I.2.2.3) en utilisant pour chacune des échelles du corridor rivulaire la donnée d'occupation du sol et les emprises spatiales les plus pertinentes.

I.2.2.1 En fonction de l'emprise spatiale

I.2.2.1.1 Avec la donnée CLC

La Figure 10.3 montre la composition moyenne (sur les 155 stations) de l'occupation du sol, selon la donnée CLC et en fonction de la typologie en 8 catégories, aux différentes échelles (BV, CRamont, CRlocal) sur les différentes emprises du corridor rivulaire (éloignée, intermédiaire et proche, cf. 0).

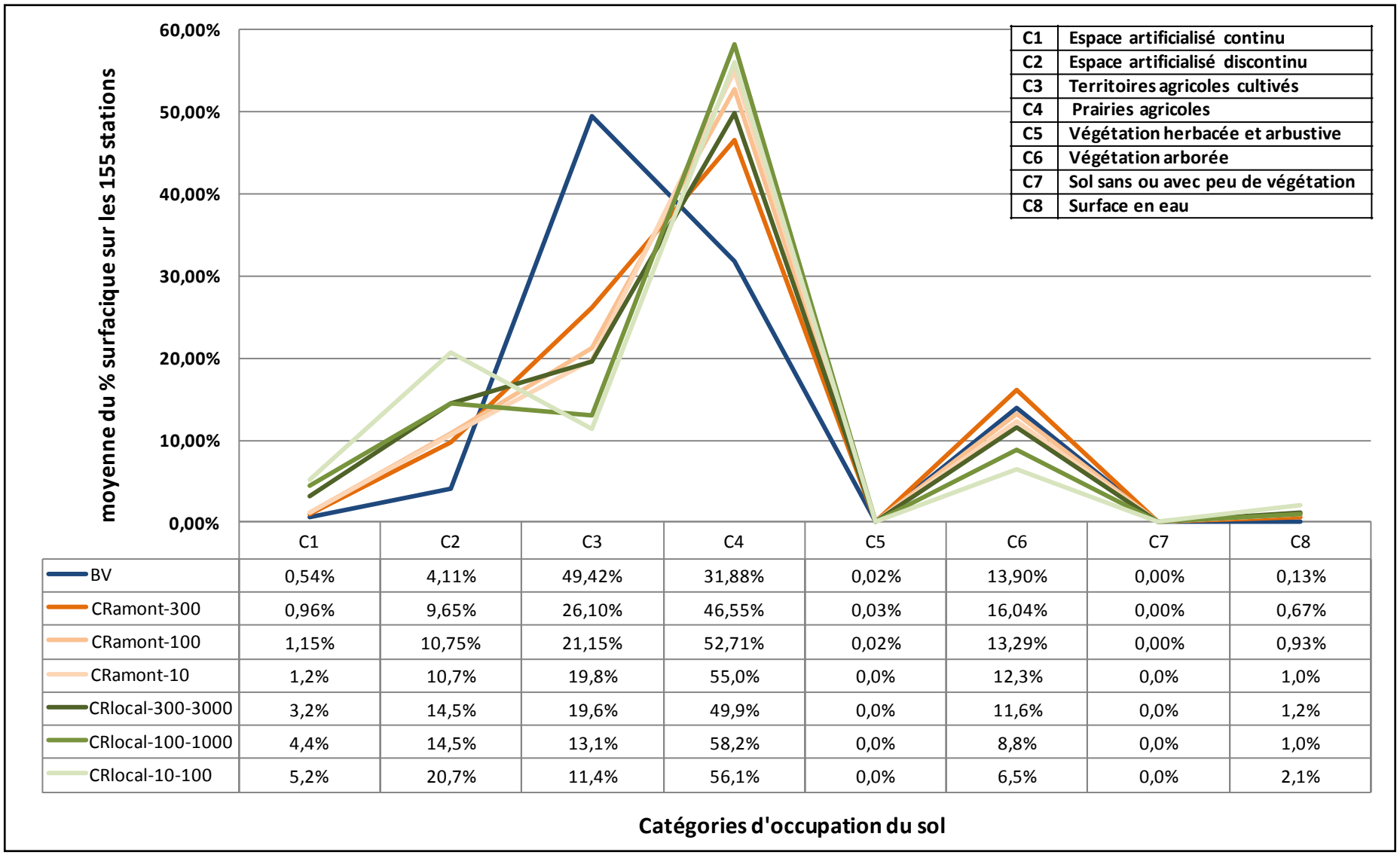
D'après cette figure, on constate tout d'abord que la composition de l'occupation du sol dans le Corridor Rivulaire (local ou global) est quasiment similaire quelque soit l'emprise considérée, avec une forte dominante en terres agricoles et prairies agricoles (classe C4). On note juste que l'emprise corridor rivulaire la plus localisée (la plus proche de la station en largeur et longueur) renforce le poids du tissu urbain discontinu (C2), alors que l'emprise la plus distante renforce le poids du territoire agricole (C3).

On remarque par ailleurs que la composition de l'occupation du sol à l'échelle du corridor rivulaire est assez proche de celle à l'échelle du BV, avec pour seule différence notable une diminution des terres arables (C3) au profit des prairies agricoles (C4) : les bassins versants amont des stations, à l'inverse des différents espaces rivulaires étudiés, sont en moyenne composés par plus de terres arables (C3) que de prairies agricoles (C4).

On observe donc d'après CLC (en moyenne sur les 155 stations de la zone d'étude):

- une composition d'occupation du sol similaire quelque soit l'emprise au niveau du corridor rivulaire ;
- la même composition de tissus urbains continus et discontinus et de végétation arborée (C1, C2 et C6) quelque soit les emprises et échelles ;
- une présence plus importante de terres arables (C3) que de prairies agricoles (C4) à l'échelle du bassin versant et un rapport inversé aux échelles du corridor rivulaire.
- les sols nus et surfaces en eau sont très faiblement représentés quelques soient les emprises.

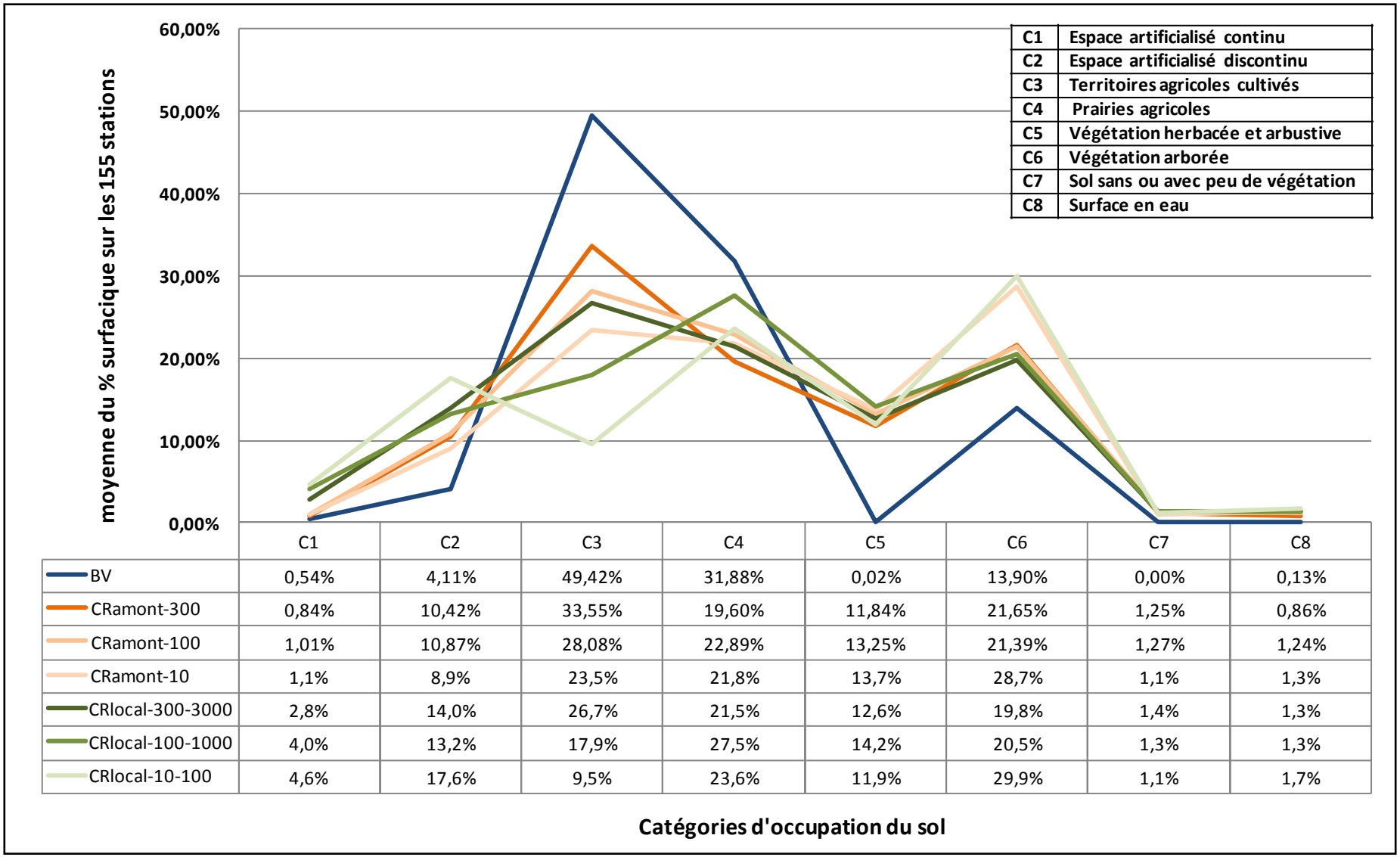
CemOA : archive ouverte d'Irstea / Cemagref



	C1	C2	C3	C4	C5	C6	C7	C8
BV	0,54%	4,11%	49,42%	31,88%	0,02%	13,90%	0,00%	0,13%
CRamont-300	0,96%	9,65%	26,10%	46,55%	0,03%	16,04%	0,00%	0,67%
CRamont-100	1,15%	10,75%	21,15%	52,71%	0,02%	13,29%	0,00%	0,93%
CRamont-10	1,2%	10,7%	19,8%	55,0%	0,0%	12,3%	0,0%	1,0%
CRlocal-300-3000	3,2%	14,5%	19,6%	49,9%	0,0%	11,6%	0,0%	1,2%
CRlocal-100-1000	4,4%	14,5%	13,1%	58,2%	0,0%	8,8%	0,0%	1,0%
CRlocal-10-100	5,2%	20,7%	11,4%	56,1%	0,0%	6,5%	0,0%	2,1%

Figure 10.3 : Composition moyenne (sur les 155 stations) de l'occupation du sol (en %) selon la donnée CLC en fonction de la typologie en 8 catégories aux différentes échelles (BV en bleu, CRamont en orange, et CRlocal en vert) et sur les différentes emprises du corridor rivulaire selon un gradient du plus foncé (emprise la plus éloigné) au plus clair (emprise la plus proche) selon la couleur de l'échelle.

CemOA : archive ouverte d'Irstea / Cemagref



	C1	C2	C3	C4	C5	C6	C7	C8
BV	0,54%	4,11%	49,42%	31,88%	0,02%	13,90%	0,00%	0,13%
CRamont-300	0,84%	10,42%	33,55%	19,60%	11,84%	21,65%	1,25%	0,86%
CRamont-100	1,01%	10,87%	28,08%	22,89%	13,25%	21,39%	1,27%	1,24%
CRamont-10	1,1%	8,9%	23,5%	21,8%	13,7%	28,7%	1,1%	1,3%
CRlocal-300-3000	2,8%	14,0%	26,7%	21,5%	12,6%	19,8%	1,4%	1,3%
CRlocal-100-1000	4,0%	13,2%	17,9%	27,5%	14,2%	20,5%	1,3%	1,3%
CRlocal-10-100	4,6%	17,6%	9,5%	23,6%	11,9%	29,9%	1,1%	1,7%

Figure 10.4 : Composition moyenne (sur les 155 stations) de l'occupation du sol (en %) selon la donnée à THRS en fonction de la typologie en 8 catégories aux différentes échelles (BV en bleu, CRamont en orange, et CRlocal en vert) et sur les différentes emprises du corridor rivulaire selon un gradient du plus foncé (emprise la plus éloigné) au plus clair (emprise la plus proche) selon la couleur de l'échelle.

CemOA : archive ouverte d'Irstea / Cemagref

I.2.2.1.2 Avec la donnée à THRS

La Figure 10.4 montre la composition moyenne (sur les 155 stations) de l'occupation du sol, selon la donnée à THRS et en fonction de la typologie en 8 catégories, aux différentes échelles (BV, CRamont, CRlocal) sur les différentes emprises du corridor rivulaire (éloignée, intermédiaire et proche, cf. 0).

D'après cette figure, on constate tout d'abord que la composition de l'occupation du sol est quasiment similaire pour l'ensemble des emprises du corridor rivulaire.

On note tout de même que l'emprise corridor rivulaire la plus localisée (la plus proche de la station en largeur et longueur) présente un pourcentage de tissu urbain discontinu (C2) plus fort et un pourcentage de terres arables (C3) plus faible que les emprises plus larges, le premier pourcentage étant supérieur au deuxième.

On note également qu'alors que les emprises larges indiquent un même pourcentage de terres arables et de prairies agricoles (~25%), les emprises les plus localisées du corridor rivulaire montrent une différenciation avec un taux de terres arables (C3) d'environ 10% et un taux de prairies agricoles (C4) d'environ 25%.

Cependant, même si cette composition présente le même profil quelque soit l'emprise, on note de grosses différences sur les pourcentages pour certaines catégories selon les emprises considérées au niveau du corridor rivulaire. C'est le cas de la catégorie terres arables (C3), pour laquelle on remarque que plus on s'approche du cours d'eau et/ou de la station plus son pourcentage diminue ; pour la catégorie végétation arborée (C6), on remarque un plus fort pourcentage (présence de ripisylves) pour les deux emprises les plus proches du cours d'eau (emprise proche à l'échelle du CRlocal et du CRamont).

Enfin et surtout, de grosses différences apparaissent entre le profil moyen de composition à l'échelle du bassin versant et les profils moyens aux différentes échelles du corridor rivulaire. On note que les pourcentages de terres arables (C3) et de prairies agricoles (C4) sont plus élevés à l'échelle du Bassin Versant, alors que les pourcentages de tissus urbains continus (C1), discontinus (C2), et de végétation semi-naturelle herbacée/arbustive (C5) et arborée (C6) sont plus élevés aux échelles du corridor rivulaire.

On observe donc d'après la donnée THRS (en moyenne sur les 155 stations de la zone d'étude):

- un profil de la composition d'occupation du sol quasiment similaire quelque soit l'emprise aux échelles du corridor rivulaire;
- des différences marquées entre le pourcentage d'occupation du sol des différentes catégories à l'échelle du bassin versant et aux échelles du corridor rivulaire (excepté pour les surfaces en eau (C8) qui sont très faiblement représentées quelque soit les emprises) ;
- une diminution de la présence des terres arables (C3) lorsqu'on s'approche du cours d'eau et ou de la station au niveau du corridor rivulaire;
- une augmentation de la présence de végétation arborée (C6) lorsqu'on s'approche du cours d'eau, ce qui indique la présence de ripisylves.

I.2.2.1.3 Comparaison entre la donnée CLC et la donnée THRS

La comparaison de la Figure 10.3 et de la Figure 10.4, montre que l'information apportée par ces deux sources de données est très significativement différente aux échelles du corridor rivulaire. Les classes C5 (végétation herbacée et arbustive) et C6 (végétation arborée), quasiment absentes de la cartographie CLC des corridors rivulaires sont fortement présentes (15% à 30%) dans la cartographie THRS, alors que la classe C4 (prairies agricoles) diminue entre CLC et THRS.

Avec la donnée CLC, l'information est similaire quelque soit l'emprise spatiale considérée au niveau du corridor rivulaire.

Par contre, avec la donnée à THRS, l'information est sensiblement différente en fonction de l'emprise spatiale, c'est le cas notamment des terres arables (C3), dont la présence diminue lorsqu'on s'approche du cours d'eau et/ou de la station, et de la végétation arborée (C6) dont la présence augmente lorsqu'on on se situe sur des espaces proches du cours d'eau (présence de ripisylves).

Cette différence prouve, comme démontré sur le système Hérault dans les parties 2 et 3 de ce mémoire, la limite de CLC du fait de son unité cartographique (25 ha minimum) et l'intérêt de la THRS pour quantifier plus précisément et de façon plus fiable les indicateurs de composition et/ou de configuration de l'occupation du sol au niveau du corridor rivulaire et les indicateurs des formations végétales le long du cours d'eau.

CemOA : archive ouverte d'Irstea / Cemagref

Elle montre aussi que lorsqu'on utilise la donnée CLC, on surestime la présence de prairies agricole (C4) au niveau du corridor rivulaire qui apparaît plus importante qu'à l'échelle du bassin versant.

On préfèrera donc utiliser, dans la suite de cette partie, la donnée THRS pour cartographier les sources de pressions aux échelles du corridor rivulaire.

I.2.2.2 En fonction du positionnement des stations sur le réseau

Nous avons jusqu'ici analysé les profils moyens de composition de l'occupation des sols pour différentes échelles et emprises, la moyenne portant sur les 155 stations. Nous allons maintenant analyser la distribution spatiale de la composition de l'occupation des sols en fonction du positionnement des stations sur le réseau. Ce positionnement est traduit non seulement en termes de coordonnées géographiques de la station (longitude, latitude) mais aussi par le rang de Strahler de la station (rang de Strahler du tronçon de cours d'eau sur lequel se trouve la station). Pour étudier cette distribution spatiale aux échelles du corridor rivulaire nous utiliserons les emprises les plus proches du cours d'eau et la donnée à THRS.

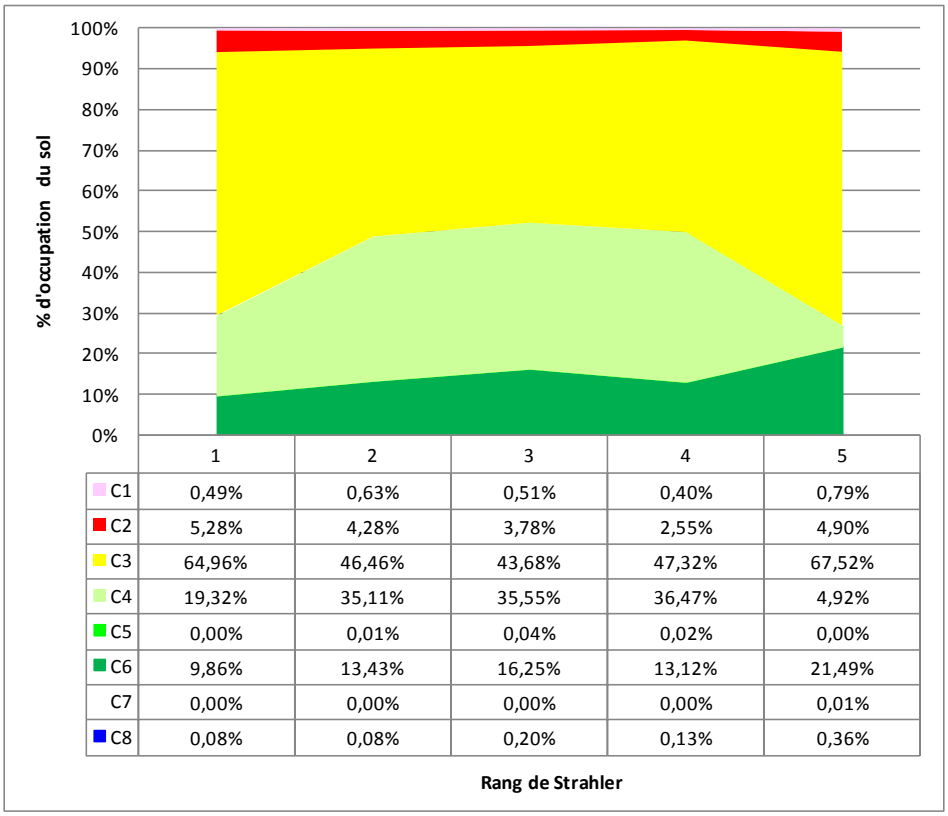
Les Figure 10.5, Figure 10.6 et Figure 10.7 résument le pourcentage moyen des différentes catégories d'occupation du sol en fonction du rang de Strahler des stations respectivement à l'échelle du BV, du CRamont et du CRlocal. On rappelle que l'information au rang 5 est peu fiable compte tenu du faible nombre de stations représentant cette taille de cours d'eau.

A l'échelle du BV (Figure 10.5), on constate en moyenne sur l'ensemble de la zone que les stations ont des bassins versant agricoles (80 % de terre arables et de prairies agricoles) avec un pourcentage de terres arables (C3) accru de +20% pour les stations en tête de bassin et les plus en aval (rang 5, moins significatif). Le taux d'urbanisation (urbain continu et discontinu) et de végétation naturelle (essentiellement arborée) reste constant sur les bassins versants quelque soit le rang du cours d'eau.

A l'échelle du CRamont (sur une bande de 5 m de part et d'autre du cours d'eau, Figure 10.6), la proportion d'espaces naturels est plus importante qu'à l'échelle du BV puisqu'en moyenne 40 % environ de cet espace est recouvert de végétation semi-naturelle (C5, C6). Par ailleurs, on note que les stations en tête de bassin (rang 1) ont un corridor amont plus urbanisé, et que les stations les plus en aval (rang 5) ont une plus forte proportion de végétation arborée dans cet espace.

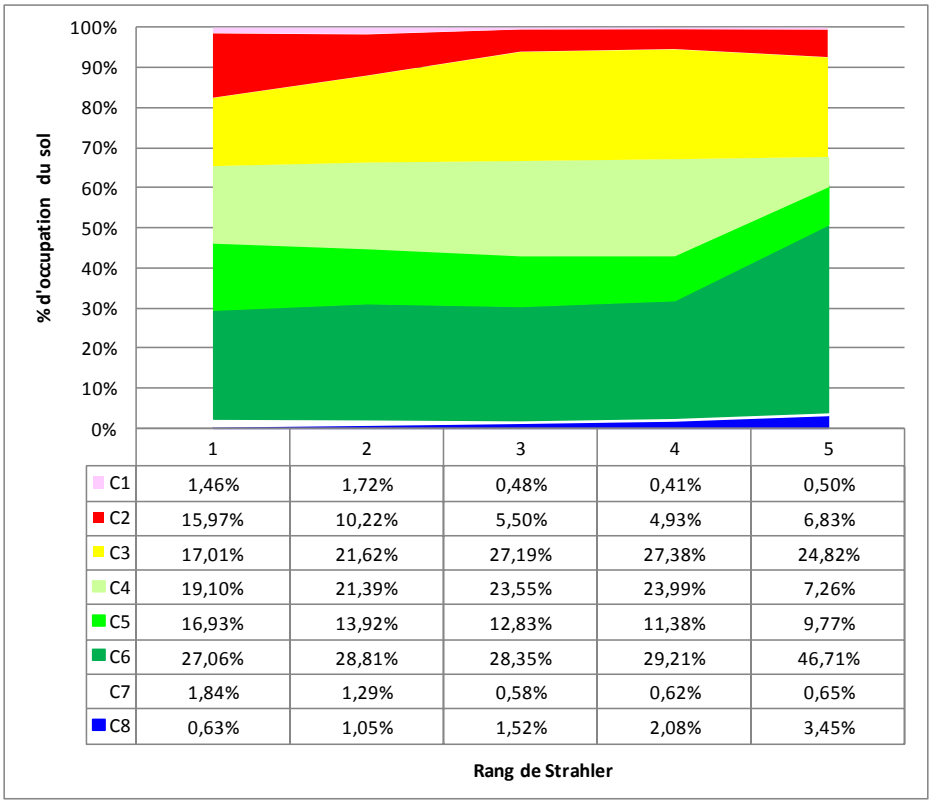
A l'échelle du CRlocal (sur une bande de 5 m de part et d'autre du cours d'eau sur 100 m en amont de la station, Figure 10.7), on observe les mêmes proportions, en moyenne, d'espaces antropisés et d'espaces semi-naturels que sur le corridor amont. Par contre, on note une plus forte proportion de territoires urbanisés (tissu urbain continu, C1, et discontinu, C2), notamment sur les cours d'eau de rang 1 et 2. On remarque par ailleurs une forte proportion de surface en eau (C8) pour les stations de rang 5, car deux des trois stations représentant ce rang sont en aval d'un plan d'eau.

CemOA : archive ouverte d'Irstea / Cemagref



	1	2	3	4	5
C1	0,49%	0,63%	0,51%	0,40%	0,79%
C2	5,28%	4,28%	3,78%	2,55%	4,90%
C3	64,96%	46,46%	43,68%	47,32%	67,52%
C4	19,32%	35,11%	35,55%	36,47%	4,92%
C5	0,00%	0,01%	0,04%	0,02%	0,00%
C6	9,86%	13,43%	16,25%	13,12%	21,49%
C7	0,00%	0,00%	0,00%	0,00%	0,01%
C8	0,08%	0,08%	0,20%	0,13%	0,36%

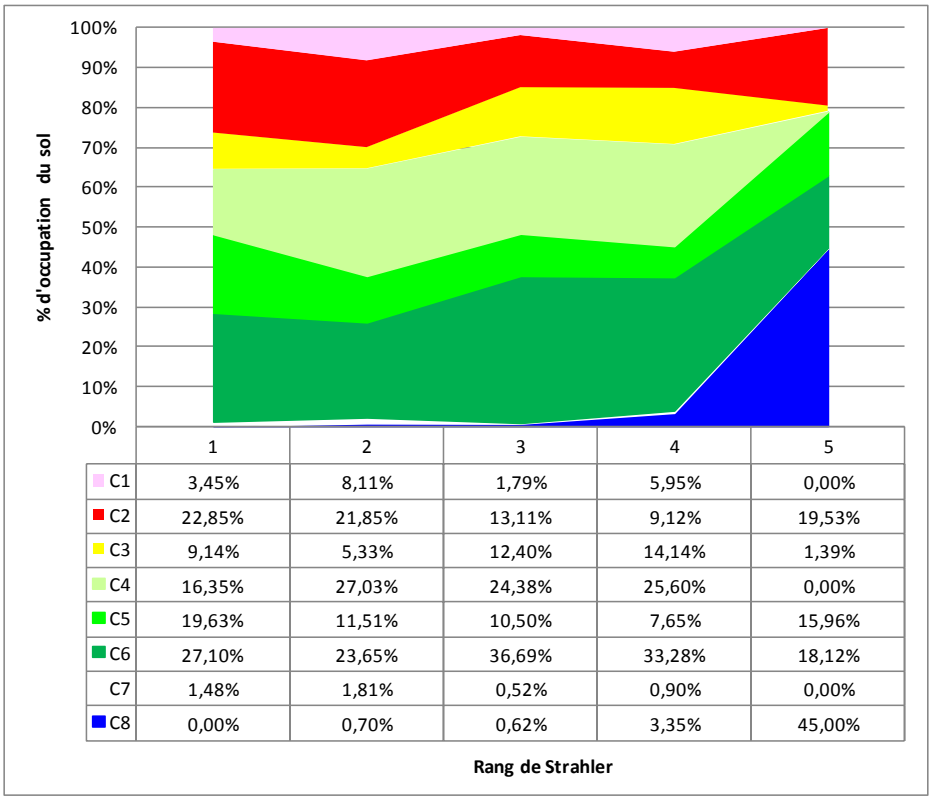
Figure 10.5 : Composition moyenne de l'occupation du sol (en %) à l'échelle du BV en fonction du rang de Strahler, selon la donnée CLC en fonction de la typologie en 8 catégories.



	1	2	3	4	5
C1	1,46%	1,72%	0,48%	0,41%	0,50%
C2	15,97%	10,22%	5,50%	4,93%	6,83%
C3	17,01%	21,62%	27,19%	27,38%	24,82%
C4	19,10%	21,39%	23,55%	23,99%	7,26%
C5	16,93%	13,92%	12,83%	11,38%	9,77%
C6	27,06%	28,81%	28,35%	29,21%	46,71%
C7	1,84%	1,29%	0,58%	0,62%	0,65%
C8	0,63%	1,05%	1,52%	2,08%	3,45%

Figure 10.6 : Composition moyenne de l'occupation du sol (en %) à l'échelle du CRamont en fonction du rang de Strahler, selon la donnée THRS en fonction de la typologie en 8 catégories.

CemOA : archive ouverte d'Irstea / Cemagref



	1	2	3	4	5
C1	3,45%	8,11%	1,79%	5,95%	0,00%
C2	22,85%	21,85%	13,11%	9,12%	19,53%
C3	9,14%	5,33%	12,40%	14,14%	1,39%
C4	16,35%	27,03%	24,38%	25,60%	0,00%
C5	19,63%	11,51%	10,50%	7,65%	15,96%
C6	27,10%	23,65%	36,69%	33,28%	18,12%
C7	1,48%	1,81%	0,52%	0,90%	0,00%
C8	0,00%	0,70%	0,62%	3,35%	45,00%

Figure 10.7 : Composition moyenne de l'occupation du sol (en %) à l'échelle du CRlocal en fonction du rang de Strahler, selon la donnée THRS en fonction de la typologie en 8 catégories.

I.2.2.3 Distribution géographique sur la zone d'étude

Dans cette partie, on présente une cartographie de la composition de l'occupation du sol en fonction des 8 sources de pressions pour chaque station à l'échelle du BV (Figure 10.8), du CRamont (Figure 10.9) et du CRlocal (Figure 10.10). La visualisation de ces cartographies permet d'identifier d'éventuelles structures spatiales, à l'échelle régionale, de cette composition d'occupation du sol. Comme précédemment, les emprises les plus proches du cours d'eau et la donnée à THRS sont utilisées pour caractériser la composition de l'occupation du sol au niveau des échelles du corridor rivulaire.

A l'échelle du BV (Figure 10.8), on distingue clairement une structure spatiale des différentes sources de pressions sur la zone d'étude. Cette répartition est liée aux orientations agricoles différenciées à l'échelle des départements : alors que les stations du centre et extrême ouest de la zone d'étude (département du Calvados et de la Manche) présentent des bassins versants couverts plutôt de prairies agricoles (C4) destinées à la production d'élevage, les autres stations de la zone ont des bassins versants couverts plutôt de terres arables pour la production de grandes cultures. On retrouve cette répartition sur la carte des orientations technico-économiques des exploitations (OTEX, Figure 4.4 cf. Chapitre 4 :II.2.2). On remarque également que les bassins versants des stations situées à l'extrême Est de la zone sont plus urbains compte tenu de leur proximité de la région parisienne.

A l'échelle du CRamont, sur une bande de 5 m de part et d'autre du cours d'eau, (Figure 10.9) on observe également une structuration spatiale régionale de la composition d'occupation du sol. La bande des 5 m est plutôt urbanisée sur l'ensemble du réseau amont des stations au Nord et à l'Est, alors qu'elle est plutôt composée de végétation arborée dans les stations du Sud et du Centre, et dominée par les prairies agricoles pour les stations de l'Ouest de la zone d'étude.

A l'échelle du CRlocal, sur une bande de 5 m de part et d'autre du cours d'eau sur 100 m en amont de station, (Figure 10.10), aucune structuration spatiale régionale ne peut être distinguée. Ceci est compréhensible car plus on travaille à une échelle locale, plus on gomme les effets régionaux au profit de la

CemOA : archive ouverte d'Irstea / Cemagref

variabilité locale. On remarque cependant, que la bande des 5 m (de part et d'autre du cours d'eau) peut être composée uniquement d'une seule catégorie d'occupation du sol sur 100 m en amont des stations. On trouve donc des stations dont le réseau amont proche (à 100 m) sont bordées essentiellement de territoires artificialisés (C1 ou C2), de terres arables (C3), de prairies agricoles (C4), et de végétation arborée (C6).

L'analyse de la distribution géographique de la composition de l'occupation des sols à trois échelles (bassin versant, corridor amont, corridor local) montre donc des signatures géographiques très différentes de ces trois échelles, ce qui renforce l'idée de leur pertinence pour analyser la variabilité de l'état écologique des cours d'eau.

CemOA : archive ouverte d'Irstea / Cemagref

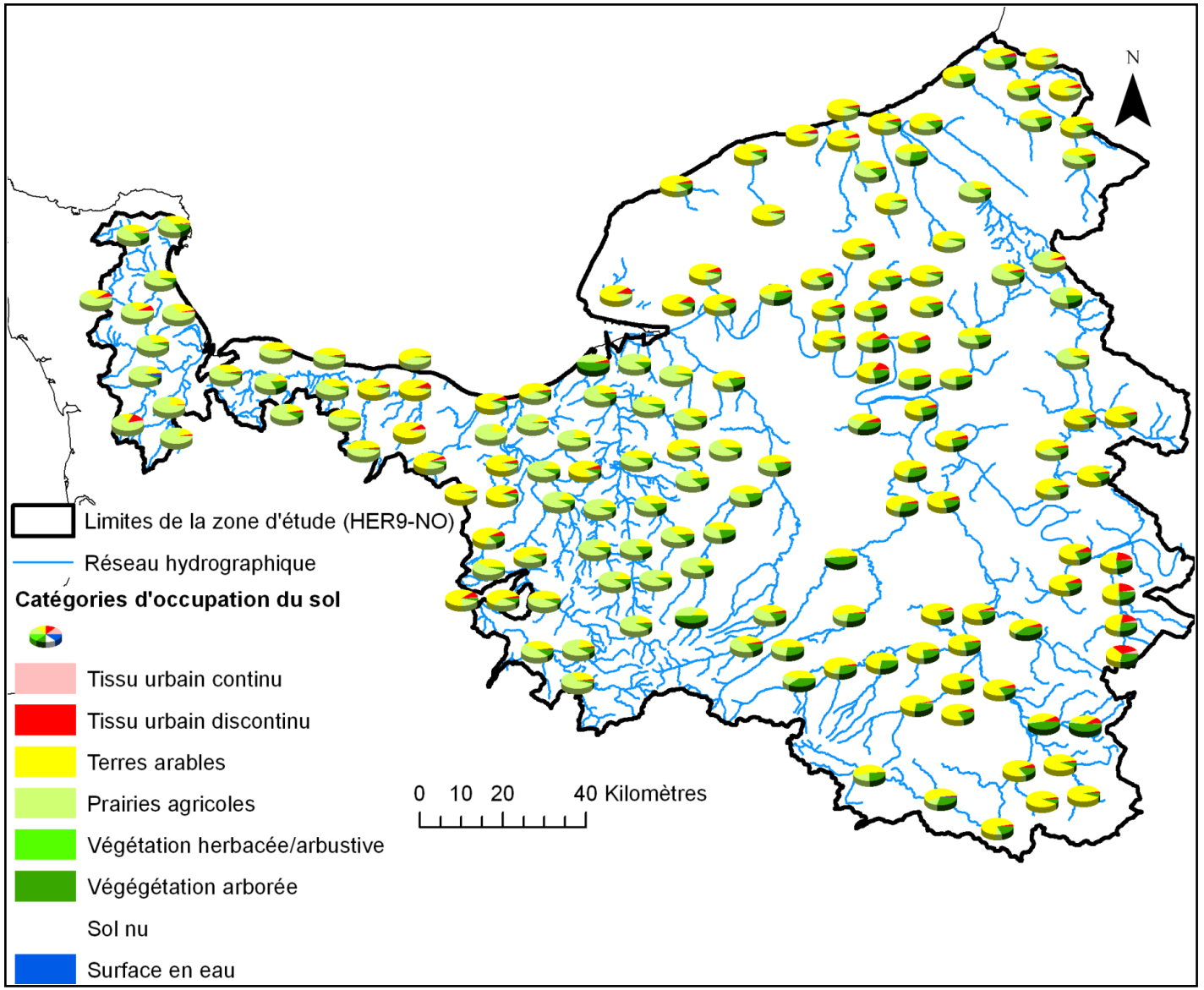


Figure 10.8 : Cartographie de la composition de l'occupation du sol en fonction des 8 sources de pressions pour chaque station à l'échelle du BV

CemOA : archive ouverte d'Irstea / Cemagref

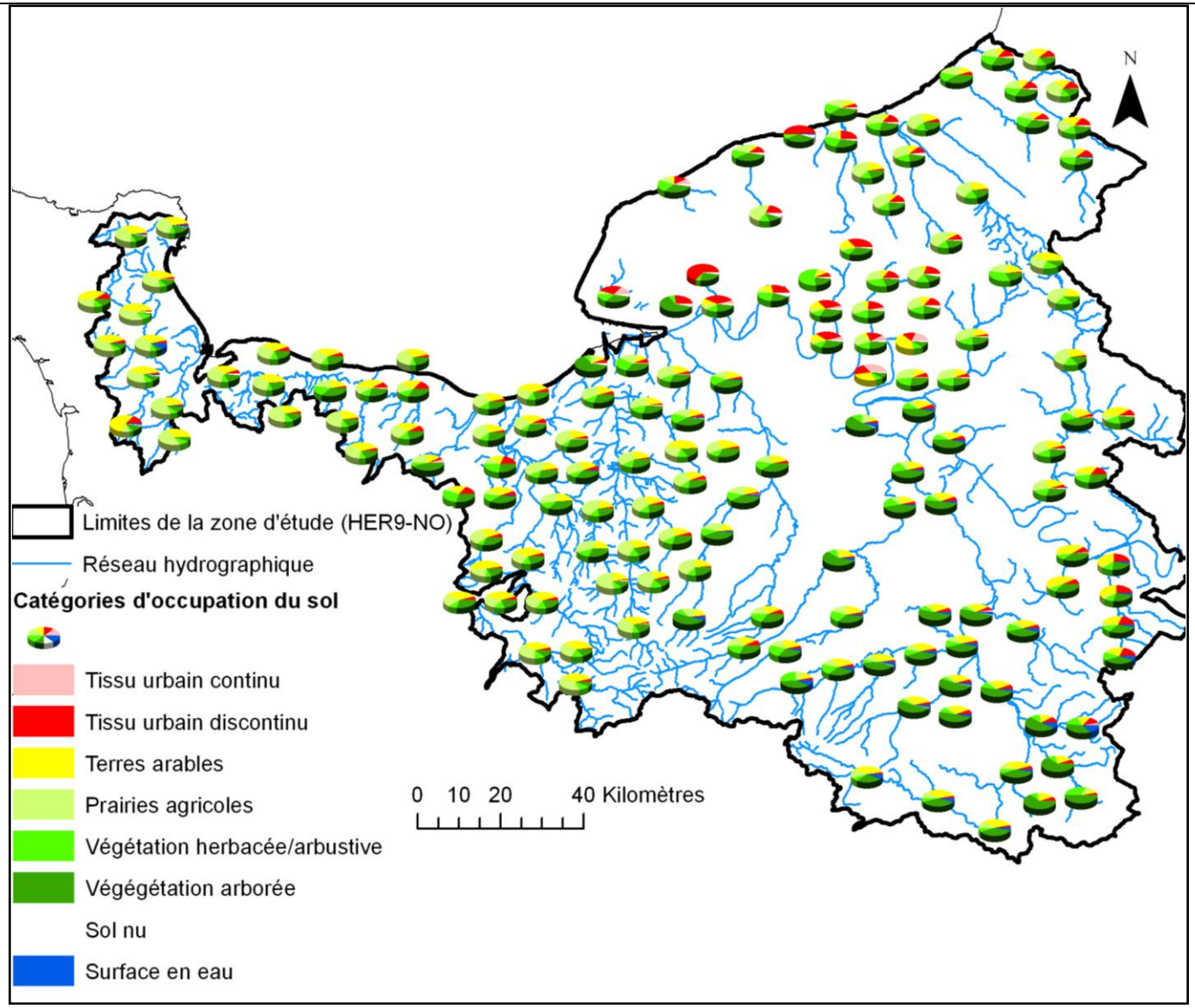


Figure 10.9 : Cartographie de la composition de l'occupation du sol à l'échelle du CRamont, pour chaque station, selon la donnée THRS en fonction des 8 sources de pressions

CemOA : archive ouverte d'Irstea / Cemagref

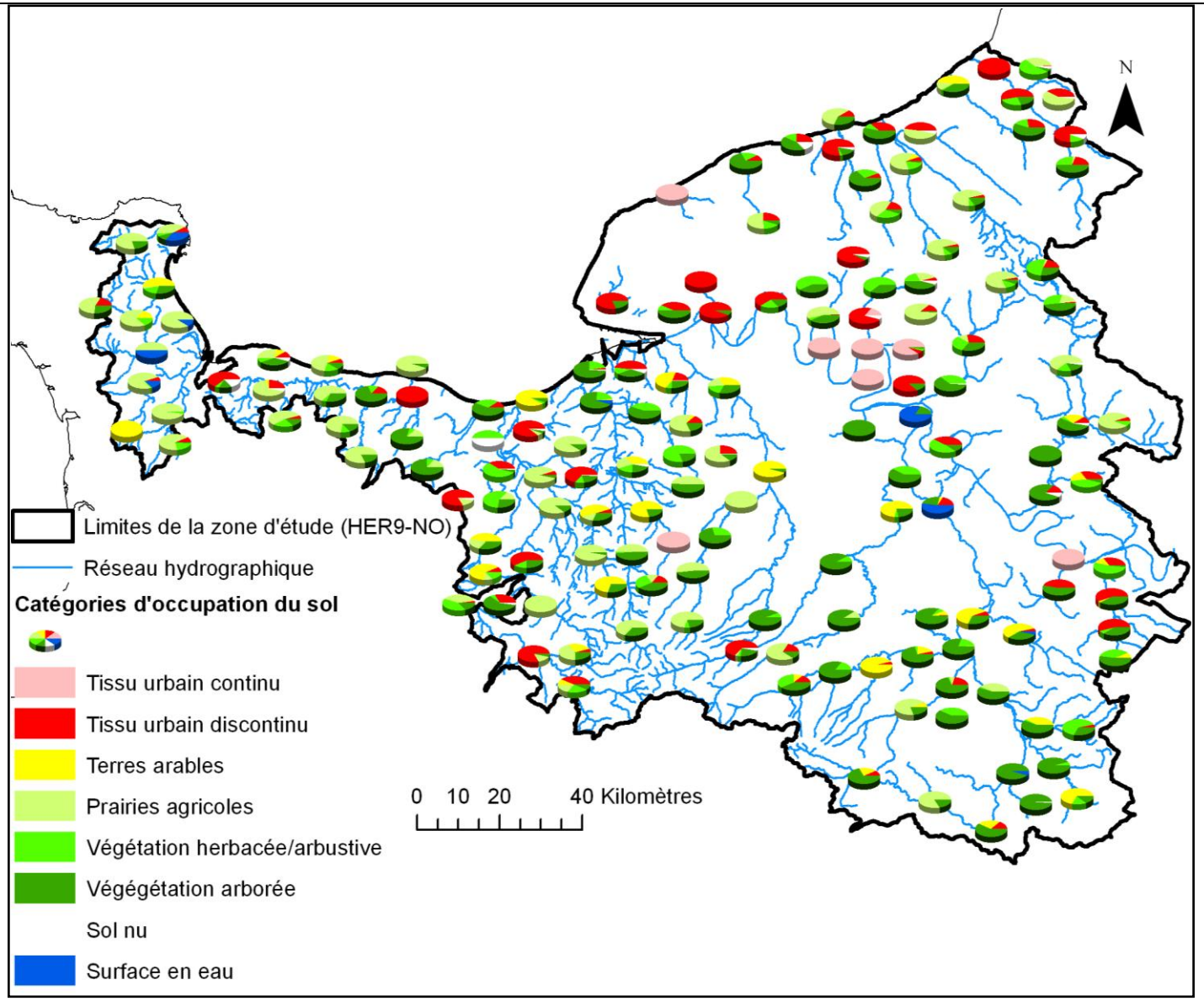


Figure 10.10 : Cartographie de la composition de l'occupation du sol à l'échelle du CRlocal, pour chaque station, selon la donnée THRS en fonction des 8 sources de pressions

CemOA : archive ouverte d'Irstea / Cemagref

I.3. Premières observations

La cartographie de l'état écologique des stations (Figure 10.1) et des pressions (occupation du sol) susceptibles de s'exercer sur celles-ci (Figure 10.8, Figure 10.9, Figure 10.10) montre sur la zone d'étude :

- des têtes de bassins (cours d'eau de rang 1) avec de plus faibles EQR-IBGN$_{moyen}$ soumises notamment à des pressions d'origine agricole sur le BV (pollution diffuse) et à des pressions d'origine urbaine (rejets directs, altération hydromorphologiques…) au niveau des espaces rivulaires à proximité du cours d'eau ;

- un état écologique qui s'améliore sur les cours d'eau de plus grande taille puisque le pourcentage de stations au dessus de la limite du bon état est plus important pour les cours de rang 2 à 4 que pour les cours d'eau de rang 1 ;

- une présence de prairies agricoles, de végétation herbacée/arbustive et arborée au niveau des espaces rivulaires plus importante à mesure que l'on se rapproche des cours d'eau, quelque soit leur rang (selon la donnée à THRS) ;

- des répartitions régionales semblables de l'état écologique et des pressions à l'échelle du BV et du CRamont. Par exemple, on observe que les stations du centre de la zone d'étude (département du calvados) sont majoritairement au dessus de la limite de bon état et que leur BV et CRamont sont dominés par les prairies agricoles. Inversement, les stations proches de la région parisienne sont majoritairement en mauvais état et ont des BV et CRamont avec de plus fortes proportions de territoires artificialisés ;

- des emprises à l'échelle du CRlocal présentant une très forte variabilité, entre couvert très anthropisé et couvert semi-naturel.

Cette analyse a permis de confirmer une nouvelle fois la limite de CLC du fait de son unité cartographique trop grossière (25 ha minimum) et l'intérêt de la THRS pour quantifier précisément et de façon fiable les indicateurs de composition et/ou de configuration de l'occupation du sol au niveau du corridor rivulaire et les indicateurs des formations végétales le long du cours d'eau. Elle montre aussi que lorsqu'on utilise la donnée CLC, on surestime la présence de prairies agricole au niveau du corridor rivulaire, présence qui apparaît plus importante qu'à l'échelle du bassin versant.

On prendra en compte ces observations lors de l'interprétation des résultats de modélisation et de la réalisation du diagnostic de l'état écologique sur la zone d'étude.

CemOA : archive ouverte d'Irstea / Cemagref

II. Résultats de la démarche de modélisation

Cette démarche est présentée dans le Chapitre 9 :IV. Elle se déroule en deux phases : une phase de construction des modèles (cf. II.1.1) et une de phase de comparaison de ces modèles (cf. II.1.2). Elle a pour double but de quantifier les voies d'amélioration (mode de spatialisation multiple, donnée à THRS) proposées dans cette thèse (objectifs méthodologiques de l'étude) et de sélectionner le modèle optimal afin d'établir un premier diagnostic de l'état écologique sur la zone d'étude (objectifs thématiques de l'étude).

II.1. Présentation des résultats

II.1.1. Construction des modèles (phase 1)

Avant de présenter les résultats des 21 modèles pressions/impacts envisagés selon différents scénarios d'échelles spatiales, de données d'occupation des sols et d'emprises (Tableau 10-6, cf. II.1.1.2) pour établir un diagnostic de l'état écologique sur la zone d'étude, on illustre à l'aide d'un exemple (Mod-3.1.4) le processus de construction d'un modèle pressions/impacts.

II.1.1.1 Illustration de la construction d'un modèle (Mod-3.1.4)

Le Mod-3.1.4 est un modèle qui étudie les relations statistiques entre des indicateurs d'occupation du sol sur les trois échelles (BV, CRamont, CRlocal) et l'EQR-IBGN$_{moyen}$. Il utilise la donnée à THRS au niveau du corridor rivulaire et le mode de spatialisation « multiple » à l'échelle du CRlocal. La méthode de construction d'un modèle (cf. Chapitre 9 :III.) nécessite deux étapes, une étape de sélection des emprises spatiales les plus significatives pour chaque type de pression (cf. II.1.1.1.1) et une étape de construction du modèle de régression PLS (cf. 0).

II.1.1.1.1 Etape 1 : Sélection des emprises spatiales

Cette sélection est effective uniquement à l'échelle du CRamont et du CRlocal car plusieurs emprises sont possibles à ces échelles pour une source de pressions donnée (options possibles grâce au mode de spatialisation du mode « multiple »). Après avoir montré un exemple de sélection pour la source de pressions « prairie permanente » (cf. II.1.1.1.1.1), on résume les résultats obtenus aux échelles du corridor rivulaire pour l'ensemble des sources de pressions étudiées (cf. II.1.1.1.1.2).

II.1.1.1.1.1 Sélection de l'emprise la plus significative pour la source de pression « prairies permanentes »

Plusieurs buffers à chacune des échelles du corridor rivulaire, CRamont, CRlocal, sont donc candidats à cette sélection pour une source de pressions donnée. 15 buffers, de largeurs différentes, au niveau du CRamont (cf. Chapitre 9 :II.2.1.5.2.3) ; 150 buffers, de longueurs et de largeurs différentes, au niveau du CRlocal (cf. Chapitre 9 :II.2.1.5.2.4.2).

Dans cet exemple, on s'intéresse à la recherche de l'emprise spatiale significative pour l'occupation du sol « prairies permanentes » (poste 231 issu de la donnée à THRS) à l'échelle du CRamont et du CRlocal. On cherche donc l'emprise (i.e., le buffer) sur laquelle cette source possède l'influence la plus significative sur la variabilité de l'EQR-IBGN$_{moyen}$.

La sélection de ce buffer, pour une échelle donnée, est basée sur la comparaison du coefficient de corrélation (R^2) du modèle linéaire simple entre l'EQR-IBGN$_{moyen}$ et chaque indicateur de cette source sur les différentes emprises spatiales (cf. Chapitre 9 :III.2).

CemOA : archive ouverte d'Irstea / Cemagref

La Figure 10.11 illustre la valeur du coefficient de corrélation (R^2) issu de la régression linéaire simple entre chaque indicateur de prairies permanentes aux différentes emprises (quelque soit l'échelle) et l'EQR-$IBGN_{moyen}$ mesurés sur les 155 stations de la zone d'étude. Ces valeurs sont présentées selon un gradient de couleur du vert, pour les valeurs les plus faibles, au jaune pour les valeurs les plus fortes.

Cette figure montre tout d'abord une variabilité du R^2 (de 0.05 à 0.12) à l'échelle du CRlocal, qui prouve qu'en fonction de l'emprise considérée l'influence d'une source de pression sur l'état écologique peut être différente. Plus les limites de l'emprise considérée à l'échelle du CRlocal s'éloignent du cours d'eau et de la station, plus la présence de prairie dans cette emprise a une influence forte sur l'EQR-$IBGN_{moyen}$.

On observe une variabilité beaucoup moins marquée à l'échelle du CRamont : plus les limites de l'emprise se rapprochent du cours d'eau, plus la présence de prairie dans cette emprise a une influence forte sur l'état écologique.

L'emprise spatiale la plus significative, pour une échelle donnée, est l'emprise possédant la plus forte relation (R^2) avec la réponse biologique :

- à l'échelle du CRlocal, elle correspond donc au buffer de 200 m de largeur (de part et d'autre du cours d'eau) et de longueur de 3500 (3000 m en amont et 500 m en aval de station écologique) ;
- à l'échelle du CRamont, elle correspond au buffer de 20 m de largeur (de part et d'autre du cours d'eau).

Pour le type de pression « prairies permanentes », ce sont donc les indicateurs correspondant à ces deux emprises qui seront utilisés dans la construction du modèle PLS (2ème étape).

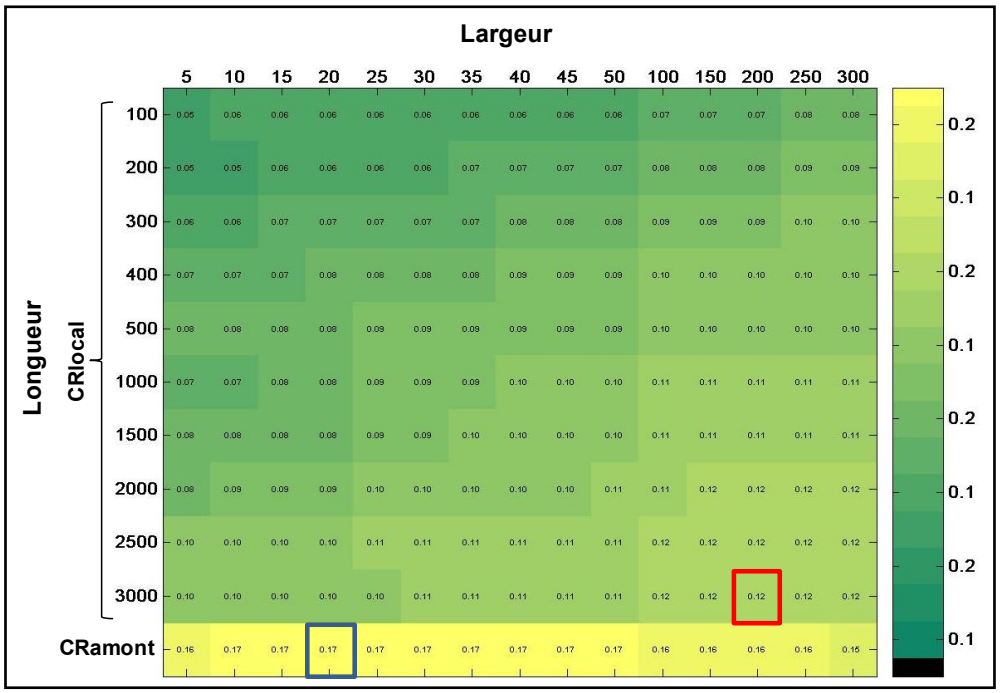


Figure 10.11 : Coefficient de corrélation (R^2) issu de la régression linéaire simple avec l'EQR-$IBGN_{moyen}$ et chaque indicateurs de prairies permanentes aux différentes emprises à l'échelle du CRlocal (buffers variant en largeur et e, longueur) et du CRamont (buffers variant en largeur).

II.1.1.1.1.2 Emprises sélectionnées aux échelles du corridor rivulaire

Le Tableau 10-4 récapitule les emprises sélectionnées pour chaque source de pressions à l'échelle du CRamont et du CRlocal pour le Mod-3.1.4. On note que l'emprise n'est pas la même pour les différentes sources de pressions. Les indicateurs correspondant à ces emprises seront donc utilisés dans la construction du modèle PLS (2ème étape).

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 10-4 : Présentation des emprises sélectionnées pour chaque source de pression étudiée par le Mod-3.4.1 à l'échelle du Cramont et du CRlocal. NP signifie qu'aucune emprise n'a été sélectionnée car la source de pression étudiée n'est pas présente sur l'ensemble des stations à l'échelle considérée.

CRamont		CRlocal-opt	
Poste OS	**Largeur**	**Largeur**	**Longueur**
111	50	300	3000
112	300	100	1000
113	300	100	3000
121	300	200	3000
122	5	5	400
124	10	5	1000
131	300	NP	
133	5	150	200
141	250	5	1500
142	15	300	100
151	300	300	400
152	150	25	100
153	150	250	400
154	5	150	300
155	300	5	2500
211	250	100	2500
212	300	150	1500
213	300	50	1000
214	250	200	2500
222	300	NP	
223	150	50	1500
231	20	200	3000
232	300	300	3000
233	300	50	1500
234	300	300	1000
241	50	300	2500
242	300	25	100
251	10	25	200
252	5	300	3000
311	250	20	200
321	250	300	3000
331	300	5	2500
412	5	300	200
414	50	100	200
415	300	200	1500
416	200	NP	
417	250	300	3000
418	300	100	2000
421	30	35	1500
427	100	200	500
428	10	10	1000

CemOA : archive ouverte d'Irstea / Cemagref

II.1.1.1.2 Etape 2 : Construction du modèle Régression PLS

L'analyse des relations pressions/impacts repose sur une méthode itérative (5 itérations minimum) et progressive (en termes de significativité) basée sur l'algorithme de régression PLS (cf. Chapitre 9 :III.3.3). Cette méthode permet d'obtenir de façon automatique le modèle PLS possédant le plus grand nombre de prédicteurs significatifs.

On présente dans le Tableau 10-5, les résultats obtenus à chaque itération lors de la construction du modèle Mod-3.1.4. Le modèle final a été obtenu au bout de la 5ème itération puisque l'ensemble des prédicteurs en entrée du modèle sont significatifs. Ce modèle explique à 37.9 % la variabilité de l'EQR-$IBGN_{moyen}$ à partir d'une combinaison de 23 indicateurs d'occupation du sol.

Tableau 10-5 Résultats obtenus à chaque itération lors de la construction du modèle Mod-3.1.4

	Nombre de prédicteurs en entrée	Intervalle de confiance	R^2	Nombre de prédicteurs significatifs
1ère itération	107	à 50 %	41.9	82
2ème itération	82	à 66 %	41.8	53
3ème itération	53	à 90 %	40.9	30
4ère itération	30	à 95 %	39.4	23
5ème itération	23	à 95 %	37.9	23

II.1.1.2 Tableau récapitulatif des différents modèles

Le Tableau 10-6 résume les résultats obtenus pour les différents (21 au total) modèles pressions/impacts développés à partir des 155 stations IBGN de la zone d'étude et construits selon différents scénarios d'échelle spatiale, de données d'occupation des sols et de mode de définition de l'emprise. La construction de chacun de ces modèles a suivi la méthodologie illustrée précédemment sur le Mod-1.3.4.

Ce tableau indique, pour chaque modèle, la capacité explicative (R^2) et la capacité descriptive : le nombre de prédicteurs (source de pressions significatives) identifiés à chaque échelle, au total et selon l'effet (dégradant ou protecteur), puis au total sur l'ensemble des échelles analysées par le modèle.

Ces résultats seront utilisés dans la phase comparative (cf. II.1.2) de la démarche. Les résultats détaillés de chaque modèle sont présentés dans l'***ANNEXE XII***.

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 10-6 : R^2 et nombre de prédicteurs sur la totalité et par échelle selon l'effet (en rose pâle, les effets dégradantes et en vert pâle les effets protecteurs) pour chacun des 21 modèles

	Echelle(s) d'analyse	Identifiant du modèle	Mode de spatialisation dans le CR	Sources d'OS dans le CR	Résultats					
					R^2	Nombre de Prédicteurs				
						Par échelle				Total
1. Modèle mono-échelle	BV	Mod-1.1.1			28.2 %	BV	3	1	4	4
	CRamont	Mod-1.2.1	multiple	CLC	30.8 %	CRamont	2	1	3	3
		Mod-1.2.2	multiple	THRS	35.7 %	CRamont	4	5	9	9
	CRlocal	Mod-1.3.1	fixe	CLC	19.1 %	CRlocal	2	1	3	3
		Mod-1.3.2	fixe	THRS	14.8 %	CRlocal	2	3	5	5
		Mod-1.3.3	multiple	CLC	23.8 %	CRlocal	2	4	6	6
		Mod-1.3.4	multiple	THRS	26.0 %	CRlocal	5	6	11	11
2. Modèle bi-échelle	BV CRamont	Mod-2.1.1	multiple	CLC	30.0 %	BV CRamont	2 0	0 1	2 1	3
		Mod-2.1.2	multiple	THRS	37.0 %	BV CRamont	3 5	1 3	4 8	12
	BV CRlocal	Mod-2.2.1	fixe	CLC	33.1 %	BV CRlocal	3 2	2 1	5 3	8
		Mod-2.2.2	fixe	THRS	29.2 %	BV CRlocal	3 2	1 2	4 4	8
		Mod-2.2.3	multiple	CLC	38.3 %	BV CRlocal	2 2	1 4	3 6	9
		Mod-2.2.4	multiple	THRS	37.8 %	BV CRlocal	3 4	1 6	4 10	14
	CRamont CRlocal	Mod-2.3.1	multiple (CRamont) fixe (CRlocal)	CLC	26.6 %	CRamont CRlocal	1 2	1 1	2 3	5
		Mod-2.3.2	multiple (CRamont) fixe (CRlocal)	THRS	34.4 %	CRamont CRlocal	3 2	4 3	7 5	12
		Mod-2.3.3	multiple	CLC	29.9 %	CRamont CRlocal	1 2	1 4	2 6	8
		Mod-2.3.4	multiple	THRS	33.8 %	CRamont CRlocal	3 5	3 6	6 11	17
3. Modèle tri-échelle	BV CRamont CRlocal	Mod-3.1.1	multiple (CRamont) fixe (CRlocal)	CLC	32.0 %	BV CRamont CRlocal	1 0 0	0 0 0	1 0 0	1
		Mod-3.1.2	multiple (CRamont) fixe (CRlocal)	THRS	36.6 %	BV CRamont CRlocal	3 3 2	1 3 3	4 6 5	15
		Mod-3.1.3	multiple	CLC	38.0 %	BV CRamont CRlocal	2 1 2	0 1 3	2 2 5	9
		Mod-3.1.4	multiple	THRS	37.9 %	BV CRamont CRlocal	3 3 6	2 3 6	5 6 12	**23**

CemOA : archive ouverte d'Irstea / Cemagref

II.1.2. Comparaison des modèles (phase 2)

Une analyse comparative pour répondre à chacun des objectifs de l'étude (objectifs méthodologiques et objectifs thématiques, cf. Chapitre 9 :I.3) est mise en œuvre dans cette phase. Cette analyse se fonde au minimum sur trois critères de comparaison : capacité explicative (critère 1), capacité descriptive (critère 2) du modèle (identification de sources de pressions significatives) et pertinence des facteurs pour les finalités de gestion (critère 3) (cf. Chapitre 9 :IV.3.1).

II.1.2.1 Évaluation de l'apport respectif des différentes voies d'amélioration

Cette analyse a pour finalité de répondre aux objectifs méthodologiques de l'étude qui sont d'évaluer (i) l'apport du mode de spatialisation « multiple » (cf. II.1.2.1.1); et (ii) l'apport de la donnée d'occupation du sol à THRS (cf. II.1.2.1.2).

II.1.2.1.1 Apport du mode de spatialisation « multiple »

On rappelle que le mode de spatialisation « fixe », mode couramment utilisé dans les modèles empiriques pressions/impacts, utilise une seule emprise (un seul buffer). Par conséquent l'influence des pressions sur l'état écologique des cours d'eau est étudiée à partir d'une même emprise quelque soit la source de pressions au niveau du corridor rivulaire. La délimitation de cette emprise (buffer) se base sur des hypothèses définies préalablement par l'analyste.

Par contre, le mode de spatialisation « multiple » consiste à considérer plusieurs emprises candidates de dimensions différentes. Par conséquent, l'influence des pressions sur l'état écologique des cours d'eau peut-être étudiée en considérant plusieurs emprises candidates pour une source de pressions donnée et en identifiant parmi elles l'emprise qui a l'influence la plus forte sur l'état écologique..

Ces deux modes de spatialisation ont été mis en place à l'échelle du CRlocal. L'emprise pour le mode « fixe » (cf. Chapitre 9 :II.2.1.5.2.4.1) correspond à un buffer de largeur dépendant du rang de strahler du cours d'eau sur 2.5 km de linaire (de même rang) en amont de la station et 0.5 km en aval. Les modèles utilisant ce mode de spatialisation à l'échelle du CRlocal ont donc analysé l'influence sur l'état écologique des sources de pressions existant au sein de cette emprise.

Pour le mode « multiple » (cf., Chapitre 9 :II.2.1.5.2.4.2), 150 buffers de couple largeur-longueur différents ont été mis en place (le couple 5-100 m étant le buffer le plus proche du cours d'eau et de la station, et 300-3000 m étant le buffer le plus éloigné). Les modèles utilisant ce mode de spatialisation à l'échelle du CRlocal ont donc analysé l'influence sur l'état écologique des sources de pressions existant au sein des emprises les plus significatives pour chacune des sources.

L'analyse comparative a été menée uniquement entre les couples de modèles utilisant au moins l'échelle du CRlocal. 8 couples de modèles ont donc été comparés (parmi les 21 modèles construits précédemment).

Le Tableau 10-7 présente les résultats de cette analyse en quantifiant l'apport du mode « multiple » par rapport mode « fixe » sur la capacité explicative (gain ou perte de R^2) et la capacité descriptive (gain ou perte du nombre de prédicteurs obtenus au total sur l'ensemble des échelles). Il indique aussi le nombre de sources de pressions gagnées et perdues à l'échelle du CRlocal.

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 10-7 : Gains (ou pertes) de pouvoir explicatif (R^2), de prédicteurs sur l'ensemble des échelles analysées, obtenus par l'application mode de spatialisation « multiple » par rapport au mode « fixe » pour chaque couple de modèles de même nombre et de combinaison d'échelles et de source de donnée d'occupation du sol utilisant au moins l'échelle du CRlocal.

Echelle(s) d'analyse	Identifiants des modèles comparées	Source OS dans le CR	Apport du mode spatialisation « multiple »/ « fixe »			
			Gain (ou perte) de R^2	Nb. de prédicteurs total	Nb. de prédicteurs à l'échelle du CRlocal	
					gagné	perdu
CRlocal	Mod-1.3.1/Mod-1.3.3	CLC	**+4,7 %**	**+3**	**+3**	**0**
	Mod-1.3.2/Mod-1.3.4	THRS	**+11,2 %**	**+6**	**+6**	**0**
BV CRlocal	Mod-2.2.1/Mod-2.2.3	CLC	**+4,7 %**	**+1**	**+1**	**0**
	Mod-2.2.2/Mod-2.2.4	THRS	**+ 8,6 %**	**+6**	**+6**	**0**
CRamont CRlocal	Mod-2.3.1/Mod-2.3.3	CLC	**+3.3 %**	**+3**	**+3**	**0**
	Mod-2.3.2/Mod-2.3.4	THRS	**-0.6 %**	**+5**	**+6**	**0**
BV CRamont CRlocal	Mod-3.1.1/Mod-3.1.3	CLC	**+6 %**	**+8**	**+5**	**0**
	Mod-3.1.2/Mod-3.1.4	THRS	**+1.3 %**	**+8**	**+8**	**0**

D'après le Tableau 10-7, on remarque que, quelque soit les échelles et la source de données utilisées dans le corridor rivulaire, le mode de spatialisation multiple permet systématiquement :

(1) un gain d'explication (critère 1) de la variabilité de l'EQR-$IBGN_{moyen}$.
Une seule exception « neutre » concerne les modèles qui couplent l'échelle du CRamont et du CRlocal avec l'utilisation de la donnée THRS, pour lesquels ce mode de spatialisation induit une très faible perte (-0.6 %) ;

(2) un gain sur la capacité descriptive (critère 2) puisqu'un plus grand nombre de prédicteurs (sources de pressions significatives) sont identifiés. On notera de plus que l'application d'emprises adaptées à chaque source de pression n'entraine pas de perte d'information : les sources de pressions identifiées par les modèles utilisant le mode « fixe » sont identifiées également par ceux utilisant le mode « multiple ».

Ainsi, cette analyse comparative montre que l'utilisation du mode « multiple » à l'échelle du CRlocal permet d'obtenir des modèles plus pertinents pour les gestionnaires : ils expliquent une plus grande part de la variabilité de la réponse biologique et identifient un plus grand nombre de sources de pressions significatives.

De plus, le mode de spatialisation multiple précise l'emprise d'influence significative de chacune des sources de pressions au niveau du corridor rivulaire : sur quelle bande rivulaire ? (distance de part et d'autre du cours d'eau) ; sur quelle longueur de cours d'eau ? Cette information est déterminante pour aider les gestionnaires à élaborer des stratégies de restauration (critère 3).

II.1.2.1.2 Apport de la donnée d'occupation du sol à THRS

On rappelle que la donnée CLC (cf. Chapitre 9 :II.2.1.3.1) est la base de données d'occupation du sol européenne issue de la photo-interprétation d'images à moyenne résolution spatiale (i.e. 20 et 30 m) avec une unité cartographique minimum de 25 ha. C'est le type de donnée d'occupation du sol couramment utilisée dans les modèles pressions/impacts à large échelle.

La donnée THRS (cf. Chapitre 9 :II.2.1.3.2) est issue d'un traitement semi-automatique qui combine l'information d'images à THRS spatiale (orthophotos, 0.5 m, et Spot5 XS, 10m) et de données auxiliaires. Plus précise, plus détaillée et plus fiable que CLC, cette donnée améliore significativement la quantification des sources de pressions dans cet espace (Tormos *et al.*, 2010 et cf. I.2.2.1.3)

CemOA : archive ouverte d'Irstea / Cemagref

Afin d'étudier l'apport de cette donnée spatiale et de sélectionner la source de données d'occupation du sol la plus efficace pour étudier les relations pressions/impacts, une analyse comparative des modèles a été mise en place. Cette analyse compare les couples de modèles utilisant les mêmes échelles d'analyse et le même mode de spatialisation. La donnée à THRS étant disponible uniquement au niveau du corridor rivulaire, 10 couples de modèles ont été comparés (20 parmi les 21 modèles construits précédemment).

Le Tableau 10-8 présente l'apport de la donnée THRS par rapport à la donnée CLC sur la capacité explicative (gain ou perte de R^2) et la capacité descriptive (nombre de prédicteurs par échelle et sur la totalité des échelles) des modèles.

D'après ce Tableau, on constate tout d'abord globalement un gain de la capacité explicative (critère 1) des différents modèles en mode de spatialisation « multiple » –dont on a vu précédemment l'intérêt- lorsqu'on utilise la THRS au niveau du corridor rivulaire.

On relève des exceptions « neutres » sur les modèles biéchelle « BVamont/CRlocal » et les modèles tri-échelle « BVamont/CRamont/CRlocal» où l'on constate une très faible perte de R^2 (-0.5 % et -0.1% respectivement). Ces modèles atteignent tous deux le maximum de variabilité de l'EQR-IBGN$_{moyen}$ qu'on est capable d'expliquer avec les indicateurs d'occupation du sol dans cette étude (environ 38 % sur ce jeu de donnée). La THRS malgré son influence positive dans les autres modèles ne permet pas de dépasser cet optimum.

Cas particulier des modèles mono-échelle « CRlocal » et bi-échelle « BV/CRlocal » :

On remarque toutefois que sur les modèles mono-échelle « CRlocal » et bi-échelle « BV/CRlocal » utilisant le mode de spatialisation « simple » à l'échelle du CRlocal, l'utilisation des données THRS entraîne une perte de R^2 significative (-4.3 % et -3.9 % respectivement) par rapport à l'utilisation de la donnée CLC. Dans ces deux cas l'emprise fixe correspond au lit majeur des cours d'eau (cf. Chapitre 9 :II.2.1.5.2.4.1) soit un buffer dont la largeur est comprise entre 50 m et 125 m de part et d'autre du cours d'eau. On est donc en limite d'utilisation de la donnée CLC sur cet espace, puisque l'unité cartographique minimum de CLC est de 25 ha (500×500 m). Une explication possible est que dans ces modèles les indicateurs du « CRlocal » construits avec CLC sont en fait fortement correllés aux indicateurs du « CRamont » qui influencent plus la variable biologique lorsqu'on compare les modèles mono-échelle. Cette explication s'appuie sur plusieurs faits : (i) l'information d'occupation du sol à l'échelle du CRlocal est très similaire à celle du CRamont avec la donnée CLC (cf. I.2.2.1.1) ; (ii) les facteurs identifiés à l'échelle du CRlocal en utilisant ce mode de spatialisation sont de même nature que les facteurs identifiés à l'échelle du CRamont. La validation de cette hypothèse nécessiterait des analyses additionnelles, non réalisées dans cette étude, sur les corrélations existantes entre les facteurs de même nature (même source de pressions) identifiés à des échelles différentes.

On constate ensuite, que les relations pressions/impacts sont mieux décrites (critère 2) par la donnée THRS que par la donnée CLC au niveau du corridor rivulaire, puisque, quelque soit la combinaison d'échelles analysées, on observe une augmentation du nombre de prédicteurs significatifs sur chacune des échelles (colonne 5 du Tableau 10-8) et sur l'ensemble des échelles (colonne 6 du Tableau 10-8).

De plus, on observe, au vu des résultats détaillés de chaque modèle (cf. ***ANNEXE XII***), que les sources de pressions identifiées en utilisant la donnée CLC au niveau du corridor rivulaire sont quasiment toutes reconnues avec la donnée THRS quelque soit l'effet de ces pressions (dégradant ou protecteur) et l'échelle ou la combinaison d'échelle analysées. Même si cette observation est difficilement quantifiable compte tenu des différences entre les deux typologies d'occupation du sol, l'information apportée par la THRS semble donc richement complémentaire à CLC au niveau du corridor rivulaire sur les différents modèles analysés.

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 10-8 : Gains (ou pertes) de pouvoir explicatif (R^2) et de prédicteurs, sur l'ensemble des échelles et par échelles, obtenus par l'application au niveau du corridor rivulaire de la THRS par rapport à CLC pour chaque couple de modèles de même(s) échelle(s) d'analyse utilisant le même mode de spatialisation. En gris les couples de modèles utilisant le mode de spatialisation « multiple » à l'échelle du CRlocal.

Echelle(s) d'analyse	Identifiants des modèles comparées	Mode de spatialisation dans le CR	Apport THRS/ CLC		
			R^2	Nombre de prédicteurs	
				Par échelle	Au Total
CRamont	Mod-1.2.1/Mod-1.2.2	multiple	**+4,9 %**	**+6**	**+6**
CRlocal	Mod-1.3.1/Mod-1.3.2	fixe	**-4,3 %**	**+2**	**+2**
	Mod-1.3.3/Mod-1.3.4	multiple	**+2,2 %**	**+5**	**+5**
BV CRamont	Mod-2.1.1/Mod-2.1.2	multiple	**+7 %**	**+2** **+7**	**+9**
BV CRlocal	Mod-2.2.1/Mod-2.2.2	fixe	**-3,9 %**	**-1** **+1**	**0**
	Mod-2.2.3/Mod-2.2.4	multiple	**-0,5 %**	**+1** **+4**	**+5**
CRamont CRlocal	Mod-2.3.1/Mod-2.3.2	multiple (CRamont) fixe (Crlocal)	**+7.8 %**	**+5** **+2**	**+7**
	Mod-2.3.3/Mod-2.3.4	multiple	**+3,9 %**	**+4** **+5**	**+9**
BV CRamont CRlocal	Mod-3.1.1/Mod-3.1.2	multiple (CRamont) fixe (Crlocal)	**+4.6 %**	**+3** **+6** **+5**	**+14**
	Mod-3.1.3/Mod-3.1.4	multiple	**-0,1 %**	**+3** **+4** **+7**	**+14**

Ainsi, cette analyse comparative, montre que l'utilisation de la donnée THRS (apportant une information plus précise et plus détaillée que la donnée CLC pour quantifier les indicateurs d'occupation du sol au niveau du corridor rivulaire) permet d'obtenir des modèles plus pertinents pour les gestionnaires puisqu'ils expliquent globalement une plus grande part de la variabilité de la réponse biologique (pour la plupart des modèles comparés) et identifient un plus grand nombre de sources de pressions.

De plus, du fait de son contenu plus informatif (détaillé) sur la nature des pressions et de sa précision géométrique plus fine que CLC, l'utilisation de la donnée THRS dans la modélisation pressions/impacts apporte des éléments pertinents pour les gestionnaires (critère 3) puisqu'elle :

(1) précise la nature et la zone d'effet des sources de pressions ayant une influence significative. C'est le cas par exemple de l'effet protecteur des prairies détecté à l'échelle du CRamont : la nature des prairies est précisée avec la THRS (prairies « permanentes ») et leur zone d'influence est précisée (bande de 20m avec THRS contre bande de 50m avec CLC) ;

(2) identifie de nouvelles forces motrices générant des pressions dégradantes au niveau du corridor rivulaire. C'est le cas par exemple des routes qui ont un impact sur les processus hydromorphologiques;

(3) met en évidence le rôle des ripisylves à l'échelle du CRlocal (lorsque le mode de spatialisation « multiple » est utilisé) puisque un effet protecteur de la végétation arborée est détecté sur une bande de 20 m de part et d'autre du cours d'eau sur 200 m en amont des stations.

CemOA : archive ouverte d'Irstea / Cemagref

II.1.2.2 Sélection du modèle optimal

La sélection du modèle optimal se fait par élimination successive de modèles en appliquant dans l'ordre les critères de comparaison.

Le premier critère de comparaison correspond à la capacité explicative des modèles (R^2). La Figure 10.12, montre l'évolution de cette capacité en fonction des différentes hypothèses de construction de modèles (21 modèles au total, cf. Tableau 10-6).

D'après cette Figure, on observe tout d'abord, que la variabilité de l'état écologique est expliquée au minimum à 14.8 % (Mod-1.3.2) et au maximum entre 35 et 38 %[49] environ pour 7 modèles d'hypothèses de construction différentes.

Cet optimum est atteint par des modèles mono-échelle, bi-échelle et tri-échelle :

(i) le modèle mono-échelle portant sur l'échelle du CRamont avec la donnée THRS au niveau du corridor rivulaire (***Mod-1.2.2***) obtient un R^2 de 35.7 %.

(ii) les modèles bi-échelle. Le modèle « BV/CRamont » utilisant la donnée THRS pour quantifier les sources de pressions à l'échelle du corridor (***Mod-2.1.2***) obtient un R^2 de 37 % ; les modèles « BV/CRlocal » qui utilisent le mode de spatialisation « multiple » (***Mod-2.2.3*** qui emploie CLC pour quantifier les sources de pressions au niveau du corridor rivulaire; ***Mod-2.2.4*** qui utilise lui la THRS) obtiennent respectivement des R^2 de 38,3% et 37,8%.

(iii) les modèles tri-échelle qui combinent les hypothèses de construction de ces modèles bi-échelle (***Mod-3.1.2*** ; ***Mod-3.1.3*** utilisant CLC ; ***Mod-3.1.4*** utilisant la THRS) obtiennent respectivement des R^2 de 36,6%, 38% et 37,9%.

L'application du premier critère a donc permis de sélectionner 7 modèles parmi les 21 modèles. Ces 7 modèles ont été comparés par la suite en fonction du deuxième critère : la capacité descriptive du modèle (le nombre de facteurs significatifs identifiés). La

Figure 10.13 présente l'évolution de cette capacité en fonction des 7 modèles.

D'après cette figure, on s'aperçoit qu'un modèle se détache de façon distincte des autres modèles : le modèle tri-échelle utilisant la donnée THRS au niveau du corridor rivulaire et le mode spatialisation « multiple » à l'échelle du CRlocal (Mod-3.1.4) identifie 23 sources de pressions au total exerçant chacune un impact significatif sur l'EQR-$IBGN_{moyen}$.

Dès l'application du deuxième critère, on peut donc conclure que le modèle optimal pour établir un diagnostic de l'état écologique le plus pertinent sur la zone d'étude est le Mod-3.1.4. Ce choix est conforté par plusieurs faits : (1) il utilise les deux voies d'amélioration (donnée THRS et mode de spatialisation multiple) qui permettent d'identifier et caractériser les facteurs les plus pertinents (nature et localisation de l'emprise d'influence) pour la gestion ; (2) il intègre les facteurs influençant les processus de fonctionnement au niveau des trois échelles de l'écosystème rivière.

[49] Ces pourcentages de R^2 sont considérés comme de bon résultats pour un modèle pressions/impacts (« forces motrices/états ») compte tenu de la distance, en terme de processus écologiques, entre les indicateurs d'occupation du sol estimés à partir de l'imagerie satellitaire et la réponse biologique, puis du bruit important contenu dans les données de réseaux collectées (différents opérateurs, plan d'échantillonnage variable et déséquilibré entre les stations) et parfois couvrant des années décalées dans le temps par rapport à l'année de la donnée d'occupation du sol, ce qui affaiblit d'autant plus les relations (Friberg, 2010 ; Wasson *et al.*, 2010).

CemOA : archive ouverte d'Irstea / Cemagref

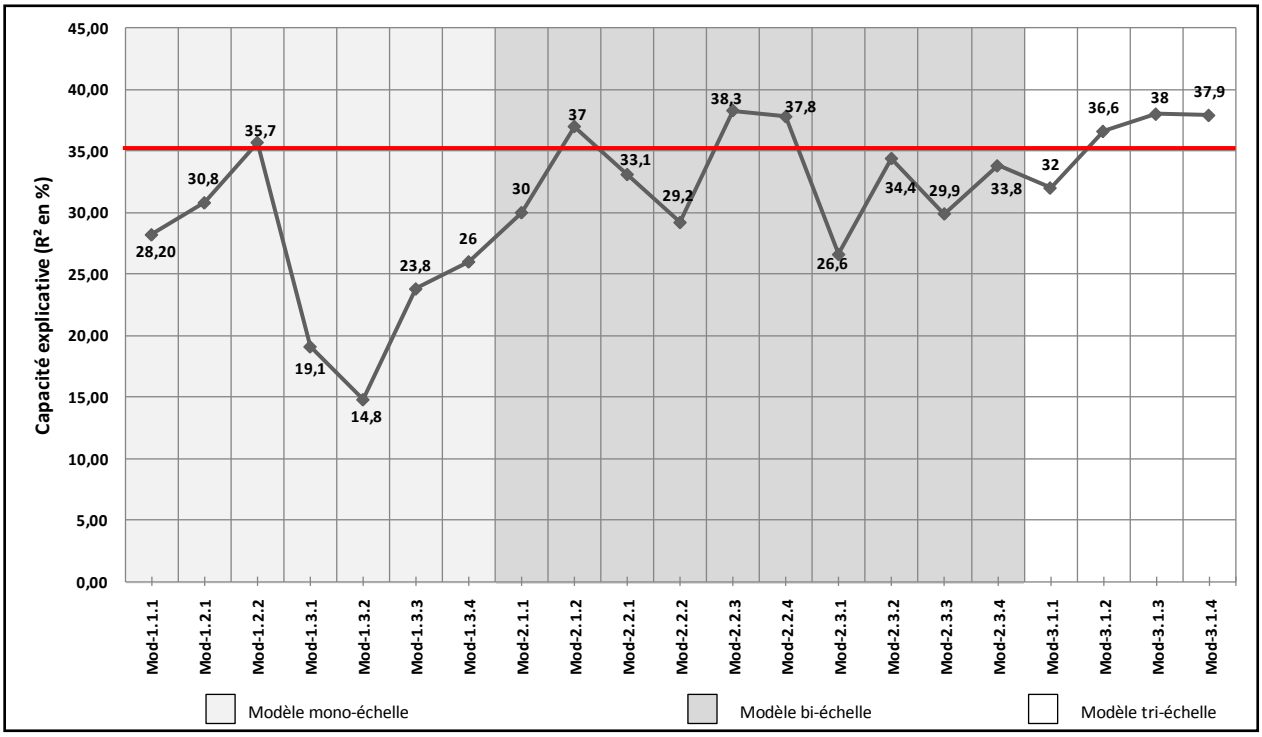


Figure 10.12 : Évolution de la capacité explicative(R^2) en fonction des différentes hypothèses de construction des modèles. La ligne en rouge correspond au seuil de sélection des modèles en fonction de la capacité explicative.

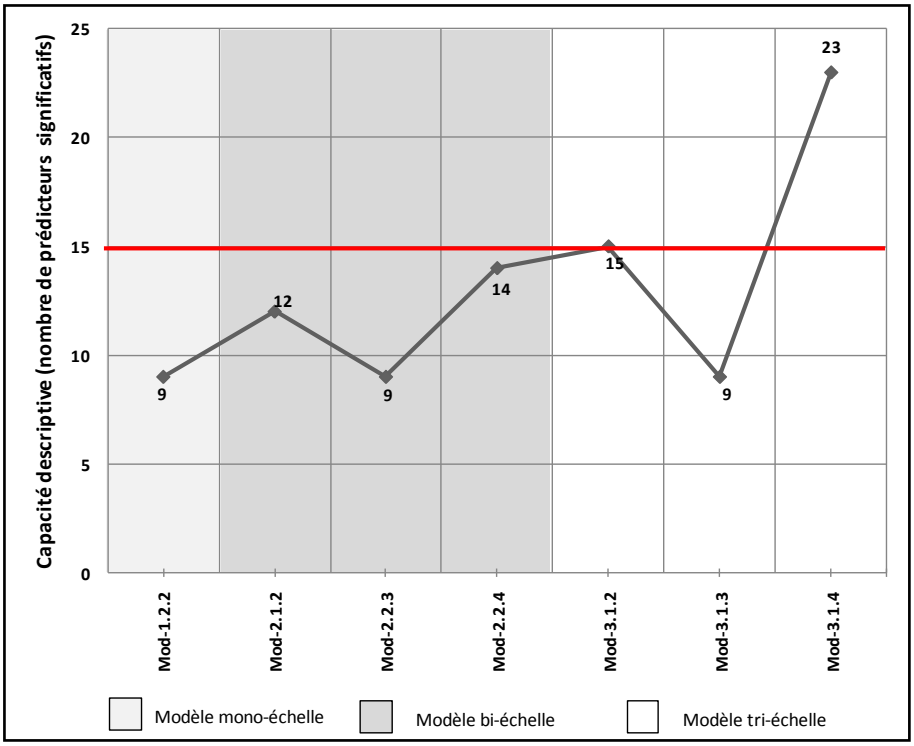


Figure 10.13 : Capacité descriptive (nombre de prédicteurs significatifs) des 7 modèles sélectionnés après l'application du premier critère (capacité explicative).

CemOA : archive ouverte d'Irstea / Cemagref

II.2.Bilan et fiabilité des résultats

II.2.1. Apport des voies d'améliorations

L'application de la démarche de modélisation a permis tout d'abord de quantifier l'apport des avancées méthodologiques mises en avant dans cette thèse pour mieux étudier les relations pressions/impacts au niveau du corridor rivulaire. L'analyse comparative des modèles montre que :

(1) quelque soit la source de donnée d'occupation du sol (CLC ou THRS), l'utilisation du mode de spatialisation « multiple » à l'échelle du CRlocal, qui permet de sélectionner pour chaque source de pressions l'emprise spatiale la plus significative, améliore les performances des modèles pressions/impacts (leur pouvoir explicatif, l'identification et la localisation des sources de pressions significatives) (cf. II.1.2.1.1).

(2) l'utilisation de la donnée THRS, apportant une information plus précise et plus détaillée que la donnée CLC pour quantifier les indicateurs d'occupation du sol au niveau corridor rivulaire, améliore les performances (leur pouvoir explicatif, l'identification et la localisation des sources de pressions significatives) des modèles pressions/impacts utilisant le mode de spatialisation « multiple » à l'échelle du CRlocal (cf. II.1.2.1.2).

La Figure 10.14 met en perspective l'apport de ces voies d'amélioration. Elle présente, en se basant sur les résultats obtenus sur la zone d'étude, l'évolution des performances des modèles pressions/impacts (modèles « forces motrices/états ») au niveau du corridor rivulaire depuis le modèle développé en France par Wasson et al. 2005 (cf. Chapitre 3 :III.2.2) jusqu'aux modèles développés dans cette étude suivant les différentes voies d'amélioration.

Le modèle de Wasson et al. (2005), modèle BV/CRlocal utilisant la donnée CLC quelque soit l'échelle et une seule et même emprise spatiale (mode « fixe ») à l'échelle du CRlocal (Mod-2.2.1, cf. ***ANNEXE XII.2.1*** pour les résultats détaillés de modèle), explique à 33.1 % la variabilité de la réponse biologique, soit environ 5 % de mieux que le modèle BV seul (Mod-1.1.1, cf. ***ANNEXE XII.1.1*** pour les résultats détaillés de modèle). Ce modèle permet donc de mettre en évidence un effet du CRlocal en identifiant 3 variables significatives à cette échelle (sur le lit majeur du cours d'eau à 2.5 km en amont et 0.5 km en aval de la station).

Ce modèle montre que la présence de tissu urbain discontinu et de zones industrielles et commerciales sur cette emprise a un impact négatif sur l'état écologique (EQR-IBGN$_{moyen}$) de la station. Ces forces motrices sont directement à l'origine de pollutions directes organiques et toxiques (rejets des stations d'épurations, déversoirs d'orages) et indirectement à l'origine de pressions hydromorphologiques (ouvrages et aménagements de protection ; dérivation ; imperméabilisation des sols…) perturbant les processus de fonctionnement et la structure du cours d'eau (cf. IV.2.1.1).

Il permet également de montrer que les cours d'eau possédant un lit majeur dominé par les prairies dans un environnement proche de la station (3 km en amont), sont en bon état écologique. Les prairies dans cet espace peuvent jouer un rôle de filtre vis-vis des flux de polluants diffus agricoles et urbains issus du bassin versant (cf. IV.2.1.3). Toutefois, compte tenu de la granularité de la donnée d'occupation du sol CLC (25 ha), on peut avoir des doutes sur la possibilité de traduire, via les indicateurs en dérivant, les mécanismes plus fins oeuvrant au niveau du corridor rivulaire.

CemOA : archive ouverte d'Irstea / Cemagref

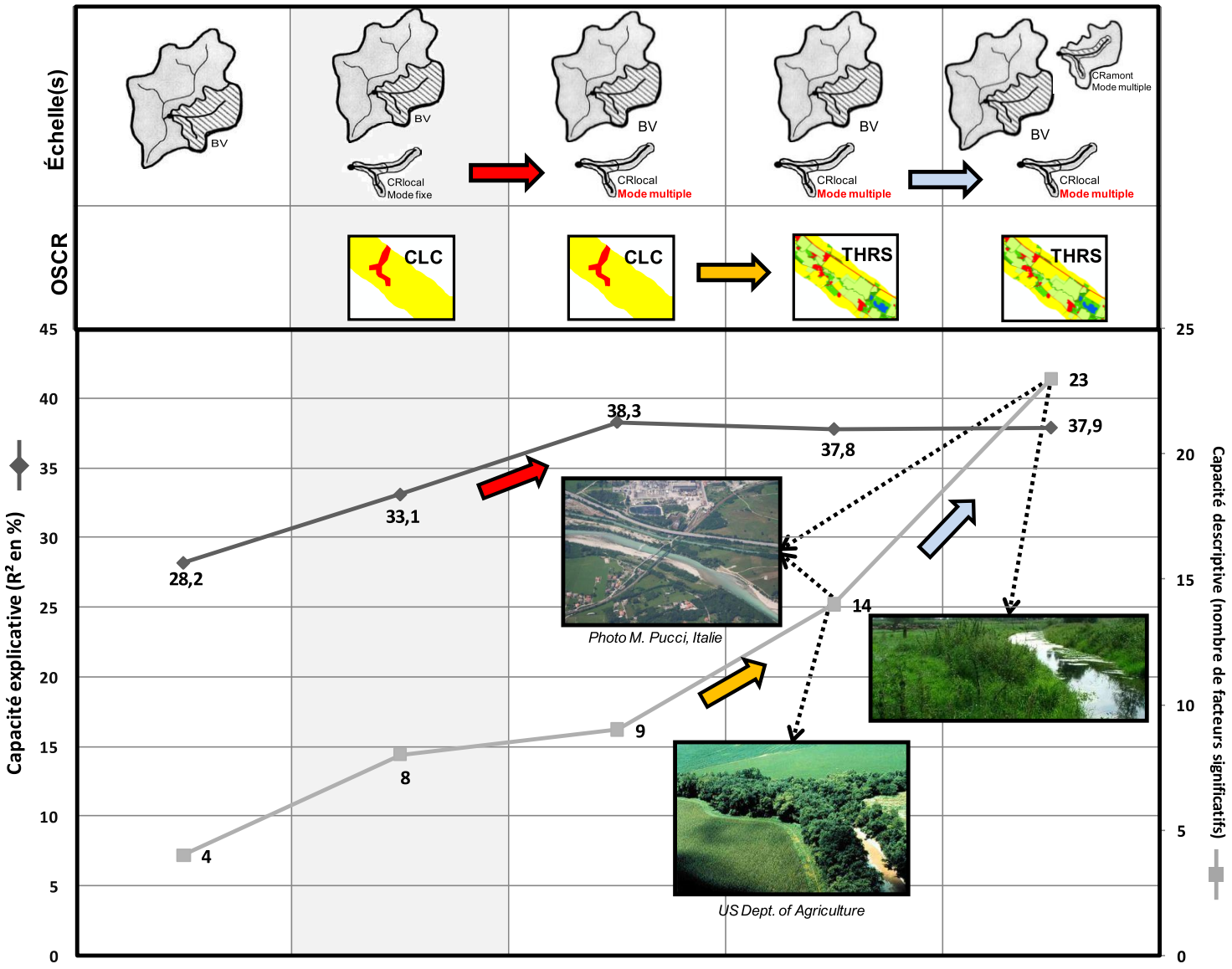


Figure 10.14 : Évolution de la performance des modèles pressions/impacts au niveau du corridor rivulaire : depuis le modèle développé en France en 2005 par Wasson et al. (en gris sur le graphique) jusqu'aux modèles développées dans cette étude suivant les différentes voies d'améliorations. OSCR : (occupation du sol dans le corridor rivulaire)

CemOA : archive ouverte d'Irstea / Cemagref

Si l'on intègre dans un premier temps, la première voie d'amélioration (mode de spatialisation « multiple ») à l'échelle du CRlocal (Mod-2.2.3, cf. ***ANNEXE XII.2.3*** pour les résultats détaillés de modèle), on explique 38.3 % de la variabilité de l'EQR-IBGN$_{moyen}$, soit environ 5 % de mieux que le modèle précédent (donc 10 % de plus que le modèle BV seul). Cette voie d'amélioration renforce donc l'effet CRlocal par le précédent modèle.

L'utilisation du mode de spatialisation multiple dans le modèle permet de préciser l'effet des différentes sources de pressions identifiées précédemment à cette échelle (sur quelle largeur de bande rivulaire ? à quelle distance de la station ?, e.g. effet du tissu urbain sur la bande des 150 m de part et d'autre du cours d'eau). Elle permet également de mettre en évidence que les cours d'eau possédant, en plus des prairies agricoles, un environnement plutôt naturel (forêt de conifères, forêts et végétation en mutation) proche de la station (sur 300 m de part et d'autre du cours d'eau sur 3 km en amont de la station) sont en bon état écologique. Comme précédemment on peut supposer que cet effet est lié au rôle de filtre que peuvent jouer ces couverts vis-à-vis des flux latéraux de polluants provenant des bassins versants très agricoles. Mais comme précédemment aussi, compte tenu de la granularité de la donnée d'occupation du sol CLC (25 ha), on peut avoir des doutes sur la possibilité de traduire, via ces indicateurs, les mécanismes plus fins œuvrant au niveau du corridor rivulaire.

Si on intègre, dans un deuxième temps, la seconde voie d'amélioration (donnée à THRS) à l'échelle du CRlocal (Mod-2.2.4, cf. ***ANNEXE XII.2.4*** pour les résultats détaillés de modèle), on explique 37.8 % de la variabilité de l'EQR-IBGN$_{moyen}$. On n'améliore donc pas la capacité explicative du modèle précédent, probablement du fait qu'on ne peut pas expliquer plus de 38 % de la variabilité de l'EQR-IBGn$_{moyen}$ avec uniquement des variables d'occupation du sol sur ce jeu de donnée. Par contre, cette voie d'amélioration permet d'accroitre significativement le nombre de sources de pressions identifiées comme significatives à l'échelle du CRlocal : le modèle utilise à l'échelle du corridor rivulaire local 4 variables explicatives de plus que modèle précédent (10 contre 6).

Au niveau des sources de pressions dégradantes, l'utilisation de la donnée THRS dans le modèle confirme les effets observés dans les modèle précédents (effet de l'urbain et des zones industrielles commerciales) et met en évidence une nouvelle force motrice, les infrastructures routières (routes départementales précisément) sur le lit majeur et au-delà (bande de 250 m de part et d'autre du cours d'eau) et à proximité de la station (400 m en amont). Ces infrastructures linéaires dans ces espaces sont à l'origine de modifications hydromorphologiques conséquentes appauvrissant la qualité des habitats et la qualité fonctionnelle du corridor fluvial (cf. IV.2.1.2). La mise en évidence de cette nouvelle source de pression ayant un impact sur EQR-IBGN$_{moyen}$ constitue un progrès dans le diagnostic de l'état écologique : elle améliore l'information des gestionnaires pour leurs prises de décision sur les stratégies de restauration et/ou de préservation.

Au niveau des sources de pressions protectrices, l'utilisation de la THRS dans le modèle vient affiner en partie les résultats obtenus précédemment. Elle confirme l'effet bénéfique de prairies dans une emprise large (bande de 200 m de chaque côté du cours d'eau sur 3 km en amont de la station) en précisant leur nature (prairies « permanentes »). Elle met en lumière l'effet protecteur (barrière contre la pollution diffuse, stabilisation des berges, zones de refuges, diversification des habitats…) de la végétation arborée en bordure du cours d'eau (les ripisylves), dans une emprise restreinte (bande de 20 m de part et d'autre du cours d'eau sur 200 m en amont de la station). La mise en évidence de l'influence des ripisylves sur l'état écologique (EQR-IBGN$_{moyen}$) confirme le rôle important de ces formations végétales, signalé dans le cadre d'études plus locales mais qui n'avait pas été confirmé à large échelle.

D'autres effets sont mis en évidence à cette échelle : des effets cohérents comme l'impact positif des gels de terre (composés de végétation herbacée et arbustive semi-naturelle) dans une bande de 25 m de part et d'autre du cours d'eau sur 200 m en amont de la station, qui peut jouer un rôle de filtre en freinant l'écoulement au même titre que les prairies ; des effets plutôt contradictoires des plans d'eau, des labours et des cultures annuelles. Ces effets sont discutés dans le § IV.2.2.3 .

Enfin dans un troisième temps on intègre la troisième échelle de fonctionnement de l'écosystème rivière, le CRamont (échelle intermédiaire entre le bassin versant et le CRlocal), en utilisant les deux voies d'amélioration (Mod-3.1.4, cf. ***ANNEXE XII.3.4*** pour les résultats détaillés de modèle). On explique 37.8 % de la variabilité de l'EQR-IBGN$_{moyen}$. On n'améliore donc pas la capacité explicative du modèle précédent (ce qui renforce l'idée qu'on ne peut pas expliquer plus de 38 % de la variabilité de l'EQR-IBGn$_{moyen}$ avec uniquement des variables d'occupation du sol sur ce jeu de donnée) mais l'ajout de cette échelle permet d'accroitre significativement le nombre de sources identifiées comme ayant un impact significatif, puisqu'on passe à 23 variables significatives contre 14 variables dans le modèle précédent.

CemOA : archive ouverte d'Irstea / Cemagref

L'introduction de cette échelle ne modifie pas les effets observés à l'échelle du BV et du CRlocal. A l'échelle du CRamont, on retrouve l'effet dégradant des routes (bande de 150 m) et du tissu urbain discontinu (bande de 300 m). Un impact positif des prairies est également mis en évidence le long du réseau à l'amont des stations sur la bande de 20 m de part et d'autre du cours d'eau. L'identification de cet effet sur l'état écologique (EQR-$IBGN_{moyen}$) à large échelle confirme le rôle des bandes enherbées à proximité du cours d'eau observé dans le cadre d'études plus locales (e.g., Patty *et al.*, 1997 ; Schmitt *et al.*, 1999). Elle conforte l'idée que le corridor rivulaire constitue un espace clé de la restauration, notamment par la présence d'une végétation herbacée le long du cours d'eau. D'autres effets protecteurs sont identifiés. Ils sont discutés dans le § IV.2.2.

Ainsi, les résultats obtenus sur cette zone d'étude (et sur ce jeu de données de 155 stations), en utilisant comme réponse biologique les macroinvertébrés benthiques (EQR-$IBGN_{moyen}$), montrent que le mode spatialisation « multiple » et la donnée à THRS améliorent très significativement les performances des modèles empiriques pressions/impacts au niveau du corridor rivulaire puisqu'ils :

(1) fiabilisent la compréhension de ces relations (augmentation du pouvoir explicatif) ;

(2) précisent pour chaque source de pressions la localisation des emprises influentes (sur quelle largeur de bande rivulaire et à quelle distance de la station ?)

(3) identifient de nouvelles forces motrices (e.g., les routes) et en précisent certaines (e.g, les prairies permanentes) ;

(4) mettent en évidence à large échelle l'influence de la végétation arborée (ripisylves) et herbacée (bandes enherbées) le long du cours d'eau.

D'autres critères seront pris en compte ultérieurement afin de mieux cerner l'apport de ces voies d'améliorations, notamment des critères permettant de quantifier leur apport dans la différenciation des effets aux différentes échelles (cf. III.2).

II.2.2. Choix du modèle optimal

La deuxième analyse comparative menée dans cette démarche (cf. II.1.2.2) a permis de sélectionner le modèle considéré comme optimal pour établir un diagnostic de l'état écologique afin d'orienter les gestionnaires dans l'élaboration de stratégies de restauration.

Le modèle choisi correspond au modèle tri-échelle 3.1.4 car il améliore, par rapport aux autres modèles expliquant le même pourcentage de variabilité de la réponse biologique, l'identification et la localisation des sources de pressions influençant l'EQR-$IBGN_{moyen}$ aux différentes échelles de fonctionnement de l'écosystème rivière.

Ce modèle utilise bien évidemment les deux voies d'amélioration (mode de spatialisation « multiple » et la donnée à THRS) à l'échelle du corridor rivulaire. Il sera donc utilisé pour identifier les facteurs dégradants et protecteurs de l'état écologique sur la zone d'étude (cf. IV.2).

Cependant, la méthode heuristique mise en place pour sélectionner ce modèle peut être discutée et critiquée. D'autres critères doivent être considérés pour juger la pertinence d'un modèle pressions/impacts, notamment des critères permettant de mieux prendre en compte les dépendances d'échelles (i.e., le rôle respectif des pressions aux différentes échelles du paysage). Ce point est discuté dans la partie III.2 de ce chapitre.

CemOA : archive ouverte d'Irstea / Cemagref

II.2.3. Approche de la fiabilité

La démarche de modélisation a été menée à partir d'un jeu de données (notre population statistique) composé des 155 stations existant sur la zone d'étude (stations IBGN disponibles à partir du réseau de surveillance RNB avec une donnée THRS sur leur corridor rivulaire amont). Toutes les analyses de significativité des variables explicatives reposent sur cet échantillon. La fiabilité des conclusions tirées des modèles, pour les objectifs méthodologiques comme pour l'objectif thématique, dépend de leur robustesse, c'est-à-dire de la sensibilité de leurs structures (variables explicatives) et de leurs paramètres à la population statistique et aux erreurs de mesure.

Trois sources de perturbations peuvent conditionner la robustesse des modèles : (1) pour un jeu de données fixé, les perturbations liées au poids des individus ou à une structure particulière du jeu de données; (2) les perturbations liées aux erreurs de mesures possibles sur les variables utilisées dans le jeu de donnée ; (3) l'effectif du jeu de données (nota : (1) et (3) sont bien sûr liés).

(1) Les modèles sont jugés robustes vis-vis du premier type de perturbations car dans leur conception (cf. Chapitre 9 :III.3.2) une validation croisée est réalisée pour confirmer le choix du nombre de composantes de la régression PLS et un jacknife est effectué afin d'identifier les prédicteurs significatifs. Le jacknife permet en même temps de quantifier l'incertitude des prédicteurs vis-vis de ce type de perturbation puisque il donne l'intervalle de confiance à 95 % des prédicteurs significatifs.

Il est par ailleurs possible d'analyser la structure spatiale des résidus d'un modèle (structure géographique ou structure en fonction du rang de Strahler) pour identifier d'éventuels biais liés à la structure des données

(2) Par contre, l'incertitude des résultats de modélisation en fonction des erreurs de mesures des différentes variables (occupation du sol et IBGN) n'est pas prise en compte et demande des analyses complémentaires (ce point est discuté dans la partie III.1.2 de ce chapitre).

(3) Enfin, pour confirmer tant l'apport des voies d'améliorations (THRS et mode de spatialisation multiple) dans l'étude des relations pressions/EQR-IBGNmoyen que le choix du modèle pour diagnostiquer l'état écologique sur la zone d'étude, il est indispensable de répéter cette démarche sur un jeu de données différent.

L'idéal serait de compléter le jeu des 155 stations par de nouveaux sites IBGN sur la zone d'étude (par exemple les sites IBGN de l'Agence Seine-Normandie non incluses dans le réseau de surveillance du RNB) afin d'améliorer la représentativité du jeu de données vis-à-vis des pressions qui s'exercent sur les macroinvertébrés. Mais aucune donnée de station supplémentaire n'a pu être récoltée durant la thèse.

De façon à mettre en évidence l'intérêt d'une telle étude, à titre d'exemple les modèles ont été reconstruits à partir d'un nouveau jeu de données composé de 100 stations tirées de façon aléatoire parmi les 155 stations disponibles sur la zone d'étude. La réduction de l'effectif réduit potentiellement la variabilité du jeu de donnée et affaiblit la pertinence statistique de la démarche, mais cette analyse réalisée à titre d'exemple permet d'éclairer la stabilité des résultats.

La Figure 10.15 présente l'évolution de la capacité explicative (R^2) des différents modèles construits sur le jeu de donnée de 155 stations et sur le jeu de donnée de 100 stations.

On observe tout d'abord que la capacité explicative est plus élevée pour les modèles élaborés à partir du jeu de donnée de 100 stations, ceci quelque soit l'hypothèse de construction. Le jeu de donnée de 100 stations, possédant une variabilité plus réduite, semble plus facilement expliqué par le jeu de prédicteurs. On constate toutefois, selon la structure de modèle considérée, de fortes différences d'évolution de cette capacité explicative en fonction du jeu de données (100 ou 155 stations).

A partir de cet exemple, on montre donc l'intérêt de réitérer une telle démarche (test sur différents jeux de données) afin de confirmer nos observations sur la robustesse des modèles au jeu de données. Le jeu de données devra être suffisamment grand pour être représentatif des différentes pressions sur la zone d'étude.

CemOA : archive ouverte d'Irstea / Cemagref

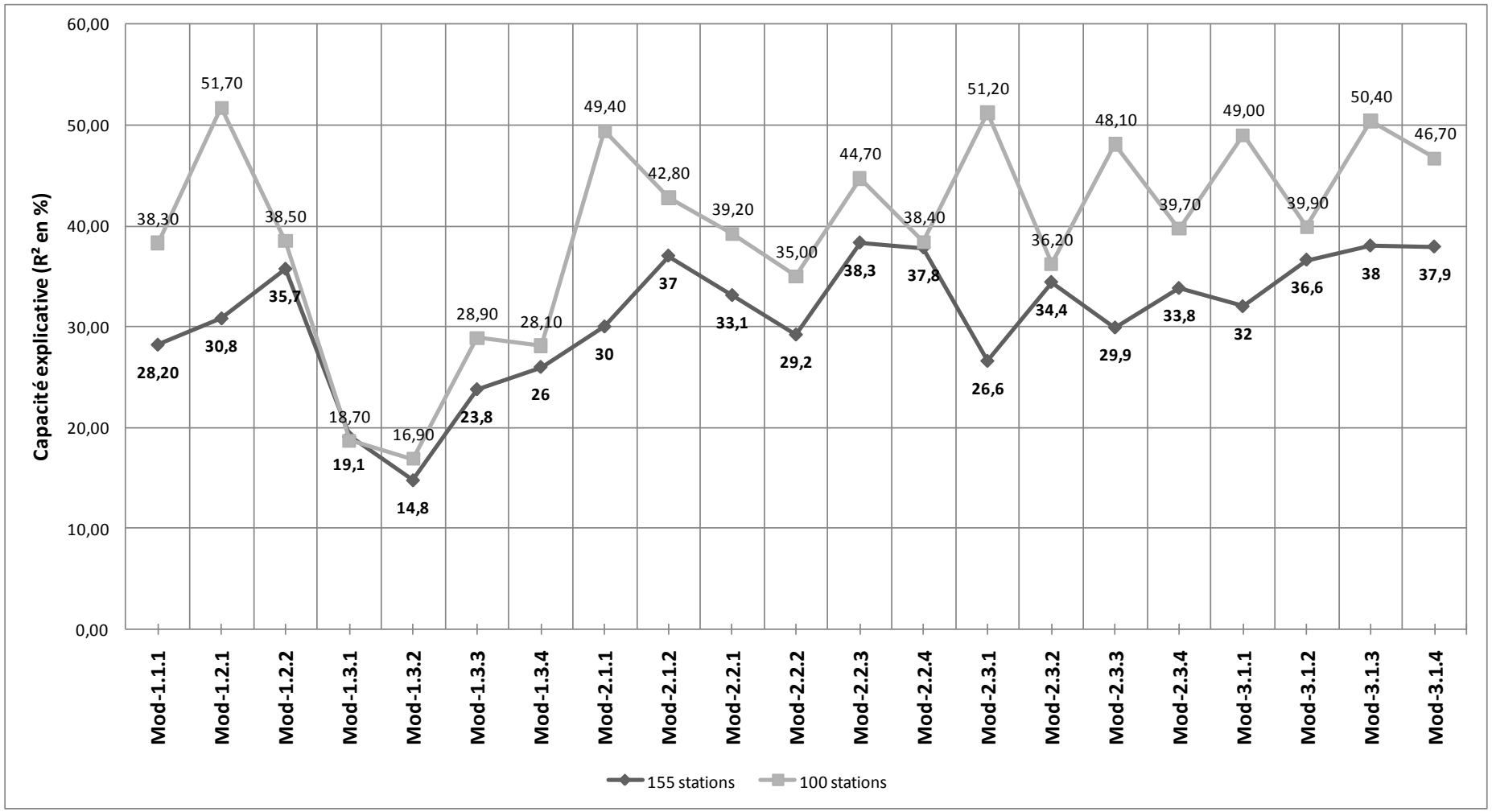


Figure 10.15 : Évolution de la capacité explicative(R^2) en fonction des différentes hypothèses de construction des modèles construits sur 155 stations (ligne foncé) et sur 100 stations (ligne claire)

CemOA : archive ouverte d'Irstea / Cemagref

III. Analyse critique des méthodes

L'approche mise en place dans cette étude a permis de donner une première réponse aux objectifs méthodologiques (apport des voies d'améliorations) et de sélectionner le modèle optimal afin d'établir un premier diagnostic de l'état écologique pour répondre aux objectifs opérationnels de l'étude. Cependant, cette approche doit être analysée de façon critique afin de mieux cerner ses limites et les alternatives ou voies d'amélioration envisageables. Nous menons ci-dessous successivement une analyse critique de la méthode de construction des modèles pressions/impacts (cf. III.1) et de la méthode de comparaison des modèles (cf. III.2).

III.1. Méthode de construction des modèles pressions/impacts

III.1.1.Sélection de l'emprise spatiale significative

La méthode mise en place dans cette étude pour construire un modèle pressions/impacts est basée sur deux étapes (cf. Chapitre 9 :III.) :

- la 1ère étape sélectionne pour chaque source de pression à une échelle donnée du corridor rivulaire (CRlocal ou CRamont) l'emprise spatiale significative ;
- la 2ème étape cherche, parmi les sources de pressions caractérisées sur leurs emprises significatives sélectionnées à l'étape 1, la combinaison de source de pressions expliquant le mieux la variabilité de la réponse biologique.

Dans notre étude, la 1ère étape a été mise en œuvre uniquement aux échelles du corridor rivulaire lorsque le mode de spatialisation « multiple » a été utilisé : ce mode propose plusieurs emprises spatiales (buffers) pour étudier l'effet d'une source de pressions sur l'état écologique.

On adopte comme emprise spatiale significative celle au sein de laquelle la source de pression considérée possède l'influence la plus significative sur l'état écologique. Pour la sélectionner, on se base sur une analyse de corrélations mono-variable : le degré de corrélation est estimé à partir du coefficient de corrélation (R^2) de la régression linéaire simple entre le pourcentage de la source de pression sur chacune des emprises spatiales candidates et la variable biologique. L'emprise sur laquelle la source de pressions possède le plus fort lien (plus grand R^2) avec la variable biologique est sélectionnée comme « emprise significative ».

Cette approche mono-variable par laquelle les pressions sont analysées séparément présente des inconvénients. On sait en effet que les sources de pressions agissent de façon « combinée » sur l'état écologique : cet effet « combiné » des sources de pressions n'est pas pris en compte par cette méthode de sélection. L'emprise spatiale significative pour une source de pressions donnée devrait dans l'absolu correspondre à l'emprise sur laquelle cette source de pression combinée aux autres (sur leurs emprises respectives) a la plus forte influence sur l'état écologique. Sa sélection impliquerait donc une analyse de corrélations multi-variables qui identifierait directement la combinaison de source de pressions et de leurs emprises respectives, favorisant l'explication de la variabilité de la réponse biologique.

Le développement d'une telle méthode n'est pas simple compte tenu du nombre de combinaisons possibles qui correspond au nombre d'emprises spatiales candidates à une échelle donnée multiplié par le nombre de source de pressions. Par exemple, 8624 combinaisons sont possibles pour un modèle tri-échelle (BV, CRlocal, CRamont) utilisant la donnée THRS et le mode de spatialisation multiple dans le corridor rivulaire.

Plusieurs pistes peuvent être envisagées : soit adapter la méthode de régression PLS de façon à sélectionner parmi les emprises candidates (aux différentes échelles) la combinaison de sources de pressions significatives la plus explicative de la réponse biologique, soit rechercher d'autres algorithmes de fouilles de données conçus pour gérer cette quantité très importante de variables et qui sont adaptés aux problèmes de multicolinéarité entre les variables explicatives.

CemOA : archive ouverte d'Irstea / Cemagref

III.1.2. Évaluation des incertitudes des résultats de modélisation

Dans le § II.2.3, on a mis en évidence l'intérêt de reproduire la démarche sur un autre jeu de données de manière à confirmer nos conclusions et fournir aux gestionnaires une indication sur leur fiabilité.

Cette fiabilité dépend aussi de la robustesse des résultats de modélisation. Dans notre étude, les perturbations liées aux erreurs de mesures sur les variables utilisées dans le jeu de données (i.e, les indicateurs d'occupation du sol, les indicateurs EQR-IBGN$_{moyen}$), qui peuvent conditionner les résultats de modélisation (la structure et les paramètres du modèle), n'ont pas été prises en compte.

Les méthodes privilégiées pour étudier la robustesse vis-vis de ces perturbations sont les techniques de ré-échantillonnage telles que le boostrap, le jacknife, la validation croisée ou encore les simulations de Monte-Carlo. Elles consistent à répéter (des centaines ou même des milliers de fois, de manière à obtenir une distribution des issues possibles), les analyses sur les différents échantillons simulés (par bruitage de l'échantillon initial) puis à étudier les fluctuations des résultats (les distributions de chaque paramètre du modèle). Les échantillons simulés sont construits aléatoirement en fonction de la distribution des erreurs de chaque variable. Ces méthodes sont d'autant plus efficaces que l'on a une bonne connaissance des distributions des erreurs de chaque variable

Or, on a vu précédemment les contraintes rencontrées pour quantifier la distribution de l'erreur d'un indicateur construit à partir d'une carte d'occupation des sols, notamment à cause de la difficulté à spatialiser correctement les erreurs des cartes d'occupation du sol (cf. Chapitre 6 :III. et Chapitre 8 :IV.).

De la même façon, les incertitudes des indicateurs biologiques, liées à la subjectivité du protocole et aux circonstances particulières lors du prélèvement comme la présence d'un stade dormant chez l'espèce au moment de la récolte (Berryman, 1990) par exemple, ne sont pas encore correctement évaluées (Archaimbault and Dumont, 2010).

Ces difficultés devront être surmontées avant d'engager une analyse des incertudes des résultats de modélisation liées aux erreurs de ces indicateurs. Ces méthodologies posent également le problème de la combinatoire (des centaines de milliers de polygones ; 155 stations ;…) et des moyens de calcul à mettre en œuvre.

III.2. Méthode de comparaison des modèles

Dans cette étude une méthode de comparaison de modèles a été mise en place pour (i) évaluer l'apport des voies d'améliorations (données à THRS et mode de spatialisation « multiple ») et (ii) sélectionner le modèle optimal parmi les modèles envisageables qui diffèrent en fonction de la combinaison d'échelles d'analyse, de la source de données d'occupation du sol (THRS ou CLC) et du mode spatialisation des sources de pressions dans le corridor rivulaire (« multiple » ou « fixe »).

Cette méthode se base sur trois critères évaluant la pertinence d'un modèle pressions/impacts pour établir un diagnostic de l'état écologique afin de cibler et de localiser les priorités d'actions. Le premier critère est relatif à la capacité explicative des modèles (R^2) car le modèle sera d'autant plus pertinent qu'il explique une part importante de la variabilité de la réponse biologique (modèle plus fiable). Le deuxième critère est relatif à la capacité descriptive (nombre de sources de pressions significatives) car le modèle sera d'autant plus pertinent qu'il identifie l'impact d'un large spectre de source de pressions. Le troisième critère s'intéresse à la nature et la localisation des sources de pressions identifiées car un modèle sera d'autant plus pertinent pour les gestionnaires que les facteurs identifiés sont de nature gérable.

Cependant d'autres critères doivent être considérés pour juger la pertinence d'un modèle pressions/impacts et notamment des critères permettant de mieux prendre en compte les dépendances d'échelles de façon (1) à quantifier l'influence relative des échelles sur la réponse biologique et (2) à vérifier si l'effet d'une source de pressions (i.e. catégorie d'occupation du sol) observée à une échelle donnée est propre à cette échelle.

CemOA : archive ouverte d'Irstea / Cemagref

(1) La mise en œuvre d'un modèle pressions/impacts étudiant les relations à plusieurs échelles permet à la fois de détecter la gamme d'échelles pour laquelle un facteur est explicatif de la réponse biologique et d'identifier un ensemble de facteurs explicatifs à une échelle donnée. Mais il ne permet pas de quantifier la part de variance expliquée par l'ensemble des facteurs d'une échelle donnée. Or cette information est essentielle pour les gestionnaires afin de prioritiser les stratégies de restauration entre les différentes échelles.

Des techniques statistiques robustes de partitionnement de la variance sont disponibles aujourd'hui pour quantifier l'influence relative des facteurs aux différentes échelles. Elles ont été mises en place dans diverses études de modélisation pressions/impacts (e.g., Richards *et al.*, 1997 ; Johnson *et al.*, 2004 ; Johnson *et al.*, 2007 ; Stoffels *et al.*, 2005 Gido *et al.*, 2006 ; Kratzer *et al.*, 2006). Une autre approche consiste à construire séparément des modèles selon différentes combinaisons d'échelles et à comparer les résultats en se basant notamment sur la variance expliquée (Lattin *et al.*, 2004, Dow *et al.*, 2006).

La Figure 10.14, qui présente l'évolution de la performance des modèles pressions/impacts au niveau du corridor rivulaire en France permet de comparer indirectement les modèles en fonction de leur capacité explicative. Elle montre que les indicateurs du CRlocal apportent, quelque soit la source de donnée d'occupation du sol et le mode de spatialisation, une information supplémentaire améliorant l'explication des relations par rapport aux seuls indicateurs du BV puisqu'on observe un gain significatif (de +5 à +10 % environ) de la capacité explicative entre le modèle BV et chacun des trois modèles BV/CRlocal (qui diffèrent en fonction de la source de donnée et du mode de spatialisation).

Par contre, elle ne permet pas de quantifier l'apport des indicateurs à l'échelle du CRamont puisque dés l'ajout de l'échelle du CRlocal on atteint le maximum de variance qu'on peut expliquer avec les indicateurs d'occupation du sol. Est-ce que les indicateurs du CRamont par rapport au CRlocal, au BV et à la combinaison BV/CRlocal apportent une information complémentaire améliorant l'explication de variabilité de la réponse biologique ? Est-ce que cet apport dépend de la source de donnée d'occupation du sol et/ou du mode de spatialisation ? Des questions auxquelles il faudrait répondre pour évaluer correctement l'apport des voies d'amélioration, mais surtout sélectionner le modèle le plus pertinent pour diagnostiquer l'état écologique.

(2) De plus, compte tenu des similarités entre les compositions d'occupation du sol entre les différentes échelles, notamment avec la donnée CLC (cf. I.2.2.1.1), il est indispensable de vérifier si l'effet d'une source de pressions (i.e. catégorie d'occupation du sol) observée à une échelle donnée est propre à cette échelle, particulièrement lorsqu'une source de pressions est identifiée sur plusieurs échelles avec la même amplitude.

Cette étude implique, pour une source de pressions donnée, d'étudier les corrélations de son pourcentage aux différentes échelles, puis de vérifier son apport relatif vis-à-vis de la réponse biologique aux différentes échelles.

Dans notre étude, par exemple, on retrouve l'effet négatif du tissu urbain discontinu (poste 112) et positif des prairies (poste 231) au niveau des trois échelles d'analyses pour les modèles utilisant le mode de spatialisation « multiple », ceci quelque soit la source de donnée d'occupation du sol. La Figure 10.16 présente, pour chacune de ces deux sources de pressions quantifiées à partir de chaque donnée d'occupation du sol (CLC et THRS), une ACP (analyse en composantes principales) construite à partir de l'observation des pourcentages au niveau des trois échelles d'analyses (BV, CRamont et CRlocal) sur les 155 stations IBGN disponibles sur la zone d'étude. L'ACP permet de visualiser si l'information d'une source de pression est différente (décorrélée) en fonction de l'échelle d'analyse.

Les quatre ACP obtenues sont significatives puisque elles permettent de représenter chacune plus de 90 % de la variabilité de leur données avec un % plus important apporté par leur premier axe (plus de 70 % environ).

On observe tout d'abord, pour le tissu urbain discontinu, que quelque soit la donnée d'occupation du sol, l'information au niveau du CRamont et l'information au niveau du CRlocal sont très proches (très corrélées), alors qu'elles sont très différentes de l'information à l'échelle du BV.

Pour les prairies, on observe qu'avec la donnée CLC l'information est sensiblement proche entre le CRlocal et le CRamont mais différente sur le BV. En revanche avec la donnée THRS, l'information est sensiblement proche entre le BV et le CRamont mais différente sur le CRlocal. Cet exemple montre bien que pour une source de pressions donnée il existe des corrélations fortes entre les pourcentages aux différentes échelles, qui peuvent varier en fonction de la donnée d'occupation du sol utilisée (exemple des prairies).

CemOA : archive ouverte d'Irstea / Cemagref

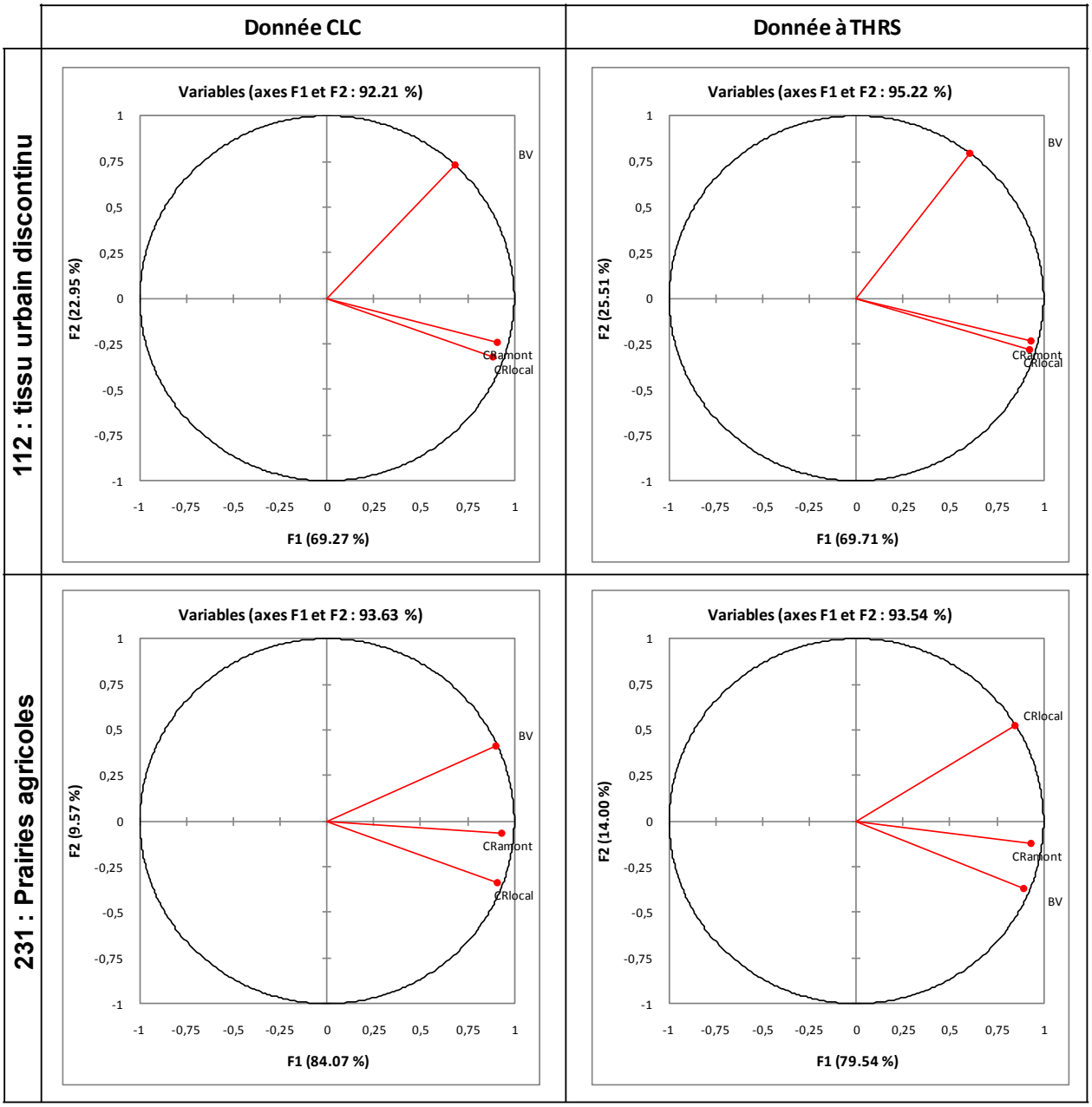


Figure 10.16 : Résultats des l'ACP construites à partir de l'observation des pourcentages au niveau des trois échelles d'analyses (BV, CRamont et CRlocal) sur les 155 stations IBGN disponibles sur la zone d'étude pour le tissu urbain discontinu (poste 112) et les prairies (poste 231) quantifiés d'une part avec la donnée CLC et d'autre part avec la donnée à THRS.

Mais quel est l'apport relatif, sur l'explication de variabilité de la réponse biologique, des pourcentages corrélés entre les échelles ? Autrement dit, est-ce que les effets d'une source de pression observés sur différentes échelles sont différents ? Est-ce que la donnée THRS et/ou le mode de spatialisation permettent d'améliorer la différenciation de ces effets ? Des questions auxquelles il faudrait également répondre pour, bien évidemment, évaluer correctement l'apport des voies d'amélioration, mais surtout pour sélectionner le modèle le plus pertinent pour diagnostiquer l'état écologique.

CemOA : archive ouverte d'Irstea / Cemagref

IV. Premier diagnostic de l'état écologique

Cette partie présente un diagnostic de l'état écologique réalisé à partir du modèle identifié lors de la mise en œuvre de la démarche de modélisation. Ce diagnostic donne un premier aperçu des principaux facteurs (anthropiques et semi-naturels) influençant les macroinvertébrés benthiques. Cependant, on devra bien garder à l'esprit que (i) ce diagnostic n'est pas représentatif de l'ensemble des sources de pressions s'exerçant sur les cours d'eau puisque il est réalisé uniquement sous le filtre des macroinvertébrés benthiques à l'aide d'un jeu de données de 155 stations IBGN sur la zone d'étude ; (ii) qu'il rend compte uniquement des causes d'altérations et des facteurs protecteurs pour des systèmes de petite et moyenne taille (qui représentent 94 % du réseau de drainage de la zone d'étude) ; et (III) qu'il est effectué uniquement à partir de l'occupation du sol, une des forces motrices générant la plupart des impacts sur le milieu aquatique, mais qui ne rend pas compte de nombreuses pressions comme les barrages, les dérivations, la navigation, l'endiguement issus d'autres forces motrices (énergie, transport, tourisme) altérant particulièrement les processus hydromorphologiques. On devra également interpréter les résultats avec une certaine prudence compte tenu (i) que la fiabilité des résultats vis-vis de l'incertitude sur les indicateurs n'a pas été évaluée ; et (ii) que les dépendances d'échelles, responsables des corrélations entre l'effet des facteurs à différentes échelles sur la variabilité de la réponse biologique et conditionnant l'effet propre d'un facteur à une échelle donnée, n'ont pas été pleinement analysées.

IV.1. Présentation du modèle « optimal »

Le modèle qui a été jugé le plus pertinent selon la démarche de modélisation, examine les relations pressions/impacts (plus précisément les relations forces motrices / EQR-IBGN$_{moyen}$) qui interagissent sur les trois échelles de fonctionnement de l'écosystèmes rivière (BV, CRamont et CRlocal) en utilisant, au niveau des échelles du corridor rivulaire, l'emprise spatiale la plus significative pour chaque source de pressions (emprise sur laquelle une source de pressions donnée possède le lien le plus fort avec l'EQR-IBGN$_{moyen}$) et en quantifiant le pourcentage de ces sources au sein des emprises avec la donnée THRS.

La procédure de modélisation a donc débuté par la construction d'un modèle PLS utilisant 107 prédicteurs (indicateurs) pour obtenir un modèle au final expliquant 37.9 % de la variabilité de l'EQR-IBGN$_{moyen}$ à partir de la combinaison de 23 sources de pressions significatives à ces échelles. Les résultats intermédiaires de la procédure de modélisation sont résumés dans l'***ANNEXE XII.3.***4

La Figure 10.17 propose une illustration de la contribution (positive ou négative) de ces 23 sources de pressions sur l'état écologique (EQR-IBGN$_{moyen}$).

Le Tableau 10-9 présente, par ordre d'influence, les coefficients de la régression PLS obtenus dans le modèle final et les statistiques associées.

On se réfèrera, dans la partie suivante, à cette figure et à ce tableau afin de hiérarchiser et identifier les principaux facteurs influençant l'état écologique (vu sous le filtre des macroinvertébrés benthiques) sur la zone d'étude.

CemOA : archive ouverte d'Irstea / Cemagref

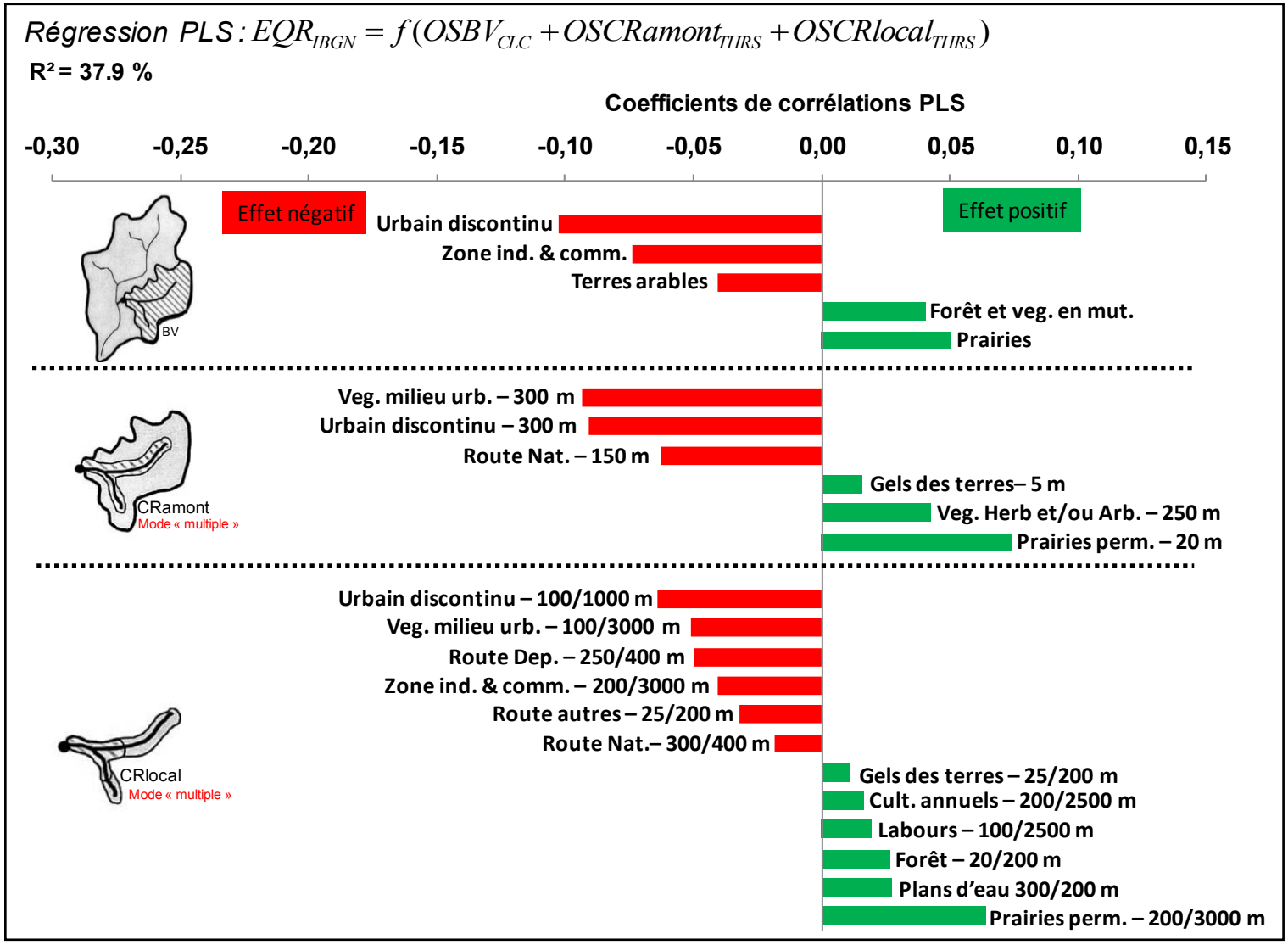


Figure 10.17 : Contribution des sources de pressions à l'échelle du BV, CRamont et du CRlocal qui agissent de façon combinée sur l'état écologique des cours d'eau (EQR-IBGN$_{moyen}$) mesurés sur les 155 stations de la zone d'étude. Ces sources expliquent 33.8 % de la variabilité de cet état. Le modèle a été conçu en utilisant le mode de spatialisation « multiple » et la donnée THRS au niveau du corridor rivulaire.

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 10-9 : Présentation par ordre d'influence des coefficients de corrélations PLS et leur statistiques (écart-type et Intervalle de confiance à 95 %) des variables significatives obtenus pour le modèle BV/CRamont/CRlocal utilisant l'emprise spatiale CRlocal-opt et la donnée THRS au niveau du corridor rivulaire. En rose pale les variables possédant un coefficient de régression négatif et en vert pale celles qui possèdent un coefficient de régression positif. Pour faciliter la lecture de l'analyse des pressions influençant l'état écologique (partie IV.2), la dernière colonne fait référence aux numéros de paragraphes où sont discutés chacune de ces variables.

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)	Référence aux num. de §
1	BV-112	-0,103	0,010	-0,122	-0,083	IV.2.1.1.
2	CR-300-amont-113	-0,093	0,022	-0,136	-0,051	IV.2.1.1.
3	CR-300-amont-112	-0,091	0,024	-0,137	-0,044	IV.2.1.1.
4	CR-20-amont-231	0,075	0,014	0,047	0,102	IV.2.2.1.
5	BV-121	-0,074	0,022	-0,118	-0,030	IV.2.1.1.
6	CR-100-1000-112	-0,064	0,026	-0,116	-0,013	IV.2.1.1.
7	CR-200-3000-231	0,064	0,023	0,019	0,109	IV.2.2.1.
8	CR-150-amont-152	-0,063	0,020	-0,101	-0,025	IV.2.1.2.
9	CR-100-3000-113	-0,051	0,021	-0,093	-0,009	IV.2.1.1.
10	BV-231	0,050	0,022	0,007	0,093	IV.2.2.1.
11	CR-250-400-153	-0,050	0,006	-0,061	-0,038	IV.2.1.2.
12	CR-250-amont-321	0,042	0,013	0,016	0,068	IV.2.2.1.
13	BV-211	-0,041	0,018	-0,077	-0,005	IV.2.1.3.
14	BV-243	0,041	0,020	0,001	0,080	IV.2.2.2.
15	CR-200-3000-121	-0,040	0,009	-0,057	-0,024	IV.2.1.1.
16	CR-150-300-154	-0,032	0,007	-0,045	-0,018	IV.2.1.2.
17	CR-300-200-412	0,027	0,009	0,010	0,045	IV.2.2.3.
18	CR-20-200-311	0,027	0,014	0,002	0,054	IV.2.2.2.
19	CR-100-2500-211	0,020	0,004	0,012	0,028	IV.2.2.3.
20	CR-300-400-151	-0,019	0,008	-0,035	-0,002	IV.2.1.2.
21	CR-200-2500-214	0,017	0,007	0,002	0,031	IV.2.2.3.
22	CR-5-amont-252	0,015	0,007	0,001	0,030	IV.2.2.3.
23	CR-25-200-251	0,011	0,005	0,001	0,021	IV.2.2.3.

CemOA : archive ouverte d'Irstea / Cemagref

IV.2. Analyse des sources de pressions influençant l'état écologique

IV.2.1.Les sources de pressions dégradantes

IV.2.1.1 L'impact de l'urbain discontinu et des zones industrielles & commerciales

On s'aperçoit à la lecture du Tableau 10-9 et en visualisant la Figure 10.18 que le tissu urbain discontinu (poste 112 ou 113[50] ; facteurs 1, 2, 3, 6 et 9 du Tableau 10-9) est le facteur influençant le plus fortement l'EQR-$IBGN_{moyen}$, en premier lieu à l'échelle du BV puis à l'échelle du CRamont sur le lit majeur du cours d'eau et au-delà (300 m de part et d'autre du cours d'eau). Il est également le facteur de pression le plus influant, quoique de façon moins marquée, à l'échelle du CRlocal.

Cet effet prépondérant du tissu urbain discontinu sur les communautés de macroinvertébrés a été observé sur la plupart des modèles mis en place sur les différentes hydro-écorégions françaises et à l'échelle nationale (Wasson et al., 2005) ainsi que dans d'autres pays européens (Wasson et al., 2010) et sur différents bassins versants américains (e.g., Stepenuck *et al.*, 2002 ; Wang *et al.*, 2003 ; Roy *et al.*, 2003a ; Snyder *et al.*, 2005 ; Potter *et al.*, 2005 ; WALSH *et al.*, 2007) aussi bien à l'échelle du BV qu'au niveau du corridor rivulaire.

Les impacts liés à la présence de tissus urbains discontinus (au même titre que du tissu urbain continu) sont bien reconnus dans la littérature (Paul and Meyer, 2001 ; Wang *et al.*, 2003 ; Allan, 2004a). La présence de ces tissus peut modifier à la fois les processus (1) biogéochimiques, (2) hydromorphologiques et (3) écologiques de l'écosystème rivière.

(1) La présence de ces tissus s'accompagne généralement de rejets directs traités (station d'épuration) ou non traités (déversoirs d'orages) et de rejets diffus (lessivage des surfaces imperméables) qui sont sources de matières organiques (MO), de matières en suspension colloïdales ou microparticulaires (MES), de sédiments fins et de substances toxiques tels que les métaux. Les actions combinées de ces apports diminuent la capacité biogène des cours d'eau en déséquilibrant l'état trophique et chimique du système puis en altérant la structure des habitats aquatiques (paramètres clés de l'écosystème rivière).

Tout d'abord, l'apport excessif de MO par ces eaux d'égouts (i) déséquilibre la chaîne alimentaire (augmentation des producteurs et des décomposeurs) ; (ii) appauvrit considérablement le milieu récepteur en oxygène (demande en oxygène forte par les bactéries aérobies dégradant la MO) ; (iii) augmente, avec les MES, la turbidité qui diminue l'activité photosynthétique (Graham, 1989) et accentue ainsi l'appauvrissement en oxygène dissous du milieu ; et (iv) augmente la quantité de polluants toxiques compte tenu de l'affinité (capacité d'adsorption) des MO et MES pour ces polluants. Au niveau des macroinvertébrés benthiques (Vivier, 1970), l'apport de MO par les eaux d'égouts se traduit directement par une diminution de la diversité taxonomique avec une augmentation (prolifération) des décomposeurs (e.g., les Chironomidées, Oligochètes) et une perte des taxons sensibles (e.g., Plécoptères, Trichoptères à fourreau et Éphéméroptères) au profit des organismes adaptés notamment aux faibles teneurs en oxygène dissout (e.g., les Tubificidées, les Syrphidées).

Ensuite, l'augmentation de l'activité des bactéries aérobies, utilisant l'oxygène pour transformer la matière organique en ammonium, nitrites puis nitrates, accroit la présence de ces substances minérales dans le milieu récepteur. L'ammonium peut se transformer en ammoniac (gaz dissous), notamment dans les eaux alcalines (ce qui est le cas dans la zone d'étude car le socle des différents cours d'eau est majoritairement calcaire), qui avec les nitrites sont des éléments très toxiques pour la vie aquatique (effets létaux ou sous-létaux). De plus, lorsque des amoncellements de boues riches en MO se constituent au fond d'un cours d'eau, ces MO peuvent être décomposées par les bactéries en anaérobiose en produisant des dégagements gazeux putrides : méthane, gaz carbonique et hydrogène sulfuré, éléments également très nocifs pour la vie aquatique (Vivier, 1970). Par ailleurs, comme les macroinvertébrés benthiques vivent en relation intime avec les sédiments (adsorbant les substances toxiques), ils subissent de façon plus importante cette pollution (Woodcock and Huryn, 2007).

[50] Le poste « végétation urbaine » de la donnée THRS correspond à la végétation dans un périmètre de 5 m autour d'un objet classé « tissu urbain discontinu ». La présence de cette végétation (herbacée, arbustive ou arborée) généralement isolée (encerclée par des objets artificialisés) s'accompagne de façon évidente de la présence de tissu urbain discontinu. Ceci explique que dans le modèle l'effet de la végétation urbaine est identique à l'effet d'un tissu urbain discontinu.

CemOA : archive ouverte d'Irstea / Cemagref

Enfin, les MO donnent également des dépôts vaseux qui participent au colmatage du substrat avec les sédiments fins issus principalement des déversoirs d'orages (Tetzlaff *et al.*, 2005). Le colmatage qui homogénéise la structure de l'habitat et réduit la teneur en oxygène dans les sédiments (par la réduction des échanges entre la surface et le milieu interstitiel) modifie la diversité taxonomique, influence la survie et certains stades de développement des macroinvertébrés et modifie l'abondance, la biomasse, la productivité du peuplement (Gayraud *et al.*, 2002). Le colmatage est un risque important au niveau de la zone d'étude (Chandesris *et al.*, 2009), accentué par la topographie plane et les faibles débits naturels caractérisant cette zone. Ces MO sont aussi à l'origine d'autres perturbations physiques mais qui agissent de façon directe sur les macroinvertébrés, par exemple, les particules de MO entrent dans le fourreau des larves des trichoptères et empêchent ainsi une circulation adéquate de l'eau ce qui réduit les échanges d'oxygène entre l'eau et l'hémolymphe du trichoptère (Borror and White, 1999.).

(2) Les altérations hydromorphologiques générées par la présence de tissu urbain discontinu ou continu sont multiples. Les actions combinées de ces altérations diminuent la capacité biogène des cours d'eau en homogénéisant la structure des habitats aquatiques et en modifiant le fonctionnement de l'écosystème rivière.

L'imperméabilisation des sols, par exemple, est à l'origine d'une augmentation des débits des crues fréquentes (Arnold and Gibbons, 1996) provoquant une perturbation de l'équilibre débit liquide/débit solide : l'augmentation du débit liquide associé à un maintien, voire une réduction du débit solide induit des phénomènes importants d'érosion, et souvent une incision du cours d'eau avec des conséquences sur les habitats aquatiques par l'augmentation fréquente des contraintes physiques (vitesses, hauteurs d'eau) et par la réduction des alternances naturelles « radiers/mouilles » à l'origine d'une banalisation des habitats et de déconnexions des annexes quand elles existent (Chandesris *et al.*, 2008).

Ce phénomène est accentué par la présence des déversoirs d'orages qui d'une part diminuent le temps de concentration des eaux et donc contribuent à l'accentuation des pics de crues (Tetzlaff *et al.*, 2005) et d'autre part participent à un accroissement brutal des forces de cisaillement, notamment sur le fond du lit, entraînant généralement une incision à l'aval des points de rejets (Borchardt and Sperling, 1997).

S'ajoutent à ces altérations hydrologiques, les altérations d'ordre morphologique. Généralement, les espaces artificialisés s'accompagnent d'aménagements et d'ouvrages conçus essentiellement pour la protection des riverains, l'assainissement et la valorisation des sites piscicoles (Wasson *et al.*, 1998). Les digues, les talus dans le lit majeur ainsi que les autres ouvrages de chenalisation (i) induisent un blocage de la dynamique latérale et une déconnexion entre lit mineur et lit majeur ; (ii) provoquent l'augmentation de débit « de plein bord » ; (iii) réduisent voire suppriment la végétation rivulaire et la sinuosité naturelle du cours d'eau. Autant de conséquences qui participent (Wasson *et al.*, 1998), à l'uniformisation des habitats et à la modification du fonctionnement (des paramètres clés de l'écosystème) altérant in fine la vie aquatique. Par exemple, si les bandes riveraines ont été détruites par l'activité humaine, la température du cours d'eau augmente, le niveau d'oxygène baisse et la quantité de débris végétaux diminue, ce qui a pour effet de favoriser les espèces tolérantes (Camargo *et al.*, 2004), ceci d'autant plus si le milieu subit une pollution organique.

(3) La présence de tissu urbain discontinu ou continu fragmente le paysage rivulaire. Cette fragmentation limite les processus de recolonisation des milieux et notamment chez les macroinvertébrés adultes (Petersen *et al.*, 1999 ; Petersen *et al.*, 2004).

Les zones industrielles et commerciales (poste 121 ; facteurs 5 et 15 duTableau 10-9) ont également un impact négatif sur l'EQR-IBGN$_{moyen}$ tant à l'échelle du BV qu'à celle du CRlocal à proximité de la station de prélèvement (200 m en amont). Cet impact est relativement plus faible que le tissu urbain discontinu sur les deux échelles. Ces espaces participent bien évidement aux altérations des processus hydromorphologiques et écologiques mais accentuent surtout l'impact existant des rejets des tissus urbains sur les communautés de macroinvertébrés benthiques par les modifications des flux biogéochimiques locaux.

Les émissaires industriels peuvent être à la source de rejets mécaniques (particules de sable,...), thermique (rejets d'effluents de refroidissement,...), organiques (surtout par les industries agroalimentaires, les fromageries,...), métalliques (mercure, cadmium, chrome,...), chimiques (produits acides ou alcalins, solvants, hydrocarbures, produits organiques de synthèse,...) et radioactives (rejets d'éléments radioactifs).

Compte tenu de l'économie industrielle du bassin (Agence de l'eau Seine Normandie, 2004, et notamment le sous-bassin Seine-aval (une partie de notre zone d'étude), qui concerne aussi bien les filières situées en aval des productions (chimie de spécialités, automobile, etc.) que celles d'amont (énergie, chimie de base, sidérurgie, etc.), on a une forte probabilité de retrouver ces différents types de rejets dans le milieu récepteur.

CemOA : archive ouverte d'Irstea / Cemagref

Selon l'Agence Seine Normandie (2004), l'industrie représente une part importante des rejets en matière organique (MO) et environ 90% des métaux toxiques. Les rejets toxiques proviennent principalement des industries électroniques, de l'ensemble sidérurgie-métallurgie-fonderie, des usines de traitement des déchets, des activités d'assemblage et de l'imprimerie.

Ainsi, les multiples pressions générées par ces occupations du sol (tissu urbain et zones industrielles et commerciales) peuvent altérer le fonctionnement et la structure de l'écosystème en agissant de façon combinée. Ils peuvent porter atteinte également directement ou indirectement à la vie aquatique en libérant dans le milieu récepteur des substances toxiques. Ces altérations se traduisent, au niveau des invertébrés par une diminution de la diversité taxonomique et une perte de taxons polluosensibles. Parmi ces pressions, la pollution organique est plus fortement suspectée car l'IBGN répond plus favorablement à ce type de pression.

IV.2.1.2 L'impact des routes dans le corridor rivulaire

Les routes (nationales, départementales ou autres) présentes dans le lit majeur à l'échelle du corridor amont (150 m de part et d'autre du cours d'eau) ou à proximité de la station (de 200 à 400 m en amont de la station sur 25 à 300 m de part et d'autre du cours d'eau) à l'échelle du CRlocal ont un impact négatif sur l'EQR-$IBGN_{moyen}$ ». Elles ont un impact relativement élevé, notamment pour les routes départementales (β = -0.063), mais inférieur à l'effet du tissu urbain discontinu aux différentes échelles. Cet impact sur les macroinvertébrés benthiques (ou sur d'autres communautés biologiques) n'a pas jusqu'à présent été mis en évidence à l'échelle régionale dans les études pressions/impacts notamment par l'absence de cette information dans les différents modèles développés.

Les routes situées dans cet espace participent, au même titre que le tissu urbain, à l'imperméabilisation des sols et à ses conséquences hydrologiques, mais génèrent également des modifications morphologiques lourdes. Leur présence s'accompagne d'ouvrages de stabilisation (pour leur propre protection ou pour la protection des zones urbaines qu'elles desservent) perturbant le fonctionnement hydraulique lors des crues morphogènes (Chandesris *et al.*, 2008) et se traduisant par un appauvrissement de la qualité d'habitats et de la qualité fonctionnelle du corridor fluvial (perte de la connectivité latérale).

La mise en évidence de cet impact, en plus de l'impact des tissus urbains, confirme que les macroinvertébrés sont également sensibles aux modifications de la structure de l'habitat aquatique.

IV.2.1.3 L'impact des terres arables

L'impact des terres arables (terres agricoles cultivées) (Postes 211; facteurs n°13 duTableau 10-9) est identifié par le modèle uniquement à l'échelle du BV. Cet impact est relativement plus faible que ceux du tissu urbain discontinu, des zones industrielles et commerciales puis des routes.

On retrouve dans les travaux menés en France au niveau des hydro-écorégions et au niveau national (Wasson *et al.*, 2005), sur d'autre pays européens (Wasson *et al.*, 2010), ainsi que sur des bassins versants américains (e.g., Roth *et al.*, 1996 ; Wang *et al.*, 1997 ; Allan *et al.*, 1997) l'impact des terres arables au niveau du BV de façon moins marquée que l'impact des terres artificialisés.

Les impacts liés à la présence des terres sont bien reconnus dans la littérature (Skinner *et al.*, 1997 et al. ; Allan, 2004a). La présence de ces terres cultivées peut modifier à la fois les processus (1) biogéochimiques et (2) hydromorphologiques de l'écosystème rivière.

(1) La présence de terres arables à l'échelle du BV est associée généralement à la pollution diffuse à l'origine d'apports de sédiments, de matières organiques (MO), de matières en suspension colloïdales ou microparticulaires (MES), de nutriments et de substances dangereuses (pesticides, insecticides, herbicides…). Ces apports modifient à la fois l'état trophique et chimique du système et altèrent la structure des habitats aquatiques (paramètres clés de l'écosystème rivière).

Par exemple, l'apport de nitrate et de phosphate par l'agriculture (engrais et lisiers) augmente le risque d'eutrophisation du cours d'eau qui est à l'origine in fine de la mortalité des communautés par asphyxie (Delong and Brusven, 1998 ; Quinn, 2000). Les MES et les MO provoquent les mêmes impacts qu'au niveau des rejets ponctuels urbains (cf. IV.2.1.1).

CemOA : archive ouverte d'Irstea / Cemagref

Les relations négatives entre la présence de terres arables sur le BV et la réponse biologique EQR-IBGN$_{moyen}$ sont certainement liées aussi aux impacts des pesticides (Croll, 1991 ; Skinner *et al.*, 1997) et de leur métabolites qui ont directement un effet nocif pour la vie aquatique, notamment des invertébrés (Schulz and Liess, 1999). Par exemple, les MES, particules très réactives agronomiquement et chimiquement, adsorbent les polluants et en particulier des métaux lourds et pesticides qui peuvent agir en synergie accentuant le caractère toxique des polluants. C'est le cas par exemple de la méthylation du mercure le rendant très bioassimilable (Meybeck *et al.*, 1998). Cependant, la mise en évidence de ces impacts provenait jusqu'ici principalement de tests toxicologiques sur des sites localisés plutôt que de recherches à l'échelle du paysage, compte tenu de la difficulté à mesurer la quantité de pesticides et leur métabolites dans le milieu récepteur (Allan, 2004a).

Les terres arables sont également à l'origine de sédiments fins provenant notamment des labours et différents travaux des sols. L'érosion aratoire entraîne une diminution des teneurs en matières organiques des sols en haut et au milieu des parcelles (Cavalié and Dressayre, 1986) provoquant ainsi une forte diminution de la biomasse lombricienne (Bouché, 1972 ; Chabert, 1996), facteur limitant l'érosion hydrique des sols (Bouché, 1972). Ces sédiments fins conduisent au colmatage des habitats aquatiques qui altèrent fortement la vie des communautés aquatiques (Waters, 1995) et notamment les macroinvertébrés benthiques vivant préférentiellement dans ces sédiments (homogénéisation du substrat et réduction des teneurs en oxygène dissous, Gayraud *et al.*, 2002).

(2) Les altérations hydromorphologiques générées par la présence de terres arables sont multiples. Les actions combinées de ces altérations participent à l'altération de la structure des habitats aquatiques et du fonctionnement de l'écosystème rivière.

L'agriculture est généralement à l'origine de changements de flux liquides dans les cours d'eau (Allan, 2004b) générant des modifications des capacités d'infiltration de l'eau dans le sol par le travail de la terre, par la mise en place des systèmes de drainage (drains, fossés, raies de curage) et de l'irrigation. La réduction des débits à l'étiage peut avoir des effets drastiques sur la vie aquatique en modifiant les conditions hydrodynamiques, donc les habitats aquatiques (réduction des profondeurs, réduction de la quantité d'habitats disponibles et donc de la capacité d'accueil), en aggravant la thermie et augmentant la concentration en polluants divers (Chandesris et al., 2008). De plus, les systèmes de drainage, rejetant plus en aval dans le système les eaux qui se chargent de façon importante de polluants au fil de leur écoulement dans les terres arables, aggravent les effets des pollutions diffuses.

Cependant, si on se réfère aux cartes de risques à large échelle établies par Chandesris et al. (2007) sur les pressions « drainage » et « irrigation » à partir des données du Recensement Général Agricole (RGA datant de 1988), la zone d'étude ne semble pas particulièrement touchée par ces pressions.

Par ailleurs, l'emprise agricole sur le territoire réduit, voire supprime, la végétation rivulaire sur la zone d'étude et la sinuosité du cours d'eau (Figure 10.18). Pendant les années 80 à 90, des programmes d'aménagement ont contribué de façon volontaire à rectifier le tracé des cours d'eau de rang 1 (Naiman *et al.*, 2005) de façon à mieux utiliser l'espace en fond de vallées alluviales (plus fertiles) et à protéger les cultures riveraines des inondations (la rectification permet entre autres d'augmenter la capacité d'écoulement du chenal en modifiant sa pente). Ces altérations morphologiques sont à l'origine d'une homogénéisation des faciès d'écoulement et des substrats, donc d'une réduction de la diversité des habitats, et d'une incision du lit mineur, avec une déconnexion des annexes hydrauliques et un abaissement de la nappe (Chandesris *et al.*, 2008). Cet impact est d'autant plus grave compte tenu du faible débit solide et liquide de la plupart des cours d'eau dans cette zone d'étude (cours de plaine à faible transport solide). La suppression de la ripisylve peut accentuer également le phénomène d'eutrophisation (Wasson *et al.*, 1998), d'érosion (berges instables, Richards *et al.*, 1996 ; Roth *et al.*, 1996 ; Wang *et al.*, 1997), et déséquilibrer l'état trophique du système : la nourriture dans le milieu, moins protégée de la lumière du soleil et augmentant par conséquent en température, est plutôt autochtone qu'allochtone, ce qui se traduit par la prolifération d'invertébrés brouteurs (Delong and Brusven, 1998 ; Quinn, 2000) et une diminution de la diversité taxonomique (Wang, 2006; Maret *et al.*, 2008).

CemOA : archive ouverte d'Irstea / Cemagref

Figure 10.18 : Exemple de rectification du tracé et suppression de la ripisylve liée à l'emprise agricole sur la zone d'étude (en bleu le réseau hydrographique). Source : BDortho® 2006 et BDCarthage® 2005 IGN

Ainsi, les multiples pressions générées par les terres arables à l'échelle du bassin peuvent altérer le fonctionnement et la structure de l'écosystème aquatique en agissant de façon combinée, ce qui se traduit par une perte de la diversité taxonomique et de taxons polluosensibles au niveau des invertébrés. Elles peuvent porter atteinte également directement ou indirectement à la vie aquatique en libérant dans le milieu récepteur des substances nocives (pesticides et métaux).

Cependant, cet impact est identifié par le modèle uniquement à l'échelle du bassin versant alors que visuellement on observe une emprise importante de cette occupation du sol au niveau du corridor rivulaire sur la zone d'étude (Figure 10.18). Plusieurs études pressions/impacts ont montré l'influence des terres arables au niveau du corridor rivulaire et notamment sur les macroinvertébrés (e.g., Sponseller *et al.*, 2001 ; Stewart *et al.*, 2001).

Dans notre cas, l'absence de mise en évidence d'un effet négatif des terres arables sur l'état écologique peut être liée au fait que les pressions générées par ce type d'occupation du sol ont été mal représentées dans le modèle au niveau du corridor rivulaire. L'élaboration des modèles a été réalisée (volontairement, cf. Chapitre 9 :II.2.1.4 selon une nomenclature détaillée et avec la donnée THRS, les terres arables sont divisés en 18 sous-catégories.

Pour autant, si on avait regroupé ces sous-catégories, la détection de cet impact n'aurait pas été si évidente. En effet, sur la zone d'étude compte tenu de l'emprise de cette occupation du sol au niveau du corridor rivulaire dans la zone d'étude et donc du manque de variation de cette variable sur le jeu de donnée, l'influence de cette occupation du sol aurait été difficile à mettre en évidence. De nombreux travaux mettent en évidence ce problème (Johnson and Host, 2010). C'est pour cette raison par exemple que les indices de structures de la communauté des macro-invertébrés ont été faiblement prédits dans les études menés sur les bassins versants agricoles dans le Minnesota et le Michigan (Richards *et al.*, 1997, Richards *et al.*, 1996).

CemOA : archive ouverte d'Irstea / Cemagref

IV.2.2. Les sources de pressions protectrices

IV.2.2.1 Le rôle des prairies et de la végétation herbacée semi-naturelle

L'influence positive des prairies (poste 231 ; Postes 4, 7 et 10 du Tableau 10-9) sur l'EQR-IBGN$_{moyen}$ se retrouve de façon marquée aux trois échelles analysées. L'influence la plus forte se trouve à l'échelle du CRamont sur une bande de 20 m de part et d'autre du cours d'eau (4ème position, cf.Tableau 10-9) : elle est relativement plus faible que l'impact des tissus urbains discontinus et équivalente à l'effet des zones industrielles et commerciales à l'échelle du BV. A l'échelle du BV et du CRlocal, cet impact est relativement plus faible qu'à l'échelle du CRamont (Tableau 10-9).

Cette influence positive des prairies a été mise en évidence par Wasson et al., (2010) à l'échelle nationale et dans diverses hydro-écorégions de la France et du Royaume-Uni aussi bien à l'échelle du BV qu'à l'échelle du corridor rivulaire (sur le lit majeur du cours d'eau à 2.5 km en amont de la station). Cependant, cette étude montre aussi sur certaines hydro-écorégions en France (e.g., Massif Armoricain), et à l'échelle nationale en Estonie, que les prairies peuvent avoir un impact négatif sur les macroinvertébrés à l'échelle du corridor rivulaire.

Les prairies à l'échelle du BV et aux échelles du corridor rivulaire ont un impact positif car comparées aux terres arables elles nécessitent moins d'apports nutritifs et de pesticides (particulièrement les prairies permanentes qui ne sont jamais labourées), ce qui limite la force polluante diffuse à cette échelle.

A l'échelle du corridor, suivant l'intensité d'élevage et sa gestion, les prairies pourront avoir un impact nuancé sur l'état écologique.

Les prairies ont une influence plutôt négative dans le cas de régions d'élevage intensif (concentration de têtes de bétails sur une plus petite surface) : les prairies peuvent engendrer une pollution diffuse (voire ponctuelle) comme les terres arables (cf. IV.2.1.3) en rejetant dans le milieu des matières organiques et des nutriments modifiant les processus biogéochimiques de l'écosystème aquatique. De plus le bétail, par le piétinement des berges du cours d'eau, peut générer une source non négligeable de matières en suspension et de sédiments fins (Macary and Paulais, 2003). Cet impact est plus ou moins nuancé suivant la fréquence de rotation de pâturage (Sovell *et al.*, 2000).

Par contre, les prairies agricoles ont une influence plutôt positive lorsque l'élevage est extensif. En bordure du cours d'eau, elles peuvent jouer un rôle de filtre vis-vis des flux de polluants diffus agricoles et urbains. Cet effet protecteur est lié à la capacité de la végétation herbacée à freiner l'écoulement, augmentant ainsi le dépôt des matières pour leur dégradation, et à stabiliser les berges réduisant ainsi les risques d'érosion donc de colmatage des habitats.

Dans la zone d'étude, on retrouve un effet protecteur des prairies permanentes aux échelles du corridor rivulaire : l'état écologique (EQR-IBGN$_{moyen}$) est favorisé par la présence de prairies à l'échelle du corridor amont sur une bande de 20 m de part et d'autre du cours d'eau et à l'échelle du corridor local sur une bande de 200 m de large sur 3km en amont de la station. Ce qui suppose à priori que l'élevage est plutôt extensif dans la zone d'étude. Une analyse des pratiques d'élevage sur les départements couvrant la zone d'étude permettrait de confirmer cette observation.

L'état écologique est également protégé par la végétation herbacée et arbustive semi-naturelle le long du corridor amont (Poste 321 ; facteur n°12) sur une bande de 250 m de part et d'autre du cours d'eau en amont des stations.

Cette végétation, prairies agricoles comme végétation herbacée et arbustive semi-naturelle, selon son emplacement, constitue autant de « puits » pour les polluants diffus (Gergel, 2005).

La Figure 10.19 illustre la bande de 20 m et la bande de 250 m (sur 3 km) en prairies qu'on retrouve en amont d'une station en très bon état soumise notamment à une pression des terres arables à l'échelle du BV.

CemOA : archive ouverte d'Irstea / Cemagref

Figure 10.19 : Illustration de la bande de 20 m et la bande de 250 m (sur 3 km) en prairies qu'on retrouve en amont d'une station en très bon état soumise notamment à une pression des terres arables à l'échelle du BV (en rouge la station, en bleu le réseau hydrographique). Source : BDortho® 2006 et BDCarthage® 2005 IGN.

Par ailleurs les gels de terres sur une bande de 5 m de part d'autre du cours d'eau sur le CRamont ou sur une bande de 25 m de large et 200m de long à proximité de la station (CRlocal), n'ont selon le modèle qu'une influence positive relativement faible sur l'EQR-IBGN$_{moyen}$. Ces gels correspondent à des parcelles laissées en jachère qui ne sont ni travaillées, ni récoltées et sont le plus souvent caractérisées par une végétation herbacée et/ou arbustive semi-naturelles. Cette végétation peut jouer un rôle de filtre en freinant l'écoulement au même titre que les prairies.

Toutefois dans ces espaces la superficie des gels de terres (quant ils sont présents) est inférieure à 1 % quelque soit les stations : des doutes sont donc permis sur la capacité de ce jeu de données à déterminer l'effet réel de cette occupation des sols au niveau du CRamont.

IV.2.2.2 Le rôle des ripisylves et de la végétation arborée

L'influence positive de la végétation arborée sur l'EQR-IBGN$_{moyen}$ se retrouve au niveau du BV (poste 243 ; facteur n°14 du Tableau 10-9) et de façon moins marquée à l'échelle du CRlocal (poste 311; facteur n°18 du Tableau 10-9) sur une bande de 20 m de part et d'autre du cours d'eau et de 200 m en amont de la station.

L'influence à l'échelle du BV (β = 0.041) est équivalente à l'influence des terres arables à cette échelle ; elle est relativement plus faible à l'échelle du corridor (β = 0.027), 3 fois moins importante environ que l'influence des prairies sur la bande 20 m à l'échelle du CRamont (Tableau 10-9).

L'influence de la végétation arborée à l'échelle du BV a été mise en évidence par de nombreux auteurs aussi bien sur les macroinvertébrés que sur les poissons (e.g., Roth *et al.*, 1996, Wang *et al.*, 1997 ; Harding *et al.*, 1998 ; Wasson *et al.*, 2010). La présence de ces surfaces dans le bassin versant limite l'apport des polluants diffus car elles peuvent, suivant leur position dans le BV, constituer des « puits » pour cette pollution (Gergel, 2005).

CemOA : archive ouverte d'Irstea / Cemagref

L'influence de la végétation arborée au niveau du corridor rivulaire a été mise en évidence également par de nombreux auteurs sur les macroinvertébrés (e.g., Sponseller *et al.*, 2001 ; Wang and Kanehl, 2003 ; Potter *et al.*, 2005 ; WALSH *et al.*, 2007) sur des emprises généralement assez larges (de 100 à 200 m de part et d'autre du cours d'eau). Ces études supposent la présence de ripisylves dans ces espaces qui protègent et favorisent la vie aquatique. Ces formations végétales le long du cours d'eau, comme on l'a vu dans le chapitre 2 de ce mémoire, constituent un compartiment clé du fonctionnement de l'écosystème rivière (elles contrôlent à la fois la nourriture, la température et la diversité des habitats dans l'écosystème), une barrière vis-à-vis des pollutions diffuses, et un stabilisateur des berges limitant leur érosion (Maridet and Souchon, 1995 ; Naiman *et al.*, 2005).

Notre étude conforte l'idée d'un rôle positif sur l'EQR-IBGN$_{moyen}$ de la présence de ripisylve dans la bande de 20 m de part et d'autre du cours d'eau à proximité de la station (200 m). La Figure 10.20 illustre la présence de cette végétation sur une station de la zone d'étude en bon état. Les ripisylves favorisent pour les macroinvertébrés la diversité taxonomique et assurent des conditions propices aux espèces polluosensibles.

Figure 10.20 : Illustration de la présence de la végétation rivulaire sur une station de la zone d'étude en bon état (en rouge la station, en bleu le réseau hydrographique). Source : BDortho® 2006 et BDCarthage® 2005 IGN.

IV.2.2.3 Les effets « contradictoires »

D'après le modèle certaines occupations du sol ont à l'échelle du CRlocal un effet protecteur sur l'EQR-IBGN$_{moyen}$, que l'on peut juger plutôt contradictoire par rapport au sens commun :

(1) les labours et les cultures annuelles sur le lit majeur du cours (bande de 100 m pour les labours et 200 m pour les cultures annuelles de part et d'autre du cours d'eau sur 2.5 km à l'amont de la station) (Postes 211 et 214 ; facteurs n°19 et 21 du Tableau 10-9)

(2) les plans d'eau à proximité de la station (bande de 300m de part et d'autre du cours d'eau sur 200 m en amont de la station) (Poste 412 ; facteur n°17 du Tableau 10-9).

CemOA : archive ouverte d'Irstea / Cemagref

Alors que les labours (poste 211) et les cultures annuelles (poste 214) font partie des prédicteurs possédant la plus faible influence sur l'EQR-IBGN$_{moyen}$ (β <= 0.02, Tableau 10-9), les plans d'eau (poste 412) ont une influence équivalente à la végétation arborée à l'échelle du CRlocal (β = 0.027).

(1) les labours et les cultures annuelles ont plutôt un effet négatif sur l'état écologique (cf. IV.2.1.3). Leur effet positif mise en évidence dans le modèle est donc contradictoire par rapport au sens commun. Cependant des doutes sont permis quant au sens réel de ces corrélations : dans les emprises identifiées par le modèle la superficie de ces occupations du sol (quant elles sont présentes) est inférieure à 2 % pour les labours et à 3% pour les cultures annuelles quelques soient les stations.

(2) Les plans d'eau dérivent de l'information contenue dans la BDCarthage de 2006. Dans la zone d'étude, ces étendues correspondent en majorité à des gravières (Agence de l'eau Seine Normandie, 2004). Les gravières à large échelle et à l'échelle d'un tronçon morphologique (Chandesris *et al.*, 2008) peuvent avoir un impact sur les flux solides. Ces gravières constituent des pièges à sédiments, ce qui se traduit par des altérations morphologiques avec une possibilité d'incision généralisée du lit mineur en aval des gravières, une réduction de la quantité de substrat alluvial grossier disponible (habitat majeur des biocénoses aquatiques), et une disparition complète dans certains cas du substrat alluvial. Ce n'est pas le cas sur la zone d'étude puisque selon le modèle les plans d'eau auraient une influence positive très locale sur l'EQR-IBGN$_{moyen}$.

Pour essayer de comprendre cette influence a priori contradictoire, on a visualisé les stations étant dans ces situations. Toutes les stations (11 au total) sont en très bon état. On observe à proximité de ces stations, un espace préservé avec la présence notamment de ripisylves et de prairies en bordure du cours d'eau. La Figure 10.21 illustre ces propos sur trois stations. On observe également que pour les stations possédant des retenues assez proches du cours d'eau, l'espace de liberté du cours d'eau a été préservé ou réaménagé : sur l'illustration de gauche dans la Figure 10.21, on s'aperçoit que la sinuosité du cours d'eau a été conservée.

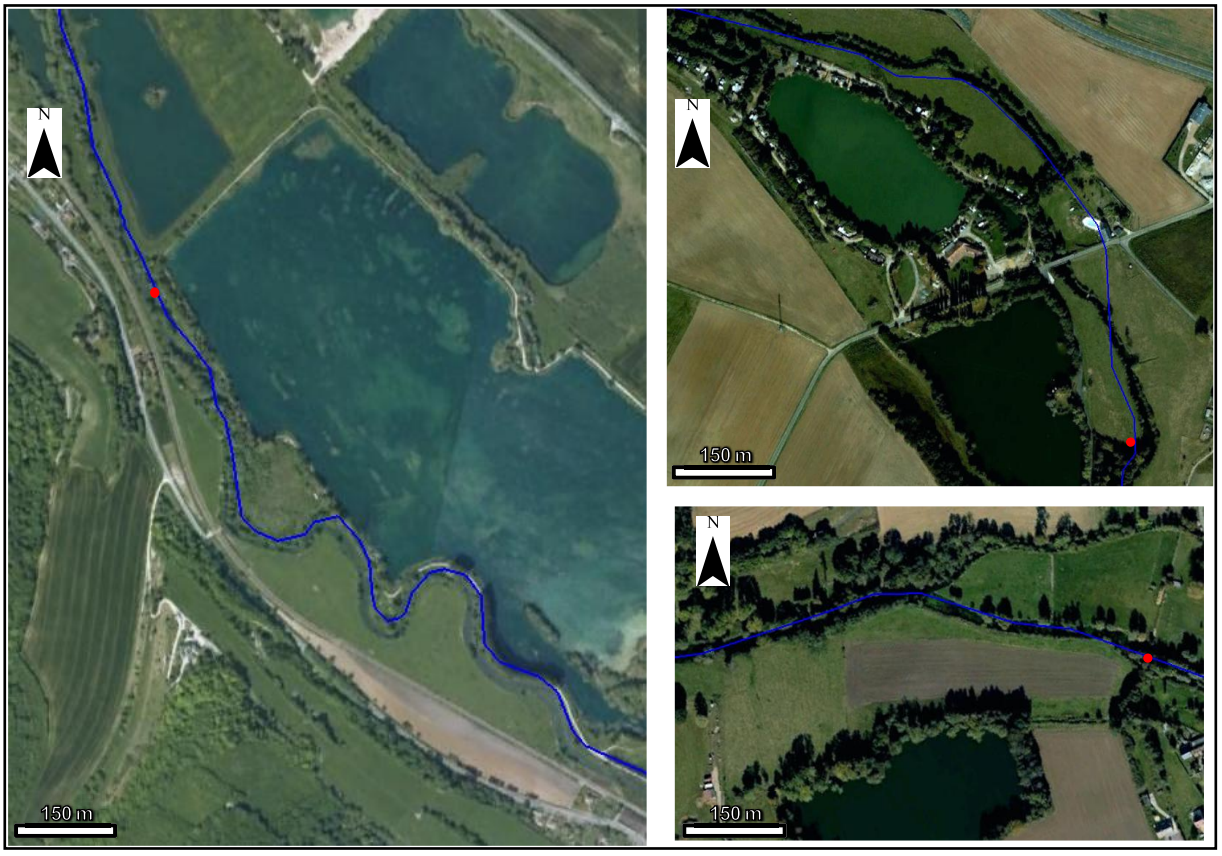


Figure 10.21 : Illustration de l'environnement qu'on observe sur les stations possédant une étendue d'eau dans leur environnement proche (en rouge la station, en bleu le réseau hydrographique). Source : BDOrtho® 2006 et BDCarthage® 2005 IGN.

CemOA : archive ouverte d'Irstea / Cemagref

IV.3. Implications pour la gestion

L'interprétation du modèle pressions/impacts étudiant les relations entre d'une part les indicateurs d'occupation du sol calculés à l'échelle du BV, du CRamont, et du CRlocal et d'autre part l'EQR-IBGN$_{moyen}$, permet de faire un premier diagnostic sur la zone d'étude (cours d'eau de rang 1 à 4) qui de façon relative identifie, hiérarchise et localise les principales structures dégradantes et les facteurs qui protègent l'état écologique (cf. Chapitre 10 :IV.2).

On peut donc à partir de l'interprétation du modèle orienter les gestionnaires sur les opérations de restauration fortement envisageables sur leur territoire. Toute tentative d'utiliser un tel modèle pour reconstituer un état écologique sur l'ensemble du réseau ou encore proposer des scénarios de restauration quantifiant l'apport d'une action de restauration donnée seraient ici caduques compte tenu du faible pourcentage de variabilité expliqué par le modèle (environ 38%) et des pressions non prises en compte dans les variables d'occupations du sol (pressions liées aux activités d'énergie, de tourisme…).

En premier lieu, l'état écologique est altéré par la présence de territoires artificialisés (tissu urbain discontinu, zones industrielles et commerciales et routes). Cet impact a été reconnu dans de nombreuses études pressions/impacts. Un seuil situé entre 6 et 15% d'urbanisation revient régulièrement dans les différentes études analysant cet impact (Valette, 2005). En dessous de ce seuil, les effets sur les peuplements aquatiques sont peu ou pas perceptibles, du moins il n'y a pas de modifications majeures des caractéristiques des populations. Alors qu'au delà de ce seuil d'urbanisation, il y a une diminution, voire une disparition, des macroinvertébrés polluosensibles au détriment d'espèces plus résistantes (Morse *et al.*, 2003 ; Freeman and Schorr, 2004) et une modification du peuplement piscicole (Booth, 1991 ; Luchetti and Fuersteburg, 1993).

L'urbain est tout d'abord source de pollution organique et de substances toxiques (métaux, substances médicamenteuses…) via les rejets ponctuels. Ces problèmes de pollution sont bien connus par les gestionnaires, et la plupart des investissements financiers sont dédiés à réduire cet l'impact au moyen de stations d'épuration (Garcia *et al.*, 2005). De ce point de vue, on peut s'interroger sur la marge de manœuvre restante pour récupérer les milieux les plus dégradés par la pollution. En effet, en 2001, plus de 95% des logements étaient raccordés à un système épuratoire collectif ou autonome, et le taux d'abattement de la pollution organique était de 92% en sortie des stations de plus de 2000 équivalent- habitant (Coutellier, 2004).

Cependant, des efforts peuvent être encore faits notamment concernant les déversoirs d'orages dans les systèmes d'assainissement (type « unitaire ») qui collectent dans une seule conduite eaux usées et eaux pluviales (ou de source). D'une part ces déversoirs diluent et refroidissent les eaux usées, ce qui perturbe le cycle d'épuration ; d'autre part, lorsque le système d'assainissement est surchargé, les eaux pluviales, très chargées en pollution (particulièrement en début de précipitations) sont directement rejetées dans l'environnement. L'article 5 de l'arrêté du 22 juin 2007 permettra dans un avenir proche de réduire cet impact puisqu'il stipule qu'aucun déversement ne pourra être admis en dessous des débits de référence des déversoirs d'orages (calculés par rapport à la pluviosité sur l'agglomération). Par ailleurs, des efforts devront également être menés à la fois sur les connaissances de l'impact et du devenir des substances toxiques (métabolites) dans le milieu naturel et sur les méthodes de réduction de ces substances (de plus en plus nombreuses) au niveau des stations d'épuration.

Viennent ensuite, de façon moins marquée, les altérations liées aux terres arables, uniquement à l'échelle du BV, sur les différents processus de fonctionnement de l'écosystème. Ces terres arables apportent de façon diffuse d'une part des nutriments et des sédiments fins qui altèrent la biodiversité des communautés en modifiant l'état chimique, trophique et la structure des habitats (colmatage) ; et d'autre part des substances toxiques (pesticides, herbicides, insecticides, …) qui ont des effets létaux ou sous-létaux sur les différents organismes.

Cependant, l'objectif de « bon état » ne semble pas irréaliste dans ces régions très agricoles. Le modèle montre que l'effet protecteur de la végétation arborée (ripisylves) dans la bande de 20 m et surtout de la végétation herbacée (les prairies permanentes ont une contribution positive sur l'EQR-IBGN$_{moyen}$ de même

CemOA : archive ouverte d'Irstea / Cemagref

amplitude que l'effet négatif des terres arables sur le BV). Cette végétation, facilement gérable dans cet espace, joue un rôle de zone tampon vis-à-vis de cette pollution, contribue à la stabilisation des berges et assure le bon fonctionnement de l'écosystème (Maridet, 1995 ; Naiman *et al.*, 2005).

Les politiques commencent à être conscients de l'intérêt de restaurer ces espaces rivulaires notamment au travers de la Politique Agricole Commune (PAC) qui impose la mise en place de bandes enherbées sur certains cours d'eau, ou encore de la Trame Verte et Bleu (Grenelle de l'Environnement) qui s'intéresse en partie aux ripisylves pour assurer la continuité écologique entre les différents milieux et espaces naturels.

Néanmoins, restaurer efficacement ces espaces demande des investigations complémentaires afin de préciser les stratégies à mettre en œuvre :

- d'une part sur le dimensionnement de la végétation rivulaire : quelles largeurs ? doivent-elles être uniformes, continues ? L'utilisation dans les modèles pressions/impacts des indicateurs de contact (cf. Chapitre 7 :II.2.3.3), renseignant sur la continuité, la largeur et l'uniformité des objets adjacents au cours d'eau, permettrait de répondre en partie à ces questions.

- d'autre part sur l'emplacement de cette végétation sur le bassin : sont-elles efficaces sur toutes les tailles de cours d'eau ? Par exemple, Potter et al. (2005) montre que les macroinvertébrés répondent plus à la présence de cette végétation en têtes de bassins. De plus, il ne faut pas oublier que ces filtres sont fonctionnels uniquement dans certaines conditions (cf. Chapitre 2 :III.2) liées à la topographie et à la géologie du milieu.

Par ailleurs, cette restauration sera d'autant plus efficace, si en parallèle des mesures sont prises au niveau des parcelles cultivées, pour réduire l'érosion des sols, la perte de nutriments et la diffusion des pesticides. Les ripisylves ou bandes enherbées peuvent être ineffectives lorsque les terres agricoles produisent plus de polluants qu'elles ne sont capables d'en filtrer (Nutter and Gaskin, 1988 ; Cooper *et al.*, 1987 ; Harding *et al.*, 1998). Le plan Ecophyto (Grenelle de l'environnement) va dans ce sens puisqu'il vise à réduire de 50 % l'usage des produits phytosanitaires en agriculture, à l'horizon 2018.

En plus la restauration des espaces rivulaires, des aménagements du territoire agricole à l'échelle du BV voire dans le lit et à proximité du lit majeur des cours d'eau peuvent être également envisagés. Le modèle montre à l'échelle du BV, du CRamont et à l'échelle du CRlocal, l'influence positive des prairies permanentes et de la végétation herbacée. Cette végétation constitue autant de pièges à polluants avant qu'ils ne s'acheminent vers le cours d'eau. Gergel (2005) montre à l'aide d'une approche heuristique que l'hétérogénéité spatiale des sources de polluants (ou des puits) expliquent une part importante des taux de nutriments retrouvés dans le cours d'eau. Il serait donc intéressant d'analyser la structure spatiale (notamment avec les indicateurs de fragmentation cf. Chapitre 7 :II.2.1.2) de ces territoires afin d'orienter les pratiques culturales à l'échelle régionale.

Toutefois, la possibilité de gagner quelques points d'IBGN ne devra pas porter uniquement sur la réduction de la pollution résiduelle et toxique, sur la végétalisation des espaces rivulaires ou encore la modification des pratiques à l'échelle du BV mais aussi sur la restauration physique des rivières. En effet, on a vu que l'EQR-$IBGN_{moyen}$ est sensible à la proximité des routes dans le lit majeur qui modifient drastiquement les processus hydromorphologiques. On peut suspecter aussi l'effet des zones urbanisées et des terres arables qui s'accompagnent généralement d'ouvrage et d'aménagements altérant ces processus.

Retrouver un bon fonctionnement hydromorphologique (reconstitution des habitats) est la pierre angulaire de la restauration de l'état écologique des cours d'eau car ces processus favorisent la biodiversité (compétition intra et inter-spécifique moins sévère), assurent la capacité auto-épuratoire de l'écosystème (Weak *et al.*, 1998), et sont garants de la viabilité d'une ripisylve effective favorisant et protégeant la vie aquatique (Maridet, 1995).

Les politiques commencent à être conscients de l'intérêt de cette restauration, même si les argumentaires sont encore trop souvent justifiés par des objectifs uniquement piscicoles (franchissement des obstacles pour les espèces migratrices). Prenons, par exemple, le cas du plan national de restauration de la continuité écologique (Trame bleu) mené en partie par l'Office National de l'Eau et des Milieux Aquatiques (ONEMA). L'ONEMA dresse actuellement au niveau national un inventaire des obstacles sur l'ensemble des cours d'eau français qui s'accompagne d'une évaluation de l'impact de chaque obstacle sur la libre circulation des espèces

CemOA : archive ouverte d'Irstea / Cemagref

migratrices et sur l'écoulement des sédiments. L'Agence de l'eau Seine-Normandie est consciente aussi de cette problématique puisqu'elle met en œuvre une politique ambitieuse en matière d'incitation à la restauration hydromorphologique des cours d'eau (Adam *et al.*, 2007). Le 9e programme (2007-2012) favorise l'émergence de travaux sur les cours d'eau où l'état hydromorphologique est considéré comme le principal facteur limitant pour l'atteinte du bon état écologique (objectif de la Directive Cadre sur l'Eau), en attribuant des aides à un taux de 60 % (majoration de 20 %).

Cependant, restaurer de façon ambitieuse l'hydromorphologie des cours d'eau nécessite encore l'acquisition de connaissances scientifiques et le développement de méthodes afin d'identifier plus précisément les facteurs modifiant ces processus et de développer une véritable ingénierie de la restauration hydromorphologique des cours d'eau. L'intégration par exemple de nouvelles variables de pression hydromorphologiques (notamment les variables homogènes au niveau national produites par le projet SYRAH, Chandesris *et al.*, 2007) dans l'étude des relations pressions/impacts permettra d'avoir une vision plus précise des mécanismes altérant ces processus ; et par conséquent d'optimiser les mesures de restauration (sur quels leviers agir en priorité et quels résultats en attendre ?).

Néanmoins, la restauration hydromorphologique des cours d'eau en zone urbaine ou proche des principaux axes de transport sera nécessairement limitée dans ses ambitions. Il sera plus difficile par exemple de remettre en cause un certain « confort » d'usage du cours d'eau et de son lit majeur (Malavoi and Adam, 2007). De plus, cet espace, chargé d'histoire en France, est difficilement mobilisable (Souchon, 2007). On pourra surtout s'attacher à améliorer l'habitat des biocénoses dans le lit mineur par de petites structures piscicoles et l'aspect paysager des berges (Malavoi, 2007).

Il faudra bien garder à l'esprit également des délais de réponse des systèmes écologiques dans les actions de restauration de ce genre (Adam *et al.*, 2007) ; en particulier sur la zone d'étude car la restructuration physique sous l'action des processus morphodynamiques naturels peut s'avérer très lente pour les rivières à faible énergie dans les régions de plaine.

Ainsi, ce diagnostic a permis de mettre en évidence à l'échelle d'une région (Nord-Ouest de l'hydro-écorégion Tables Calcaires), l'impact des terres arables, des routes, et des terres artificialisées (ce dernier étant plus marqué) sur les macroinvertébrés benthiques des cours d'eau de petite à moyenne taille agissant de façon combinée aux différentes échelles de fonctionnement de l'écosystème rivière. Il montre aussi l'intérêt de restaurer et/ou de préserver la bande de 20 m au niveau du corridor rivulaire en prairie et/ou en végétation arborée qui protègent le milieu des impacts anthropiques et favorisent la diversité taxonomique des macroinvertébrés et la viabilité des taxons polluosensibles.

Bien évidemment, ce diagnostic doit être (i) validé en prenant en compte les incertitudes liées aux erreurs de mesure des indicateurs et les dépendances d'échelles ; (ii) confirmé en réitérant cette analyse sur un autre jeu de donnée et enrichit en utilisant d'autres variables biologiques (poissons, diatomées) ; et (iii) précisé en faisant appel à des variables de pressions d'origine agricole, de pressions hydromorphologiques et d'état de la végétation le long des cours d'eau.

CemOA : archive ouverte d'Irstea / Cemagref

Synthèse de la quatrième partie

- ***21 modèles pressions/impacts qui diffèrent*** soit par la ***combinaison d'échelles d'analyse*** (bassin versant et/ou corridor rivulaire amont et/ou corridor rivulaire local), soit par ***la source de donnée d'occupation du sol mobilisée au niveau des deux échelles du corridor rivulaire (***donnée CORINE Land Cover ou donnée à THRS), soit par le ***mode de spatialisation de l'emprise au niveau du corridor rivulaire local*** (fixe ou multiple), ont été construits sur la zone d'étude (partie Nord-Ouest de l'hydro-écorégion Tables Calcaires).

- Tous ces modèles cherchent à expliquer, en se basant sur ***une population de 155 stations*** écologiques, ***la variabilité de l'écart entre la référence*** et ***la moyenne de l'Indice Biologique Global Normalisé*** observé à la station ***sur la période de 1992 à 2004***. Ils sont construits à partir d'une ***procédure en deux étapes***. Une ***première étape*** qui ***sélectionne*** pour chaque source de pression à une échelle donnée ***l'emprise spatiale*** sur laquelle cette source a ***l'influence la plus significative*** sur l'état écologique. Une ***deuxième étape*** basée sur ***l'algorithme de régression Partial Least Square*** qui cherche de façon automatique la ***combinaison des différentes sources de pressions*** (s'exprimant sur leur emprise spatiale la plus significative) ***aux différentes échelles*** expliquant au mieux la variabilité de la réponse biologique.

- La ***comparaison des résultats des différents modèles*** montre que ***les avancées méthodologiques*** (donnée à THRS, mode de spatialisation multiple) permettent ***d'améliorer significativement le pouvoir explicatif*** des modèles et ***d'identifier un plus grand nombre de sources de pressions*** influençant la réponse biologique tout ***en précisant leur emprise spatiale significative***. Elles ***fiabilisent*** donc ***la compréhension des relations pressions/impacts*** et ***apportent de nouvelles informations*** qui peuvent être ***pertinentes pour la gestion***.

- Cette comparaison permet de sélectionner ***le modèle le plus pertinent pour établir un diagnostic de l'état écologique***. Ce modèle utilise les avancées méthodologiques et ***étudie les impacts de l'occupation du sol qui interagissent sur l'ensemble des trois échelles***. Il explique environ ***38 % de la variabilité de la réponse biologique*** et identifie une ***combinaison*** de ***23 sources de pressions*** significatives et les emprises spatiales associées.

- Son interprétation confirme ***l'impact marqué des terres artificialisées*** aussi bien à l'échelle du ***bassin versant*** (BV) qu'aux échelles du ***corridor rivulaire local et amont, ainsi que l'impact des terres arables*** à l'échelle du ***BV***. ***L'objectif de « bon état*** » ne semble ***pas irréaliste*** dans ces régions très agricoles puisque le modèle met en évidence ***les effets protecteurs dans une bande rivulaire gérable*** (20 m de part et d'autre du cours d'eau) de la ***végétation arborée et de la végétation herbacée permanente*** qui ont une ***contribution positive cumulée*** sur l'EQR-IBGNmoyen ***de même amplitude*** que ***l'effet négatif des terres arables*** sur le BV.

- Toutefois, la possibilité de gagner quelques points d'IBGN ne devra pas s'appuyer uniquement sur la ***restauration/préservation*** des espaces rivulaires (même s'ils conditionnent en partie les processus de fonctionnement) mais également sur celles de ***l'état hydromorphologique des rivières*** compte tenu de la nature des impacts mis en évidence (***impacts des routes*** notamment).

- L'ensemble de ces résultats doit être interprété avec ***une certaine prudence*** compte tenu (i) que leur ***fiabilité vis-vis des incertitudes des indicateurs n'a pas été évaluée*** ; et (ii) que ***les dépendances d'échelles***, *permettant de vérifier l'effet propre d'un facteur à une échelle donnée sur la variabilité de la réponse biologique, et de confronter l'effet des facteurs à une échelle par rapport à une autre, n'ont pas été analysées*

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

CONCLUSION & PERSPECTIVE

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

La Directive Cadre européenne sur l'Eau (DCE) requiert d'atteindre d'ici 2015 le « bon état » (i.e., le « bon fonctionnement ») des cours d'eau. Afin d'orienter efficacement les gestionnaires dans l'élaboration des mesures de préservation et/ou de restauration, il est indispensable de développer des méthodes et des modèles capables d'identifier, de hiérarchiser et de localiser, à large échelle (nationale, régionale, grands bassin versants), les causes d'altération ou de protection de cet état, et de quantifier leur influence.

Dans cette thèse, nous nous sommes consacrés aux modèles pressions/impacts empiriques régionalisés (ou statistiques) qui, contrairement aux modèles déterministes, permettent d'étudier sur de grands territoires les relations entre l'occupation du sol, à l'origine de la plupart des pressions modifiant les processus fonctionnels des cours d'eau, et les biocénoses aquatiques répondant à ces processus fonctionnels. Les modèles pressions/impacts empiriques régionalisés permettent d'étudier ces relations aux différentes échelles de fonctionnement de l'écosystème rivière : bassin versant, corridor rivulaire amont et corridor local. Dans ces modèles, nous nous sommes particulièrement intéressés aux relations existant à l'échelle du corridor rivulaire vu comme l'espace clé de restauration (Naiman *et al.*, 2005 ; CSPNB, 2008) puisque d'une part le foncier est plus facilement négociable dans ces espaces qu'à l'échelle du bassin et d'autre part ils abritent les formations végétales rivulaires qui protègent et régulent les processus fonctionnels de l'écosystème rivière.

Cependant la littérature scientifique montre que deux principales difficultés empêchent d'étudier correctement ces relations au niveau des corridors rivulaires :

(i) la première difficulté réside dans le manque d'information adaptée à cette échelle : la résolution de la donnée spatiale actuellement disponible (Corine Land Cover) est trop grossière pour permettre de quantifier précisément la composition et la configuration de l'occupation des sols dans ces espaces et de caractériser l'état des ripisylves. De ce fait l'analyse et l'interprétation de ces relations sur la base de données à moyenne résolution sont limitées (Hollenhorst *et al.*, 2006 ; Wasson *et al.*, 2005 ; Wang, 2006 ; Roy *et al.*, 2007) ;

(ii) la deuxième difficulté réside dans la spatialisation des sources de pressions (Frimpong *et al.*, 2005b ; Van Sickle and Johnson, 2008) et l'identification de leur emprise géographique d'influence : généralement, que ce soit au niveau du corridor rivulaire local ou du corridor rivulaire amont, les sources de pressions, quelqu'elles soient, sont étudiées sur une emprise spatiale unique définie par expertise. Cette difficulté limite la compréhension des relations pressions/impacts car l'effet d'une catégorie d'occupation du sol sur la réponse biologique du cours d'eau est plus ou moins perceptible selon l'emprise spatiale au sein de laquelle cette pression agit. Par ailleurs l'utilisation d'une emprise spatiale unique réduit pour les gestionnaires l'information concernant la localisation de ces sources de pressions au niveau du corridor.

L'enjeu de cette thèse a donc été d'une part de développer des solutions méthodologiques permettant de dépasser ces difficultés, en testant les potentialités des outils de la télédétection (classification à haute résolution spatiale de l'occupation du sol le long des cours d'eau) et de l'analyse spatiale (construction d'indicateurs spatiaux du corridor rivulaire), et d'autre part d'élaborer des modèles pressions/impacts rendant compte de l'influence des différents facteurs de pression.

La classification de l'occupation du sol dans le long des cours d'eau

Dans un premier temps (partie 2), nous nous sommes donc intéressés à développer une procédure de classification opérationnelle pour extraire finement et efficacement l'occupation du sol dans les corridors rivulaires sur de grands territoires à partir de l'information accessible à coûts réduits pour les gestionnaires. Nous avons mobilisé pour cela l'information issue (i) des images de télédétection à Très Haute Résolution Spatiale (THRS) aéroportées et/ou satellitaires qui fournissent une information spatiale homogène sur de grandes surfaces, appropriée à la caractérisation des systèmes rivulaires (Müller, 1997) ; et/ou (ii) des données thématiques (i.e, données auxiliaires) d'une précision métrique voire décamétrique, afin d'enrichir et d'améliorer l'information utile pour classer les images.

Cette procédure de classification a été conçue sur le bassin versant de l'Hérault, avec pour objectif minimum d'extraire six catégories principales d'occupation du sol : les surfaces en eau (cours d'eau, plans d'eau…) et les sols nus naturels (sable, galets, roches nues…) afin de délimiter précisément le lit de la rivière ; les sols artificialisés (tissus urbains continus, discontinus, industriels …) et sols agricoles (cultures annuelles,

CemOA : archive ouverte d'Irstea / Cemagref

permanentes…) considérés dans la littérature (Allan, 2004b) comme les principales sources de pressions altérant l'état écologique des cours d'eau ; les zones de végétation arborée et de végétation herbacée constituant les principaux types de végétation naturelle assurant le rôle de préservation des corridors rivulaires (Maridet, 1994; Naiman *et al.*, 2005).

La procédure de classification est structurée en 7 étapes :

- l'***étape 1*** sélectionne les sources de données images et les sources de données thématiques (ou auxiliaires) adaptées à la caractérisation de l'occupation du sol dans les corridors rivulaires et disponibles à moindre coût (pour les gestionnaires de l'eau) sur la zone d'étude ;
- l'***étape 2*** délimite les corridors rivulaires (espace à classer) ;
- l'***étape 3*** stratifie la zone d'étude en régions homogènes en termes de types d'occupation du sol (contexte géographique naturel : relief, climat, géologie) et de période d'acquisition des images ;
- l'***étape 4*** découpe les sources de données selon les contours des différentes zones ainsi délimitées (corridors rivulaires et régions homogènes) et applique des prétraitements afin d'améliorer le contenu des sources de données images et auxiliaires pour la classification de l'occupation du sol ;
- ***l'étape 5*** classe l'occupation du sol dans chaque région homogène (chaque strate) selon une approche orientée objet multi-échelle définissant une hiérarchie de quatre niveaux d'objets emboîtés et un arbre de classification descendant ajustable, basé sur des règles expertes booléennes ou floues combinant l'information de l'ensemble des sources de données prétraitées ;
- l'***étape 6*** produit une cartographie sur l'ensemble de la zone d'étude selon une nomenclature structurée en réunissant les résultats de classification de chaque strate ;
- l'***étape 7*** évalue la qualité sémantique (types d'occupation des sols des objets) de cette cartographie à l'aide de matrices de confusions, puis sa fiabilité et sa stabilité en utilisant les résultats de la classification floue.

La conception et l'optimisation de cette procédure de classification sur le bassin versant de l'Hérault a permis de mettre en évidence la combinaison optimale de sources de données : (1) les orthophotos de la base de données IGN (0.5 m de résolution spatiale avec une information radiométrique dans le visible) nécessaires notamment pour délimiter les objets fins et étroits le long des cours d'eau ; (2) les images satellites Spot5 XS de Spotimage© (10 m de résolution spatiale avec une information dans le vert, le rouge, le proche infra rouge et le moyen infra rouge) indispensable pour différencier les différentes classes de végétation ; et (3) les données auxiliaires : les postes artificialisés de Corine Land Cover (CLC), les îlots de parcelles déclarées à la Politique Agricole Commune du Registre Parcellaire Graphique (RPG), les surfaces en eau de la BDCarthage® de l'IGN et les infrastructures linéaires de la BDroutes500® de l'IGN. Cette combinaison des sources de données disponibles à coûts réduits sur le territoire français est la plus efficace aussi bien pour extraire les catégories d'occupation du sol d'intérêt que pour caractériser les ripisylves (la végétation arborée rivulaire).

Cette optimisation a permis également de sélectionner la technique de classification la plus appropriée et d'améliorer le processus de validation par matrice de confusion.

Au niveau du processus de classification orientée objet, la technique « supervisée » s'est avérée efficace et robuste avec un jeu d'apprentissage optimal de 50 parcelles/classe, ceci quelque soit la taille de la région homogène. Néanmoins, notre choix s'est porté sur la technique « expertisée » qui opère la classification des objets par fonctions d'appartenance appliquées à différents attributs. Cette technique est à la fois relativement plus efficace que la technique « supervisée » et moins lourde à mettre en œuvre sur de grands territoires car elle est modulable (i.e., on peut intervenir dans le processus de classification sans remettre en cause l'ensemble du processus) et distribuable sur un serveur de calcul.

Au niveau du processus de validation, la taille de l'échantillon de validation a été optimisée afin d'obtenir une valeur robuste (stable) des estimateurs de qualité de la classification. Ces estimateurs se sont avérés stables (sur cette zone d'étude) lorsque la matrice de confusion est construite à partir d'un échantillon constitué de 50 objets/classe, sélectionnés de façon aléatoire et géographiquement répartie sur la zone d'étude.

La mise en œuvre de la procédure de classification sur le bassin versant de l'Hérault, au contexte géographique contrasté, a permis de tester son opérationnalité et sa généricité. Cette procédure s'est avérée :

(i) ***efficace*** globalement puisque 89 % des objets de l'échantillon de validation se sont avérés être correctement classés selon la typologie d'occupation du sol minimum exigée (six catégories) ;

CemOA : archive ouverte d'Irstea / Cemagref

(ii) ***générique*** (transférable) car le raisonnement de classification orientée objet descendant, favorisant la description de règles simples, robustes et facilitant l'intégration de nouvelles classes, a pu être appliqué sur l'ensemble du bassin (aussi bien en plaine qu'en zone de relief) sans (1) avoir à modifier l'architecture de l'arbre de classification et (2) avoir à définir de nouvelles règles de classification mais simplement en ajustant les seuils des règles existantes (excepté dans le cas de nouvelles classes spécifiques aux régions étudiées) ;

(iii) ***applicable sur de grands territoires*** : trois jours de traitement orienté objet (segmentation, ajustement des règles de classification et classification) ont suffi pour caractériser l'occupation du sol le long des 1150 km du réseau hydrographique du bassin versant de l'Hérault (560 km² de corridor rivulaire), ceci à l'aide d'un cluster de calcul constitué de 8 cœurs.

Par ailleurs cette mise en œuvre de la procédure sur le bassin de l'Hérault, tout comme sa mise en oeuvre sur les corridors rivulaires de la Basse Normandie dans la partie 4 (partie Nord-Ouest de l'hydro-écorégion Tables Calcaires, HER9-NO, 5600 km² de corridor), a permis de mettre évidence l'intérêt des données auxiliaires pour améliorer le résultat de classification (en terme de fiabilité et de nombre de classes) et pour augmenter la faisabilité technique de cette procédure en réduisant considérablement les temps de traitements.

C'est le cas notamment de la donnée RPG (Registre Parcellaire Graphique) : sur l'HER9-NO, cette donnée s'est avérée très intéressante car elle a permis (1) de détailler les classes d'agriculture ; (2) d'améliorer les résultats de la segmentation des images ; et (3) de réduire le temps de traitement des images car elle occupe plus de 46 % de la superficie des corridors rivulaires analysés.

Cette thèse a donc permis de forger une procédure de classification capable d'extraire finement l'occupation du sol le long des cours d'eau sur de grands territoires, en combinant l'information de données spatiales multi-sources (images et thématiques) disponibles sur la zone considérée.. Elle fournit une information globalement fiable sur les principales catégories d'occupation du sol exerçant des pressions sur l'état écologique des cours d'eau. Cette méthode est opérationnelle et peut-être proposée aujourd'hui aux gestionnaires de l'eau et des milieux aquatiques.

Perspectives

Toutefois, certains aspects de cette procédure de classification peuvent être améliorés afin de fournir une information plus précise pour l'étude des relations pressions/impacts sur de grands territoires. Ces perspectives concernent d'une part l'amélioration de la classification et d'autre part la spatialisation de l'incertitude associée aux résultats

Si l'on analyse par exemple la matrice de confusion résultant de la classification de l'occupation du sol sur les corridors rivulaires du bassin versant de l'Hérault, on s'aperçoit que seulement 35% des objets « sols nus semi-naturels » et 60% des objets « sols artificialisés » sont bien classés, en partie à cause de confusions entre ces deux classes. Confondre une catégorie susceptible d'altérer l'état écologique avec une catégorie non dégradante peut introduire des biais importants dans l'étude de ces relations. Des efforts sont donc nécessaires pour améliorer la caractérisation de ces deux catégories d'occupation des sols.

Par ailleurs cette procédure (telle qu'elle a été optimisée) ne permet pas aujourd'hui d'obtenir une donnée instantanée de l'occupation du sol. L'application de la procédure sur les corridors rivulaires de l'HER-NO couvrant 11 départements illustre bien ce problème. Compte tenu du plan de mise à jour des orthophotos IGN (mise à jour tous les 5 ans par département), et du choix d'acquisition des données Spot5 XS (images d'archives), la zone d'étude a été couverte avec une mosaïque d'images très hétérogènes du point de vue temporel (un écart maximum de 9 ans, de 2002 à 2007, sur la zone d'étude). Cette hétérogénéité peut limiter la compréhension des relations pressions/impacts notamment lorsque des mesures de préservation et/ou de restauration ont été mises en œuvre sur le territoire. Une réflexion devra donc être menée sur le choix des images (et des données auxiliaires) afin d'acquérir une information d'occupation du sol temporellement homogène. Par ailleurs il sera nécessaire de concevoir un pas de temps de renouvellement approprié pour évaluer l'efficacité des divers plans de restauration mis en œuvre dans ces espaces.

Cependant, du fait de la capacité de cette procédure à intégrer facilement des sources de données (grâce aux propriétés de la logique floue qui permettent de gérer les conflits entre ces sources), de l'évolution des conditions d'accessibilité aux données spatiales existantes (données IGN) et de l'apparition de nouvelles

CemOA : archive ouverte d'Irstea / Cemagref

données dans le paysage français (couverture satellitaire annuelle de la France 2010 dans le cadre de GEOSUD) ces difficultés pourront être aisément surmontées dans un futur proche.

Les tableaux A et B présentent respectivement les caractéristiques des données images et thématiques potentielles qui sont (ou seront dans un futur proche) accessibles à coûts réduits sur l'ensemble du territoire français pour les gestionnaires. D'après ces tableaux, on conseillera, parmi les pistes à explorer, de s'intéresser tout d'abord à la donnée auxiliaire BDtopo® de l'IGN, d'une précision métrique, compte tenu de l'information et du gain de temps de traitement qu'elle pourrait apporter ; cette donnée est d'ailleurs largement utilisée dans le projet SYRAH[51] de caractérisation des cours d'eau (Chandesris *et al.*, 2007). Ensuite, on conseille de tester les potentialités de la couverture satellitaire d'été 2010 de la France en images satellitaires RapidEye (5 m de résolution spatiale avec une information dans le bleu, le vert, le rouge, le « Red-Edge »[52] et le proche infra rouge) qui apporte une information temporelle homogène d'une part (donc réduit l'hétérogénéité temporelle de la donnée) et requiert des coûts de traitement moins élevées (en terme de stockage et de temps de traitement) que les orthophotos IRC (infra rouge couleur, avec une information dans le proche infra rouge). Cette couverture satellitaire haute résolution 2010 pourra en outre être plus facilement mise en correspondance avec la donnée RPG de la même année.

Dans un futur plus lointain il sera intéressant d'intégrer l'information sur la dimension verticale des objets, issue des données LIDAR. Elle permettra notamment d'améliorer la discrimination des classes de végétation (Whited *et al.*, 2002 ; Johansen *et al.*, 2010) et de séparer les « sols artificialisés » des « sols semi-naturels » à partie de l'information sur les hauteurs de bâtis. Cette approche a été expérimentée en Australie (Johansen *et al.*, 2010) et à Mayotte (Laine *et al.*, 2010) sur de grands territoires (plus de 350 km²).

En plus de l'apport de nouvelles sources de données, l'exploitation des données utilisées dans la thèse peut être améliorée. Par exemple, l'information contenue dans la donnée RPG n'a été que partiellement exploitée dans cette étude car seul le type de culture occupant majoritairement la superficie des différents îlots a été attribuée à l'espace délimité par ces îlots. Même si cette classification fournit déjà une vision globale détaillée et intéressante des principales pressions agricoles protectrices ou dégradantes s'exerçant sur le milieu aquatique, il semble intéressant de délimiter et de classer les différentes parcelles au sein des îlots RPG afin de préciser l'information liée aux activités agricoles, de mieux localiser les parcelles de prairies participant aux mécanismes d'interception et d'épuration des polluants diffus, de détecter à l'intérieur de ces îlots les haies arborée ou arbustives qui contribuent également activement à ces mécanismes (Lebert, 1993).

En parallèle de ces futurs développements susceptibles d'améliorer la classification de l'occupation des sols dans les corridors rivulaires, un enjeu fort réside dans l'amélioration de la quantification spatialisée des erreurs de classification qui est indispensable pour correctement évaluer l'incertitude des indicateurs spatialisés.
Dans cette thèse deux techniques ont été mises en œuvre pour quantifier la qualité sémantique (i.e., nature des objets) de la carte d'occupation du sol dans le corridor rivulaire : (1) la matrice de confusion qui, en confrontant les résultats de classification à une vérité terrain, apporte sur l'ensemble de la zone d'étude analysée une information synthétique sur la manière dont certaines confusions entre classes affectent les résultats de la classification ; (2) la logique floue qui, en mettant en oeuvre des ensembles flous de règles de classification définis par l'opérateur indépendamment de toute vérité terrain, apporte pour chaque objet une information différenciée sur la manière dont les classes se mélangent au sein de cet objet.

Les informations apportées par les deux techniques de validation sont complémentaires : la première, issue de la matrice de confusion, est primordiale car elle est construite à partir d'une vérité terrain ; la deuxième, issue des propriétés de la logique floue, apporte une information locale et est indispensable pour correctement propager les erreurs de classification au niveau des indicateurs spatialisés car elle permet de spatialiser l'incertitude sur la carte d'occupation du sol. L'enjeu est donc de trouver des méthodes permettant de combiner ces informations complémentaires afin d'avoir une évaluation spatialisée de l'incertitude de la carte d'occupation liée à une vérité terrain.

Cependant, si l'on veut correctement tirer profit des avantages de la logique floue il est indispensable d'améliorer le processus de définition des ensembles flous de chaque règle de classification. L'analyse de fiabilité de la carte d'occupation du sol dans les corridors rivulaires du bassin versant de l'Hérault, menée à partir des propriétés de la logique floue, a évalué comme très fiable la classification alors que la matrice de confusion a mis en évidence des erreurs sur certaines catégories (notamment « sols artificialisés » et « sols nus

[51] SYRAH : SYstème Relationnel d'Audit de l'Hydromorphologie

[52] Le red-edge est un mot d'origine anglaise, signifiant littéralement bord ou marge rouge, correspondant a une portion de spectre dans le proche infrarouge. Le red-edge, renvoyant à la région de changement rapide de la réflectance de la chlorophylle, est très souvent utiliser dans les études d'agriculture de précision (Lebourgeois, 2009)

CemOA : archive ouverte d'Irstea / Cemagref

semi-naturels »). Dans cette étude, l'opérateur par excès de confiance ou manque d'expertise, a défini des ensembles flous trop étroits sur certaines classes, sous-estimant l'ambigüité existante entre les classes (par exemple celles se confondant avec les « sols nus semi-naturels »). Ainsi « fiabilité de la classification » en logique floue signifie « absence d'ambiguïté » interne au mode de classification, mais ne présume pas de l'absence d'erreurs par rapport à la réalité. Deux solutions peuvent être envisagées pour améliorer la définition de ces ensembles flous : soit (1) de redéfinir *a posteriori* les ensembles flous après analyse des résultats de la matrice de confusion, soit (2) de définir ces ensembles à partir d'échantillons d'apprentissage (Glorennec, 1999 ; Guillaume, 2001) pris dans la zone d'étude (lors de définition des règles de classification). Dans les deux cas (plus particulièrement pour la 2ème solution), ces processus sont consommateurs de temps.

Tableau A : Caractéristiques des données images potentiellement accessibles par les gestionnaires à coûts réduits sur le territoire français (avec MEEDM : Ministère de l'Écologie de l'Énergie et du Développement durable et de la Mer ; IGN : Institut Géographique National ; et GEOSUD : GEOInformation for SUstainable Development »).

	Résolution spatiale	Résolution spectrale	Couverture française actuelle	Accessibilité	Producteur
BDortho IRC® (photographies aériennes infra rouge couleur)	0.5 m	Bleu Vert Rouge PIR	Couverture quasi complète avec un renouvellement des photographies tous les 5 ans par département	Licence d'utilisation pour le MEEDM (Convention IGN-MEEDM)	IGN
Images satellites RapidEyes	5 m	Bleu Vert Rouge Red-edge	Couverture 2010 disponible à partir de janvier 2011	Gratuites pour les établissements publics (Projet GEOSUD)	RapidEyes AG

Tableau B : Caractéristiques des données auxiliaires potentiellement accessibles par les gestionnaires à coûts réduits sur le territoire français

	Précision	Contenu	Mise à jour	Accessibilité	Producteur
BDTopo®	métrique	Réseaux routiers, ferroviaires, de transport d'énergie et hydrographiques ; ruptures de pentes artificielles (orographie),	en continu (décalage maximum d'un an)	Licence d'utilisation le MEEDM (Convention IGN-MEEDM)	IGN
		végétation arborée (types de forêts, haies, vignes, vergers…)	décalage maximum 10 ans (cycle de prise de vues + délai de constitution de la BD TOPO®)		

Construction des Indicateurs spatialisés du corridor rivulaire

Dans un deuxième temps (Partie 3), nous nous sommes intéressés à définir et à construire des indicateurs spatialisés de l'occupation du sol dans le corridor rivulaire à partir des outils d'analyse spatiale fournis par les Systèmes d'Information Géographique (SIG), indicateurs spatialisés susceptibles (i) d'expliquer une partie de la variabilité de l'état écologique des cours d'eau à l'échelle d'une station, (ii) d'améliorer la localisation de l'emprise géographique au sein de laquelle ces sources de pressions ont une influence significative sur l'état écologique de la station, et (iii) d'apporter une information quantifiée, synthétique et explicite, pour guider les gestionnaires dans leurs stratégies de restauration et notamment dans les stratégies relatives à la végétation rivulaire.

On a défini de façon générique dans cette thèse l'indicateur spatialisé du corridor rivulaire comme un attribut de structure (composition et/ou configuration) calculé pour une catégorie d'occupation du sol donnée

CemOA : archive ouverte d'Irstea / Cemagref

(selon une certaine typologie) dans une emprise spatiale rivulaire d'une dimension donnée représentée par un buffer délimité autour d'un tronçon de cours d'eau d'une certaine longueur en amont et/ou aval à la station et selon une distance latérale à ce tronçon (largeur de part et d'autre du cours d'eau). A partir de cette définition, un grand nombre d'indicateurs peuvent être construits.

Deux attributs de structure ont été sélectionnés pour traduire l'influence de l'occupation du sol sur l'état écologique des cours d'eau à l'échelle d'une station : (i) le pourcentage surfacique qui traduit la présence et l'intensité des pressions susceptibles d'être générées par une catégorie d'occupation du sol (OS); et (ii) la fragmentation qui nuance l'intensité des pressions générées pour certaines de ces catégories dont l'influence est susceptible de varier en fonction de leur emplacement et de leur connexion au cours d'eau. Par exemple, une bande de végétation arborée fragmentée le long du cours d'eau aura une capacité de rétention plus faible qu'une bande continue.

Ces métriques peuvent être extraites dans plusieurs emprises spatiales traduisant ainsi le rayon d'influence de ces pressions, plus ou moins proche du cours d'eau et de la station. Un grand nombre d'indicateurs peut donc être construit en fonction de la taille de ces emprises. L'utilisation de plusieurs emprises candidates (mode de spatialisation « multiple ») permettra d'améliorer l'identification, dans les modèles pressions/impacts, de l'emprise au sein de laquelle une catégorie d'OS a une influence significative sur la variabilité de la réponse biologique du cours d'eau.

Parmi ces possibilités d'indicateurs, trois familles traduisant des mécanismes d'impact spécifiques et/ou des localisations particulières dans le corridor rivulaire, ont été détaillées. Ces familles d'indicateurs diffèrent en fonction de la méthode et de la distance latérale au cours d'eau prise en compte pour délimiter le buffer d'analyse.

La famille des « indicateurs du linéaire », pour lesquels l'emprise est définie par un buffer aux limites régulières (i.e. qui suivent la linéarité du cours d'eau) d'une distance fixe de 5 m de part et d'autre du cours d'eau, et qui traduisent les pressions d'une catégorie d'occupation du sol donnée qui s'exercent au plus proche du cours d'eau et apportent une information sur l'état des berges rivulaires.

La famille des « indicateurs du lit majeur », pour lesquels l'emprise est définie par un buffer régulier à distance variable (en fonction du rang du cours d'eau), et qui traduisent les pressions exercées par une catégorie d'occupation du sol donnée présente au sein du lit majeur, espace de liberté théorique du cours d'eau. Ces indicateurs caractérisent ainsi l'état global de l'espace rivulaire.

La famille des « indicateurs de contact », pour lesquels l'emprise est définie à partir d'un buffer irrégulier (i.e. dont les limites ne suivent pas la linéarité du cours d'eau) constitué par la fusion des objets au contact du cours d'eau, et qui traduisent les pressions exercées par les objets d'une catégorie d'occupation du sol donnée directement au contact du cours d'eau. Ces indicateurs permettent de mettre en évidence la présence de ruptures dans la bande de végétation rivulaire et de préciser l'intensité de ces pressions au niveau de ces ruptures.

Six types d'indicateurs spécifiques au rôle des ripisylves ont été également définis : (1) la longueur et (2) la fragmentation de la ripisylve qui appartiennent à la famille des indicateurs du « linéaire » ; (3) la superficie de la ripisylve qui appartient à la famille des indicateurs de « contact » ; (4) la densité de ruptures (nombre de patchs distincts par km de cours d'eau), (5) la largeur moyenne (en m) et (6) l'uniformité (en fonction de la variation de la largeur des patchs) de la ripisylve définies à partir du buffer irrégulier de contact. Ces indicateurs permettent de quantifier de façon générale la présence et l'intensité de cette végétation le long du cours (longueur et superficie), sa connectivité (fragmentation et densité de ruptures) caractéristique essentielle de son fonctionnement, mais aussi plus spécifiquement sa capacité à jouer son rôle de barrière vis-vis des polluants diffus (fragmentation, densité de ruptures, largeur moyenne et uniformité).

Tous ces indicateurs sont construits de façon automatique. La délimitation automatisée des différents buffers a nécessité le développement d'algorithmes SIG spécifiques notamment pour définir la distance amont et ou/aval à la station. La définition de cette distance se base sur les outils de la segmentation dynamique utilisant les systèmes de repérage curviligne et les informations topologiques du réseau hydrographique.

Cette thèse apporte donc aujourd'hui une définition générique et des méthodologies automatisées pour construire, à partir d'une donnée d'occupation du sol, des indicateurs spatialisés du corridor rivulaire susceptibles d'expliquer la variabilité de l'état écologique au niveau d'une station et caractérisant l'état de

CemOA : archive ouverte d'Irstea / Cemagref

l'occupation du sol dans le corridor et notamment de la végétation le long du cours d'eau. Elle a démontré par la mise en œuvre de certains de ces indicateurs sur le bassin de l'Hérault : (i) l'intérêt déterminant de la donnée à THRS pour quantifier correctement ces indicateurs au niveau du corridor rivulaire ; et (ii) la capacité de ces indicateurs à fournir une information cartographique explicite pour guider les gestionnaires dans leurs stratégies de restauration sur le corridor rivulaire.

Perspectives

À partir de cette définition générique et de ces méthodologies automatisées, il sera dès lors intéressant d'identifier de nouveaux indicateurs spatialisés du corridor rivulaire, relatifs aux mécanismes d'influence sur l'état écologique des cours d'eau, en collaboration avec des experts du domaine et les gestionnaires. L'utilisation de ces indicateurs permettra d'intégrer dans les modèles pressions/impacts leurs connaissances, ce qui facilitera par la suite l'interprétation des modèles et leur utilisation pour prioriser et localiser les actions à mener sur leurs territoires. Ces indicateurs pourront être construits à partir de la donnée d'occupation à THRS dans ces espaces mais aussi en combinant des sources d'information complémentaires qui sont ou seront disponibles sur leurs territoires afin de préciser les mécanismes d'impacts. Par exemple, l'utilisation de l'information à une résolution adaptée à ces systèmes sur les pentes des versants (issue des données LIDAR ou Radar à une résolution métrique ou décamétrique) et la géologie de fonds de vallée (issue des cartes du BRGM[53] au 1 / 50 000), permettront de préciser les pressions potentielles exercées par les terres arables en terme de pollution diffuse et de caractériser plus précisément l'efficacité des zones tampons le long des cours d'eau.

Bien évidemment, une information spatialisée sur les pratiques agricoles par types de culture sur le territoire permettrait de traduire plus précisément l'apport de pollution diffuse par substances (pesticides, nitrates, phosphates…), mais une telle donnée reste encore aujourd'hui difficilement mobilisable à l'échelle nationale.

Par ailleurs, au-delà de l'enjeu du bon état écologique des cours d'eau de la DCE, il sera intéressant d'utiliser ces méthodologies pour préparer la politique « trames vertes et bleues » concernant les corridors de biodiversité à l'échelle nationale. La construction d'indicateurs d'évaluation de l'état des corridors rivulaires pertinents à large échelle est un enjeu majeur pour la mise œuvre de cette politique. De tels indicateurs sont nécessaires à la fois pour dresser un « état des lieux » national et suivre les effets des actions entreprises.

L'étude de la sensibilité de ces indicateurs aux incertitudes des cartes d'occupation du sol est indispensable pour en tenir compte dans la définition des stratégies de restauration. Les recherches menées à ce sujet restent inachevées : les méthodologies pour étudier l'influence de l'incertitude des cartes d'occupation des sols sont clairement identifiées (tirages de Monte-Carlo) mais les difficultés rencontrées pour correctement spatialiser les erreurs de la donnée d'occupation du sol, au-delà de la seule matrice de confusion qui n'est pas spatialisée, demeurent un point de blocage à lever.

Prise en compte des corridors rivulaires dans l'élaboration des modèles pressions/impacts régionalisés

Enfin dans un troisième temps (partie 4), une démarche de construction de modèles pressions/impact régionalisés a été mise en place et une analyse comparative de différentes structures de modèles a été menée afin (1) de quantifier l'apport respectif des avancées méthodologiques développées dans cette thèse (cartographie à THRS de l'occupation des sols, indicateurs d'occupation du sol calculés sur différentes emprises du corridor rivulaire) dans l'étude des relations pressions/impacts et dans la prise en compte du rôle du corridor rivulaire ; et (2) de sélectionner le modèle le plus pertinent afin de donner des éléments de réponse aux gestionnaires sur l'identification et la localisation des facteurs dégradant et des facteurs protégeant l'état écologique sur la zone d'étude.

Pour cela plusieurs modèles pressions/impacts ont été construits sur la zone d'étude (partie Nord-Ouest de l'hydro-écorégion Tables Calcaires), ces modèles différant soit par la combinaison d'échelles d'analyse (bassin versant et/ou corridor rivulaire amont et/ou corridor rivulaire local), soit par la source de donnée d'occupation du sol mobilisée au niveau des deux échelles du corridor rivulaire (donnée CORINE Land Cover ou donnée à THRS), soit par le mode de spatialisation de l'emprise au niveau du corridor rivulaire local (emprise unique pour le mode fixe ou plusieurs emprises pour le mode multiple).

[53] BRGM : Bureau des Ressources Géologiques et Minières

CemOA : archive ouverte d'Irstea / Cemagref

Tous les modèles développés avaient pour finalité d'expliquer statistiquement, en se basant sur une population de 155 stations écologiques, la variabilité de l'EQR-IBGN$_{moyen}$ (écart entre la référence et la moyenne de l'Indice Biologique Global Normalisé sur la période de 1992 à 2004) observé aux stations à partir d'indicateurs surfaciques d'occupation du sol sur le BV amont et le corridor rivulaire amont et local de ces stations.

Une procédure automatisée en deux étapes a été mise en place pour construire ces modèles qui peuvent intégrer en entrée plus de 8000 variables explicatives dont certaines sont fortement colinéaires. La première étape sélectionne pour chaque source de pression à une échelle donnée, selon une approche comparative mono-variable, l'emprise spatiale sur laquelle cette source a l'influence la plus significative sur l'état écologique. Cette étape est indispensable pour réduire la dimension des modèles. Ensuite, la deuxième étape basée sur l'algorithme de régression Partial Least Square (adapté aux problèmes de colinéarités entre variables et aux problèmes de surparamétrisation) cherche de façon automatique la combinaison des différentes sources de pressions (s'exprimant sur leur emprise spatiale la plus significative) aux différentes échelles expliquant au mieux la variabilité de la réponse biologique aux stations.

La comparaison des résultats des différents modèles a montré que les avancées méthodologiques (donnée à THRS, mode de spatialisation multiple) permettent d'améliorer significativement le pouvoir explicatif des modèles, d'identifier un plus grand nombre de sources de pressions influençant significativement la réponse biologique et de préciser leur emprise spatiale significative. Elles fiabilisent donc la compréhension des relations pressions/impacts et apportent de nouvelles informations qui peuvent être pertinentes pour la gestion.

Cette comparaison a permis également de sélectionner le modèle le plus pertinent pour établir un diagnostic de l'état écologique sur la zone d'étude. Ce modèle utilise les avancées méthodologiques et étudie les impacts de l'occupation du sol qui interagissent sur l'ensemble des trois échelles (bassin versant, corridor rivulaire amont et corridor rivulaire local). Il explique environ 38 % de la variabilité de la réponse biologique (ce qui est un bon résultat compte tenu de la distance, en termes de processus écologiques, entre les indicateurs d'occupation du sol estimés à partir de l'imagerie satellitaire et la réponse biologique) et identifie une combinaison de 23 sources de pressions significatives et les emprises spatiales associées.

L'interprétation de ce modèle confirme, par rapport aux études équivalentes menées sur le territoire français, l'impact marqué des terres artificialisées aussi bien à l'échelle du bassin versant (BV) qu'aux échelles du corridor rivulaire local et amont, ainsi que l'impact des terres arables à l'échelle du BV sur la diversité et la polluosensibilité des macro-invertébrés benthiques. Ce modèle montre également que l'objectif de « bon état » ne semble pas irréaliste dans ces régions très agricoles puisqu'il met en évidence les effets protecteurs dans une bande rivulaire gérable (20 m de part et d'autre du cours d'eau) de la végétation arborée et de la végétation herbacée permanente qui ont une contribution positive cumulée sur l'EQR-IBGNmoyen de même amplitude que l'effet négatif des terres arables sur le BV. Toutefois, il indique aussi que la possibilité de gagner quelques points d'IBGN (quelques dixièmes d'EQR-IBGN) ne devra pas s'appuyer uniquement sur la restauration/préservation des espaces rivulaires (même si ces espaces conditionnent en partie les processus de fonctionnement) mais également sur la restauration/préservation de l'état hydromorphologique des rivières compte tenu de la nature des impacts mis en évidence (impacts des infrastructures liées au réseau routier notamment).

Cette thèse apporte donc aujourd'hui une méthodologie automatisée de construction et sélection des modèles pressions/impacts cherchant la meilleure combinaison d'indicateurs d'occupation du sol explicatifs de la variabilité de l'état écologique, parmi un grand nombre d'indicateurs calculés aux différentes échelles de fonctionnement de l'écosystème rivière. La mise en œuvre de ces modèles sur la partie Nord-Ouest de l'hydro-écorégion Tables Calcaires montre l'intérêt de disposer d'une donnée à THRS et d'indicateurs calculés sur différentes emprises spatiales pour améliorer l'identification et la localisation des sources de pressions dans ces espaces. Grâce à ces avancées méthodologiques l'influence significative sur l'état écologique des cours d'eau des formations végétales rivulaires arborées et herbacées locales, dans une bande de 20 m de part et d'autre du cours d'eau, a été démontrée à l'échelle régionale.

Toutefois, ces résultats restent à confirmer puisque cette analyse statistique a été mise en œuvre uniquement sur un jeu de donnée (155 stations), étudiant un seul compartiment biologique (les

CemOA : archive ouverte d'Irstea / Cemagref

macroinvertébrés). Il sera donc nécessaire de réitérer une telle analyse dans différentes hydro-écorégions, sur des jeux de données assez grands pour être représentatifs de l'ensemble des sources de pressions s'exerçant sur les cours d'eau, et en étudiant la réponse de différents indicateurs de l'état écologique des cours d'eau (macroinvertébrés, poissons, diatomées…).

De plus, ces résultats sont à interpréter avec une certaine prudence. D'une part, la fiabilité des résultats vis-vis de l'incertitude sur les indicateurs n'a pas été évaluée : tant pour les indicateurs d'occupation du sol, du fait des difficultés rencontrées pour spatialiser l'incertitude de la carte d'occupation du sol dans le corridor rivulaire, que pour les indicateurs biologiques dont les méthodes de caractérisation de l'incertitude restent encore à développer (Archaimbault and Dumont, 2010). D'autre part, les dépendances d'échelles, permettant de vérifier l'effet propre d'un indicateur d'occupation du sol à une échelle donnée sur la variabilité de la réponse biologique, et de confronter l'effet des indicateurs d'occupation du sol à une échelle par rapport à une autre, n'ont pas été pleinement analysées. Des techniques statistiques robustes de partitionnement de la variance sont disponibles aujourd'hui pour quantifier l'influence relative des facteurs aux différentes échelles. Elles ont été mises en place dans diverses études de modélisation pressions/impacts (e.g., e.g., Richards *et al.*, 1997 ; Johnson *et al.*, 2004 ; Johnson *et al.*, 2007 ; Stoffels *et al.*, 2005 Gido *et al.*, 2006 ; Kratzer *et al.*, 2006). Une autre approche consiste à construire séparément des modèles selon différentes combinaisons d'échelles et à comparer les résultats en se basant notamment sur la variance expliquée (Lattin *et al.*, 2004, Dow *et al.*, 2006).

Perspectives

En parallèle de ces analyses, de nombreuses questions de recherche sont ouvertes pour améliorer la capacité des modèles pressions/impacts à comprendre et prédire l'état écologique des cours d'eau :

(1) Améliorer ***l'information sur les sources de pressions*** : l'occupation du sol, une des forces motrices générant la plupart des impacts sur le milieu aquatique, (i) ne rend pas compte de l'ensemble des pressions comme les barrages, les dérivations, la navigation, l'endiguement issus d'autres forces motrices (énergie, transport, tourisme) altérant particulièrement les processus hydromorphologiques de l'écosystème rivière et (ii) ne permet pas de quantifier précisément les apports de polluants diffus modifiant les processus biogéochimiques de cet écosystème. Il est donc important de se doter de méthodes capables de quantifier ces sources de pressions. On pourra explorer la possibilité d'utiliser, en complément des indicateurs d'occupation du sol, (i) les indicateurs d'altération hydromorphologiques fournis par le projet SYRAH (Chandesris *et al.*, 2007) complété par le Référentiel national des Obstacles à l'Ecoulement (ROE mise en place par l'ONEMA) ; et (ii) les concentrations et les flux de polluants (tels que les nitrates et les phosphates) que sont capables de calculer les modèles de flux biogéochimiques fonctionnant à l'échelle des grands bassins hydrographiques (tels que SENEQUE Billen *et al.*, 1994 ; Garnier *et al.*, 1995, et PEGASE Smitz *et al.*, 1997) ;

(2) Améliorer ***l'information des indicateurs biologiques*** : dans notre approche on a utilisé l'Indice Biologique Global Normalisé basé sur des métriques répondant spécifiquement à un faible spectre de pressions. Il sera intéressant d'utiliser des métriques biologiques (de structure et de fonctionnement) répondant spécifiquement à un grand nombre de pressions anthropiques, ce qui facilitera la recherche des causes d'altérations (amélioration de leur valeur de diagnostic) pour orienter les stratégies de restauration. Ces travaux sont en cours notamment au niveau des macroinvertébrés benthiques puisque le protocole IBGN a été modifié au niveau du prélèvement et du traitement des échantillons afin d'améliorer l'information contenue dans cet indice (Usseglio Polatera *et al.*, 2007). Il permettra, en plus de l'IBGN, d'évaluer des métriques plus fonctionnelles, utilisant les traits bio/écologiques des organismes et des indices de structure des peuplements qui permettront la mise en œuvre d'un nouvel indice multi-métrique d'évaluation de l'état écologique. En plus de ces améliorations sur le contenu informatif, des efforts devront être menés (i) sur la quantification des incertitudes que les différents indices véhiculent afin de mieux appréhender leurs limites d'utilisation et d'interprétation ; et (ii) sur leur adaptation aux grands cours d'eau notamment pour les macroinvertébrés aussi bien au niveau du protocole de prélèvement que sur la constitution des indices (Archaimbault and Dumont, 2010).

(3) Améliorer ***la sélection des emprises spatiales significatives*** : l'approche de sélection mono-variable développée présente des inconvénients. L'effet « combiné » des sources de pressions n'est pas pris en compte par cette méthode de sélection. L'emprise spatiale significative pour une source de pressions donnée devrait dans l'absolu correspondre à l'emprise sur laquelle cette source de pression combinée aux autres (sur leurs emprises respectives) a la plus forte influence sur l'état écologique. Sa sélection impliquerait donc une analyse de corrélations multi-variables qui identifierait directement la combinaison de source de pressions et de leurs emprises respectives, favorisant l'explication de la variabilité de la réponse biologique. Plusieurs pistes peuvent

CemOA : archive ouverte d'Irstea / Cemagref

être envisagées : soit adapter la méthode de régression PLS de façon à sélectionner parmi les emprises candidates (aux différentes échelles) la combinaison de sources de pressions significatives la plus explicative de la réponse biologique, soit rechercher d'autres algorithmes de fouille de données conçus pour gérer cette quantité très importante de variables et qui sont adaptés aux problèmes de multicolinéarité entre les variables explicatives.

(4) Prendre en compte ***les « effets de nomenclature »*** : dans notre étude un seul niveau de nomenclature (donc d'information) dans la construction des indicateurs d'occupation du sol a été testé pour expliquer la réponse biologique (niveau le plus détaillé). Le regroupement de certaines catégories d'occupation peut changer l'information apportée par l'indicateur et améliorer la compréhension de l'état écologique. En plus de la sélection des emprises spatiales significatives, se rajoutent alors la question du niveau de nomenclature optimal pour la construction des modèles pressions/impacts.

(5) Prendre en compte ***« l'effet d'héritage » (***Allan, 2004a ; Johnson and Host, 2010): les réponses biologiques n'apparaissent souvent qu'après un temps de latence plus ou moins long, plusieurs années s'avérant nécessaires pour observer, par exemple, les effets d'un dépôt acide, d'une espèce introduite ou d'une pollution des eaux souterraines. La prise en compte de cet effet nécessite un historique de données de pressions conséquent, actuellement inexistant, sachant que ces données sont déjà très difficiles à acquérir sur une zone d'étude à un instant donné ;

(6) Prendre en compte ***la « non-linéarité »*** des réponses biologiques à ces perturbations (Allan, 2004a ; Décamps, 2006 ; Johnson and Host, 2010): les effets s'ajoutent-ils les uns aux autres de façon linéaire, s'amplifient-ils, introduisent-ils des seuils plus sévères ? La réponse à ces questions est loin d'être triviale et demande l'utilisation d'outils mathématiques complexes (difficiles à maitriser pour l'analyste) et surtout l'identification précise des pressions s'exerçant sur le milieu, l'occupation du sol informant uniquement sur un panel non exhaustif de pressions potentielles pouvant s'exercer sur le milieu ;

(7) Prendre en compte ***« l'autocorrélation spatiale »*** : les observations d'état écologique en différentes stations peuvent être corrélées du fait de leur répartition spatiale dans la zone d'étude. C'est un phénomène bien ancré en écologie du paysage et de nouvelles techniques existent pour répondre à cette difficulté (Legendre, 1993 ; Thompson et al., 1996). L'enjeu est de taille au niveau de l'écosystème rivière, puisque cette dépendance spatiale est liée à un réseau arborescent orienté (amont-aval) : peu d'études (King *et al.*, 2005; Maillard and Pinheiro Santos, 2006) traitent cette difficulté dans les modèles pressions/impacts.

(8) Prendre en compte les ***« effets du réchauffement climatique »*** : le réchauffement climatique est une force motrice à part entière qui influe sur l'ensemble des variables du système aussi bien naturelles qu'anthropiques. Il impacte directement (i) la température d'un cours d'eau qui joue un rôle fondamental dans la dynamique des écosystèmes aquatiques puisqu'elle influe sur la distribution spatiale des organismes vivants, leur croissance, leur reproduction, leur comportement ou encore leur état sanitaire ; et (ii) sur le régime des précipitations régulant la dynamique fluviale moteur de l'écosystème. Ces bouleversements attendus dans un futur proche vont impacter la vie dans les milieux aquatiques et probablement modifier leur sensibilité aux pressions. De plus, le réchauffement climatique va engendrer des modifications sur les activités humaines (modifications des pratiques agricoles par exemple avec des prélèvements d'eau plus sévères) et donc à la fois sur la nature et l'intensité des pressions qui s'exerceront sur l'état écologique.

Malgré ce large spectre de perspectives et de questions de recherche ouvertes, le délai imposé par la Directive Cadre européenne sur l'Eau, qui vise à ramener les masses d'eau à un « bon état » écologique d'ici 2015, impose de se limiter dans un premier temps aux sources majeures de pressions (l'occupation du sol) aux différentes échelles de fonctionnement, en mobilisant les connaissances des experts et des gestionnaires afin de définir des indicateurs d'occupation du sol pertinents pour orienter, prioriser et localiser les actions à mettre en œuvre sur les territoires.

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

LISTE DES FIGURES

FIGURE 0.1 : SCHEMATISATION DES FLUX SELON LES QUATRE DIMENSIONS DE L'HYDROSYSTEME (D'APRES AMOROS AND PETTS, 1993) :..... 20
FIGURE 0.2 : LES PRINCIPALES PROBLEMATIQUES DE RECHERCHES LIEES A LA MISE EN ŒUVRE DCE A PARTIR DE 2000 (WASSON, 2001)..... 22
FIGURE 0.3 : INFLUENCE DES CORRIDORS RIVULAIRES SUR LES PROCESSUS FONCTIONNELS DE L'ECOSYSTEME D'EAU COURANTE (PHYSIQUES, BIOGEOCHIMIQUES ET ECOLOGIQUES) (D'APRES MARIDET AND COLLIN HUET, 1995). 23
FIGURE 0.4 : DEMARCHE SCIENTIFIQUE DERIVEE DE L'ECOLOGIE DU PAYSAGE ADOPTEE DANS LES MODELES PRESSIONS / IMPACTS : LE PAYSAGE AQUATIQUE EST LA RESULTANTE DE LA DYNAMIQUE DU MILIEU ET DE LA SOCIETE QUI S'Y EST DEVELOPPEE. LA STRUCTURE, L'ORGANISATION, LA DYNAMIQUE DU PAYSAGE AQUATIQUE (L'HYDROSYSTEME) SONT EN INTERACTION CONSTANTE AVEC LES PROCESSUS FONCTIONNELS QUI S'Y DEROULENT (MODIFIEE DE BUREL AND BAUDRY, 1999). 25
FIGURE 0.5 : COMPARAISON D'IMAGES A MOYENNE RESOLUTION, LANDSAT ETM+ (30 M), A HAUTE RESOLUTION, SPOT 5 XS (10 M) ET A TRES HAUTE RESOLUTION SPATIALE, ORTHOPHOTO IGN (0.5 M) SUR UNE PORTION DU FLEUVE HERAULT AU NIVEAU DE SA PLAINE. L'IDENTIFICATION DES STRUCTURES RIVULAIRES EST NETTEMENT PLUS VISIBLE SUR LES IMAGES A HAUTE ET TRES HAUTE RESOLUTION SPATIALE QUI FOURNISSENT UNE INFORMATION PLUS RICHE ET PLUS DETAILLEE..... 26
FIGURE 1.1 : LES TROIS CONCEPTS CARACTERISANT LE « BON ETAT » : PROCESSUS, DYNAMIQUE, REVERSIBILITE (D'APRES WASSON *ET AL.*, 2001)..... 39
FIGURE 1.2 : SCHEMA D'ORGANISATION DE LA PHASE DE D'EVALUATION DE LA DCE (D'APRES WASSON, 2001)..... 41
FIGURE 1.3 : MODELE CONCEPTUEL DE L'ECOSYSTEME D'EAU COURANTE D'APRES WASSON, 1993..... 44
FIGURE 1.4 : METHODE D'ORDINATION DES COURS D'EAU SELON STRAHLER, 1957. L'ORDINATION CONSISTE A ATTRIBUER LE RANG 1 AUX COURS D'EAU EN TETE DE BASSIN, PUIS, EN PROGRESSANT VERS L'AVAL, CHAQUE COURS D'EAU RECEVANT UN COURS D'EAU DU MEME RANG VOIT SON RANG AUGMENTER D'UNE UNITE. 44
FIGURE 1.5 : CARTE DES HYDRO-ECOREGIONS (HER-1) DE FRANCE METROPOLITAINE D'APRES WASSON *ET AL.*, 2002 46
FIGURE 1.6 : METRIQUES FONCTIONNELLES REQUISES PAR LA DCE EN FONCTION DU PEUPLEMENT AQUATIQUE ETUDIE ET METRIQUES PRISES EN COMPTE PAR L'INDICE BIOLOGIQUE GLOBAL NORMALISE (IBGN) (WASSON *ET AL.*, 2004B)..... 48
FIGURE 1.7 : CARTOGRAPHIE DES STATIONS IBGN EN FRANCE D'APRES LA BASE DE DONNEES GIRAFE COUVRANT LA PERIODE DE 1992 A 2004 50
FIGURE 1.8 *: PRINCIPE DE CALCUL DE LA LIMITE DU BON ETAT POUR L'IBGN*..... 51
FIGURE 1.9 : DESCRIPTION DU PROCESSUS DE MODELISATION ET D'EXTRAPOLATION SPATIALE QUI SE DEROULE EN QUATRE TEMPS (WASSON *ET AL.*, 2006) :..... 52
FIGURE 2.1 : LES CORRIDORS RIVULAIRES REMPLISSENT DEUX FONCTIONS ESSENTIELLES POUR LA PROTECTION DE LA BIODIVERSITE : DANS LE SENS LONGITUDINAL (FLECHE NOIRE), ILS METTENT EN CONNEXION DES HABITATS ELOIGNES, COMME DEUX MASSIFS FORESTIERS ; DANS LE SENS TRANSVERSAL (FLECHES ROUGES), CE SONT DES ECOTONES ENTRE DES FORMATIONS BOISEES ET OUVERTES, ET ENTRE LE MILIEU AQUATIQUE ET LE MILIEU TERRESTRE (CSPNB, 2008)..... 62
FIGURE 2.2 : INFLUENCE DE LA RIPISYLVE SUR LE FONCTIONNEMENT ECOLOGIQUE DES ECOSYSTEMES D'EAU COURANTE. FONCTIONS ASSUREES PAR LA RIPISYLVE. EFFET SPECIFIQUE SUR L'HABITAT AQUATIQUE (D'APRES MARIDET, 1995)..... 64
FIGURE 2.3 : INFLUENCE DE LA RIPISYLVE SUR LE FONCTIONNEMENT TROPHIQUE DANS LES ECOSYSTEMES D'EAU COURANTE (D'APRES MARIDET, 1995). 66
FIGURE 2.4 : INFLUENCE D'UN CORRIDOR RIVULAIRE BOISE SUR LA TEMPERATURE MAXIMALE DE PETITS COURS D'EAU, EN FONCTION DE LA LARGEUR DU CORRIDOR ET DE LA LONGUEUR PROTEGEE A L'AMONT DU SITE ; D'APRES L'ETUDE DE 40 SITES EN ONTARIO, CANADA (D'APRES NAIMAN *ET AL.*, 2005). 66
FIGURE 2.5 : PRINCIPAUX PROCESSUS D'EPURATION DES EAUX POLLUEES EN NITRATES PAR LES RIPISYLVES : (A) EN PERIODE DE HAUTES EAUX : DENITRIFICATION BACTERIENNE ANAEROBIE ; ET (B) EN PERIODE DE CROISSANCE VEGETALE : ABSORPTION RACINAIRE (D'APRES MARIDET, 1995). 68
FIGURE 2.6 : RELATIONS ENTRE LES RIPISYLVES, LE COURS D'EAU ET LE VERSANT EN FONCTION DE LEUR POSITION LE LONG DU RESEAU HYDROGRAPHIQUE (D'APRES TABACCHI *ET AL.*, 1998). 70
FIGURE 3.1 : SCHEMA DPSIR POUR LA REPRESENTATION DE L'ACTION HUMAINE SUR L'ETAT ECOLOGIQUE DES COURS D'EAU 78
FIGURE 3.2 : PRINCIPES DU MODELES SENEQUE RUELLAND, 2004. 84
FIGURE 3.3 : CARACTERISATION DE LA ZONE RIVULAIRE DANS LE RIPARIAN ECOSYSTEM MANAGEMENT MODEL (REMM). CE SCHEMA A ETE CONÇU A PARTIR DU MODELE DE BUFFER EN TROIS BANDES D'APRES WELSCH, 1991. 85
FIGURE 3.4 : RELATIONS EMPIRIQUE ENTRE LE TAUX D'ELIMINATION D'AZOTE ET LA LARGEUR DE BUFFER D'APRES LES RESULTATS ISSUS DE MODELE MECANISTES OU D'EXPERIENCE IN-SITU D'APRES MAYER *ET AL.*, 2007*:* 87
FIGURE 3.5 : LES MANQUES DE CONNAISSANCES DANS LES APPROCHES MECANISTES POUR EVALUER L'ETAT ECOLOGIQUE DES COURS D'EAU AU SENS DE LA DCE D'APRES WASSON *ET AL.*, 2003B. 88

CemOA : archive ouverte d'Irstea / Cemagref

FIGURE 3.6 : PRINCIPE DES MODELES PRESSIONS IMPACTS EMPIRIQUES. L'ENSEMBLE DES VOIES DOIVENT ETRE PRISES EN COMPTE POUR ETABLIR DES RELATIONS SIGNIFICATIVES ENTRE UN INDICATEUR DU PAYSAGE ET UNE REPONSE BIOLOGIQUE D'APRES ALLAN, 2004A. 90
FIGURE 3.7 : HABITAT ET MORPHOLOGIE : DES ECHELLES EMBOITEES D'APRES WASSON *ET AL.*, 1998. 91
FIGURE 3.8 : ILLUSTRATION DES DISTANCES EUCLIDIENNES A LA STATION (A GAUCHE) ET AU COURS D'EAU (A DROITE), (D'APRES KING *ET AL.*, 2005). 96
FIGURE 3.9 : ILLUSTRATION DE LA MODELISATION DES DISTANCES D'UNE CATEGORIE D'OCCUPATION DU SOL A UN POINT DONNE DU BASSIN VERSANT AU COURS D'EAU (DIRECTION DES FLUX ESTIMEE A PARTIR D'UN MNT) ET A LA STATION (DISTANCE CURVILIGNE DU RESEAU) (D'APRES VAN SICKLE AND JOHNSON, 2008). 96
FIGURE 3.10 : ÉCHELLES DE TRAVAIL GENERALEMENT UTILISEES DANS LES MODELES PRESSIONS/IMPACTS D'APRES ALLAN, 2004B... 97
FIGURE 3.11 : COMPARAISON ENTRE UNE PHOTOGRAPHIE AERIENNE (0.5 M DE RESOLUTION SPATIALE) ET LA DONNEE CORINE LAND COVER (CLC) SUR UN BUFFER DE 600 M DE PART ET D'AUTRES D'UN TRONÇON DE COURS D'EAU (FLEUVE HERAULT AU NIVEAU DE SA PLAINE). L'UNITE CARTOGRAPHIQUE DE CLC (25 HA MIN) NE PERMET PAS DE QUANTIFIER PRECISEMENT LA STRUCTURE DE L'OCCUPATION DU SOL DANS CES ESPACES. LES RIPISYLVES (VEGETATION RIVULAIRE) NE SONT PAS DETECTEES. 103
FIGURE 4.1 : LOCALISATION DE LA ZONE PILOTE : LE BASSIN VERSANT DE L'HERAULT ET PRESENTATION DE SES CARACTERISTIQUES HYDROGRAPHIQUES ET PAYSAGERES (HYDRO-ECOREGION). 112
FIGURE 4.2 : PRESENTATION DE LA ZONE DE MODELISATION, NORD OUEST DE L'HER TABLES-CALCAIRES (HER9-NO) 114
FIGURE 4.3 : OCCUPATION DU SOL SUR LA ZONE D'ETUDE EN FONCTION DE 5. SOURCE : DONNEE CORINE LAND COVER 2000 (IFEN) 115
FIGURE 4.4 : PRINCIPALES ACTIVITES AGRICOLES SUR LA ZONE DE MODELISATION. SOURCE : DONNEE OTEX (ORIENTATION TECHNICO-ÉCONOMIQUE DES EXPLOITATIONS) 18 POSTES (RGA) 116
FIGURE 5.1 : SIGNATURES SPECTRALES TYPIQUES (EAU, SOL, VEGETATION) EN REFLECTANCE (RAPPORT DE LA LUMINANCE SUR L'ENERGIE INCIDENTE) ET EXEMPLES DES BANDES MULTI-SPECTRALES RETENUES POUR LE SPOT5 UN SATELLITE DEDIE A L'OBSERVATION DE LA TERRE (D'APRES UN DOCUMENT CNES IN CARON *ET AL.*, 2003) 128
FIGURE 5.2 : COMPORTEMENT SPECTRALE DE LA VEGETATION (MODIFIE DE GIRARD AND GIRARD, 1999) 131
FIGURE 5.3 : LES SEPT ETAPES DE LA PROCEDURE DE CLASSIFICATION DE L'OCCUPATION DU SOL DANS LES CORRIDORS RIVULAIRE. .. 136
FIGURE 5.4 : CARACTERISTIQUES DES IMAGES SATELLITE SPOT5 XS : 140
FIGURE 5.5 : CARACTERISTIQUES DES IMAGES SATELLITE SPOT5 PAN: 140
FIGURE 5.6 : CARACTERISTIQUES DES ORTHOPHOTOS : 140
FIGURE 5.7 : SOURCES DE DONNEES OPTIMISEES POUR LA CLASSIFICATION DE L'OCCUPATION DANS LES CORRIDORS RIVULAIRES (ETAPE 1) 144
FIGURE 5.8 : DELIMITATION DES CORRIDORS RIVULAIRES (ETAPE 2) EN CREANT DES BUFFERS AVEC UNE LARGEUR DEPENDANT DU RANG DE STRAHLER. 146
FIGURE 5.9 : STRATIFICATION DE LA ZONE D'ETUDE EN REGIONS DE TYPES D'OCCUPATION DU SOL HOMOGENES (ETAPE 3) A PARTIR DES LIMITES DES HYDRO-ECOREGIONS ET DES EMPRISES D'IMAGES DE MEME DATE. 147
FIGURE 5.10 : PRESENTATION DES PRETRAITEMENTS APPLIQUES (ETAPE 4) SUR LES SOURCES DE DONNEES IMAGES ET EXOGENES UTILISEES POUR CLASSER L'OCCUPATION DU SOL DANS LES CORRIDORS RIVULAIRES. 148
FIGURE 5.11 : CONSTRUCTION DU PLAN THEMATIQUE EN FONCTION DE LA PRECISION DE CHAQUE SOURCE DE DONNEES EXOGENES (CLC=CORINE LAND COVER, BDC=BDCARTHAGE®, BDR=ROUTE 500®, RPG=REGISTRE PARCELLAIRE GRAPHIQUE). .. 150
FIGURE 5.12 : ARBRE DE CLASSIFICATION DESCENDANT. LES CLASSES FILLES HERITENT DES PROPRIETES DE CLASSE MERE. 152
FIGURE 5.13 : REPRESENTATION ABSTRAITE D'UNE IMAGE D'APRES UN RESEAU HIERARCHIQUE D'OBJETS POSSEDANT TROIS NIVEAUX DE SEGMENTATION MODIFIES DE BENZ *ET AL.*, 2004 152
FIGURE 5.14 : AUTOMATISATION (OU FORMALISATION DE LA CONNAISSANCE) DE L'INTERPRETATION VISUELLE D'UNE IMAGE (A) PAR L'ANALYSE ORIENTEE OBJET (B). L'EXEMPLE PRESENTEE DANS CETTE ILLUSTRATION CONCERNE LA CLASSIFICATION DES DIFFERENTS TYPES DE FORETS (FEUILLUS, CONIFERES, MIXTES). 153
FIGURE 5.15 : DEFINITION DU PARAMETRE D'HETEROGENEITE F DE L'ALGORITHME DE SEGMENTATION MULTI-RESOLUTION 156
FIGURE 5.16 : EXEMPLE DE FONCTION D'APPARTENANCE SIGMOÏDE POUR DECRIRE UN ENSEMBLE FLOU : LA VALEUR D'APPARTENANCE A UNE CLASSE EN FONCTION DU CRITERE X ENTRE LA VALEUR 50 ET 100. SI LA VALEUR X DE L'OBJET = 95 ALORS SA VALEUR D'APPARTENANCE = 0.9 POUR LA CLASSE. PAR CONTRE SI X = 80 ALORS SA VALEUR D'APPARTENANCE = 0.2. 157
FIGURE 5.17 : PRINCIPE DE LA CLASSIFICATION PAR PLUS PROCHE VOISIN (DEFINIENS, 2008) 159
FIGURE 5.18 : FONCTION D'APPARTENANCE GENEREE AUTOMATIQUEMENT PAR LE LOGICIEL (DEFINIENS, 2008) 159
FIGURE 5.19 : REPRESENTATION DE LA COMBINAISON DES FONCTIONS D'APPARTENANCE GENEREE AUTOMATIQUEMENT PAR LE LOGICIEL DANS UN ESPACE A UNE DIMENSION (DEFINIENS, 2008) 160
FIGURE 5.20 : REPRESENTATION DE LA COMBINAISON DES FONCTIONS D'APPARTENANCE GENEREE AUTOMATIQUEMENT PAR LE LOGICIEL DANS UN ESPACE A DEUX DIMENSIONS (DEFINIENS, 2008) 160

CemOA : archive ouverte d'Irstea / Cemagref

FIGURE 5.21 : EVOLUTION DES MOYENNES D'OBJETS BIEN CLASSES (PRECISION GLOBALE) EN FONCTION DE L'AUGMENTATION DE LA TAILLE DE L'ECHANTILLON D'APPRENTISSAGE (EN NOMBRE D'OBJETS PAR CLASSE) POUR PLUSIEURS TECHNIQUES DE CLASSIFICATION SUPERVISEE (AVEC OU SANS METHODE D'OPTIMISATION DE VARIABLE) APPLIQUEES SUR UN CORRIDOR RIVULAIRE DE L'AIN (12 KM, 1846 OBJETS) D'APRES DURRIEU *ET AL.*, 2007. 161
FIGURE 5.22 : CETTE FIGURE PRESENTE L'EVOLUTION DES MOYENNE D'OBJETS BIEN CLASSES (PRECISION GLOBALE) EN FONCTION DE L'AUGMENTATION DE LA TAILLE DE L'ECHANTILLON EN NOMBRE D'OBJETS PAR CLASSE POUR L'ANALYSE DISCRIMINANTE LINEAIRE (SANS METHODE D'OPTIMISATION DE VARIABLES) APPLIQUE SUR UN CORRIDOR RIVULAIRE DE L'HERAULT (15667 OBJETS). 162
FIGURE 5.23 : CETTE FIGURE PRESENTE L'EVOLUTION DES MOYENNE D'OBJETS BIEN CLASSES (PRECISION GLOBALE) EN UTILISANT L'ANALYSE DISCRIMINANTE LINEAIRE (SANS METHODE D'OPTIMISATION DE VARIABLES) AVEC 50 PARCELLES/CLASSE EN FONCTION DE L'AUGMENTATION DE LA TAILLE DE LA ZONE D'ETUDE (NOMBRES D'OBJETS)..... 162
FIGURE 5.24 : LES 3 PHASES DE CLASSIFICATION ORIENTEE OBJET PAR REGION (ETAPE 5) 164
FIGURE 5.25 : ILLUSTRATIONS DE LA SEGMENTATION (A), DE LA CLASSIFICATION (B) ET PRESENTATION DU DEVELOPPEMENT DE L'ARBRE DE CLASSIFICATION (C) A LA PHASE 1..... 165
FIGURE 5.26 : ILLUSTRATIONS DE LA SEGMENTATION (A), DE LA CLASSIFICATION (B) ET PRESENTATION DU DEVELOPPEMENT DE L'ARBRE DE CLASSIFICATION (C) A LA PHASE 2..... 166
FIGURE 5.27 : EXEMPLE DE LA DEFINITION D'UNE FONCTION FLOUE POUR EXTRAIRE LES OBJETS DE LA CLASSE « COURS D'EAU » SUR UNE REGION HOMOGENE DONNEE. CETTE FONCTION UTILISE L'INFORMATION DU PIR DU SATELLITE SPOT5 XS PUISQUE L'EAU POSSEDE UNE FAIBLE REFLECTANCE DANS CETTE BANDE SPECTRALE. 167
FIGURE 5.28 : CREATION DU NIVEAU DE SEGMENTATION « MACRO » (C) : CE NIVEAU SE BASE A LA FOIS SUR LES OBJETS FINS CREES AU NIVEAU « MICRO » (A) ET LE NIVEAU « MICRO FUSIONNE » (B) RESULTANT D'UNE COPIE DU NIVEAU « MICRO » ET D'UNE FUSION DES OBJETS CONTIGUS DE MEME CLASSE. CE SYSTEME PERMET DE GARDER LES PETITS OBJETS ISOLES DANS LA ZONE ETUDIEE AU NIVEAU « MACRO ».... 168
FIGURE 5.29 : ILLUSTRATIONS DE LA SEGMENTATION (A), DE LA CLASSIFICATION (B) ET PRESENTATION DU DEVELOPPEMENT DE L'ARBRE DE CLASSIFICATION POUR LA CLASSE « SOL SANS OU AVEC PEU DE VEGETATION » (C) A LA PHASE 3 168
FIGURE 5.30 : RESEAU HIERARCHIQUE D'OBJETS DE LA DEMARCHE DE CLASSIFICATION. IL SE COMPOSE DE QUATRE NIVEAUX DE SEGMENTATION EMBOITES 169
FIGURE 5.31 : PRESENTATION DES POST-TRAITEMENTS APPLIQUES (ETAPE 6) SUR LES RESULTATS DE CLASSIFICATION OBTENUS DANS CHAQUE REGION..... 170
FIGURE 5.32 : VALIDATION DU RESULTAT DE CLASSIFICATION (ETAPE 7)..... 172
FIGURE 5.33 : SENS DE DEPLACEMENT DE L'ALGORITHME D'ECHANTILLONNAGE DES PARCELLES DE VALIDATION. L'ALGORITHME COMMENCE LA SELECTION PAR LA STRATE ROUGE SELECTIONNEE ALEATOIREMENT. 174
FIGURE 5.34 : CETTE FIGURE PRESENTE L'EVOLUTION DES ECARTS TYPES DE POURCENTAGE D'OBJETS BIEN CLASSE (PRECISION GLOBALE) EN FONCTION DE L'AUGMENTATION DE LA TAILLE DE L'ECHANTILLON DE VALIDATION..... 175
FIGURE 5.35 : SCHEMATISATION DE LA DEMARCHE DE CLASSIFICATION..... 178
FIGURE 7.1 : LES CATEGORIES D'ELEMENTS DU PAYSAGE D'APRES BUREL ET BAUDRY, 1999 237
FIGURE 7.2 : LES DEUX TYPES DE MODELES EXPLOITES PAR LES SYSTEMES D'INFORMATION GEOGRAPHIQUE POUR REPRESENTER LE MONDE REEL (ESRI FRANCE). 238
FIGURE 7.3 : BUFFER D'ANALYSE POUR LE CALCUL D'UN INDICATEUR SPATIALISE AU NIVEAU D'UNE STATION ECOLOGIQUE SELON UNE LARGEUR L FIXE SUR CETTE ILLUSTRATION ET UNE LONGUEUR L DETERMINEE PAR UNE DISTANCE AMONT DAMONT ET/OU UNE DISTANCE AVAL DAVAL A LA STATION. 242
FIGURE 7.4 : SCHEMATISATION DES EMPRISES SPATIALES DES TROIS FAMILLES D'INDICATEURS SPECIFIQUES D'INDICATEURS D'OCCUPATION DU SOL 246
FIGURE 7.5 : SCHEMATISATION DE L'INDICATEUR DE CONTACT AGRICOLE. CET INDICATEUR TRADUIT LA PRESENCE ET L'INTENSITE AU NIVEAU DES RUPTURES DE LA BANDE ARBOREE, DONC LES PRESSIONS AGRICOLES DIRECTEMENT EN CONTACT AVEC LE COURS D'EAU. EN JAUNE, LES TERRITOIRES AGRICOLES, EN VERT LA VEGETATION ARBOREE RIVULAIRE, EN BLEU L'EAU. 247
FIGURE 7.6 : ILLUSTRATION DE LA DONNEE HYDROGRAPHIQUE OBTENUE PAR LA COMBINAISON DES OBJETS CLASSES « COURS D'EAU » DE LA CARTOGRAPHIE DE L'OCCUPATION DU SOL A THRS ET D'UN BUFFER FIN (1 M DE LARGEUR) AUTOUR DE LA BDCARTHAGE® NIVERAU 2. 251
FIGURE 7.7: ILLUSTRATION DU DECALAGE ENTRE LA BDCARTHAGE° ET LE RESEAU HYDROGRAPHIQUE « REEL » PHOTO-INTERPRETES SUR UNE ORTHOPHOTO. 252
FIGURE 7.8 : DIFFERENCE (EN VALEUR ABSOLUE) ENTRE L'INDICATEUR DE LONGUEUR DE RIPISYLVE CONSTRUIT A PARTIR DU « RESEAU NON CORRIGE » ET LE MEME INDICATEUR CONSTRUIT A PARTIR DU « RESEAU CORRIGE » EN FONCTION DU DECALAGE MOYEN ENTRE CES DEUX RESEAUX POUR CHAQUE COURS D'EAU (16) DU SOUS-BASSIN VERSANT DE L'HERAULT. 252
FIGURE 7.9 : DIFFERENCE (EN VALEUR ABSOLUE) ENTRE L'INDICATEUR DE LONGUEUR DE RIPISYLVE CONSTRUIT A PARTIR DU « RESEAU NON CORRIGE » ET LE MEME INDICATEUR CONSTRUIT A PARTIR DU « RESEAU CORRIGE » EN FONCTION DU DECALAGE MAXIMUM ENTRE CES DEUX RESEAUX POUR CHAQUE COURS D'EAU (16) DU SOUS-BASSIN VERSANT DE L'HERAULT. 252

CemOA : archive ouverte d'Irstea / Cemagref

FIGURE 7.10 : EXEMPLE DE LA DELIMITATION D'UN BUFFER REGULIER « FIXE » EN A ET D'UN BUFFER IRREGULIER S'INTERESSANT AUX OBJETS AU CONTACT DU COURS D'EAU EN B, D'APRES TORMOS ET AL., 2010. 254
FIGURE 8.1 : BUFFER D'ANALYSE POUR LE CALCUL D'UN INDICATEUR SPATIALISE AU NIVEAU D'UN TRONÇON DE COURS D'EAU (POUR LA MASSE D'EAU A DANS CET EXEMPLE) SELON UNE LARGEUR L FIXE SUR CETTE ILLUSTRATION. 263
FIGURE 8.2 : PRESENTATION DE LA PARTIE AVAL DU FLEUVE HERAULT (80 KM) 277
FIGURE 8.3 : CARTOGRAPHIE DE LA LARGEUR MOYENNE ET DE L'UNIFORMITE DE LA RIPISYLVE SUR CHAQUE TRONÇON (A->J) DE LA PARTIE AVAL DE L'HERAULT. 278
FIGURE 8.4 : CARTOGRAPHIE DE TROIS INDICATEURS DE CONTACT : SUPERFICIE DE LA RIPISYLVE, PRESSIONS AGRICOLE (AGRI.) ET URBAINE (URB.) DE CONTACT SUR CHAQUE TRONÇON DE LA PARTIE AVAL DE L'HERAULT. 279
FIGURE 8.5 : PRESENTATION DES TREIZE MASSES D'EAU DU BASSIN VERSANT DE L'HERAULT 281
FIGURE 8.6 : CARTOGRAPHIE DE LA LONGUEUR (ICR-REGFIXE-LIN-%-FORET) ET DE LA DENSITE DE RUPTURES (ICR-IREG-CON-DRUP-FORET) DE MA RIPISYLVE DE CHAQUE MASSE D'EAU DU BASSIN VERSANT DE L'HERAULT 282
FIGURE 8.7 : CARTOGRAPHIE DE LA PRESSION AGRICOLE SUR LE LIT MAJEUR (ICR-REGVAR-MAJ -%-AGRI) DE CHAQUE MASSE D'EAU DU BASSIN VERSANT DE L'HERAULT 283
FIGURE 8.8 : CARTOGRAPHIE DE LA PRESSION URBAINE SUR LE LIT MAJEUR (ICR-REGVAR-MAJ -%-URB) DE CHAQUE MASSE D'EAU DU BASSIN VERSANT DE L'HERAULT 284
FIGURE 9.1 : NOMBRE DE STATIONS PAR NOMBRE DE CAMPAGNES DE PRELEVEMENT SUR LA PERIODE DE 1992 A 2004. 301
FIGURE 9.2 : CARTOGRAPHIE DE L'OCCUPATION DU SOL DANS L'HER9-NO AU NIVEAU DU CORRIDOR RIVULAIRE SELON LA DONNEE A CLC AGREGEE SUIVANT LA TYPOLOGIE DE COMPARAISON EN 8 CATEGORIES 306
FIGURE 9.3 : CARTOGRAPHIE DE L'OCCUPATION DU SOL DANS L'HER9-NO AU NIVEAU DU CORRIDOR RIVULAIRE SELON LA DONNEE A THRS AGREGEE SUIVANT LA TYPOLOGIE DE COMPARAISON EN 8 CATEGORIES 307
FIGURE 9.4 : EMPRISE SPATIALE A L'ECHELLE DU BASSIN VERSANT. EXEMPLE DE CALCUL A PARTIR D'UNE STATION (EN ROUGE) ET DES PRESSIONS ANTHROPIQUES DERIVEES DE L'ANALYSE DE CORINE LAND COVER (D'APRES PELLA *ET AL.*, 2004) 308
FIGURE 9.5 : EMPRISES SPATIALES MISES EN PLACE DANS LE CADRE DU MODE DE SPATIALISATION « MULTIPLE » A L'ECHELLE DU CORRIDOR RIVULAIRE AMONT. POUR CHAQUE STATION ECOLOGIQUE (EN ROUGE) CES BUFFERS SONT CONSTRUITS SUR LA TOTALITE DU RESEAU AMONT AINSI QUE SUR UNE LONGUEUR DE RESEAU AVAL DE 500 M. 310
FIGURE 9.6 : EMPRISE SPATIALE DEFINIE PAR LE MODE DE SPATIALISATION « FIXE » A L'ECHELLE DU CRLOCAL : EXEMPLE DE DELIMITATION POUR UNE STATION ECOLOGIQUE (EN ROUGE) DE L'EMPRISE SUR LAQUELLE SERONT CALCULES LES INDICATEURS DE PRESSIONS ANTHROPIQUES DERIVEES DE L'ANALYSE DE CORINE LAND COVER (PELLA *ET AL.*, 2004) 311
FIGURE 9.7 : EMPRISES SPATIALES MISES EN PLACE DANS LE CADRE DU MODE DE SPATIALISATION « MULTIPLE » A L'ECHELLE DU CRLOCAL. CES BUFFERS, DE LONGUEURS ET DE LARGEURS DIFFERENTES, SONT CONSTRUITS AUTOUR DU RESEAU HYDROGRAPHIQUE EN AMONT ET EN AVAL DE LA STATION ECOLOGIQUE (EN ROUGE). EN A : LES DIFFERENTES LONGUEURS ETUDIEES ; EN B : LES DIFFERENTES LARGEURS SUR UN EXEMPLE DE LONGUEUR (ICI 2500 M) 312
FIGURE 10.1 : CARTOGRAPHIE DE L'ETAT ECOLOGIQUE DES STATIONS DE L'HER9-NO SELON L'EQR-IBGN MOYENNE SUR LA PERIODE DE 1992 A 2004 EN FONCTION DES 5 MODALITES D'ETAT. 333
FIGURE 10.2 : ÉVOLUTION DE LA QUALITE DES STATIONS, EVALUEE A PARTIR DE L'$EQR\text{-}IBGN_{MOYEN}$ (SUR LA PERIODE DE 1992 A 2004) SELON LES 5 MODALITES D'ETAT (TRES BON, BON, MOYEN, MEDIOCRE, MAUVAIS), EN FONCTION DU RANG DE STRAHLER. 334
FIGURE 10.3 : COMPOSITION MOYENNE (SUR LES 155 STATIONS) DE L'OCCUPATION DU SOL (EN %) SELON LA DONNEE CLC EN FONCTION DE LA TYPOLOGIE EN 8 CATEGORIES AUX DIFFERENTES ECHELLES (BV EN BLEU, CRAMONT EN ORANGE, ET CRLOCAL EN VERT) ET SUR LES DIFFERENTES EMPRISES DU CORRIDOR RIVULAIRE SELON UN GRADIENT DU PLUS FONCE (EMPRISE LA PLUS ELOIGNE) AU PLUS CLAIR (EMPRISE LA PLUS PROCHE) SELON LA COULEUR DE L'ECHELLE. 337
FIGURE 10.4 : COMPOSITION MOYENNE (SUR LES 155 STATIONS) DE L'OCCUPATION DU SOL (EN %) SELON LA DONNEE A THRS EN FONCTION DE LA TYPOLOGIE EN 8 CATEGORIES AUX DIFFERENTES ECHELLES (BV EN BLEU, CRAMONT EN ORANGE, ET CRLOCAL EN VERT) ET SUR LES DIFFERENTES EMPRISES DU CORRIDOR RIVULAIRE SELON UN GRADIENT DU PLUS FONCE (EMPRISE LA PLUS ELOIGNE) AU PLUS CLAIR (EMPRISE LA PLUS PROCHE) SELON LA COULEUR DE L'ECHELLE. 338
FIGURE 10.5 : COMPOSITION MOYENNE DE L'OCCUPATION DU SOL (EN %) A L'ECHELLE DU BV EN FONCTION DU RANG DE STRAHLER, SELON LA DONNEE CLC EN FONCTION DE LA TYPOLOGIE EN 8 CATEGORIES. 341
FIGURE 10.6 : COMPOSITION MOYENNE DE L'OCCUPATION DU SOL (EN %) A L'ECHELLE DU CRAMONT EN FONCTION DU RANG DE STRAHLER, SELON LA DONNEE THRS EN FONCTION DE LA TYPOLOGIE EN 8 CATEGORIES. 341
FIGURE 10.7 : COMPOSITION MOYENNE DE L'OCCUPATION DU SOL (EN %) A L'ECHELLE DU CRLOCAL EN FONCTION DU RANG DE STRAHLER, SELON LA DONNEE THRS EN FONCTION DE LA TYPOLOGIE EN 8 CATEGORIES. 342
FIGURE 10.8 : CARTOGRAPHIE DE LA COMPOSITION DE L'OCCUPATION DU SOL EN FONCTION DES 8 SOURCES DE PRESSIONS POUR CHAQUE STATION A L'ECHELLE DU BV 344
FIGURE 10.9 : CARTOGRAPHIE DE LA COMPOSITION DE L'OCCUPATION DU SOL A L'ECHELLE DU CRAMONT, POUR CHAQUE STATION, SELON LA DONNEE THRS EN FONCTION DES 8 SOURCES DE PRESSIONS 345
FIGURE 10.10 : CARTOGRAPHIE DE LA COMPOSITION DE L'OCCUPATION DU SOL A L'ECHELLE DU CRLOCAL, POUR CHAQUE STATION, SELON LA DONNEE THRS EN FONCTION DES 8 SOURCES DE PRESSIONS 346

CemOA : archive ouverte d'Irstea / Cemagref

FIGURE 10.11 : COEFFICIENT DE CORRELATION (R^2) ISSU DE LA REGRESSION LINEAIRE SIMPLE AVEC L'EQR-IBGN$_{MOYEN}$ ET CHAQUE INDICATEURS DE PRAIRIES PERMANENTES AUX DIFFERENTES EMPRISES A L'ECHELLE DU CRLOCAL (BUFFERS VARIANT EN LARGEUR ET E, LONGUEUR) ET DU CRAMONT (BUFFERS VARIANT EN LARGEUR). 349
FIGURE 10.12 : ÉVOLUTION DE LA CAPACITE EXPLICATIVE(R^2) EN FONCTION DES DIFFERENTES HYPOTHESES DE CONSTRUCTION DES MODELES. LA LIGNE EN ROUGE CORRESPOND AU SEUIL DE SELECTION DES MODELES EN FONCTION DE LA CAPACITE EXPLICATIVE. 358
FIGURE 10.13 : CAPACITE DESCRIPTIVE (NOMBRE DE PREDICTEURS SIGNIFICATIFS) DES 7 MODELES SELECTIONNES APRES L'APPLICATION DU PREMIER CRITERE (CAPACITE EXPLICATIVE). 358
FIGURE 10.14 : ÉVOLUTION DE LA PERFORMANCE DES MODELES PRESSIONS/IMPACTS AU NIVEAU DU CORRIDOR RIVULAIRE : DEPUIS LE MODELE DEVELOPPE EN FRANCE EN 2005 PAR WASSON ET AL. (EN GRIS SUR LE GRAPHIQUE) JUSQU'AUX MODELES DEVELOPPEES DANS CETTE ETUDE SUIVANT LES DIFFERENTES VOIES D'AMELIORATIONS. OSCR : (OCCUPATION DU SOL DANS LE CORRIDOR RIVULAIRE)..... 360
FIGURE 10.15 : ÉVOLUTION DE LA CAPACITE EXPLICATIVE(R^2) EN FONCTION DES DIFFERENTES HYPOTHESES DE CONSTRUCTION DES MODELES CONSTRUITS SUR 155 STATIONS (LIGNE FONCE) ET SUR 100 STATIONS (LIGNE CLAIRE)..... 364
FIGURE 10.16 : RESULTATS DES L'ACP CONSTRUITES A PARTIR DE L'OBSERVATION DES POURCENTAGES AU NIVEAU DES TROIS ECHELLES D'ANALYSES (BV, CRAMONT ET CRLOCAL) SUR LES 155 STATIONS IBGN DISPONIBLES SUR LA ZONE D'ETUDE POUR LE TISSU URBAIN DISCONTINU (POSTE 112) ET LES PRAIRIES (POSTE 231) QUANTIFIES D'UNE PART AVEC LA DONNEE CLC ET D'AUTRE PART AVEC LA DONNEE A THRS. 368
FIGURE 10.17 : CONTRIBUTION DES SOURCES DE PRESSIONS A L'ECHELLE DU BV, CRAMONT ET DU CRLOCAL QUI AGISSENT DE FAÇON COMBINEE SUR L'ETAT ECOLOGIQUE DES COURS D'EAU (EQR-IBGN$_{MOYEN}$) MESURES SUR LES 155 STATIONS DE LA ZONE D'ETUDE. CES SOURCES EXPLIQUENT 33.8 % DE LA VARIABILITE DE CET ETAT. LE MODELE A ETE CONÇU EN UTILISANT LE MODE DE SPATIALISATION « MULTIPLE » ET LA DONNEE THRS AU NIVEAU DU CORRIDOR RIVULAIRE..... 370
FIGURE 10.18 : EXEMPLE DE RECTIFICATION DU TRACE ET SUPPRESSION DE LA RIPISYLVE LIEE A L'EMPRISE AGRICOLE SUR LA ZONE D'ETUDE (EN BLEU LE RESEAU HYDROGRAPHIQUE). SOURCE : BDORTHO® 2006 ET BDCARTHAGE® 2005 IGN 376
FIGURE 10.19 : ILLUSTRATION DE LA BANDE DE 20 M ET LA BANDE DE 250 M (SUR 3 KM) EN PRAIRIES QU'ON RETROUVE EN AMONT D'UNE STATION EN TRES BON ETAT SOUMISE NOTAMMENT A UNE PRESSION DES TERRES ARABLES A L'ECHELLE DU BV (EN ROUGE LA STATION, EN BLEU LE RESEAU HYDROGRAPHIQUE). SOURCE : BDORTHO® 2006 ET BDCARTHAGE® 2005 IGN. 378
FIGURE 10.20 : ILLUSTRATION DE LA PRESENCE DE LA VEGETATION RIVULAIRE SUR UNE STATION DE LA ZONE D'ETUDE EN BON ETAT (EN ROUGE LA STATION, EN BLEU LE RESEAU HYDROGRAPHIQUE). SOURCE : BDORTHO® 2006 ET BDCARTHAGE® 2005 IGN. 379
FIGURE 10.21 : ILLUSTRATION DE L'ENVIRONNEMENT QU'ON OBSERVE SUR LES STATIONS POSSEDANT UNE ETENDUE D'EAU DANS LEUR ENVIRONNEMENT PROCHE (EN ROUGE LA STATION, EN BLEU LE RESEAU HYDROGRAPHIQUE). SOURCE : BDORTHO® 2006 ET BDCARTHAGE® 2005 IGN. 380

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

LISTE DES TABLEAUX

TABLEAU 1-1 : CLASSES DE TAILLE POUR LES COURS D'EAU EN FONCTION DES RANGS DE STRAHLER (CHANDESRIS ET AL., 2006)..... 46
TABLEAU 1-2 : RESULTATS DE L'ETAT ECOLOGIQUE IBGN AU NIVEAU DES ZONES HYDROGRAPHIQUES (PREDIT PAR LES MODELES D'EXTRAPOLATION SPATIALE) ET AU NIVEAU DES RESEAUX DE SUIVI SELON LES DEUX HYPOTHESES DE BON ETAT (D'APRES WASSON *ET AL.*, 2006)..... 53
TABLEAU 3-1 : SYNTHESE BIBLIOGRAPHIQUE DES RECHERCHES SUR LE POUVOIR EPURATEUR DES ZONES ALLUVIALES VIS-VIS DES NUTRIMENTS, AZOTE ET PHOSPHORE D'APRES RUFFINONI *ET AL.*, 2003. 86
TABLEAU 5-1 : ÉTUDES CARACTERISANT LES CORRIDORS RIVULAIRES A PARTIR D'IMAGES OPTIQUES MULTI-SPECTRALES MONO-DATE A THRS (TRES HAUTE RESOLUTION SPATIALE) ET A L'AIDE DE TRAITEMENTS AUTOMATIQUES. (B=BLEU ; V=VERT ; R=ROUGE ; PIR=PROCHE INFRAROUGE ; MIR=MOYEN INFRAROUGE ; IRTL=INFRAROUGE THERMIQUE LOINTAIN ; PAN=PANCHROMATIQUE)..... 134
TABLEAU 5-2 : ESTIMATEURS GLOBAUX DE LA QUALITE DES CLASSIFICATIONS DE L'OCCUPATION DU SOL DANS LES CORRIDORS RIVULAIRES SELON LES TROIS COMBINAISONS (C1, C2, C3) DE SOURCES DE DONNEES DIFFERENTES (DA = DONNEES AUXILIAIRES). 143
TABLEAU 5-3 : PRECISION POUR L'UTILISATEUR PAR CATEGORIE D'OCCUPATION DU SOL SELON LES TROIS COMBINAISONS (C1, C2, C3) DE SOURCES DE DONNEES DIFFERENTES (DA = DONNEES AUXILIAIRES)..... 143
TABLEAU 5-4 : POURCENTAGE DE VEGETATION ARBOREE RIVULAIRE CORRECTEMENT DETECTE POUR TROIS TAILLES DE LARGEUR DE VEGETATION SELON ON LES TROIS COMBINAISONS (C1, C2, C3) DE SOURCES DE DONNEES DIFFERENTES (DA = DONNEES AUXILIAIRES). 144
TABLEAU 5-5 : LARGEUR DE CORRIDOR RIVULAIRE PRIS EN COMPTE POUR EXTRAIRE L'INFORMATION D'OCCUPATION DU SOL DANS CES ESPACES..... 145
TABLEAU 5-6 : NOTATION SUR UNE MATRICE DE CONFUSION..... 172
TABLEAU 6-1 : PRESENTATION DES DONNEES IMAGES UTILISEES SUR LE BASSIN VERSANT DE L'HERAULT..... 185
TABLEAU 6-2 : PRESENTATION DES DONNEES AUXILIAIRES UTILISEES SUR LE BASSIN VERSANT DE L'HERAULT (MAJ = MISE A JOUR) 185
TABLEAU 6-3 : STATISTIQUES DE LA FIABILITE DE LA STABILITE DES SIX CATEGORIES D'OCCUPATION DU SOL EXTRAITES DANS LE CORRIDOR RIVULAIRE DE L'HERAULT..... 224
TABLEAU 7-1 : LARGEUR DE LIT MAJEUR PRIS EN COMPTE EN FONCTION DU RANG DE STRAHLER POUR CALCULER LES INDICATEURS DU LIT MAJEUR..... 246
TABLEAU 7-2 : RECAPITULATIF DES FAMILES ET TYPES SPECIFIQUES D'INDICATEURS SPATIALISES DU CORRIDOR RIVULAIRE 256
TABLEAU 9-1 : RECAPITULATIF DES INDICATEURS D'OCCUPATIONS DU SOL GENERES POUR CHAQUE STATION ECOLOGIQUE DE LA ZONE D'ETUDE 313
TABLEAU 9-2 : PRESENTATION DES HYPOTHESES DE CONSTRUCTION DES 21 MODELES PRESSIONS/IMPACTS. CES HYPOTHESES DIFFERENT EN FONCTION DE L'ECHELLE(S) D'ANALYSE ; DE LA SOURCE DE DONNEE D'OCCUPATION DU SOL (CLC OU THRS) ET DU MODE DE SPATIALISATION AU NIVEAU DU AU NIVEAU DU CORRIDOR RIVULAIRE. CE TABLEAU INDIQUE AU FINAL LE NOMBRE DE PREDICTEURS (D'INDICATEURS D'OCCUPATION DU SOL) QUI SERONT UTILISES PAR CHACUN DES MODELES. 323
TABLEAU 9-3 : PRESENTATION DES 8 COUPLES DE MODELES DONT LES RESULTATS SERONT COMPARES POUR ETUDIER L'APPORT DU MODE DE SPATIALISATION « MULTIPLE » PAR RAPPORT AU MODE « FIXE ». EN GRIS LES COUPLES DE MODELES UTILISANT LA DONNEE A THRS. 326
TABLEAU 9-4 : PRESENTATION DES 10 COUPLES DE MODELES DONT LES RESULTATS SERONT COMPARES POUR ETUDIER L'APPORT DE LA DONNEE A THRS PAR RAPPORT A CLC. EN GRIS LES COUPLES DE MODELES UTILISANT LE MODE DE SPATIALISATION « MULTIPLE » A L'ECHELLE DU CRLOCAL. 326
TABLEAU 10-1 : REPARTITION DES STATIONS DE L'HER9-NO EN FONCTION DE LA DIMENSION DU COURS D'EAU (RANG DE STRAHLER) 332
TABLEAU 10-2 : REGLES D'AGREGATION (POUR CHACUNE DES SOURCES DE DONNEES) POUR OBTENIR LA TYPOLOGIE COMMUNE COMPOSEE DE 8 CATEGORIES. 335
TABLEAU 10-3 : CARACTERISTIQUES DES EMPRISES SELECTIONNEES AU NIVEAU DU CORRIDOR RIVULAIRE POUR L'ANALYSE DE L'ETAT DES PRESSIONS. 336
TABLEAU 10-4 : PRESENTATION DES EMPRISES SELECTIONNEES POUR CHAQUE SOURCE DE PRESSION ETUDIEE PAR LE MOD-3.4.1 A L'ECHELLE DU CRAMONT ET DU CRLOCAL. NP SIGNIFIE QU'AUCUNE EMPRISE N'A ETE SELECTIONNEE CAR LA SOURCE DE PRESSION ETUDIEE N'EST PAS PRESENTE SUR L'ENSEMBLE DES STATIONS A L'ECHELLE CONSIDEREE..... 350
TABLEAU 10-5 RESULTATS OBTENUS A CHAQUE ITERATION LORS DE LA CONSTRUCTION DU MODELE MOD-3.1.4 351
TABLEAU 10-6 : R^2 ET NOMBRE DE PREDICTEURS SUR LA TOTALITE ET PAR ECHELLE SELON L'EFFET (EN ROSE PALE, LES EFFETS DEGRADANTES ET EN VERT PALE LES EFFETS PROTECTEURS) POUR CHACUN DES 21 MODELES 352

CemOA : archive ouverte d'Irstea / Cemagref

TABLEAU 10-7 : GAINS (OU PERTES) DE POUVOIR EXPLICATIF (R^2), DE PREDICTEURS SUR L'ENSEMBLE DES ECHELLES ANALYSES, DE TYPES DE SOURCES DE PRESSION (I.E. CATEGORIE D'OCCUPATION DU SOL) OBTENUS PAR L'APPLICATION DE L'EMPRISE SPATIALE CRLOCAL-OPT PAR RAPPORT A CRSTR POUR CHAQUE COUPLE DE MODELES DE MEME NOMBRE ET DE COMBINAISON D'ECHELLES ET DE SOURCE DE DONNEE D'OCCUPATION DU SOL UTILISANT AU MOINS L'ECHELLE DU CRLOCAL. 354

TABLEAU 10-8 : GAINS (OU PERTES) DE POUVOIR EXPLICATIF (R^2) ET DE PREDICTEURS, SUR L'ENSEMBLE DES ECHELLES ET PAR ECHELLES, OBTENUS PAR L'APPLICATION AU NIVEAU DU CORRIDOR RIVULAIRE DE LA THRS PAR RAPPORT A CLC POUR CHAQUE COUPLE DE MODELES DE MEME(S) ECHELLE(S) D'ANALYSE UTILISANT LE MEME MODE DE SPATIALISATION. EN GRIS LES COUPLES DE MODELES UTILISANT LE MODE DE SPATIALISATION « MULTIPLE » A L'ECHELLE DU CRLOCAL. 356

TABLEAU 10-9 : PRESENTATION PAR ORDRE D'INFLUENCE DES COEFFICIENTS DE CORRELATIONS PLS ET LEUR STATISTIQUES (ECART-TYPE ET INTERVALLE DE CONFIANCE A 95 %) DES VARIABLES SIGNIFICATIVES OBTENUS POUR LE MODELE BV/CRAMONT/CRLOCAL UTILISANT L'EMPRISE SPATIALE CRLOCAL-OPT ET LA DONNEE THRS AU NIVEAU DU CORRIDOR RIVULAIRE. EN ROSE PALE LES VARIABLES POSSEDANT UN COEFFICIENT DE REGRESSION NEGATIF ET EN VERT PALE CELLES QUI POSSEDENT UN COEFFICIENT DE REGRESSION POSITIF. POUR FACILITER LA LECTURE DE L'ANALYSE DES PRESSIONS INFLUENÇANT L'ETAT ECOLOGIQUE (PARTIE IV.2), LA DERNIERE COLONNE FAIT REFERENCE AUX NUMEROS DE PARAGRAPHES OU SONT DISCUTES CHACUNE DE CES VARIABLES. 371

CemOA : archive ouverte d'Irstea / Cemagref

BIBLIOGRAPHIE

Adam, P., Malavoi, J.-R. & Debiais, N. (2007) Manuel de restauration hydromorphologique des rivières. Agence de l'Eau Seine-Normandie.

Agence De L'eau Seine Normandie (2004) Etat des lieux : Bassin Seine et cours d'eau côtiers normands. p. 200.http://www.eau-seine-normandie.fr/index.php?id=2258.

Agences De L'eau (1999) Les outils d'évaluation de la qualité des cours d'eau (S.E.Q). Principes généraux. *Etudes des Agences de l'eau,* **72,** 13.

Allan, J.D., Erickson, D. & Fay, J. (1997) The influence of catchment land use on stream integrity across multiple spatial scales. *Freshwater Biology,* **37,** 149-161.

Allan, J.D. (2004a) Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. *Annual Review of Ecology, Evolution, and Systematics,* **35,** 257-284.

Allan, J.D. (2004b) Influence of land use and landscape setting on the ecological status of rivers. *Limnetica,* **23,** 187-198.

Allan, J.D. & Castillo, M.M. (2007) *Stream Ecology,* Springer, Dordrecht, Netherdlands.

Allen, T.F.H. & Starr, T.B. (1982) *Hierarchy perspectives for ecological complexity,* University of Chicago Press, Chicago, USA, p. 310.

Altier, L.S., Lowrance, R., Williams, R.G., Inamdar, S.P., Bosch, D.D., Sheridan, J.M., Hubbard, R.K. & Thomas, D.L. (2002) Riparian Ecosystem Management Model: Simulator
for ecological processes in riparian zones. USDA-ARS, Conservation Research Report, Washington, D.C.: GPO, p. 46.

Ambroise, B. (1991) Hydrologie des petits bassins versants ruraux en milieu tempéré - Processus et modèles. In: *Séminaire du Conseil Scientifique du Département "Science du sol" de l'INRA* (Dijon, 26-27/03/1991), p. 34.

Amoros, C. & Bravard, J.P. (1985) L'intégration du temps dans les recherches méthodologiques appliquées à la gestion écologique des vallées fluviales. *Rev. Française Sciences de l'eau,* **4,** 349-364.

Amoros, C., Roux, A.L., Reygrobellet, J.L., Bravard, J.P. & Pautou, G. (1987) A method for applied ecological studies of fluvial hydrosystems. *Regulated Rivers : Research and Management,* **11,** 17-36.

Amoros, C. & Petts, G.E. (1993) *Hydrosystèmes fluviaux,* Collection d'écologie 24, Masson, Paris.

Angermeier, P.L. & Karr, J.R. (1984) Relationships between woody débris and fish habitat in a small warmwater stream. *Trans. Am. Fish. Soc,* **113,** 716-726.

Aplin, P. & Smith, G.M. (2008) Advances in Object-Based Image Classification. In: *ISPRS Congress, Proceedings of Commission VII* (Beijing.

Archaimbault, V. & Dumont, B. (2010) L'indice biologique global normalisé (IBGN) : principes et évolution dans le cadre de la directive cadre européenne sur l'eau. *Revue SET,* **1**.

Arnold, C.L. & Gibbons, C.J. (1996) Impervious surface coverage - The emergence of a key environmental indicator. *Journal of the American Planning Association,* **62,** 243-258.

Arnold, J.G., Williams, J.R., Srinivasan, R., King, K.W. & Griggs, R.H. (1994) SWAT(Soil and Water Assessment Tool) User's Manual. USDA, Agricultural Research Service, Grassland, Soil and Water Research Laboratory, Temple, USA.

Ashraf, S., Brabyn, L., Hicks, B.J. & Collier, K. (2010) Satellite remote sensing for mapping vegetation in New Zealand freshwater environments: A review. *New Zealand Geographer,* **66,** 33-43.

Asp *Mise à disposition du registre parcellaire graphique anonyme* http://www.asp-public.fr/?q=node/856, 5 Novembre.

Baatz, M. & Schäpe, A. (2000) Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung,* **12,** 12-23.

Balestrat, M., Chéry, J.-P. & Valette, E. (2008) Suivi des changements d'occupation et d'utilisation du sol pour la compréhension des dynamiques périurbaines. In: *XLV colloque de l'Association Régionale de Langue Française* (Rimouski (Québec).

Barbour, M.T., Gerritsen, J., Snyder, B.D. & Stribling, J.B. (1999) Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish. *Second edition, EPA 841-B-99-002,.* U.S. Environmental Protection Agency, Office of water, Washington, D.C, p. 339.

Barge, O. & Joliveau, T. (1996) Démarche territoriale et Systèmes d'Information Géographique pour une gestion concertée de l'eau. *Revue Géographique de Lyon.* pp. 297-309.f.

Barling, R. & Moore, I.D. (1994) Role of buffer of strips io management of waterway pollution : a review. *Enrvironmental Management,* **18,** 543-558.

CemOA : archive ouverte d'Irstea / Cemagref

Behmer, D.J. & Hawkins, C.P. (1986) Effects of overhead canopy on macroinvertebrate production in a Utah stream. *Freshwat. Biol.,* **16,** 287-300.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I. & Heynen, M. (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing,* **58,** 239-258.

Berryman, D. (1990) Sélection des nouveaux indicateurs de la qualité des cours d'eau du Québec. Ministère de l'Environnement du Québec, Direction de la qualité de l'eau des cours d'eau, p. 77.

Beven, K.J. (1989) Changing ideas in hydrology: the case of physically based models. *J. Hydrol.,* **105,** 157-172.

Billen, G., Garnier, J. & Hanset, P. (1994) Modelling phytoplankton development in whole drainage networks: The RIVERSTRAHLER model applied to the Seine river système. *Hydrobiologia,* **289,** 119-137.

Billen, G. & Garnier, J. (1999) Nitrogen transfers through the Seine drainage network: a budget based on the application of the 'Riverstrahler model'. *Hydrobiologia,* **410,** 139-150.

Birkinshaw, S.J. & Ewen, J. (2000) Nitrogen transformation component for SHETRAN catchment nitrate modelling. *Journal of Hydrology,* **230,** 1-17.

Bis, B., Zdanowicz, A. & Zalewski, M. (2000) Effects of catchment properties on hydrochemistry, habitat complexity and invertebrate community structure in a lowland river. *Hydrobiologia,* **422-423,** 369-387.

Bisson, P.A., Nielsen, J.L., Palmason, R.A. & L.E., G. (1981) A System of naming habitat types in small streams, with examples of habitat utilization by salmonids during low streamflow. *In ARMANTROUT N.B. (éd.), Acquisition and utilization of aquatic habitat inventory information, Portiand, Oregon, American fisheries society***,** 62-73.

Bisson, R., Bilby, R.E., Bryant, M.D., Dolloff, C.A., Grette, G.B., House, R.A., Murphy, M.L., Koski, K.V. & Sedell, J.R. (1987) Large woody débris in forested streams in the Pacific Northwest : past, présent and future. In CUNDY T.W. (éd.). *In CUNDY T.W. (éd.), Streamside Management : Forestry and Fishery Interactions, Seattle, Collège of Forest Resources, University of Washington***,** 143-190.

Blair, J.B., Hofton, M.A. & Rabine, D.L. (2006) Processing of NASA LVIS Elevation and Canopy (LGE, LCE and LGW) Data Products. Version 1.2. Available at http://lvis.gsfc.nasa.gov. Accessed in January 2006.

Blaschke, T. & Hay, G.J. (2001) Object-oriented Image Analysis and Scale-space: Theory and Methods for Modeling and Evaluating Multiscale Landscape Structures. *International Archives of Photogrammetry and Remote Sensing,* **34,** 22-29.

Blondel, J. (2003) L'avifaune des ripisylves méditerrranéennes. *Forêt Méditerranéenne.,* **24,** 249-256.

Bonada, N., Prat, N., Resh, V.H. & Statzner, B. (2006) Developments in aquatic insect biomonitoring: A comparative analysis of recent approaches. *Annual Review of Entomology,* **51,** 495-523.

Bonn, F. & Rochon, G. (1992) *Précis de télédétection, volume 1 : principes et méthodes,* Presse de l'Université du Québec/AUPELF**,** p. 485

Booth, D.B. (1991) Urbanization and the natural drainage system: impacts, solutions, and prognoses. *The Northwest Environmental Journal,* **7,** 93-118.

Borchardt, D. & Sperling, F. (1997) Urban stormwater discharges: Ecological effects on receiving waters and consequences for technical measures. *Water Science and Technology,* **36,** 173-178.

Borchardt, D. & Richter, S. (2003) Identification of significant pressures and impacts upon receiving waters. *Water Science and Technology,* **48,** 33-38.

Bordin, P. (2002) *SIG concepts, outils et données,* Hermès - Lavoisier**,** p. 260.

Borror, D.J. & White, R.E. (1999.) *Le guide des insectes du Québec et de l'Amérique du Nord,* Les guides Peterson, Québec, Canada.

Bossard, M., Feranec, J. & Otahel, J. (2000) CORINE land cover technical guide-Addeendum 2000. *Technical report No 40.* European Environment Agency.

Bouché, M.B. (1972) *Lombriciens de France. Ecologie et systématique*
INRA**,** p. 671

Boulton, A.J. (1999) An overview of river health assessment: philosophies, practice, problems and prognosis. *Freshwater Biology,* **41,** 469-479.

Boussu, M.F. (1954) Relationship between trout populations and cover on a small stream. *J. Wild. Manag,* **18** 229-239.

Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M.H., Ruget, F., Nicoullaud, B., Gate, P., Devienne, F., Antonioletti, R., Dürr, C., Richard, G., Beaudoin, N., Recous, S., Tayot, X., Plénet, D., Cellier, P., Machet, J.M., Meynard, J.M. & Delécolle, R. (1998) STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parametrization applied to wheat and corn. *Agronomie,* **18,** 311-346.

Brooks, A. (1998) Channelized rivers. Perspectives for environnemental management.

CemOA : archive ouverte d'Irstea / Cemagref

Brown, J.H. (1995) *Macroecology,* University of Chicago Press, Chicago**,** p. 284.

Brusven, M.A., Meehan, W.R. & Ward, J. (1986) Summer use of simulated undercut banks by juvénile chinook salmon in an artificial Idaho channel

North Am. J. of Fish Man, **6,** 32-37.

Bryant, M.D. (1981) Organic débris in salmonid habitat in Southeast Alaska : measurement and effects. In ARMANTROUT N.B. (éd.). Acquisition and utilization of aquatic habitat inventory information, Portiand, Oregon. *American fisheries society***,** 259-265.

Bryant, R.G. & Gilvear, D.J. (1999) Quantifying geomorphic and riparian land cover changes either side of a large flood event using airborne remote sensing: River Tay, Scotland. *Geomorphology,* **- 29,** - 321.

Buffagni, A., Erba, S. & Furse, M.T. (2007) A simple procedure to harmonize class boundaries of assessment systems at the pan-European scale. *Environmental Science & Policy,* **10,** 709-724.

Burel, F. & Baudry, J. (1999) *Ecologie du paysage: Concepts, méthodes et applications***,** p. 348.

Camargo, J.A., Alonso, A. & De La Puente, M. (2004) Multimetric assessment of nutrient enrichment in impounded rivers based on benthic macroinvertebrates. pp. 233-249. Environmental Monitoring and Publishers.

Carleer, A.P., Debeir, O. & Wolff, E. (2005) Assessment of very high spatial resolution satellite image segmentations

Photogrammetric Engineering and Remote Sensing, **71,** 1285-1294.

Caron, J.-M., Gauthier, A., Lardeaux, J.-M., Schaaf, A., Ulysse, J. & Wozniak, J. (2003) *Comprendre et enseigner la planète Terre,* Ophrys, Gap ; Paris.

Carrascal, L.M. & Seoane, J. (2009) Factors affecting large-scale distribution of the Bonelli's eagle Aquila fasciata in Spain. *Ecological Research,* **24,** 565-573.

Catterall, C.P., Piper, S.D., Bunn, S.E. & M. Arthur, F. (2001) Flora and fauna assemblages vary with local topography in a subtropical eucalypt forest. *Austral Ecology,* **26,** 56-69.

Cavalié, J. & Dressayre, E. (1986) Etude érosion. Chambre d'Agriculture de la Haute-Garonne SUAD.

Cbp (2003) (Chesapeake Bay Program), Recommendations on the 2003 Directive on Expanded Riparian Buffer Goals in the Chesapeake Bay Watershed. *USEPA Directive 03-01.* U.S. Environmental Protection Agency, Annapolis, Maryland. Available at www.chesapeakebay.net/info/pressreleases/ec2003/RFB_Report_EC_Meeting.pdf. Accessed in January 2006.

Cemagref *Dossier thématique : Des méthodes et des modèles,* http://www.cemagref.fr/Informations/DossiersThematiques/MethodesModeles/Enjeu.htm#, 10 Octobre 2009.

Chabert, A. (1996) Les lombriciens des sols cultivés. Effet de quelques pratiques culturales sur les peuplements. ACTA Point n° 4, Association de Coordination Technique Agricole, Paris, p. 28

Chandesris, A., Wasson, J.G., Pella, H., Sauquet, E. & Mengin, N. (2006) Appui scientifique à la mise en oeuvre de la Directive cadre européenne sur l'eau. Typologie des cours d'eau de France métropolitaine. CemOA Publications - Publications scientifiques et techniques du Cemagref http://cemadoc.cemagref.fr/cemoa/PUB00021163

Chandesris, A., J.R., M., Y., S., J.G., W. & N., M. (2007) Le système relationnel d'audit de l'hydromorphologie des cours d'eau (SYRAH CE): un outil multi-échelles d'aide à la décision pour la gestion des cours d'eau. . *Ingénieries - Eau Agriculture & Territoires,,* **50,** 77-80.

Chandesris, A., Mengin, N., Malavoi, J.R., Souchon, Y., Pella, H. & Wasson, J.G. (2008) Système relationnel d'audit de l'hydromorphologie des cours d'eau : principes et méthodes. Cemagref Lyon BEA/LHQ, p. 64.

Chandesris, A., Mengin, N., Malavoi, J.R., Souchon, Y. & Wasson, J.G. (2009) Système Relationnel d'Audit de l'Hydromorphologie des Cours d'Eau - Atlas à large échelle. Cemagref Lyon BEA/LHQ, p. 48.

Comeleo, R.L., Paul, J.F., August, P.V., Copeland, J., Baker, C., Hale, S.S. & Latimer, R.W. (1996) Relationships between watershed stressors and sediment contamination in Chesapeake Bay estuaries. *Landscape Ecology,* **11,** 307-319.

Congalton, R. (2004) Putting the map back in map accuracy assessment. A peer-reviewed chapter IN: Lunetta, R.S., and J.G. Lyon (Eds.), Remote Sensing and GIS Accuracy Assessment, CRC Press, Boca Raton, FL 304p.

Congalton, R.G. (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment,* **37,** 35-46.

Congalton, R.G., Birch, K., Jones, R. & Schriever, J. (2002) Evaluating remotely sensed techniques for mapping riparian vegetation. *Computers and Electronics in Agriculture,* **37,** 113-126.

CemOA : archive ouverte d'Irstea / Cemagref

Cook, R.D. (1977) Detection of Influential Observations in Linear Regression. *Technometrics (American Statistical Association),* **19,** 15-18.

Cooper, J.R., Gilliam, J.W., Daniels, R.B. & Robarge, W.P. (1987) Riparian Areas as Filters for Agricultural Sediment. *Soil Science Society of America Journal,* **51,** 416-420.

Corgne, S. (2004) *Modélisation prédictive de l'occupation des sols en contexte agricole intensif : Application à la couverture hivernale des sols en Bretagne.* Université de Rennes 2, p. 230.

Cormier, S.M., Smith, M., Norton, S. & Neiheisel, T. (2000) Assessing ecological risk in watersheds: A case study of problem formulation in the Big Darby Creek watershed, Ohio, USA. *Environmental Toxicology and Chemistry,* **19,** 1082-1096.

Corp, G.E., Olivier, J.M., Penaz, M. & Roux, A.L. (1992) Juvénile fishes as functional describers of fluvial ecosystem dynamics : applications on the River Rhône, France. *Reg. Rivers Res. Manag.,* **6,** 135-145.

Correll, D.L. (2000) The current status of our knowledge of riparian buffer water quality functions. *American Water Resources Association*.

Coulter, L., Stow, D., Hope, A., O'leary, J., Turner, D., Longmire, P., Peterson, S. & Kaiser, J. (2000) Comparison of high spatial resolution imagery for efficient generation of GIS vegetation layers. *Photogrammetric Engineering and Remote Sensing* **66,** 1329-1335.

Coutellier, A. (2004) L'épuration des eaux usées urbaines. *Les données de l'environnement,* **98,** 4.

Cramer, E.M. (1985) *Multicollinearity,* John Wiley and Sons. **,** pp. 639-644.

Croll, B.T. (1991) Pesticides in Surface Waters and Groundwaters. *Journal of the Institution of Water and Environmental Management,* **5,** 389-395.

Cspnb (2008) *L'arbre, la rivière et l'homme*, Paris**,** p. 63.

Cuffney, T.F., Meador, M.R., Porter, S.D. & Gurtz, M.E. (2000) Responses of Physical, Chemical, and Biological Indicators of Water Quality to a Gradient of Agricultural Land Use in the Yakima River Basin, Washington. *Environmental Monitoring and Assessment,* **64,** 259-270.

Cugier, P. & Le Hir, P. (2002) Development of a 3D Hydrodynamic Model for Coastal Ecosystem Modelling. Application to the Plume of the Seine River (France). *Estuarine, Coastal and Shelf Science,* **55,** 673-695.

Davis, P.A., Staid, M.I., Plescia, J.B. & Johnson, J.R. (2002) Evaluation of airborne image data for mapping riparian vegetation within the Grand Canyon. U.S. Department of the Interior, U.S. Geological Survey, p. 65.

Davy-Bowker, J., Clarke, R., Johnson, R., Kokes, J., Murphy, J. & Zahrádková, S. (2006) A comparison of the European Water Framework Directive physical typology and RIVPACS-type models as alternative methods of establishing reference conditions for benthic macroinvertebrates. *Hydrobiologia,* **566,** 91-105.

Décamps, H. (2002) Du bon usage des ripisylves. *La houille blanche,* **3,** 49-53.

Décamps, H. (2006) Ecosystèmes aquatiques. In: *in de Marsily, G. Les Eaux Continentales.* (Ed^Eds E.E. Sciences), p. 328. Rapport sur la science et la technologie n°25.

Décamps, H. (2008) Au bord du cours d'eau : des écosystèmes utiles. *Dossier pour la science.* pp. 42-47.

Definiens (2008) Definiens Developper 7.0® User Guide. Definiens AG.

Delong, M.D. & Brusven, M.A. (1998) Macroinvertebrate community structure along the longitudinal gradient of an agriculturally impacted stream. *Environmental Management,* **22,** 445-457.

Devaux, N. (2005) Détection automatique et cartographie de l'habitat rural dans les pays en développement par télédétection et analyse spatiale. In: *7èmes Rencontres de Théo Quant.* (Besançon, janvier 2005), pp. 26-28.

Devaux, N. (2006) *Images très haute résolution et analyse spatiale pour localiser l'habitat rural des pays en développement.* Université d'Orléans, p. 322.

Dion, M. (2009) *Le suivi de l'intégrité écosystémique sur le territoire québécois: emplacement et avancement sur l'échiquier environnemental international* Master thesis, Université de Sherbrooke, Sherbrooke, Québec, Canada, p. 182.

Dovciak, A.L. & Perry, J.A. (2002) In Search of Effective Scales for Stream Management: Does Agroecoregion, Watershed, or Their Intersection Best Explain the Variance in Stream Macroinvertebrate Communities? *Environmental Management,* **30,** 365-377.

Dow, C.L., Arscott, D.B. & Newbold, J.D. (2006) Relating major ions and nutrients to watershed conditions across a mixed-use, water-supply watershed. *Journal of the North American Benthological Society,* **25,** 887-911.

Dowling, R. & Accad, A. (2003) Vegetation classification of the riparian zone along the Brisbane River, Queensland, Australia, using light detection and ranging (lidar) data and forward looking digital video. *Canadian Journal of Remote Sensing,* **29,** 556-563.

CemOA : archive ouverte d'Irstea / Cemagref

Ducharne, A., Brunstein, D., Billen, G. & Garnier, J. (2001) Couplage du modèle biogéochimique Sénèque avec un modèle hydrologique à bases physiques. *Rapport d'activité 2000 du Programme PIREN-Seine.* pp. B1-B14.

Dudgeon, D., Arthington, A., Gessner, M.O., Kawabata, Z.-I., Knowler, D.J., Leveque, C., Naiman, R.J., Prieur-Richard, A.-H., Soto, D., Stiassny, M.L.J. & Sullivan, C.A. (2006) Freshwater biodiversity: importance, threats, status, and conservation challenges.

Dupuis, M.F. & Fischesser, B. (1997) Richesse, diversité et enjeux des paysages de rivières. *Ingénieries - E A T,* **n° spécial Rivières et paysages,** 11-31.

Durieux, L., Lagabrielle, E. & Nelson, A. (2007) A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. *ISPRS Journal of Photogrammetry and Remote Sensing,* **63,** 399-408.

Durrieu, S., Tormos, T., Kosuth, P. & Golden, C. (2007) nfluence of training sampling protocol and of feature space optimization methods on supervised classification results. In: *Geoscience and Remote Sensing Symposium - IGARSS 2007* (Barcelona, pp. 2030-2033. IEEE International.

Ecostat Group (2007) WFD Intercalibration technical report, Rivers, Benthic macroinvertebrates. *Edited by Olsauskyte V. & W. van de Bund.* Rapport JRC, p. 51.

Eea, E.E.A. (2003) Europe´s Water : an indicator-based assesment. Office for Official Publications of the European Communities, Luxembourg, p. 97.

El Hajj, M. (2008) *Exploitation des séries temporelles d'images satellite à haute résolution spatiale par fusion d'informations multi-sources pour le suivi des opérations culturales - Application à la détection des coupes de canne à sucre à la Réunion.* AgroParisTech, p. 234.

Esri France *Comment fonctionne un SIG,* http://www.esrifrance.fr/sig3.asp, 20 novembre.

European Commission (2000) Directive 2000/60/EC, Establishing a framework for community action in the field of water policy. *3639/1/100.* European Commission PE-CONS, Luxemborg.

Even, S., Poulin, M., Garnier, J., Billen, G., Servais, P., Chesterikoff, A. & Coste, M. (1998) River ecosystem modelling : application of the PROSE model to the Seine river (France). *Hydrobiologia,* **373,** 27-37.

Even, S., Billen, G., Bacq, N., Théry, S., Ruelland, D., Garnier, J., Cugier, P., Poulin, M., Blanc, S., Lamy, F. & Paffoni, C. (2007) New tools for modelling water quality of hydro-systems : an application in the Seine River basin in the frame of the Water Framework Directive. *The Science of The Total Environment,* **37,** 274-291.

Fauch, K.D. & Northcote, T.G. (1992) Large woody débris and Salmonid habitat in a small coastal British Columbia stream. Can. *J. Fish. Aquat. Sci.,* **49,** 682-693.

Ferreol, M. (2010 (in press)) Variabilité temporelle de l'indice biologique global normalisé. *Analyse descriptive.* Pole ONEMA Cemagref Lyon, p. 23.

Fitzpatrick, F.A. (2001) A comparison of multi-disciplinary methods for measuring physical condition of streams. *Water Science and Application,* **4,** 7-18.

Flipo, N. (2005) *Modélisation intégrée des transferts d'azote dans les aquifères et les rivières: Application au bassin du Grand Morin.* p. 244.

Foody, G.M. (2002) Status of land cover classification accuracy assessment. *Remote Sensing of Environment,* **80,** 185-201.

Forman, R.T.T. (1983) Corridors in a landscape their ecological structure and function. *Ekologia (CSSR),* **2,** 375-387.

Foucart, T. (2000) Colinéarité et instabilité numérique dans le modèle linéaire. *RAIRO Operations research,* **34(2)2,** 199-212.

Foucart, T. (2006) Colinéarité et régression linéaire. *Mathématiques et sciences humaines,* **n° 173,** 5-25.

Franklin, J.F. (1988) Structural and functional diversity in temperate forests. *In WILSON E.O. (éd.), Biodiversity, National Academy Press, Washington, DC*, 166-174.

Freeman, P.L. & Schorr, M.S. (2004) Influence of watershed urbanization on fine sediment and macroinvertebrate assemblage characteristics in Tennessee Ridge and Valley Streams. *Journal of Freshwater Ecology,* **19,** 353-362.

Friberg, N. (2010) Pressure-response relationships in stream ecology: introduction and synthesis. *Freshwater Biology,* **55,** 1367-1381.

Frimpong, E., Sutton, T., Engel, B. & Simon, T. (2005a) Spatial-Scale Effects on Relative Importance of Physical Habitat Predictors of Stream Health. *Environmental Management,* **36,** 899-917.

Frimpong, E.A., Sutton, T.M., Lim, K.J., Hrodey, P.J., Engel, B.A., Simon, T.P., Lee, J.G. & Le Master, D.C. (2005b) Determination of optimal riparian forest buffer dimensions for stream biota- landscape

CemOA : archive ouverte d'Irstea / Cemagref

association models using multimetric and multivariate responses. *Canadian Journal of Fisheries and Aquatic Sciences,* **62,** 1-6.

Frissell, C.A., Liss, W.J., Warren, C.E. & Hurley, M.D. (1986) A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. *Environmental Management,* **10,** 199-214.

Fuller, R.L. & Fry, T.J. (1991) The influence of température and food quality on the growth of Hydropsyche betteni (Trichoptera) and Simulium vittatum (Diptera). *J. of Freshwat.*

Ecol.,, **6,** 75-86.

Furse, M., Hering, D., Brabec, K., Buffagni, A., Sandin, L. & Verdonschot, P. (2006) The ecological status of European rivers: evaluation and intercalibration of assessment methods. *Hydrobiologia,* **566,** 1-2.

Gaillard, S., Bendjoudi, H., Brunstein, D., Amezal, A., Théry, S., Bravard, J., Abou Dagga, N. & Joyeux, E. (2005) Les zones humides riveraines dans le bassin de la Seine : délimitation, caractérisation hydrogéomorphologique et évaluations fonctionnelles. *Rapport PIREN.* p. 16.

Gambrell, R.P., Gilliam, J.W. & Weed, S.B. (1975) Denitrification in subsoils of the North Carolina coastal plain as affected by soil drainage. *J. Environ. Qual.,* **4,** 311-316.

Garcia, A., Villeneuve, B. & Wasson, J.G. (2005) Combined pressures. In: REBECCA D4. *Linkages between chemical and biological quality of surface waters. Current knowledge on indicators and methods for Water Framework Directive ecological status assessment.*

Garcia, B., Bertin, C., Ricard, J., Bourg, A., Lavandier, P. & Labroue, L. (1994) Effet de berge, effet de vase, deux facteurs différents de mobilisation du manganèse : un exemple dans un champ captant de la vallée du Lot (France). *Annls Limnol.,* **30,** 67-85.

Garnier, J., Billen, G. & Coste, M. (1995) Seasonal succession of diatoms and chlorophyecae in the drainage network of the River Seine : Observations and modelling. *Limnol. & Oceanogr,* **40,** 750-765.

Gayraud, S., Herouin, E. & Philippe, M. (2002) The clogging of stream beds: A review of mechanisms and consequences on habitats and macroinvertebrate communities. *Bulletin Francais De La Peche Et De La Pisciculture***,** 339-355.

Geladi, P. & Kowalski, B.R. (1986) Partial Least-Squares Regression - a Tutorial. *Analytica Chimica Acta,* **185,** 1-17.

Gergel, S.E., Turner, M.G., Miller, J.R., Melack, J.M. & Stanley, E.H. (2002) Landscape indicators of human impacts to riverine systems. *Aquatic Sciences,* **34,** 118-128.

Gergel, S.E. (2005) Spatial and non-spatial factors: When do they affect landscape indicators of watershed loading? *Landscape Ecology,* **20,** 177-189.

Gergel, S.E., Stange, Y., Coops, N.C., Johansen, K. & Kirby, K.R. (2007) What is the value of a good map? An example using high spatial resolution imagery to aid riparian restoration. *Ecosystems,* **10,** 688-702.

Gido, K.B., Falke, J.A., Oakes, R.M. & Hase, K.J. (2006) Fish-habitat relations across spatial scales in prairie streams. *Landscape Influences on Stream Habitats and Biological Assemblages,* **48,** 265-285.

Gineste, P. (1998) *Contribution de l'imagerie satellitale radar ERS à la détection des zones saturées et à la modélisation hydrologique d'un petit bassin versant agricole par TOPMODE.* Montpellier, p. 224.

Girard, M.C. & Girard, C.M. (1999) *Traitement des données de télédétection*, Paris**,** p. 390.

Gleick, P.H. (1996) Basic water requirements for human activities: Meeting basic needs. *Water International,* **21,** 83-92.

Gleick, P.H. (2008) La valeur d'une goutte d'eau. *Dossier pour la science.* pp. 70-74.

Glorennec, P.-Y. (1999) *Algorithmes d'apprentissage pour systèmes d'inférence floue,* Editions Hermès, Paris.

Goetz, S.J., Wright, R., Smith, A.J., Zinecker, E. & Schaub, E. (2003) Ikonos imagery for resource management: Tree cover, impervious surfaces and riparian buffer analyses in the mid-Atlantic region. *Remote Sens. Environ,* **88,** 195-208.

Goetz, S.J. (2006) Remote sensing of riparian buffers: Past progress and future prospects. *Journal of the American Water Resources Association,* **42,** 133-143.

Golden, C. (2006) *Influence du choix des parcelles d'apprentissage et de validation sur la qualité des résultats de classifications supervisées orientées objet d'images de télédétection.* Autres diplômes thesis p. 50 + annexes pp.

Gorman, T.O. & Karr, J.R. (1978) Habitat structure and stream fish communities. *Ecology,* **59,** 507-515.

Graham, A.A. (1989) *The effects of siltation and flow characteristics on the productivity and composition of lotic periphyton.* Université d'Otago, p. 253

Graham, M.H. (2003) Confronting multicollinearity in ecological multiple regression. *Ecology,* **84,** 2809-2815.

Gregory, K.J. & Walling, D.E. (1973) *Drainage Basin Form and Process, a Geomorphological Approach,* Londres, Ed. Arnold**,** p. 456.

Gregory, K.J. & Davis, R.J. (1992) Coarse woody débris in stream channels in relation to

CemOA : archive ouverte d'Irstea / Cemagref

river channel management in woodiand areas. *Reg. Rivers Res. Manag.,* **7,** 117-136.

Grenier, M. (2007) Identification des communautés de référence de macro-invertébrés pour l'évaluation du niveau d'intégrité écologique des écosystèmes aquatiques : Comparaison des approches a priori et a posteriori et proposition d'un indice multivarié. Institut national de la recherche scientifique (INRS-ETE), Québec, p. 73.

Grenier, M., Labrecque, S., Benoit, M. & Allard, M. (2008) Accuracy assement method for wetland object-based classification In: *Geographic Object Based Image Analysis (GEOBIA 2008), ISPRS Commision IV* (Calgary, Alberta, Canada.

Grüber, T.R. (1993) A translation approach to portable ontologies. *Knowledge Acquisition,* **5,** 199-220.

Guenat, C., Bureau, F. & Charnet, F. (2003) Genèse, diversité et fonctionnement des sols sous la dépendance d'un cours d'eau: écologie, fonctions et gestion. In: *Forêts riveraines des cours d'eau.* (Ed^Eds, pp. 46-71. Institut pour le Développement Forestier, Paris.

Guerrini, M.C., Mouchel, J.M., Meybeck, M., Penven, M.J., Hubert, G., Muxart, T. & Meybeck, M.E., De Marsily G., Fustec E. (Coordinateurs Du Programme). La Seine En Son Bassin, Elsevier, 750 P (1998) Chap 1, Le bassin de la seine : confronntation du rural et de l'urbain. In: *La seine en son bassin : Fonctionnement écologique d'un système fluvial anthropisé.* (Ed^Eds M. Meybeck), pp. 23-75. Elsevier, Paris, Amsterdam, Lausane, New York, Oxford, Tokyo.

Guillaume, S. (2001) Designing fuzzy inference systems from data : an interpretability-oriented review. *IEEE Transactions on Fuzzy Systems,* **9,** 426-443.

Guyot, G. (1997) *Climatologie de l'environnement. De la plante aux écosystèmes,* Masson, Paris**,** p. 505.

Haralick, R.M., Shanmugam, K. & Dinstein, I. (1973) Textural features for image classification. *IEEE Trans. on Systems, Man, and Cybernetics,* **SMC-3,** 610-621.

Harding, J.S., Benfield, E.F., Bolstad, P.V., Helfman, G.S. & Jones, E.B.D. (1998) Stream biodiversity: The ghost of land use past. *Proc. Natl. Acad. Sci. USA,* **95,** 14843–14847.

Harmon, M.E., Franklin, J.F., Swanson, F.J., Sollins, P., Gregory, S.V., Lattin, J.D., Anderson, N.H., Cline, S.P., Aumen, N.G., Sedell, J.R., Lienkaemper, G.W., Cromack Jr, K. & Cumming, K.W. (1986) Ecology of coarse woody debris in temperate ecosystems. *Advances in ecological research.* London, pp. 133-302.

Haupt, H.F. & Kidd, W.J. (1965) Good Logging Practices Reduce Sedimentation in Central Idaho. *Journal of Forestry,* **63,** 664-670.

Hawkins, G.R., Murphy, M.L., Anderson, N.H. & Wilzbach, M.A. (1983) Density of fish and salamanders in relation to riparian canopy and physical habitat in streams of the Northwestern United States. *Can. J. Fish. Aquat. Sci.,* **40,** 1173-1185.

Haycock, N.E., Pinay, G. & Walker, C. (1993) Nitrogen retention in river corridors: European perspective. *Ambio,* **22,** 340-346.

Hering, D., Moog, O., Sandin, L. & Verdonschot, P. (2004) Overview and application of the AQEM assessment system. *Hydrobiologia,* **516,** 1-20.

Hoerl, A.E. & Kennard, R.W. (1970) Ridge regression: biased estimation for nonorthogonal problems *Technometrics 12***,** 55-67.

Hollenhorst, T., Host, G. & Johnson, L. (2006) Scaling issues in mapping riparian zones with remote sensing data: quantifying errors and sources of uncertainty. . In: *Scaling and uncertainty analysis in ecology.* (Ed^Eds J. Wu & B. Jones & H. Li & O. Loucks), pp. 275-295. Springer, Dordrecht, The Netherlands.

House, R.A. & Boehne, P. (1986) Effects of instream structures on salmonid habitat and populations in Tobe Creek, Oregon. *Northi Am. J. of Fish. Man.,* **6,** 38-46.

Hubert-Moy, L., Michel, K., Corpetti, T. & Cle?Ment, B. (2006) Object-oriented mapping and analysis of wetlands using SPOT 5 data. In: *International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 3447-3450.

Huet, M. (1949) Appréciation de la valeur piscicole des eaux douces. *Travaux de la station de Recherches de Groenendaal, série D,* **10**.

Hughes, R.M., Wang, L. & Seelbach, P.W. (2006) *Landscapes influences on Stream Habitats and Biological Assemblages,* American Fisheries Society, Bethesda, Maryland.

Hunter, C.J. (1991) *Better trout habitat, a guide to stream restoration and management,* Palmer T. (éd.), Montana Land Reliance**,** p. 320.

Hupp, C.R. & Osterkamp, W.R. (1985) Bottomiand végétation distribution along Passage Creek, Virginia, in relation to fluvial landforms. *Ecology,* **63 6,** 670-681.

Hupp, C.R. & Osterkamp, W.R. (1996) Riparian vegetation and fluvial geomorphic processes. *Geomorphology,* **14,** 277-295.

Hynes, H.B.N. (1975) The stream and its valley. *Verh. Internat. Verein. Limnol.,* **19,** 1-15.

CemOA : archive ouverte d'Irstea / Cemagref

Ign *BD ORTHO® V2, Descriptif de contenu,* http://professionnels.ign.fr/DISPLAY/000/506/445/5064454/BDORTHO_specification.pdf,

Illies, J. & Botosaneanu, L. (1963) Problèmes et méthodes de la classification et de la zonation écologiques des eaux courantes. *Association internationale de Limnologie***,** 12.

Inamdar, S.P. (2006) Challenges in Modeling Hydrologic and Water Quality Processes in Riparian Zones. *Journal of the American Water Resources Association (JAWRA),* **42,** 5-14.

Infante, D.M., Wiley, M.J. & Seelbach, P.W. (2006) Relationships among channel shape, catchment characteristics, and fish in lower Michigan streams. *Landscape Influences on Stream Habitats and Biological Assemblages,* **48,** 339-357.

Irons, J.R., Markham, B.L., Nelson, R.F., Toll, D.L., Williams, D.L., Latty, R.S. & Stauffer, M.L. (1985) The effects of spatial resolution on the classication of Thematic Mapper data. *International Journal of Remote Sensing,* **6,** 1385-1403.

Ivits, E. & Koch, B. (2002) Object-Oriented Remote Sensing Tools for Biodiversity Assessment: a European Approach. *Millpress Science***,** 8.

Jacobs, T.C. & Gilliam, J.W. (1985) Riparian Losses of Nitrate from Agricultural Drainage Waters. *J Environ Qual,* **14,** 472-478.

Jamieson, D.G. & Fedra, K. (1996) The « Waterware » decision-support system for river-basin planning. 1. Conceptual design. *Journal of Hydrology,* **177,** 163-175.

Johansen, K. & Phinn, S. (2006a) Mapping Structural Parameters and Species Composition of Riparian Vegetation Using IKONOS and Landsat ETM+ Data in Australian Tropical Savannahs. *Photogram. Eng. Remote Sensing,* **72,** 71-79.

Johansen, K. & Phinn, S. (2006b) Linking riparian vegetation spatial structure in Australian tropical savannas to ecosystem health indicators: Semi-variogram analysis of high spatial resolution satellite imagery. *Canadian Journal of Remote Sensing,* **32,** 228-243.

Johansen, K., Phinn, S., Dixon, I., Douglas, M. & Lowry, J. (2007) Comparison of image and rapid field assessments of riparian zone condition in Australian tropical savannas. *Forest Ecology and Management,* **240,** 42-60.

Johansen, K., Phinn, S. & Witte, C. (2010) Mapping of riparian zone attributes using discrete return LiDAR, QuickBird and SPOT-5 imagery: Assessing accuracy and costs. *Remote Sensing of Environment,* **114,** 2679-2691.

Johnson, L. & Gage, S. (1997) Landscape approaches to the analysis of aquatic ecosystems. *Freshwater Biology,* **37,** 113-132.

Johnson, L., Richards, C., Host, G. & Arthur, J. (1997) Landscape influences on water chemistry in Midwestern stream ecosystems. *Freshwater Biology,* **37,** 193-208.

Johnson, L.B. (1990) Analyzing spatial and temporal phenomena using geographical information systems. *Landscape Ecology,* **4**.

Johnson, L.B. & Host, G.E. (2010) Recent developments in landscape approaches for the study of aquatic ecosystems. *Journal of the North American Benthological Society,* **29,** 41-66.

Johnson, R.K., Goedkoop, W. & Sandin, L. (2004) Spatial scale and ecological relationships between the macroinvertebrate communities of stony habitats of streams and lakes. *Freshwater Biology,* **49,** 1179-1194.

Johnson, S. & Covich, A. (1997) Scales of observation of riparian forests and distributions of suspended detritus in a prairie river. *Freshwater Biology,* **37,** 163-175.

Johnson, T.E., Mcnair, J.N., Srivastava, P. & Hart, D.D. (2007) Stream ecosystem responses to spatially variable land cover: an empirically based model for developing riparian restoration strategies. *Freshwater Biology,* **52,** 680-695.

Karr, J.R. (1981) Assessment of biotic integrity using fish communities. *Fisheries (Bethesda),* **6,** 21-27.

Kehmeier, J.W. (2000) *A spatially explicit method for determining the effects of watershed scale land use on stream conditions.* Master's thesis, Oregon State University

Kelsey, K.A. & West, S.D. (1998) River Ecology and Management: Lessons From the Pacific Coastal Ecoregion New York. In: *Riparian wildlife.* (Ed^Eds B.R. In Naiman Rj, Eds), pp. 235-258. Springer.

Kerans, B.L. & Karr, J.R. (1994) A Benthic Index of Biotic Integrity (B-Ibi) for Rivers of the Tennessee Valley. *Ecological Applications,* **4,** 768-785.

Khorram, S. (2003) High Resolution Mapping Land Cover Classification of the Hominy Creek Watershed. CEO Technical Report, p. 220.

CemOA : archive ouverte d'Irstea / Cemagref

King, R.S., Baker, M.E., Whigham, D.F., Weller, D.E., Jordan, T.E., Kazyak, P.F. & Hurd, M.K. (2005) Spatial considerations for linking watershed land cover to ecological indicators in streams. *Ecological Applications,* **15,** 137-153.

Klein, R.D. (1979) Urbanization and Stream Quality Impairment. *Water Resources Bulletin,* **15,** 948-963.

Klemm, D.J., Blocksom, K.A., Fulk, F.A., Herlihy, A.T., Hughes, R.M., Kaufmann, P.R., Peck, D.V., Stoddard, J.L., Thoeny, W.T., Griffith, M.B. & Davis, W.S. (2003) Development and evaluation of a Macroinvertebrate Biotic Integrity Index (MBII) for regionally assessing Mid-Atlantic Highlands streams. *Environmental Management,* **31,** 656-669.

Kratzer, E.B., Jackson, J.K., Arscott, D.B., Aufdenkampe, A.K., Dow, C.L., Kaplan, L.A., Newbold, J.D. & Sweeney, B.W. (2006) Macroinvertebrate distribution in relation to land use and water chemistry in New York City drinking-water-supply watersheds. *Journal of the North American Benthological Society,* **25,** 954-976.

Kressler, F.P., Kim, Y.S. & Steinnocher, K.T. (2003) Object-oriented land cover classification of panchromatic KOMPSAT-1 and SPOT-5 data. *Geoscience and remote sensing,* **19,** 263-269.

Lâchât, B. (1991) Le cours d'eau, Conservation, Entretien et Aménagement. *Conseil de l'Europe, Comité directeur pour la protection et la gestion de l'environnement et du milieu naturel.* Strasbourg, p. 84 p.

Lâchât, B. (1994) Guide de protection des berges de cours d'eau en techniques végétales. Ministère de l'Environnement, Paris, p. 143.

Laine, G., Dupuy, S. & Tassin, J. (2010) Typologie des espaces à dominante boisée de Mayotte. Cirad-UMR TETIS - UPR BESF, Montpellier, p. 110.

Lake, L.A. (1991) Analyses de méthodes texturales appliquées au traitement des images. Essai, Département de Géographie, Université de Sherbrooke, p. 38.

Lammert, M. & Allan, J.D. (1999) Assessing biotic integrity of streams : Effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. *Environmental Management,* **23,** 257-270.

Lange-Bertalot, H. (1979) Pollution tolerance of diatoms as a criterion for water quality estimation. *Nova Hedwigia,* **64,** 285-304.

Lattin, P., Wiginton, P., Moser, T.J., Peniston, B.E., Lindeman, D.R. & Oetter, D.R. (2004) Influence of remote sensing imagery source on quantification of riparian land cover/land use. *Journal of the American Water Resources Association,* **40,** 215-227.

Latty, R.S. & Hoffer, R.M. (1981) Computer-based Classification Accuracy Due to the Spatial Resolution Using Per-point Versus Per-field Classification Technique. In: *Symposium of Machine Processing of Remotely Sensed Data,* (pp. 384-392.

Le Hir, P., Ficht, A., Silva Jacinto, R., Lesueur, P., Dupont, J.P., Lafite, R., Brenon, I., Thouvenin, B. & Cugier, P. (2001) Fine Sediment Transport and Accumulations at the Mouth of the Seine Estuary (France). *Estuaries,* **24,** 950-963.

Lebert, F. (1993) *Modélisation des connaissances et raisonnements pour l'analyse des paysages agraires à partir de données satellitaires.* Université de Nancy 1, p. 211.

Lebourgeois, V. (2009) *Utilisation d'un système léger d'acquisitions aéroportées dans les domaines optiques réflectif et thermique pour la caractérisation de l'état hydrique et nutritionnel des cultures.* Université de la Réunion-Cirad, p. 174.

Leckie, D.G., Cloney, E., Jay, C. & Paradine, D. (2005) Automated mapping of stream features with high-resolution multispectral imagery: An example of the capabilities. *Photogrammetric Engineering and Remote Sensing,* **71,** 145-155.

Ledoux, E., Gomez, E., Monget, J., Viavattene, C., Viennot, P. & Benoit, M. (2007) Agriculture and Groundwater Nitrate Contamination in the Seine Basin. The STICS-MODCOU modelling chain. *Science of the Total Environment,* **375,** 33-47.

Lefsky, M.A., Cohen, W.B., Parker, G.G. & Harding, D.J. (2002) Lidar Remote Sensing for Ecosystem Studies. *BioScience,* **52,** 19-30.

Lennartz, S.P. (2004) Classifying and mapping forest cover types using ikonos imagery in the northeastern United States. In: *Proceedings of the ASP R S* (Denver.

Lévêque, C. (2001) *Ecologie. De l'écosystème à la biosphère.,* Dunod, Paris**,** p. 502.

Li, H. & Reynolds, J.F. (1995) On Definition and Quantification of Heterogeneity. *Oikos,* **73,** 280-284.

Lonard, R.I., Judd, F.W., Everitt, J.H., Escobar, D.E., Davis, M.R., Crawford, M.M. & Desai, M.D. (2000) Evaluation of color-infrared photography for distinguishing annual changes in riparian forest vegetation of the lower Rio Grande in Texas. *Forest Ecology and Management,* **128,** 75-81.

CemOA : archive ouverte d'Irstea / Cemagref

Lotspeich, F.B. (1980) Watersheds as the basic ecosystem : this conceptual framework provides a basis for a natural classification system. *Water Resources Bulletin,* **16,** 581-586.

Lotspeich, F.B. & Platts, W.S. (1982) An integrated land-aquatic classification system. *North American Journal of Fisheries Management,* **2,** 138-149.

Lowrance, R., Altier, L.S., Newbold, J.D., Schnabel, R.R., Groffman, P.M., Denver, J.M., Correll, D.L., Gilliam, J.W., Robinson, J.L., Brinsfield, R.B., Staver, K.W., Lucas, W. & Todd, A.H. (1995) Water Quality Functions of Riparian Forest Buffer Systems in the Chesapeake Bay Watershed. *DC. EPA 903-R-95-004/CBP/TRS* U.S. Environmental Protection Agency, ,Washington.

Lowrance, R., Altier, L.S., Williams, R.G., Inamdar, S.P., Sheridan, J.M., Bosch, D.D., Hubbard, R.K. & Thomas, D.L. (2000) REMM: The Riparian Ecosystem Management Model (REMM) *Journal of Soil and Water Conservation,* **55,** 27-34.

Lowrance, R.R., Todd, R.L. & Asmussen, L.E. (1983) Waterborne Nutrient Budgets for the Riparian Zone of an Agricultural Watershed. *Agriculture, Ecosystems, and Environment,* **10,** 371-384.

Lowrance, R.R., Todd, R.L., Jr, J.F., Hendrickson, J.O., Leonard, R. & Asmussen, L. (1984) Riparian forests as nutrient filters in agricultural watersheds. *BioScience,* **34,** 374-377.

Luchetti, G. & Fuersteburg, R. (1993) Relative fish use in urban and non-urban streams. . In: *Conference on Wild Salmon* (Vancouver, British Columbia.

Ma, Z.C. (1991) *Modélisation du transfert des nitrates : du bassin de recherche au grand bassin.* Strasbourg, p. 183.

Macary, F. & Paulais, J. (2003) Méthode d'identification de zones prédisposées aux émissions et aux transferts particulaires. *Ingénieries - Eau Agriculture & Territoires,***,** 3-47.

Maddock, I. (1999) The importance of physical habitat assessment for evaluating river health. *Freshwater Biology,* **41,** 373-391.

Magaud, H., Migeon, B., Morfin, P., Garric, J. & Vindimian, E. (1997) Modelling fish mortality due to urban storm run-off: Interacting effects of hypoxia and un-ionized ammonia. *Water Research,* **31,** 211-218.

Maillard, P. & Pinheiro Santos, N.A. (2006) A spatial-statistical approach for modeling the effect of non-point source pollution on different water quality parameters in the Velhas river watershed - Brazil. *Journal of Environmental Management,* **In Press, Corrected Proof**.

Malavoi, J.R. (2007) Impacts de la chenalisation des riviéres et principes de restauration. *Compte-rendus de l'Académie d'Agriculture de France,* **93,** 54 - 61.

Malavoi, J.R. & Adam, P. (2007) Préservation et restauration physique des cours d'eau. Aspects techniques. *Techniques, Sciences, Méthodes***,** 39-53.

Maret, T.R., Maccoy, D.E. & Carlisle, D.M. (2008) Long-term water quality and biological responses to multiple best management practices in Rock Creek, Idaho. *Journal of the American Water Resources Association,* **44,** 1248-1269.

Maridet, L. (1994) *La végétation rivulaire, facteur de contrôle du fonctionnement écologique des cours d'eau : influence sur les communautés benthiques et hyporhéiques et sur les peuplements de poissons dans trois cours d'eau du Massif Central.* Thèse de doctorat, Université Claude Bernard LYON I, p. 275.

Maridet, L. & Piegay, H. (1994) Formations végétales arborées riveraines des cours d'eau et potentialités piscicoles. *Bull. Fr. Pêche Piscic.,* **333,** 125-147.

Maridet, L. (1995) Rôle des formations végétales riveraines. Recommandations pour une gestion régionalisée. *Rapport final.* Convention n°1/94. Cemagref, Bely/LHQ. Ministère de l'Environnement. SDMAP PARIS, p. 69.imprimé.

Maridet, L. & Collin Huet, M.P. (1995) La végétation aux abords des rivières : source de vie et d'équilibre. Ministère de l'Environnement Direction de l'Eau.http://cemadoc.cemagref.fr/cemoa/PUB00003149

Maridet, L. & Souchon, Y. (1995) Habitat potentiel de la truite fario (Salmo trutta fario, L. 1758) dans trois cours d'eau du Massif Central : approche méthodologique et premiers résultats sur le rôle de la végétation rivulaire arborée. CemOA Publications - Publications scientifiques et techniques du Cemagref http://cemadoc.cemagref.fr/cemoa/PUB00012477

Martens, H. & Martens, M. (2000) Modified Jack-knife Estimation of Parameter Uncertainty in Bilinear Modelling by Partial Least Squares Regression (PLSR). *Food Quality and Preference***,** 5-16.

Massy, W.F. (1965) Principal components regression in exploratory statistical research. *J. Amer. Stat. Assoc,* **60,** 234-256.

Mayer, P.M., Reynolds, S.K., Jr., Mccutchen, M.D. & Canfield, T.J. (2007) Meta-Analysis of Nitrogen Removal in Riparian Buffers. *J Environ Qual,* **36,** 1172-1180.

Mcgarigal, K. & Marks, B.J. (1995) FRAGSTATS:Spatial Pattern Analysis Program for Quantifying Landscape Structure. *General Technical Report.* United States Deparment of Agriculture, Coastal Oregon, p. 132.

CemOA : archive ouverte d'Irstea / Cemagref

Meador, M.R. & Goldstein, R.M. (2003) Assessing Water Quality at Large Geographic Scales: Relations Among Land Use, Water Physicochemistry, Riparian Condition, and Fish Community Structure. *Environmental Management,* **31,** 504-517.

Medd/De & Aquascop (2003) Mise en œuvre de la DCE Identification des pressions et impacts. *Guide méthodologique version 4.1.*

Meehan, W.R., Swanson, F.J. & Sedell, J.R. (1977) Influence of riparian végétation on aquatic ecosystems with particular références to salmonid fishes and their food supply. USDA Forest Sen/ice, Gen. Tech. Rep. RM 43, pp. 137-145.

Meehan, W.R., Brusven, M.E. & Ward, J.R. (1987) Effects of artificial shading on distribution and abundance of juvénile chinook salmon. *Great Basin Naturalist,* **47** 22-31.

Meybeck, M., De Marsily, G. & Fustec, E. (1998) *La Seine en son bassin. Fonctionnement écologique d'un système fluvial anthropisé,* Paris, Elsevier**,** p. 749.

Meygret, A. (2007) SPOT absolute calibration : synthesis, CNES technical report, S5-NT-0-2880-CN, Ed2, rev 0.

Mollot, L.A., Munro, D. & Bilby, R.E. (2007) Classifying fine-scale spatial structure of riparian forests using hyperspectral high-resolution remotely sensed imagery at the Cedar River municipal watershed in western Washington, USA. *Canadian Journal of Remote Sensing,* **33,** 99-108.

Morse, C.C., Huryn, A.D. & Cronan, C. (2003) Impervious surface area as a predictor of the effects of urbanization on stream insect communities in Maine, U.S.A. *Environmental Monitoring and Assessment,* **89,** 95-127.

Müller, E., Décamps, H. & Dobson, M. (1993) Contribution of space remote sensing to river studies. *Freshwater Biology,* **29,** 301-312.

Müller, E. (1997) Mapping riparian vegetation along rivers: old concepts and new methods. *Aquatic Botany,* **58,** 437.

Nagler, P.L., Glenn, E.P. & Huete, A.R. (2001) Assessment of spectral vegetation indices for riparian vegetation in the Colorado River delta, Mexico. *Journal of Arid Environments,* **49,** 91-110.

Naiman, R.J. & Décamps, H. (1990) *Ecology and Management of Aquatic-Terrestrial Ecotones*, Paris**,** p. 316.

Naiman, R.J., Lonzarich, D.G., Beechie, T.J. & Ralph, S.C. (1992) General principles of classification and the assessment of conservation potential in rivers. *In : Boon P., Calow P. & Petts G. eds., River Conservation and Management. Wiley & Sons***,** 93-123.

Naiman, R.J. & Décamps, H. (1997) The ecology of interfaces:Riparian Zones. *Annual Review of Ecology and Systematics,* **28,** 621-658.

Naiman, R.J., Décamps, H. & Mcclain, M.E. (2005) *Riparia : ecology, conservation, and management of streamside communities,* Elsevier Academic, ©2005, Amsterdam ; Boston.

Neale, C.M.U. (1997) Classification and mapping of riparian systems using airborne multispectral videography. *Restoration Ecology,* **5,** 103-112.

Nf T90-344 (2004-05-01) Qualité de l'eau - Détermination de l'indice poissons rivière (IPR).

Nf T90-350 (2004-03-01) Qualité de l'eau - Détermination de l'indice biologique global normalisé (IBGN).

Nf T90-354 (2007-12-01) Qualité de l'eau - Détermination de l'Indice Biologique Diatomées (IBD).

Nf T90-395 (2003-10-01) Qualité de l'eau - Détermination de l'indice biologique macrophytique en rivière (IBMR).

Niemi, G.J., Devore, P., Detenbeck, N., Taylor, D., Lima, A., Pastor, J., Yount, J.D. & Naiman, R.J. (1990) Overview of case studies on recovery of aquatic systems from disturbance. *Environmental Management,* **14,** 571-587.

Noonan, M. & Chafer, C. (2007) A method for mapping the distribution of willow at a catchment scale using bi-seasonal SPOT5 imagery. *Weed Research,* **47,** 173-181.

Nutter, W.L. & Gaskin, J.W. (1988) Role of streamside management zones in controlling discharges to wetlands. In: *The Forested Wetlands of the Southern United States.* (Ed^Eds, pp. 81-84. USDA Forest Service., Orlando, FL.

O'connor, N.A. (1992) Quantification of submerged wood in a lowland Australian stream. *System. Freshwat. Biol.,* **27,** 387-395.

O'neill, R.V., De Angelis, D.L., Waide, J.B. & Allen, T.H.F. (1986) *A hierarchical concept of ecosystems,* Princeton University Press.

Omernik, J.M. (1987) Ecoregions of the conterminous United States. *Annals of the Association of American Geographers,* **77,** 118-125.

Pan, Y., Herlihy, A., Kaufmann, P., Wigington, J., Van Sickle, J. & Moser, T.J. (2004) Linkages among land-use, water quality, physical habitat conditions and lotic diatom assemblages: A multi-spatial scale assessment. *Hydrobiologia,* **515,** 59-73.

CemOA : archive ouverte d'Irstea / Cemagref

Paparoditis, N., Souchon, J.P., Martinoty, G. & Pierrot-Deseilligny, M. (2006) High-end aerial digital cameras and their impact on the automation and quality of the production workflow. *ISPRS Journal of Photogrammetry and Remote Sensing,* **60,** 400-412.

Parsons, M. & Norris, R. (1996) The effect of habitat-specific sampling on biological assessment of water quality using a predictive model. *Freshwater Biology,* **36,** 419-434.

Parsons, M., Thoms, M.C. & Norris, R.H. (2004) Development of a Standardised Approach to River Habitat Assessment in Australia. *Environmental Monitoring and Assessment,* **98,** 109-130.

Patty, L., Real, B. & Gril, J.J. (1997) he use of grassed buffer strips to remove pesticides, nitrate and soluble phosphorus compounds from runoff water. *T. Pestic. Sci.,* **49,** 25139:22105.

Paul, M.J. & Meyer, J.L. (2001) Streams in the urban landscape. *Annual Review of Ecology and Systematics,* **32,** 333-365.

Payraudeau, S. (2002) *Modélisation distribuée des flux d'azote sur des petits bassins versants Méditerranéens.* Montpellier, p. 424.

Pella, H., Chandesris, A. & Wasson, J.-G. (2004) Constitution d'un système d'information à référence spatiale dans le contexte de la Directive Cadre sur l'Eau. *Ingénieries-EAT,* **40,** 3-10.

Pella, H., Sauquet, E. & Chandesris, A. (2006) Construction d'un réseau hydrographique simplifié à partir de la BD Carthage®. *Ingénieries - E A T,* **46,** 3-14.

Perrez Correa, M. (2004) Classification orientée objet d'images à très haute résolution spatiale: application à la cartographie de l'occupation des sols le long des cours d'eau. Cemagref Montpellier, p. 40.

Peterjohn, W.T. & Correll, D.L. (1984) Nutrient Dynamics in an Agricultural Watershed: Observations in the Role of the Riparian Forest. *Ecology,* **65,** 1466-1475.

Petersen, I., Winterbottom, J.H., Orton, S., Friberg, N., Hildrew, A.G., Spiers, D.C. & Gurney, W.S.C. (1999) Emergence and lateral dispersal of adult Plecoptera and Trichoptera from Broadstone Stream, UK. *Freshwater Biology,* **42,** 401-416.

Petersen, I., Masters, Z., Hildrew, A.G. & Ormerod, S.J. (2004) Dispersal of adult aquatic insects in catchments of differing land use. *Journal of Applied Ecology,* **41,** 934-950.

Piegay, H., Pautou, G. & Ruffinoni, C. (2003) Les forets riveraines des cours d'eau: Ecologie, fonctions et gestion. I.P.L.D. Forestier). IDF, Paris.

Piégay, H., Arnaud, D. & Souchon, Y. (2003) Effets de la végétation riveraine sur la géométrie des lits fluviaux : études de cas dans le Massif Central. *Géomorphologie : relief processus environnement,* **2,** 111-128.

Pinay, G. (1986) Impact of a riparian forest on the nitrogen content of phreatic water in the Garonne basin. *LAUGA J., DÉCAMPS H., ROLLAND M.M. (éd.), Land use impacts on aquatic ecosystemsMAB UNESCO, PIREN-CNRS*, 303-317.

Pinay, G. & Décamps, H. (1988) The role of riparian woods in regulating nitrogen fluxes between the alluvial aquifer and surface water: a conceptual model. . *Regulated Rivers : Research and Management,* **2,** 507-516.

Pinay , G., Ruffinoni, C., Wondzell, S. & Gazelle, F. (1998) Change in groundwater nitrate concentration in a large river floodplain: denitrification, uptake or mixing ? *Journal of North American Benthological Society,* **17,** 179-189.

Piren Seine *Qu'est-ce que le PIREN Seine ?,* http://www.sisyphe.upmc.fr/piren/presentation-piren-seine, 20 juillet.

Planty-Tabacchi, A.M., Tabacchi, E., Naiman, R.J., Deferrari, C. & Décamps, H. (1996) Invasibility of species-rich communities in riparian zones. . *Conservation Biology,* **10,** 598-607.

Potter, K.M., Cubbage, F.W. & Schaberg, R.H. (2005) Multiple-scale landscape predictors of benthic macroinvertebrate community structure in North Carolina. *Landscape and Urban Planning,* **71,** 77-90.

Prochazka, K. & Stewart, B.A. (1991) Leaf litter rétention and its implications for shredder distribution in two headwater streams. *Arch. Hydrobiol.,* **120** 315-325.

Provencher, L. & Dubois, J.M. (2007) *Précis de télédétection, volume 4 : Photointerprétation,* Presses Universite Du Quebec p. 540.

Puissant, A. & Weber, C. (2004) Demarche orientée ''Objets - Attributs'' et classification d'images THRS. *Revue Française de Photogrammétrie et de Télédétection,* **1-2,** 123-134.

Pyle, E. (1999) The use of macroinvertebrates in water management. Ministry for the environnement, Nouvelle-Zélande.http://www.mfe.govt.nz/publications/water/macroinvertebrates-in-water-mgmt-mar99.pdf.

Quenouille, M.H. (1949) Approximate tests of correlation in time series. *Journal of The Royal Statistical Society,* **Series B** 68-84.

CemOA : archive ouverte d'Irstea / Cemagref

Quinn, J.M., Williamson, R.B., Smith, R.K. & Vickers, M.L. (1992) Effects of riparian grazing and channelisation on streams in Southland, New Zealand. 2. Benthic invertebrates. *New Zealand Journal of Marine & Freshwater Research,* **26,** 259-273.

Quinn, J.M. (2000) Effects of pastoral development. In: *New Zealand Stream Invertebrates: Ecology and Implications for Management.* (Ed^Eds K.J. Collier & M.J. Winterbourn), pp. 208-229. Caxton, Christchurch, NZ.

Raven, P., Fox, J.P., Everard, M., Holmes, N.T.H. & Dawson, F.H. (1997) River habitat survey: a new system for classifying rivers according to their habitat quality. In: *Freshwater quality: defining the indefinable The Stationery Office (Scottish Natural Heritage).* (Ed^Eds P.J. Boon & D.L. Howell), pp. 215-234 Edinburgh, Scotland.

Repaka, S.R. (2004) *Comparing spectral and object based approaches for classification and transportation feature extraction from high resolution multispectral imagery.* Mississippi State University, Mississippi, p. 162.

Reynoldson, T.B., Bailey, R.C., Day, K.E. & Norris, R.H. (1995) Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. *Austral Ecology,* **20,** 198-219.

Richards, C., Johnson, L.B. & Host, G.E. (1996) Landscape-scale influences on stream habitats and biota. *Journal canadien des sciences halieutiques et aquatiques,* **53**.

Richards, C., Haro, R., Johnson, L. & Host, G. (1997) Catchment and reach-scale properties as indicators of macroinvertebrate species traits. *Freshwater Biology,* **37,** 219-230.

Robin, M. (2002) *Télédétection. Des satellites aux SIG.,* Nathan Université, Paris, France.

Roche, P.-A., Billen, G., Bravard, J.-P., Decamps, H., Pennequin, D., Vindimian, E. & Wasson, J.-G. (2005) Les enjeux de recherche liés à la directive-cadre européenne sur l'eau. *Geosciences,* **337,** 243-267.

Roche, P.A. & Zimmer, D. (2006) Eau, aménagement et usages. In: *in de Marsily, G. Les Eaux Continentales.* (Ed^Eds E.E. Sciences), p. 328. Rapport sur la science et la technologie n°25, Paris.

Rot, B.W., Naiman, R.J. & Bilby, R.E. (2000) Stream channel configuration, landform, and riparian forest structure in the Cascade Mountains Washington. *Canadian Journal of Fisheries and Aquatic Sciences,* **57,** 699-707.

Roth, N.E., Allan, J.D. & Erickson, D.L. (1996) Landscape influences on stream biotics integity asseced at multiple spatial scales. *Landscape Ecology,* **11,** 16.

Roy, A.H., Rosemond, A.D., Leigh, D.S., Paul, M.J. & Wallace, J.B. (2003a) Habitat-specific responses of stream insects to land cover disturbance: biological consequences and monitoring implications. *Journal of the North American Benthological Society,* **22,** 292-307.

Roy, A.H., Rosemond, A.D., Paul, M.J., Leigh, D.S. & Wallace, J.B. (2003b) Stream macroinvertebrate response to catchment urbanisation (Georgia, U.S.A.). *Freshwater Biology,* **48,** 329-346.

Roy, A.H., Freeman, B.J. & Freeman, M.C. (2007) Riparian influences on stream fish assemblage structure in urbanizing streams. *Landscape Ecology,* **22,** 385-402.

Royer, A., Becker, F. & Gagnon, P. (2007) Trente-cinq ans d'observations spatiales de la terre : de la photographie à l'électromagnétométrie. *Revue Télédétection,* **7**.

Ruelland, D. (2004) SENEQUE, logiciel SIG de modélisation prospective de la qualité de l'eau. *Revue Internationale de Géomatique,* **14,** 97-117.

Ruelland, D., Billen, G., Brunstein, D. & Garnier, J. (2007) A multiscaling GIS interface to the Riverstrahler model of the biogeochemical functioning of river systems. *Science of the Total Environment,* **375,** 257-273.

Ruffinoni, C., Tremolières, M. & Zanchez-Perez, J.M. (2003) Végétation alluviale et Flux de nutriments : des liens interactifs. In: *Les forêts riveraines des cours d'eau, écologie, fonctions et gestion.* (Ed^Eds, pp. 134-154. Institut pour le Développement Forestier, Paris.

Sala, O.E., Chapin Iii, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., Poff, N.L., Sykes, M.T., Walker, B.H., Walker, M. & Wall, D.H. (2000) Global biodiversity scenarios for the year 2100. *Science,* **287,** 1770-1774.

Sandin, L., Hering, D., Buffagni, A., Lorenz, A., Moog, O., Rolauffs, P. & Stubaufer, I. (2001) The Development and Testing of an Integrated Assessment System for the Ecological Quality of Streams and Rivers Throughout Europe Using Benthic Macroinvertebrates. *3rd Deliverable, Experiences with Different Stream Assessment Methods and Outlines of an Integrated Method for Assessing Streams Using Benthic Macroinvertebrates.* EVK1-CT1999- 00027, Allemagne, p. 109.

CemOA : archive ouverte d'Irstea / Cemagref

Sandre *BDCarthage / Zones hydrographiques,* http://sandre.eaufrance.fr/geonetwork/srv/fr/metadata.show?id=367&currTab=simple, 10 septembre.

Sauquet, E. (2005) Cartographie des écoulements annuels moyens en France. p. 39.

Schmitt, T.J., Dosskey, M.G. & Hoagland, K.D. (1999) Filter Strip Performance and Processes for Different Vegetation, Widths, and Contaminants. *J Environ Qual,* **28,** 1479-1489.

Schneider, W. & Steinwender, J. (1999) Landcover mapping by interrelated segmentation and classification of satellite images. *International Archives of Photogrammetry and Remote Sensing,* **32,** Part 7-4-3 W6.

Schöpfer, E. & Lang, S. (2006) Object fate analysis - A virtual overlay method for the categorisation of object transition and object-based accuracy assessment. In: *Proceedings of the 1st International Conference on Object Based Image Analysis (OBIA 2006)* (Salsburg, Austria.

Schuler, T. (1995) The Architecture of urban stream buffers. *Watershed Protection Techniques,* **1,** 155-163.

Schulz, R. & Liess, M. (1999) A field study of the effects of agriculturally derived insecticide input on stream macroinvertebrate dynamics. *Aquatic Toxicology,* **46,** 155-176.

Schumm, S. (1977) *The fluvial system*, New York**,** p. 338.

Sedell, J.R. & Luchessa, K.J. (1982) Using the historical record as an aid to salmonid habitat enhancement. *Proc. of a Symp. on Acquisition and Utilization of Aquatic Habitat Inventory Information Held, Portiand, Oregon (U.S.A.)***,** 210-223.

Sedell, J.R., Yuska, J.E. & Speaker, R.W. (1984) Habitats and salmonid distribution in pristine, sédiment rich river valley Systems : South Fork Hoh and Queets River, Olympic National Park. *Fish and Wildiife relationships in oidgrowth forests, American Institute of Fishery Research Biologists,***,** 33-46.

Sedell, J.R., Everest, R. & Gibbons, D.R. (1987) Streamside végétation management for aquatic habitat. *Proc. of the National Silviculture Workshop, Sacramento, California***,** 115 - 125.

Shapiro, S.S. & Wilk, M.B. (1965) An analysis of variance test for normality. *Biometrika,* **52,** 591-611.

Shearer, K.S. & Xiang, W.N. (2007) The Characteristics of Riparian Buffer Studies. *Journal of Environmental Informatics,* **9,** 41-55.

Shuft, M.J., Moser, T.J., Wigington Jr, P.J., Stevens Jr, D.L., Mcallister, L.S., Chapman, S. & Ernst, T.L. (1999) Development of landscape metrics for characterizing riparian-stream networks. *PE&RS - Photogrammetric Engineering and Remote Sensing,* **65,** 1157-1167.

Sioli, H. (1982) Tropische Flusse in ihren Beziehungen zur terrestrischen Umgebung und im Hinblick auf menschliche Eingriffe. *Arch. l-lydrobiol,* **95,** 463-485.

Skinner, J.A., Lewis, K.A., Bardon, K.S., Tucker, P., Catt, J.A. & Chambers, B.J. (1997) An overview of the environmental impact of agriculture in the UK. *Journal of Environmental Management,* **50,** 111-128.

Smith, R.A., Schwarz, G.E. & Alexander, R.B. (1997) Regional interpretation of water-quality monitoring data. *Water Resources Research,* **33,** 2781-2798.

Smitz, J., Everbecq, E., Deliège, J.F., Descy, J.P., Wollast, R. & Vanderborght, J.P. (1997) PEGASE, une méthodologie et un outil de simulation prévisionnelle pour la gestion de la qualité des eaux de surface. *Tribune de l'Eau,* **588,** 73-82.

Snyder, C.D., Young, J.A., Villella, R. & Lemarié, D.P. (2003) Influences of upland and riparian land use patterns on stream biotic integrity. *Landscape Ecology,* **18,** 647-664.

Snyder, M.N., Goetz, S.J. & Wright, R.K. (2005) Stream Health Rankings Predicted by Satellite Derived Land Cover Metrics. *AWRA,* **41,** 659-677.

Souchon, Y., Andriamahefa, H., Cohen, P., Breil, P., Pella, H., Lamouroux, N., Malavoi, J.-R. & Wasson, J.-G. (2000) Régionalisation de l'habitat aquatique dans la bassin de la Loire. *Synthése.* Cemagref Lyon BEA/LHQ, Agence de l'eau Loire Bretagne, p. 291.f.

Souchon, Y. (2002) L'habitat des cours d'eau dans tous ses états. *Méroire pour l'obtention et l'habilitation à dirriger des recherches.* Cemagref Lyon BEA/LHQ et MATE/DE, p. 141.

Souchon, Y. (2007) La restauration hydraulique et biologique des cours d'eau : pourquoi s'impose-t-elle et comment la mettre en œuvre ? *Compte-rendus de l'Académie d'Agriculture de France,* **93,** 54 - 61.

Sovell, L.A., Vondracek, B., Frost, J.A. & Mumford, K.G. (2000) Impacts of Rotational Grazing and Riparian buffers on Physicochemical and Biological Characteristics of Southeastern Minnesota, USA, Streams. *Environmental Management,* **26,** 629-641.

Sponseller, R.A., Benfield, E.F. & Valett, H.M. (2001) Relationships between land use, spatial scale and stream macroinvertebrate communities. *Freshwater Biology,* **46,** 1409-1424.

Spotimage *Le Supermode® Spot 5 (Fiche technique),* http://www.spotimage.com/automne_modules_files/standard/public/p551_fileLINKEDFILE_supermode-fr.pdf, 10 Janvier

CemOA : archive ouverte d'Irstea / Cemagref

Stauffer, J.C., Goldstein, R.M. & Newman, R.M. (2000) Relationship of wooded riparian zones and runoff potential to fish community composition in agricultural streams. *Canadian Journal of Fisheries and Aquatic Sciences,* **57,** 307-316.

Steedman, R.J. (1988) Modification and assesment of an index of biotic integrity to quantify stream quality in southern Ontario. *Canadian Journal of Fisheries and Aquatic Sciences,* **45,** 492-501.

Stepenuck, K.F., Crunkilton, R.L. & Wang, L.Z. (2002) Impacts of urban landuse on macroinvertebrate communities in southeastern Wisconsin streams. *Journal of the American Water Resources Association,* **38,** 1041-1051.

Stewart, J.S., Wang, L., Lyons, J., Horwatich, J.A. & Bannerman, R. (2001) Influences of watershed, riparian-corridor, and reach-scale characteristics on aquatic biota in agricultural watersheds. *Journal of the American Water Resources Association,* **37,** 1475-1487.

Stoddard, J.L., Peck, D.V., Paulsen, S.G., Van Sickle, J., Hawkins, C.P., Herlihy, A.T., Hughes, R.M., Kaufmann, P.R., Larsen, D.P., Lomnicky, G., Olsen, A.R., Peterson, S.A., Ringold, P.L. & Whittier, T.R. (2005) An ecological assessment of western streams and rivers. *EPA/620/R-05/005.* U.S. Environmental Protection Agency, Washington, DC.

Stoffels, R.J., Clarke, K.R. & Closs, G.P. (2005) Spatial scale and benthic community organisation in the littoral zones of large oligotrophic lakes: potential for cross-scale interactions. *Freshwater Biology,* **50,** 1131-1145.

Strahler, A.N. (1957) Quantitative analysis of watershed geomorphology. *Transactions of the American Geophysical Union,* **38,** 913-920.

Strayer, D.L. (2006) Challenges for freshwater invertebrate conservation *J. N. Am. Benthol. Soc.,* **25,** 271–287.

Tabacchi, E., Correll, D., Hauer, R., Pinay, G., Planty-Tabacchi, A.M. & Wissmar, R. (1998) Development, maintenance and role of riparian vegetation in the landscape. *Freshwater Biology,* **40,** 1-21.

Tachet, H., Richoux, P., Ournaud, M. & Usseglio Polatera, P. (2000) *Invertébrés d'eau douce : Systématique, biologie, écologie,* CNRS Editions, Paris**,** p. 588.

Tenenhaus, M. (1998) *La régression PLS, théorie et pratique,* Editions Technip.

Tetzlaff, D., Uhlenbrook, S., Grottker, M. & Leibundgut, C. (2005) Hydrological assessment of flow dynamic changes by storm sewer overflows and combined sewer overflows on an event-scale in an urban river. *Urban Water,* **2,** 201-214.

Thouvenin, B., Billen, G., Even, S., Fischer, J.C., Gonzalez, J.L., Le Hir, P., Loizeau, V., Mouchel, J.M., Olivier, C. & Jacinto, S.R. (1999) Les modèles: outils de connaissance et de gestion. *Fascicule Seine-Aval N° 16.* Editon Ifremer, Plouzané, France, p. 32.

Tiede, D., Lang, S. & Hoffmann, C. (2006) Supervised and forest type-specific multi-scale segmentation for a one-level-representation of single trees. In: *1st International Conference on Object Based Image Analysis (OBIA 2006)* (Salzburg Austria.

Tormos, T. (2005) Méthode de caractérisation par télédétection des ripisylves pour l'étude de leur impact sur la qualité écologique des cours d'eau. *Rapport de DEA Sciences de l'Eau dans L'Environnement Continentale.* Université de Montpellier II, Montpellier, p. 81.

Tormos, T., Kosuth, P., Durrieu, S., Villeneuve, B. & Wasson, J.-G. (2010) Improving the quantification of land cover pressures on stream ecological status at riparian scale using High Spatial Resolution Imagery. *Physics and Chemistry of the Earth,* **XX,** XX-XX.

Townsend, C.R., Downes, B.J., Peacock, K. & Arbuckle, C.J. (2004) Scale and the detection of land-use effects on morphology, vegetation and macroinvertebrate communities of grassland streams. *Freshwater Biology,* **49,** 448-462.

Triska, F.J. & Cromack, K. (1980) The role of wood débris in forests and streams. In WARING R.H. (éd.), Forests, Fresh perspectives from ecosystem analysis. *Oregon State University Press, Corvallis, Oregon (U.S.A.)***,** 171-190.

Troll, C. (1939) Luftbildplan und ökologische Bodenforschung (Aerial photography and ecological studies of the earth). *Zeitschrift der Gesellschaft für Erdkunde Berlin***,** 241-298.

Tukey, J.W. (1958) Bias and confidence in not quite large samples. *Annals of Mathematical Statistics***,** 614.

Turner, M.G., Gardner, R.H. & O'neill, R.V. (2001) *Landscape Ecology in Theory and Practice,* Springer-Verlag, New York, NY, USA.

Usseglio-Polatera, P., Richoux, P., Bournaud, M. & Tachet, H. (2001) A functional classification of benthic macroinvertebrates based on biological and ecological traits: application to river condition assessment and stream management. *Archiv für Hydrobiologie,* **Suppl. 139,** 53-83.

CemOA : archive ouverte d'Irstea / Cemagref

Usseglio Polatera, P., Wasson, J.G. & Archaimbault, V. (2007) Protocole de prélèvement et de traitement des échantillons des invertébrés sur le réseau de contrôle de surveillance. *BO du MEDAD.*http://cemadoc.cemagref.fr/cemoa/PUB00022001

Valette, L. (2005) *Recherche de quelques indicateurs attachés à l`état de pression sur le bassin versant de l`Yzeron.* Master 2, Université Lumière Lyon 2, p. 84.

Valette, L., Chandesris, A., Mengin, N., Malavoi, J.R., Souchon, Y. & Wasson, J.G. (2008) SYstème Relationnel d'Audit de l'Hydromorphologie des Cours d'Eau SYRAH CE. Principes et méthodes de la sectorisation hydromorphologique. Cemagref LHQ-BEA Lyon.http://cemadoc.cemagref.fr/cemoa/PUB00025844

Van Dam, H., Mertens, A. & Sinkeldam, J. (1994) A coded checklist and ecological indicator values of freshwater diatoms from The Netherlands. *Neth. J. Aquat. Ecol.,* **28,** 117-133.

Van Der Sande, C.J., De Jong, S.M. & De Roo, A.P.J. (2003) A segmentation and classification approach of IKONOS-2 imagery for land cover mapping to assist flood risk and flood damage assessment. *International Journal of Applied Earth Observation and Geoinformation,* **4,** 217-229.

Van Sickle, J. (2003) Analyzing Correlations Between Stream and Watershed Attributes. *Journal of the American Water Resources Association,* **39,** 717-726.

Van Sickle, J., Baker, J., Herlihy, A., Bayley, P., Gregory, S., Haggerty, P., Ashkenas, L. & Lib, J. (2004) Projecting the biological condition of stream under alternative scenarios of human land use. *Ecological Applications,* **14,** 368–380.

Van Sickle, J. & Johnson, B.C. (2008) Parametric distance weighting of landscape influence on streams. *Landscape Ecology,* **23,** 427-438.

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,* **37,** 130-137.

Vermote, E.F., Tanre, D., Deuze, J.L., Herman, M. & Morcrette, J.-J. (1997) Second simulation of the satellite signal in the solar spectrum, 6s: an overview. *EEE Transactions on Geoscience and Remote Sensing,* **35,** 675-686.

Verneaux, J. (1976) Biotypologie de l'écosystème “eau courante”. La structure biotypologique. *C. R. Acad. Sc. Paris,* **283**.

Verneaux, J., Galmiche, P., Janier, F. & Monnot, A. (1982) Une méthode pratique d'évaluation de la qualité des eaux courantesUn indice biologique de qualité générale. *Annales Scientifiques de l'Université de Franche-Comté,* **4,** 11-21.

Vivier, P. (1970) Influence de la pollution organique sur la faune aquatique des eaux courantes. *BFP***,** 89-104.

Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K. & Van Driel, N. (2001) Completion of the 1990’s national land cover dataset for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *PE&RS - Photogrammetric Engineering and Remote Sensing,* **67,** 650-662.

Volkman, S., Barkdoll, B. & Bingner, R. (2005) Quickbird Satellite Imagery for Riparian Management: Characterizing Riparian Filter Strips and Detecting Concentrated Flow in an Agricultural Watershed. In: *Conference Proceeding Of World Water & Environmental Resources Congress 2005* (Anchorage, AK 173, 487

Vondracek, B., Blann, K., Cox, C., Nerbonne, J., Mumford, K.G., Nerbonne, B., Sovell, L. & Zimmerman, J. (2005) Land Use, Spatial Scale, and Stream Systems: Lessons from an Agricultural Region. *Environmental Management,* **36,** 775-791.

Vought, L.B.-M., Dahl, J., Pedarson, C.L. & Lacoursiere, J.O. (1994) Nutrient Retention in Riparian Ecoton. *Ambio,* **23,** 342-348.

Walker, B.H., Carpenter, S.R., Anderies, J.M., Abel, N., Cumming, G.S., Janssen, M.A., Lebel, L., Norberg, J., Peterson, G.D. & Pritchard, R. (2002) Resilience management in social-ecological systems: a working hypothesis for a participatory approach. *Conservation Ecology,* **6,** 14.

Wallace, J.B. & Benke, A.C. (1984) Quantification of wood habitat in subtropical coastal plain streams. *Can. J. Fish. Aquat. Sci.,* **41,** 1643-1652.

Wallin, M., Wiederholm, T. & Johnson, R.K. (2003) Guidance on establishing reference conditions and ecological status class boundaries for inland surface waters. CIS Working Group 2.3 - REFCOND. 7th Version.

Walser, C.A. & Bart, H.L. (1999) Influence of agriculture on in-stream habitat and fish community structure in Piedmont watersheds of the Chattahoochee River System. *Ecology of Freshwater Fish,* **8,** 237-246.

Walsh, C.J., Waller, K.A., Gehling, J. & Nally, R.M. (2007) Riverine invertebrate assemblages are degraded more by catchment urbanisation than by riparian deforestation. *Freshwater Biology,* **52,** 574-587.

CemOA : archive ouverte d'Irstea / Cemagref

Wang, L., Lyons, J., Kanehl, P. & Gatti, R. (1997) Influences of Watershed Land Use on Habitat Quality and Biotic Integrity in Wisconsin Streams. *Fisheries,* **22,** 6-12.

Wang, L., Lyons, J. & Kanehl, P. (2001) Impacts of urbanization on stream habitat and fish across multiple spatial scales. *Environmental Management,* **28,** 255-256.

Wang, L. & Kanehl, P. (2003) Influences of watershed urbanization and instream habitat on macroinvertebrates in cold water streams. *Journal of the American Water Resources Association,* **35,** 1181-1196.

Wang, L., Lyons, J., Rasmussen, P., Seelbach, P., Simon, T., Wiley, M., Kanehl, P., Baker, E., Niemela, S. & Stewart, P.M. (2003) Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest Ecoregion, U.S.A. *Canadian Journal of Fisheries and Aquatic Sciences,* **60,** 491-505.

Wang, L. (2006) Introduction to landscape influences on stream habitats and biological assemblages. In: *Landscape Influences on Stream Habitats and Biological Assemblages.* (Ed^Eds R.M. Hughes & L. Wang & P.W. Seelbach), pp. 1-23. American Fisheries Society, Bethesda, Maryland.

Ward, G.M., Cummins, K.W., Gregory, S.V., Speaker, R.W., Dudley, T. & Ward, A.K. (1982) Habitat and food resources for invertebrate communities in South Fork Hoh River, Olympic National Park, Washington. *National Parks, Washington*, 35-46.

Ward, J.V. (1989) The Four-Dimensional Nature of Lotic Ecosystems. *Journal of the North American Benthological Society,* **8,** 2-8.

Wasson, J.-G., Malavoi, J.-R., Maridet, L., Souchon, Y. & Paulin, L. (1998) *Impacts écologiques de la chenalisation des rivières,* Cemagref éditions, coll. Etudes, série Gestion des Milieux Aquatiques n°14**,** p. 158.

Wasson, J.-G. (2001) Les questions de recherche posées par la Directive Cadre Européeenne sur l' Eau : problématique pour les eaux de surfaces continentales. *Hydroecol. Appl.,* **13,** 1-19.

Wasson, J.-G., Chandesris, A., Pella, H. & Souchon, Y. (2001) Méthodologie de détermination des conditions de référence au sens de la Directive Cadre pour la gestion des eaux. *Rapport final.* Cemagref Lyon BEA/LHQ, p. 69.

Wasson, J.-G., Chandesris, A., Pella, H. & Blanc, L. (2002) Les hydro-écorégions en France métropolitaine approche régionale de la typologie des eaux courantes et éléments pour la définition des peuplements de référence d'invertébrés. *Rapport interne.* Cemagref Lyon BEA/LHQ, p. 190.

Wasson, J.-G., Chandesris, A., Pella, H., Blanc, L., Villeneuve, B. & Mengin, N. (2003a) Détermination des valeurs de Référence de l'IBGN et propositions de valeurs limites du " Bon Etat ". *Appui scientifique à la mise en œuvre de la DCE.* Cemagref Lyon BEA/LHQ, p. 74.imprimé.

Wasson, J.-G., Tusseau-Vuillemin, M.H., Andréassian, V., Perrin, C., Faure, J.-B., Barreteau, O., Bousquet, M. & Chastan, B. (2003b) What kind of water models are needed for the implementation of the European Water Framework Directive ? *International Journal of River Basin Management,* **1,** 1-11.

Wasson, J.-G., Chandesris, A., Pella, H. & Blanc, L. (2004a) Les hydro-écorégions : une approche fonctionnelle de la typologie des rivières pour la Directive cadre européenne sur l'eau. *Ingénieries-EAT,* **40,** 3-10.

Wasson, J.-G., Villeneuve, B., Mengin, N., Pella, H. & Chandesris, A. (2005) Modèles pressions/impacts : Approche méthodologique, modèles d'extrapolation spatiale et modèles diagnostics. *Appui scientifique à la mise en œuvre de la DCE.* Cemagref Lyon BEA/LHQ, p. 55.

Wasson, J.G. (1989) Eléments pour une typologie fonctionnelle des eaux courantes. Revue critique de quelques approches existantes. *Bull. Ecol.,* **20,** 109-127.

Wasson, J.G. (1992) La rivière et l'homme : vers une gestion par bassin intégrant la dimension écologique. *CemOA Publications - Publications scientifiques et techniques du Cemagref*.

Wasson, J.G., Bethemont, J., Degorce, J.N., Dupuis, B. & Joliveau, T. (1993) Approche écosystémique du bassin de la Loire : éléments pour l'élaboration des orientations fondamentales de gestion. *Phase I : Etat initial - Problématique.* Cemagref Lyon BEA/LHQ et Université St-Etienne CNRS CRENAM, p. 102.

Wasson, J.G., Pella, H., Mengin, N., Chandesris, A., Blanc, L. & Villeneuve, B. (2004b) Concepts and methods for the typology and definition of reference conditions for rivers in France. In: *Communication au congrès Typologies and reference conditions in the Rhine, Meuse and Schelde basins* (Luxembourg, 30th November 2004).

Wasson, J.G., Villeneuve, B., Mengin, N., Pella, H. & Chandesris, A. (2006) Quelle limite de « bon état écologique » pour les invertébrés benthiques en rivières ? Apport des modéles d'extrapolation spatiale reliant l'indice biologique global normalisé à l'occupation du sol. *Ingénieries - E A T,* **47,** 3-15.

Wasson, J.G., Villeneuve, B., Iital, A., Dobiasova, M., Bacikova, S., Pella, H., Mengin, N., Murray-Bligh, J., Timm, H. & Chandesris, A. (2007) Large scale relationships between basin and riparian land cover and

CemOA : archive ouverte d'Irstea / Cemagref

ecological status of European rivers: examples with invertebrate indices from France, Estonia, Slovakia and United Kingdom. *Freshwater Biology,* **(accepté)**.

Wasson, J.G., Villeneuve, B., Iital, A., Dobiasova, M., Bacikova, S., Pella, H., Mengin, N., Murray-Bligh, J., Timm, H. & Chandesris, A. (2010) Large scale relationships between basin and riparian land cover and ecological status of European rivers: examples with invertebrate indices from France, Estonia, Slovakia and United Kingdom. *Freshwater Biology,* **55,** 1465 - 1482.

Waters, T.F. (1995) *Sediment in streams: sources, biological effects and control,* Am. Fish. Soc., Betheda**,** p. 251.

Weak, D.N., Turner, M.G. & Naiman, R.J. (1998) Land cover along an urban-rural gradient: Implications for water quality. *Ecological Applications,* **8,** 6619-6630.

Weber, R.M. & Dunno, G.A. (2001) Riparian vegetation mapping and image processing techniques, Hopi Indian Reservation, Arizona. *Photogrammetric Engineering & Remote Sensing (PE&RS),* **67,** 179-186.

Weller, D.E., Jordan, T.E. & Correll, D.L. (1998) Heuristic models for material discharge from landscapes with riparian buffers. *Ecological Applications,* **8**.

Welsch, D.J. (1991) Riparian forest buffers: function and design for protection and enhancement of water resources. USDA For. Servo Northeastern Area, State, and Private Forestry. Radnor, P A. NA-PR-07-91.

White, E.M. & Knights, B. (1997) Environmental factors affecting the migration of the European eel in the Rivers Severn and Avon, England. *Journal of Fish Biology,* **50,** 1104- 1116.

Whited, D., Stanford, J.A. & Kimball, J.S. (2002) Application of airborne multispectral digital imagery to quantify riverine habitats at different base flows. *River Research and Applications,* **18,** 583-594.

Wiens, J.A. (1989) Spatial scaling in ecology. *Functional Ecology,* **3,** 385-397.

Wiens, J.A. (2002) Riverine landscapes: taking landscape ecology into the water. *Freshwater Biology,* **47,** 501-515.

Williams, R. *REMM: Riparian Ecosystem Management Model,* http://www.tifton.uga.edu/remmwww/, 15 septembre.

Wilson, L.G. (1967) Sediment removal from flood water by grass filtration. *Transactions of the. 4. ASAE,* **10,** 35-37.

Wilzbach, M.A. (1985) Relative rôles of food abundance and cover in determining the habitat distribution of stream Dwelling Cutthroat trout (Salmo clarki). *an. J. Fish. Aquat Sci,* **42,** 1668-1672.

Witte, C., Dowling, R., Weller, D., Denham, R. & Rowland, T. (2001) Quantifying riparian vegetation and stream bank form through the use of airborne laser scanning and digital video data. In: *International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1545-1548.

Wold, H. (1966) Estimation of Principal Components and Related Models by Iterative Least Squares.
analysis In: *Multivariate Analysis.* (Ed^Eds P.R. Krishnaiaah), pp. 391-420. Academic Press, New York.

Wold, H. (1982) Soft modeling: the basic design and some extensions. In: *Systems under indirect observation.* (Ed^Eds K.G. Jö Reskog & H. Wold), pp. 1-54. North-Holland, Amsterdam, New York, Oxfor.

Wold, S., Sjostrom, M. & Eriksson, L. (2001) PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems,* **58,** 109-130.

Woodcock, T.S. & Huryn, A.D. (2007) The response of macroinvertebrate production to a pollution gradient in a headwater stream. *Freshwater Biology,* **52,** 177-196.

Wright, J.F., Furse, M.T. & Armitage, P.D. (1993) RIVPACS: A technique for evaluating the biological quality of rivers in the UK. *European Water Pollution Control,* **3,** 15-25.

Wright, J.F., Furse, M.T. & Moss, D. (1998) River classification using invertebrates: RIVPACS applications. *Aquatic Conservation: Marine and Freshwater Ecosystems,* **8,** 617-631.

Yang, X. (2007) Integrated use of remote sensing and geographic information systems in riparian vegetation delineation and mapping. *International Journal of Remote Sensing,* **28,** 353 - 370.

Young, R.A., Huntrods, T. & Anderson, W. (1980) Effectiveness of vegetated buffer strips in controlling pollution from feedlot runoff. *J. Environ. Qual.,* **9,** 483-487.

Zadeh, L.A. (1965) Fuzzy sets. *Information and Control,* **8,** 338-353.

Zadeh, L.A. (1978) Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems 1***,** 3-28.

Zhan, Q., Molenaar, M., Tempfli, K. & Shi, W. (2005) Quality assessment for geo-spatial objects derived from remotely sensed data. *International Journal of Remote Sensing,* **26,** 2953-2974.

Zhu, Z., Yang, L., Stehman, S.V. & Czaplewski, R.L. (2000) Accuracy Assessment for the U.S. Geological Survey Regional Land-Cover Mapping Program: New York and New Jersey Region. *Photogrammetric Engineering and Remote Sensing,* **66,** 1425-1435.

Zorn, T.G. & Wiley, M.J. (2006) Influence of landscape characteristics on local habitat and fish biomass in streams of Michigan's Lower Peninsula. *Landscape Influences on Stream Habitats and Biological Assemblages,* **48,** 375-393.

CemOA : archive ouverte d'Irstea / Cemagref

LISTE DES ANNEXES

ANNEXE I. Carte des hydro-écorégions de niveau 2 pour la France métropolitaine

ANNEXE II. Questionnaire soumis au DIREN pour la sélection des sites de références

ANNEXE III. Caractéristiques des HER2 de l'HER9-NO

ANNEXE IV. Nomenclatures des données d'occupation

ANNEXE V. Récapitulatif des informations apportées par les données auxiliaires

ANNEXE VI. Présentation synthétique des attributs disponibles dans la plate-forme logicielle orientée objet

ANNEXE VII. Optimisation de la procédure de classification

ANNEXE VIII. Influence de l'échantillonnage d'apprentissage sur les résultats de classification supervisée

ANNEXE IX. Sensibilité des indicateurs aux erreurs de décalage de la BDCarthage®

ANNEXE X. Résultats de l'analyse de la stabilité des chroniques IBGN

ANNEXE XI. Extraction de l'information de l'occupation du sol dans les corridors rivulaires de l'HER9-NO à partir des images à THRS

ANNEXE XII. Résultats zdes 21 modèles pressions/impacts développés sur l'HER9-NO

CemOA : archive ouverte d'Irstea / Cemagref

DOCUMENT ANNEXE

CemOA : archive ouverte d'Irstea / Cemagref

THÈSE

pour l'obtention du grade de
DOCTEUR D'AGROPARISTECH
Institut des Sciences et Industries du Vivant et de l'Environnement
Spécialité : Eaux Continentales et Société

présentée et soutenue publiquement le *6 Décembre 2010* par

Thierry Tormos

ANALYSE À L'ÉCHELLE RÉGIONALE DE L'IMPACT DE L'OCCUPATION DU SOL DANS LES CORRIDORS RIVULAIRES SUR L'ÉTAT ÉCOLOGIQUE DES COURS D'EAU

DOCUMENT ANNEXE

Membres du jury :

M. John THOMPSON	Directeur de recherches, CNRS, Montpellier	Président du jury
M. Yves SOUCHON	Directeur de recherches, Cemagref, Lyon	Rapporteur
Mme Laurence HUBERT-MOY	Professeur d'Université, Rennes 2	Rapporteur
M. Etienne MULLER	Ingénieur de recherches, CNRS, Toulouse	Examinateur
M. Philip ROCHE	Directeur de recherches, Cemagref, Aix-en-	Examinateur
M. Hervé PIEGAY	Directeur de recherches, CNRS, Lyon	Examinateur
M. Pascal KOSUTH	ICPEF, Cemagref, Montpellier	Directeur de Thèse

L'Institut des Sciences et Industries du Vivant et de l'Environnement (Agro Paris Tech) est un Grand Etablissement dépendant du Ministère de l'Agriculture et de la Pêche, composé de l'INA PG, de l'ENGREF et de l'ENSIA (decret n° 2006-1592 du 13 decembre 2006)

CemOA : archive ouverte d'Irstea / Cemagref

CemOA : archive ouverte d'Irstea / Cemagref

SOMMAIRE

ANNEXE I. Carte des hydro-écorégions de niveau 2 pour la France métropolitaine ____ *4*

ANNEXE II. Questionnaire soumis à la DIREN pour la sélection des sites de références ____ *5*

ANNEXE III. Caractéristiques des HER2 de la zone d'étude dédiée à la modélisation (HER9-NO) ____ *6*

ANNEXE IV. Nomenclatures des données d'occupation ____ *7*

ANNEXE IV.1 CORINE Land Cover (CLC) ____ **7**

ANNEXE IV.2 Données d'occupation du sol à THRS sur l'HER9_NO ____ **8**

ANNEXE V. Récapitulatif des informations apportées par les données auxiliaires ____ *11*

ANNEXE VI. Présentation synthétique des attributs disponibles dans la plate-forme logicielle orientée objet *12*

ANNEXE VII. Optimisation de la procédure de classification ____ *14*

ANNEXE VII.1 Sélection de la technique de classification ____ **15**

ANNEXE VII.2 Choix optimisé des sources de données ____ **17**

ANNEXE VII.3 Identification de la taille d'échantillonnage de validation optimale ____ **24**

ANNEXE VIII. Influence de l'échantillonnage d'apprentissage sur les résultats de classification supervisée ____ *26*

ANNEXE VIII.1 Communication acceptée aux IGARSS de 2007 ____ **27**

ANNEXE VIII.2 Analyse à l'échelle d'une région homogène ____ **31**

ANNEXE IX. Sensibilité des indicateurs aux erreurs de décalage de la BDCarthage® ____ *36*

ANNEXE X. Résultats de l'analyse de la stabilité des chroniques IBGN ____ *39*

ANNEXE XI. Extraction de l'information de l'occupation du sol dans les corridors rivulaires de l'HER9-NO à partir des images à THRS ____ *45*

ANNEXE XII. Résultats détaillés des 21 modèles pressions/impacts développés sur l'HER9-NO ____ *53*

ANNEXE XII.1 Modèles mono-échelle ____ **53**

ANNEXE XII.1.1 Modèle BV (Mod-1.1.1) ____ 53

ANNEXE XII.1.2 CRamont-CLC (Mod-1.2.1) ____ 53

ANNEXE XII.1.3 CRamont-THRS (Mod-1.2.2) ____ 54

ANNEXE XII.1.4 CRlocal-fixe-CLC (Mod- 1.3.1) ____ 55

ANNEXE XII.1.5 CRlocal-fixe-THRS (Mod- 1.3.2) ____ 55

ANNEXE XII.1.6 CRlocal-mult-CLC (Mod-1.3.3) ____ 56

ANNEXE XII.1.7 CRlocal-mult-THRS (Mod-1.3.4) ____ 57

ANNEXE XII.2 Modèles bi-échelle ____ **58**

ANNEXE XII.2.1 BV/CRamont-CLC (Mod-2.1.1) ____ 58

ANNEXE XII.2.2 BV/CRamont-THRS (Mod-2.1.2) ____ 58

ANNEXE XII.2.3 BV/CRlocal-fixe-CLC (Mod-2.2.1) ____ 59

ANNEXE XII.2.4 BV/CRlocal-fixe-THRS (Mod-2.2.2) ____ 60

ANNEXE XII.2.5 BV/CRlocal-mult-CLC (Mod-2.2.3) ____ 61

ANNEXE XII.2.6 BV/CRlocal-mult-THRS (Mod-2.2.4) ____ 62

ANNEXE XII.2.7 CRamont /CRlocal-fixe-CLC (Mod-2.3.1) ____ 63

ANNEXE XII.2.8 CRamont /CRlocal-fixe -CLC (Mod-2.3.2) ____ 64

ANNEXE XII.2.9 CRamont /CRlocal-mult-CLC (Mod-2.3.3) ____ 65

ANNEXE XII.2.10 CRamont /CRlocal-mult-THRS (Mod-2.3.4) ____ 66

ANNEXE XII.3 Modèles tri-échelles ____ **68**

ANNEXE XII.3.1 BV/CRamont /CRlocal-fixe-CLC (Mod-3.1.1) ____ 68

ANNEXE XII.3.2 BV/CRamont /CRlocal-fixe-THRS (Mod-3.1.2) ____ 69

ANNEXE XII.3.3 BV/CRamont /CRlocal-mult-CLC (Mod-3.1.3) ____ 70

ANNEXE XII.3.4 BV/CRamont /CRlocal-mult-THRS (Mod- 3.1.4) ____ 71

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE I. Carte des hydro-écorégions de niveau 2 pour la France métropolitaine

L'analyse des informations géographiques disponibles sur le territoire français aboutit à l'identification de 107 unités différentes représentée sur la carte ci-dessus. Le descriptif détaillé des polygones est fourni dans le rapport intitulé « 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 » du laboratoire « Hydro-écologie Quantitative » du Cemagref de Lyon de 2002.

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE II. Questionnaire soumis à la DIREN pour la sélection des sites de références

Le tableau ci-dessous constitue un guide résumant l'ensemble des critères sur lesquels un site de prélèvement (une station) sera jugé comme un site de référence ou non. Ces critères sont détaillés dans le rapport « Détermination des valeurs de référence de l'IBGN et propositions de valeurs limites du « Bon État » du laboratoire « Hydro-écologie Quantitative » du Cemagref de Lyon de 2004.

CARACTERISATION D'UNE STATION IBGN EXISTANTE EN VUE D'EVALUER SA CANDIDATURE AU STATUT DE SITE DE REFERENCE	Cours d'eau : Nom station :	Code Station :		
Echelle				
BASSIN VERSANT	**Occupation du sol dominante sur le BV** à l'amont du site			
	Forêt naturelle, friches, landes, prairies, paturâges …	*Dominante*	*Secondaire*	*Déclassant*
	Culture intensive, vigne	*Absence*	*Secondaire*	*Déclassant*
	Urbanisation	*Absence*	*Secondaire*	*Déclassant*
	Régime Hydrologique			
	Modification des crues	*Absence*	*Non signif.*	*Déclassant*
	Modification des étiages	*Absence*	*Non signif.*	*Déclassant*
	Régime sédimentaire			
	Blocage du flux sédimentaire (signes d'incision du lit)	*Absence*	*Non signif.*	*Déclassant*
	Erosion généralisée des sols (sédimentation dans le lit)	*Absence*	*Non signif.*	*Déclassant*
TRONCON (masse d'eau)	**Rejets polluants**			
	Industriels (et urbains avec industries)	*Absence*	*Non signif.*	*Déclassant*
	Urbains domestiques	*Absence*	*Non signif.*	*Déclassant*
	Rejets 'diffus' (et dispersés)	*Absence*	*Non signif.*	*Déclassant*
	Optionnel (si connue) — *Classe de qualité SEQ-Eau (B/V/J/O/R)*: *MOOX*; *MA*; *Nitrates*; *MP*; *EPRV*			*Moox : matières organiques et oxydables* *MA : matières azotées (hors nitrates)* *MP : matières phosphorées* *EPRV : effet des proliférations végétales*
	Modifications du régime hydrologique			
	Eclusées	*Absence*	*Non signif.*	*Déclassant*
	Transferts	*Absence*	*Non signif.*	*Déclassant*
	Prélèvements	*Absence*	*Non signif.*	*Déclassant*
	Ralentissement des écoulements *(influence seuil ou barrage en % du linéaire du tronçon)*	*≤ 10 %*	*10-30%*	*>30%*
	Modifications 'lourdes' du tracé en plan et profil en long *(en % linéaire)* *(rectification, recalibrage, chenalisation, digues sur les 2 rives, extractions anciennes en lit mineur …)*	*Absence*	*≤ 10 %*	*> 10 %*
	Travaux' légers' de protection et d'entretien *(stabilisation des berges, curages d'entretien)*	*≤ 10 %*	*10-30%*	*>30%*
	Occupation du fond de vallée (corridor fluvial)			
	Forêt naturelle, landes, zones naturelles, prairies …	*Dominante*	*Secondaire*	*Déclassant*
	Culture intensive en fond de vallée	*Non*	*Oui*	
	Si oui, largeur du corridor végétal rivulaire	*Forte* [1]	*Faible* [2]	*Quasi-nulle* [3]
	[1] : *corridor végétal large et quasi-continu,* [2] : *corridor étroit et/ou discontinu,* [3] : *culture arrivant jusque sur la berge*			
STATION (=site) longueur de 10 à 15 fois la largeur du lit mineur	**REPRESENTATIVITE DE LA STATION / TYPE DU TRONCON**	*Très bonne*	*Bonne*	*Mauvaise*
	Rejet connu à l'amont immédiat	*Non*	*Négligeable*	*Déclassant*
	Station dans l'emprise d'un aménagement 'lourd' *(rectification, recalibrage, chenalisation, digues sur les 2 rives, extractions anciennes en lit mineur, ralentissement lié à un seuil ou un barrage …)*	*Non*	*Non signif.*	*Déclassant*
	Effets cumulés de travaux 'légers' (stabilisation des berges, curages d'entretien)	*Non*	*Non signif.*	*Déclassant*
	Végétation des berges	*'Naturelle'* [1]	*Modifiée* [2]	
	[1]: ripisylve typique du type de cours d'eau présente; [2]:ripisylve réduite ou formée d'essences allochtones			
APPRECIATION GLOBALE DE LA STATION PAR L'EXPERT		**Référence**		**Perturbée**

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE III. Caractéristiques des HER2 de la zone d'étude dédiée à la modélisation (HER9-NO)

L'HER Tables-Calcaires (HER9) constituent une région de plaine, d'altitude inférieure à 200 m, couvrant le bassin parisien et le nord du bassin aquitain ; le sous-sol est formé de roches sédimentaires tabulaires en majorité carbonatées.

Cependant, ces roches présentent en superficie des caractéristiques variables de perméabilité induisant des différences dans la densité du réseau de drainage, ou des différences de dureté et de résistance à l'érosion produisant localement des formes de reliefs plus marquées (côtes et collines).

Sur ces différences 13 HER de niveau 2 (HER2) ont été définies. La zone d'étude en couvre 7 : les HER2 Pays de Caen (30) et Cotentin Est (37) qui sont complètement contenu, et les HER2 Pays de Bray (35), Auréole crétacé du Bassin Parisien (38), Nord Loire et Perche (54) et Haute Normandie Picardie (57) recouvertes en partie.

Les caractéristiques de chacune de ces HER2 sont présentées ci-dessous :

Pays de caen (30) : Cette zone se différencie par la perméabilité de son sous sol, attestée par la densité de drainage, les formations de calcaires du jurassique (Bathonien, Bajocien) étant affleurantes dans ce secteur côtier, par ailleurs très plat.

Pays de bray (35) : Cette zone se différencie par la perméabilité de son sous sol, attestée par la densité de drainage, les formations marneuses du jurassique constituant un contraste au milieu de la craie, la différence de densité de drainage justifie une telle différenciation.

Ile de France (36): Cette zone reste très hétérogène autant en perméabilité, qu'en nature de sous sol, et correspond à la zone centrale sur le plan géologique du Bassin parisien (sables, dépôts argilo-sableux, calcaires, marnes, gypses, alluvions quaternaires….).

Cotentin Est (37) : Cette zone se différencie par rapport au massif armoricain et par rapport au secteur plus calcaire et perméable de la région de Caen, avec des terrains assez imperméables (densité de réseau).

Tables Calcaires Auréole crétacé du Bassin Parisien (38) : Ce secteur se différencie de ses voisins par une assez faible densité de drainage, et par un relief tabulaire marqué par quelques côtes. Le relief peu élevé (inférieur à 300 m) avec des pentes marquées au niveau des versants de vallées. Le climat est océanique.

Nord Loire et Perche (54) : Cet ensemble constitué de formations sédimentaires du jurassique et du crétacé, se distingue par une densité de drainage supérieure à l'auréole crétacé du bassin parisien, liée à l'imperméabilité des placages superficiels.

Haute Normandie Picardie (57): Cette partie de l'auréole crétacé du bassin parisien est caractérisée par sa proximité du littoral de la Manche d'une part (induisant un climat à tendance fraîche) et par le drainage de la nappe de la craie dans les vallées incisées où l'essentiel du réseau hydrographique est représenté (eau froide et carbonatée et régime hydrologique soutenu).

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE IV. Nomenclatures des données d'occupation

ANNEXE IV.1 *CORINE Land Cover (CLC)*

Le programme CORINE land cover repose sur une nomenclature standard hiérarchisée à **3 niveaux** et **44 postes** répartis selon **5 grands types d'occupation du territoire** :

1. Territoires artificialisés

1.1. Zones urbanisées

1.1.1. Tissu urbain continu
1.1.2. Tissu urbain discontinu

1.2. Zones industrielles ou commerciales et réseaux de communication

1.2.1. Zones industrielles et commerciales
1.2.2. Réseaux routier et ferroviaire et espaces associés
1.2.3. Zones portuaires
1.2.4. Aéroports

1.3. Mines, décharges et chantiers

1.3.1. Extraction de matériaux
1.3.2. Décharges
1.3.3. Chantiers

1.4. Espaces verts artificialisés, non agricoles

1.4.1. Espaces verts urbains
1.4.2. Equipements sportifs et de loisirs

2. Territoires agricoles

2.1. Terres arables

2.1.1. Terres arables hors périmètres d'irrigation
2.1.2. Périmètres irrigués en permanence
2.1.3. Rizières

2.2. Cultures permanentes

2.2.1. Vignobles
2.2.2. Vergers et petits fruits
2.2.3. Oliveraies

2.3. Prairies

2.3.1. Prairies

2.4. Zones agricoles hétérogènes

2.4.1. Cultures annuelles associées aux cultures permanentes
2.4.2. Systèmes culturaux et parcellaires complexes
2.4.3. Territoires principalement occupés par l'agriculture,avec présence de végétation naturelle importante
2.4.4. Territoires agro-forestiers

3. Forêts et milieux semi-naturels

3.1. Forêts

3.1.1. Forêts de feuillus
3.1.2. Forêts de conifères
3.1.3. Forêts mélangées

3.2. Milieux à végétation arbustive et/ou herbacée

3.2.1. Pelouses et pâturages naturels
3.2.2. Landes et broussailles
3.2.3. Végétation sclérophylle
3.2.4. Forêt et végétation arbustive en mutation

3.3. Espaces ouverts, sans ou avec peu de végétation

3.3.1. Plages, dunes et sable
3.3.2. Roches nues
3.3.3. Végétation clairsemée
3.3.4. Zones incendiées
3.3.5. Glaciers et neiges éternelles

4. Zones humides

4.1. Zones humides intérieures

4.1.1. Marais intérieurs
4.1.2. Tourbières

4.2. Zones humides maritimes

4.2.1. Marais maritimes
4.2.2. Marais salants
4.2.3. Zones intertidales

5. Surfaces en eau

5.1. Eaux continentales

5.1.1. Cours et voies d'eau
5.1.2. Plans d'eau

5.2. Eaux maritimes

5.2.1. Lagunes littorales
5.2.2. Estuaires
5.2.3. Mers et océans

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE IV.2 *Données d'occupation du sol à THRS sur l'HER9 NO*

Le résultat de classification (cf. ANNEXE XI) produit une couche vecteur d'occupation du sol selon une typologie détaillée constituée de 62 postes (classes sémantiques). Ces postes correspondent soit directement à une classe de l'arbre de classification ou à un regroupement de plusieurs de ces classes.

Ces 62 postes ont été regroupés selon leur sémantique à l'aide d'une typologie hiérarchique construite sur quatre niveaux. Le quatrième niveau est défini spécifiquement pour les cultures annuelles issues du RPG (cf. Tableau 5).

Le premier niveau répartit ces postes en fonction de 4 grands types d'occupation du territoire :

- Territoires artificialisés (Tableau 1)
- Territoires agricoles (Tableau 2)
- Forêts et milieux semi-naturels (Tableau 3)
- Surface en eau (Tableau 4)

Dans chacun de ces tableaux on indique également pour les classes du troisième niveau à partir de quelle(s) source(s) de données elles ont été définies, avec a – les données images ; b – CORINE Land Cover ; c – Registre Parcellaire Graphique ; d – BD Routes ; e – BD Carthage.

Tableau 1 : Territoires artificialisés

Premier niveau	*Deuxième niveau*	*Troisième niveau*	*Source*
1. Territoires artificialisés	1.1. Tissu urbain	1.1.1 Tissu urbain résidentiel continu	b
		1.1.2 Tissu urbain résidentiel discontinu et diffus	a
		1.1.3 Végétation urbaine discontinu	a
	1.2. Zones industrielles ou commerciales et réseau de communication	1.2.1 Zones industrielles et commerciales	b
		1.2.2 Réseau routier et ferroviaires et esp. Associés	b
		1.2.3 Aéroports	b
		1.2.4 Zones portuaires	b
	1.3. Mines, Décharges	1.3.1 Décharges	b
		1.3.3 Extraction de matériaux	b
	1.4 Espaces verts artificialisés non agricoles	1.4.1. Espaces verts urbains	b
		1.4.2. Zones de sport et de loisirs	b
	1.5 Routes & Chemins	1.5.1 Autoroutes	d
		1.5.2 Nationales	d
		1.5.3 Départementales	d
		1.5.4 Non Définies (ND)	d
		1.5.5 Chemins	a

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 2 : Territoires agricoles

Premier niveau	*Deuxième niveau*	*Troisième niveau*	*Source*
2. Territoires agricoles	2.1. Cultures annuelles	2.1.1 Labours	a
		2.1.2. Cultures actives	a
		2.1.3. et 2.1.4 RPG annuelles (cf. quatrième niveau)	c
	2.2. Cultures permanentes	2.2.1. Vignes	c
		2.2.2. Arbres fruitiers coques	c
		2.2.3 Vergers	c
	2.3. Pâtures & Prairie	2.3.1 Prairie permanente	c
		2.3.2 Prairie temporaire	c
		2.3.3 Estives, landes	c
		2.3.4 Fourrages	c
	2.4. Zones agricoles hétérogènes	2.4.1 Cultures non définies	a
		2.4.2 Exploitation RPG complexes	c
	2.5 Gels des terres	2.5.1 Gels	c
		2.5.2 Gels industriels	c
		2.5.3 Autres gels	c

Tableau 3 : Forêts et milieux semi-naturels

Premier niveau	*Deuxième niveau*	*Troisième niveau*	*Source*
3. Forêts et milieux semi-naturels	3.1. Forêts	3.1.1 Zones arborées	a
	3.2. Milieux à végétation arbustive et/ou herbacée (Landes, Garrigues, Broussailles)	3.2.1 Zones arbustive et/ou herbacée	a
	3.3. Espaces ouverts avec ou peu de végétation associée	3.3.1 Sols nues et avec peu de végétation	a

Tableau 4 : Surfaces en eau

Premier niveau	*Deuxième niveau*	*Troisième niveau*	*Source*
4. Surfaces en eau	4.1 Eaux continentales	4.1.1 Bassins portuaires maritimes	e
		4.1.2 plans d'eau, bassins et réservoirs	e
		4.1.3 Zones sableuses	e
		4.1.4 Traitements des eaux	e
		4.1.5 Zones couvertes d'eau	e
		4.1.6 Zones temporaires	e
		4.1.7 Cours d'eau	a
		4.1.8 Ensemble de petits plans d'eau	e
	4.2 Eaux maritimes	4.2.1 Bassin portuaires maritimes	e
		4.2.2 Ecoulement eau salée	e
		4.2.3 Marais maritimes	e
		4.2.4 Océans	e
		4.2.5 Zones de graviers maritimes	e
		4.2.6 Zones rocheuses et sableuses maritimes	e
		4.2.7 Zones rocheuses maritimes	e
		4.2.8 Zones sableuses maritimes	e
		4.2.9 Zones vaseuses	e

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 5: Quatrième niveau (RPG annuelles)

Quatrième niveau (RPG annuelles)
2.1.3.0 Autres céréales
2.1.3.1 Autres cultures industrielles
2.1.3.2 Autres oléagineux
2.1.3.3 Blé tendre
2.1.3.4 Colza
2.1.3.5 Divers
2.1.3.6 Légumes fleurs
2.1.3.7 Maïs
2.1.3.8 Orge
2.1.3.9 Plantes à fibres
2.1.4.0 Protéagineux
2.1.4.1 Semences
2.1.4.2 Tournesol

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE V. Récapitulatif des informations apportées par les données auxiliaires

Quatre sources de données exogènes disponibles à coût réduit à l'échelle du territoire national ont été sélectionnées : CORINE Land Cover (CLC), les bases de données IGN Route 500® (BDR) et BDCarthage® (BDC), et le Registre Parcellaire Graphique (RPG). Les informations apportées par chacune de ces sources sont résumées dans le tableau ci-dessous.

BDR	11	Autoroute
	12	Nationale
	13	Départementale
	14	ND
CLC	21	Tissu urbain continu
	22	Zone Ind. & Com.
	23	Res. routiers, Gares
	24	Zones portuaires
	25	Aéroports
	26	Extarction de matériaux
	27	Décharges
	28	Espaces verts urbains
	29	Equ. Sport. & Lois.
RPG	41	BLE TENDRE
	42	MAIS GRAIN ET ENSILAGE
	43	ORGE
	44	AUTRES CEREALES
	45	COLZA
	46	TOURNESOL
	47	AUTRES OLEAGINEUX
	48	PROTEAGINEUX
	49	PLANTES A FIBRES
	51	SEMENCES
	52	GEL (SURFACES GELEES SANS PRODUCTION)
	53	GEL INDUSTRIEL
	54	AUTRES GELS
	55	RIZ
	56	LEGUMINEUSES A GRAINS
	57	FOURRAGE
	58	ESTIVES LANDES
	59	PRAIRIES PERMANENTES
	61	PRAIRIES TEMPORAIRES
	62	VERGERS
	63	VIGNES
	64	FRUITS A COQUE
RPG	65	OLIVIERS
	66	AUTRES CULTURES INDUSTRIELLES
	67	LEGUMES-FLEURS
	68	CANNE A SUCRE
	69	ARBORICULTURE
	71	DIVERS
BDC	84	Mers et océans
	85	Marais maritimes
	86	Plan d'eau, bassin réservoir
	87	Ensemble de petits plans d'eau
	88	Traitements des eaux, station de pompage
	89	Bassin portuaire fluvial
	91	Ecoulement d'eau salée
	92	Bassin portuaire ES
	93	Zone rocheuse ES
	94	Zone mixtes rochers et sables ES
	95	Zone de sable humide ES
	96	Zone de vase ES
	97	Zone de graviers et des galets ES
	98	Autre zone couverte d'eau
	99	Zone temporairement recouverte d'eau douce
	15	Sable et graviers dans le lit d'un cours d'eau

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE VI. Présentation synthétique des attributs disponibles dans la plate-forme logicielle orientée objet

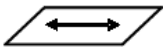
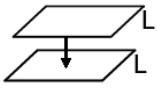
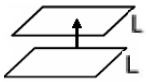
Avec l'approche orientée objet, on dispose de nombreux attributs de classification. En effet, un grand nombre de nouveaux attributs, au-delà des seules valeurs spectrales, peuvent être employés pour décrire les objets de l'image, tels que la forme, la texture ou le contexte. Les objets résultants sont organisés dans un réseau hiérarchique, ce qui rend possible la description d'un objet par les attributs d'autres objets extraits au même et à différents niveaux d'organisation. Ainsi, *Definiens Developper 7.0* ® fournit un certain nombre d'attributs qui peuvent être utilisés dans les processus de classification (Figure 25).

Les attributs d'objets (object features) relatifs aux caractéristiques radiométriques (moyenne, brillance, ratio…), de forme (superficie, longueur, largeur…), de texture (texture Haralick notamment) et de hiérarchie de l'objet sont présentés dans le tableau ci-dessous.

<table>
<tr><td>Critères radiométriques</td><td>Critères de forme</td></tr>
<tr><td>
Object features
Customized
Layer Values
Mean
Brightness
MIR
Max. diff.
PIR
Panchro
Rouge
Vert
Standard deviation
Pixel-based
To neighbors
To super-object
To Scene
Hue, Saturation, Intensity
</td><td>
Object features
Customized
Layer Values
Shape
Texture
Variables
Create new 'Object Variable'
Hierarchy
Level
Number of higher levels
Number of neighbors
Number of sub-objects
Number of sublevels
Thematic attributes
</td></tr>
<tr><td>Critères de texture</td><td>Critère de hiérarchie</td></tr>
<tr><td>
Object features
Customized
Layer Values
Shape
Texture
Layer value texture based on sub-objects
Shape texture based on sub-objects
Texture after Haralick
GLCM Homogeneity
GLCM Contrast
GLCM Dissimilarity
GLCM Entropy
GLCM Ang. 2nd moment
GLCM Mean
GLCM StdDev
GLCM Correlation
GLDV Ang. 2nd moment
GLDV Entropy
GLDV Mean
GLDV Contrast
GLCM Homogeneity (quick 8/11)
GLCM Contrast (quick 8/11)
GLCM Dissimilarity (quick 8/11)
GLCM Entropy (quick 8/11)
GLCM Ang. 2nd moment (quick 8/11)
GLCM Mean (quick 8/11)
GLCM StdDev (quick 8/11)
GLCM Correlation (quick 8/11)
GLDV Ang. 2nd moment (quick 8/11)
GLDV Entropy (quick 8/11)
GLDV Mean (quick 8/11)
GLDV Contrast (quick 8/11)
</td><td>
Object features
Customized
Layer Values
Shape
Generic
Area
Asymmetry
Border index
Border length
Compactness
Density
Elliptic Fit
Elliptic Fit (legacy f
Length
Length/Width
Main direction
Radius of largest e
Radius of largest e
Radius of smallest
Radius of smallest
Rectangular Fit
Rectangular Fit (leg
Roundness
Roundness (legacy
Shape index
Width
Position
To super-object
Based on Polygons
Based on Skeletons
</td></tr>
</table>

CemOA : archive ouverte d'Irstea / Cemagref

Puis, les attributs liés aux classes relatifs au contexte de l'objet. Ces attributs se rapportent à la classification d'autres objets de l'image situés à n'importe quel endroit de la hiérarchie des objets de l'image. Cet endroit peut être défini par une distance verticale dans la hiérarchie des objets de l'image (vers des super-objets ou des sous-objets) ou par une distance horizontale (vers des objets voisins), cf. tableau ci-dessous.

Relation aux objets voisins 	Ces attributs sont employés pour décrire un objet de l'image par ses rapports mutuels avec d'autres objets de l'image, assignés à une classe donnée au même niveau (distance horizontale).
Relation aux sous-objets 	Ces attributs sont utilisés pour décrire un objet de l'image par ses relations à d'autres objets d'image, assignés à une classe donnée à un niveau inférieur (distance verticale). Puisque la résolution augmente lorsqu'on descend dans la hiérarchie des objets de l'image, on peut prendre en compte des informations à des échelles inférieures en utilisant ces attributs.
Relation aux super-objets 	Ces attributs sont utilisés pour décrire un objet de l'image par ses relations à d'autres objets d'image, assignés à une classe donné à un niveau supérieur (distance verticale). Comme avec les relations aux sous-objets, il est possible ici de prendre en compte l'information à une échelle supérieure.

Au-delà de ces critères proposés par le logiciel, des indices spécifiques peuvent être calculés. Notamment les indices spectraux comme les indices de végétation (NDVI), les indices la Brillance, les ratios, etc…

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE VII. Optimisation de la procédure de classification

Dans cette annexe sont présentés les protocoles d'analyses et les résultats obtenus lors de la démarche d'optimisation de la procédure de classification. Cette démarche cherche à optimiser trois étapes:

- L'étape 1 (Sélection des données) afin d'identifier les sources de données images (ou la combinaison de ces sources) permettant d'obtenir les meilleurs résultats de classification parmi les images optiques à THRS disponibles sur le territoire français.

- L'étape 5 (Classification de chaque région) afin de sélectionner la technique de classification orientée objet (« expertisée » vs. « supervisée ») adaptée pour extraire efficacement l'occupation du sol dans les corridors rivulaires sur de grands territoires (en termes de fiabilité et d'applicabilité).

- L'étape 7 (Validation du résultat de classification) afin d'identifier le nombre de parcelles de validation nécessaire pour évaluer de façon fiable la qualité du résultat de classification sur l'ensemble de la zone d'étude à partir des matrices de confusion.

Cette démarche itérative (étant donné le lien existant entre ces trois étapes) s'intéresse à optimiser :
1- l'étape 5 (sélection de la technique de classification, cf. ANNEXE VII.1)
2- l'étape 1 (choix optimisé des sources de données, cf. ANNEXE VII.2)
3- l'étape 7 (identification de l'échantillonnage optimal de validation, cf. ANNEXE VII.3)

Afin d'optimiser ces différentes étapes, des choix ont été faits sur les matériels, les méthodes et la zone d'analyse pour chacune des étapes à optimiser. Ces choix sont présentés dans le tableau 1.

Tableau 1 : Présentation des choix sur les matériels, méthodes et la zone d'analyse utilisés pour optimiser tout d'abord l'étape 5, ensuite l'étape 1 et enfin l'étape 7. DA = Données auxiliaires.

	Technique de classification (étape 5)	Sources de données (étape 1)	échantillonnage de validation (étape 7)	Zone d'analyse
1- Optimisation étape 5	Choix à optimiser	Orthophotos + Spot5 XS + DA	50 parcelles/classe numérisées et vérifiées par un photo-interprète indépendant	Région homogène du bassin versant de l'Hérault HER6-sud
2- Optimisation étape 1	Choix optimisé de l'étape 5	Choix à optimiser	50 parcelles/classe numérisées et vérifiées par un photo-interprète indépendant	Région homogène du bassin versant de l'Hérault HER6-sud
3- Optimisation étape 7	Choix optimisé de l'étape 5	Choix optimisé de l'étape 1	Choix à optimiser	Bassin versant de l'Hérault

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE VII.1 *Sélection de la technique de classification*

Un protocole d'analyse a été mis en place sur une région homogène de l'Hérault (HER6-Sud, 350 Km de réseau hydrographique) afin de sélectionner la technique de classification la plus facile à mettre en œuvre et la plus fiable à la phase 3 de la démarche orientée objet développée (cf. chapitre 5 partie III.5.5).

On rappelle que deux techniques de classification existent dans le logiciel :

- La technique « expertisée » : l'opérateur définit les règles de classification à l'aide de fonctions d'appartenances floues ou non floues.
- La technique « supervisée » : ces règles sont définies automatiquement à l'aide de l'algorithme du Plus Proche Voisin (PPV) à partir d'un échantillonnage d'apprentissage (sélection d'un nombre de parcelles d'entraînement par classe). L'espace d'attributs est défini soit par l'opérateur ou soit automatiquement en utilisant un algorithme d'optimisation de variables.

Ce protocole utilise l'information combinée de l'orthophoto, de Spot5 XS et des données auxiliaires. A partir de ces sources de données, la démarche orientée objet développée a été appliqué jusqu'à la phase 2. A la phase 3, deux méthodes ont été envisagés :

- Méthode 1 : Une classification par technique « expertisée » en continuant à décrire les classes définies dans l'arbre de classification en utilisant des fonctions d'appartenance floues (L'arbre et les critères de classification définis sur cette zone d'analyse sont présentés en ANNEXE VII.2 Figure 2).

- Méthode 2 : Une classification « supervisée » en appliquant sur ce niveau de segmentation (niveau « macro ») un algorithme de PPV afin de différencier directement 6 classes :
    - Sols artificialisés
    - Terres agricoles en végétation
    - Terres agricoles nues (labours)
    - Vignes
    - Végétation herbacée
    - Végétation arbustive semi-naturelle

  Pour chacune de ces classes 50 parcelles d'apprentissage (taille d'échantillonnage optimale pour entraîner un algorithme de classification supervisée sur une région homogène, cf. ANNEXE VIII) ont été sélectionnées de façon à être réparties de manière homogène sur la zone d'étude. L'espace d'attributs a été défini par expérience. Il contient 22 critères (cf. Figure 1).

Pour comparer la filialité du résultat de classification obtenu à partir de ces deux méthodes, des matrices de confusion ont été construites selon la typologie exigée :

- Cultures
- Végétation arborée
- Végétation herbacée
- Sols nus
- Surfaces en eau

Ces matrices sont construites à partir d'un échantillonnage de validation constitué de 50 parcelles par classe qui ont été numérisées (manuellement) par photo-interprétation (photo-interprète indépendant).

CemOA : archive ouverte d'Irstea / Cemagref

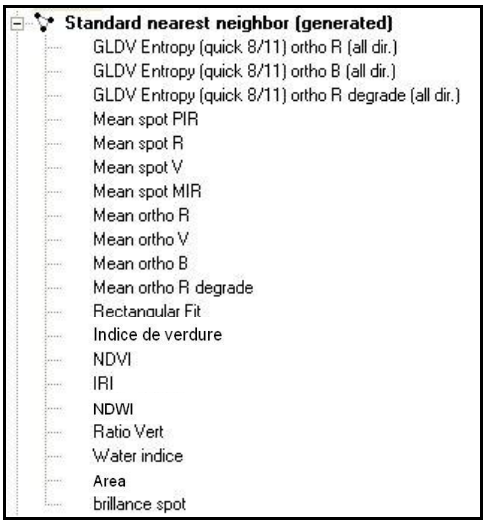


Figure 1 : présentation des 22 critères utilisés pour la classification supervisée par PPV. Ils ont été choisis par expérience.

Le tableau 2 présente le pourcentage de bien classé (précision totale) et l'indice de Kappa obtenus à partir de ces deux méthodes.

La méthode 1 donne de meilleurs résultats par rapport à la méthode 2 (un gain de 10 % environ quelque soit l'estimateur global).

Tableau 2 : Estimateurs de qualité globaux issus des matrices de confusions construites pour valider le résultat de classification de la méthode 1 et la méthode 2.

Méthode de classification utilisée à la phase 3	**Qualité du résultat de classification**	
	Précision totale	Indice de Kappa
Méthode 1 : Technique « expertisée »	92.82 %	90.46 %
Méthode 2 : Technique « supervisée »	80.26 %	73.84 %

En plus d'être moins fiable, la méthode 2 (utilisant la technique « supervisée ») s'est avérée très difficile à mettre en œuvre comparée à la méthode 1 (utilisant la « technique expertisée »). Alors que cette dernière peut être distribuée (sur un serveur de calcul), la méthode 1 impose de travailler sur l'ensemble de la zone étudiée ce qui ralentit considérablement les temps de traitements. A titre de comparaison (uniquement sur le temps de calcul) le résultat de classification avec la méthode 1 a été obtenu en 5 h 41 min et 34 s, le résultat avec la méthode 2 a été obtenu au bout de 2 jours 15 h 50 min et 5 s (soit 10 fois plus de temps de calcul environ).

Même si la méthode 2 est intéressante car elle est plus appropriable par un nouvel l'opérateur (elle nécessite moins d'expérience car les règles sont définies automatiquement), cette difficulté limite son utilisation sur de grands territoires.

On utilisera donc la méthode 1 (la technique « expertisée ») pour classer l'occupation du sol à la phase 3 de la démarche orientée objet développée.

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE VII.2 *Choix optimisé des sources de données*

Un protocole d'analyses a été mis en place afin de sélectionner les sources de données images optimales pour classer l'occupation du sol le long des cours d'eau sur de grands territoires parmi les sources présélectionnées sur le territoire français. Trois combinaisons (C1, C2 et C3), contenant au minimum l'information dans l'infrarouge proche du satellite Spot 5 XS et l'information thématique des données auxiliaires, ont été explorées :

- C1 : Spot5 XS + les données auxiliaires
- C2 : Spot5 XS + Spot PAN + les données auxiliaires
- C3 : Spot5 XS + orthophotos + les données auxiliaires

A partir de ces trois combinaisons de sources de données, l'occupation du sol a été classé sur une région homogène de l'Hérault (Her6-Sud, 350 Km de réseau hydrographique) selon une même base de traitements de classification orientée objet (démarche descendante utilisant les fonctions d'appartenance) et de validation.

Le tableau 3 présente les différents paramètres utilisés pour créer les niveaux de segmentation utilisant l'algorithme de multi-résolution (niveau « thématique », niveau « micro », niveau « macro ») pour chacune des classifications utilisant ces trois combinaisons. La hiérarchie de classes et les critères de classification sont présentés Figure 2 pour C1, Figure 3 pour C2, Figure 4 pour C3.

Les classifications ont été validées en construisant des matrices de confusion démontrant leur fiabilité :

(1) à obtenir la typologie exigée en six catégories :
- Sols artificialisés
- Cultures
- Végétation arborée
- Végétation herbacée
- Sols nus
- Surfaces en eau

(2) à extraire trois largeurs de végétation rivulaire arborée :
- < 15 m
- entre 15 et 30 m
- > 30 m

L'échantillonnage de validation collecté permet d'analyser aussi bien la qualité sémantique que géométrique. En effet, chacune des parcelles a été numérisée manuellement (à l'échelle du 1 : 1000) et photo-interprétées par un opérateur indépendant.

Les matrices de confusion selon la typologie exigée sont présentées dans le tableau 4 pour la classification utilisant C1, le tableau 5 pour la classification utilisant C2 et le tableau 6 pour la classification utilisant C3.

Les résultats du pourcentage de végétation rivulaire arborée (selon les trois largeurs de végétation définies) correctement détecté par la classification en « zone de végétation arborée » sont présentés tableau 7.

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 3: Paramètres de l'algorithme de segmentation multi-résolution utilisé lors de l'application de la démarche orientée objet pour chacune des combinaisons de sources de données

		Paramètre de la segmentation multi-résolution				
	Niveau	***scale***	***shape***	***comp.***	***Bandes spectrales***	***plan thématique***
C1	« thématique »	10	0	0	aucune	oui
	« micro »	10	0.1	0.5	Spot5 XS	non
	« macro »	15	0.1	0.8	Spot5 XS	non
C2	« thématique »	10	0	0	aucune	oui
	« micro »	10	0.1	0.5	Spot5 PAN (poids de 2) PIR Spot5 XS	non
	« macro »	60	0.1	0.8	Spot5 XS	non
C3	« thématique »	10	0	0	aucune	oui
	« micro »	100	0.1	0.5	orthophotos (poids de 2) PIR Spot5 XS	non
	« macro »	300	0.1	0.8	Spot5 XS	non

CemOA : archive ouverte d'Irstea / Cemagref

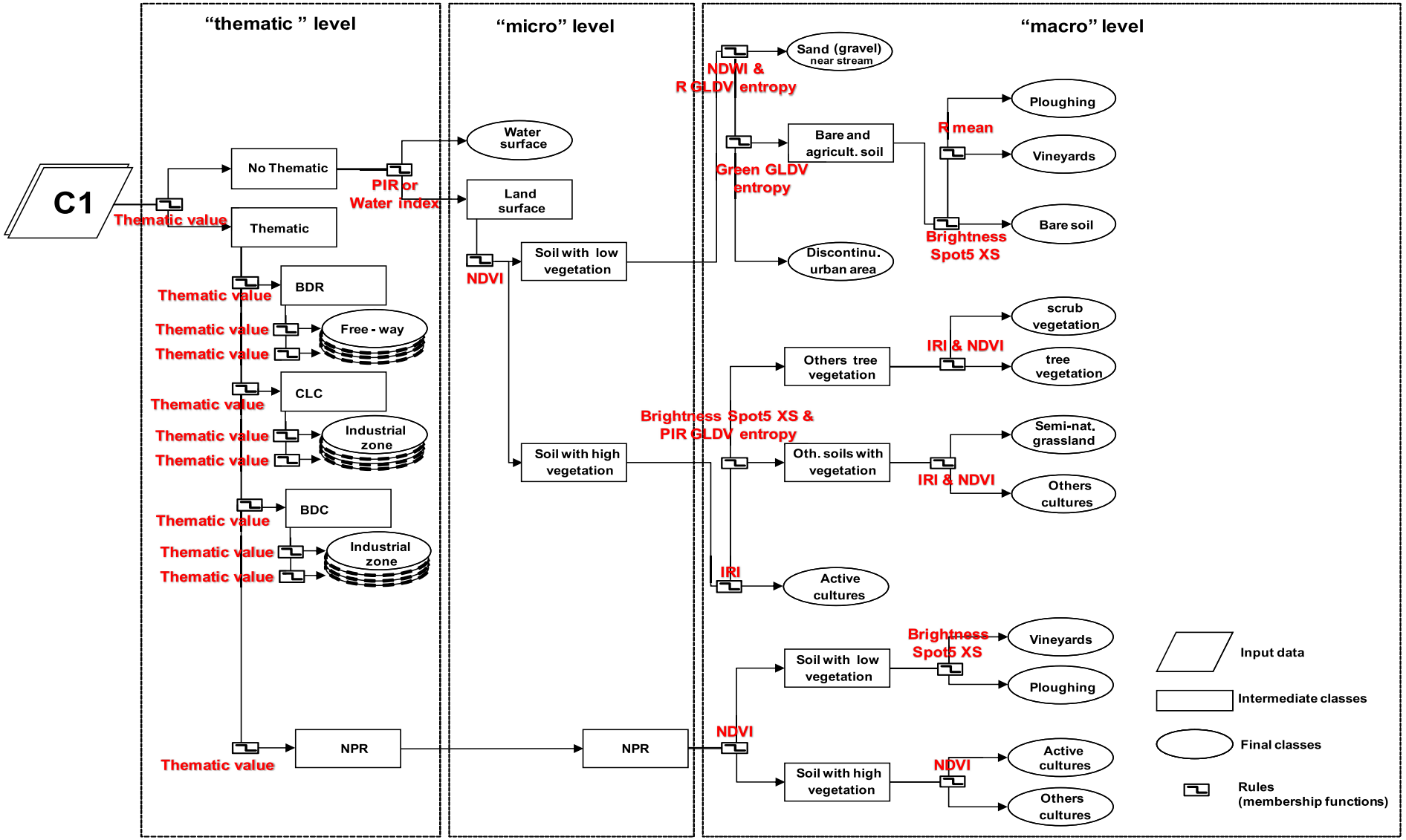


Figure 2 : Arbre et critères de classification de la démarche orientée objet utilisant C1

CemOA : archive ouverte d'Irstea / Cemagref

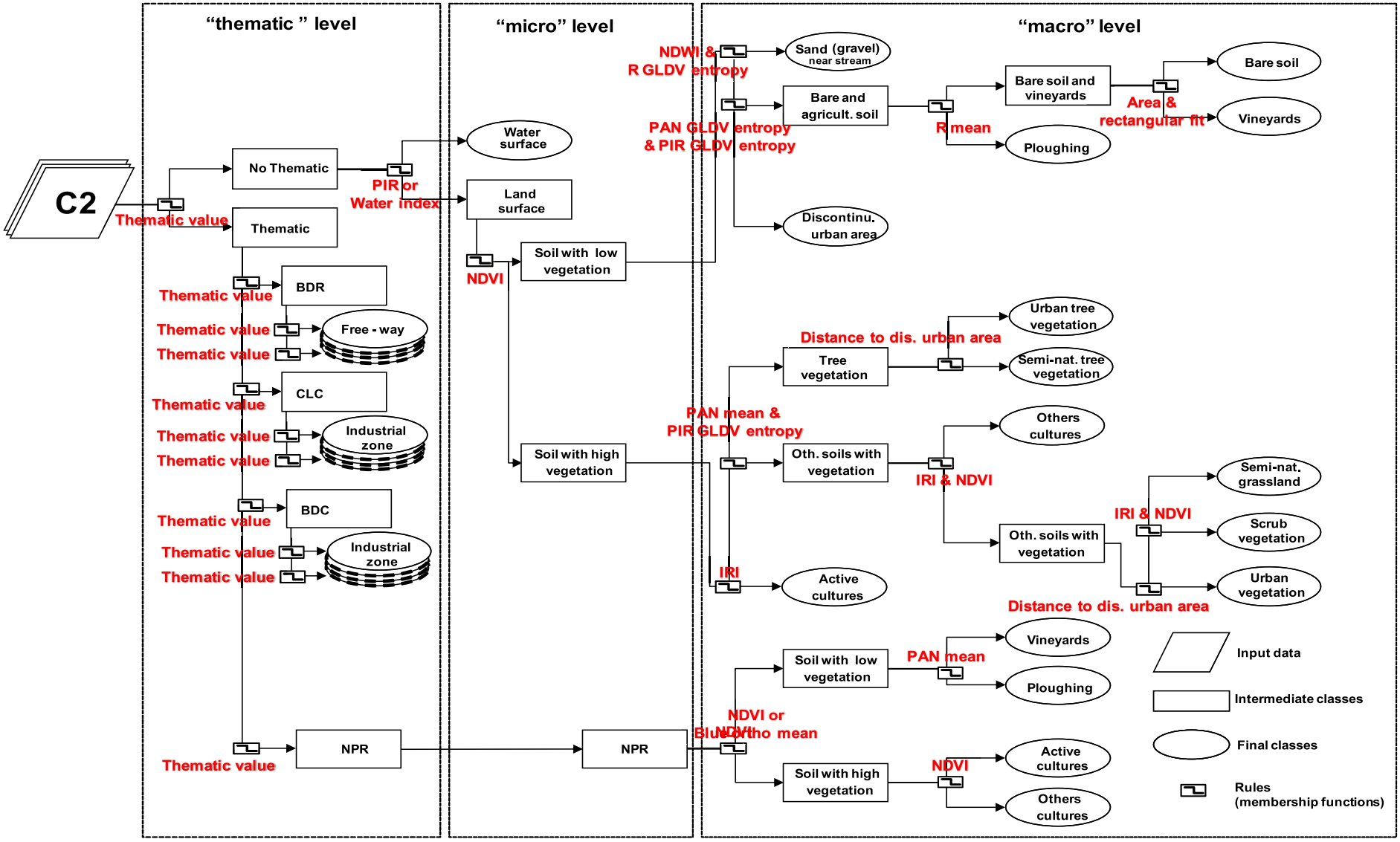


Figure 3 : Arbre et critères de classification de la démarche orientée objet utilisant C2

CemOA : archive ouverte d'Irstea / Cemagref

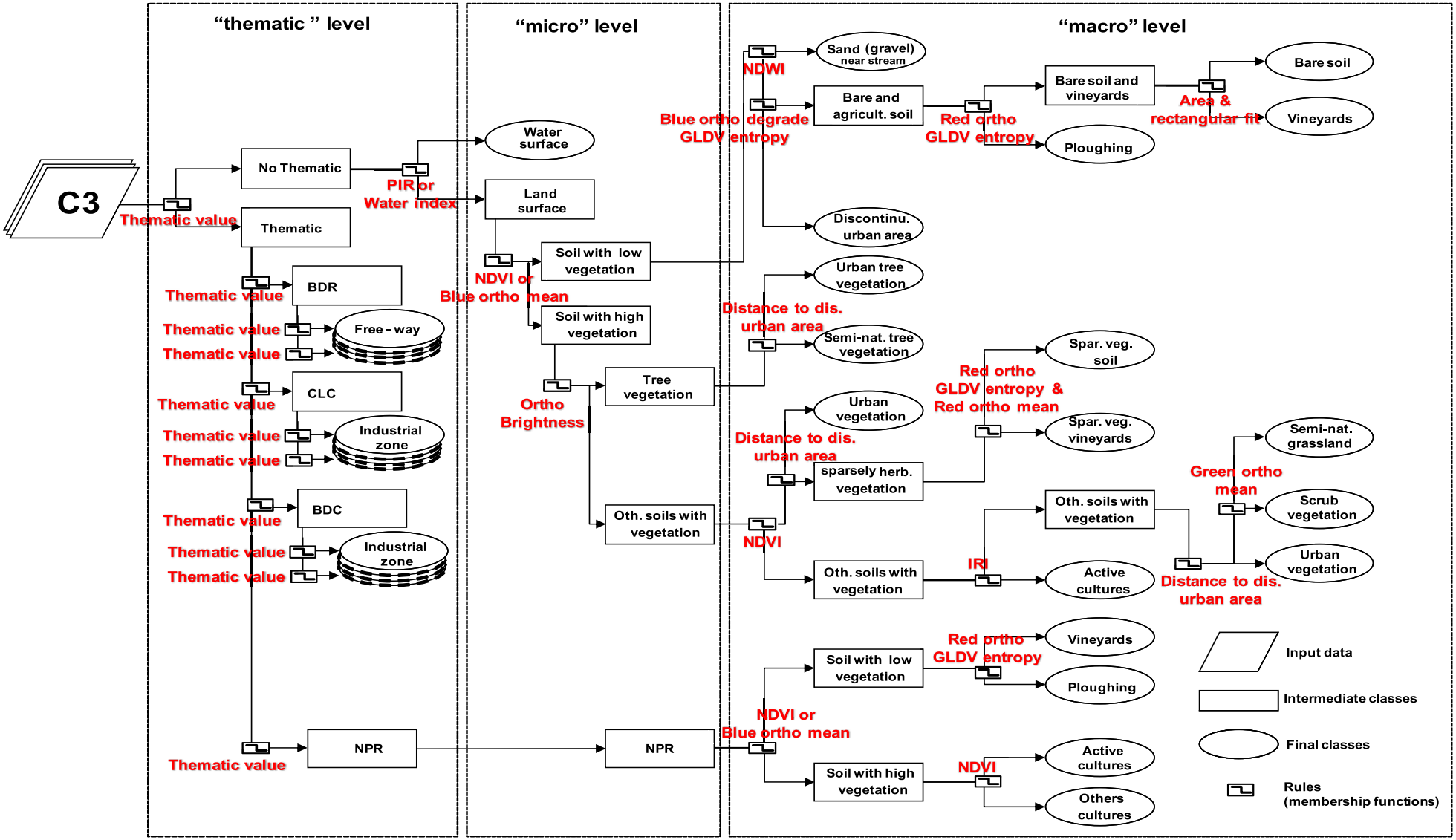


Figure 4 : Arbre et critères de classification de la démarche orientée objet utilisant C3

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 4 : Matrice de confusion résultant de la classification utilisant C1 selon la typologie exigée

		Classification						
		Sols artificialisés	Cultures	Végétation arborée	Végétation herb. et arb. nat.	Sols nus nat.	Surfaces en eau	Total
Référence	Sols artificialisés	231 306	11 072	1 237	8 313	108 731	0	360 659
	Cultures	74 972	3 756 262	32 810	128 549	0	0	3 992 593
	Végétation arborée	9 069	139 061	1 642 848	266 532	2 200	75 916	2 135 626
	Végétation herb. et arb. nat.	26 805	175 114	502 496	1 157 297	1 172	0	1 862 884
	Sols nus nat.	20 440	74 561	9 658	56 525	230 879	78 044	470 107
	Surfaces en eau	2 911	6 975	57 612	29 159	21 921	1 592 892	1 711 470
	Total	365 503	4 136 045	2 246 661	1 646 375	364 903	1 746 852	10 533 339

Précision utilisateur	63,28%	90,23%	73,12%	70,29%	63,27%	91,19%
Précision producteur	64,13%	94,08%	76,93%	62,12%	49,11%	93,07%
Précision globale	81,75%					
Indice de Kappa	75,48%					

Tableau 5 : Matrice de confusion résultant de la classification utilisant C2 selon la typologie exigée

		Classification						
		Sols artificialisés	Cultures	Végétation arborée	Végétation herb. et arb. nat.	Sols nus nat.	Surfaces en eau	Total
Référence	Sols artificialisés	242 376	7 172	3 120	4 843	100 996	2164	360 671
	Cultures	36 323	3 829 220	32 992	67 644	27020	0	3 993 199
	Végétation arborée	14 794	182 373	1 270 481	590 776	6 367	71 263	2 136 054
	Végétation herb. et arb. nat.	17 512	265 840	733 445	842 061	4 588	0	1 863 446
	Sols nus nat.	5 792	835	3 620	61 539	366 008	32 385	470 179
	Surfaces en eau	2 980	25 646	26 172	0	85 940	1 570 732	1 711 470
	Total	319 777	4 311 086	2 069 830	1 566 863	590 919	1 676 544	10 535 019

Précision utilisateur	75,80%	88,82%	61,38%	53,74%	61,94%	93,69%
Précision producteur	67,20%	95,89%	59,48%	45,19%	77,84%	91,78%
Précision globale	77,08%					
Indice de Kappa	69,25%					

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 6 : Matrice de confusion résultant de la classification utilisant C3 selon la typologie exigée

		Classification						
		Sols artificialisés	Cultures	Végétation arborée	Végétation herb. et arb. nat.	Sols nus nat.	Surfaces en eau	Total
Référence	Sols artificialisés	277 481	0	5 422	4 822	72 813	0	360 538
	Cultures	3 112	3 764 564	3 055	112 273	109 875	6	3 992 885
	Végétation arborée	28	15 226	1 930 112	116 082	1 763	72 524	2 135 735
	Végétation herb. et arb. nat.	8 382	13 638	16 320	1 822 825	2 100	359	1 863 624
	Sols nus nat.	1 568	0	113	57 817	393 724	16 465	469 687
	Surfaces en eau	33	464	109 938	11 653	95	1 588 678	1 710 861
	Total	290 604	3 793 892	2 064 960	2 125 472	580 370	1 678 032	10 533 330

Précision utilisateur	95,48%	99,23%	93,47%	85,76%	67,84%	94,68%	
Précision producteur	76,96%	94,28%	90,37%	97,81%	83,83%	92,86%	
Précision globale	92,82%						
Indice de Kappa	90,46%						

Tableau 7 : Pourcentage de végétation arborée rivulaire correctement détecté pour chacune des trois largeurs de végétation à partir des trois combinaisons (C1, C2, C3) de sources de données différentes (DA = Données Auxiliaires).

	C1 : Spot 5 XS + DA	**C2 : Spot 5 XS + Spot PAN + DA**	**C3 : Spot 5 XS + orthophotos + DA**
< 15 m	23,09%	61,83%	73,10%
entre 15 et 30 m	75,70%	83,19%	82,61%
> 30 m	86,11%	80,24%	83,63%

D'après ces résultats, la combinaison C3 permet d'obtenir les meilleurs résultats de la classification si l'on compare les différents estimateurs de qualité obtenus selon la typologie exigée et les trois classes de largeurs de végétation rivulaire arborée.

On utilisera donc la combinaison C3 : les orthophotos, les images satellites Spot5 XS et les données auxiliaires pour extraire l'occupation du sol dans les corridors rivulaires.

La classification utilisant C1, moins performante, est limitée notamment pour détecter et classer correctement les « sols artificialisés » (la précision pour l'utilisateur est de 63.27 %). L'information à 10 m de résolution spatiale contenue dans Spot5 XS n'est pas adaptée pour délimiter correctement les bâtiments diffus et n'est pas assez texturée pour séparer la classe « sols artificialisés » de la classe « sols nus naturels » et « Cultures » (terres labourées et vignes). Du fait de cette résolution, les bandes fines (largeur < 15 m) de végétation rivulaire arborée sont faiblement détectées (seulement 23.09 %).

La classification utilisant C2 qui profite d'une information plus précise (2.5 m) issue de Spot5 PAN améliore la qualité de la classe « sol artificialisés » (75.80 %). Cependant, cette information dégrade la qualité des classes « végétation arborée » (61.38 %) et « végétation herbacée et arbustives »

CemOA : archive ouverte d'Irstea / Cemagref

(53.74%) car les objets contigus appartenant à l'une de ces deux classes ne sont pas segmentés correctement.

Par contre la classification utilisant C3 qui apporte une information encore plus précise (0.5 m) issue des orthophotos améliore la caractérisation de la classe « sols artificialisés » (95.48 %) mais permet aussi une bonne différenciation de la classe « végétation arborée » (93.47 %) et « végétation herbacée et arbustives » (85.76%).

Cependant, on observe que quelque soit la combinaison des sources de données utilisées des difficultés pour extraire correctement la classe « sols nus naturels ». Cette classe possède un comportement similaire à la fois aux classes « sols artificialisés » et « Cultures » (terres labourées et vignes).

ANNEXE VII.3 *Identification de la taille d'échantillonnage de validation optimale*

Un protocole d'analyse a été mis en place sur l'ensemble du bassin versant de l'Hérault pour identifier la taille d'échantillonnage (nombre de parcelles de validation par classe) permettant d'obtenir des résultats de validation robustes à partir des matrices de confusions.

Pour cela, un jeu de 300 parcelles de contrôle par classe, réparti sur la zone d'étude de façon homogène, a été sélectionné aléatoirement pour valider le résultat de classification obtenu selon la démarche orientée objet utilisant les fonctions d'appartenances et les sources de données optimales (orthophotos, les images satellites Spot5 XS et les données auxiliaires).

La typologie retenue pour cette analyse contient 9 classes :

- Sols artificialisés
- Terres agricoles en végétation
- Terres agricoles nues (labours)
- Vignes
- Végétation arborée
- Végétation herbacée
- Végétation arbustive semi-naturelle
- Sols nus naturels
- Surfaces en eau

A partir de ce jeu de parcelles, plusieurs tailles d'échantillonnage (nombres de point de contrôle par classe) on été testés (de 10 parcelles/classe jusqu'à 100 parcelles/classe au pas de 10) pour construire les matrices de confusions. Ces points de contrôle ont été sélectionnés de manière aléatoire et répartie sur la zone d'étude de façon homogène. Ce protocole a été itéré 100 fois.

Le tableau 8 présente la moyenne, l'écart-type, le maximum et le minimum de pourcentage de bien classé (précision totale) obtenus pour les différents nombres de parcelles de validation (PV) par classe testées.

On constate d'après ce tableau que quelque soit le nombre de parcelles par classe, on obtient en une précision globale moyenne similaire (84 % environ).

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 8 : Statistiques du pourcentage de bien classé (moyenne, écart-type, minimum, maximum) obtenues pour différentes tailles d'échantillonnage de validation. PV = Parcelles de Validation.

Nombre de PV / classe	Pourcentage de bien classé			
	Moyenne	Ecart type	Maximum	Minimum
10	83.541	7.6552	94.544	55.508
20	84.106	5.2206	93.029	68.551
30	84.224	3.5482	92.142	72.374
40	84.301	3.2238	91.398	74.516
50	83.874	2.5382	89.644	76.282
60	84.411	2.4616	89.649	78.648
70	84.232	2.4792	90.149	76.901
80	84.699	1.8095	90.141	80.284
90	84.19	1.7671	88.542	80.929
100	84.664	1.6938	88.844	79.622

Par contre (Figure 8), plus on augmente le nombre de parcelle, plus le résultat de classification est robuste (faible écart-type). La robustesse de l'évaluation de la qualité du résultat se stabilise environ lorsqu'on utilise une taille d'échantillon de 50 parcelles de validation par classe.

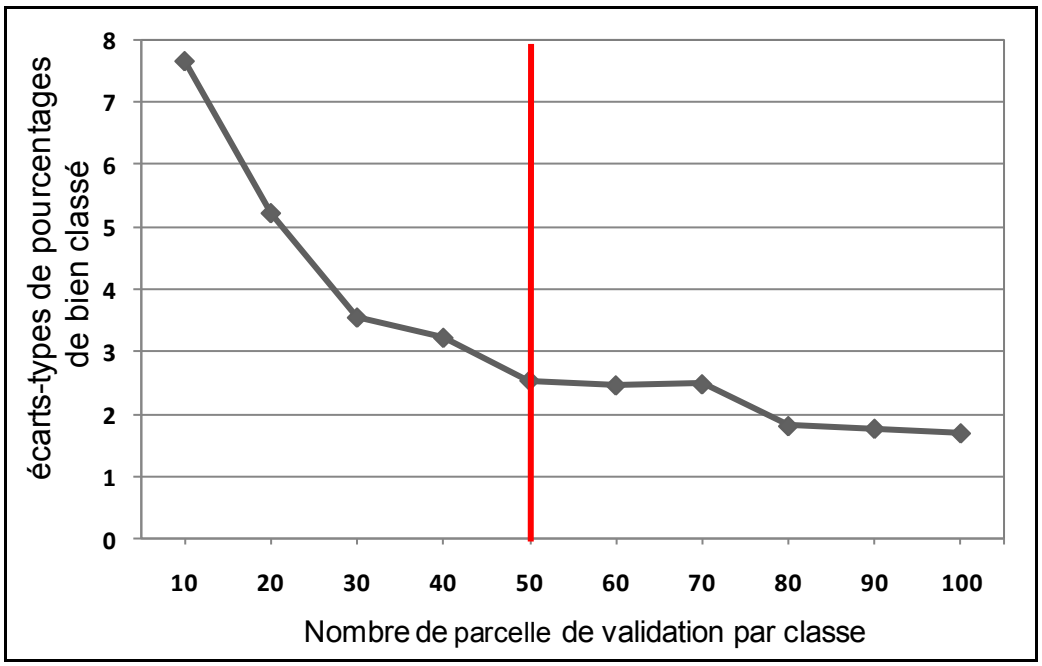


Figure 8 : Cette figure présente l'évolution des écarts types de pourcentage d'objets bien classés (précision globale) en fonction de l'augmentation de la taille de l'échantillon de validation.

On utilisera donc pour évaluer la qualité du résultat de classification sur la zone d'étude, des matrices de confusions construites avec un échantillonnage de 50 parcelles de validation par classe.

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE VIII. Influence de l'échantillonnage d'apprentissage sur les résultats de classification supervisée

Dans cette annexe sont présentées les deux analyses dédiées à l'étude de l'influence de l'échantillonnage d'apprentissage sur les résultats de classification supervisée.

La première analyse (Golden, 2006) a été réalisée sur un corridor rivulaire de l'Ain (12 km) en utilisant l'information d'une orthophoto. Ce travail a fait l'objet d'une communication aux IGARSS[1] (Durrieu et al., 2007) présentée en ANNEXE VIII.1.

La deuxième analyse a été réalisée sur une zone d'étude plus importante (région homogène du bassin versant de l'Hérault, Her6-Sud, 350 Km) en utilisant l'information combinée des orthophotos, de Spot5 XS et des données auxiliaires. Ce travail est présenté en ANNEXE VIII.2.

CemOA : archive ouverte d'Irstea / Cemagref

[1] International Geoscience and Remote Sensing Symposium

ANNEXE VIII.1 ***Communication acceptée aux IGARSS de 2007***

Influence of training sampling protocol and of feature space optimization methods on supervised classification results

Durrieu S., Tormos T., Kosuth P.
of UMR TETIS Cemagref-Cirad-ENGREF
Cemagref
Maison de la Télédétection en Languedoc-Roussillon,
Montpellier, France
sylvie.durrieu@teledetection.fr

Golden C.
of UMR G-EAU
Cemagref
Montpellier, France
catherine.golden@cemagref.fr

***Abstract*— Land cover map are produced from remote sensing images using per-pixel or, more recently, object-based classifications. Various trainable classifiers and feature space optimization methods can be used to that aim. The choice of both training and control samples is liable to influence the results according to the classification method employed but little is known about the way of choosing an appropriate sampling set. This makes thus the focal point of our study. Using three sampling methods and four discriminative classifiers we compared various classification procedures, some of them including a feature space optimization step. The one that led to the best results was LDA preceded by its feature pre-selection algorithm. Generally, for training samples, class numbers of 40 were necessary to get the best results.**

Keywords : remote sensing, classification accuracy, sampling, feature space optimization, descriminative classifier

I. CONTEXT AND OBJECTIVES

Accurate digital land cover mapping is interesting for many applications related to land management. Several per-pixel or, more recently, object-based classification methods are commonly used to produce such maps from remote sensing images [1],[2]. When proceeding to an object-based classification, segmentation is first performed and, in addition to radiometric characteristics, textural, shape and contextual related attributes (or features) are computed for each segment. Most of the classifiers [3] used are trainable, and training sample characteristics are used as a reference to build up the decision rules that will be applied to the whole data set.

Although there is no accepted standard method of accuracy reporting, the trend is to base accuracy assessment on the confusion (or error) matrix [4] which is built using a control sample, independent of the training one. Quantitative metrics of classification accuracy can be derived such as overall accuracy, commission and omission errors, or kappa coefficient [5]. However, one must keep in mind that the basic assumptions underlying accuracy assessment using confusion matrix may not be satisfied (e.g. when working at coarse scale with mixed pixels) [4].

Two steps of the classification process require a particular attention. The first one lies in the definition of proper ground-reference data sets. Generally little information is given in the literature concerning the way of choosing an appropriate training set and the reliability of results accuracy assessment, depending on control sample, is poorly discussed [6], [4]. Another critical step of the process is feature space optimization. The features characterizing pixels or objects (pixel clumps) might not be all informative; they can be correlated or non-relevant for classification. Therefore, feature space optimization aims to extract and/or select the most appropriate features for class discrimination [7], [8].

The two main objectives of this work are 1/ to study influence of training and control samples on the result of classification and 2/ to compare various discriminative classifiers and feature space optimization methods. The final goal is to supply users with recommendations for optimizing classification processes and producing reliable quality assessment.

The structure of this paper is as follow. Section 2 contains a presentation of the data set, the methods used for sample selection, classification and feature space optimization as well as the experimental design. The results are shown in section 3. Section 4 discusses the results and finally concludes this paper.

II. DATA SET AND METHOD

A. Data set

The data set was obtained by analyzing a multispectral aerial image along the river Ain in France. Segmentation using eCognition© software was first performed resulting in homogeneous land-cover plots characterized by a large set of attributes [9]. The initial feature set was reduced to 42 attributes by purposive selection. 1846 objects were thus created, each of them being characterized by 42 features. Eight classes of interest were defined: C1- water, C2-sand, C3-forest, C4-soil, C5-superficial water, C6-crops, C7-roads or paths, C8-buildings. The real land-cover class was obtained by photo interpreting every plot. Since real class is known for all plots, various training and control sets can be selected. However it is

CemOA : archive ouverte d'Irstea / Cemagref

worth noticing that classes presented a large heterogeneity in size. As shown on Table 1 a class (C3) was over-represented and two classes under-represented (C5 and C8).

TABLE I. CLASS FREQUENCY FOR THE EIGHT LAND COVER CLASSES

Class	C1	C2	C3	C4	C5	C6	C7	C8	Total
Class frequency	107	165	1320	122	12	54	52	14	1846

B. Sampling method

Validation and training sets are independent (no common element). Three sampling methods were proposed. The first one is a stratified sampling, with a class-by-class random selection according to a given sampling rate. Let x% and y% be respectively the training and control sampling rates, with (x+y) ≤100. Method 1 first randomly selects y% of the objects in class 1, y% of the objects in class 2,… to constitute the control sample. These selected elements are taken away of the original data set and x% objects of each class are then chosen among the remaining elements in order to build up the training sample. In this study the maximum value of y was set to 50. The resulting samples were unbalanced which resulted in giving too much weight to over-represented classes during decision rules establishment and validation. Second and third sampling methods aim to re-balance the per-class sampling size. The classes are first separated into two datasets, according, as previously, to a given percentage y% (with y_{max}= 50). Sampling method 2 and 3 are then applied to each dataset in order to define both validation and training samples. For method 2, we just set an allowed maximum number of objects per class and no replacement was done; thus the heterogeneity between classes was reduced when working with small sample sizes but reappeared when size increased. Method 3, inspired from the bootstrap method, allowed replacement when creating samples in order to reach an equal sample size by class. In the case of the same object could appear several times within a sample. Table 2 provides an illustration of the composition of training sets obtained with the methods 2 and 3 when the maximum number of objects per class was set at 30.

TABLE II. TRAINING PLOTS REPARTITION AMONG THE 8 CLASSES WITH 30 OBJECTS MAXIMUM PER CLASS FOR 2ND AND 3RD SAMPLING METHODS

Class	C1	C2	C3	C4	C5	C6	C7	C8
Numbers of training plots (2nd method-selected)	30	30	30	30	6	27	26	7
Numbers of training plots (3rd method/bootstrap-selected)	30	30	30	30	30	30	30	30

C. Classification and Feature Space Optimization methods

Since some feature data distributions inside classes do not follow parametric distribution law, discriminative classifiers were used [11]. Four classifiers were compared: Classification And Regression Trees (CART), K Nearest Neighbours (KNN), Support Vector Machines (SVM) and Linear Discriminant Analysis (LDA).

CART builds up binary classification trees. It consists in dividing efficiently the objects of the training sample with tests on attributes. The tree establishment procedure stops when each subset of training plots contains objects of the same (or almost the same) class. KNN algorithm is one of the simplest statistical nonparametric classifiers that have been studied extensively from both the theoretical and the practical point of view [6]. It attributes to each object the major class among those of its k-closest training samples (Euclidian distance meaning). After preliminary tests K was set to 10. SVM is one widespread recent, robust and efficient kernel method for data classification. We used a Gaussian kernel. At least, LDA, which is well adapted to classification problems, allows 1/ to elaborate a Discriminant Function Analysis, that identifies directions in space along which classes are well separated and then 2/ to build a classifier so that observations can be assigned to classes.

These classification methods were applied with and without an associated feature space optimization method. Principal component analysis (PCA), widely used in remote sensing studies, was chosen to precede KNN and CART classifications. Although LDA already includes its own feature extraction process (it creates new features from the original ones by linear combination), an additional feature selection algorithm inspired from sequential forward selection algorithms [9], was also developed. It was built as follow:

- LDA was first performed with only one feature. This step was repeated k times, for each of the k initial features. The feature leading to the best overall accuracy was kept.
- LDA was then performed with two features, the previously selected one plus one of the (k-1) remaining features. This step was repeated k-1 times to find the best variable that, coupled with the first one, enabled to reach the highest overall accuracy.
- These stages were repeated k times, each time adding one variable to the former pre-selected set. At least, j features (j<k) providing the maximal overall accuracy were kept.

D. Experimental design

A sampling procedure is defined by the sampling method and rates. After a preliminary study the size of the control sample was set to a 50% per class rate when using sampling method 1. When using method 2 and 3 an additional constraint, consisting in limiting the maximum number of objects per class to 60, was imposed. This led to 60 or less and respectively exactly 60 objects by class for validation. The training set was chosen among remaining plots and varied from 5% to 50% for sampling method 1, and from 5 to 100 elements for sampling methods 2 and 3 (corresponding to the maximum and respectively exact numbers of plots selected by class, e.g. table 2). When studying the influence of training sample protocol, the whole classification procedure was repeated 400 times for a given sampling procedure. Only one control sample was first selected and both the selection of a training sample and the classification process were repeated 20 times. Means and standard deviation of overall accuracies were calculated. Then,

CemOA : archive ouverte d'Irstea / Cemagref

in order to avoid bias from a hypothetical unrepresentative control sample, the previous process was repeated 20 times, with various control samples. 400 results were thus obtained for a single sampling procedure. Means of previous means, means of previous standard deviations (later referred as global standard deviation), min and max values of overall accuracies were used to assess the whole procedure performance. The mean overall accuracy indicates the mean performance of the procedure. Standard deviation, min and max are indicators of procedure robustness and of its sensitivity to changes in chosen plots. Results by class were also available. The procedures were implemented using R© software.

This experimental protocol allowed us to study the influence of the training sample size on classification results and to compare various classification methods.

III. RESULTS

Figures 1 and 2 illustrate the results obtained with the different classifiers when sample size increased for the second sampling method, which is the closest from usual practice. Figure 1 shows mean overall accuracy. It can be seen that LDA, with or without optimization procedure, lead to a mean overall accuracy that could exceed 80 %. The SVM method provided also high quality results. The results obtained with CART were of poorer quality with mean overall accuracies always below 65%, which is insufficient from a user point of view. For all the methods, the results were improved when the number of training plots by class increased from 10 to 40 or 50. No significant improvement was noticed beyond these values.

Figure 2 presents the standard deviations of overall accuracy. For LDA, low global standard deviation revealed the robustness of the method. The highest global standard deviations of overall accuracies, estimated for the 400 repetitions, were obtained for CART and CART+ACP. CART appeared as the most sensitive method to a change in training set.

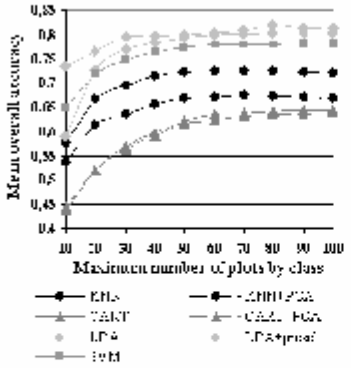


Figure 1. This figure presents evolution of mean overall accuracy (expressed in percentage of well-classified plots) with increase in training sample size (with sampling method 2) for various classification methods, applied with or without its associated feature space optimization methods.

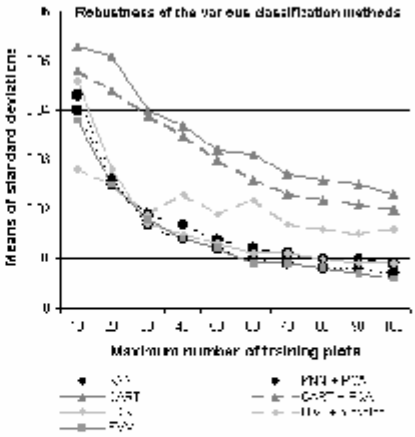


Figure 2. Evolution of 'global' standard deviations when training sample size increases (with sampling method 2) for various methods of classification, applied with or without its associated feature space optimization methods.

Analysis of min and max overall accuracy values confirmed both LDA robustness and CART sensitivity: higher (max-min) differences were obtained with CART compared to LDA (e.g. 0.23 against 0.10 when maximum sampling class number was set to 40).

Almost all methods (except LDA preceded by the optimization procedure) presented high standard deviations when working with small-size training plots and became more robust when considering large-size training plots (Fig 2). This could be confirmed by the (max-min) differences that were always higher when working with small-size training plots than when considering larger-size samples. This trend was less obvious for LDA preceded by the optimization procedure.

Concerning feature space optimization, overall accuracies obtained with PCA are lower or equal than without PCA (CART and KNN) (Fig. 1). As for the pre-selection method developed for LDA classification, it appears very efficient.

Sampling method 2 has been developed in order to re-balance the per-class sampling size and to avoid giving too much weight to over-represented classes during decision rules establishment. It is thus interesting to compare the per-class results obtained with each of the three methods. We present on Figure 3 the commission errors, which are for each class the percentage of classified objects that have been misclassified compared to the reference. We chose to analyze the per-class results obtained with one of the most performing methods, i.e. LDA without optimization procedure. For methods 2 and 3, the number of objects in each class was chosen equal to 10, which is equivalent to a sampling rate of 5% for method 1.

When the first method was employed, several underrepresented classes were extremely misclassified (81% and 88% commission errors for classes 5 and 8 respectively) whereas the over-represented class was

CemOA : archive ouverte d'Irstea / Cemagref

particularly well classified (7% of commission errors only) (Fig. 3). By-class analysis of results confirmed thus that overall accuracy was representative of the over-represented class (C3) when classic stratified sampling procedure was applied.

Regarding results provided by method 2, commission errors never exceeded 54% but always remained equal or higher than 30%. The error rates for the under-represented classes C5 and C8 (46% and 49%) were lower than those obtained with method 1 (Fig 3). Method 2 gave just about the same weight to each class, which contributed to reduce the variation of the commission errors between classes. Concerning the third method (bootstrap), the error rates were close to those obtained with the second method but remained however higher.

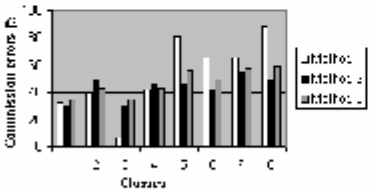


Figure 3. Percentage of commission errors (% of the classified objects that are misclassified in relation to the reference) obtained with LDA without the optimization procedure for sampling methods 1 (with a 5 % sampling), 2 and 3 (with a maximum class number of 10).

IV. Discussion and Conclusion

This study allowed us to compare various classification methods. The best performing and most robust classifier was LDA preceded by the optimization procedure with high overall accuracy rates (80% overall accuracy when training sample had a maximum of 50 objects by class; low standard deviations and the lowest (max–min) differences). This can be easily understood because LDA extracts some features that discriminate at best between classes. The optimization procedure allows pre-selecting the most discriminative features before to create the final features by linear combination. SVM produced also high quality results. However it requires a tricky and critical step: the choice of a kernel function and its parameters. KNN, which is a simple and widely used method, gave average results.

Classification methods that had their own feature extraction process produced the best results. Adding a step of feature pre-selection could enhance even more their efficiency. PCA, which is widely used in remote sensing studies, must be used cautiously when proceeding to feature space optimization. As PCA maximizes inter-object and not inter-class distances it might deteriorate classification results.

For all the sampling methods, the results were improved when the number of training plots by class increases from 10 to 40 or 50. No significant improvement was noticed beyond these values. Only LDA preceded by the optimization procedure allowed reaching a high overall accuracy with a small number of training sites.

In this study, training plots were randomly selected. It would be interesting to compare results obtained with random selection to those obtained with an operator purposive selection, which is commonly realized in remote sensing studies. The selection by an expert of the most representative objects for each class; might allow to reach the same performances as those obtained with random sampling with a smaller training sample size. The choice of a sampling procedure depends on user objectives: if the same thematic importance is attributed to each class, even the spatially underrepresented ones, second sampling procedure appears more suitable. This method allowed to decrease the error rates in the under-represented classes. These errors remained however quite high (>40%) what can be explained by a high heterogeneity of these classes and/or by unrepresentative plot selection. Class definition remains a critical step. It is often difficult to choose interesting classes from thematic point of view that can be correctly discriminate according to their radiometric, texture and shape attributes. Although results are likely to be partially case-dependant, further experiments are in progress on other sites to complete this first set of results in order to supply users with recommendations for optimizing classification processes.

References

[1] M.C. Girard, C.M. Girard, Traitement des données de télédétection ("Processing of Remote sensing data"), DUNOD editions, Paris, 1999, 529 p.

[2] M. Bock, P. Xofis, J. Mitchley, G. Rossner, M. Wissen, "Object-oriented methods for habitat mapping at multiple scales – Case studies from Northern Germany and Wye Downs" Journal for Nature Conservation vol 13, UK, 2005, pp 75-89.

[3] T. Hastie, R. Tibshirani, J. Friedman, "The elements of statistical learning." Data Mining, Inference and Prediction, Springer editions, 2001, pp 84-95.

[4] G.M. Foody , "Status of land cover classification accuracy assessment" Remote Sensing of Environment 80, 2002, pp.185-201.

[5] Gilbert Saporta, Ganamo Youness, "With regard to indexes for comparing two classifications", "Sur des indices de comparaison de deux classifications", 10èmes rencontres de la Société Francophone de Classification, Neuchatel, September, the 10-12th. 2003, pages 177-180.

[6] M. Chi, L. Bruzzone, "An ensemble-driven k-NN approach to ill-posed classification problems", Pattern Recognition Letters 27, 2006, pp 301–307.

[7] François Leduc, "Feature space optimisation prior to fuzzy image classification", the 7th International Conference on Information Fusion, Stockholm, Sweden, June the 28th – July the first 2004, 8p.

[8] G. Saporta, Probabilités, Analyse des Données et Statistique ("Probabilities, Data analysis and statistics"), TECHNIP editions, 1990.

[9] M. Perez Correa, "Object-based classification of very high spatial resolution images; application to land cover mapping along river corridors", "Classification orientée objet d'images à très haute résolution spatiale: application à la cartographie de l'occupation des sols le long des cours d'eau", Rapport Mastère SILAT, 2004, 43 p.

[10] D. Semani, C. Frélicot, and P. Courtellemont, "Fuzzy/possibilistic labels combination for feature selection", "Combinaison d'étiquettes floues/possibilistes pour la sélection de variables" 14ieme Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, vol 2, Toulouse, January 2004, 10p

[11] C. Van der Walt and E. Barnard, "Data characteristics that determine classifier performance" in Proceedings of the Sixteenth Annual Symposium of the Pattern Recognition Association of South Africa, 2006, pp.160-165.

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE VIII.2 *Analyse à l'échelle d'une région homogène*

Cette annexe présente l'analyse dédiée à l'étude de l'influence de l'échantillonnage d'apprentissage sur les résultats de classification supervisée sur une zone d'étude plus importante (région homogène du bassin versant de l'Hérault, Her6-Sud, 350 Km) en utilisant l'information combinée des orthophotos, de Spot5 XS et des données auxiliaires.

Le jeu de données d'objets utilisé pour cette analyse provient du résultat de classification obtenu lors de la mise en œuvre de la démarche orientée objet basée sur la technique de classification « expertisée » (par fonction d'appartenances). L'arbre et les critères de classification définis sur cette zone sont présentés en ANNEXE VII.2 Figure 4.

Une typologie à 9 classes a été retenue pour l'analyse :

- Sols artificialisés
- Terres agricoles en végétation
- Terres agricoles nues (labours)
- Vignes
- Végétation herbacée
- Végétation arbustive semi-naturelle
- Végétation arborée
- Surfaces en eau

Cette classification a été préalablement validée suivant cette typologie à partir d'une matrice de confusion construite avec 50 parcelles de validation par classe (taille échantillon optimale cf. ANNEXE VII.3) Ces parcelles ont été sélectionnés aléatoirement et de façon à être réparties de manière homogène sur la zone d'étude puis photo-interprétés. 90% des objets ont été correctement classés.

Dans cette couche vecteur d'occupation du sol, seuls les objets classés à partir de l'information des images ont été sélectionnés constituant ainsi un jeu de données de 31334 objets. Pour chacun de ces objets 22 attributs (cf. Figure 1) ont été extraits.

A partir de ce jeu de données, deux protocoles d'étude ont été définis. Ils utilisent uniquement l'analyse discriminante linéaire (ADL) comme technique de classification supervisée. L'ADL est la technique qui a permis d'obtenir les meilleurs résultats de classification selon la première analyse (cf. ANNEXE VIII.1). Le mode d'échantillonnage testé se base sur un raisonnement de sélection en termes de x nombres de parcelles par classe. Ce raisonnement est le plus souvent utilisé en télédétection (Girard and Girard, 1999).

Pour limiter le biais lié aux erreurs de classifications, les protocoles d'étude ont été itérés un grand nombre de fois.

Premier protocole

L'objectif de ce protocole est d'identifier la taille d'échantillonnage optimale (i.e, le nombre de parcelles d'apprentissage par classe permettant d'obtenir les meilleurs résultats de classification).

Pour cela, une chaîne de traitements (Figure 1) composée de sept étapes a été mise en place pour analyser l'influence du nombre (X) de parcelles d'apprentissage sur la qualité du résultat de la classification supervisée utilisant l'ADL. Plusieurs nombres (X) de parcelles par classe ont été testés allant de 10 jusqu'à 100 au pas de 10. Ce travail a été effectué sur la moitié de la zone d'étude (la moitié des objets soit 15667 objets). Dans cette analyse, les échantillonnages d'apprentissage et de validation sont indépendants.

CemOA : archive ouverte d'Irstea / Cemagref

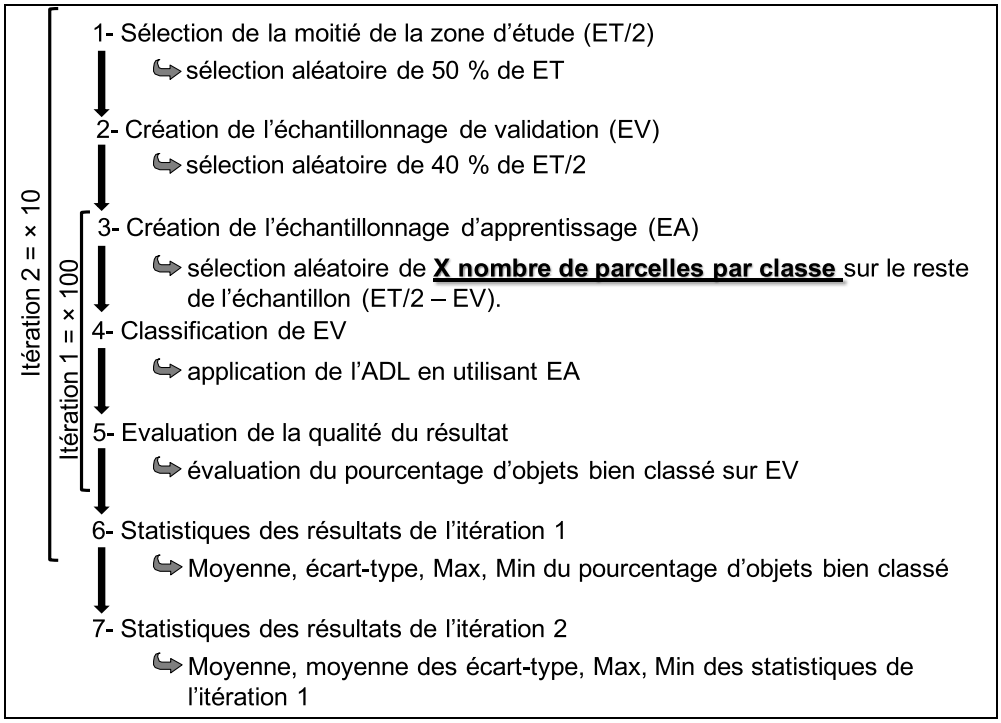


Figure 1 : Chaîne de traitements mise en place pour étudier l'influence du nombre (X) de parcelles d'apprentissage sur la qualité du résultat de classification avec ET = Echantillon total (jeu de données, soit 31334 objets) et ADL = Analyse Discriminante Linéaire.

Le tableau 1 présente les résultats obtenus lors de ce protocole.

Tableau 1 : Statistiques du pourcentage de bien classé (moyenne, moyenne des écart-types, minimum, maximum) obtenues pour différentes tailles d'échantillonnage d'apprentissage. PA = Parcelles d'Apprentissage.

Nombre de PA	Moyenne	Min	Max	Moyenne Ecart-type
10	69,2454	64,1433	75,4786	3,3417
20	72,8757	69,2884	76,2843	2,4627
30	74,5421	71,6656	77,0980	1,9062
40	75,9405	72,5191	79,9378	1,8755
50	77,1323	74,2661	79,5150	1,6734
60	77,0557	73,8034	80,4483	1,5031
70	77,9834	75,3031	79,1560	1,4468
80	77,5375	74,6729	79,7862	1,4250
90	78,0967	75,5185	80,1851	1,4236
100	77,0716	75,1595	79,1560	1,3502

Comme pour la première analyse (cf. ANNEXE VIII.1), on observe les meilleurs résultats (pourcentage de bien classé) à partir de 50 parcelles d'apprentissage par classe. Après 50 parcelles, on n'observe pas d'amélioration significative du pourcentage de bien classé (Figure 2). On note aussi que plus le nombre de parcelles est grand, plus la classification supervisée utilisant l'ADL est robuste (faible écart-type). Elle se stabilise pour un nombre compris entre 50 et 60 parcelles (Figure 3).

CemOA : archive ouverte d'Irstea / Cemagref

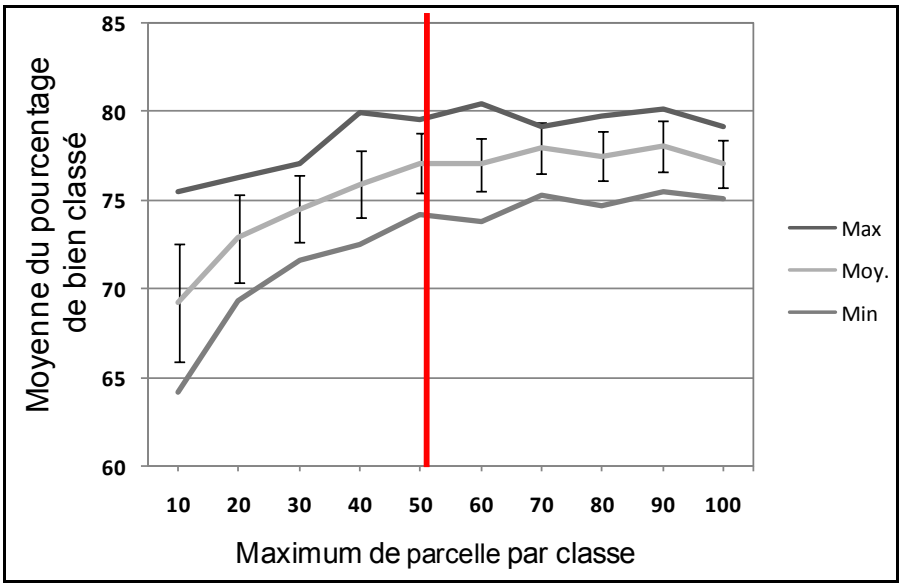


Figure 2: Cette figure présente l'évolution des moyennes du pourcentage d'objets bien classé (précision globale) en fonction de l'augmentation de la taille de l'échantillon en nombre d'objets par classe pour l'analyse discriminante linéaire (sans méthode d'optimisation de variables) appliqué sur un corridor rivulaire de l'Hérault (15667 objets).

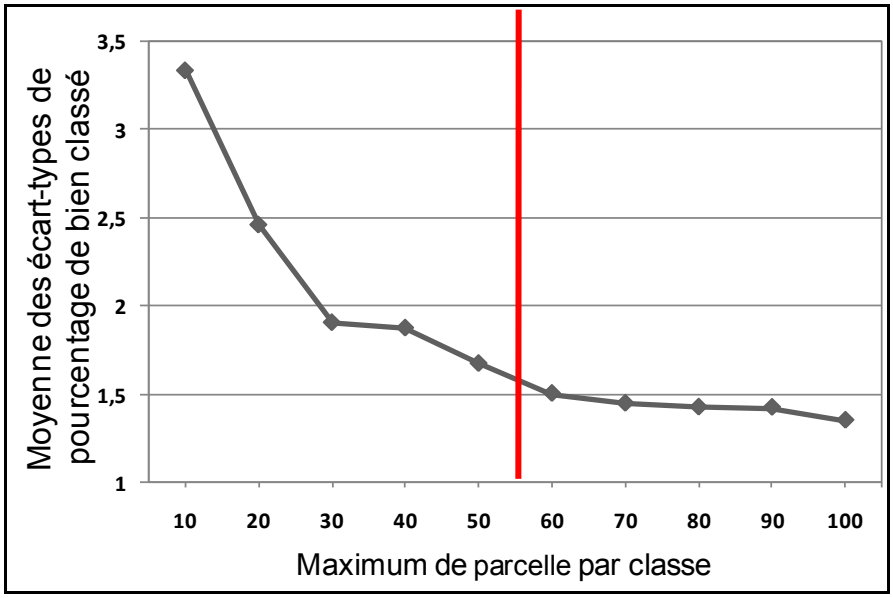


Figure 3: Cette figure présente l'évolution des moyennes des écart-types du pourcentage d'objets bien classé (précision globale) en fonction de l'augmentation de la taille de l'échantillon en nombre d'objets par classe pour l'analyse discriminante linéaire (sans méthode d'optimisation de variables) appliquée sur un corridor rivulaire de l'Hérault (15667 objets).

Deuxième protocole

L'objectif de ce protocole est de tester si le nombre de parcelles d'apprentissage optimal (50 parcelles par classe) est robuste quelque soit la taille de la zone d'étude (i.e, le nombre d'objets).

CemOA : archive ouverte d'Irstea / Cemagref

Pour cela, une chaîne de traitements (Figure 4) composée de sept étapes a été mise en place pour analyser l'influence de la taille de la zone d'étude (en fonction d'un % de Y sur le nombre total d'objets) sur la qualité du résultat de la classification supervisée utilisant l'ADL avec l'échantillonnage d'apprentissage optimal. Plusieurs tailles de zone d'étude ont été examinées : pour Y = 25% (le quart de la zone), 33 % (le tiers de la zone), 50% (la moitié de la zone) et 100% (la zone complète). Dans cette analyse, les échantillonnages d'apprentissage et de validation sont indépendants.

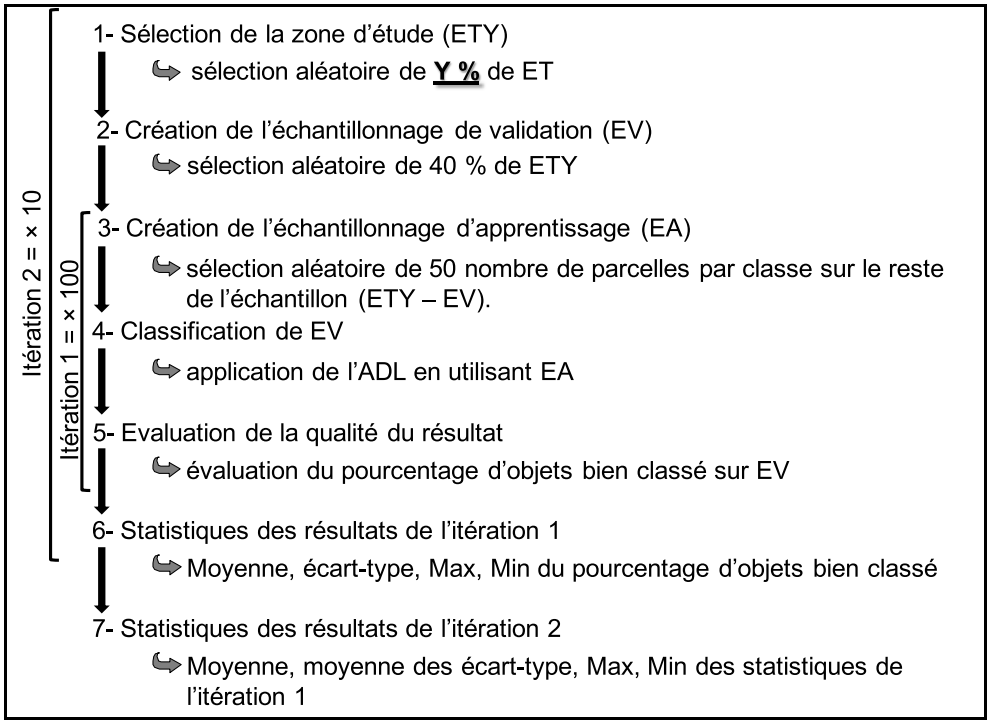


Figure 4 : Chaîne de traitements mise en place pour étudier l'influence de la taille de la zone d'étude (en fonction d'un % de Y sur le nombre total d'objets) sur la qualité du résultat de la classification supervisée utilisant l'ADL avec l'échantillonnage d'apprentissage optimal. ET = Echantillon total (jeu de données, soit 31334 objets) et ADL = Analyse Discriminante Linéaire.

Le tableau 2 présente les résultats obtenus lors de ce protocole.

Tableau 2 : Moyenne du pourcentage de bien classé pour différentes tailles de zone d'étude à partir d'une classification supervisée utilisant l'analyse discriminante linéaire avec la taille d'échantillonnage optimale (50 parcelles par classe).

Zones étudiées	Nombres d'objets	Moyenne du pourcentage de bien classé
le quart de la zone	7834	76,9650
le tiers de la zone	10445	77,5058
la moitié de la zone	15667	77,1323
la zone complète	31334	77,1025

Cette analyse montre que le nombre de 50 parcelles d'apprentissage est robuste pour mener une classification supervisée quelque soit la taille de la zone d'étude puisqu'il n'existe qu'un lien très faible entre le pourcentage de bien classé et la taille de la zone d'étude (en nombre d'objets) (R^2 = 0.0279, Figure 5).

CemOA : archive ouverte d'Irstea / Cemagref

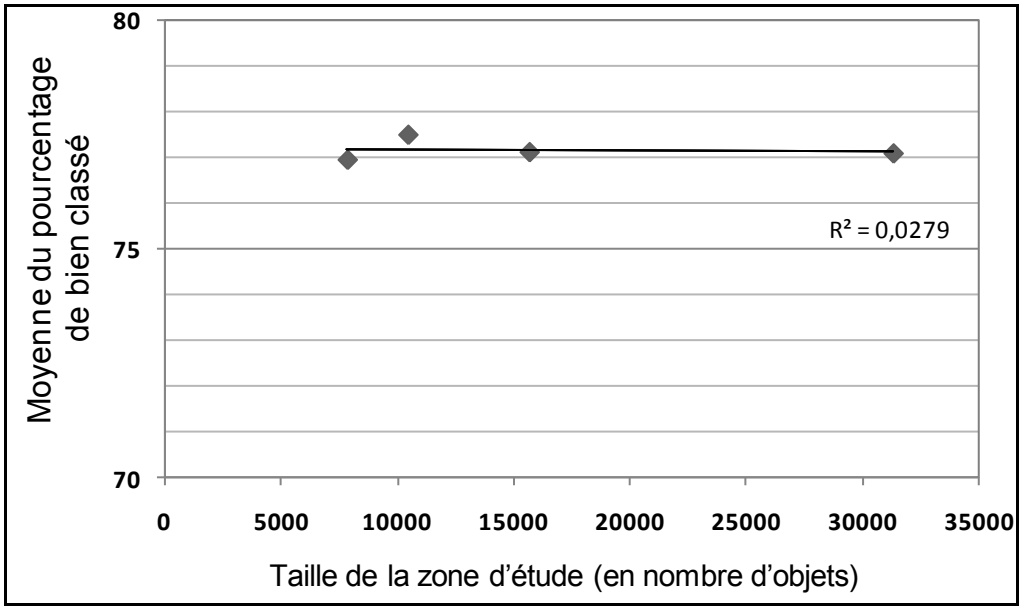


Figure 5: Cette figure présente l'évolution des moyennes de pourcentage d'objets bien classés (précision globale) en utilisant l'analyse discriminante linéaire (sans méthode d'optimisation de variables) avec 50 parcelles/classe en fonction de l'augmentation de la taille de la zone d'étude (nombres d'objets).

Ces deux analyses (sur un tronçon de l'Ain et sur une région homogène de l'Hérault) indiquent clairement que l'entraînement d'un algorithme de classification supervisée dédié à l'extraction de l'occupation du sol dans les corridors rivulaires appliqué dans une région de types d'occupation du sol homogène, nécessite 50 parcelles d'apprentissage par classe quelque soit sa taille (i.e son nombre d'objets).

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE IX. Sensibilité des indicateurs aux erreurs de décalage de la BDCarthage®

La BDCarthage® issue de la numérisation des entités hydrographiques sur les cartes au 1/25 000ème de l'IGN peut être décalée par rapport aux orthophotos plus précises. Ces décalages peuvent engendrer des erreurs dans la quantification des indicateurs, notamment ceux qui sont calculés à proximité immédiate du cours d'eau (ex. indicateurs linéaires ou de contact).

Comme il n'est pas envisageable dans le cadre de ce travail de « corriger » la BDCarthage® (niveau 2) sur l'ensemble du réseau hydrographique, une analyse a été menée sur un sous bassin-versant de l'Hérault (Figure 1) pour étudier la sensibilité des indicateurs aux erreurs de décalage de cette donnée.

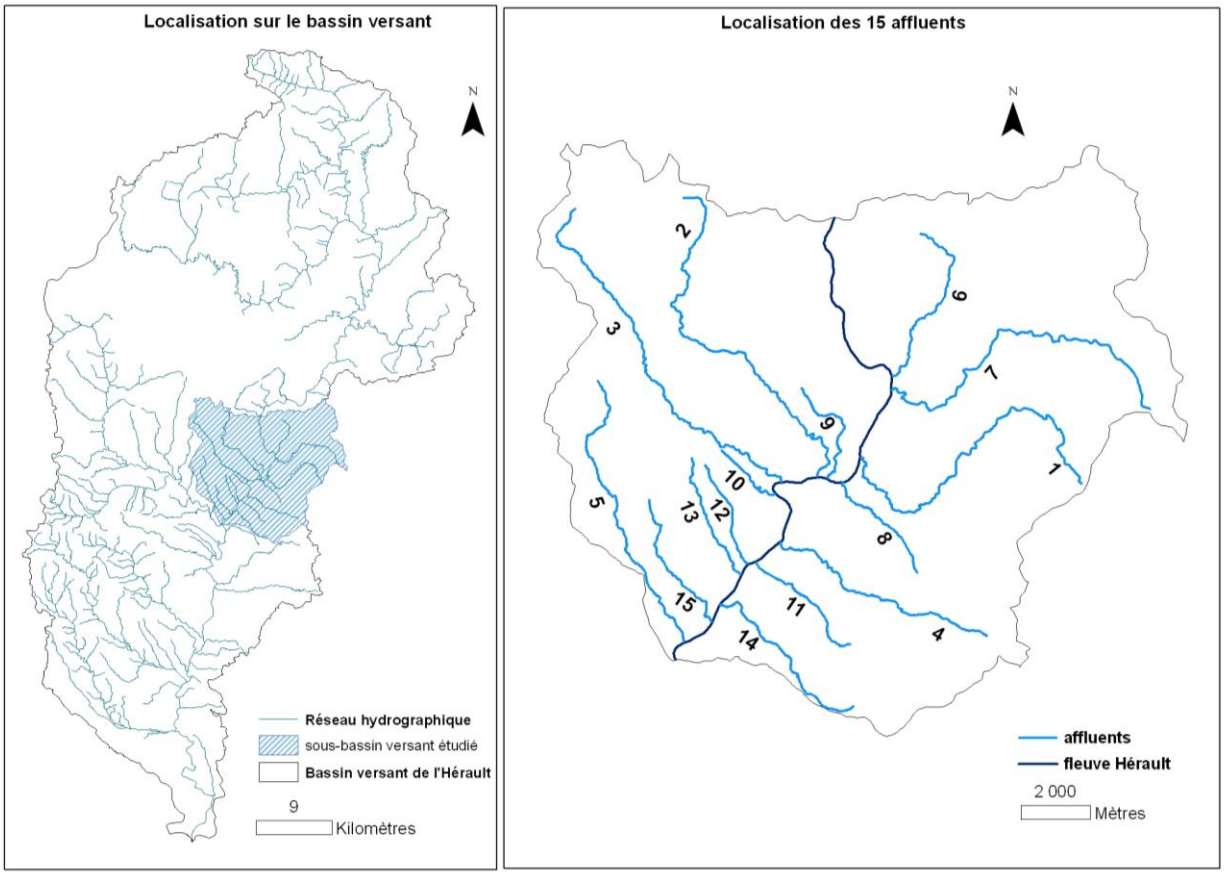


Figure 1 : Présentation du sous-bassin versant utilisé pour l'analyse de la sensibilité des indicateurs aux erreurs de décalage de la BDCarthage®.

Pour cela, nous avons dans un premier temps construit un réseau hydrographique corrigé par photo-interprétation sur la base des orthophotos.

Dans un deuxième temps, nous avons quantifié le décalage entre ce réseau corrigé et le réseau BDCarthage®. Les deux réseaux (vecteur) ont été pour cela tout d'abord convertis en raster (0.5 m de résolution spatiale). Ensuite une matrice de distance euclidienne (fonction incrémentale) a été construite (exemple de cette matrice Figure 2) à partir de la BDCarthage® (raster). Enfin la couche raster du réseau corrigé a été croisée avec cette matrice de distance euclidienne (fonction de statistique zonales) afin d'extraire les statistiques de décalage (distance par rapport au réseau BDCarthage®) pour chaque tronçon du réseau corrigé. Les résultats sont présentés dans le tableau 1.

CemOA : archive ouverte d'Irstea / Cemagref

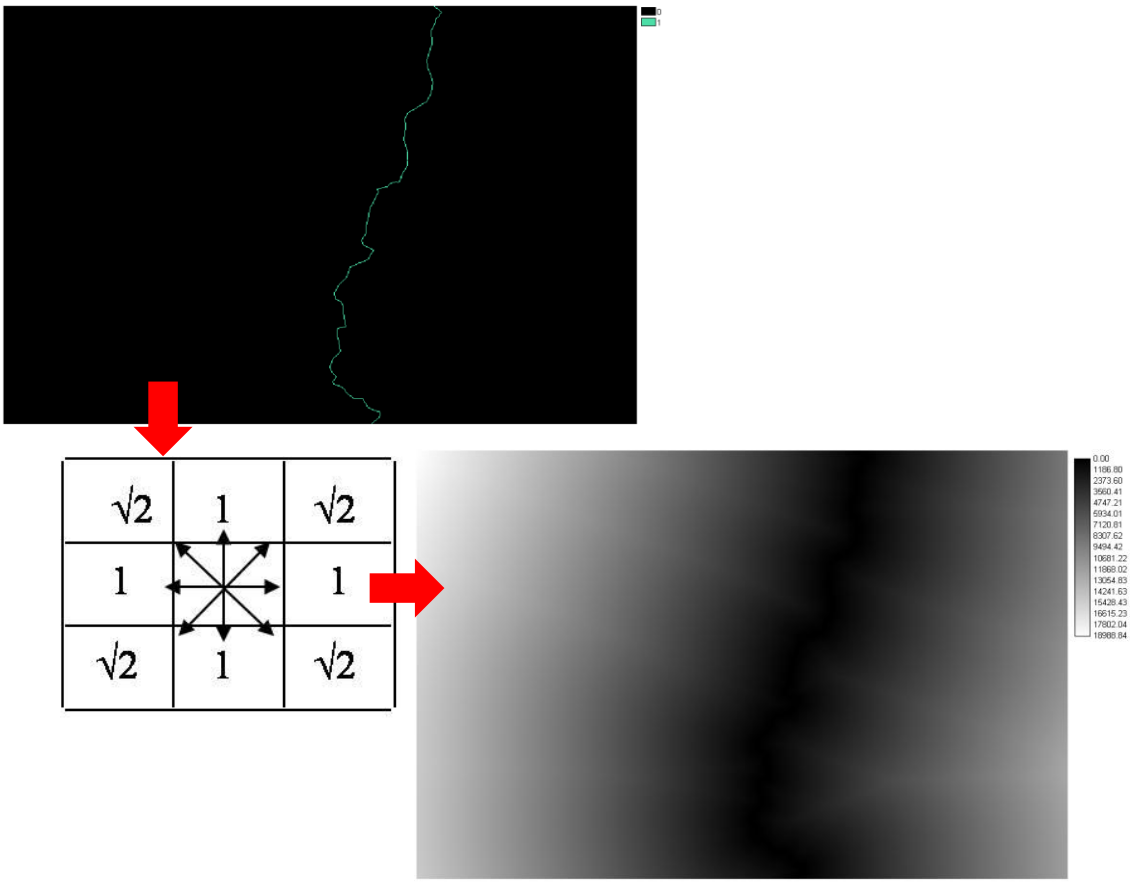


Figure 2 : exemple d'application d'une matrice euclidienne sur une partie d'un réseau hydrographique

Enfin, dans un troisième temps, nous avons construit trois indicateurs : (1) l'indicateur de longueur de ripisylve (ICR-reg-Lin-%-forêt), l'indicateur de pourcentage de prairie à 10 m du cours d'eau (ICR-reg-10m-%-prairie) et l'indicateur linéaire de pourcentage surfacique urbain (ICR-reg-Lin-%-Urbain) à partir de deux versions de donnée hydrographique « polygone » :

- La version 1 (« non corrigée ») construite à partir de la combinaison de la fusion des objets classés « cours d'eau » de la cartographie d'occupation du sol à THRS et d'un buffer très fin autour de la BDCarthage® ;

- La version 2 (« corrigée ») construite à partir de la combinaison de ces mêmes objets fusionnés et d'un buffer très fin autour du réseau hydrographique corrigé.

La différence de résultats entre une construction utilisant la version « non corrigée » et la version « corrigée » de la donnée hydrographique « polygone » pour ces trois indicateurs sur chaque tronçon du sous-bassin versant de l'Hérault sont présentés dans le tableau 1.

CemOA : archive ouverte d'Irstea / Cemagref

Tableau 1: Statistiques du décalage entre le réseau corrigé et le réseau BDCarthage® pour chaque tronçon du sous bassin-versant.

	Décalage (m)		Différence entre les deux versions (%)		
Cours d'eau	**maximum**	**moyen**	**I-lin-%-forêt**	**I-10m-%-prairie**	**I-lin-%-urbain**
fleuve Hérault	103,72	9,19	3,11	1,80	0,95
affluent 1	86,12	9,55	6,08	2,18	0,16
affluent 2	90,14	11,24	0,23	0,17	2,58
affluent 3	37,70	4,91	2,38	2,19	0,07
affluent 4	63,29	9,88	1,43	0,28	0,94
affluent 5	102,61	18,94	4,31	0,08	0,17
affluent 6	65,09	14,57	0,67	1,00	0,79
affluent 7	60,69	10,09	6,56	4,16	1,56
affluent 8	27,78	6,86	0,07	0,51	0,02
affluent 9	48,28	11,06	0,97	0,21	0,45
affluent 10	45,80	13,22	1,39	0,99	0,43
affluent 11	61,36	10,21	1,44	1,28	0,68
affluent 12	34,66	12,67	1,19	1,48	0,58
affluent 13	42,03	12,79	19,17	0,56	0,66
affluent 14	52,00	13,30	0,83	1,51	0,78
affluent 15	76,49	15,77	2,75	1,72	0,42
Moyenne	**62,36**	**11,52**	**3,29**	**1,26**	**0,70**

D'après cette analyse on remarque :

(1) Le décalage entre la BDCarthage® et le réseau corrigé est de 11.52 m en moyenne et de 62.36 m au maximum sur le sous-bassin versant de l'Hérault.

(2) Pour ces trois indicateurs (construits à proximité du cours d'eau), la différence en moyenne sur le sous-bassin versant de l'Hérault est relativement faible pour ICR-reg-Lin-%-forêt (3.29 % de différence), et très faible pour ICR-reg--10m-%-prairie (1.26 % de différence) et pour ICR-reg-Lin-%-Urbain (0.70 % de différence).

(3) La valeur des indicateurs est faiblement liée aux erreurs de décalage moyen et maximum de la BDCarthage® (exemple Figure3 pour ICR-reg--Lin-%-forêt).

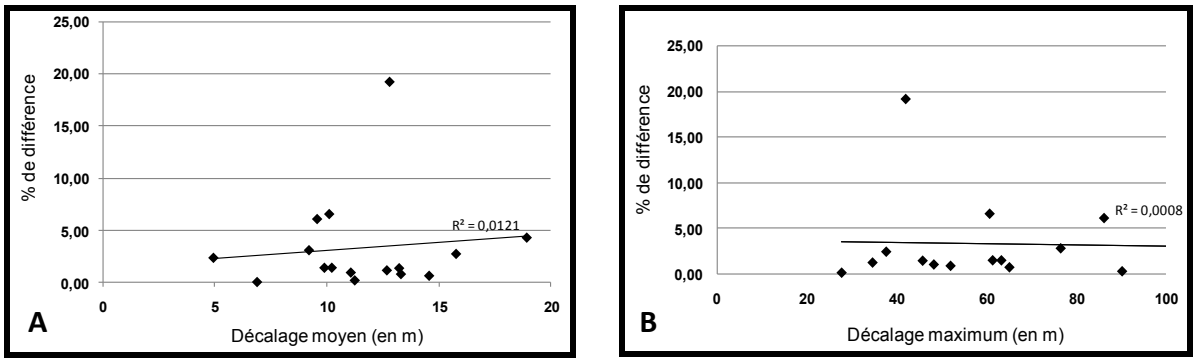


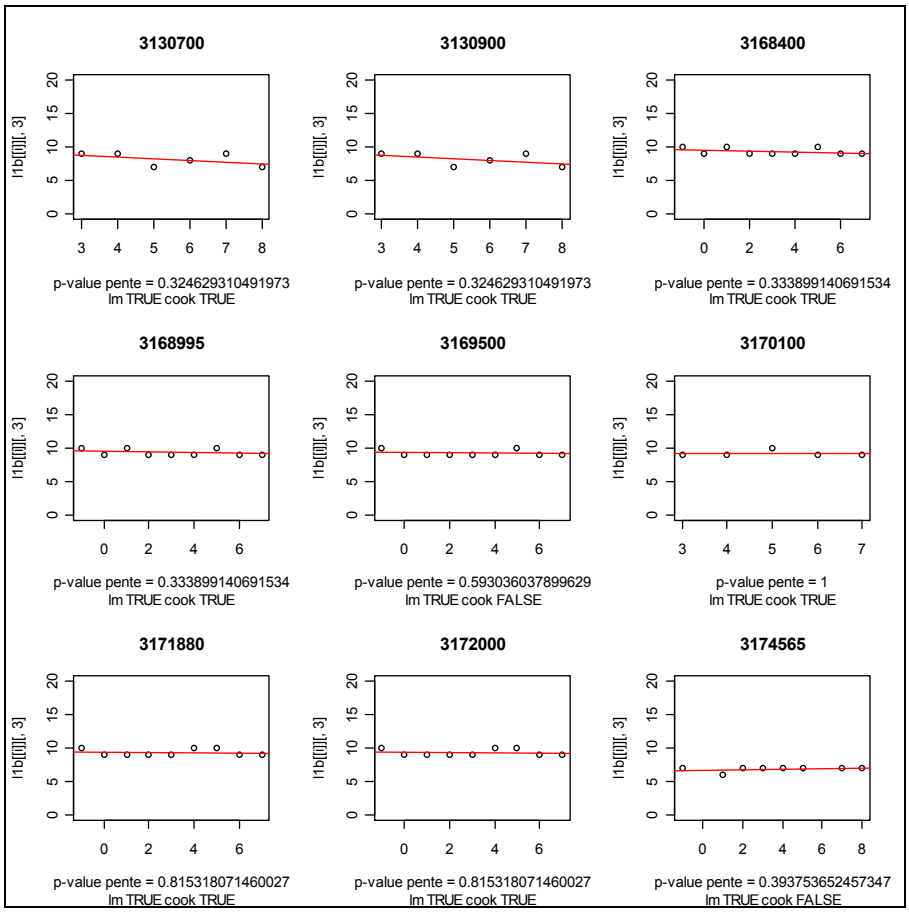
Figure 3 : Différence (en valeur absolue) entre l'indicateur de longueur de ripisylve construit à partir du « réseau non corrigé » et le même indicateur construit à partir du « réseau corrigé » en fonction du décalage moyen (A) et maximum (B) entre ces deux réseaux pour chaque cours d'eau du sous-bassin versant de l'Hérault.

CemOA : archive ouverte d'Irstea / Cemagref

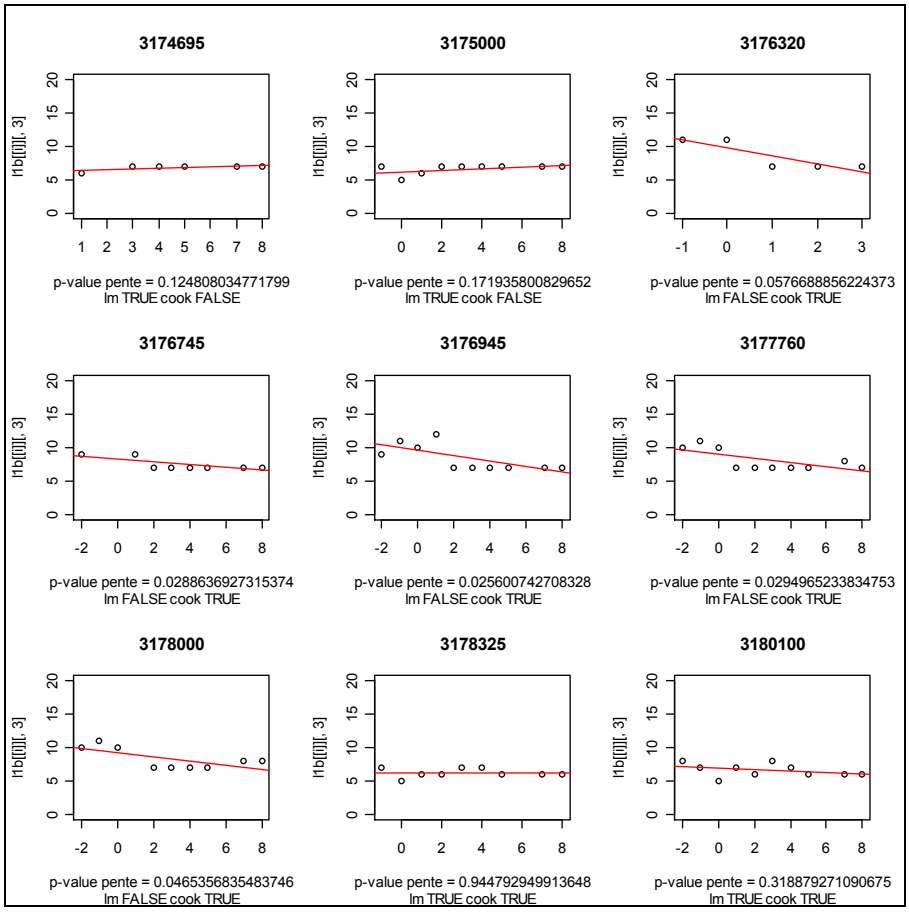
ANNEXE X. Résultats de l'analyse de la stabilité des chroniques IBGN

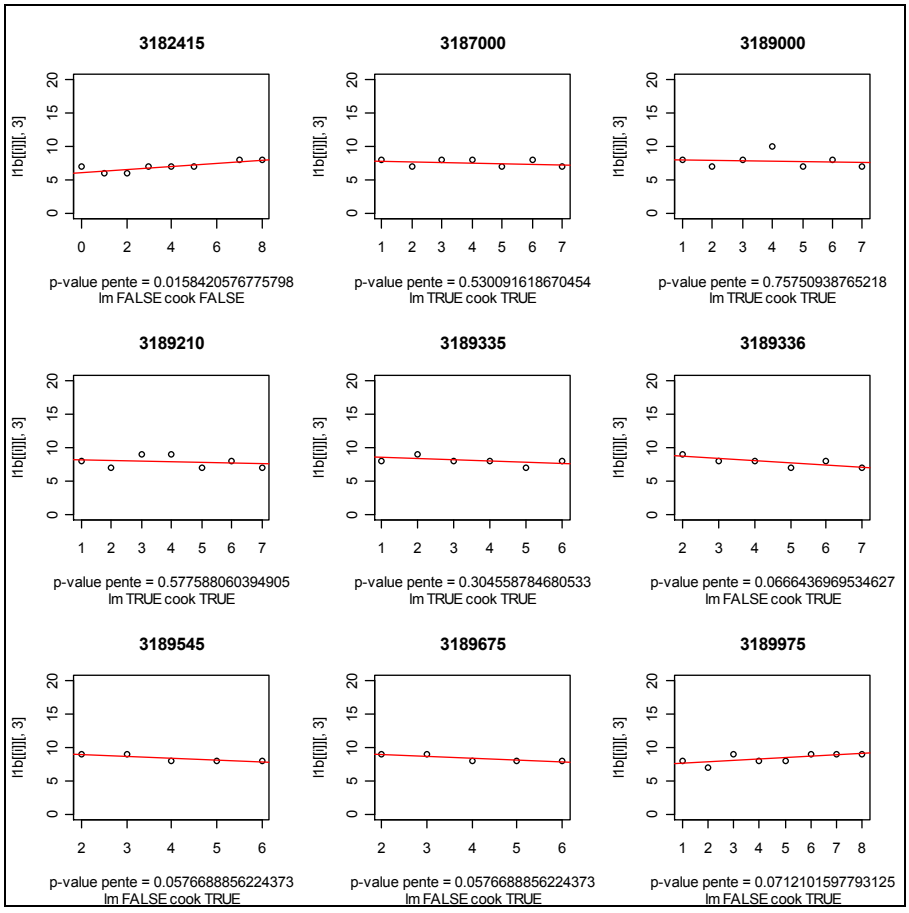
Cette annexe présente les résultats obtenus lors de l'analyse de la stabilité des chroniques IBGN. Dans cette analyse, le critère minimum retenu est une absence de tendance linéaire. La mise en évidence de ce critère dépend de la p-value de la pente du modèle linéaire simple de chaque site ; lorsque cette valeur est supérieure à 0.1, on considère que la pente n'est pas significative et que, par conséquent, la chronique considérée ne montre pas de tendance significative. Afin d'éviter les écueils habituels des modèles linéaires simples, la normalité des résidus est vérifiée et les points influents éliminés (Jacknife sur la pente et les distances de Cook).

On présente ci-dessous, pour chaque station l'évolution des valeurs IBGN sur la période étudiée (1992 à 2004) et la p-value de la pente du modèle linéaire simple en indiquant si des points influents ont été éliminés par le test de Cook.

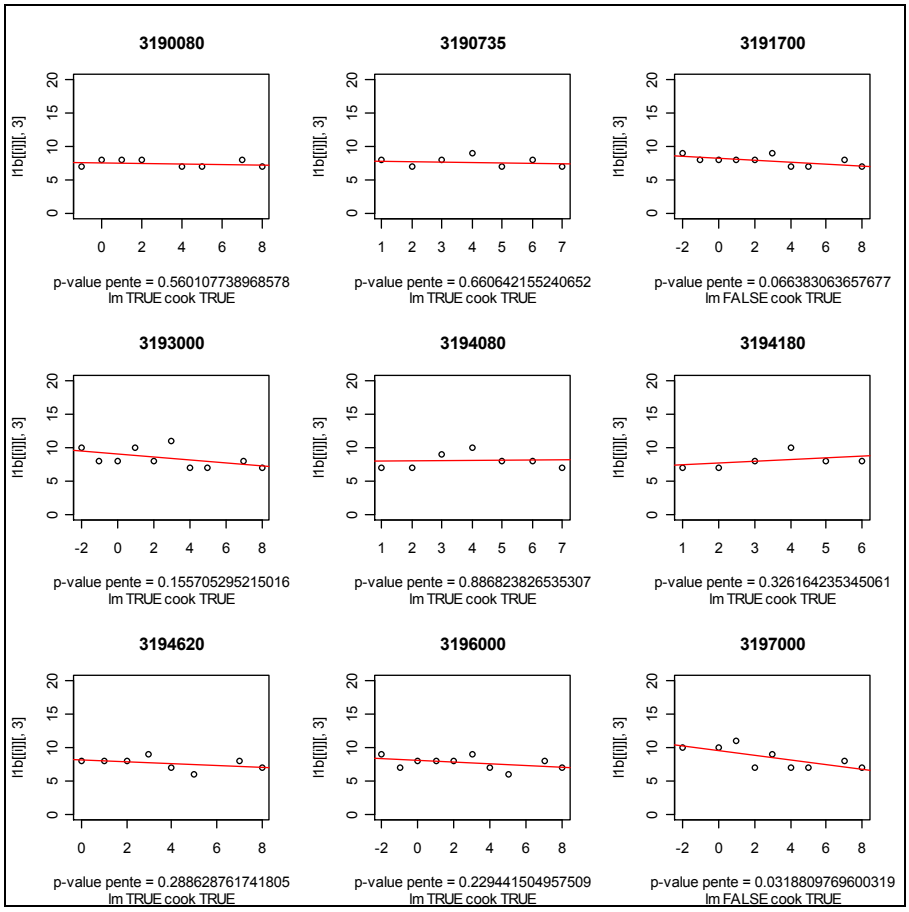


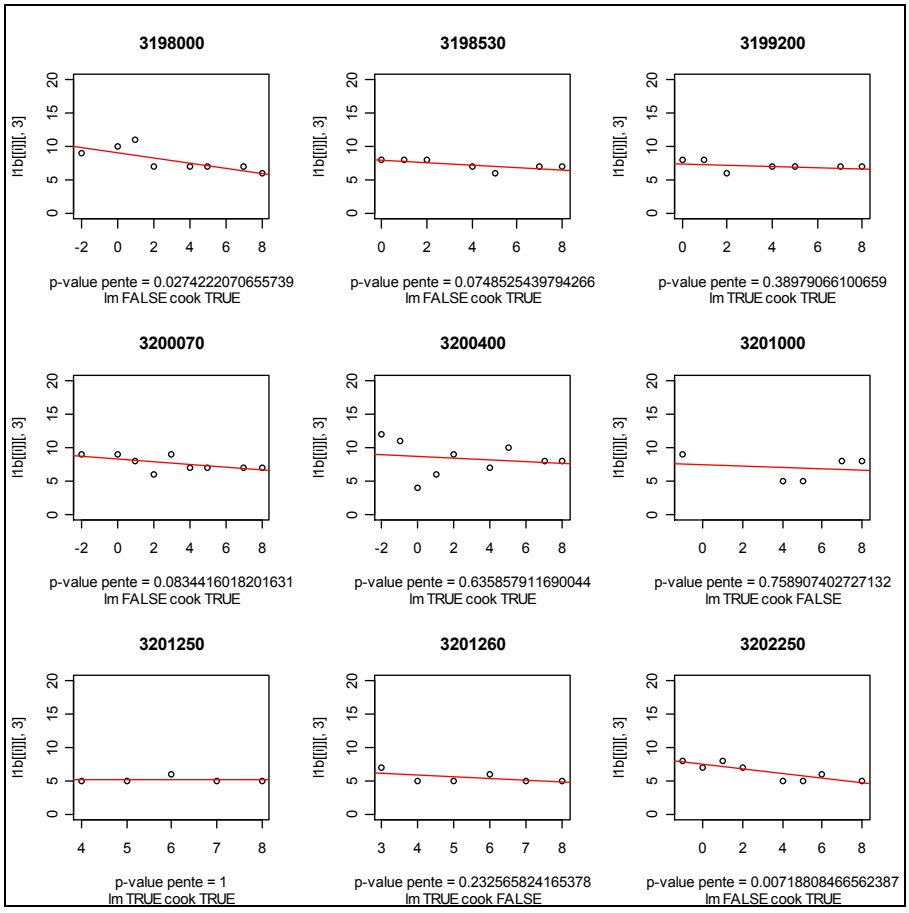
CemOA : archive ouverte d'Irstea / Cemagref



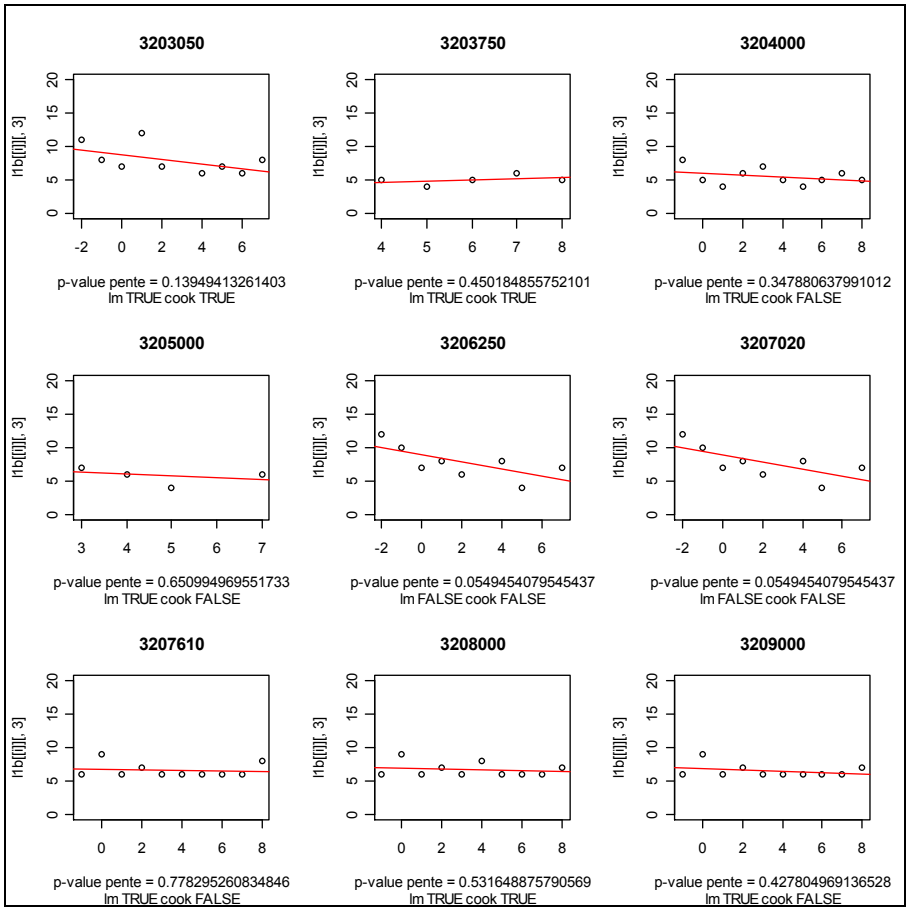


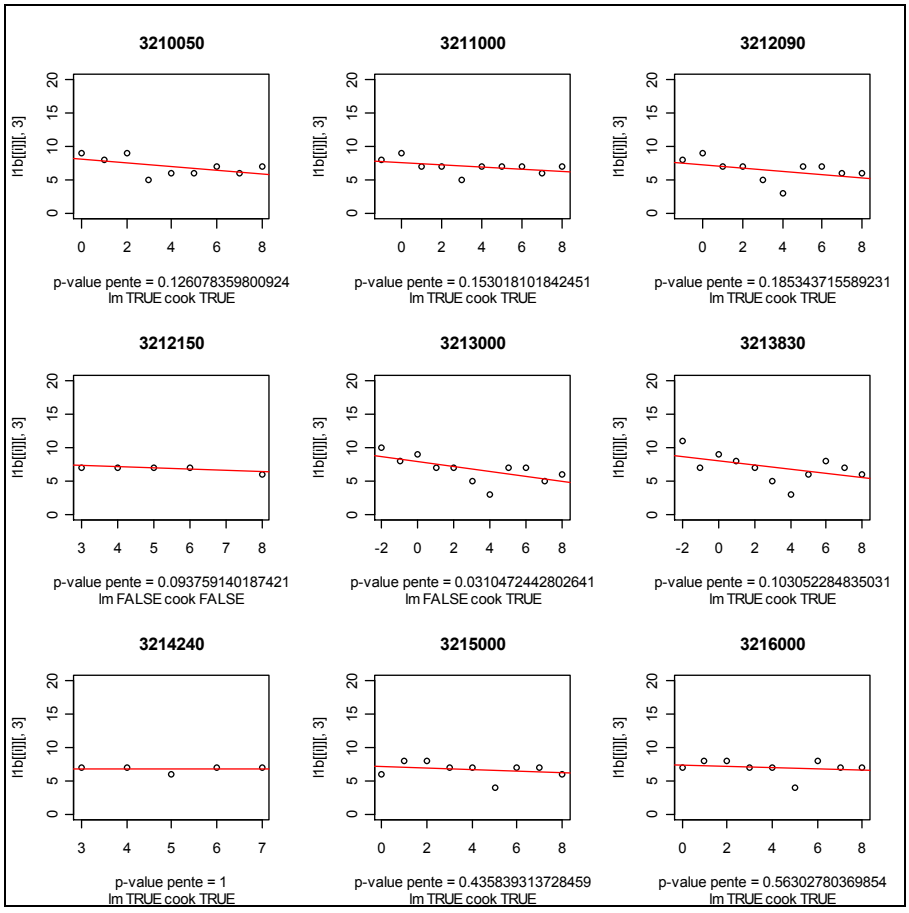
CemOA : archive ouverte d'Irstea / Cemagref



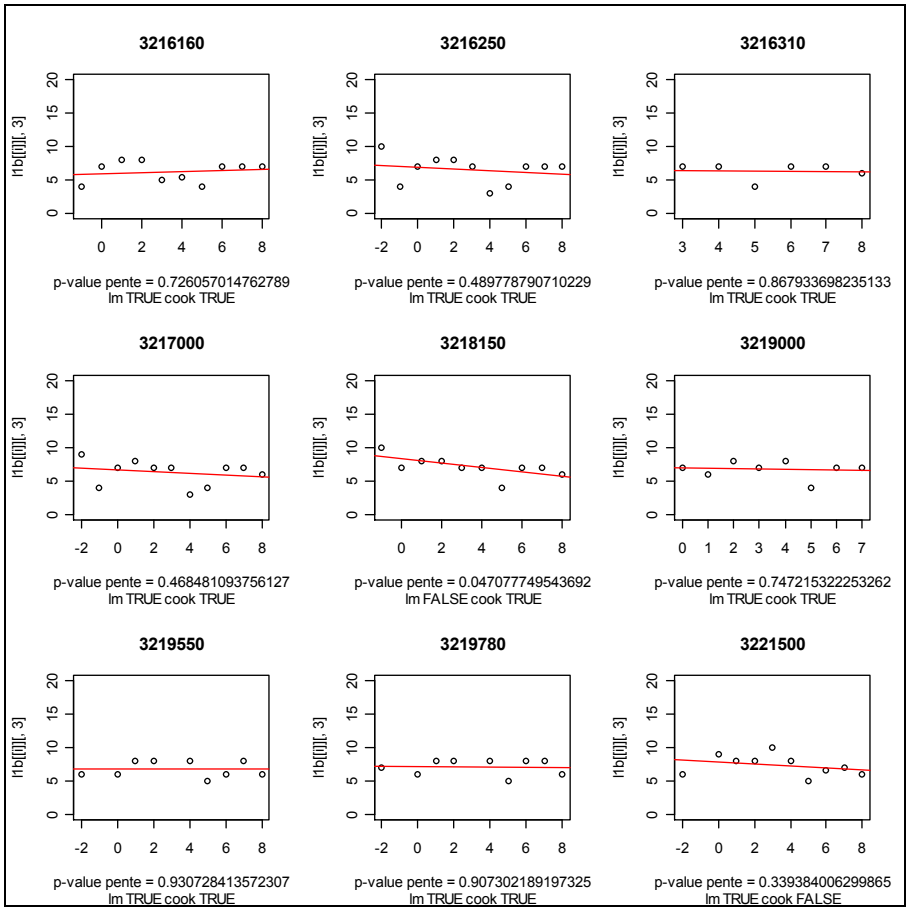


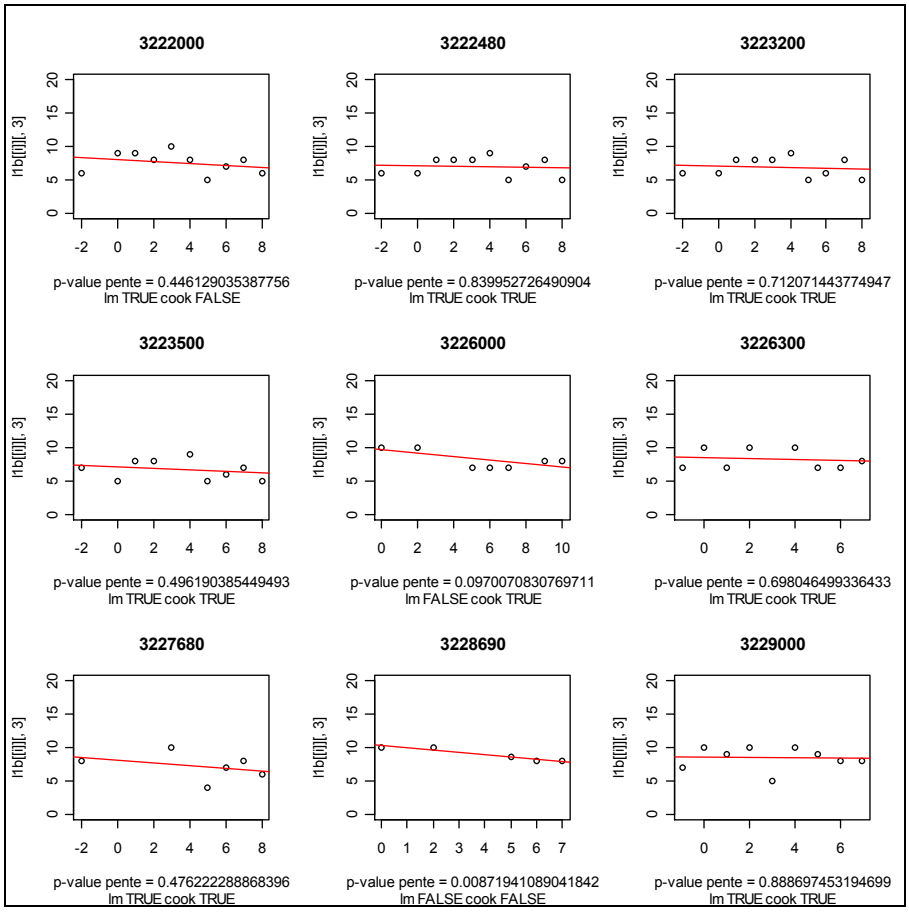
CemOA : archive ouverte d'Irstea / Cemagref



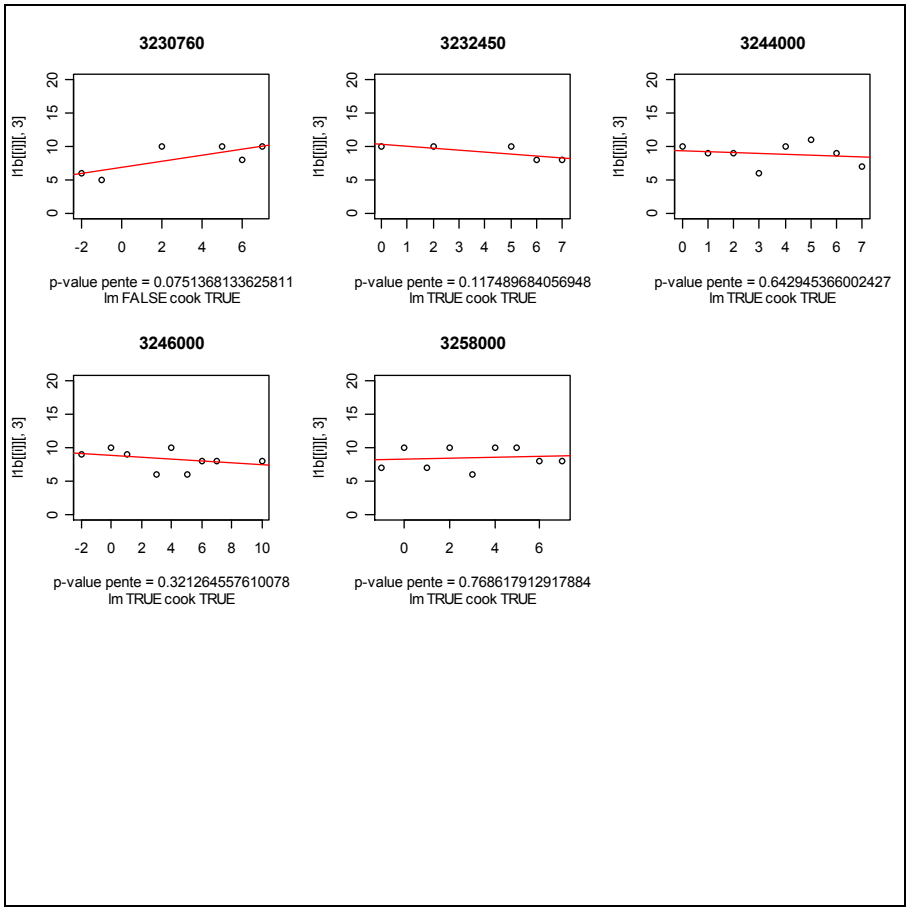


CemOA : archive ouverte d'Irstea / Cemagref





CemOA : archive ouverte d'Irstea / Cemagref



CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE XI. Extraction de l'information de l'occupation du sol dans les corridors rivulaires de l'HER9-NO à partir des images à THRS

L'extraction de l'occupation du sol dans le corridor rivulaire de la partie Nord-Ouest de l'Hydro-Ecorégion Tables Calcaires (HER9-NO) a été réalisée à l'aide de la procédure de classification en 7 étapes développée dans la partie 2 de ce mémoire. Cette annexe présente l'application des six premières étapes de la procédure. L'étape 7 (validation du résultat de classification), n'a pas pu être menée dans le cadre de la thèse.

Étape 1 : Sélection des données

La méthode repose à la fois sur a) des données images et b) des données auxiliaires.

a) Les données images

Deux sources de données images sont nécessaires : 1) les orthophotos (0.5 m) et les images satellites Spot5 XS (10m).

1) Les orthophotos

2455 orthophotos (Figure 1) ont été sélectionnés parmi 11 départements (le 14, 27, 28, 50, 60, 61, 76, 78, 80, 91, 95) pour recouvrir la zone d'étude. Ces images ont été acquises à des périodes différentes allant de 2002 jusqu'à 2006 (Figure 2).

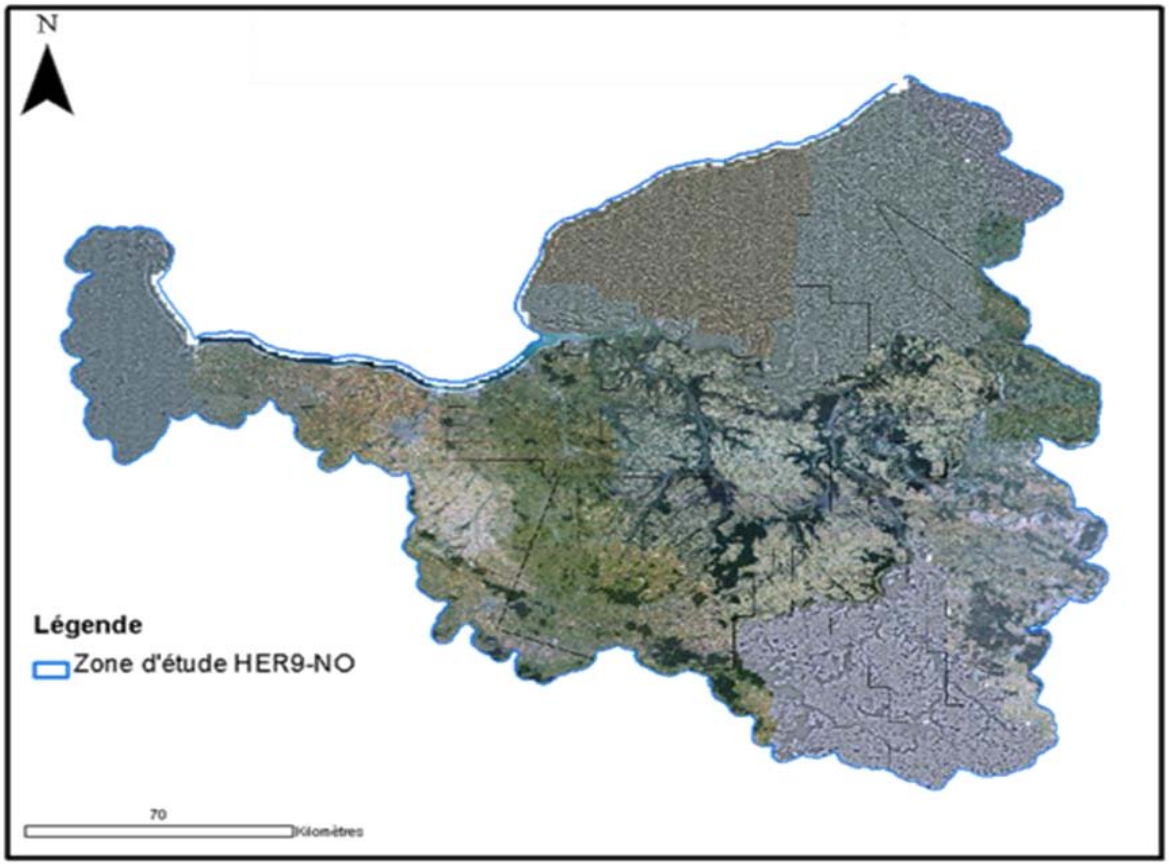


Figure 1 : 2455 orthophotos recouvrant la zone d'étude

CemOA : archive ouverte d'Irstea / Cemagref

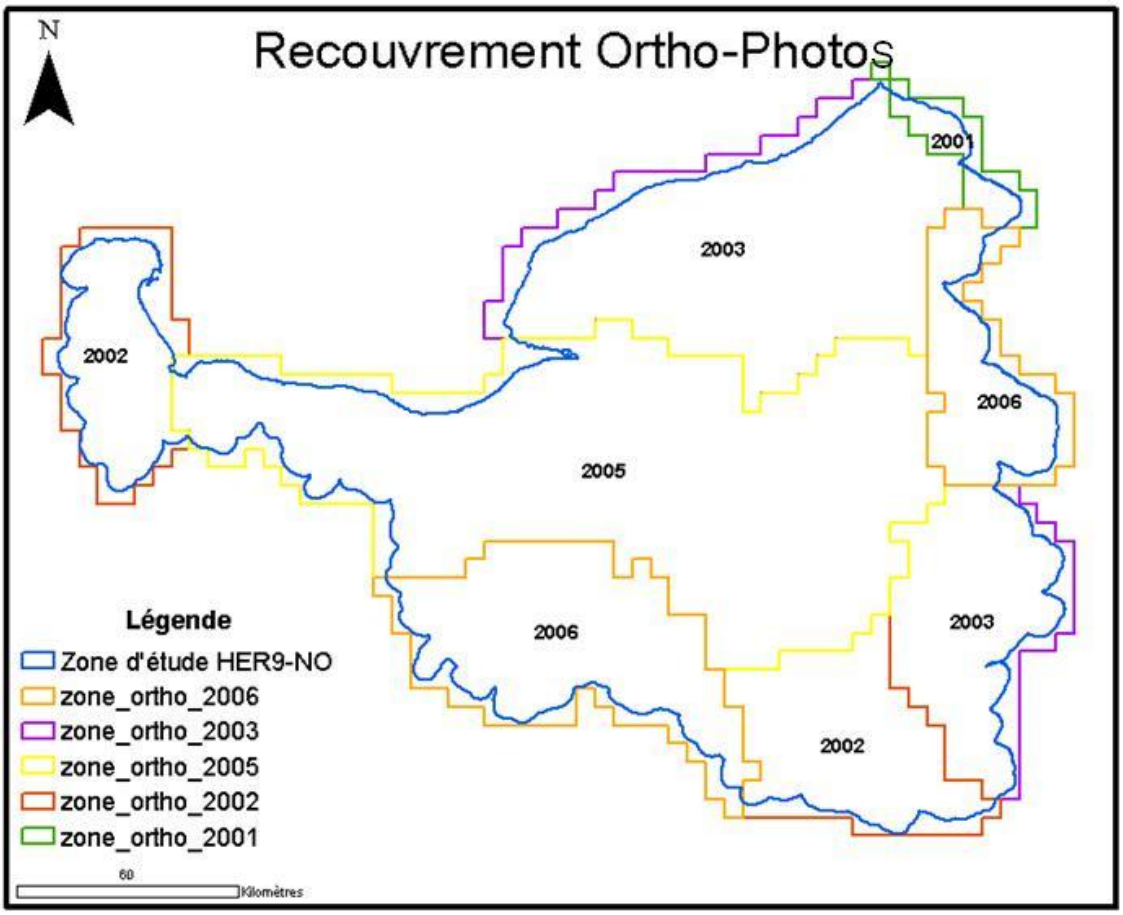


Figure 2 : Date d'acquisition des images orthophotos sur la zone d'étude

2) Les images Spot5 XS

11 images Spot5 XS d'archives (Figure 3) ont été sélectionnées pour recouvrir la zone d'étude. Ces images ont été acquises à des périodes différentes allant de 2003 jusqu'à 2007 (Tableau 1 et Figure 4). Le choix de ces images s'est fait en fonction de trois critères (dans l'ordre de priorité) : (1) la couverture nuageuse, (2) le recouvrement de la zone d'étude, (3) la date optimale pour séparer les prairies des cultures.

Tableau 1 : Date d'acquisition et coordonnées K_J des 11 images satellites Spot5 XS

Cordonnées K_J	Date d'acquisition
32_250	23/08/05
33_250	15/02/03
34_250	19/09/05
35_249	27/05/05
35_250	10/08/06
36_251	22/06/05
37_249	05/04/07
37_250	29/04/07
38_250	05/04/07
38_251	08/04/07
38_252	08/04/07

CemOA : archive ouverte d'Irstea / Cemagref

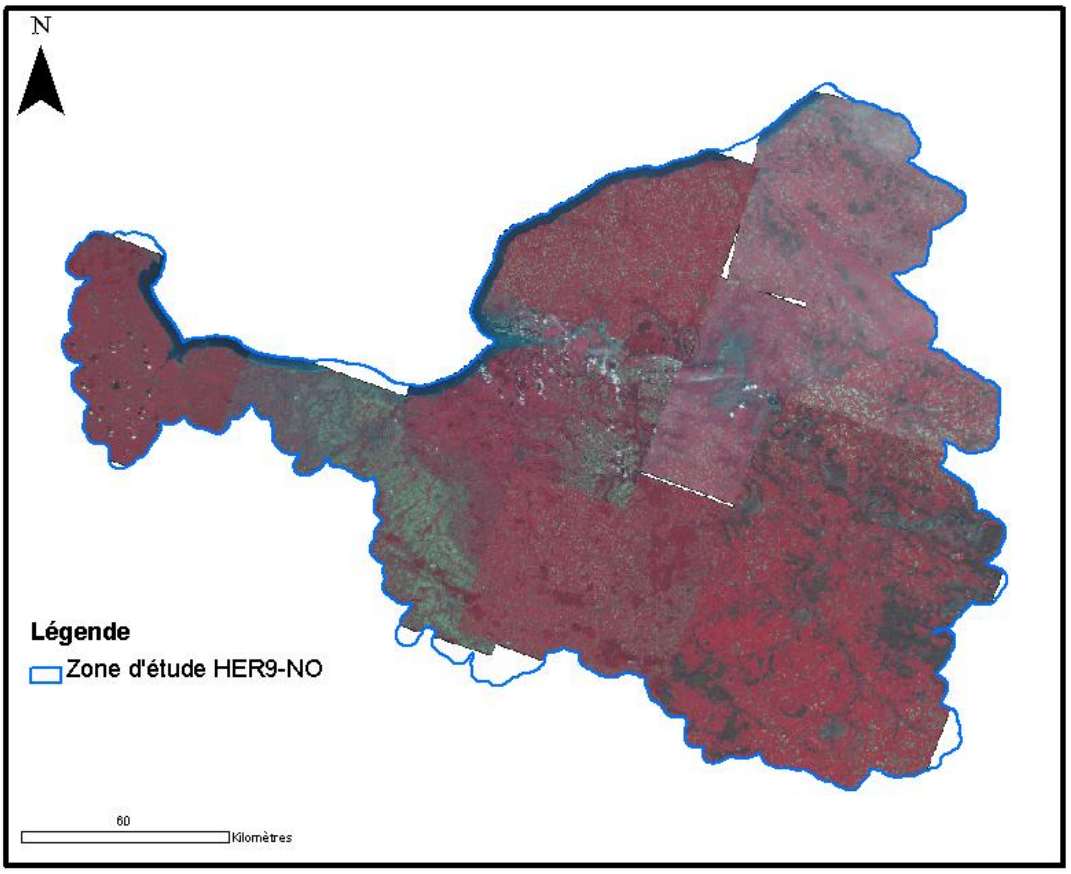


Figure 3 : 11 images Spot5 XS recouvrant la zone d'étude

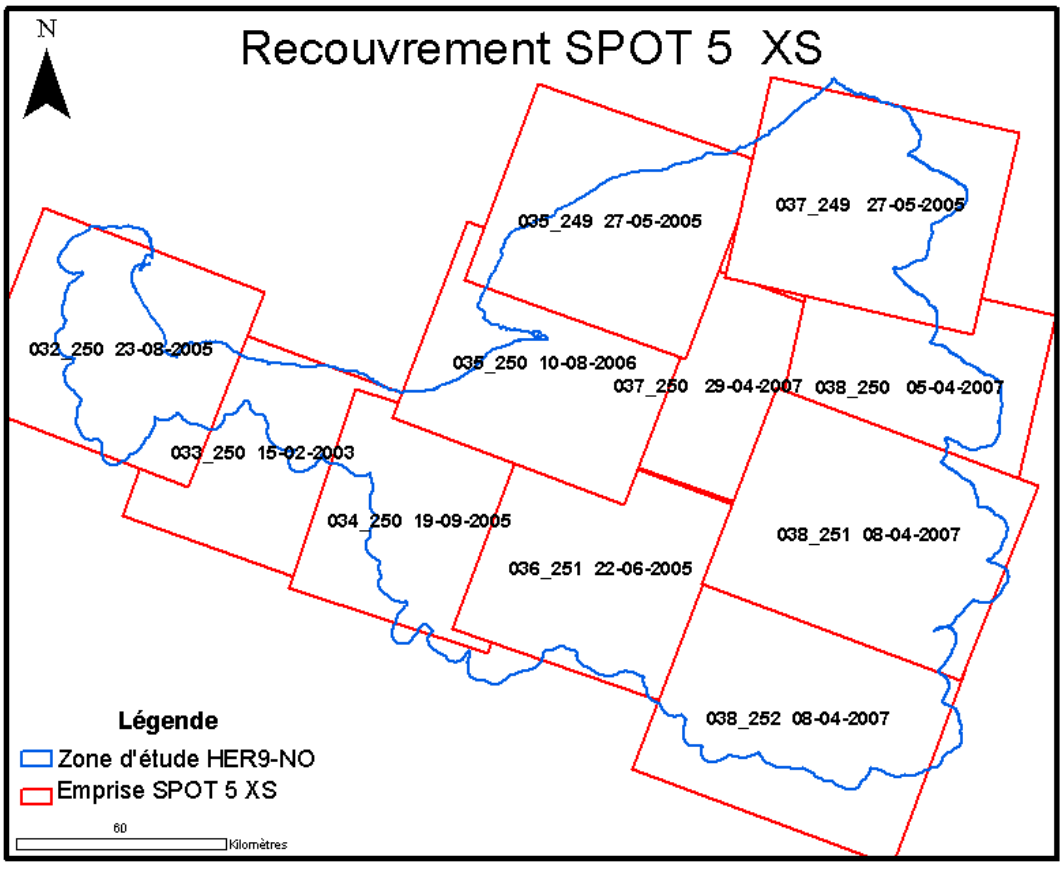


Figure 4 : Emprises des 11 images SPOT 5 XS d'archives acquises sur la zone d'étude

CemOA : archive ouverte d'Irstea / Cemagref

b) Les données auxiliaires

Les postes artificialisés de CORINE Land Cover (2000) (cf. ANNEXE V), les routes de la BDRoute 500® (2006), les entités surfaciques de la BDCarthage® (2006) et les parcelles de cultures éligibles à la Politique Agricole Commune du Registre Parcellaire Graphique (RPG 2007) ont été collectées sur la zone d'étude. Dans cette étude l'information sur le type de culture à l'échelle de l'îlot était disponible.

Étape 2 : Délimitation des corridors rivulaires

Les corridors rivulaires ont été délimités en traçant un buffer autour de la donnée hydrographique BDCarthage® niveau 2. La largeur du buffer dépend tout naturellement de la taille du cours d'eau (ordination de Strahler). Plus de 5600 Km² de corridor rivulaire ont été délimités sur la zone d'étude (Figure 5).

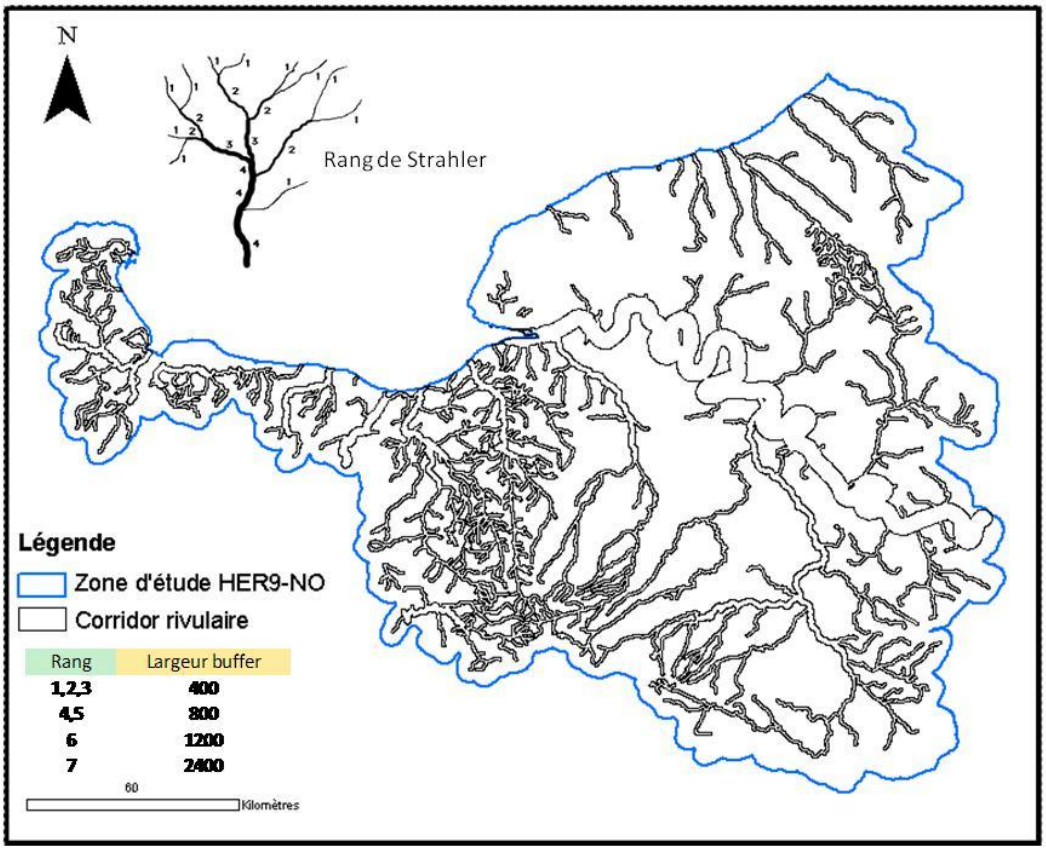


Figure 5 : Corridors rivulaires de la zone d'étude estimés à partir du rang de Strahler

Étape 3 : Stratification en zones homogènes

A partir des limites des HER niveau 2 (HER2, Figure 6), des dates d'acquisition des orthophotos (Figure 2) et des images Spot5 XS (Figure 4) sur la zone d'étude, 30 zones de type d'occupation du sol homogène ont été délimités (Figure 7).

CemOA : archive ouverte d'Irstea / Cemagref

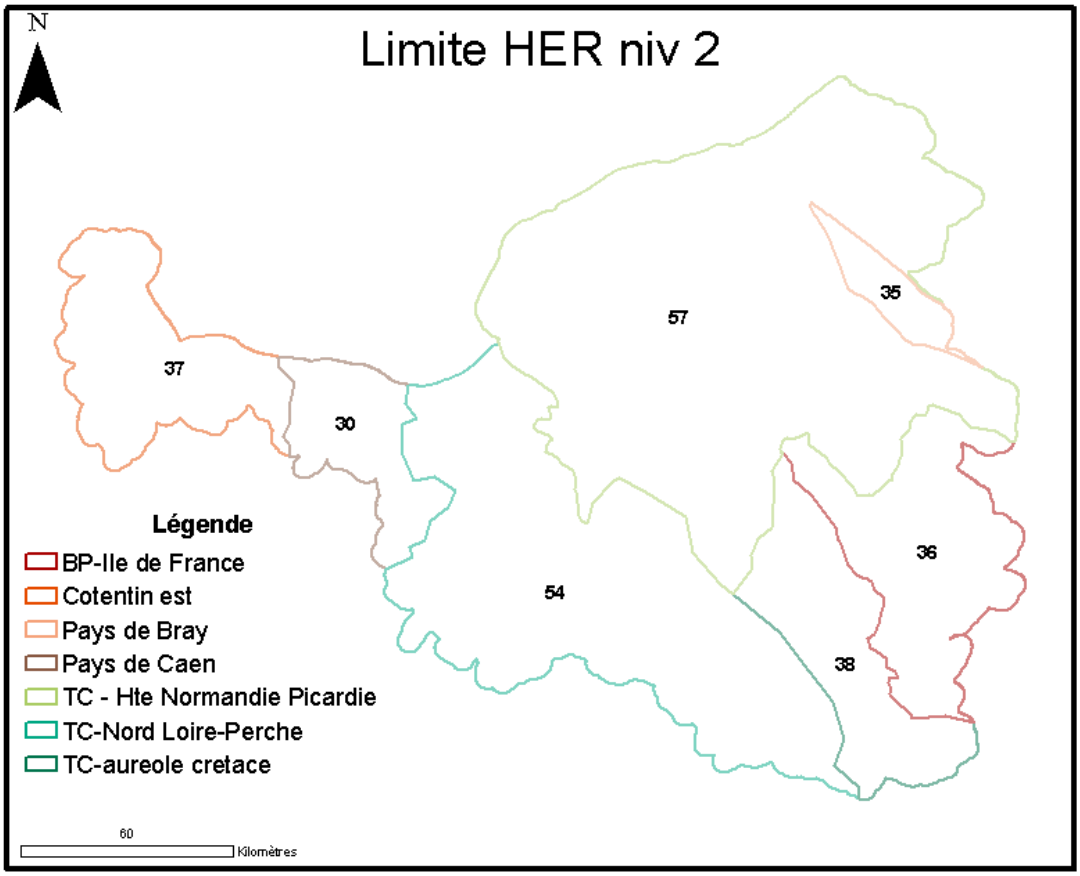


Figure 6 : Limites des HER2 de la zone d'étude

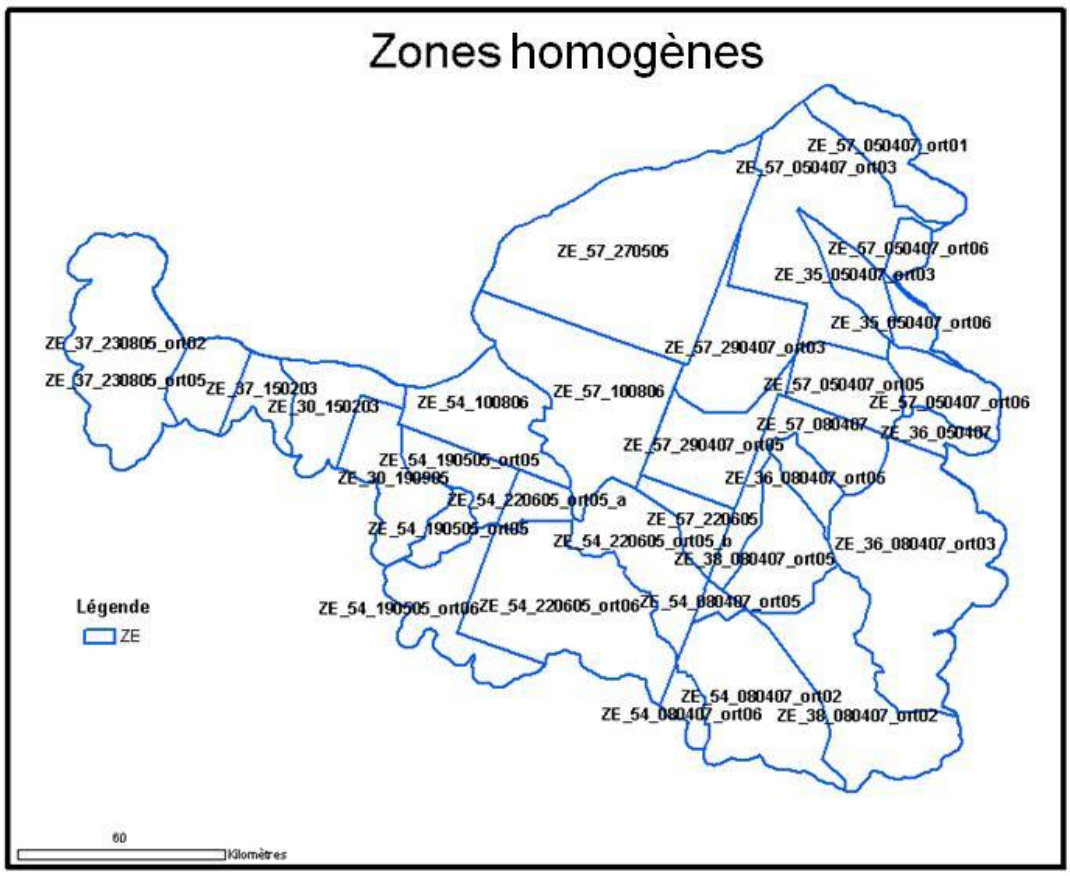


Figure 7: Délimitation des 30 zones homogènes

CemOA : archive ouverte d'Irstea / Cemagref

Étape 4 : Application de prétraitements

Les images Spot5 XS on été tout d'abord transformées en réflectance « Top of Atmosphere » (TOA). Ensuite ces images, ainsi que les orthophotos et les données auxiliaires ont été découpées et mosaïquées (si nécessaire) en fonction des limites dans chaque région homogène et dans les corridors rivulaires (Figure 8). Enfin, pour chaque zone homogène, une mosaïque orthophoto dégradée à 6 m et un plan thématique raster (0.5 m) construit à partir des données auxiliaires ont été créés.

Étape 5 : Classification dans chaque région

La démarche de classification orientée objet en trois étapes a été appliquée sur chaque région homogène. L'arbre de classification utilisé pour chacune de ces régions est présentée Figure 9.

Étape 6 : Application de post-traitements

Les résultats d'occupation du sol issus du traitement orienté objet de chaque région homogène sont réunis de manière à obtenir une donnée continue sur l'ensemble de la zone d'étude. Cette donnée continue est ensuite structurée dans une nomenclature hiérarchisée à quatre niveaux (cf. ANNEXE IV.2).

CemOA : archive ouverte d'Irstea / Cemagref

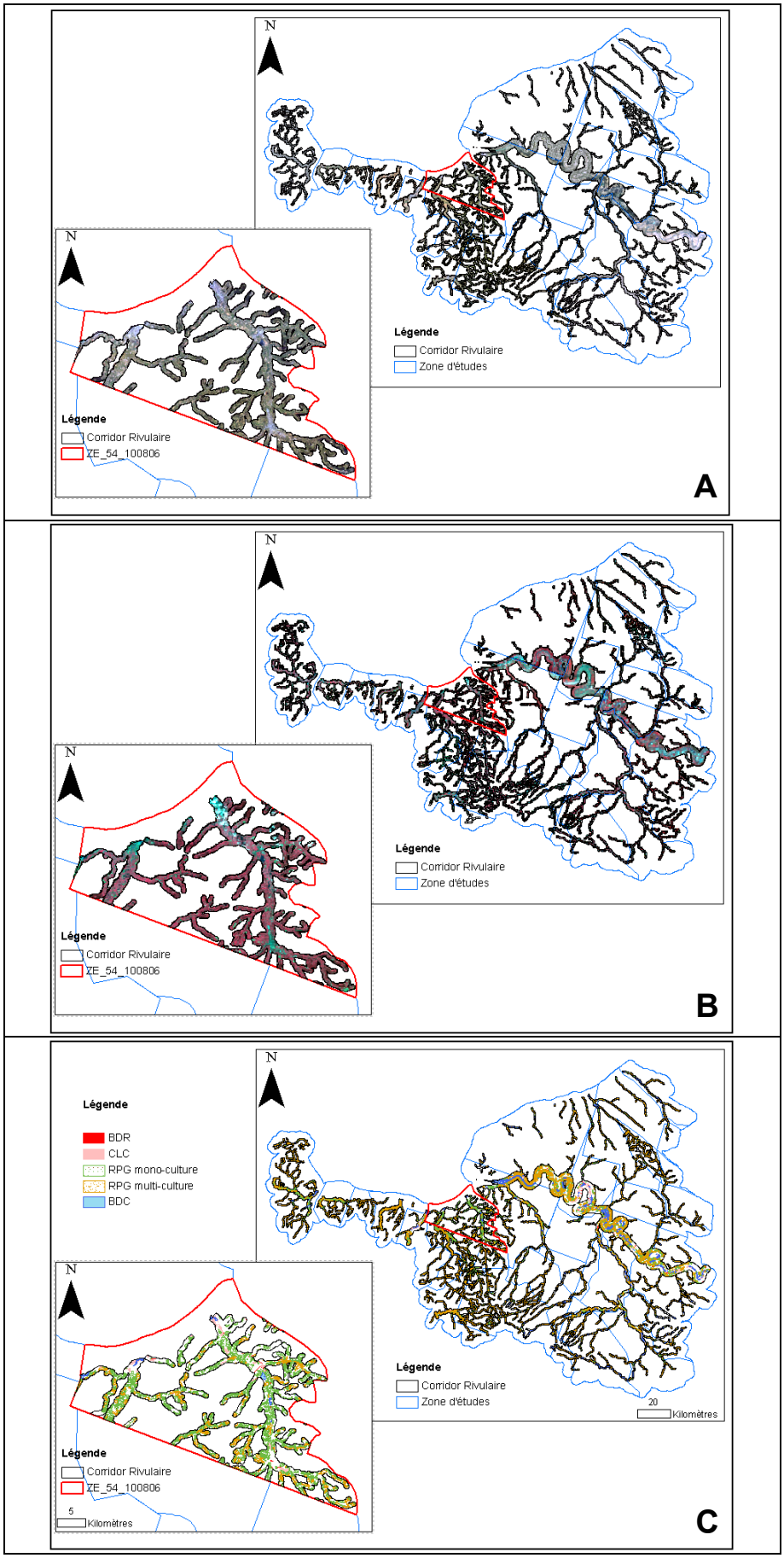


Figure 8 : Découpage des sources de données dans les zones homogènes et dans les corridors rivulaires (en A les orthophotos, en B les images Spot5 XS en TOA, et en C les données auxiliaires)

CemOA : archive ouverte d'Irstea / Cemagref

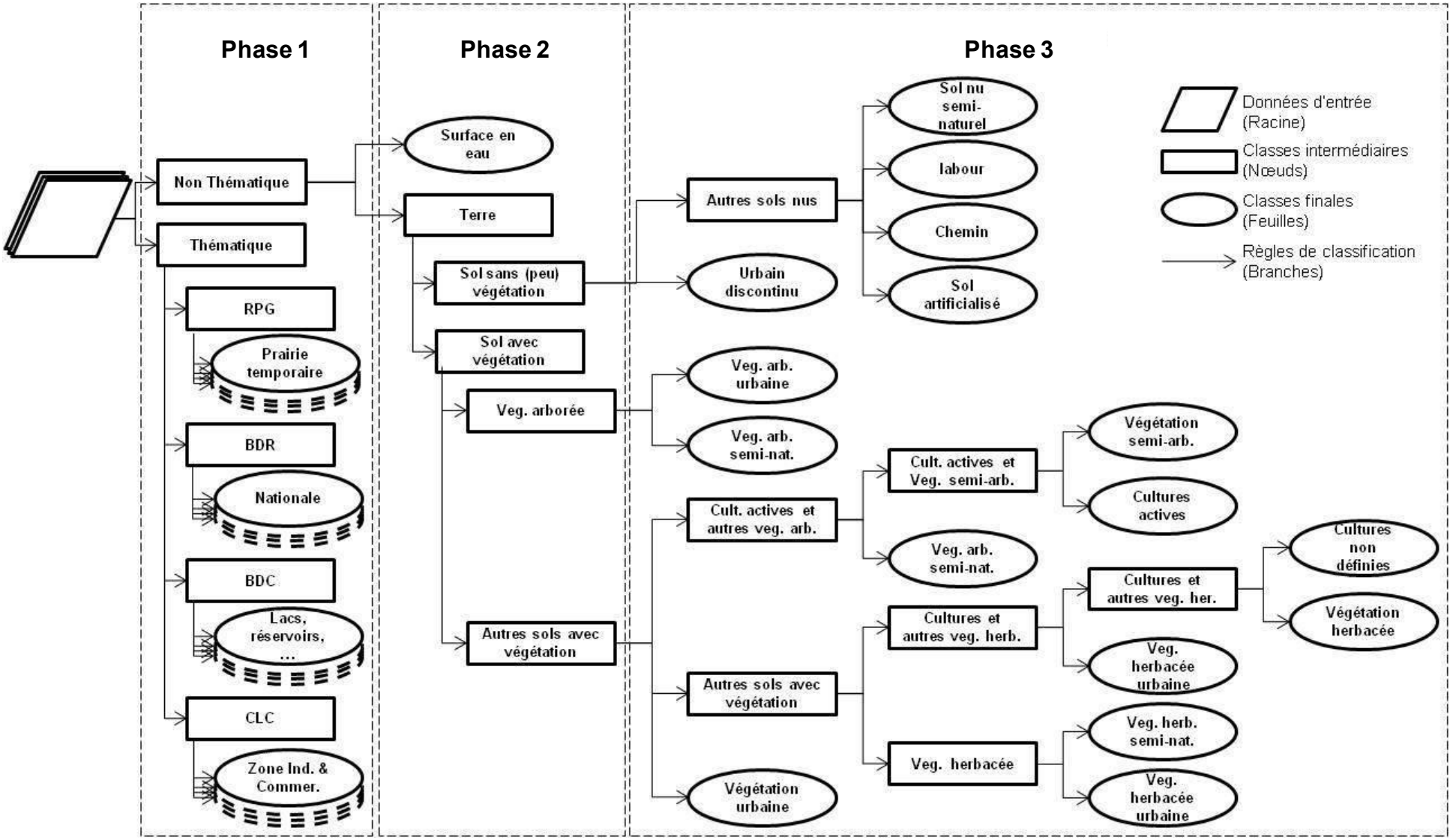


Figure 9 : Arbre de classification utilisé pour caractériser l'occupation du sol dans le corridor rivulaire de l'HER9-NO

CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE XII. Résultats détaillés des 21 modèles pressions/impacts développés sur l'HER9-NO

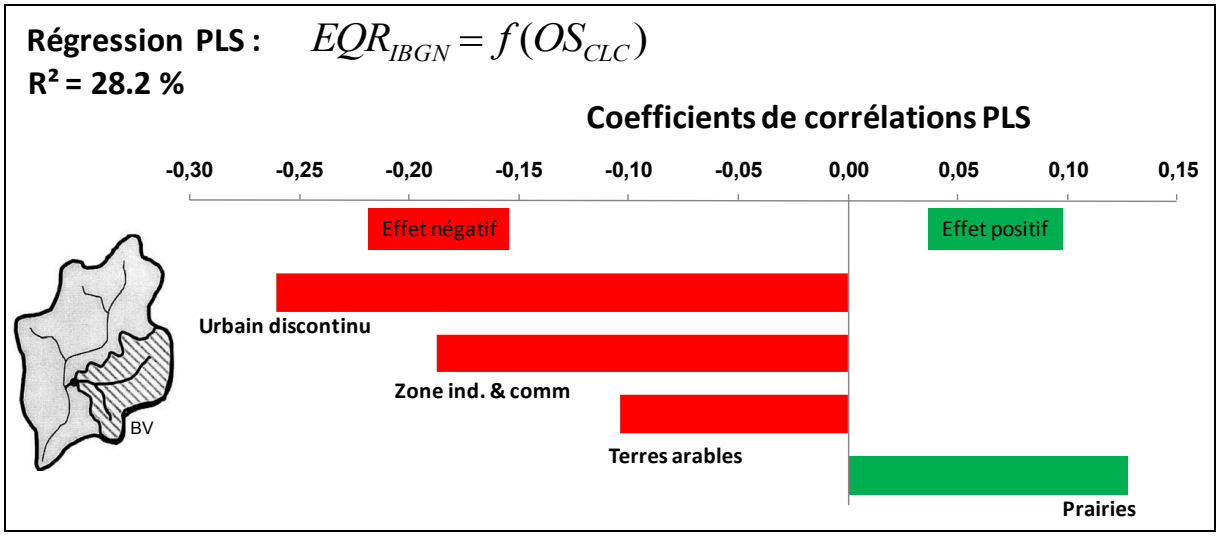
ANNEXE XII.1 *Modèles mono-échelle*

ANNEXE XII.1.1 Modèle BV (Mod-1.1.1)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,261	0,037	-0,334	-0,187
2	BV-121	-0,188	0,051	-0,288	-0,087
3	BV-231	0,127	0,051	0,028	0,227
4	BV-211	-0,104	0,045	-0,192	-0,016

Illustration de la contribution des sources de pressions significatives dans ce modèle



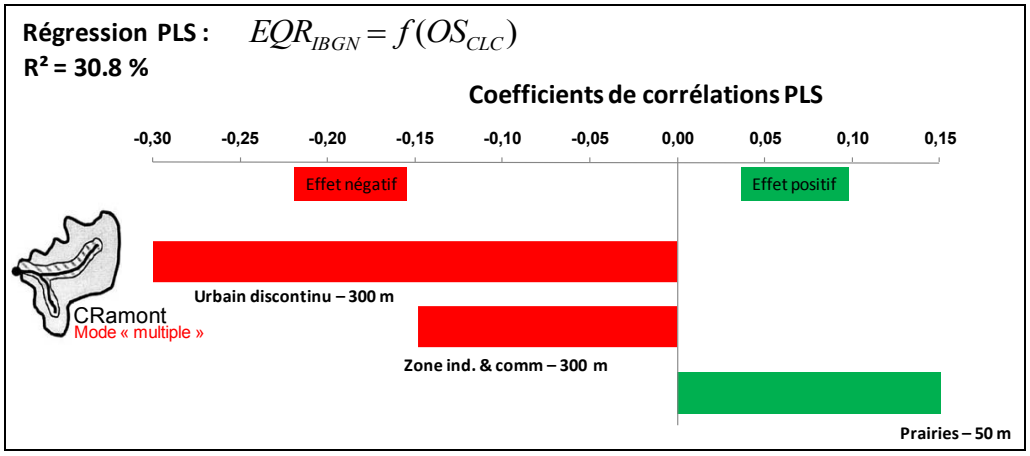
ANNEXE XII.1.2 CRamont-CLC (Mod-1.2.1)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	CR-300-amont-112	-0,316	0,116	-0,543	-0,089
2	CR-50-amont-231	0,252	0,074	0,106	0,397
3	CR-300-amont-121	-0,149	0,068	-0,281	-0,016

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle

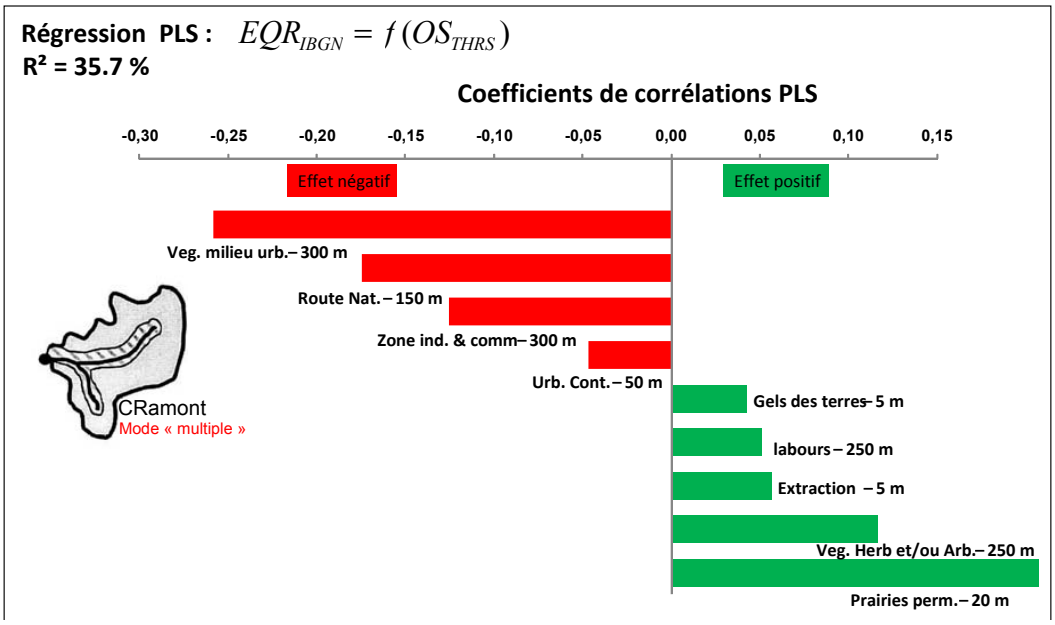


ANNEXE XII.1.3 CRamont-THRS (Mod-1.2.2)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	CR-300-amont-113	-0,258	0,076	-0,408	-0,108
2	CR-20-amont-231	0,207	0,052	0,106	0,308
3	CR-150-amont-152	-0,175	0,048	-0,268	-0,081
4	CR-300-amont-121	-0,126	0,057	-0,237	-0,014
5	CR-250-amont-321	0,117	0,034	0,050	0,184
6	CR-5-amont-133	0,057	0,026	0,006	0,108
7	CR-250-amont-211	0,051	0,025	0,003	0,100
8	CR-50-amont-111	-0,047	0,023	-0,091	-0,002
9	CR-5-amont-252	0,043	0,019	0,006	0,079

Illustration de la contribution des sources de pressions significatives dans ce modèle



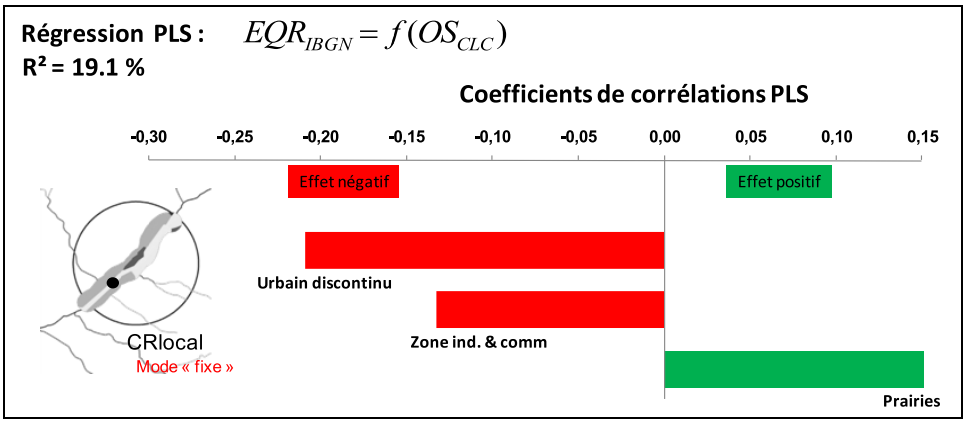
CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE XII.1.4 CRlocal-fixe-CLC (Mod- 1.3.1)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	CRstr-231	0,224	0,055	0,117	0,331
2	CRstr-112	-0,209	0,098	-0,402	-0,017
3	CRstr-121	-0,132	0,020	-0,171	-0,094

Illustration de la contribution des sources de pressions significatives dans ce modèle



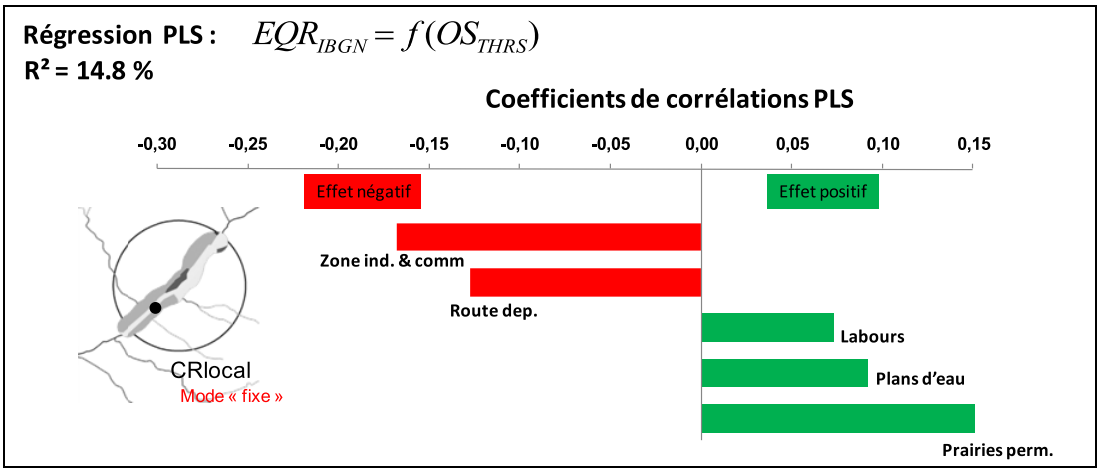
ANNEXE XII.1.5 CRlocal-fixe-THRS (Mod- 1.3.2)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	CRstr-231	0,233	0,103	0,031	0,435
2	CRstr-121	-0,168	0,026	-0,219	-0,118
3	CRstr-153	-0,128	0,058	-0,241	-0,014
4	CRstr-412	0,092	0,036	0,020	0,163
5	CRstr-211	0,073	0,032	0,012	0,135

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle

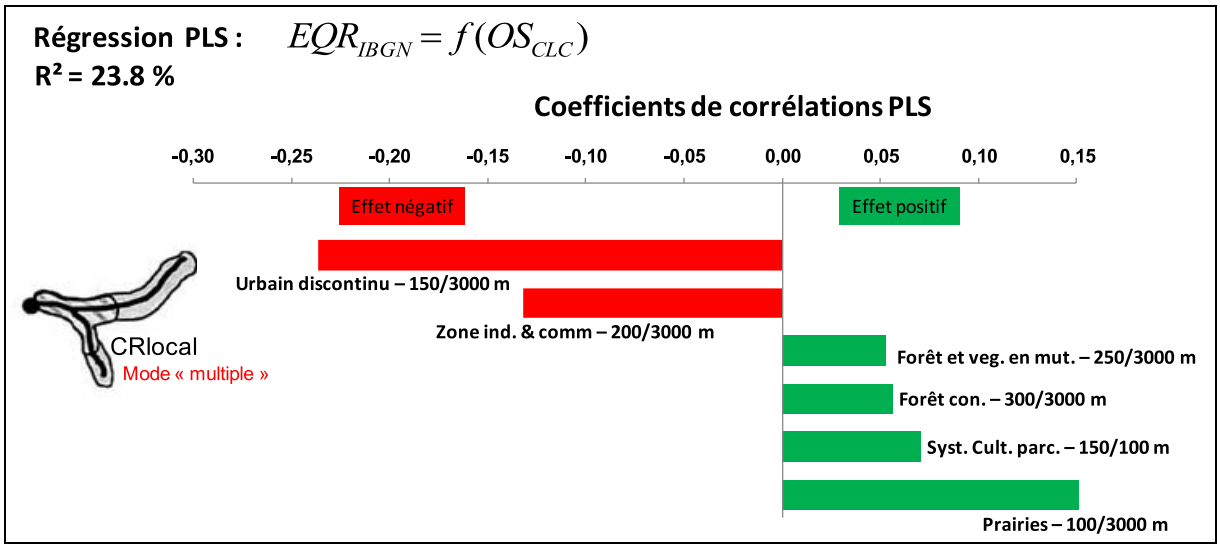


ANNEXE XII.1.6 CRlocal-mult-CLC (Mod-1.3.3)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	CR-150-3000-112	-0,236	0,077	-0,387	-0,086
2	CR-100-3000-231	0,234	0,053	0,130	0,339
3	CR-200-3000-121	-0,132	0,027	-0,185	-0,079
4	CR-150-100-242	0,071	0,035	0,002	0,139
5	CR-300-1000-312	0,056	0,021	0,016	0,096
6	CR-250-3000-324	0,053	0,012	0,030	0,076

Illustration de la contribution des sources de pressions significatives dans ce modèle



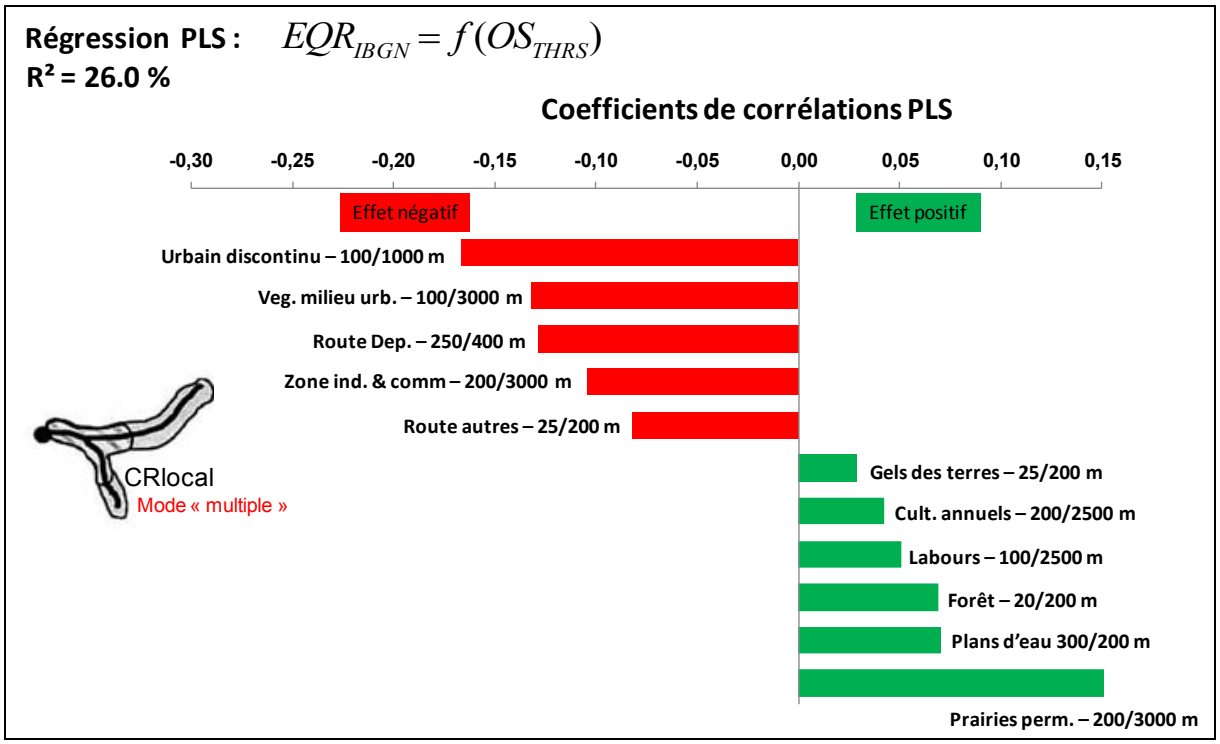
CemOA : archive ouverte d'Irstea / Cemagref

ANNEXE XII.1.7 CRlocal-mult-THRS (Mod-1.3.4)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	CR-100-1000-112	-0,167	0,070	-0,303	-0,030
2	CR-200-3000-231	0,165	0,069	0,031	0,300
3	CR-100-3000-113	-0,132	0,053	-0,235	-0,028
4	CR-250-400-153	-0,129	0,023	-0,175	-0,083
5	CR-200-3000-121	-0,105	0,019	-0,142	-0,067
6	CR-150-300-154	-0,082	0,019	-0,120	-0,045
7	CR-300-200-412	0,071	0,020	0,031	0,110
8	CR-20-200-311	0,069	0,034	0,002	0,136
9	CR-100-2500-211	0,051	0,008	0,035	0,067
10	CR-200-2500-214	0,043	0,019	0,006	0,080
11	CR-25-200-251	0,029	0,012	0,006	0,051

Illustration de la contribution des sources de pressions significatives dans ce modèle



CemOA : archive ouverte d'Irstea / Cemagref

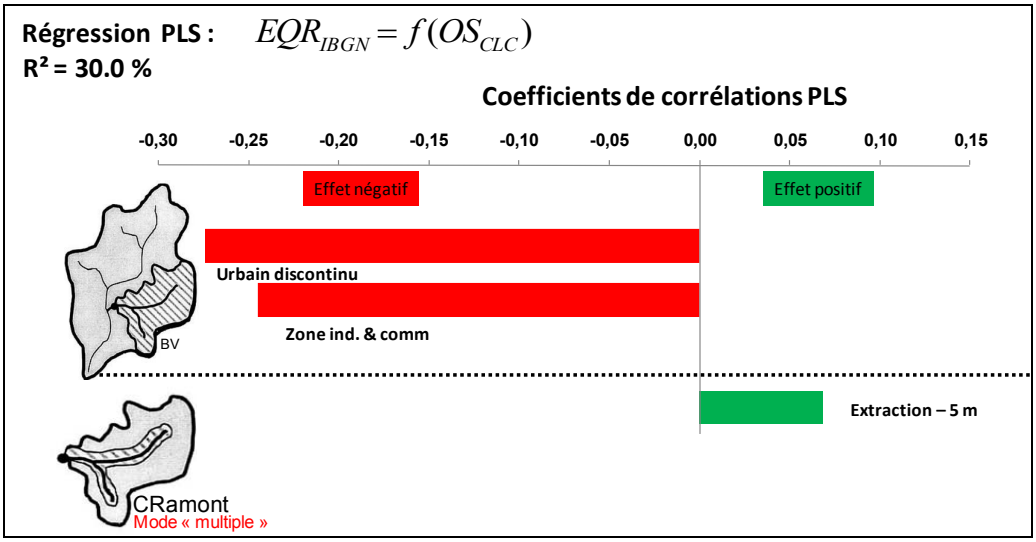
ANNEXE XII.2 *Modèles bi-échelle*

ANNEXE XII.2.1 BV/CRamont-CLC (Mod-2.1.1)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,340	0,073	-0,483	-0,197
2	BV-121	-0,245	0,034	-0,312	-0,177
3	CR-5-amont-131	0,069	0,034	0,002	0,136

Illustration de la contribution des sources de pressions significatives dans ce modèle



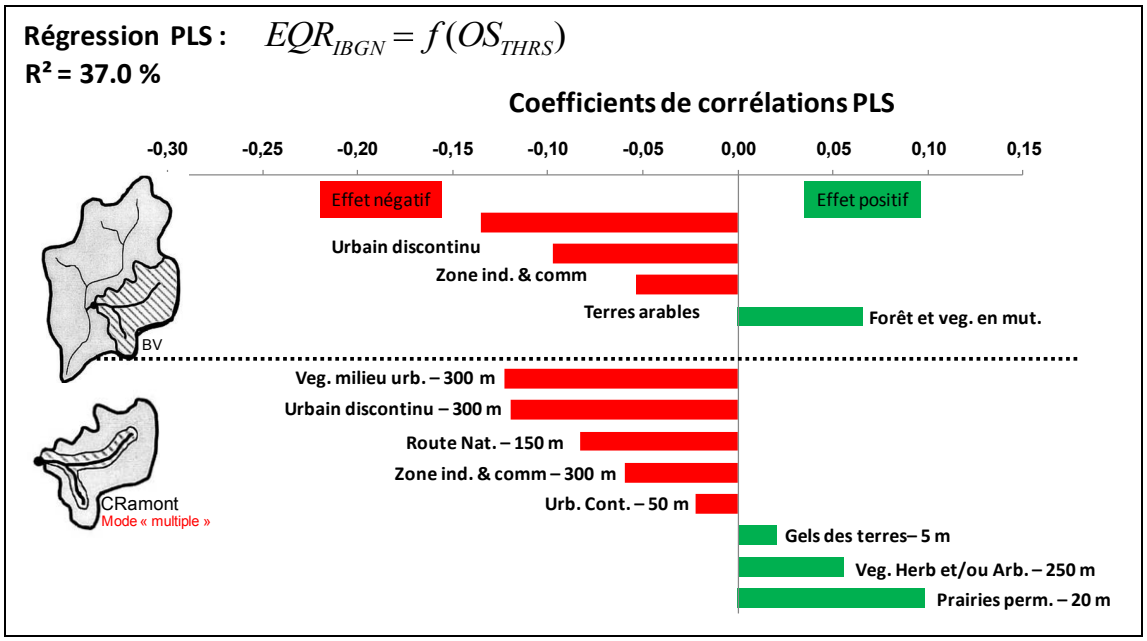
ANNEXE XII.2.2 BV/CRamont-THRS (Mod-2.1.2)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,135	0,015	-0,165	-0,105
2	CR-300-amont-113	-0,123	0,036	-0,194	-0,052
3	CR-300-amont-112	-0,119	0,039	-0,197	-0,042
4	CR-20-amont-231	0,098	0,024	0,052	0,145
5	BV-121	-0,097	0,022	-0,141	-0,053
6	CR-150-amont-152	-0,083	0,029	-0,140	-0,027
7	BV-231	0,066	0,030	0,007	0,125
8	CR-300-amont-121	-0,060	0,028	-0,115	-0,005
9	CR-250-amont-321	0,056	0,016	0,023	0,088
10	BV-211	-0,054	0,026	-0,104	-0,003
11	CR-50-amont-111	-0,022	0,011	-0,043	-0,001
12	CR-5-amont-252	0,020	0,010	0,001	0,040

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle



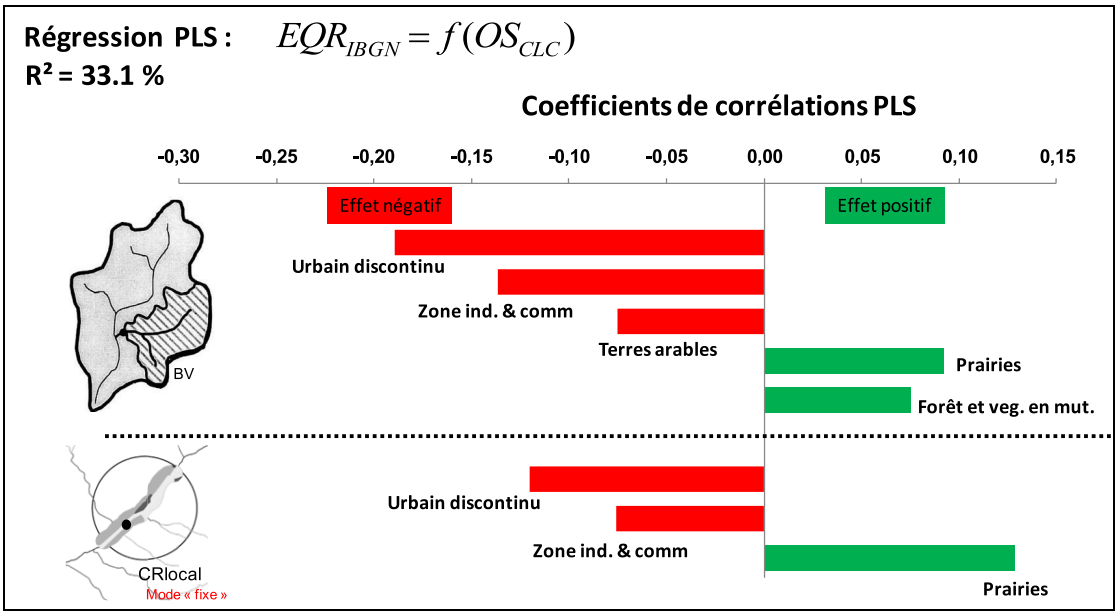
ANNEXE XII.2.3 BV/CRlocal-fixe-CLC (Mod-2.2.1)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,189	0,024	-0,236	-0,143
2	BV-121	-0,136	0,048	-0,231	-0,041
3	CRstr-231	0,129	0,022	0,085	0,172
4	CRstr-112	-0,120	0,056	-0,229	-0,011
5	BV-231	0,092	0,039	0,016	0,168
6	CRstr-121	-0,076	0,007	-0,089	-0,063
7	BV-211	-0,075	0,030	-0,133	-0,017
8	BV-243	0,075	0,037	0,002	0,147

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle



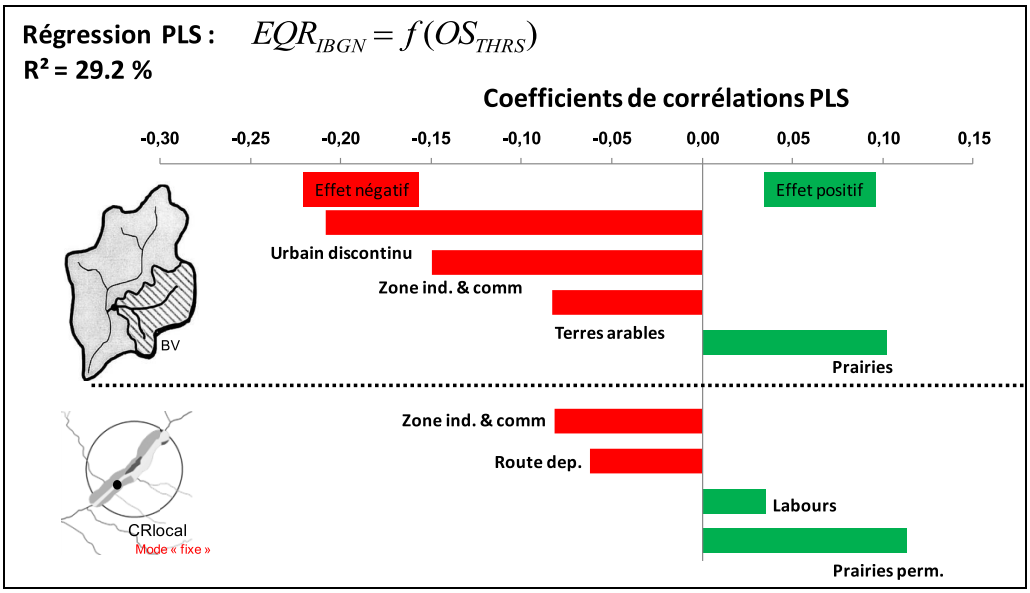
ANNEXE XII.2.4 BV/CRlocal-fixe-THRS (Mod-2.2.2)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

Ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,208	0,029	-0,265	-0,151
2	BV-121	-0,150	0,051	-0,250	-0,050
3	CRstr-231	0,113	0,039	0,037	0,188
4	BV-231	0,102	0,042	0,020	0,183
5	BV-211	-0,083	0,033	-0,148	-0,017
6	CRstr-121	-0,081	0,007	-0,095	-0,068
7	CRstr-153	-0,062	0,027	-0,115	-0,009
8	CRstr-211	0,035	0,016	0,004	0,067

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle



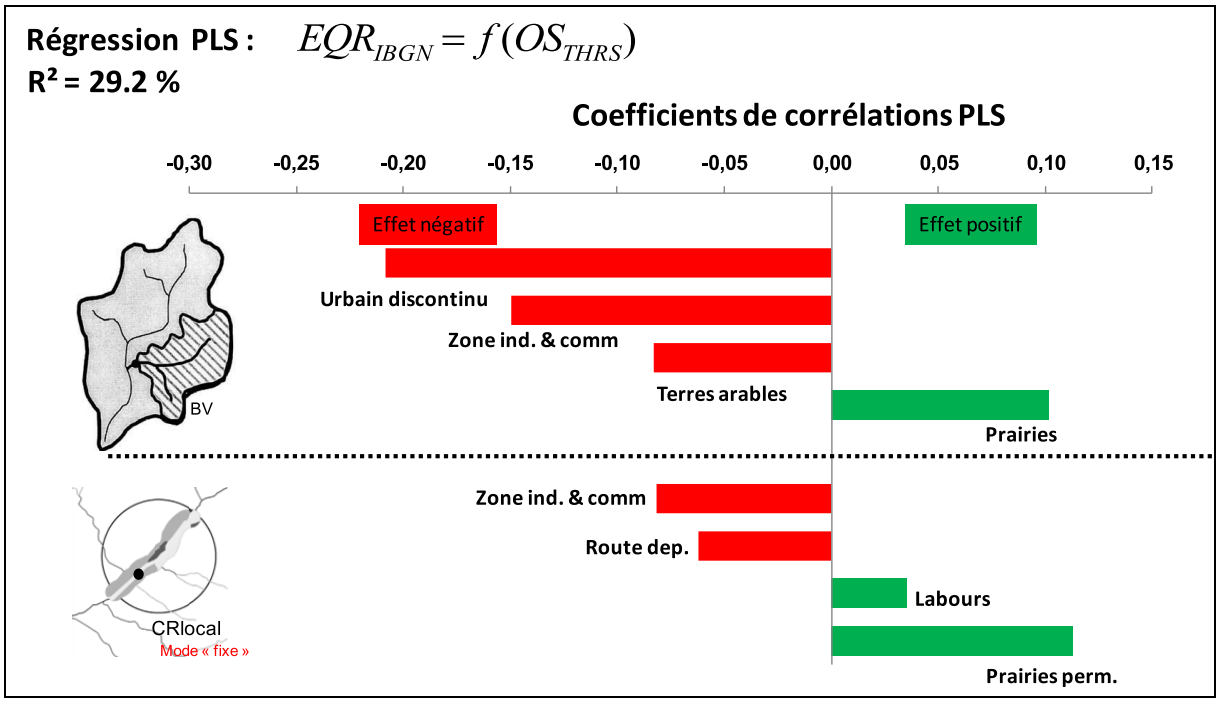
ANNEXE XII.2.5 BV/CRlocal-mult-CLC (Mod-2.2.3)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,229	0,011	-0,250	-0,208
2	BV-121	-0,165	0,051	-0,265	-0,064
3	CR-150-3000-112	-0,163	0,053	-0,267	-0,059
4	CR-100-3000-231	0,162	0,036	0,091	0,232
5	CR-200-3000-121	-0,091	0,020	-0,130	-0,052
6	BV-243	0,090	0,043	0,006	0,175
7	CR-150-100-242	0,049	0,024	0,001	0,097
8	CR-300-1000-312	0,039	0,013	0,014	0,064
9	CR-250-3000-324	0,037	0,009	0,019	0,054

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle



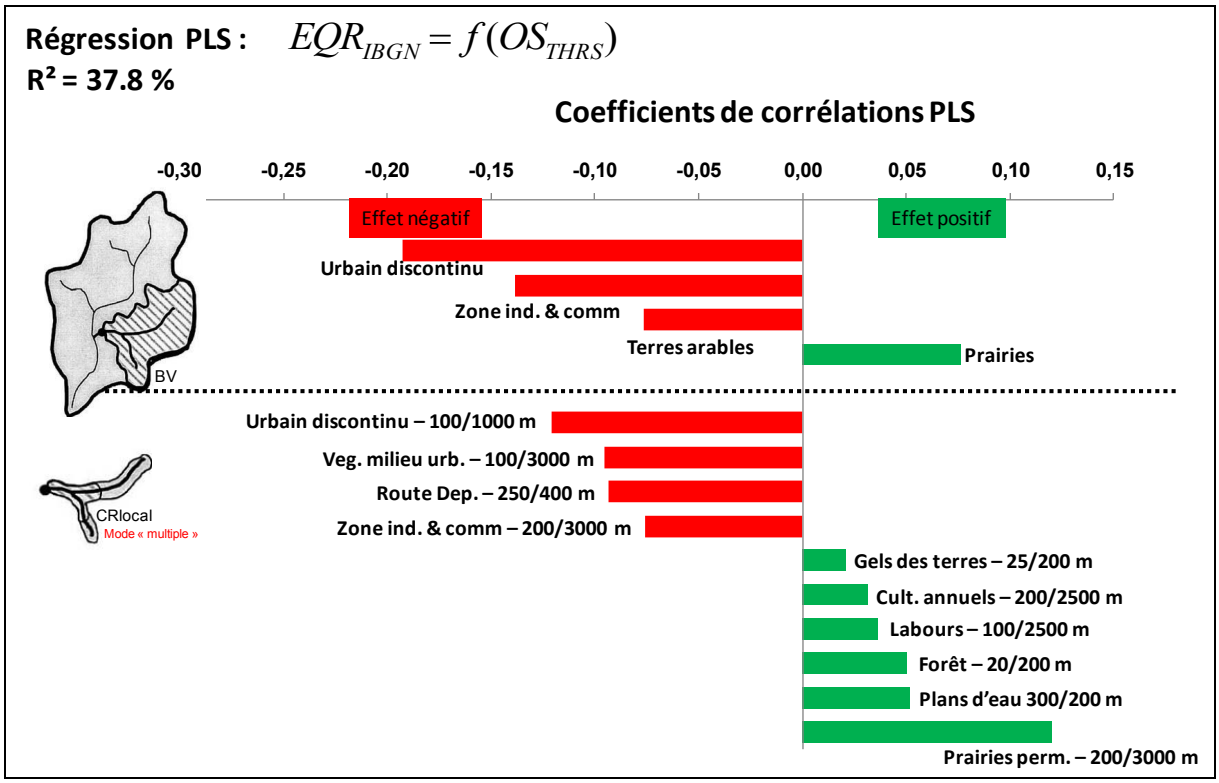
ANNEXE XII.2.6 BV/CRlocal-mult-THRS (Mod-2.2.4)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,193	0,016	-0,223	-0,162
2	BV-121	-0,139	0,047	-0,231	-0,046
3	CR-100-1000-112	-0,121	0,049	-0,216	-0,025
4	CR-200-3000-231	0,120	0,045	0,032	0,208
5	CR-100-3000-113	-0,096	0,039	-0,171	-0,020
6	CR-250-400-153	-0,093	0,015	-0,122	-0,065
7	BV-211	-0,077	0,034	-0,143	-0,010
8	BV-243	0,076	0,036	0,005	0,147
9	CR-200-3000-121	-0,076	0,014	-0,104	-0,048
10	CR-300-200-412	0,051	0,016	0,019	0,084
11	CR-20-200-311	0,050	0,026	0,000	0,101
12	CR-100-2500-211	0,037	0,006	0,025	0,049
13	CR-200-2500-214	0,031	0,014	0,004	0,058
14	CR-25-200-251	0,021	0,009	0,003	0,039

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle



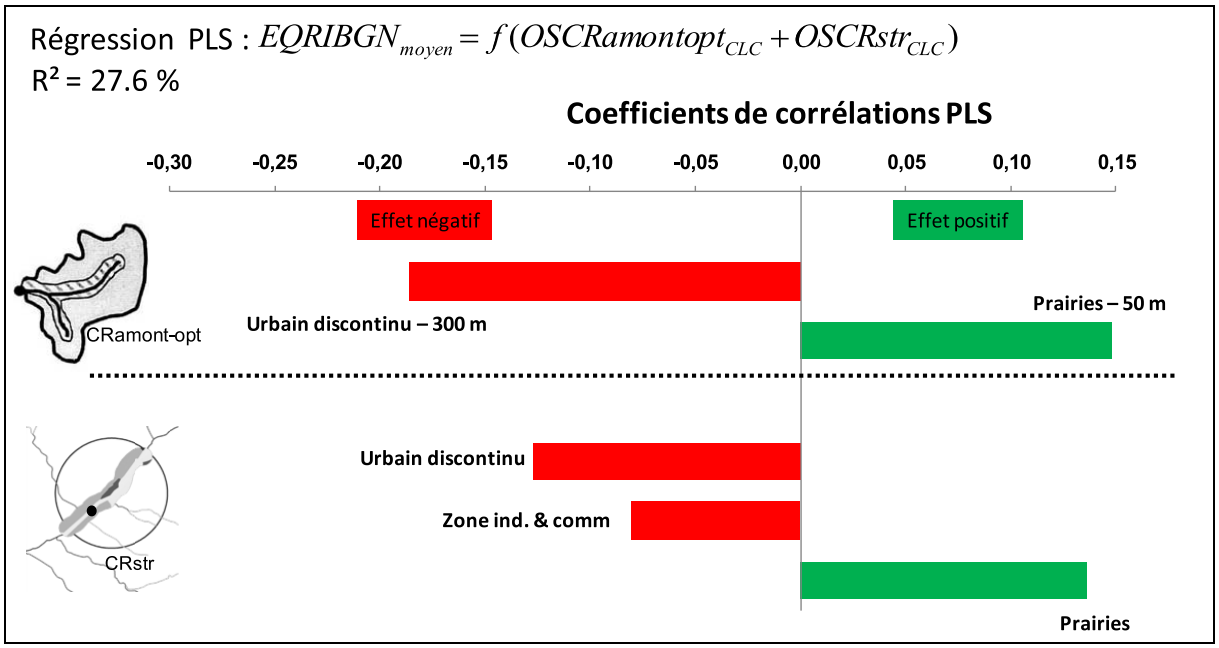
ANNEXE XII.2.7 CRamont /CRlocal-fixe-CLC (Mod-2.3.1)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	CR-300-amont-112	-0,186	0,051	-0,286	-0,086
2	CR-50-amont-231	0,148	0,038	0,074	0,222
3	CRstr-231	0,136	0,032	0,073	0,200
4	CRstr-112	-0,127	0,062	-0,249	-0,005
5	CRstr-121	-0,080	0,012	-0,104	-0,057

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle



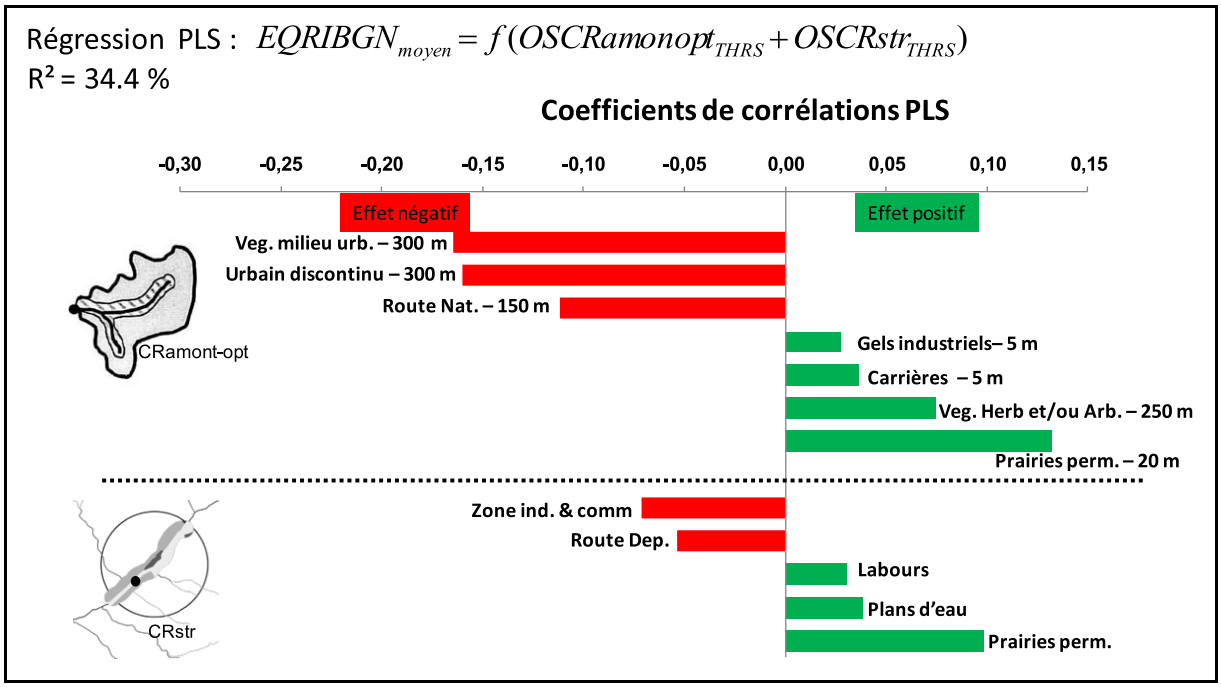
ANNEXE XII.2.8 CRamont /CRlocal-fixe -CLC (Mod-2.3.2)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	CR-300-amont-113	-0,165	0,029	-0,221	-0,108
2	CR-300-amont-112	-0,160	0,038	-0,234	-0,087
3	CR-20-amont-231	0,132	0,027	0,080	0,184
4	CR-150-amont-152	-0,111	0,030	-0,170	-0,053
5	CRstr-231	0,098	0,037	0,026	0,171
6	CR-250-amont-321	0,075	0,027	0,022	0,128
7	CRstr-121	-0,071	0,008	-0,087	-0,055
8	CRstr-153	-0,054	0,023	-0,099	-0,009
9	CRstr-412	0,039	0,016	0,007	0,070
10	CR-5-amont-133	0,036	0,018	0,000	0,073
11	CRstr-211	0,031	0,014	0,003	0,058
12	CR-5-amont-252	0,027	0,011	0,005	0,050

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle



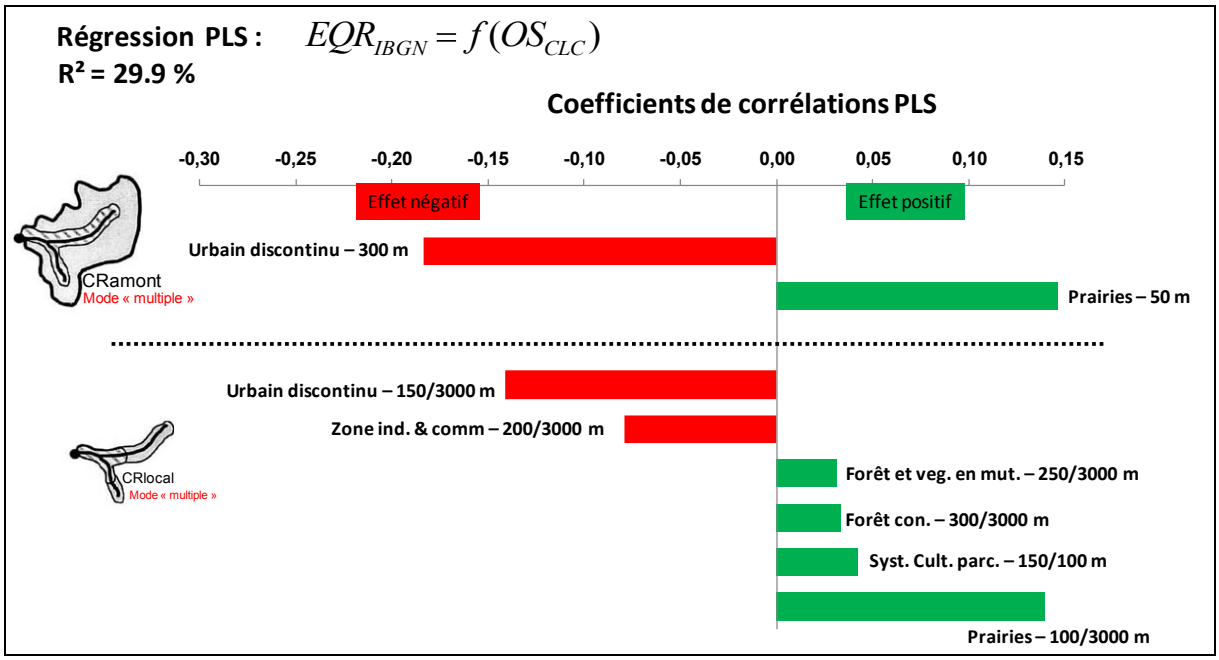
ANNEXE XII.2.9 CRamont /CRlocal-mult-CLC (Mod-2.3.3)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	CR-300-amont-112	-0,184	0,055	-0,291	-0,076
2	CR-50-amont-231	0,146	0,038	0,072	0,220
3	CR-150-3000-112	-0,141	0,051	-0,240	-0,042
4	CR-100-3000-231	0,140	0,035	0,072	0,208
5	CR-200-3000-121	-0,079	0,019	-0,116	-0,041
6	CR-150-100-242	0,042	0,020	0,002	0,082
7	CR-300-1000-312	0,033	0,011	0,011	0,056
8	CR-250-3000-324	0,032	0,008	0,016	0,047

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle



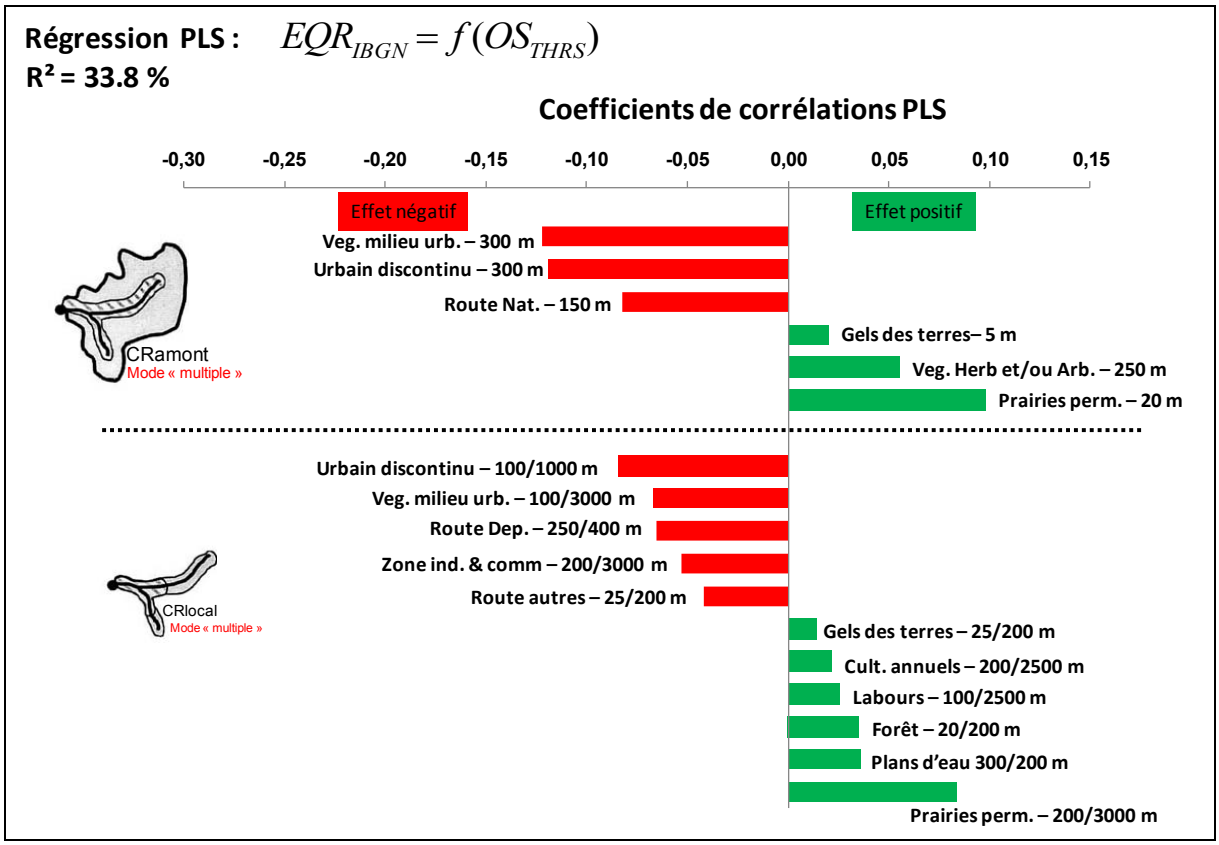
ANNEXE XII.2.10 CRamont /CRlocal-mult-THRS (Mod-2.3.4)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	CR-300-amont-113	-0,122	0,021	-0,163	-0,082
2	CR-300-amont-112	-0,119	0,027	-0,172	-0,066
3	CR-20-amont-231	0,098	0,020	0,058	0,138
4	CR-100-1000-112	-0,084	0,033	-0,149	-0,020
5	CR-200-3000-231	0,084	0,033	0,020	0,148
6	CR-150-amont-152	-0,083	0,022	-0,125	-0,040
7	CR-100-3000-113	-0,067	0,026	-0,117	-0,016
8	CR-250-400-153	-0,065	0,011	-0,087	-0,043
9	CR-250-amont-321	0,055	0,020	0,016	0,094
10	CR-200-3000-121	-0,053	0,009	-0,071	-0,035
11	CR-150-300-154	-0,042	0,009	-0,058	-0,025
12	CR-300-200-412	0,036	0,010	0,015	0,056
13	CR-20-200-311	0,035	0,017	0,002	0,068
14	CR-100-2500-211	0,026	0,004	0,017	0,034
15	CR-200-2500-214	0,022	0,009	0,003	0,040
16	CR-5-amont-252	0,020	0,009	0,003	0,037
17	CR-25-200-251	0,014	0,006	0,003	0,026

CemOA : archive ouverte d'Irstea / Cemagref

Illustration de la contribution des sources de pressions significatives dans ce modèle



CemOA : archive ouverte d'Irstea / Cemagref

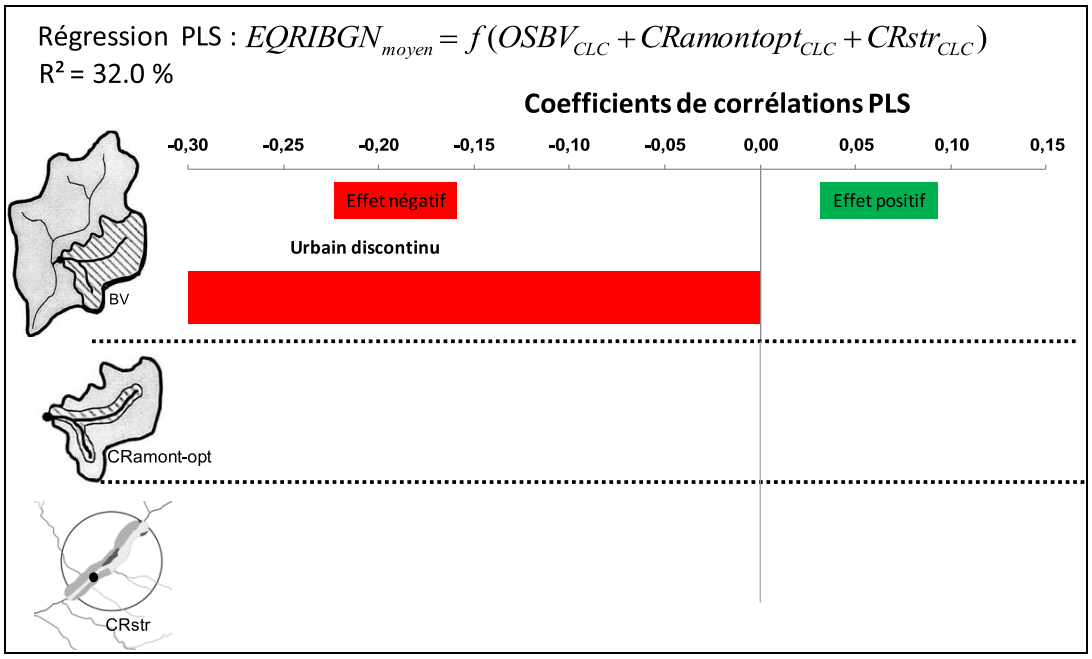
ANNEXE XII.3 *Modèles tri-échelles*

ANNEXE XII.3.1 BV/CRamont /CRlocal-fixe-CLC (Mod-3.1.1)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,566	0,040	-0,644	-0,487

Illustration de la contribution des sources de pressions significatives dans ce modèle



CemOA : archive ouverte d'Irstea / Cemagref

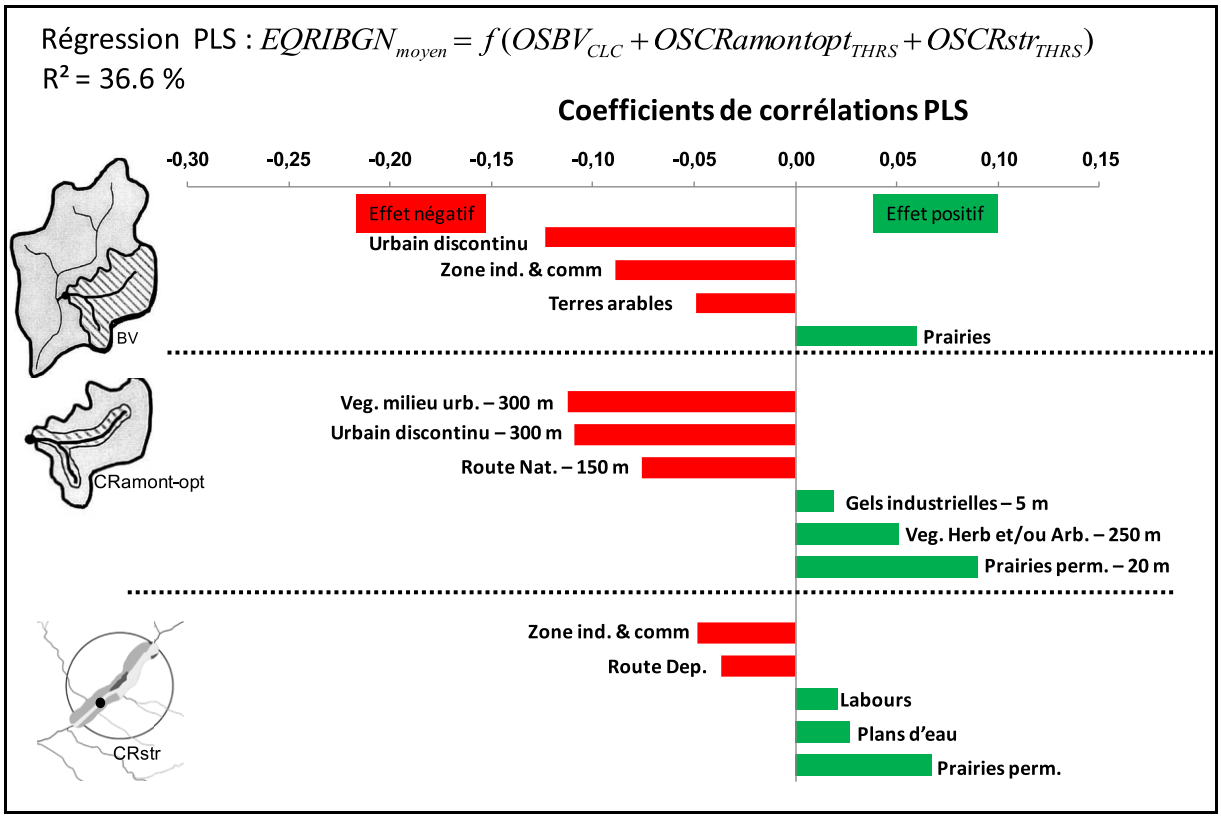
ANNEXE XII.3.2 BV/CRamont /CRlocal-fixe-THRS (Mod-3.1.2)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,124	0,013	-0,149	-0,098
2	CR-300-amont-113	-0,112	0,030	-0,171	-0,054
3	CR-300-amont-112	-0,109	0,032	-0,172	-0,047
4	CR-20-amont-231	0,090	0,018	0,054	0,126
5	BV-121	-0,089	0,025	-0,137	-0,041
6	CR-150-amont-152	-0,076	0,027	-0,128	-0,024
7	CRstr-231	0,067	0,025	0,017	0,117
8	BV-231	0,060	0,026	0,009	0,112
9	CR-250-amont-321	0,051	0,016	0,019	0,083
10	BV-211	-0,049	0,022	-0,093	-0,006
11	CRstr-121	-0,048	0,006	-0,059	-0,038
12	CRstr-153	-0,037	0,017	-0,070	-0,003
13	CRstr-412	0,026	0,013	0,000	0,052
14	CRstr-211	0,021	0,009	0,003	0,039
15	CR-5-amont-252	0,019	0,009	0,001	0,036

Illustration de la contribution des sources de pressions significatives dans ce modèle

CemOA : archive ouverte d'Irstea / Cemagref



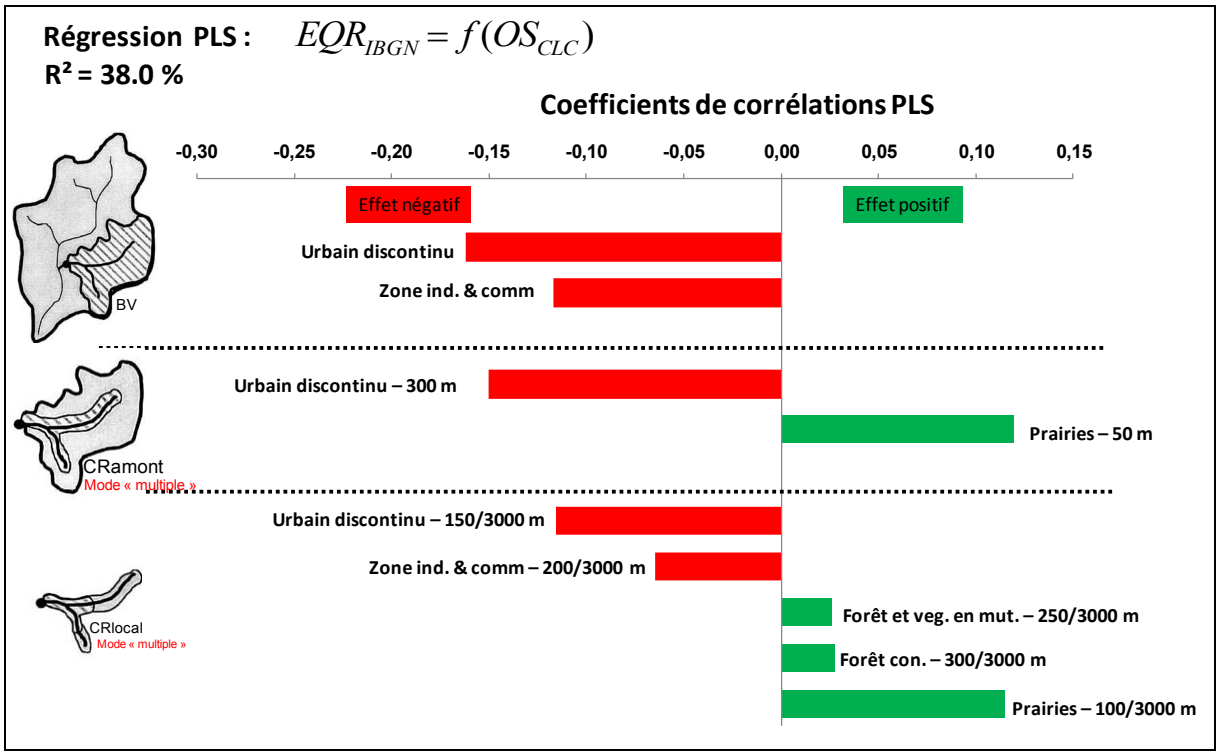
ANNEXE XII.3.3 BV/CRamont /CRlocal-mult-CLC (Mod-3.1.3)

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,162	0,009	-0,181	-0,144
2	CR-300-amont-112	-0,150	0,044	-0,236	-0,065
3	CR-50-amont-231	0,120	0,031	0,060	0,180
4	BV-121	-0,117	0,034	-0,183	-0,051
5	CR-150-3000-112	-0,116	0,040	-0,193	-0,038
6	CR-100-3000-231	0,115	0,027	0,063	0,167
7	CR-200-3000-121	-0,065	0,015	-0,095	-0,035
8	CR-300-1000-312	0,027	0,009	0,010	0,045
9	CR-250-3000-324	0,026	0,007	0,012	0,039

Illustration de la contribution des sources de pressions significatives dans ce modèle

CemOA : archive ouverte d'Irstea / Cemagref



ANNEXE XII.3.4 BV/CRamont /CRlocal-mult-THRS (Mod- 3.1.4)

Résultats de la première itération (modèle avec l'ensemble des variables – IC à 50 %)

Caractéristiques du modèle :

Observations	155 stations
Echelles d'analyse	BV/CRamont/CRlocal
Source d'OS	BV : CLC CRamont: THRS CRlocal-mult: THRS
Variables à expliquer	$EQR\text{-}IBGN_{moyen}$
Variables explicatives	107
Intervalle de confiance	50 %

Résultats de l'itération :

R^2	0.419
Variables significatives (50 %)	82

CemOA : archive ouverte d'Irstea / Cemagref

Coefficients normalisés et IC à 50 % (0 : variables non significatives ; 1 : variables significatives) :

Variable	Coefficient	Ecart-type	Borne inférieure (50%)	Borne supérieure (50%)	Variables significatives
BV-111	-0,009	0,013	-0,017	0,000	1
BV-112	-0,075	0,019	-0,088	-0,062	1
BV-121	-0,054	0,008	-0,060	-0,048	1
BV-123	0,008	0,008	0,002	0,014	1
BV-124	0,013	0,006	0,009	0,017	1
BV-132	-0,016	0,017	-0,028	-0,005	1
BV-133	-0,023	0,024	-0,039	-0,007	1
BV-141	-0,029	0,040	-0,056	-0,002	1
BV-142	-0,030	0,036	-0,054	-0,005	1
BV-211	-0,030	0,017	-0,041	-0,018	1
BV-231	0,037	0,016	0,026	0,047	1
BV-242	0,009	0,012	0,001	0,017	1
BV-243	0,030	0,020	0,016	0,043	1
BV-312	0,013	0,012	0,006	0,021	1
BV-313	0,018	0,018	0,006	0,030	1
BV-322	-0,029	0,040	-0,056	-0,002	1
BV-333	-0,016	0,016	-0,027	-0,005	1
BV-412	-0,008	0,008	-0,014	-0,002	1
BV-512	0,014	0,020	0,001	0,027	1
CR-50-amont-111	-0,012	0,007	-0,017	-0,008	1
CR-300-amont-112	-0,066	0,026	-0,084	-0,049	1
CR-300-amont-113	-0,068	0,025	-0,085	-0,051	1
CR-300-amont-121	-0,033	0,014	-0,043	-0,023	1
CR-300-amont-131	-0,016	0,017	-0,028	-0,005	1
CR-5-amont-133	0,015	0,008	0,010	0,021	1
CR-250-amont-141	-0,031	0,034	-0,054	-0,008	1
CR-300-amont-151	0,010	0,007	0,005	0,015	1
CR-150-amont-152	-0,046	0,016	-0,057	-0,035	1
CR-150-amont-153	-0,032	0,021	-0,047	-0,018	1
CR-5-amont-154	-0,019	0,017	-0,030	-0,008	1
CR-250-amont-211	0,014	0,008	0,008	0,019	1
CR-300-amont-212	0,010	0,013	0,001	0,018	1
CR-250-amont-214	0,012	0,012	0,004	0,020	1
CR-300-amont-222	0,006	0,006	0,002	0,010	1
CR-150-amont-223	0,007	0,007	0,002	0,012	1
CR-20-amont-231	0,055	0,013	0,046	0,063	1
CR-300-amont-232	0,010	0,015	0,000	0,020	1
CR-300-amont-233	0,009	0,011	0,002	0,017	1
CR-300-amont-234	-0,013	0,016	-0,024	-0,002	1
CR-50-amont-241	0,007	0,010	0,001	0,014	1
CR-10-amont-251	-0,016	0,018	-0,028	-0,004	1

CemOA : archive ouverte d'Irstea / Cemagref

CR-5-amont-252	0,011	0,006	0,007	0,015	1
CR-250-amont-311	-0,028	0,023	-0,043	-0,012	1
CR-250-amont-321	0,031	0,011	0,024	0,038	1
CR-300-amont-331	-0,022	0,016	-0,033	-0,011	1
CR-5-amont-412	-0,018	0,015	-0,028	-0,008	1
CR-50-amont-414	-0,007	0,008	-0,012	-0,002	1
CR-300-amont-415	-0,022	0,011	-0,030	-0,015	1
CR-250-amont-417	0,016	0,016	0,005	0,027	1
CR-300-amont-418	0,015	0,017	0,004	0,027	1
CR-300-3000-111	-0,014	0,011	-0,021	-0,006	1
CR-100-1000-112	-0,047	0,024	-0,063	-0,031	1
CR-100-3000-113	-0,037	0,021	-0,051	-0,023	1
CR-200-3000-121	-0,030	0,011	-0,037	-0,022	1
CR-5-400-122	-0,007	0,007	-0,012	-0,002	1
CR-150-200-133	-0,017	0,018	-0,029	-0,005	1
CR-5-1500-141	-0,016	0,016	-0,026	-0,005	1
CR-300-100-142	-0,016	0,017	-0,028	-0,005	1
CR-300-400-151	-0,014	0,004	-0,016	-0,011	1
CR-25-100-152	0,016	0,017	0,004	0,028	1
CR-250-400-153	-0,036	0,005	-0,040	-0,033	1
CR-150-300-154	-0,023	0,006	-0,027	-0,019	1
CR-5-2500-155	-0,018	0,015	-0,028	-0,008	1
CR-100-2500-211	0,014	0,006	0,011	0,018	1
CR-150-1500-212	0,014	0,011	0,007	0,022	1
CR-50-1000-213	-0,023	0,014	-0,033	-0,014	1
CR-200-2500-214	0,012	0,005	0,009	0,015	1
CR-50-1500-223	0,017	0,017	0,005	0,028	1
CR-200-3000-231	0,047	0,019	0,034	0,059	1
CR-300-3000-232	0,010	0,008	0,005	0,015	1
CR-50-1500-233	-0,007	0,008	-0,012	-0,002	1
CR-300-1000-234	-0,011	0,009	-0,017	-0,005	1
CR-25-100-242	-0,023	0,013	-0,032	-0,015	1
CR-25-200-251	0,008	0,004	0,006	0,011	1
CR-20-200-311	0,020	0,012	0,012	0,027	1
CR-300-3000-321	0,025	0,016	0,014	0,035	1
CR-5-2500-331	-0,015	0,021	-0,029	0,000	1
CR-300-200-412	0,020	0,008	0,015	0,025	1
CR-100-200-414	-0,007	0,008	-0,012	-0,002	1
CR-200-1500-415	-0,016	0,016	-0,027	-0,005	1
CR-300-3000-417	0,011	0,013	0,002	0,020	1
CR-100-2000-418	0,007	0,008	0,002	0,013	1
BV-122	-0,001	0,016	-0,012	0,010	0
BV-131	-0,004	0,013	-0,013	0,004	0
BV-222	0,007	0,022	-0,008	0,022	0
BV-311	-0,006	0,018	-0,018	0,007	0
BV-324	0,009	0,014	-0,001	0,019	0

CemOA : archive ouverte d'Irstea / Cemagref

BV-411	-0,002	0,012	-0,011	0,006	0
BV-423	0,000	0,002	-0,001	0,002	0
BV-511	-0,002	0,005	-0,006	0,001	0
BV-523	0,000	0,002	-0,001	0,002	0
CR-5-amont-122	-0,011	0,026	-0,029	0,006	0
CR-10-amont-124	-0,001	0,002	-0,002	0,001	0
CR-15-amont-142	-0,007	0,016	-0,017	0,004	0
CR-300-amont-155	-0,008	0,022	-0,022	0,007	0
CR-300-amont-213	-0,006	0,020	-0,019	0,008	0
CR-300-amont-242	0,018	0,029	-0,002	0,037	0
CR-200-amont-416	-0,006	0,020	-0,020	0,007	0
CR-30-amont-421	-0,001	0,002	-0,002	0,001	0
CR-100-amont-427	0,000	0,002	-0,001	0,002	0
CR-10-amont-428	0,000	0,002	-0,001	0,002	0
CR-5-1000-124	-0,001	0,002	-0,002	0,001	0
CR-300-2500-241	0,010	0,022	-0,005	0,025	0
CR-300-3000-252	-0,016	0,027	-0,034	0,002	0
CR-35-1500-421	-0,001	0,002	-0,002	0,001	0
CR-200-500-427	0,000	0,002	-0,001	0,002	0
CR-10-1000-428	0,000	0,002	-0,001	0,002	0

Résultats de la deuxième itération (modèle avec l'ensemble des variables – IC à 66 %)

Caractéristiques du modèle :

Observations	155 stations
Echelles d'analyse	BV/CRamont/CRlocal
Source d'OS	BV : CLC CRamont: THRS CRlocal-mult: THRS
Variables à expliquer	EQR-IBGN$_{moyen}$
Variables explicatives	82
Intervalle de confiance	66 %

Résultats de l'itération :

R^2	0.418
Variables significatives (66 %)	53

Coefficients normalisés et IC à 66 % (0 : variables non significatives ; 1 : variables significatives) :

CemOA : archive ouverte d'Irstea / Cemagref

Variable	Coefficient	Ecart-type	Borne inférieure (66%)	Borne supérieure (66%)	Variables significatives
BV-112	-0,077	0,019	-0,095	-0,058	1
BV-121	-0,055	0,008	-0,063	-0,047	1
BV-124	0,013	0,006	0,008	0,019	1
BV-133	-0,024	0,024	-0,047	-0,001	1
BV-211	-0,030	0,018	-0,048	-0,013	1
BV-231	0,037	0,017	0,021	0,054	1
BV-243	0,030	0,020	0,011	0,049	1
BV-312	0,014	0,012	0,003	0,025	1
BV-313	0,018	0,018	0,001	0,035	1
CR-50-amont-111	-0,013	0,007	-0,019	-0,006	1
CR-300-amont-112	-0,068	0,026	-0,092	-0,043	1
CR-300-amont-113	-0,070	0,025	-0,094	-0,046	1
CR-300-amont-121	-0,034	0,016	-0,049	-0,019	1
CR-5-amont-133	0,015	0,008	0,008	0,023	1
CR-300-amont-151	0,011	0,008	0,003	0,018	1
CR-150-amont-152	-0,047	0,015	-0,062	-0,032	1
CR-150-amont-153	-0,033	0,022	-0,054	-0,012	1
CR-5-amont-154	-0,019	0,018	-0,037	-0,002	1
CR-250-amont-211	0,014	0,008	0,006	0,022	1
CR-250-amont-214	0,012	0,012	0,000	0,024	1
CR-150-amont-223	0,007	0,007	0,000	0,014	1
CR-20-amont-231	0,056	0,013	0,043	0,068	1
CR-5-amont-252	0,012	0,006	0,006	0,017	1
CR-250-amont-311	-0,028	0,024	-0,051	-0,005	1
CR-250-amont-321	0,031	0,011	0,021	0,042	1
CR-300-amont-331	-0,023	0,017	-0,039	-0,007	1
CR-5-amont-412	-0,018	0,016	-0,034	-0,003	1
CR-300-amont-415	-0,023	0,012	-0,034	-0,011	1
CR-250-amont-417	0,016	0,016	0,001	0,032	1
CR-300-3000-111	-0,014	0,011	-0,025	-0,003	1
CR-100-1000-112	-0,048	0,025	-0,072	-0,024	1
CR-100-3000-113	-0,038	0,021	-0,058	-0,018	1
CR-200-3000-121	-0,030	0,011	-0,041	-0,020	1
CR-5-1500-141	-0,016	0,016	-0,031	-0,001	1
CR-300-100-142	-0,017	0,017	-0,033	0,000	1
CR-300-400-151	-0,014	0,004	-0,018	-0,010	1
CR-250-400-153	-0,037	0,005	-0,042	-0,033	1
CR-150-300-154	-0,024	0,006	-0,029	-0,018	1
CR-5-2500-155	-0,018	0,015	-0,033	-0,003	1
CR-100-2500-211	0,015	0,006	0,009	0,020	1
CR-150-1500-212	0,015	0,011	0,004	0,025	1
CR-50-1000-213	-0,024	0,014	-0,037	-0,011	1

CemOA : archive ouverte d'Irstea / Cemagref

CR-200-2500-214	0,012	0,005	0,008	0,017	1
CR-50-1500-223	0,017	0,018	0,000	0,034	1
CR-200-3000-231	0,048	0,019	0,029	0,066	1
CR-300-3000-232	0,010	0,008	0,002	0,018	1
CR-300-1000-234	-0,011	0,009	-0,020	-0,003	1
CR-25-100-242	-0,024	0,013	-0,037	-0,011	1
CR-25-200-251	0,008	0,004	0,005	0,012	1
CR-20-200-311	0,020	0,012	0,009	0,031	1
CR-300-3000-321	0,025	0,016	0,011	0,040	1
CR-300-200-412	0,020	0,008	0,013	0,028	1
CR-200-1500-415	-0,016	0,017	-0,032	-0,001	1
BV-123	0,008	0,009	0,000	0,016	0
BV-132	-0,017	0,017	-0,033	0,000	0
BV-141	-0,030	0,041	-0,069	0,010	0
BV-142	-0,030	0,037	-0,066	0,005	0
BV-242	0,009	0,012	-0,002	0,021	0
BV-322	-0,030	0,041	-0,069	0,009	0
BV-333	-0,016	0,017	-0,032	0,000	0
BV-412	-0,008	0,009	-0,016	0,000	0
BV-512	0,014	0,020	-0,005	0,034	0
CR-300-amont-131	-0,017	0,017	-0,033	0,000	0
CR-250-amont-141	-0,031	0,035	-0,065	0,002	0
CR-300-amont-212	0,010	0,013	-0,002	0,023	0
CR-300-amont-222	0,006	0,006	0,000	0,012	0
CR-300-amont-232	0,010	0,015	-0,004	0,025	0
CR-300-amont-233	0,010	0,011	-0,001	0,020	0
CR-300-amont-234	-0,013	0,017	-0,029	0,003	0
CR-50-amont-241	0,007	0,010	-0,002	0,017	0
CR-10-amont-251	-0,017	0,019	-0,034	0,001	0
CR-50-amont-414	-0,007	0,008	-0,015	0,000	0
CR-300-amont-418	0,016	0,017	0,000	0,032	0
CR-5-400-122	-0,007	0,008	-0,014	0,000	0
CR-150-200-133	-0,018	0,018	-0,035	0,000	0
CR-25-100-152	0,016	0,018	0,000	0,033	0
CR-50-1500-233	-0,007	0,008	-0,015	0,000	0
CR-5-2500-331	-0,015	0,022	-0,036	0,006	0
CR-100-200-414	-0,007	0,008	-0,015	0,000	0
CR-300-3000-417	0,011	0,014	-0,002	0,025	0
CR-100-2000-418	0,007	0,008	0,000	0,015	0

Résultats de la troisième itération (modèle avec l'ensemble des variables – IC à 90 %)

Caractéristiques du modèle :

Observations	155 stations
Echelles d'analyse	BV/CRamont/CRlocal
Source d'OS	BV : CLC

CemOA : archive ouverte d'Irstea / Cemagref

	CRamont: THRS CRlocal-mult: THRS
Variables à expliquer	$EQR\text{-}IBGN_{moyen}$
Variables explicatives	53
Intervalle de confiance	90 %

Résultats de l'itération :

R^2	0.409
Variables significatives (90 %)	30

Coefficients normalisés et IC à 90 % (0 : variables non significatives ; 1 : variables significatives) :

Variable	Coefficient	Ecart-type	Borne inférieure (90%)	Borne supérieure (90%)	Variables significatives
BV-112	-0,086	0,010	-0,103	-0,069	1
BV-121	-0,062	0,015	-0,086	-0,038	1
BV-124	0,015	0,006	0,005	0,025	1
BV-211	-0,034	0,016	-0,061	-0,008	1
BV-231	0,042	0,018	0,012	0,071	1
BV-243	0,034	0,019	0,003	0,065	1
CR-50-amont-111	-0,014	0,007	-0,025	-0,003	1
CR-300-amont-112	-0,076	0,024	-0,116	-0,036	1
CR-300-amont-113	-0,078	0,022	-0,114	-0,042	1
CR-300-amont-121	-0,038	0,019	-0,069	-0,007	1
CR-5-amont-133	0,017	0,009	0,003	0,032	1
CR-150-amont-152	-0,053	0,017	-0,081	-0,024	1
CR-150-amont-153	-0,037	0,022	-0,073	-0,001	1
CR-250-amont-211	0,016	0,009	0,001	0,030	1
CR-20-amont-231	0,063	0,012	0,043	0,082	1
CR-5-amont-252	0,013	0,006	0,003	0,023	1
CR-250-amont-321	0,035	0,011	0,017	0,053	1
CR-100-1000-112	-0,054	0,025	-0,094	-0,014	1
CR-100-3000-113	-0,043	0,020	-0,076	-0,010	1
CR-200-3000-121	-0,034	0,009	-0,048	-0,019	1
CR-300-400-151	-0,016	0,006	-0,026	-0,005	1
CR-250-400-153	-0,042	0,004	-0,048	-0,036	1
CR-150-300-154	-0,027	0,006	-0,037	-0,016	1
CR-100-2500-211	0,016	0,004	0,010	0,023	1
CR-200-2500-214	0,014	0,005	0,005	0,023	1
CR-200-3000-231	0,054	0,020	0,021	0,086	1
CR-25-200-251	0,009	0,004	0,002	0,016	1
CR-20-200-311	0,022	0,013	0,002	0,043	1
CR-300-3000-321	0,028	0,017	0,001	0,056	1
CR-300-200-412	0,023	0,007	0,011	0,035	1

CemOA : archive ouverte d'Irstea / Cemagref

BV-133	-0,027	0,027	-0,072	0,018	0
BV-312	0,015	0,013	-0,006	0,036	0
BV-313	0,021	0,020	-0,012	0,053	0
CR-300-amont-151	0,012	0,008	-0,001	0,025	0
CR-5-amont-154	-0,022	0,023	-0,059	0,016	0
CR-250-amont-214	0,013	0,014	-0,009	0,036	0
CR-150-amont-223	0,008	0,008	-0,005	0,021	0
CR-250-amont-311	-0,032	0,025	-0,073	0,010	0
CR-300-amont-331	-0,026	0,017	-0,053	0,002	0
CR-5-amont-412	-0,021	0,020	-0,053	0,012	0
CR-300-amont-415	-0,025	0,016	-0,052	0,001	0
CR-250-amont-417	0,018	0,017	-0,010	0,047	0
CR-300-3000-111	-0,016	0,012	-0,036	0,005	0
CR-5-1500-141	-0,018	0,018	-0,048	0,012	0
CR-300-100-142	-0,019	0,019	-0,050	0,013	0
CR-5-2500-155	-0,020	0,015	-0,045	0,004	0
CR-150-1500-212	0,017	0,011	-0,002	0,035	0
CR-50-1000-213	-0,027	0,017	-0,055	0,001	0
CR-50-1500-223	0,019	0,020	-0,013	0,051	0
CR-300-3000-232	0,011	0,010	-0,005	0,027	0
CR-300-1000-234	-0,013	0,011	-0,030	0,005	0
CR-25-100-242	-0,027	0,017	-0,054	0,000	0
CR-200-1500-415	-0,018	0,019	-0,049	0,013	0

Résultats de la quatrième itération (modèle avec l'ensemble des variables – IC à 95 %)

Caractéristiques du modèle :

Observations	155 stations
Echelles d'analyse	BV/CRamont/CRlocal
Source d'OS	BV : CLC CRamont: THRS CRlocal-mult: THRS
Variables à expliquer	EQR-IBGN$_{moyen}$
Variables explicatives	30
Intervalle de confiance	95 %

Résultats de l'itération :

R^2	0.394
Variables significatives (95 %)	24

CemOA : archive ouverte d'Irstea / Cemagref

Coefficients normalisés et IC à 95 % (0 : variables non significatives ; 1 : variables significatives) :

Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)	Variables significatives
BV-112	-0,097	0,009	-0,114	-0,080	1
BV-121	-0,070	0,022	-0,113	-0,027	1
BV-124	0,017	0,007	0,003	0,030	1
BV-211	-0,039	0,017	-0,072	-0,005	1
BV-231	0,048	0,021	0,007	0,088	1
BV-243	0,038	0,019	0,001	0,076	1
CR-100-1000-112	-0,061	0,025	-0,110	-0,012	1
CR-100-2500-211	0,019	0,004	0,012	0,026	1
CR-100-3000-113	-0,048	0,020	-0,088	-0,009	1
CR-150-300-154	-0,030	0,007	-0,043	-0,017	1
CR-150-amont-152	-0,060	0,019	-0,096	-0,023	1
CR-200-2500-214	0,016	0,007	0,002	0,029	1
CR-200-3000-121	-0,038	0,008	-0,054	-0,023	1
CR-200-3000-231	0,061	0,022	0,018	0,103	1
CR-20-200-311	0,025	0,013	0,002	0,051	1
CR-20-amont-231	0,071	0,013	0,045	0,097	1
CR-250-400-153	-0,047	0,006	-0,059	-0,036	1
CR-250-amont-321	0,040	0,013	0,015	0,065	1
CR-25-200-251	0,010	0,005	0,001	0,020	1
CR-300-200-412	0,026	0,008	0,010	0,042	1
CR-300-400-151	-0,018	0,008	-0,034	-0,002	1
CR-300-amont-112	-0,086	0,023	-0,130	-0,042	1
CR-300-amont-113	-0,088	0,020	-0,128	-0,049	1
CR-5-amont-252	0,015	0,007	0,001	0,028	1
CR-150-amont-153	-0,042	0,023	-0,086	0,002	0
CR-250-amont-211	0,018	0,011	-0,003	0,038	0
CR-300-3000-321	0,032	0,018	-0,003	0,067	0
CR-300-amont-121	-0,043	0,026	-0,094	0,008	0
CR-50-amont-111	-0,016	0,009	-0,033	0,001	0
CR-5-amont-133	0,020	0,010	-0,001	0,040	0

Résultat final

Présentation par ordre d'influence des coefficients de la régression obtenus dans ce modèle

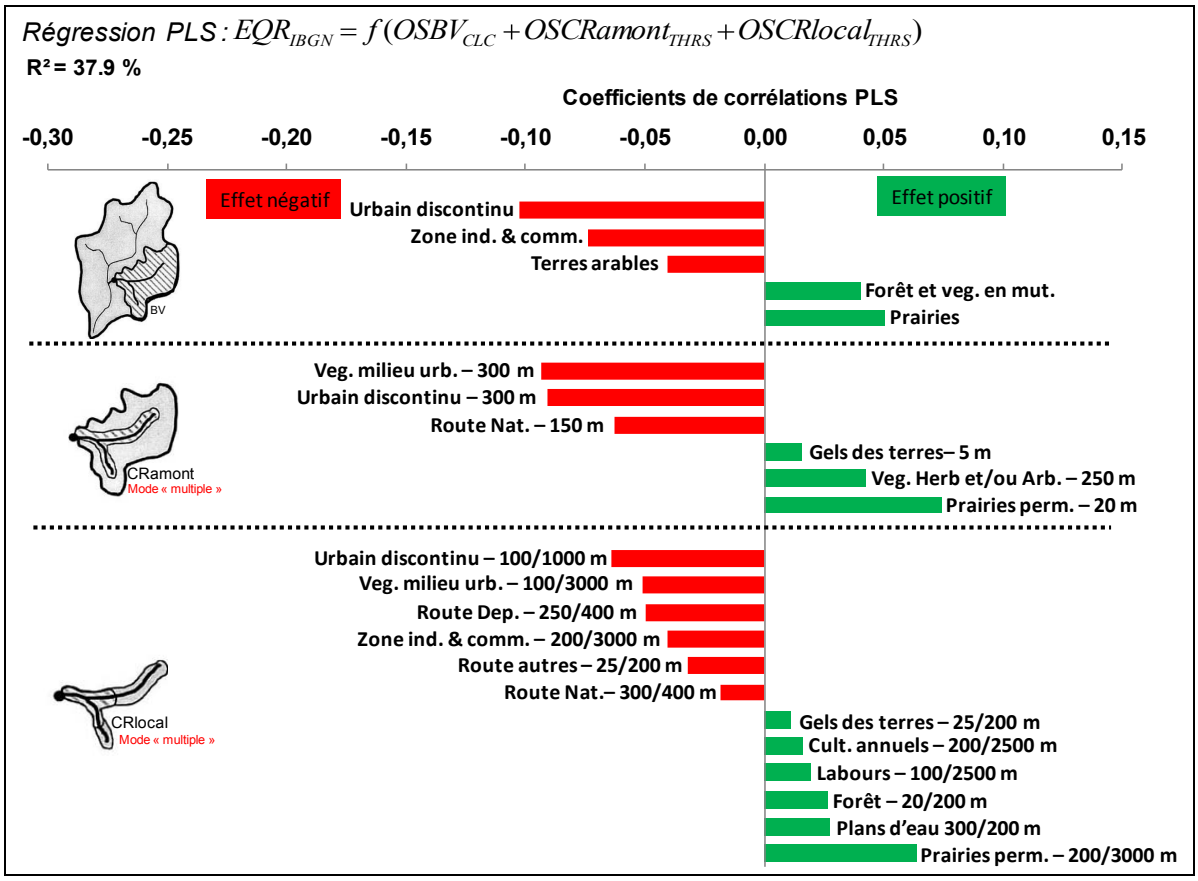
ordre d'influence	Variable	Coefficient	Ecart-type	Borne inférieure (95%)	Borne supérieure (95%)
1	BV-112	-0,103	0,010	-0,122	-0,083
2	CR-300-amont-113	-0,093	0,022	-0,136	-0,051
3	CR-300-amont-112	-0,091	0,024	-0,137	-0,044
4	CR-20-amont-231	0,075	0,014	0,047	0,102

CemOA : archive ouverte d'Irstea / Cemagref

5	BV-121	-0,074	0,022	-0,118	-0,030
6	CR-100-1000-112	-0,064	0,026	-0,116	-0,013
7	CR-200-3000-231	0,064	0,023	0,019	0,109
8	CR-150-amont-152	-0,063	0,020	-0,101	-0,025
9	CR-100-3000-113	-0,051	0,021	-0,093	-0,009
10	BV-231	0,050	0,022	0,007	0,093
11	CR-250-400-153	-0,050	0,006	-0,061	-0,038
12	CR-250-amont-321	0,042	0,013	0,016	0,068
13	BV-211	-0,041	0,018	-0,077	-0,005
14	BV-243	0,041	0,020	0,001	0,080
15	CR-200-3000-121	-0,040	0,009	-0,057	-0,024
16	CR-150-300-154	-0,032	0,007	-0,045	-0,018
17	CR-300-200-412	0,027	0,009	0,010	0,045
18	CR-20-200-311	0,027	0,014	0,002	0,054
19	CR-100-2500-211	0,020	0,004	0,012	0,028
20	CR-300-400-151	-0,019	0,008	-0,035	-0,002
21	CR-200-2500-214	0,017	0,007	0,002	0,031
22	CR-5-amont-252	0,015	0,007	0,001	0,030
23	CR-25-200-251	0,011	0,005	0,001	0,021

Illustration de la contribution des sources de pressions significatives dans ce modèle

CemOA : archive ouverte d'Irstea / Cemagref



CemOA : archive ouverte d'Irstea / Cemagref