Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Discipline ou spécialité :
Écologie

Présentée et soutenue par :
Clément Tisseuil

le : 4 Décembre 2009

Titre :

MODÉLISER L'IMPACT DU CHANGEMENT CLIMATIQUE SUR LES ÉCOSYSTÈMES AQUATIQUES PAR APPROCHE DE DOWNCALING

JURY

M. Martin Daufresne : Chercheur au Cemagref d'Aix-en-Provence
M. Yves Souchon : Directeur de recherche au Cemagref de Lyon
M. Philippe Naveau : Chercheur CNRS au LSCE à Paris
M. Sébastien Brosse : Professeur à l'Université Paul Sabatier à Toulouse

---

Ecole doctorale :
Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

Unité de recherche :
Laboratoire Evolution et Diversité Biologique (UMR 5174)

Directeur(s) de Thèse :
M. Sovan Lek : Professeur à Université de Toulouse (France)
M. Andrew Wade : Professeur à l'Université de Reading (UK)
Dr. Mathieu Vrac : Chercheur CNRS au LSCE à Paris (France)

Rapporteurs :
M. Taha Ouarda : Professeur à l'INRS au Québec (Canada)
M. Emili García-Berthou : Professeur à l'université de Girone (Espagne)

THESE

Pour obtenir le grade de
DOCTEUR DE L'UNIVERSITE DE TOULOUSE
Délivré par l'université Toulouse III – Paul Sabatier
Spécialité : Ecologie
Présentée par :

Clément TISSEUIL

MODELISER L'IMPACT DU CHANGEMENT CLIMATIQUE SUR LES ECOSYSTEMES AQUATIQUES PAR APPROCHE DE DOWNSCALING

Co-dirigée par :
M. Sovan Lek (Professeur à Université de Toulouse, France)
M. Andrew Wade (Professeur à l'Université de Reading, UK)
Dr. Mathieu Vrac (Chercheur CNRS au LSCE à Paris, France)

Soutenue publiquement le 4 décembre 2009, devant le jury composé de :

M. Taha Ouarda : Professeur à l'INRS au Québec (Canada)	Rapporteur
M. Emili García-Berthou : Professeur à l'université de Girone (Espagne)	Rapporteur
M. Martin Daufresne : Chercheur au Cemagref d'Aix-en-Provence	Examinateur
M. Yves Souchon : Directeur de recherche au Cemagref de Lyon	Examinateur
M. Philippe Naveau : Chercheur CNRS au LSCE à Paris	Examinateur
M. Sébastien Brosse : Professeur à l'université Paul Sabatier à Toulouse	Examinateur

« Les statistiques, c'est comme le bikini.

Ce qu'elles révèlent est suggestif.

Ce qu'elles dissimulent est essentiel »

(Aaron Levenstein)

AUTEUR : Clément Tisseuil

TITRE : Modéliser l'impact du changement climatique sur les écosystèmes aquatiques par approche de downscaling.

DIRECTEURS DE THÈSE : Sovan Lek, Andrew J. Wade, Mathieu Vrac

LIEU ET DATE DE SOUTENANCE : Toulouse, vendredi 4 Décembre 2009

RÉSUMÉ :

L'objectif de ma thèse était d'évaluer l'impact du changement global sur les écosystèmes aquatiques au cours du 21ème siècle, dans le bassin Adour Garonne (S-O France). Une approche de « downscaling » a été développée à l'interface entre les sciences du climat, de l'hydro-chimie et de l'écologie. Les résultats suggèrent une augmentation globale des débits hivernaux et une diminution des débits d'étiage. Les concentrations en nitrate ainsi que la distribution des espèces de poisson thermophiles pourraient également augmenter. Toutefois, des scénarios de diminution des gaz à effet de serre ainsi qu'une modification des pratiques agricoles (ex. diminution des fertilisants azotés) pourraient limiter l'intensité des perturbations écologiques. Cette thèse offre une contribution originale, notamment pour la gestion future des ressources hydriques et écologiques.

TITRE et résumé en anglais au recto de la dernière page

MOTS CLÉS : assemblages d'espèces, changement climatique, distribution d'espèces, gradients environnementaux, incertitudes, modélisation statistique, modélisation mécanistique, niche écologique, poissons d'eau douce, projections futures, nitrates, régime hydrologique, régionalisation, downscaling, variabilité spatio-temporelle, scénarios climatiques.

DISCIPLINE : Ecologie

ADRESSE DU LABORATOIRE DE RATTACHEMENT : Laboratoire Evolution & Diversité Biologique Bâtiment 4R3 Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 4, France

REMERCIEMENTS

J’aurais aimé écrire une petite chansonnette pour remercier à la volée tous ces amis, collègues et Amour, qui ont largement contribué à mon équilibre affectif et professionnel durant ces années de thèse. Mais étant donné que je me sens un peu sec après les dernières semaines de labeur à rédiger ce mémoire, je me contenterai d’un remerciement plus traditionnel…

Je tiens à remercier tout particulièrement Sovan Lek pour la confiance et la patience qu’il m’a accordées, en toutes circonstances. L’affaire n’était pas gagnée d’avance mais je dois dire qu’on ne s’en est pas trop mal tiré ! Merci à mon deuxième directeur de thèse, Andrew Wade, pour m’avoir ouvert au monde « INCA », version modèle *hydro-chimique* (tout de suite, ça perd de son charme)… Un grand merci à Mathieu Vrac, mon troisième directeur de thèse adopté en cours de route, pour m’avoir ouvert aux joies transcendantales du downscaling… Merci au projet EUROLIMPACS pour m’avoir permis de manger à ma faim.

Quant aux collègues de travail, tous sont aujourd’hui de véritables amis pour qui je garde une affection profonde et des petits souvenirs avec lesquels on pourrait écrire au moins deux ou trois articles dans Nature : « Muriel et ses friandises » ; « boucles d’or de Madame Rosy » ; « Géraldine et son monstrueux chien (Simon ?) » ; « Petit guide des piqures de raies par Gaël, aux éditions Papataki (j’ai pas trouvé mieux) » ; « Laetitia et Chabot, l’histoire de toute une vie » ; « La pêche au thon en dix leçons avec Simon » ; « Le cri perçant du Bobby reptilien » ; « Petites discussions de comptoir avec Guillaume » ; « Conseils de Séb pour bien réussir son brushing ». Sans oublier, « Maman Christine et ses gâteaux » ; « Sithan et ses bambous » ; « Les fous rires de Dominique », et pour finir, « Les aventures de Candida à suivre sur www.aquarium-passion@candi.fr »... Je pense aussi bien fort à tous les EDBiens qui peuplent le laboratoire et aux fabuleux membres et organisateurs de la Beer Party du jeudi soir ! Merci à Bertrand qui a été le meilleur stagiaire que j’ai eu (même si il a été le seul ;).

Merci tout particulièrement aux collègues qui ont participé à la relecture de ce mémoire, notamment à Laetitia pour sa patience et ses corrections.

Le travail, c’est la santé… C’est bien mignon mais il n’y a pas de santé sans amour. Merci à Rhéa, ma douce compagne pour m’avoir rafistolé après mes galipettes en montagne, pour son amour fidèle et confiant en toute circonstance. Merci enfin à la famille qui, je l’espère, sera bien fière de sa progéniture ;)

SOMMAIRE

INTRODUCTION : 9

1 Impact du changement global sur les écosystèmes d'eau douce 9

2 Enjeux et défis de la modélisation en hydro-écologie: 10

3 Développement d'un modèle hydro-écologique conceptuel 11

4 Objectifs généraux de la thèse 13

1ière PARTIE : CONCEPTS ET MÉTHODOLOGIE 15

1 Introduction 15

2 Description des données 17

2.1 Données régionales et locales: hydrologie, climat, biologie, physico-chimie, géomorphologie 17

2.2 Processus atmosphériques, modèles climatiques et scénarios futurs 17

3 Modélisation statistique versus mécanistique, statique versus dynamique 19

4 Downscaling des conditions hydro-climatiques locales: 20

4.1 Principes du downscaling 20

4.2 Développement d'un modèle de downscaling statistique 21

5 Modèle de downscaling hydro-biologique 25

5.1 Downscaling saisonniers des débits et des températures 25

5.2 Modèle statistique et statique de distribution d'espèce (niche-based models) 27

5.3 Validation des projections hydro-biologiques sur la période contrôle 29

6 Modèle de downscaling hydro-chimique 31

6.1 Downscaling des précipitations et températures journalières 31

6.2 Modèle hydro-chimique HBV/INCA-N 31

6.3 Validation des projections hydro-chimiques 35

2ième PARTIE : PROJECTIONS FUTURES ET INCERTITUDES 36

1 Méthode 36

1.1 Indicateurs de biodiversité et de changements hydro-chimiques 36

1.2 Partitionnement de la variabilité dans les projections 37

1.3 Patrons de variation spatio-temporelle dans les projections 39

2 Changements dans la biodiversité des peuplements de poissons 41

3 Modification de la dynamique hydro-chimique sur la Garonne 45

3ième PARTIE : DISCUSSION 49

1 Considérations méthodologiques 49

1.1 Crédibilité des projections futures, variabilité et incertitudes 49

1.2 Downscaling hydro-climatique 50

2 Considérations écologiques 52

2.1 Perturbations inévitables des écosystèmes ? 52

2.2 Atténuations possibles des impacts du changement climatique ? 54

CONCLUSIONS ET PERSPECTIVES.....55

1.1 Synthèse des résultats.....55

1.2 Vers une modélisation statistico-dynamique plus réaliste.....56

RÉFÉRENCES.....59

LISTE DES ARTICLES

1. **Article n°1** : **Tisseuil, C**., Wade, A.J., Tudesque, L. and Lek, S., 2008. Modeling the Stream Water Nitrate Dynamics in a 60,000-km^2 European Catchment, the Garonne, Southwest France. J Environ Qual, 37: 2155-2169.
2. **Article n°2 : Tisseuil C**., Vrac M., Wade AJ., Lek S. Statistical downscaling of river flow. *Moderate revisions in Journal of Hydrology.*
3. **Article n°3 : Tisseuil C.**, Vrac M, Wade AJ, Grenouillet G, Gevrey M, Lek S. Validating a hydro-ecological model to project fish community structure from general circulation models using downscaling techniques *(in preparation).*
4. **Article n°4 : Tisseuil C.**, Vrac M, Wade AJ, Grenouillet G, Gevrey M, Lek S. Spatio-temporal impacts of climate change on biodiversity: strengthen the link between downscaling and bioclimatic models *(in preparation).*

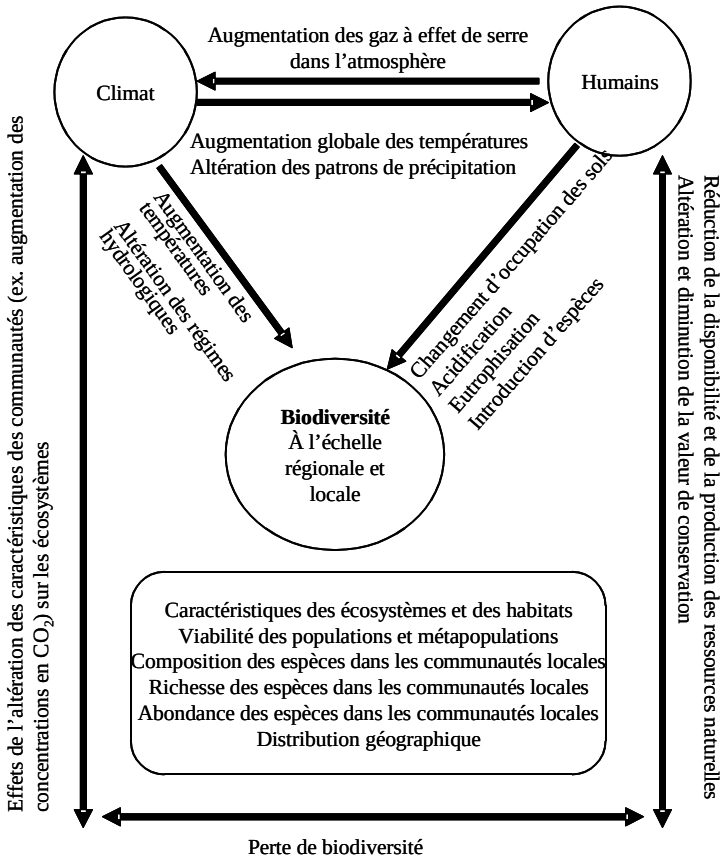


Figure 1. Relations entre le changement climatique et certaines perturbations anthropiques et leurs effets sur la biodiversité. Les deux facteurs principaux résultant directement du changement climatique et des facteurs anthropiques majeurs ont des effets à la fois individuels et interactifs sur la biodiversité des écosystèmes d'eau douces. Adaptée de Heino *et al.* (2009)

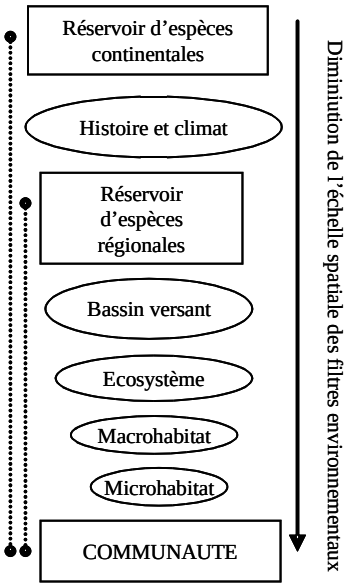


Figure 2. Modèle schématique des différents filtres environnementaux affectant les communautés régionales et locales. Le pool continental des espèces est déterminé par les processus d'extinction et de spéciation à très larges échelles spatiales et temporelles. Le filtre supérieur caractérisé par l'histoire (ex. spéciation, extinction, dispersion) et le climat (ex. température, précipitation, énergie) détermine le pool d'espèces régional. Au sein du pool d'espèces régional, quatre niveaux de filtres environnementaux déterminent les communautés locales: (i) le bassin versant (ex. occupation du sol, régimes hydrologiques), l'écosystème (ex. température, chimie de l'eau), le macrohabitat (ex. % occupation de macrophytes) et le microhabitat (ex. substrat et granulométrie). Ces filtres déterminent la diversité et la composition des communautés au travers des traits biologiques des espèces. Adaptée de Poff (1997)

INTRODUCTION :

1 IMPACT DU CHANGEMENT GLOBAL SUR LES ÉCOSYSTÈMES D'EAU DOUCE

Bien qu'il soit souvent restreint à des considérations d'ordre climatique, le terme changement global se réfère à une série de changements naturels ou d'origine anthropique de la structure biologique et physique de la Terre, qui dans leur ensemble ont des effets significatifs à échelle globale (Pachauri & Reisinger 2007). Les modèles numériques de circulation générale (GCM) calibrés sur les 100 dernières années, projettent que le réchauffement climatique devrait s'accentuer dans les années à venir et que les températures pourraient augmenter de 1.4°C à 5.8°C d'ici à la fin du 21^{ième} siècle, selon que l'atmosphère sera plus ou moins chargée en gaz à effet de serre. Quant aux précipitations, les GCM sont assez discordants dans leurs projections selon les régions. Certains GCM suggèrent que les précipitations pourraient augmenter de façon très variable, particulièrement au niveau des tropiques avec une intensification à la fois des extrêmes pluviométriques et des sécheresses. Si le changement climatique est global, ses impacts sont en revanche perceptibles à l'échelle locale car c'est en réalité la modification des combinaisons entre les conditions climatiques, hydrologiques et géomorphologiques locales qui est susceptible d'altérer le fonctionnement des écosystèmes.

Au cours des dix dernières années, le nombre de projets internationaux visant à mieux comprendre l'impact du changement global sur les écosystèmes terrestres (ex. ALARM ; GOCE-CT-2003-506675) ou aquatiques (ex. EUROLIMPACS ; GOCE-CT-2003-505540) s'est multiplié, à différents niveaux d'organisation biologique (gènes, populations, espèces, communautés et écosystèmes) et à différentes échelles spatiales (habitat, locale, régionale et continentale) (Heino *et al.* 2009 ; Figure 1). Concernant les écosystèmes aquatiques, les changements climatiques pourraient avoir un effet de cascade à différents niveaux de l'écosystème, depuis le cycle hydrologique jusqu'à l'occupation des territoires, en influençant la mobilité des éléments physiques (sédiments, nutriments), la structuration de l'habitat des rivières et, *in fine*, l'organisation des communautés biologiques (Wrona *et al.* 2006 ; Ormerod 2009 ; Palmer 2009). En Europe, l'augmentation présumée des températures, en conjugaison avec une diminution des débits, pourraient intensifier les processus d'acidification et d'eutrophisation des cours d'eau (Heino *et al.* 2009). La modification des débits pourrait inexorablement influencer la morphologie des cours d'eau, les flux de matière et, en conséquence, les conditions d'habitat qui soutiennent la biodiversité actuelle des rivières.

Les effets projetés du changement climatique sur la biodiversité sont sans appel et font état de 15 à 37% d'extinction possible chez les espèces terrestres au cours des 50 prochaines années (Thomas *et al.* 2004). Plusieurs études ont déjà pu observer/prédire des changements significatifs dans la structure et le fonctionnement des communautés biologiques en réponse au changement global : déplacements des organismes vers de plus hautes latitudes et altitudes en accord avec leurs préférences thermiques (Parmesan & Yohe 2003 ; Root *et al.* 2003), changement dans la phénologie (décalage saisonnier dans le cycle biologique ; Walther *et al.* 2002) ou diminution de la taille des organismes pouvant affecter les paramètres démographiques des populations (fertilité, interactions compétitives ; Daufresne *et al.* 2009). Concernant les organismes aquatiques d'eau douce, les poissons sont probablement les plus étudiés dans les études du changement global. Les projections futures suggèrent l'expansion de la distribution spatiale des espèces d'eau chaude vers l'amont des rivières, au détriment de celle des espèces d'eau froide en réponse à l'augmentation globale des températures (Matulla *et al.* 2007; Buisson *et al.* 2008 ; Heino *et al.* 2009). Le réchauffement global pourrait également favoriser l'introduction d'espèces invasives et exotiques, qui pourraient alors nuire aux espèces natives et modifier profondément le fonctionnement des réseaux trophiques (Rahel & Olden 2008 ; Leprieur *et al.* 2009).

2 Enjeux et défis de la modélisation en hydro-écologie:

C'est à l'échelle du bassin versant que se situent les véritables enjeux de la modélisation en hydro-écologie (Palmer *et al.* 2009). Le bassin versant est en effet l'unité géographique sur laquelle se base la réalisation des flux de matière et l'expression du vivant (biodiversité) (Statzner *et al.* 1988 ; Noss 1990), dont la structure complexe met en perspective quatre dimensions interconnectées : les dimensions longitudinale (continuum rivière ou gradient amont-aval), verticale (zones hyporhéiques de transition entre les eaux de surface et souterraines), latérale (connectivité entre le cours principal et les annexes hydrauliques) et temporelle (variations saisonnières dans les régimes d'écoulement). Face aux changements globaux, les enjeux majeurs de la modélisation en hydro-écologie sont à la fois d'ordre écologique, économique et social afin de contribuer à une gestion durable des ressources en eau et une préservation de la biodiversité. Ces enjeux sont également d'ordre scientifique car ils favorisent l'interdisciplinarité entre les sciences du climat, de l'hydrologie et de l'écologie, dans le but de développer des modèles intégrés qui aident à mieux anticiper (prédire), comprendre et mesurer les conséquences du changement climatique sur les écosystèmes aquatiques et l'incertitude qui leur est associée.

Face à ces enjeux, un des défis majeurs de la modélisation hydro-écologique consiste à intégrer les projections climatiques issues des GCM en entrée de modèles d'impact hydro-écologiques. Cette opération délicate est confrontée à deux difficultés majeures : (i) la qualité des sorties des GCM peut être très variable selon les modèles, et plus particulièrement la difficulté des GCM à modéliser correctement et de manière consensuelle les processus hydrologiques dans l'atmosphère; (ii) d'un point de vue technique, un modèle hydro-écologique implique intuitivement au moins trois niveaux de modélisation (modèle climatique, hydrologique et écologique) qu'il convient d'agencer de manière judicieuse, tout en limitant l'expansion inévitable de l'incertitude au fur et à mesure des niveaux de modélisation. Dans ce contexte, une étape déterminante connue sous le terme de 'downscaling' constitue un élément clé pour favoriser le transfert de l'information climatique vers les niveaux hydrologiques et écologiques inférieurs. En outre, grâce au transfert de l'information climatique contenue à large échelle spatiale dans les GCM (~250 km x 250 km) vers une résolution spatiale plus fine, régionale (~50 km x 50 km) ou locale, le processus de downscaling permet de prendre en compte de manière tangible les variabilités régionales et saisonnières du changement climatique, ce qui s'avère indispensable pour la plupart des modèles d'impacts.

3 Développement d'un modèle hydro-écologique conceptuel

Comme le suggère la structure des bassins versants, les différents processus hydrologiques, chimiques et biologiques sont organisés de manière hiérarchique dans l'espace. Les conditions climatiques constituent un premier filtre à large échelle spatiale, alors que l'occupation des sols et la géomorphologie qui sont enchevêtrées à plus fine résolution, sont susceptibles d'influencer les processus hydrologiques et physico-chimiques à travers le bassin versant. A l'échelle de l'habitat ou du micro-habitat, la structure et le fonctionnement des communautés aquatiques sont la résultante des processus opérant à des échelles supérieures (Poff *et al.* 1997 ; Heino *et al.* 2009 ; Figure 2). Dans un contexte de changement climatique, la conception d'un modèle hydro-écologique peut donc s'envisager comme le couplage de différents modèles prédictifs en chaîne assurant le transfert de la variabilité climatique vers des modèles hydro-écologiques. Dans le cadre des travaux réalisés au cours de ma thèse, deux approches conceptuelles ont été développées et appliquées au bassin Adour Garonne (sud-ouest de la France) : une approche hydro-biologique et une approche hydro-chimique. Ces deux approches sont constituées à la base d'un modèle de downscaling qui projette les conditions hydro-climatiques futures, sur lesquelles se greffe le modèle d'impact hydro-

biologique ou hydro-chimique en vue de projeter les perturbations hydro-écologiques potentielles futures.

Le modèle hydro-biologique se base sur la projection de la distribution des poissons d'eau douce. Les poissons sont des organismes poïkilothermes dont la distribution spatiale est fortement structurée le long du gradient amont-aval (continuum) des rivières (Vannote *et al.* 1980). Les poissons constituent donc des modèles biologiques particulièrement adaptés pour l'étude des impacts du changement global. Bien que plusieurs études aient déjà exploré les impacts potentiels futurs du changement global sur la structure et le fonctionnement des communautés de poisson, que ce soit en Europe (Matulla *et al.* 2007 ; Buisson *et al.* 2008) ou en Amérique du Nord (Jackson & Mandrak, 2002 ; Mohseni *et al.* 2003 ; Chu *et al.* 2005 ; Sharma *et al.* 2007), peu d'entre elles ont explicitement inclus la composante hydrologique dans les projections futures. L'hydrologie est pourtant l'un des moteurs fondamentaux du fonctionnement global des écosystèmes aquatiques et joue un rôle particulièrement important dans le cycle de vie des poissons d'eau douce (Statzner *et al.* 1988 ; Poff *et al.* 1997 ; Cattanéo 2005 ; Lamouroux & Cattanéo 2006). Le modèle hydro-chimique a quant à lui été conçu pour quantifier l'impact du changement global sur la variabilité saisonnière des débits et des concentrations en nitrates le long du gradient amont-aval de la Garonne. Les concentrations en nitrates sont fortement influencées par les activités anthropiques sur les bassins versants. Aussi, le modèle hydro-chimique a été développé en vue de comparer l'intensité des changements hydro-chimiques futurs selon différents scénarios climatiques et de changement d'occupation des sols.

4 Objectifs généraux de la thèse

L'objectif général de ma thèse est à la fois d'ordre méthodologique et écologique :

(i) **L'objectif méthodologique** est de proposer une approche de modélisation intégrée des bassins versants qui favorise le lien entre les projections hydro-climatiques futures issues du downscaling des modèles climatiques et des modèles écologiques. Cette approche permettra d'aider à mieux modéliser l'impact du changement global sur le fonctionnement hydro-écologique des bassins versants. Cette composante s'est donc articulée autour de trois disciplines que sont la climatologie, l'hydrologie et l'écologie pour la construction et la compréhension des modèles.

(ii) **L'objectif écologique** est d'évaluer l'impact potentiel futur du changement global sur la biodiversité des poissons d'eau douce ainsi que sur la dynamique hydrologique et hydro-chimique des nitrates sur les bassins versants. L'analyse des projections hydro-écologiques futures s'intéressera à trois aspects principaux : (a) dans un objectif de conservation des écosystèmes, l'impact du changement global sur les écosystèmes aquatiques pourrait-il être plus marqué dans certaines zones des bassins versants et/ou à des périodes futures particulières ? (b) dans un souci d'aide à la gestion et à la décision, quelle crédibilité peut-on accorder aux projections, compte-tenu de leurs nombreuses sources de variabilité et d'incertitude ? (c) dans un contexte socio-économique, comment l'impact du changement global sur les écosystèmes aquatiques peut-il évoluer, s'intensifier ou s'atténuer en fonction des différentes orientations socio-économiques et politiques (ex. émissions de gaz à effet de serre, modification des pratiques agricoles) ?

Pour répondre à ces deux objectifs principaux, mon manuscrit se divisera en trois parties. Une première partie, essentiellement méthodologique, s'attachera à la description des modèles mis en oeuvre et à leur validation sur les conditions climatiques actuelles. S'appuyant sur cette validation, une deuxième partie analysera les projections hydro-chimiques et hydro-biologiques dans le futur, en se focalisant sur la quantification des patrons de variabilité spatiaux et temporels ainsi que leur incertitude. La dernière partie nous ramènera aux deux objectifs initiaux de ma thèse, en discutant des limites et des forces de la méthodologie mise en œuvre et des impacts potentiels du changement global sur le fonctionnement hydro-écologique des bassins versants. La conclusion fera une synthèse des résultats, afin d'en faire émerger des perspectives de recherche future.

Tableau 1Synthèse des données utilisées pour la modélisation hydro-chimique et hydro-biologique

MODELES	TYPE DE DONNEES	DESCRIPTION DES DONNEES	ORIGINE DES DONNEES
MODELISATION HYDRO-CHIMIQUE	Réseaux de surveillance des concentrations en NO_3 and NH_4 (mg N l^{-1})	Séries mensuelles entre 1992 et 2005 sur 16 stations localisées sur la Garonne, pour la calibration et la validation d'INCA-N.	AEAG
	Concentrations en NO3 and NH4 (mg N l-1) et débits à la sortie des stations d'épuration sur la Garonne entre 1990 et 2000	Moyenne annuelle théorique calculée par sous-bassin et estimation des flux journaliers moyens sur la période d'étude.	AEAG
	Séries de débits journaliers moyens ($m^3 s^{-1}$)	Débits moyens journaliers mesurés en niveau de 7 stations le long de la Garonne entre 1992 et 2005.	MEEDM
	Précipitations et températures (mm)	Interpolation des données journalières à partir d'environ 150 stations climatiques, pour 7 groupes de sous-bassins climatiques définis pour le fonctionnement du modèle hydrologique HBV.	METEOFRANCE
	Précipitations hydrologiques efficaces (HER) et déficit en eau dans le sol (SMD) (mm)	Estimations jounalières par le modèle HBV au niveau des 7 groupes de sous-bassin climatiques, utilisées en entrée du modèle hydro-chimique INCA-N.	HBV (Bergström, 1992)
	Pratiques agricoles et taux d'application des fertilisants azotés (kg N ha^{-1}année $^{-1}$)	Taux estimé en fonction des variétés de culture (céréales, oléagineux) et pratiques régionales.	MAP (statistiques de l'Agreste)
	Dépositions atmosphériques sèches et humides en NH_4 et NO_3 (mgN l^{-1})	Moyenne annuelle des dépositions totales de NH_4 et NO_3 à partir de carte digitalisées, réparties équitablement en dépositions humides et sèches à partir des dépositions totales.	Réseaux RENECOFOR (Croisé et al., 2002)
	Concentration en NO_3 and NH_4 dans les eaux souterraines (mg N l^{-1})	Données purement informatives estimées pour chaque sous-bassin à partir de 21 stations à proximité du linéaire de la Garonne.	BRGM (base de donnée ADES)
	Occupation des sols (km^2)	Recoupement entre les cartes digitalisées de l'occupation des sols de Corine et la répartition des cultures régionales	IFEN (2000, mise à jour en 2005) et MAP
	Pratiques agricoles et périodicité dans l'application des fertilisants	Date approximative de début et de fin d'application des fertilisants azotés en fonction des types de cultures (céréales, oléagineux)	ARVALIS
MODELISATION HYDRO-BIOLOGIQUE	Sorties de modèles climatiques (GCM) et de scénarios climatiques	Sorties mensuelles à l'échelle du monde, et journalières à l'échelle de l'Europe, pour 13 GCMs et approximativement 20 variables atmosphériques en fonction de 4 scénarios climatiques: 20c3m, A2, A1B et B1.	Serveurs du GIEC [https://esg.llnl.gov:8443/index.jsp]
	Réanalyses NCEP/NCAR	Données de réanalyses NCEP/NCAR journalières à l'échelle de l'Europe pour 21 variables atmosphériques, de 1948 à 2005.	NCEP/NCAR [http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html]
	Séries de débits journaliers moyens ($m^3 s^{-1}$)	Débits journaliers moyens mesurés approximativement entre 1950 et 2000, sur une cinquantaine de stations hydrologiques	MEEDM [http://www.hydro.eaufrance.fr/]
	Séries journalières de température (°C)	Températures journalières moyennes interpolées par krigeage au niveau des 50 stations d'étude, à partir d'un réseau d'approximativement 150 stations climatologiques, entre 1970 et 2005.	METEOFRANCE
	Cractéristiques géomorphologiques	Caractéristiques géomorphologiques des 50 stations d'étude: pente moyenne, largeur du lit des cours d'eau, localisation géographique altitudinale, longitudinale, et latitudinale, surface du bassin versant.	ONEMA
	Présence-absence de poissons	Inventaire piscicole annuel estimant l'abondance des espèces de poisson d'eau douce sur la France métropolitaine. Les données issues des 50 stations d'étude ont été extraites et utilisées en terme de présence-absence dans les modèles de distribution d'espèces.	ONEMA

1IÈRE PARTIE : CONCEPTS ET MÉTHODOLOGIE

1 INTRODUCTION

L'objectif de ce chapitre est de synthétiser le principe et les étapes de validation des modèles hydro-chimiques et hydro-biologiques. La conceptualisation d'un modèle hydro-écologique est classiquement élaborée de manière 'ascendante', selon laquelle le modèle d'impact écologique est contraint de s'adapter à la disponibilité et à la nature des données climatiques modélisées dans le futur. Les projections du modèle écologique manquent ainsi parfois de pertinence et de précision car les données climatiques fournies en entrée manquent parfois elles-mêmes de pertinence et de précision. Au cours de cette thèse, une conceptualisation 'descendante', tout à fait complémentaire à la précédente, a été privilégiée. Son principe est de fournir en entrée des modèles d'impact des variables climatiques de qualité optimale et adaptées au besoin du modèle écologique grâce à l'utilisation de techniques de downscaling.

L'ensemble des concepts et modèles développés dans le cadre de cette thèse a été expérimenté sur le bassin Adour Garonne, couvrant la partie sud-ouest de la France sur approximativement 160 000 km², et caractérisé par une large de gamme de conditions environnementales : hydrologiques (du régime nival de montagne au régime pluvial de plaine), climatiques (influence continentale au nord, méditerranéenne au sud-est, océanique à l'ouest) et topographiques (Massif Central au nord-est et Pyrénées au sud). Cette variabilité des conditions environnementales favorisent la diversification et la richesse des écosystèmes aquatiques.

La première partie de ce chapitre fait l'inventaire des données utilisées pour la construction des modèles et obtenues grâce à des collaborations avec de nombreux organismes nationaux et internationaux (Tableau 1). La deuxième partie fait le point sur les différentes approches de modélisation couramment utilisées en climatologie, hydrologie et biogéographie afin de justifier le choix des modèles utilisés au cours de cette thèse. La troisième partie développera les concepts et outils de downscaling qui ont été appliqués en entrée des modèles hydro-biologiques et hydro-chimiques et détaillés dans les deux dernières parties. Au fil de la construction des modèles, mon attention s'est tout particulièrement portée sur la compréhension des processus. Aussi, une démarche rigoureuse a été mise en œuvre pour comprendre et valider les modèles sur le climat présent.

Tableau 2 : Variables atmosphériques à large échelle utilisées en fonction des GCM , selon la modélisation hydro-biologique et hydro-chimique

Variables atmosphériques			Disponibilité des données / GCMs													
Nom complet	Nom court	Unités	*csiro_mk3_0*	*ncar_ccsm3_0*	bccr_cm2_0	*cccma_cgcm3_1*	***cnrm_cm3***	*csiro_mk3_5*	***gfdl_cm2_0***	***gfdl_cm2_1***	*giss_model_e_r*	inmcm3_0	*ipsl_cm4*	***miroc3_2_medres***	*mpi_echam5*	***mri_cgcm2_3_2a***
Température moyenne de l'air à la surface	air.2m	K		×	×	×	×		×	×	×	×		×	×	×
Température moyenne de l'air à 500 hPa	air.500	K		×	×	×	×		×	×	×	×		×		×
Température moyenne de l'air à 850 hPa	air.850	K		×	×	×	×		×	×	×	×		×		×
Précipitation moyenne convective à la surface	cprat	kg m^{-2} s^{-1}	×					×	×	×	×			×	×	×
Radiation de longue longueur d'onde descente à la surface, par ciel dégagé	csdlf	W m^{-2}		×	×	×	×		×	×	×	×		×		×
Radiation solaire de courte longueur d'onde ascendante, par ciel dégagé	csusf	W m^{-2}					×		×	×	×				×	
Radiation de longue longueur d'onde descente à la surface	dlwrf	W m^{-2}	×	×	×	×	×	×	×	×	×	×		×	×	×
Radiation de courte longueur d'onde descente à la surface	dswrf	W m^{-2}					×							×		
Géopotentiel moyen à 500 hPa	hgt.500	m		×	×	×			×	×	×	×		×		
Géopotentiel moyen à 850 hPa	hgt.850	m		×	×	×			×	×	×	×	×	×		
Précipitations moyennes à la surface	prate	kg m^{-2} s^{-1}	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Pression moyenne de surface	pres	Pa	×	×	×	×	×	×	×	×				×	×	×
Humidité relative moyenne à 500 hPa	rhum.500	%		×	×	×			×		×			×		×
Humidité relative moyenne à 850 hPa	rhum.850	%		×	×	×			×	×	×			×		×
Humidité spécifique moyenne à 500 hPa	shum.500	kg kg^{-1}								×		×	×	×		
Humidité spécifique moyenne à 850 hPa	shum.850	kg kg^{-1}		×	×	×				×		×		×		
Température moyenne du sol	skt	K					×		×	×	×	×		×		×
Niveau de pression de la mer	slp	Pa	×					×	×							
Couverture moyenne des nuages	tcdc	%		×	×	×							×	×	×	×
Radiation de longue longueur d'onde ascendante	ulwrf	W m^{-2}					×		×	×		×		×		×
Radiation de courte longueur d'onde ascendante	uswrf	W m^{-2}		×	×	×	×		×	×	×			×	×	×

En italique, données utilisées pour la partie de downscaling du modèle hydro-chimique
En gras, données utilisées pour la partie downscaling du modèle hydro-biologique

2 Description des données

2.1 Données régionales et locales: hydrologie, climat, biologie, physico-chimie, géomorphologie

Les données de débits journaliers ont été utilisées pour une cinquantaine de stations d'étude pour la période 1970-2000. Elles ont été fournies par le Ministère de l'Ecologie, de l'Energie et du Développement durable et de la Mer (MEEDM ; base de données Hydro2). Les données climatologiques journalières pour près de 150 stations réparties sur l'ensemble du bassin Adour Garonne, ont été fournies par Météofrance sur la période 1950-2000. Les concentrations en azote et ammonium, mesurées mensuellement entre 1990 et 2005 dans 16 stations de la Garonne, ainsi que l'estimation des flux de rejets azotés en provenance des stations d'épuration répertoriées sur la Garonne, ont été fournies par l'Agence de l'Eau Adour Garonne (AEAG). L'occupation des sols sur le bassin de la Garonne, en relation avec les pratiques agriculturales (type de cultures, fréquence et quantité de fertilisants) renseignées par le Ministère de l'Agriculture et de la Pêche (MAP), a été extraite de la couche vectorielle Corine (Institut Français de l'Environnement ; 2001, 2005). Les inventaires piscicoles annuels entre 1992 et 2005, fournis par l'Office National de l'Eau et des Milieux Aquatiques (ONEMA), ont été utilisés en terme de présence-absence pour les 13 espèces de poissons les plus fréquentes sur les 50 sites d'études. Les données d'abondance n'ont pas été considérées en raison d'un certain biais relatif aux différents protocoles d'échantillonnage utilisés lors des campagnes de pêche.

2.2 Processus atmosphériques, modèles climatiques et scénarios futurs

La circulation atmosphérique est le mouvement à l'échelle planétaire de la couche d'air entourant la Terre qui redistribue la chaleur provenant du soleil. En conjonction avec la circulation océanique, elle contribue ainsi à la variabilité spatiale et temporelle des climats. La dynamique de la circulation atmosphérique est généralement mesurée ou modélisée dans les trois dimensions spatiales et dans le temps, au travers de différents processus atmosphériques (ex. température, précipitations, pression, ensoleillement, humidité et vitesse du vent). Dans le cadre des travaux de ma thèse, deux types de données ont été utilisées : les réanalyses du National Centre for Environmental Prediction and the National Centre for Atmospheric Research (NCEP/NCAR ; Kalnay *et al.* 1996) et les sorties de plusieurs modèles de circulation générale (GCM). Vingt et une variables atmosphériques ont été utilisées couvrant

géographiquement la zone d'étude et caractérisant les principaux descripteurs atmosphériques (Tableau 2).

Les réanalyses peuvent être considérées comme des pseudo-observations reconstituant l'évolution de la circulation atmosphérique depuis plus d'un demi-siècle. Elles résultent de l'assimilation de différentes sources de mesures pouvant provenir de stations météorologiques locales ou d'observations satellitaires. Les réanalyses NCEP/NCAR sont disponibles à l'échelle du globe, à un pas de temps journalier ou inférieur, et caractérisées par une large résolution spatiale d'approximativement 2.5° x 2.5°. Au cours de ma thèse, les réanalyses NCEP/NCAR journalières ont été utilisées à l'échelle de la France, pour la période 1970-2000. Elles ont été utilisées pour comprendre les relations entre les processus atmosphériques à large échelle spatiale avec la variabilité locale et saisonnière du climat et de l'hydrologie. Cette étape était par conséquent indispensable en vue de calibrer et de valider les modèles hydro-climatiques de downscaling (Figure 3).

Les GCM fournissent globalement le même type de variables atmosphériques et à la même résolution spatiale et temporelle que les réanalyses. Toutefois, les GCM sont des modèles numériques complexes résolvant explicitement les équations primitives de la mécanique des fluides géophysiques et de la thermodynamique. Environ 25 GCM existent à travers le monde, dont le principal désaccord porte sur le bilan hydrique et radiatif de la planète. Les GCM génèrent des simulations de climats transitoires pour projeter le climat futur selon différents scénarios développés dans les travaux de groupe d'experts intergouvernemental sur l'évolution du climat (GIEC). Dans ma thèse les données journalières et mensuelles de 13 GCM ont été utilisées respectivement à l'échelle du globe et de l'Europe. Néanmoins, pour les besoins de l'étude, ces données n'ont été concrètement exploitées que pour une zone géographique réduite à la moitié Sud de la France.

Quatre scénarios climatiques modélisés par les différents GCM ont également été utilisés selon les besoin de l'étude : (i) le scénario 20c3m, dit de contrôle pour chacun des GCM, est une reconstitution numérique du climat présent selon l'évolution observée des forçages naturels et anthropiques depuis le siècle dernier ; (ii) les scénarios futurs sont basés sur le Rapport Spécial des Scénarios d'Emission (SRES) publié par le GIEC (Pachauri & Reisinger 2007), caractérisant l'évolution potentielle future du climat en fonction des orientations sociales, politiques et économiques qui pourraient être prises au cours du 21^{ième} siècle et qui détermineraient les émissions de gaz à effet de serre (GES). Le scénario A2 suppose une augmentation globale de la population ainsi qu'une croissance économique régionale importante et plus fragmentée que dans les autres scénarios. Le scénario A1B suppose une

croissance économique et démographique très rapide jusqu'à un pic au milieu du 21ième siècle, suivie d'une décroissance relative conjuguée avec l'introduction rapide de nouvelles technologies énergétiques plus efficaces et moins polluantes. Le scénario B1 est le plus optimiste et se base sur une transformation rapide et globale des fonctionnements économiques avec l'introduction généralisée de nouvelles technologies propres et efficaces.

3 Modélisation statistique versus mécanistique, statique versus dynamique

Sans rentrer dans un débat qui dépasse largement le cadre de cette thèse - et qui plus est dont les terminologies sont parfois différentes en climatologie, biogéographie et hydrologie - une classification des différents types de modèles pourrait se faire selon les deux critères suivants : statistique *versus* mécanistique, statique *versus* dynamique.

L'approche mécanistique se base sur des considérations physiques en climatologie (ex. bilan radiatif), démographiques en écologie (ex. taux de fertilité) ou encore biochimiques en hydrologie (ex. dénitrification) qui régulent les processus ('*process-based models*'). Au contraire, l'approche statistique établit une relation empirique entre le (ou les) processus à modéliser et un (ou plusieurs) prédicteur(s) supposé(s). Rien ne permet d'affirmer la supériorité d'une approche par rapport à l'autre et les deux approches présentent parfois des avantages très complémentaires. L'approche statistique explore probablement de manière plus intuitive et simplifiée les relations entre un processus et son ensemble de prédicteurs. En outre, s'ils sont paramétriques, les modèles statistiques permettent de tester de manière robuste un certain nombre d'hypothèses sur l'effet ou non d'une variable prédictrice et la nature de sa relation avec le processus. Quant à l'approche mécanistique, elle peut se révéler plus réaliste en intégrant explicitement des équations et paramètres de la physique, de l'écologie ou de l'hydrologie. En revanche, il est fréquent que le paramétrage des modèles mécanistiques requière de grosses quantités de données, ce qui les rend souvent plus coûteux que des modèles statistiques, en termes de temps de calcul, et moins facilement applicables sur de grandes échelles spatiales et temporelles.

La différence entre modèles statiques et dynamiques réside principalement dans leur façon d'intégrer l'information. En biogéographie, les modèles dynamiques tentent généralement de donner une représentation de la niche fondamentale de l'espèce (Hutchinson, 1957), selon un état de non-équilibre entre l'espèce et son milieu en décrivant explicitement dans l'espace et/ou dans le temps des processus démographiques et écologiques de l'espèce (ex. compétition, capacité de dispersion, taux de croissances). En climatologie, les modèles dynamiques prennent généralement en compte les processus rétroactifs du climat qui peuvent

avoir des conséquences en différé dans l'espace et/ou dans le temps. A l'opposé, les modèles statiques supposent une relation directe entre un ensemble de prédicteurs et le processus à modéliser. En biogéographie, les modèle statiques basés sur les niches réalisées des espèces ('niche-based models') reposent sur le postulat que les espèces sont à l'équilibre (ou quasi-équilibre) avec leur milieu, ce qui constitue une hypothèse souvent nécessaire pour la prédiction de la distribution des espèces à grande échelle spatiale (Guisan & Zimmermann 2000; Pearson & Dawson 2003; Guisan & Thuiller 2005). De la même façon en climatologie, les modèles statiques de downscaling reposent sur le postulat que les processus climatiques locaux résultent directement de la variabilité des processus atmosphériques à large échelle, et/ou des contraintes géographiques régionales, sans par exemple prendre en compte les évènements climatiques des jours précédents.

Globalement au cours de ma thèse, une approche statistique a été privilégiée pour construire les différents modèles en raison d'une plus grande flexibilité et rapidité de calcul par rapport à l'approche mécanistique. Par exemple, la reproductibilité des projections selon différents GCM et scénarios climatiques est plus facile, compte tenu de la quantité de données et de l'échelle spatiale étudiée. Pour plus de détails sur la nature et la spécificité des différents modèles utilisées, une synthèse est fournie dans les sections suivantes : modèle statistique et statique de downscaling (Section 4), modèle statistique et statique de distribution d'espèces (Section 5.2), modèle hydro-chimique mécanistique et dynamique (Section 6.2).

4 DOWNSCALING DES CONDITIONS HYDRO-CLIMATIQUES LOCALES:

4.1 PRINCIPES DU DOWNSCALING

Le principe du downscaling consiste à augmenter la résolution spatiale des sorties des GCM afin de prendre en compte la variabilité régionale ou locale liée par exemple à la topographie, ou l'occupation des sols (Wilby *et al.* 2002 ; Fowler *et al.* 2007). Dans le downscaling mécanistique, les modèles de climats régionaux (RCM) sont nichés à plus forte résolution spatiale (approximativement 50 km x 50 km), à l'intérieur des mailles de faible résolution des GCM (approximativement 250 km x 250 km). Les RCMs sont ainsi à l'échelle régionale ce que les GCM sont à l'échelle globale : une représentation mécanistique faisant interagir les processus atmosphériques modélisés par les GCM et les variabilités géomorphologiques et physiques de la région.

Le downscaling statistique établit une relation statistique entre une (ou plusieurs) variable(s) atmosphérique(s) des GCM modélisée(s) à large échelle spatiale, et une variable hydro-climatique locale, en se basant sur trois hypothèses fondamentales : (i) les variables

GCM sont des variables appropriées pour le problème étudié (climat régional/local), leur lien avec le climat régional est fort et la zone sur laquelle on les considère est pertinente ; (ii) les variables climatiques sont simulées de façon réaliste par les GCM à l'échelle où on les considère et doivent représenter correctement le signal du changement climatique ; (iii) l'hypothèse de stationnarité suppose que la relation établie entre les variables des GCM et la variable locale à prédire a été validée pour le climat présent et reste valable pour le climat futur perturbé par les forçages anthropiques et naturels.

4.2 Développement d'un modèle de downscaling statistique

Le modèle de downscaling statistique, composé essentiellement de deux parties, a été compilé en langage R (R Development Core team 2009) dans la librairie DWS (disponible sur demande) dont les différents étapes sont résumées ci-après. Une étape dite de « régionalisation » établit la relation statistique entre la circulation atmosphérique à large échelle spatiale et la variable hydro-climatique locale ou régionale. Une deuxième étape s'appuie sur la méthode de transformation de la fonction de distribution cumulée (CDFt, Michelangeli *et al.* 2009). Cette dernière a été utilisée pour répondre à deux objectifs dans cette thèse : (i) la correction du biais statistique dans les projections régionales par un ajustement saisonnier des projections, spécifiquement à chaque station (voir Section 5.1 ; Article n°3 ; Figure 3c); (ii) en tant que méthode de downscaling à part entière (voir Section 6.1 ; Figure 4b), en faisant directement le lien entre la probabilité de distribution d'une variable climatique à large échelle et celle d'une variable locale (voir Michelangeli *et al.* 2009).

4.2.1 Processus atmosphériques à large échelle

Dix des 21 variables atmosphériques ont été présélectionnées afin de synthétiser les principaux processus atmosphériques supposés influencer la variabilité hydro-climatique locale (Table 2; Figure 3b en gras). La méthode de présélection des variables est détaillée dans l'Article n°3 sur la validation des projections hydro-biologiques sur le climat présent. La synthèse des principaux processus atmosphériques se fait en deux temps : (i) les variables atmosphériques les plus proches, en terme de similarité dans leurs patrons de variabilité journalière, sont regroupées à l'aide d'une méthode de classification hiérarchique en quatre groupes de processus atmosphériques : précipitations, température (incluant les radiations de grande longueur d'onde émise dans l'infrarouge), radiations de courte longueur d'onde (émissions directes du soleil) et pression (Figure 3b) ; (ii) pour chaque groupe, le premier axe d'une analyse en composantes principales (ACP) est ensuite extrait, synthétisant plus de 90%

de l'information dans chaque ACP, afin de caractériser de manière synthétique le processus atmosphérique. Cette représentation de l'information présente le double avantage de réduire le nombre de prédicteurs en entrée du modèle de downscaling, tout en identifiant spécifiquement leur nature et en limitant la corrélation entre eux (colinéarité).

4.2.2 Régionalisation

Cinq méthodes statistiques provenant de différentes librairies R on été regroupées dans la librairie DWS pour créer le lien statistique entre les processus atmosphériques à large échelle et la variabilité hydro-climatique locale (Figure 3c). Ces méthodes incluent les modèles linéaires généralisés (librairie *stat*), les modèles additifs généralisés (librairie *mgcv*), les réseaux neuronaux (librairie *amore*), les forêts d'arbres aléatoires (random forest, librairie *randomForest*) et les forêts adaptatives (boosted tree, librairie *gbm*). Ces différentes méthodes reposent sur des principes algorithmiques spécifiques qui sous-tendent des relations plus ou moins complexes entre les prédicteurs et la réponse.

Dans les modèles linéaires généralisés (GLM ; McCullagh 1984) et les modèles additifs généralisés (GAM ; Hastie & Tibshirani 1990; Wood 2008), la variable réponse qui suit une loi de distribution statistique connue ou hypothétique (ex : loi normale, binomiale, poisson) est reliée au prédicteurs par une fonction de lien de type paramétrique dans le cas des GLM (identité, logit, log-vraisemblance) ou non paramétrique de lissage dans les cas des GAM (« smooth spline »). Les réseaux neuronaux apprennent à prédire la variable réponse de manière itérative en pondérant les prédicteurs jusqu'à parfaire la prédiction de la variable réponse en utilisant un algorithme, le plus communément utilisé étant le 'back-propagation network' (Rumelhart *et al.* 1986 ; Reed & Marks 1998; Lek & Guégan 1999). Les forêts adaptatives (boosted tree) et aléatoires (random forest) sont deux méthodes dérivées des arbres de classification dont le principe de base est d'expliquer la variation d'une variable continue (régression) ou qualitative (classification) en différenciant successivement les données en groupes homogènes (De'ath & Fabricius 2000). Les forêts adaptatives génèrent une succession d'arbres où chaque nouvel arbre diminue l'erreur du précédent (De'ath 2007 ; Elith *et al.* 2008). Les forêts aléatoires génèrent également une série d'arbre, chaque arbre résultant de l'échantillonnage aléatoire des observations et des prédicteurs, pour finalement moyenner le résultat de tous ces arbres (Breiman 2001).

Dans l'Article n°2 sur le downscaling des débits, la capacité des modèles à prédire la variabilité hydrologique régionale a été comparée entre modèles linéaires, modèles additifs généralisés, réseaux de neurones et forêts adaptatives. Les forêts adaptatives ont montré une

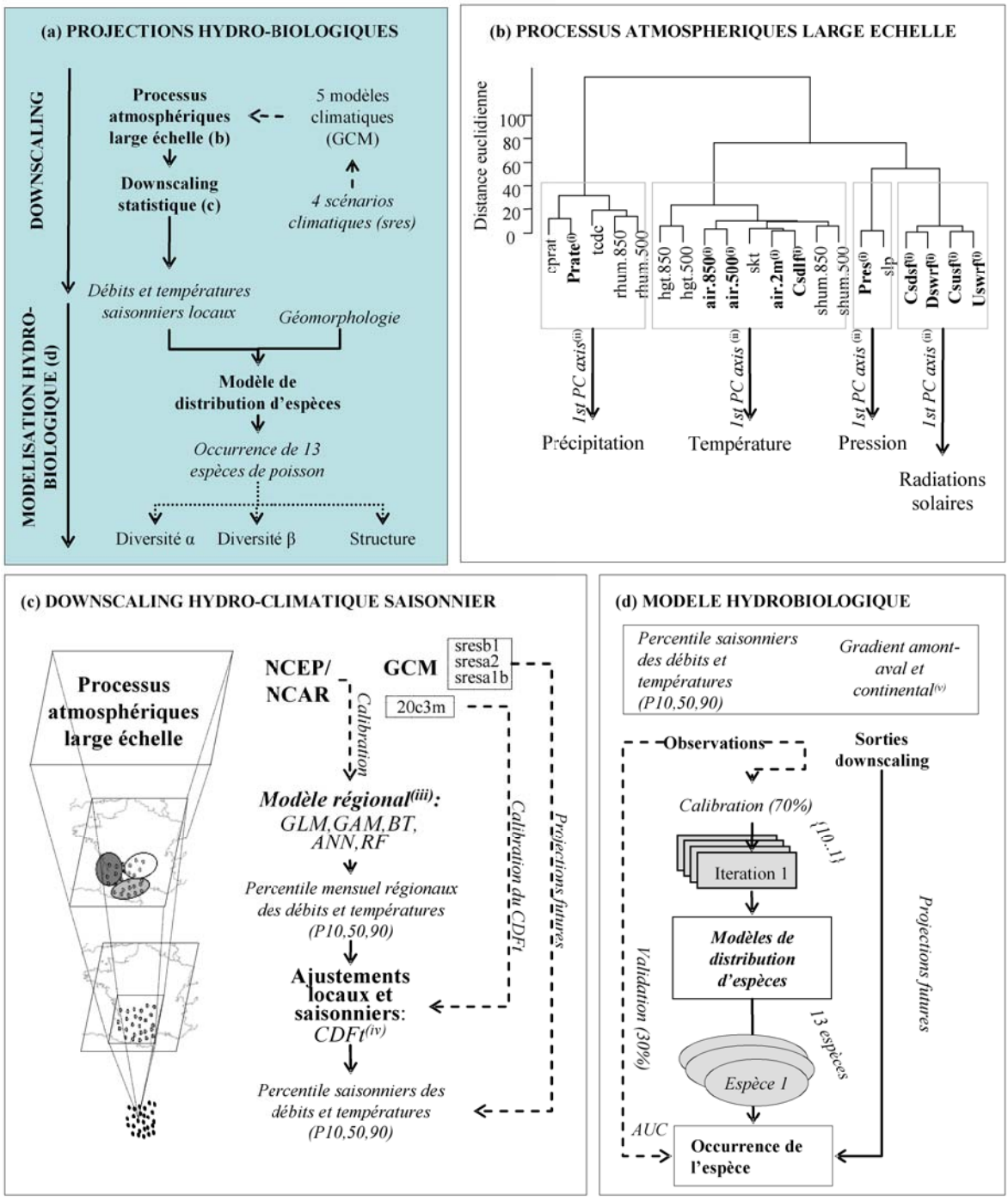
meilleure capacité à projeter la variabilité hydrologique régionale à partir des processus atmosphériques à large échelle. Cette meilleure performance peut être éventuellement du à trois raisons : (i) la prise en compte de la non-linéarité entre les processus atmosphériques et la variabilité hydrologique ; (ii) la structure hiérarchique héritée des arbres de régression intègre implicitement des interactions possibles entre processus atmosphériques ; (iii) leur principe qui est de classifier les données dans l'intervalle de valeur des observations utilisées pour leur calibration, les expose moins au risque de projeter des valeurs extrêmes de manière erratique. Par la suite, les forêts adaptatives ont été utilisées comme seule méthode statistique de régionalisation pour la projection des conditions hydro-climatiques futures sur la région d'étude.

4.2.3 Ajustement des projections hydro-climatiques

La fonction de transformation de la distribution cumulée (CDFt) est une méthode proche de la méthode quantile-quantile (Deque 2007) dont le principe est de corriger un certain biais statistique dans les projections par rapport à des données observées ou théoriques. CDFt a pour objectif de transformer la distribution de probabilité des projections de manière à l'ajuster à celle de la variable réponse observée. La particularité de CDFt est donc de pouvoir prendre en compte l'évolution de la probabilité de distribution d'une variable. En se basant sur la transformation établie entre les données à grande et petite échelle sur le climat présent (typiquement le scénario 20c3m), CDFt permet de transposer dans le futur les projections à petite échelle à partir des projections futures à large échelle.

4.2.4 Stationnarité des projections hydro-climatiques

L'hypothèse de stationnarité des débits est globalement transgressée sur les 30 années approximatives d'étude. Afin d'y remédier pour la calibration des modèles de downscaling, une procédure de validation croisée a été utilisée dont le principe est de : (i) découper la série de données en trois séries temporellement distinctes (a, b, c) ; (ii) chaque série est alors utilisée tour à tour pour la calibration du modèle régional (ex. sur a), la calibration des paramètres du CDFt (ex. le projections sur b du modèle régional issu de a) et la projection ajustée sur la période de validation (ex. sur c). Les projections réalisées sur chaque période de validation sont ensuite moyennées, permettant ainsi de reconstruire une série temporelle dont une part de variabilité liée à la non-stationnarité des données observées est atténuée.



(i) : Présélection des variables atmosphériques les mieux modélisées par les GCMs (Voir Annexe 3)

(ii) : 1° axe d'Analyse en Composante Principale (PCA) appliquée à chaque groupe de variables atmosphériques.

(iii) : Choix de régionalisation statistique basée sur les modèles linéaires (GLM) ou additifs (GAM) généralisés, les forêts adaptatives (BT) ou aléatoires (RF) et les réseaux de neurones (ANN).

(iv) : Fonction de Transformation de la Distribution Cumulée.

(v) : Extraction des 2 premiers axes d'une ACP appliquée au caractéristiques géomorphologiques du bassin versant: altitude, pente, distance à la source, surface du bassin, localisation géographique.

Figure 3. (a) Schéma conceptuel du modèle hydro-biologique pour la projection future des conditions hydro-climatiques locales et la distribution des 13 espèces les plus abondantes sur les 50 sites d'étude, décomposé en trois parties principales: (b) synthèse des processus atmosphériques agissant à large échelle caractérisant le climat présent et futur; (c) downscaling de la variabilité hydro-climatique locale et saisonnière à partir des processus atmosphériques large échelle, selon les étapes de régionalisation et d'ajustements locaux; (d) modèle statique de distribution d'espèces, appliqué individuellement pour chaque espèce, projetée à partir des conditions hydro-climatiques downscalées et des caractéristiques géomorphologiques des sites d'étude.

5 Modèle de downscaling hydro-biologique

Le downscaling hydro-biologique (Figure 3a) fait référence au modèle hydro-climato-écologique (HCE) présenté dans l'Article n°3, qui vise à coupler les projections hydro-climatiques issues de modèles de downscaling (Figure 3a, b, c) avec des modèles statiques de distribution pour 13 espèces de poisson sur le bassin Adour Garonne (Figure 3d).

5.1 Downscaling saisonniers des débits et des températures

Le downscaling hydro-climatique s'est focalisé sur l'optimisation des projections saisonnières des débits et des températures servant de prédicteurs en entrée des modèles de distribution d'espèces. Les quatre variables synthétisant les processus atmosphériques (précipitation, température, pression et radiation solaires de courte longueur d'onde) ont été utilisées comme prédicteurs de la variabilité hydro-climatique saisonnière. Bien que le downscaling de l'hydrologie ait été réalisé indépendamment de celui des températures, le principe méthodologique reste le même. L'étape de régionalisation s'est faite en deux temps :

(i) cinq régions hydrologiques et quatre régions thermiques ont été identifiées séparément à l'aide de méthodes de classification hiérarchique afin de regrouper les stations ayant une dynamique hydrologique (Figure 4a) ou de température saisonnière (Figure 4b) similaire ;

(ii) pour chacune des régions, les forêts adaptatives on été calibrées afin d'assurer la connexion entre les prédicteurs atmosphériques et chacun des trois percentiles mensuels 10, 50 et 90% des débits et des températures (P10, P50 et P90), qui caractérisent le profil mensuel minimum, moyen et maximum des débits et températures.

Au total, 27 modèles de régionalisation ont donc été construits incluant 15 modèles hydrologiques (3×5) et 12 modèles de température (3×4). Une description des connexions reliant la variabilité hydrologique régionale avec les descripteurs atmosphérique est discutée dans l'Article n°2 sur le downscaling des débits. Dans cette étude, l'influence probable des radiations solaires sur le déclenchement de la fonte des neiges printanières est mise évidence dans les régimes nivaux, alors que la température atmosphérique apparaît comme un prédicteur majeur de variabilité hydrologique dans les régimes pluviaux, très certainement au travers du processus d'évaporation (Figure 5a). L'étape d'ajustement des projections hydro-climatiques issues des 27 modèles de régionalisation a été appliquée pour chacune des 50 stations et chacune des trois saisons biologiques définies pour les modèles statiques de distribution des espèces (Figure 3c).

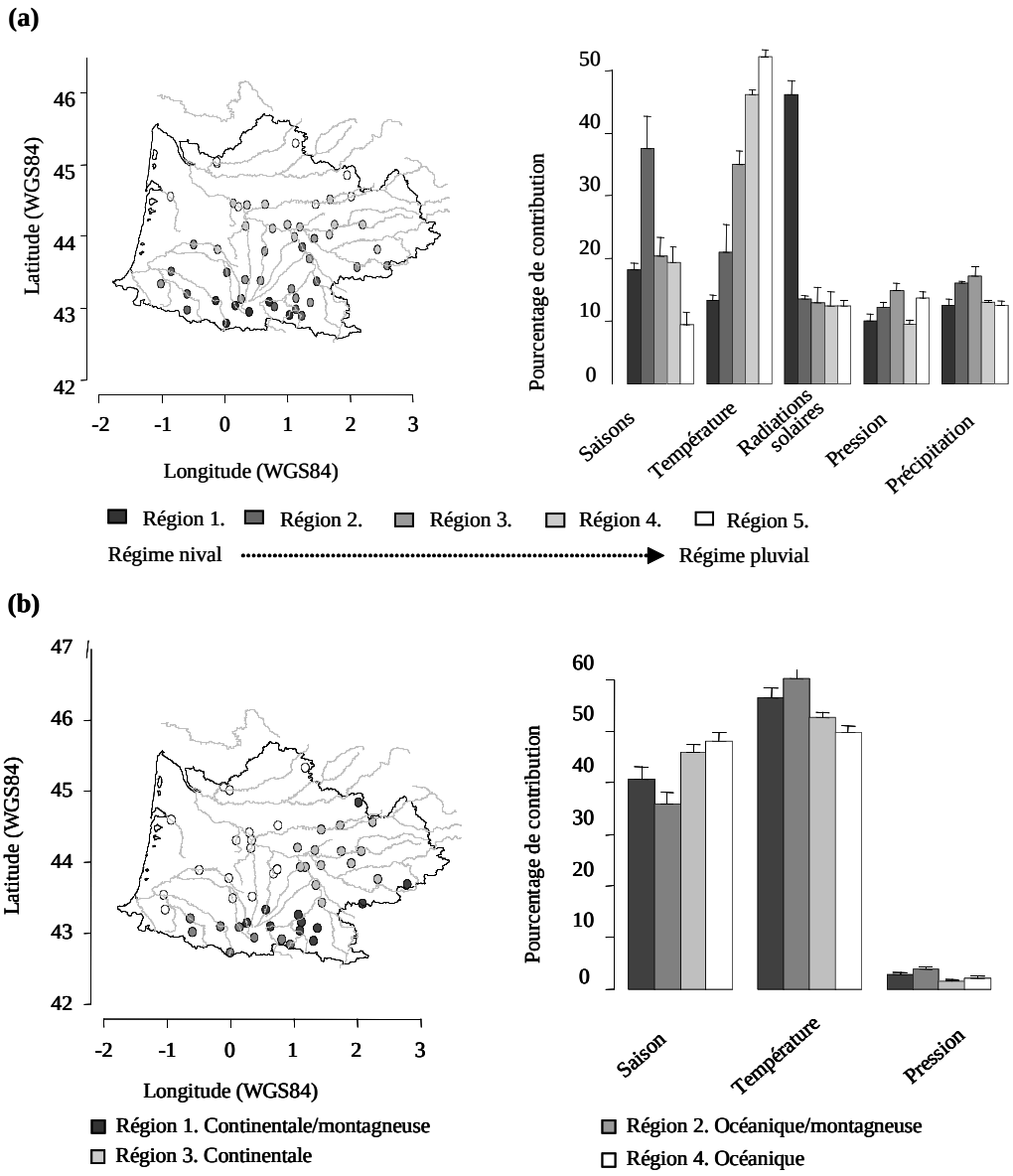


Figure 4. Description des régions hydrologiques (a) et climatiques (b) pour l'étape de downscaling régional, identifiées par classification hiérarchique. Pour chaque région hydro-climatique, la contribution à la variabilité hydro-climatique régionale expliquée par chaque processus atmosphérique synthétique (température, radiations solaires de courte longueur d'onde, pression et précipitations) ainsi que par le cycle mensuel a été calculée à l'aide de l'indice de Gini au travers de la méthode statistique des forêts adaptatives (boosted tree).

5.2 Modèle statistique et statique de distribution d'espèce (*niche-based models*)

5.2.1 Choix des modèles

La niche réalisée d'une espèce est plus petite que sa niche fondamentale car elle ne comprend que les portions de niche fondamentale que l'organisme occupe réellement, résultant de l'exclusion compétitive et autres paramètres liés à la dynamique de l'espèce (Hutchinson, 1957). Compte tenu de la disponibilité des données et de l'échelle spatiale considérée, des modèles statistiques et statiques basés sur la niche réalisée des espèces ont été développés afin d'expliquer et de projeter la probabilité d'occurrence des 13 espèces de poisson les plus communes sur la région d'étude, à partir des caractéristiques hydro-climatiques et géomorphologiques des sites d'étude (Figure 3d).

5.2.2 Choix des prédicteurs hydro-climatiques et géomorphologiques

Deux types de descripteurs environnementaux, interagissant à différentes échelles spatiales ont été définis pour décrire la niche réalisée et individuelle de chaque espèce de poisson.

Les descripteurs géomorphologiques de l'habitat constituent les limites biogéographiques des espèces à large échelle spatiale et résultent de l'extraction des deux premiers axes d'une ACP appliquée aux caractéristiques géomorphologiques des sites d'étude comme : la distance à la source, la surface du bassin versant, l'altitude et les coordonnées géographiques. Le premier axe (A1 ; 60 % de variance expliquée) positionne les sites d'étude le long du continuum amont-aval alors que le deuxième axe (A2 ; 20% de variance expliquée) caractérise un gradient continental sud-ouest/nord-est.

Les descripteurs hydro-climatiques caractérisent les conditions saisonnières de variabilité des débits et des températures. Les saisons considérées représentent les périodes clés dans l'accomplissement du cycle de vie de la majorité des poissons étudiés : la période de faible activité hivernale (octobre – mars), de reproduction (avril – juin) et de croissance (juillet – septembre). Dans le cas de la truite commune (*Salmo trutta*), cette classification saisonnière reste valable mais n'a pas la même signification biologique car l'espèce fraye durant la période hivernale. Pour chacune des saisons, les conditions de variabilité hydrologique et de température sont caractérisées par quatre variables statistiques. Les percentiles 10%, 50% et 90% (P10, P50 et P90) soulignent le profil saisonnier minimum, moyen et maximum des débits et températures. Une variable hydro-climatique de variation saisonnière a également été définie comme la différence entre le profil saisonnier maximum et minimum.

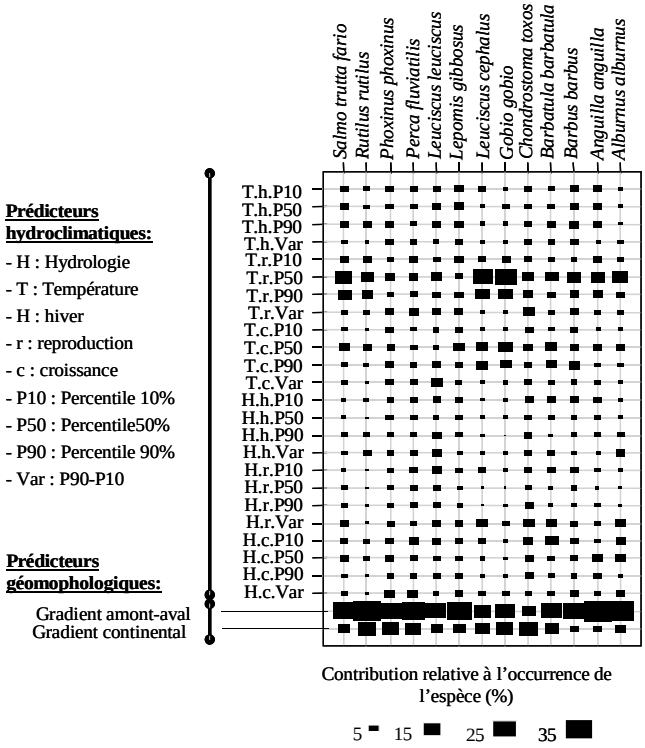


Figure 5. Contribution relative (proportionnelle à la grosseur des carrés) de chaque prédicteur hydro-climatique et géomorphologique pour expliquer la probabilité d'occurrence de chacune des 13 espèces. La contribution relative a été calculée lors de la calibration des modèles de distribution d'espèce basée sur les forêts adaptatives.

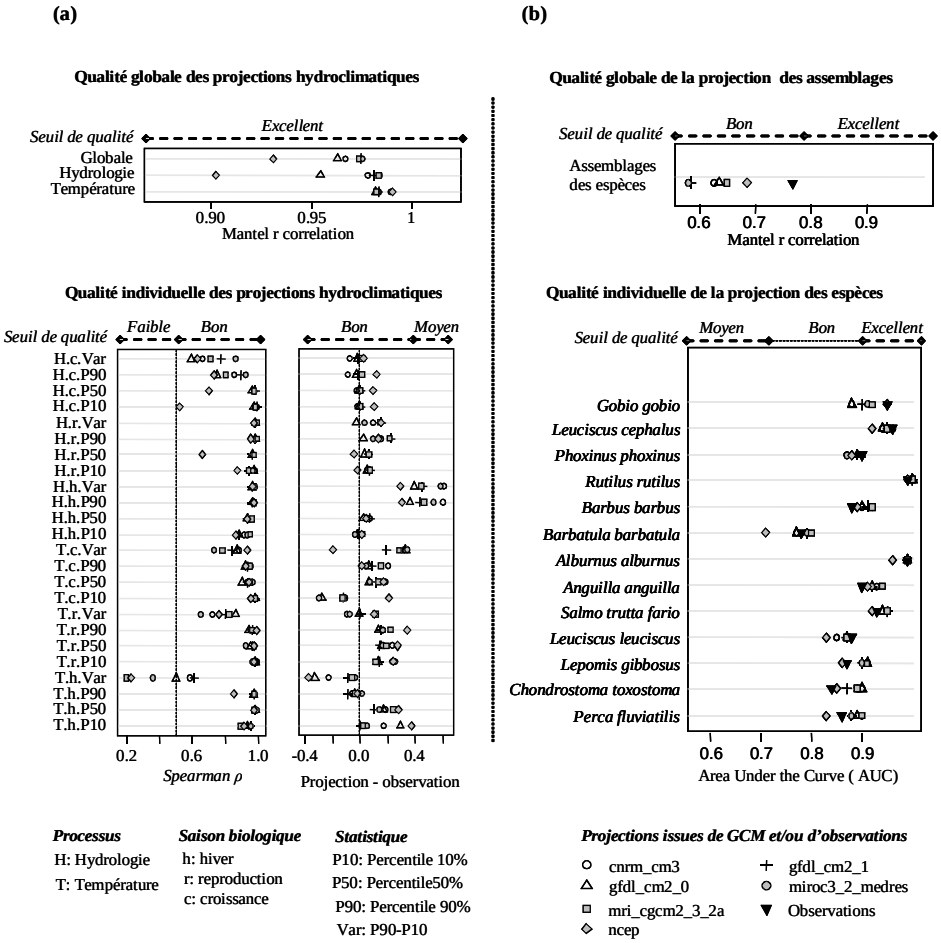


Figure 6. Validation des projections hydro-biologiques sur le climat présent (scénario 20c3m) en fonction de cinq modèles de circulation générale (GCM) : (a) validation globale (Mantel r) et individuelle (Spearman ρ) pour chaque projection, du downscaling des conditions hydro-climatiques saisonnières; (b) validation globale (Mantel r) et individuelle pour chacune espèce (critère AUC) des projections des modèles de distribution d'espèces.

Au total, les 12 variables hydrologiques saisonnières (3 saisons × 4 statistiques), les 12 variables de température saisonnières (3 saisons × 4 statistiques) ainsi que les deux variables géomorphologiques (les 2 axes A1 et A2) ont été utilisées comme prédicteurs en entrée d'un modèle statique basé sur les forêts adaptatives, calibré individuellement pour chaque espèce afin de prédire leur probabilité d'occurrence sur les 276 sites annuels (50 sites × 5.5 années).

La structure hiérarchique des forêts adaptatives semble particulièrement adaptée à la nature des descripteurs environnementaux du modèle biologique étudié, eux-mêmes structurés de manière hiérarchique dans l'espace. Par ailleurs, les forêts adaptatives offrent la possibilité de mieux comprendre, quantitativement et qualitativement, la nature des relations entre les prédicteurs et la réponse de chaque espèce individuellement. Les résultats des modèles de distribution d'espèces illustrent les différences de sensibilité des poissons aux différents descripteurs environnementaux, ce qui justifie d'autant plus la construction de modèles individuels pour chaque espèce (Figure 5a).

5.3 Validation des projections hydro-biologiques sur la période contrôle

Dans l'Article n°3 sur la validation du modèle hydro-biologique, les projections ont été validées sur la période contrôle selon 5 GCM. La question principale était de savoir si le downscaling des GCM était capable de reproduire les patrons actuels de variabilité spatiale, observés dans l'hydrologie, les températures saisonnières ainsi que dans la distribution des espèces. Bien que très largement négligée dans la plupart des études sur l'impact du changement climatique sur la biodiversité, cette étape de validation des GCM sur le climat présent est fondamentale et indispensable avant toute extrapolation dans le futur.

Les deux résultats principaux tirés de cette étude montrent la bonne capacité des modèles de downscaling hydro-climatique (Figure 6a ; Spearman $\sigma > 0.6$ et Mantel $r > 0.9$) et de distribution d'espèces (Figure 6b ; AUC > 0.7 et Mantel $r > 0.6$) à reproduire les patrons actuels de variabilité spatiale. Ces résultats permettent ainsi de valider l'utilisation du modèle hydro-biologique en vue de projeter la variabilité hydro-climatique ainsi que la distribution potentielle des 13 espèces dans le futur.

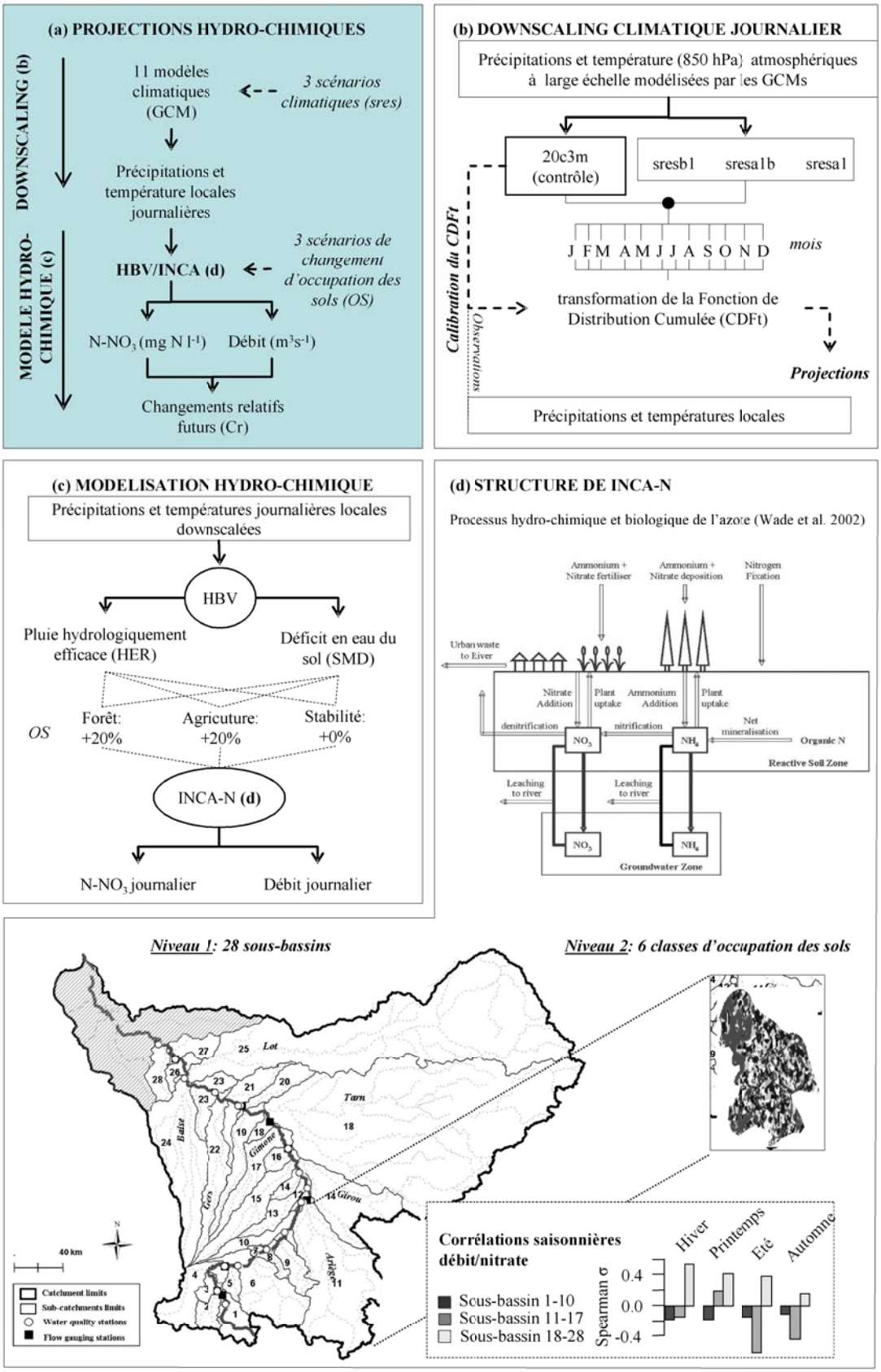


Figure 7. (a) Schéma conceptuel de la modélisation hydro-chimique pour les projections futures des débits et des concentrations en nitrates sur le bassin versant de la Garonne, dont la structure est composée: (b) d'un modèle de downscaling des conditions climatiques journalières (précipitations, température) sur le bassin à l'aide de la méthode CDFt; (c) du modèle hydrologique HBV couplée au modèle INCA-N pour les projections hydro-chimiques selon trois scénarios d'occupation des sols (OS) prévoyant d'ici 2100, soit une augmentation progressive de 20% du territoire en surfaces boisées (Forêt) ou agricoles (Agriculture), soit une stabilité de l'occupation des sols actuelle; (d) L'ensemble des processus hydro-chimiques et biologiques de l'azote est intégré dans INCA-N au sein d'une structure à 2 niveaux d'échelle spatiale. Au niveau 1, le bassin Adour Garonne est décomposé en 28 sous-bassins. Au niveau 2, chaque sous-bassin est subdivisé en 6 classes d'occupation des sols (forêt, prairies, céréales, oléagineux, autres cultures et urbanisme).

6 Modèle de downscaling hydro-chimique

Le modèle de downscaling hydro-chimique appliqué sur la Garonne (Figure 7a) repose sur le couplage entre un modèle de downscaling des températures et précipitations journalières futures (Figure 7b), avec le modèle dynamique HBV/INCA-N (Figure 7c et d) pour la projection des débits et des concentrations en nitrates sur la Garonne. La calibration et la validation du modèle HBV/INCA-N sur la Garonne à partir de données historiques ont fait l'objet d'un Article n°1 publié dans *Journal of Environmental Quality*. Quant à la validation et à la projection future des débits et des nitrates, les résultats présentés dans cette thèse sont préliminaires et aucun article n'est pour l'instant en cours de préparation.

6.1 Downscaling des précipitations et températures journalières

Le modèle HBV/INCA-N nécessite des séries journalières de température et de précipitation comme variables d'entrées. Le downscaling de ces deux variables climatiques a été réalisé à l'aide de la méthode du CDFt, en ajustant leur distribution de probabilité issue des GCM sur la période de contrôle (scénario 20c3m) à celle des variables localement mesurées entre 1970 et 2005. Afin d'ajuster plus finement les projections saisonnières, le CDFt a été appliqué séparément selon chaque mois de l'année. L'ensemble de la procédure a été appliqué pour chacune des sept stations climatiques du bassin de la Garonne qui déterminent le fonctionnement hydrologique du modèle HBV/INCA-N (Figure 7b).

6.2 Modèle hydro-chimique HBV/INCA-N

6.2.1 Principe

Le modèle HBV/INCA-N résulte du couplage entre le modèle hydrologique HBV (Lindstrom *et al.* 1997) et le modèle hydro-chimique INCA-N (Whitehead *et al.* 1998; Wade *et al.* 2002) (Figure 7c). HBV assure la balance hydrique au sein du bassin versant au travers des processus de précipitation liquide ou neigeuse, d'interception, d'infiltration, d'évaporation et de ruissellement. Le cycle de l'azote est modélisé par le modèle INCA-N qui prend en compte l'occupation des sols et les processus biologiques de transformation et de fixation de l'azote, ainsi que les apports ponctuels dans la rivière ou diffus sur le bassin. La structure de HBV/INCA-N est dite « semi-distribuée » car la modélisation journalière des flux (hydrologie, azote) n'est pas calculée de manière spatialement continue le long du linéaire de la Garonne, mais au niveau de l'exutoire de 28 sous-bassins identifiés sur la Garonne (Niveau 1, Figure 7d)

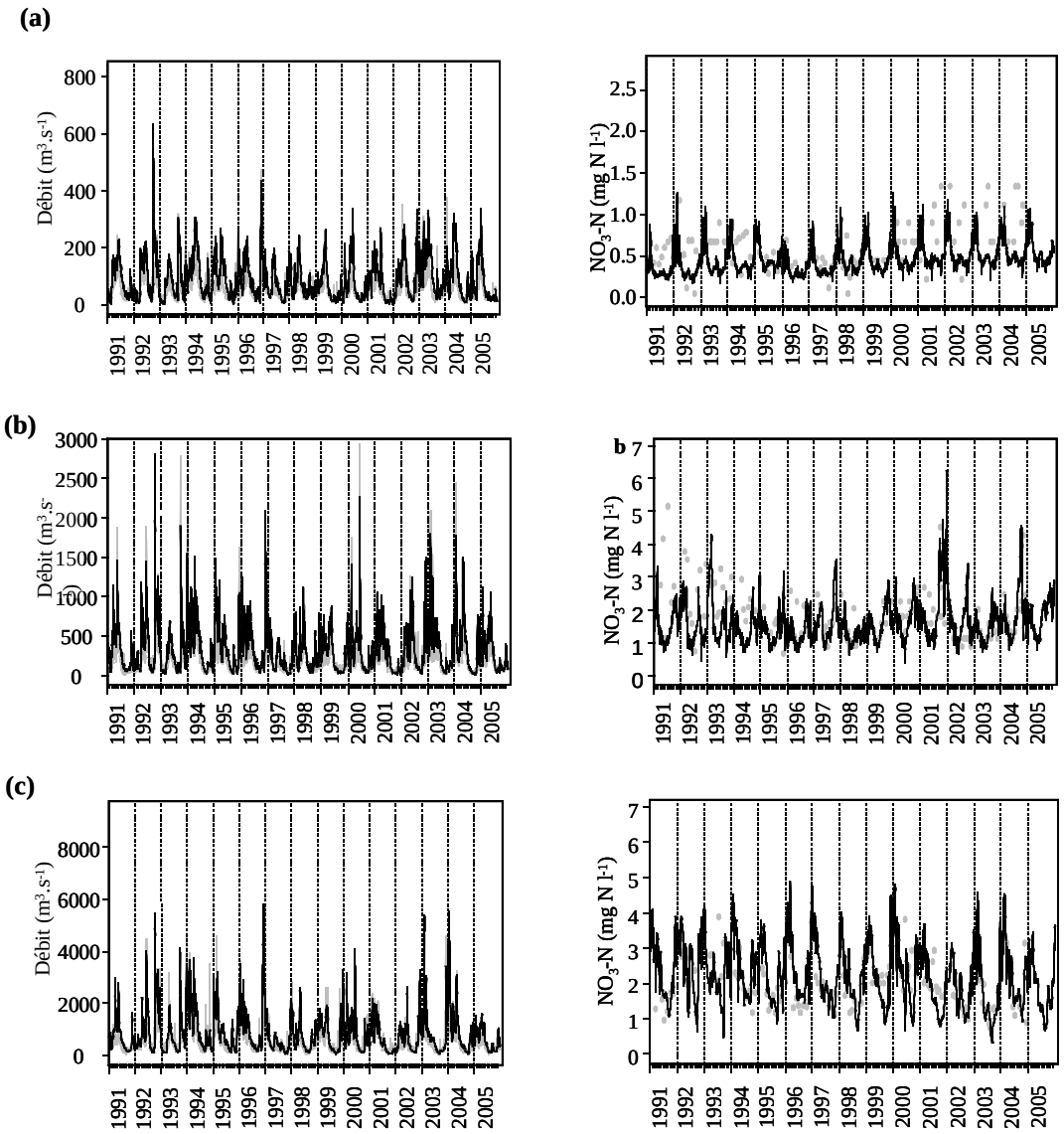


Figure 8. Calibration du modèle HBV/INCA-N sur le bassin de la Garonne. Les débits et les concentrations en nitrates sont modélisés (noir) sur un pas de temps journalier et comparés aux observations (gris) au niveau des zones amont (a; station 4), médiane (b; station 16) et aval (c; station 28) de la Garonne.

.

Pour chacun des 28 sous-bassins, six classes d'occupation des sols ont été définies : les cultures de céréales, cultures d'oléagineux et autres cultures, forêts, prairies et territoires urbanisés. Pour chaque classe d'occupation des sols de chacun des sous-bassins, l'ensemble des flux de matières et des processus biochimiques sont intégrés et synthétisés au travers du bassin versant (Niveau 2, Figure 7d). Une interface de commandes a été développée en langage R pour faciliter la création des fichiers d'entrée pour HBV et INCA-N (version 1.11), le contrôle des exécutables de chacun des modèles, ainsi que l'extraction des résultats des simulations (librairie 'HBV-INCA', disponible sur demande).

6.2.2 Dynamique de l'azote et simulations du modèle INCA-N

La compréhension de la dynamique entre l'hydrologie et l'azote est un pré-requis à la calibration et la validation de tout modèle hydro-chimique (Article n°1). Avec une surface de l'ordre de 60 000 km², le bassin de la Garonne se caractérise par un régime hydrologique et une occupation des sols très hétérogènes, ce qui rend la dynamique de l'azote elle aussi très hétérogène en terme de variations saisonnières. Au niveau du piémont pyrénéen (Figure 7d ; sous-bassins 1-10), le régime nivo-pluvial de la Garonne se traduit par une forte augmentation des débits au moment de la fonte des neiges printanières. Forêt et prairies dominent essentiellement la surface du bassin, ce qui limite les apports azotés diffus vers les rivières. Dans cette zone, une forte augmentation des débits s'accompagne donc généralement d'une dilution des concentrations en nitrates (Figure 7d ; corrélation débit-nitrates négative). A l'inverse, la partie de plaine en aval de la Garonne (Figure 7d ; sous-bassins 18-28) se caractérise par un régime hydrologique de type pluvial où les débits maximums annuels ont lieu en hiver. Le territoire est fortement dominé par l'agriculture (>60%) et les fortes précipitations hivernales sont en grande partie responsable d'un intense processus de lessivage des sols vers les rivières. L'azote épandu sur les cultures entre l'automne et le printemps est à exporté par ruissellement vers les rivières et percole également vers les couches inférieures du sol et les nappes phréatiques, comme en témoigne la corrélation positive en hiver entre les débits et les concentrations en azote (Figure 7d).

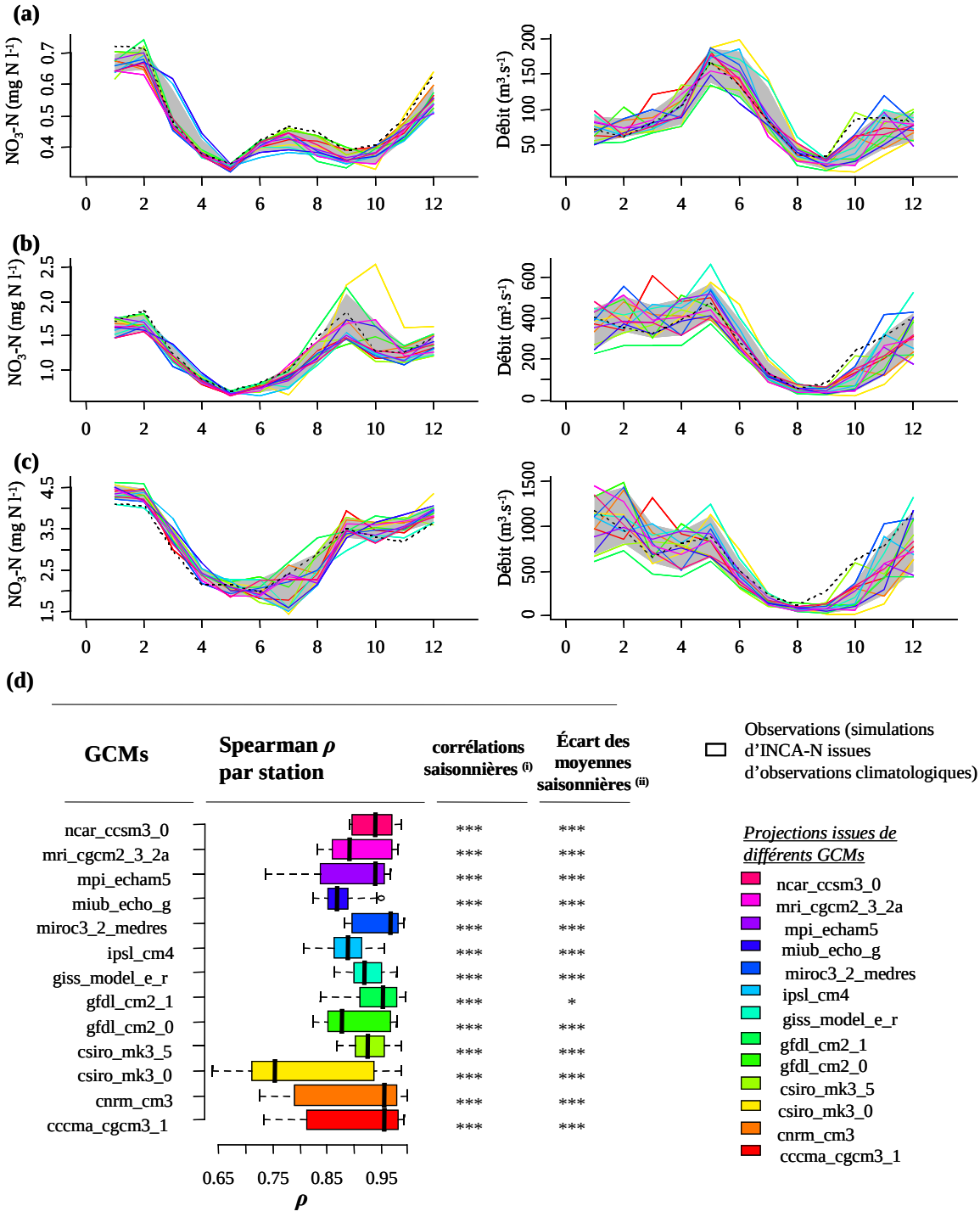


Figure 9. Validation des projections hydro-chimiques du modèle INCA-N sur la période contrôle (scénario 20c3m) issues du downscaling des condition climatiques pour 13 GCM : comparaison des moyennes mensuelles des débits et des nitrates sur les zones amonts (a ; station 4), médianes (b ; station 16) et aval (c ; station 26) de la Garonne ; (d) Pour chaque GCM, les corrélations de Spearman ρ entre les projections mensuelles des GCM et les simulations d'INCA issues d'observations climatologiques, ont été calculées sur les 28 stations du linéaire de la Garonne.

6.3 Validation des projections hydro-chimiques

La calibration et la validation de HBV/INCA-N ne peuvent se faire sans une bonne connaissance des processus saisonniers dans la dynamique débit-nitrates évoquée précédemment. Le modèle hydro-chimique a été calibré entre 1990 et 1999 en s'assurant de la cohérence des résultats au regard des bilans azotés calculés pour chaque type d'occupation de sol ainsi que dans la dynamique saisonnière modélisée (Figure 8 ; Article n°1).

Les conditions hydro-chimiques ont été projetées sur la période contrôle (scénario 20c3m) de 13 GCM sélectionnés (Figure 9). Globalement, les tendances mensuelles de débits et de nitrates sont projetées de manière relativement consensuelles entre les différents GCM (Figure 9a,b,c ; couleurs). Les projections sont aussi significativement concordantes par rapport aux simulations d'INCA basées sur les observations climatologiques (Figure 9a, b, c ; pointillés noirs) au niveau des stations en amont (station 4 ; Figure 9a), médianes (station 16, Figure 9b) et en aval de la Garonne (station 26, Figure 9c). En moyenne, les corrélations de Spearman calculées au niveau des 28 stations par rapport aux observations sont significatives pour les 13 GCM, et les moyennes sont égales (Figure 9d ; $p>0.01$). Ces résultats permettent ainsi de valider l'utilisation du modèle hydro-chimique en vue de projeter et d'analyser la variabilité des débits et des nitrates dans le futur.

2IÈME PARTIE : PROJECTIONS FUTURES ET INCERTITUDES

1 MÉTHODE

Dans cette partie, les modèles hydro-biologiques et hydro-chimiques décrits et validés dans la partie précédente, ont été utilisés afin d'explorer la manière dont le changement climatique futur pourrait modifier les débits et les concentrations saisonnières en nitrates, ainsi que la biodiversité des poissons sur la région étudiée. Conceptuellement, l'ensemble des projections a tout d'abord été analysé de manière à quantifier les sources de variations potentielles dans les projections futures (ex : les projections issues des différents GCM ou scénarios climatiques sont-elles consensuelles dans le futur ? Des zones et/ou des périodes sont-elle plus exposées que d'autres à des perturbations écologiques ?) En se basant sur la compréhension des différentes sources de variation, une deuxième étape s'est intéressée à représenter de manière synthétique, spatialement (ex gradient amont-aval) et temporellement (ex années ou saisons) explicite, les patrons de variation dans les projections futures (ex : peut-on mettre en évidence différents patrons de variation en fonction des scénarios climatiques ou d'occupation des sols ?)

1.1 INDICATEURS DE BIODIVERSITÉ ET DE CHANGEMENTS HYDRO-CHIMIQUES

Au travers des projections de la distribution potentielle future des 13 espèces de poisson étudiées, trois types d'indicateurs ont été utilisés afin de caractériser les changements de biodiversité, au niveau de la communauté des poissons (combinaison des projections individuelles de chaque espèce). Pour chaque site/année de projection, la diversité alpha (richesse spécifique) a été calculée, c'est-à-dire le nombre d'espèces potentiellement présentes dans le futur. La diversité beta a quant à elle été calculée pour mesurer la similarité entre les sites dans le futur, au niveau de leur composition d'espèces. Pour ce faire, la mesure de dissimilarité de Jaccard (1901), variant entre 0 et 1, a été calculée par année entre chaque paire de sites, indiquant si les sites sont faiblement (valeur égale à 0) ou fortement (valeur égale à 1) similaires entre eux dans leur composition en espèces (Sax & Gaines 2003). Le 1ier axe d'une analyse de redondance (RDA) a permis de synthétiser la structure des communautés le long du gradient du gradient amont-aval (1ier axe = 70%). La RDA réalise un couplage linéaire entre une matrice de prédicteurs *X* (ici, le gradient géomorphologique amont-aval et les années) et une matrice de réponse *Y* (ici, la matrice de probabilité

d'occurrence des espèces projetée), en se basant sur des calculs de distance Euclidienne pour les deux matrices.

Les projections futures du modèle INCA-N ont été analysées en terme de changement relatif (CR) dans les débits et les concentrations en nitrates futur (Ftr), par rapport à des conditions de référence (Réf ; 2005-2010):

$$CR = \frac{(Ftr - Réf)}{Réf}$$

Par exemple, une valeur du CR de +0.5 dans les changements relatifs en nitrates, signifie une augmentation de 50% des concentrations par rapport concentrations actuelles.

1.2 Partitionnement de la variabilité dans les projections

Pour chaque indicateur de biodiversité (ex. diversité beta) ou de changement hydro-chimique (ex. CR des nitrates) projeté, une étape de partitionnement de la variation a été réalisée afin de quantifier l'influence relative des différents facteurs potentiels (ex. scénarios climatiques, GCM). Cette analyse est essentielle car elle permet de discuter la part d'incertitude, et donc de crédibilité, que l'on peut accorder à l'analyse des projections futures (ex. est-ce que tous les GCM sont d'accord entre eux ?)

Cinq facteurs ont été considérés dans le cadre du partitionnement de la variabilité des projections de chacun des indices de biodiversité (ex. diversité beta): (i) la structure spatiale des peuplements caractérisée par le positionnement des sites sur le gradient amont-aval (1^{ier} axe d'ACP des caractéristiques géomorphologiques des bassins; voir Partie 1, Section 5.2.2) ; (ii) les années, de 2005 à 2100, caractérisant la tendance interannuelle du changement climatique ; (iii) cinq GCM; (iv) trois scénarios climatiques du plus au moins pessimistes, A2, A1B et B1 ; (v) la répétition aléatoire (10 fois) de la construction des forêts aléatoires lors de la modélisation statique de la distribution d'espèces.

Quant au partitionnement des changements relatifs dans les concentrations en nitrates et les débits, six facteurs potentielles de variation ont été considérés : (i) le gradient amont-aval caractérisé par la surface cumulée des 28 bassins versants définissant la structure du modèle INCA-N (voir Figure 7 ; Partie 1, Section 6.2.1) ; (ii) le cosinus et sinus de chacun des 12 mois de l'année reflétant les cycles mensuels; (iii) trois périodes de temps, 2005-2010 (présent), 2048-2052 et 2095-2100 ; (iv) 13 GCM différents ; (v) trois scénarios climatiques (A2, A1B et B1) ; (vi) trois scénarios futurs de changement d'occupation des sols d'ici à

2100 : (a) stabilité de l'occupation des sols actuelle, (b) augmentation progressive et homogène de 20% des surfaces agricoles, (c) augmentation de 20% des zones pastorales et forestières. Etant donné que les projections hydro-chimiques sont des résultats encore préliminaires à l'heure actuelle, il est important de noter que les trois périodes de temps considérées sont vraisemblablement trop courtes (cinq années) pour prétendre analyser de manière robuste leurs variations saisonnières (au moins 20 années seraient requises). Par ailleurs, les scénarios d'occupation des sols sont basés sur des hypothèses arbitraires, sans réels fondements socio-économiques, et dont le seul intérêt est d'appréhender l'amplitude des réponses possibles d'INCA-N dans un contexte de climat futur.

1.2.1 Quantifier l'influence relative des différents facteurs

Le partitionnement de la variabilité dans les projections futures a été effectué en utilisant l'approche de partitionnement hiérarchique (Chevan & Satherland 1991 ; librairie *hier.part* dans R). Elle consiste à estimer la variabilité d'une réponse en fonction de l'effet indépendant (marginal) et joint d'un (ou plusieurs) prédicteur(s). Le partitionnement hiérarchique est classiquement construit sur une succession de modèles linéaires généralisés (GLM), supposant ainsi une relation linéaire entre les prédicteurs et la réponse. Ici, la méthode a été adaptée aux modèles additifs généralisés (GAM) afin de pouvoir modéliser la non-linéarité possible entre la réponse (ex. la diversité beta) et les prédicteurs (les années et le gradient amont-aval). La variance expliquée par l'effet joint et indépendant de chaque prédicteur se traduit par une valeur de R^2 comprise entre 0 (faible effet) et 1 (toute la variance est expliquée par ce prédicteur). La variance totale expliquée par un facteur (ex. scénario climatique) dans le modèle hiérarchique peut être perçue comme la variance 'inter-groupe', c'est-à-dire, la variance quantifiée entre les moyennes de chaque modalité (ex. B1, A1B, A2). Au contraire, la fraction inexpliquée (résidus) du partitionnement hiérarchique peut être attribuée, soit à l'erreur du modèle GAM, soit à la variance 'intra-groupe'. La variance 'intra-groupe' correspond à la variabilité au sein de chaque modalité de ce facteur (ex. scénarios climatiques), comme par exemple des patrons de variabilité spatio-temporelle différents selon les scénarios B1, A1B et A2.

1.2.2 Quantifier la variabilité spatiale et temporelle des différents facteurs

Estimer l'influence relative de différents facteurs sur la variabilité des projections est essentiel. Par ailleurs, une étape supplémentaire consiste à analyser des patrons de variation spatio-temporelle relatifs à chacun de ces facteurs, étant donné le contexte spatial et temporel

dans lequel les projections ont été générées. Cette analyse peut par exemple aider à évaluer si des secteurs le long du gradient-amont et/ou des époques particulières, enregistrent plus de variabilité que d'autres au sein d'un facteur (ex. les projections sont-elles consensuelles entre les différents GCM le long du gradient amont-aval ? L'impact du changement global est-il le même entre les différentes périodes ?).

Pour évaluer cette variabilité spatiale et temporelle, le coefficient de variation (CV) de chaque projection (ex. richesse) a été calculé séparément pour chaque facteur (ex. scénarios climatiques), de manière spatialement et temporellement explicite (ex. à chaque site/année dans le cas des projections hydro-biologiques et à chaque station/mois dans le cas des projections hydro-chimiques). Le coefficient de variation résulte de la division entre l'écart-type et la moyenne de la projection étudiée (ex. diversité beta). Dans la situation où la moyenne et l'écart-type de la projection (ex. diversité beta) peuvent varier fortement dans l'espace et dans le temps, le CV présente l'avantage de normaliser cette variation et de fournir une information moins biaisée que celle fournie par l'écart-type. Le calcul du CV a été réalisé en deux temps: (i) la projection (ex. richesse) est d'abord moyennée pour chaque modalité de ce facteur (ex B1, A1B, A2) ; (ii) le CV est ensuite calculé entre les différentes modalités de ce facteur.

1.3 Patrons de variation spatio-temporelle dans les projections

Une image synthétique de l'hétérogénéité spatiale et temporelle des changements pouvant affecter la biodiversité (ex. diversité beta) et les processus hydro-chimiques (ex. CR des nitrates), ainsi que la variabilité (CV) liée à chacun de leurs facteurs de variation (ex. GCM), a été représentée à l'aide de modèles additifs généralisés (GAM). Les projections ont pour cela été définies dans un plan factoriel à trois dimensions (x, y, z), où x (gradient amont-aval des rivières) et y (années ou mois de l'année) caractérisent les dimensions spatiales et temporelles de la projection z (ex. diversité beta). Le modèle GAM a ensuite été appliqué, en utilisant x et y comme prédicteurs afin de lisser la projection z à l'aide d'une fonction spline de lissage, s, spécifique à chaque prédicteur (ici de type « thin plate ») et de dimension k adéquate (généralement k=4) tel que :

$$z = s_x(x|k) + s_y(y|k)$$

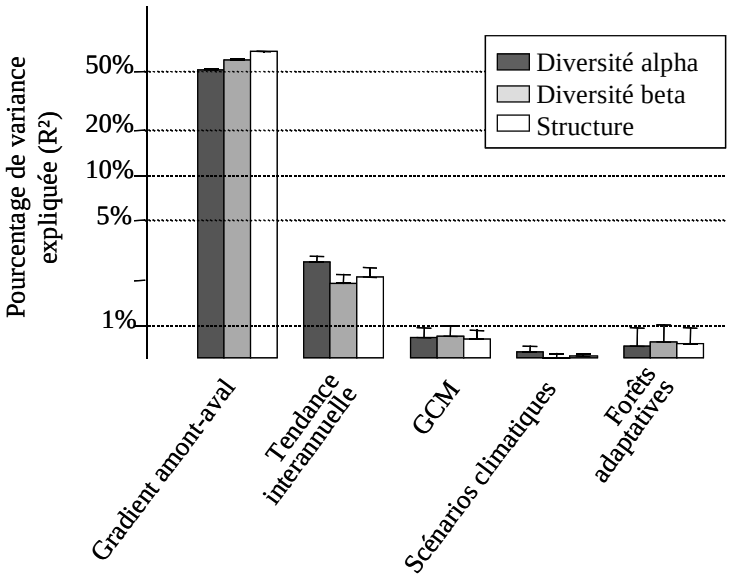


Figure 10. Partitionnement hiérarchique de la variabilité dans les projections hydro-biologiques en fonction du gradient amont-aval, de la tendance interannuelle, de cinq GCMs, trois scénarios climatiques et des 10 modèles itératifs (forêts adaptatives). Chaque diagramme caractérise le % de variances expliquée individuellement par chaque facteur (R^2) dans les projections de diversité alpha (richesse; gris foncé), diversité beta (similitude des communautés entre sites; gris clair) et la structure des communautés (1° axe d'une analyse de redondance; blanc).

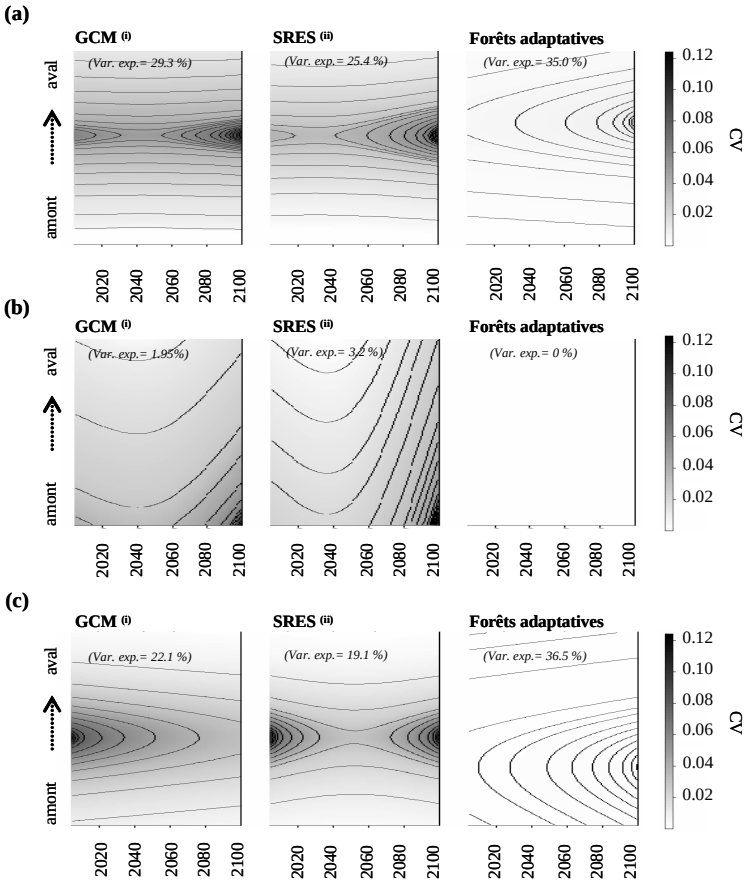


Figure 11. Variabilité spatiale (gradient amont-aval) et temporelle (années de 2005 à 2100) de l'incertitude entre les GCM, scénarios climatiques et forêts adaptatives, mesurée à l'aide du coefficient de variation (CV) dans les projections hydro-biologiques : (a) diversité alpha (richesse) ; (b) diversité beta (similitude des communautés entre sites) ; (c) structure des communautés (1° axe d'une analyse de redondance).

2 Changements dans la biodiversité des peuplements de poissons

Globalement, les cinq facteurs de variabilité étudiés expliquent plus de 60% de la variabilité dans les projections biologiques de diversité alpha (Figure 10 ; gris foncé), beta (Figure 10 ; gris clair) et de la structure des communautés (Figure 10 ; blanc). Les résultats du partitionnement hiérarchique (non montrés) soulignent que chacun de ces cinq facteurs possède un effet nettement indépendant des autres facteurs. Les patrons de variation dans les projections sont fortement spatialisés le long du gradient amont-aval des rivières (Figure 10 ; plus de 50% de la variance expliquée) et les tendances interannuelles du changement climatique expliquent en moyenne 4% de la variabilité totale (Figure 10 ; Tendance interannuelle). En revanche, les différences entre GCM, scénarios climatiques et forêts adaptatives expliquent globalement moins de 1% de la variabilité totale dans les projections de biodiversité.

Les 40% de variabilité inexpliquée par les cinq facteurs peuvent résulter, entre autres, de la variabilité spatiale et temporelle au sein de chacun des facteurs (variabilité intra-groupe ; Figure 11). Concernant les projections de diversité alpha (Figure 11a) et de structure des communautés (Figure 11c), la variabilité (CV) au sein des GCM et des scénarios climatiques est la plus élevées au niveau des zones intermédiaires du gradient amont-aval et tendent à augmenter continuellement avec le temps, surtout à partir de la deuxième moitié du siècle. Quant aux projections de diversité beta, les patrons de variabilité au sein chaque facteur (GCM, scénarios climatiques et forêts adaptatives) n’apparaissent que faiblement structurés dans l’espace et dans le temps (Figure 11b).

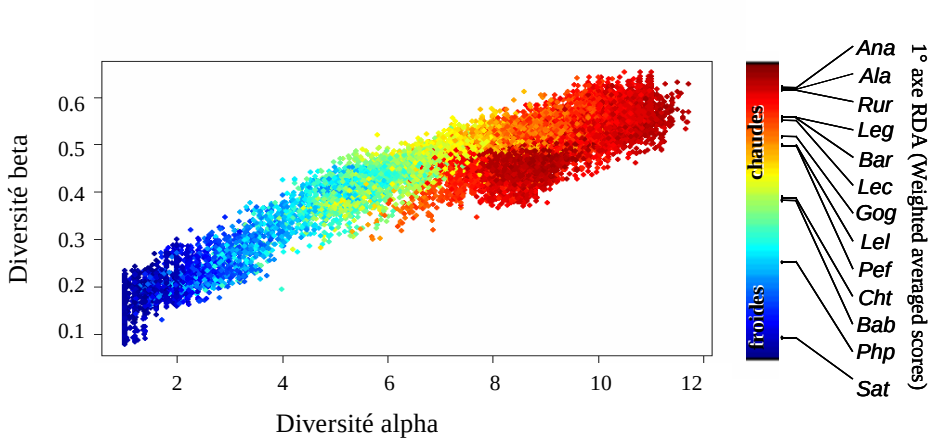


Figure 12. Relation entre la diversité alpha (richesse), beta (similarité des communautés entre sites) et la composition des communautés (1° axe d'une analyse de redondance) établie à partir des projections de la distribution potentielle future de 13 espèces de poisson.

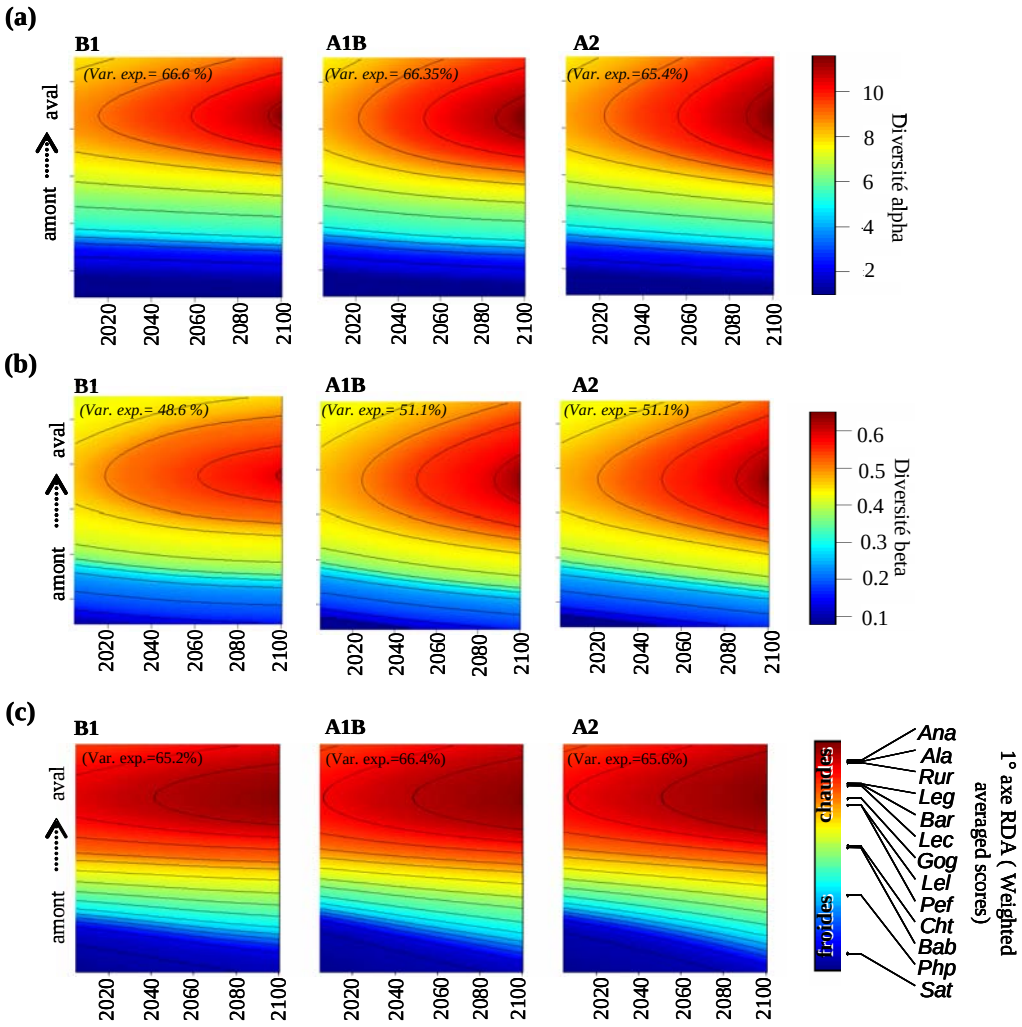


Figure 13. Patrons de variabilité spatiale (le long du gradient amont-aval) et temporelle (de 2005 à 2100) dans les projections hydro-biologiques en fonction de trois scénarios climatiques (sres) d'émission de gaz à effet de serre d'intensité croissante, respectivement B1, A1B et A2, calculés pour :(a) diversité alpha (richesse) ; (b) diversité beta (similarité des communautés entre sites) ; (c) structure des communautés (1° axe d'une analyse de redondance).

Les patrons spatio-temporels entre les différents indices de biodiversité projetés (diversité alpha (richesse), beta, et la structure des communautés) sont fortement corrélés entre eux (Figure 12 ; R Pearson = 0.92 ±0.04), et mettent en évidence deux résultats principaux (Figure 13).

Premièrement, tous les indices de biodiversité s'accordent à dire que des changements sont incontournables, quels que soient les scénarios climatiques. Ces changements de biodiversité pourraient se traduire par une augmentation globale de la richesse, en particulier sur la moitié aval du gradient où le nombre moyen d'espèces pourrait passer de 8 à 11. A l'inverse, les zones à l'amont pourraient perdre en moyenne une espèce, passant de 5 à 4 (Figure 13a). La diversité beta aurait tendance à diminuer, notamment sur la moitié aval du gradient, indiquant que les assemblages d'espèces pourraient devenir plus similaires entre les sites. (Figure 13b). Cette tendance traduit un phénomène d'homogénéisation taxonomique que les changements dans la structure des communautés permettent de mieux éclairer (Figure 13c). Les modifications hydrologiques et climatiques auraient tendance à favoriser l'expansion des espèces d'eau chaude, comme l'anguille (*Ana)*, l'ablette (*Ala)* ou le gardon (*Rur*), vers des sites situés plus en amont du gradient (Figure 11c, en rouge). Au contraire, les espèces d'eau froide comme la truite commune (*Sat*) ou le vairon (*Php*), pourraient voir leur aire distribution se retreindre car elles pourraient être incapables de trouver des sites où l'habitat deviendrait favorable hydrologiquement et thermiquement (Figure 11c, en bleu).

Le deuxième résultat important montre que l'intensité des changements sur la biodiversité est sensiblement équivalente entre les trois scénarios climatiques, durant la première moitié du 21^{ième} siècle. En revanche, ce n'est qu'à partir de la deuxième moitié du siècle que des différences apparaissent entre scénarios climatiques. L'intensité des perturbations se poursuit de manière plus marquée et rapide dans le cas du scénario le plus pessimiste (A2), que pour les deux autres scénarios, A1B et surtout B2.

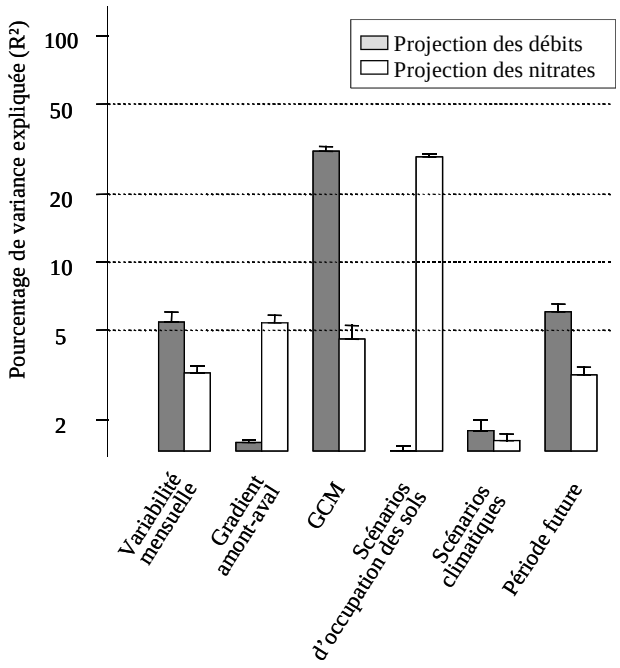


Figure 14. Partitionnement hiérarchique de la variabilité dans les projections des débits (gris) et des nitrates (blanc) issues du modèle INCA-N en fonction de la variabilité mensuelle, du gradient spatial amont-aval, de 13 GCMs, trois scénarios d'occupation de sol, trois scénarios climatiques et deux périodes de temps (2048-2052 et 2095-2100). Chaque diagramme caractérise le % de variance expliquée (R^2) individuellement par chaque facteur de variation.

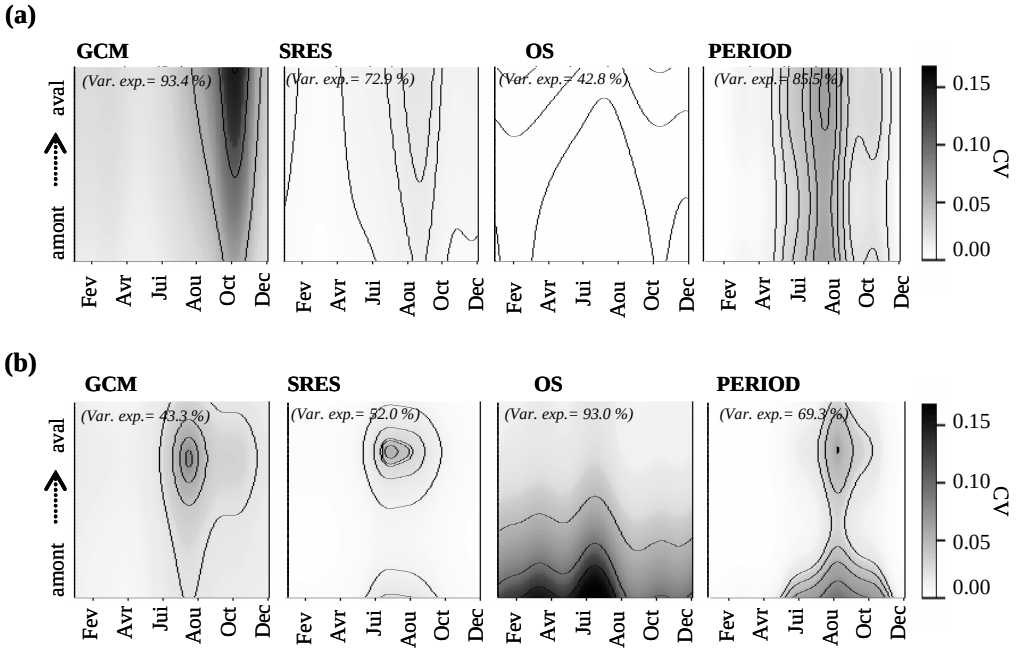


Figure 15. Variabilité spatiale (gradient amont-aval) et temporelle (mois) du coefficient de variation (CV) dans le changements relatif des débits (a) et des nitrates (b), calculé entre 13 modèles climatiques (GCM), trois scénarios climatiques (SRES), trois scénarios d'occupation des sols (OS) et deux périodes (2048-2052 et 2095-2100).

3 Modification de la dynamique hydro-chimique sur la Garonne

Le partitionnement de la variabilité dans les changements relatifs des débits et nitrates suggère que les changements futurs pourraient tout aussi toucher les zones amont que les zones aval du bassin de la Garonne, compte tenu que seulement 3% de la variation totale est expliquée par le gradient amont-aval (Figure 14). En revanche le cycle mensuel dans la dynamique de l'hydrologie et des nitrates pourrait être sensiblement perturbé car la variabilité mensuelle explique près de 5% de la variabilité totale (Figure 14). Concernant les changements relatifs des débits, la variabilité entre les 13 GCM explique plus de 25% de la variation totale, ce qui constitue la plus grande source d'incertitude dans la projection des changements relatifs des débits (Figure 14 ; gris). Les changements relatifs des nitrates sont quant à eux particulièrement sensibles aux scénarios de changement d'occupation des sols qui expliquent approximativement 30% de la variabilité totale (Figure 14 ; blanc). Par ailleurs, les patrons de changement relatif dans les débits et les nitrates sont sensiblement différents entre les périodes 2048-2050 et 2095-2100 ainsi qu'entre les trois scénarios climatiques, expliquant respectivement 5 et 3% de la variabilité totale (Figure 14).

L'analyse des patrons de variation spatiale (gradient amont-aval) et temporelle (mois) dans la variabilité des changements relatifs des débits et des nitrates (CV) indique que les différences entre les deux périodes futures étudiées seraient sensiblement plus marquées en été (Figure 15). La variabilité dans les changements relatifs des débits souligne l'incertitude particulièrement forte entre les GCM au moment de l'étiage (été-automne), notamment au niveau des parties aval de la Garonne (Figure 15a). Quant aux changements relatifs en nitrates, ce sont les secteurs du piémont pyrénéen (parties amont de la Garonne) qui enregistrent le plus de sensibilité aux scénarios de changement d'occupation des sols (Figure 15b).

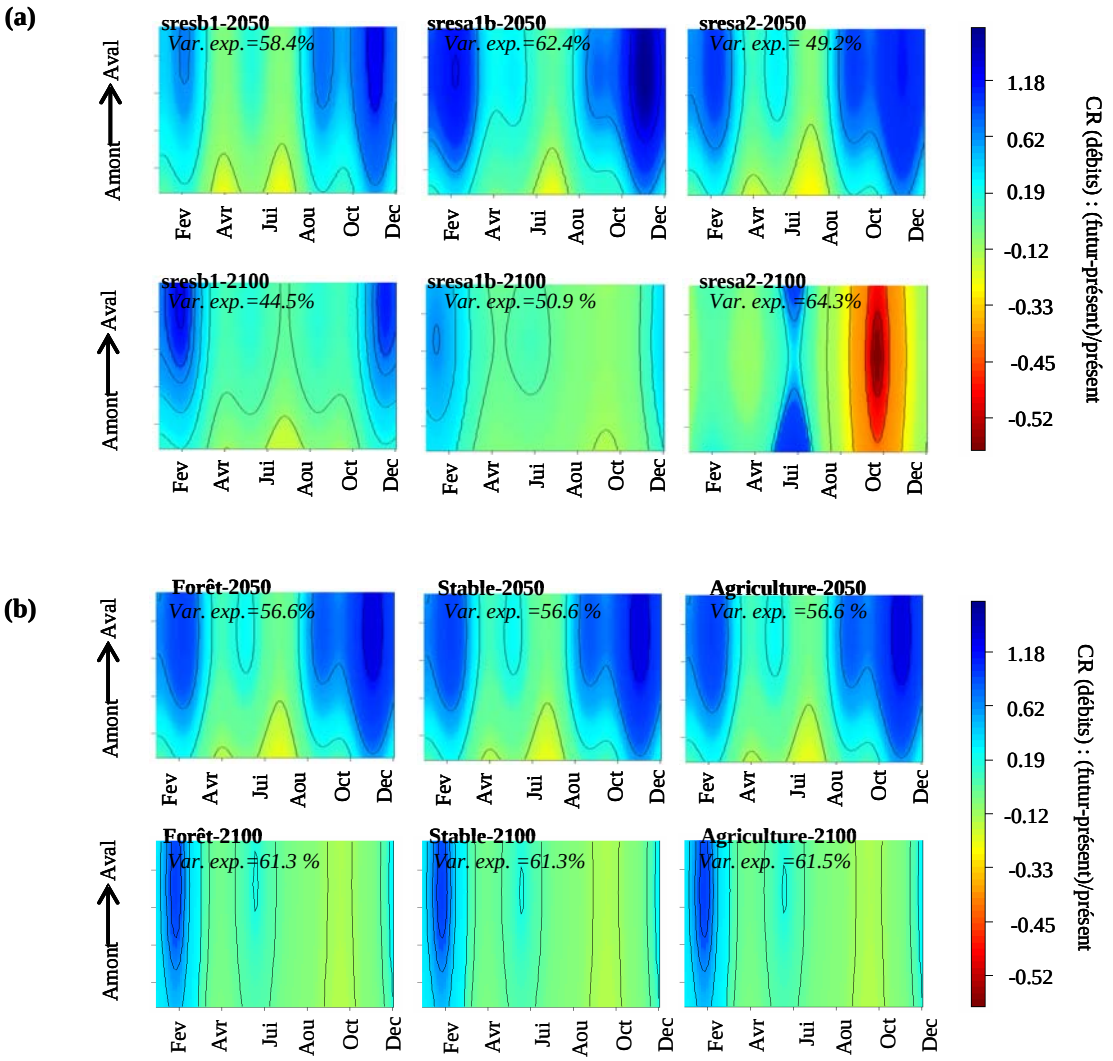


Figure 16. Patrons de variabilité spatiale (gradient amont-aval de la Garonne) et temporelle (mois de l'année) dans les changement relatifs de débit selon deux périodes de temps considérées, 2048-2052 et 2095-2100, en fonction de : (a) trois scénarios climatiques d'émission de gaz à effet de serre ; (b) trois scénarios d'occupation des sols. Voir texte pour la description des scénarios

En 2048-2052, l'analyse des changements relatifs (CR) met en évidence une augmentation globale des débits hivernaux (environ +60%) ainsi qu'une diminution des débits estivaux (globalement -10% et jusqu'à -20% dans les zones amonts). Ces patrons de changement sont relativement consensuels entre les différents scénarios climatiques (Figure 16a). En revanche, les changements hydrologiques deviennent plus contrastés entre les scénarios climatiques en 2095-2100. Le scénario 'optimiste' (B1) décrit une tendance relativement similaire à celle de 2048-2052, avec une diminution généralisée des débits printaniers et estivaux (-10%) ainsi qu'une augmentation des débits hivernaux (jusqu'à +100% dans les zones aval). A l'opposé, le scénario climatique le plus 'pessimiste' (A2) projette une augmentation des débits estivaux (+20%) suivie d'un important déficit en eau durant l'automne (-50%). Toutefois, il est à noter que c'est également en automne que l'incertitude liée aux GCM est la plus forte (Figure 15a). Le niveau de confiance accordé à ces résultats doit donc être ajusté. Par ailleurs, les scénarios d'occupation des sols ne semblent pas particulièrement affecter le changement relatif des débits (Figure 16b).

Les patrons de variation spatiaux et saisonniers dans les changements relatifs en nitrates mettent en évidence des différences à la fois entre périodes futures, scénarios climatiques et scénarios d'occupation des sols (Figure 17). L'augmentation des nitrates est la plus critique en 2095-2100 (en moyenne + 50%) lorsqu'elle se conjugue avec le scénario climatique le plus pessimiste (Figure 17a ; scénario A2 ;) et le scénario d'augmentation des surfaces agricoles (Figure 17b ; scénario 'Agriculture'). L'augmentation de 20% des surfaces agricoles pourrait notamment entraîner un doublement des concentrations actuelles en nitrates (+100%) sur les secteurs amont, alors que cette augmentation serait plus modérée (+10%) en aval de la Garonne (Figure 17b ; scénario 'Agriculture'). Le maintien de l'occupation actuelle des sols dans le futur ou l'augmentation des surfaces pastorales et boisées (Figure 17b ; scénario 'Stable' ou 'Forêt') pourrait favoriser la stabilité, voire une diminution généralisée, des concentrations actuelles en nitrates, particulièrement en hiver (jusqu'à -30%) et de manière plus marquée en 2048-2052 qu'en 2095-2100 (Figure 17b). La dynamique des nitrates étant très liée à la dynamique de l'hydrologie, les fortes périodes d'étiage automnal projetées pour la fin du 21^{ième} siècle selon le scénario climatique A2 (Figure 16a) peuvent expliquer l'augmentation parallèle des concentrations en nitrates (Figure 17a). De manière comparable, l'augmentation relativement importante des débits hivernaux projetée en 2048-2052 (Figure 16a) pourrait favoriser la dilution des nitrates, ce qui expliquerait des changements relatifs en nitrates moindres (Figure 17a).

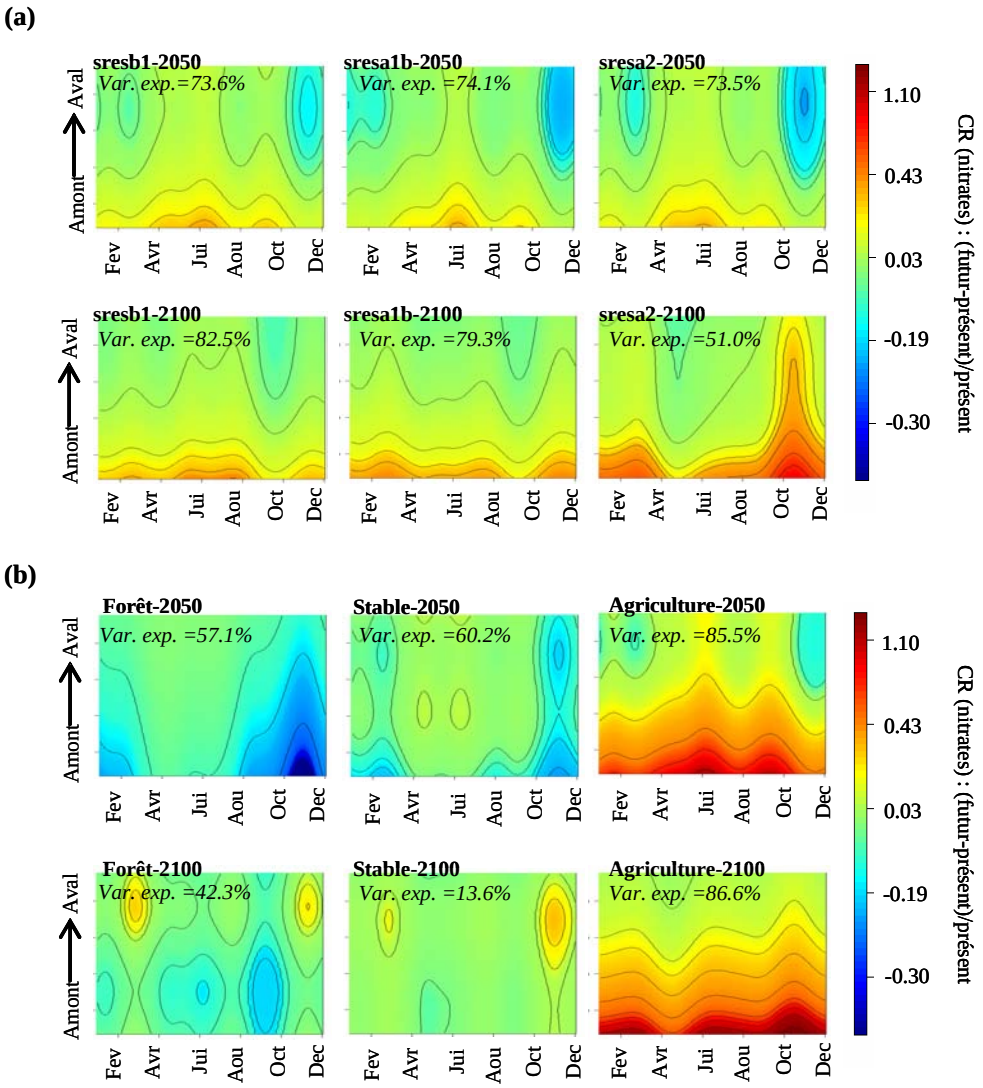


Figure 17. Patrons de variabilité spatiale (gradient amont-aval de la Garonne) et temporelle (mois de l'année) dans les changements relatifs en nitrates selon deux périodes de temps considérées, 2048-2052 et 2095-2100, en fonction de: (a) trois scénarios climatiques d'émission de gaz à effet de serre; (b) trois scénarios d'occupation des sols. Voir texte pour la description des scénarios

3IÈME PARTIE : DISCUSSION

Les deux objectifs de ma thèse étaient : (i) de proposer une méthodologie permettant de mieux intégrer le signal de changement climatique dans des modèles hydro-écologiques ; (ii) d'appliquer cette méthodologie afin d'évaluer l'impact potentiel futur du changement global sur la biodiversité des poissons de rivière et la dynamique hydro-chimique des nitrates. Ces deux objectifs seront discutés successivement afin que la compréhension des principales forces et limites méthodologiques des approches mises en œuvre aident à évaluer la crédibilité des projections hydro-biologiques et hydro-chimiques potentielles futures.

1 CONSIDÉRATIONS MÉTHODOLOGIQUES

1.1 CRÉDIBILITÉ DES PROJECTIONS FUTURES, VARIABILITÉ ET INCERTITUDES

Tout d'abord, il est important de rappeler que le modèle hydro-biologique utilisé dans le cadre de cette thèse est un modèle statistique et statique. Il ne projette donc que la distribution potentielle des espèces. En outre, les projections de ce modèle n'intègrent aucun paramètre démographique (fécondité, mortalité, croissance) et dynamique des populations (dispersion, migration, compétition inter- ou intra- spécifique). Par conséquent, les futurs habitats potentiels d'une espèce ne peuvent en aucune manière être assimilés à sa future distribution réalisée. A l'inverse, les projections hydro-chimiques intègrent une interprétation vraisemblablement plus réaliste de l'impact des changements globaux car, du fait de sa nature dynamique, le modèle HBV/INCA-N intègre spatialement et temporellement un ensemble d'interactions hydro-chimiques interagissant entre l'atmosphère, l'occupation des sols et le sous-sol.

Identifier et quantifier les sources de variabilité dans les projections futures est une étape indispensable. L'utilisation d'une méthode de partitionnement hiérarchique a l'intérêt de considérer les effets joints et indépendants de plusieurs facteurs pouvant expliquer la variabilité des projections. Un des résultats majeurs de cette étude montre que l'importante variabilité spatiale et temporelle des changements écologiques et de leurs incertitudes associées. En moyenne pour l'ensemble des projections, la variabilité entre GCM explique près de 30% de l'incertitude dans les projections hydro-chimiques (13 GCM considérés) et 1% à peine dans les projections hydro-biologiques (5 GCM considérés). Ces valeurs d'incertitude relativement modérée compte tenu du nombre de GCM considéré, soulignent certainement la bonne aptitude des modèles de downscaling à fournir de robustes projections

hydro-climatiques en entrée des modèles hydro-biologiques et hydro-chimiques. Par ailleurs, l'incertitude totale dans les projections hydro-biologiques et hydro-chimiques aurait tendance à augmenter progressivement dans le futur, à mesure que les différences entre scénarios apparaissent dans la seconde moitié du 21^{ième} siècle et que les divergences entre GCM s'accentuent.

Une fraction de la variabilité dans les projections qui n'a pas été prise en compte dans cette thèse, pourrait être associée à l'incertitude inhérente aux modèles écologiques eux-mêmes. Même si cette part de variabilité a été partiellement prise en compte dans le cas de la modélisation hydro-biologique par la construction de plusieurs forêts adaptatives, il serait préférable de comparer différentes méthodes statistiques pour deux raisons au moins : (i) bien que les forêts adaptatives présentent des performances explicatives et prédictives correctes (Elith *et al.* 2008 ; De'Ath 2008), rien ne justifie que cette méthode soit supérieure à une autre ; (ii) plusieurs études ont montré que les projections futures de distribution d'espèces pouvaient être très variables selon les modèles statistiques utilisés (Thuiller *et al.* 2004 ; Lawler *et al.* 2006 ; Buisson *et al.* 2009). Il convient aussi de rappeler qu'une des faiblesses de nombreux modèles mécanistiques et/ou dynamiques, notamment chez le modèle HBV/INCA-N, renvoie au paradigme 'd'équifinalité' (Beven & Freer 2001). En raison du nombre souvent important de paramètres à ajuster dans ces modèles (plus de 150 paramètres dans le cas d'INCA-N), le concept 'd'équifinalité' démontre que différents paramétrages du modèle peuvent conduire au même résultat. Cette variabilité dans le paramétrage constitue une part d'incertitude inhérente au modèle lui-même qui, en supposant qu'elle n'affecte que modérément la qualité des projections sur le climat actuel, peut en revanche avoir des conséquences importantes sur les projections futures.

1.2 Downscaling hydro-climatique

L'idée d'intégrer des projections climatiques issues de modèles de downscaling dans des modèles hydrologiques et écologiques n'est pas nouvelle en soit. Dans leur revue, Fowler *et al.* (2007) font une synthèse des récentes avancées dans ce domaine. Au cours de ma thèse, deux types d'approches différentes ont été utilisées pour la projection hydrologique. La modélisation hydro-chimique s'est inspirée de l'approche de downscaling la plus couramment utilisée et promue par la littérature (Xu *et al.* 1999 ; Fowler *et al.* 2007). Dans cette approche, le downscaling des conditions climatiques est d'abord appliqué avant que les projections ne soient intégrées en entrée d'un modèle hydrologique, le modèle HBV dans ma thèse. Cette

procédure en deux temps à l'avantage de reconstruire le cycle hydrologique au travers de processus fondamentaux (précipitation, évaporation, interception, infiltration, ruissellement, etc.). En revanche, cette procédure est généralement contrainte dans l'espace et dans le temps par la complexité, le domaine de calibration et la résolution temporelle du modèle hydrologique. Aussi, ces contraintes rendent les sorties du modèle hydrologique difficilement utilisables pour d'autres modèles à large échelle spatiale, comme le modèle de distribution des poissons dans cette thèse. Certains modèles hydrologiques globaux, comme le modèle WaterGap (Alcamo *et al.* 2003; Doll *et al.* 2003), peuvent offrir une alternative intéressante aux projections hydrologiques à large échelle spatiale. Par exemple, Xenopoulos *et al.* (2005) ont utilisé WaterGap à l'échelle du globe en relation avec des modèles prédictifs pour projeter la richesse future potentielle de poissons dans plusieurs grands bassins hydrographiques du monde.

Dans la modélisation hydro-biologique, une approche de downscaling des conditions hydrologiques a été développée directement à partir des processus atmosphériques à large échelle spatiale. Etant donné la difficulté avérée des GCM à modéliser correctement les composantes essentielles du cycle de l'eau, cette approche de downscaling a été moins considérée que la précédente dans le passé (Xu *et al.* 1999, Fowler *et al.* 2007). Pourtant, la compréhension des connexions entre les processus climatiques agissant à large échelle spatiale et la variabilité hydrologique locale a nettement progressé au cours des dernières années (Phillips *et al.* 2003 ; Kingston *et al.* 2006, Kingston *et al.* 2007). De plus, l'utilisation de méthodes statistiques non-linéaires pour modéliser ces connexions, comme les vecteurs de machine (Ghosh & Mujumdar 2008) ou les réseaux de neurones (Cannon & Whitfield 2002), a permis d'augmenter l'aptitude des modèles de downscaling à projeter la variabilité hydrologique régionale et locale à partir des processus atmosphériques à large échelle.

Comme toute approche de downscaling statistique, la méthode développée au cours de cette thèse peut être exposée à certaines limites. En effet, l'hypothèse de stationnarité sur laquelle repose cette approche suppose que les connexions et ajustements établis sur le climat présent restent valides dans le futur. Dans cette étude, malgré le développement d'une approche de validation croisée pour pallier la non-stationnarité des séries hydrologiques, il est difficile de réfuter ou d'affirmer cette hypothèse. A l'heure actuelle, peu d'études se sont intéressées à tester la robustesse de cette hypothèse de stationnarité dans le futur. Toutefois, Vrac *et al.* (2007) se sont par exemple appuyés sur une comparaison entre le climat présent et futur, modélisé par des modèles climatiques de circulation générale (GCM) et des modèles

régionaux (RCM), afin de valider cette hypothèse. De futures recherches restent cependant essentielles pour répondre plus finement à cette question.

2 Considérations écologiques

2.1 Perturbations inévitables des écosystèmes ?

2.1.1 Perte de biodiversité?

La notion de biodiversité peut être assez subjective selon l'échelle spatiale considérée et le type d'indice utilisé pour la caractériser (Moss *et al.* 2009). D'après les résultats de cette thèse, si l'on considère une perte de biodiversité comme la diminution du nombre d'espèces (diminution de la diversité alpha), le changement climatique pourrait avoir un impact relativement positif sur les communautés de poissons, étant donné que l'on pourrait assister à une augmentation globale du nombre d'espèces le long du gradient amont-aval. Cette augmentation de la richesse pourrait particulièrement être due à l'expansion des espèces d'eau chaude dans les parties amont, ce qui rejoint les résultats de certaines études réalisées en Europe (Daufresne & Boet 2007 ; Matulla *et al.* 2007 ; Buisson *et al.* 2008) et en Amérique (Jackson & Mandrak 2002; Mohseni *et al.* 2003; Chu *et al.* 2005; Sharma *et al.* 2007). De manière globale, l'intensité des modifications dans la structure des assemblages, en faveur du développement des espèces d'espèce d'eau chaude, pourrait être aussi importante, voire plus, que ceux projetés pour d'autres organismes (Peterson *et al.* 2002 ; Thuiller *et al.* 2005 ; Broennimann *et al.* 2006).

En revanche, cette augmentation de la richesse spécifique pourrait s'accompagner d'une diminution de la diversité beta, c'est-à-dire une homogénéisation taxonomique des communautés. Dans une autre étude, Buisson & Grenouillet (2009) ont également souligné cette tendance à l'homogénéisation sur des aspects fonctionnels (traits biologiques) des communautés. A ce jour, la question de l'homogénéisation a été largement considérée dans le cadre des invasions par des espèces exotiques introduites par les activités humaines (Rahel 2000 ; McKinney 2004 ; Olden 2006 ; Olden & Rooney 2006). Cependant, à notre connaissance, peu d'étude ont réellement mis en évidence si les assemblages projetés dans le futur sous l'effet des changements climatiques pourraient être plus similaires que ceux présents actuellement. De plus, il faut tenir compte du fait que les changements climatiques pourraient amplifier l'homogénéisation des assemblages de poissons d'eau douce causée par l'introduction d'espèces exotiques (e.g., Rahel 2000 ; Leprieur et al. 2008 ; Olden et al. 2008). Dans ce contexte, la gestion future des peuplements piscicoles n'apparaît pas si simple, ce qui

prouve la nécessité de considérer plusieurs aspects de la biodiversité à différentes échelles, au delà de la simple notion de richesse spécifique (Sax & Gaines 2003 ; Jurasinski & Kreyling 2007).

2.1.2 Vulnérabilité des zones de montagne

Si les zones de montagne semblent les plus vulnérables à une forte perturbation de la biodiversité, ces zones pourraient également subir des modifications hydro-chimiques très drastiques. Dans ces zones, la sévérité des étiages pourrait s'accentuer (jusqu'à -20% des débits actuels) sans pouvoir être compensée par une augmentation des débits hivernaux, comme cela pourrait éventuellement être le cas des parties aval de la Garonne (+50% des débits hivernaux). En effet, l'hydrologie des régimes nivaux étant fortement liée à la fonte du manteau neigeux accumulé pendant l'hiver, il est probable qu'une augmentation globale des températures et/ou une diminution des précipitations hivernales favorisent la diminution globale du manteau neigeux. Nos résultats sont relativement concordants avec ceux de Caballero *et al.* (2007) qui ont réalisé des projections hydrologiques sur plusieurs stations du bassin Adour Garonne à l'aide du modèle SAFRAN-ISBA-MODCOU. De manière similaire, les patrons de diminution globale des débits hivernaux et d'augmentation des étiages prédits dans cette thèse recoupent les projections réalisées par Boe *et al.* (2009) sur l'ensemble de la France métropolitaine.

C'est également dans les zones amont que les changements relatifs en nitrates seraient susceptibles d'être les plus importants, notamment du fait de la diminution des débits estivaux et des changements d'occupation des sols. En effet, les sols exportent l'azote vers les rivières et le calendrier saisonnier des pratiques agricoles (ex. fertilisation) peut avoir de grandes répercussions sur la disponibilité et le transfert des nutriments (ex. lessivage des sols en hiver). Une augmentation des concentrations en azote dans les écosystèmes aquatiques, comme source d'apport en nutriments, entraîne généralement une augmentation de la productivité primaire phytoplanctonique menant au processus d'eutrophisation. Il est largement admis que l'eutrophisation excessive des milieux peut perturber l'ensemble des processus fonctionnels de l'écosystème aquatique (réseaux trophiques), depuis l'acidification jusqu'à l'anoxie des milieux (Heino *et al.* 2009 ; Whitehead *et al.* 2009). Plusieurs études d'ailleurs ont pu mettre en évidence qu'une diminution de la biodiversité pouvait résulter d'une augmentation des concentrations de nutriments dans le milieu, en favorisant le développement d'espèces généralistes et compétitives au détriment d'un grand nombre d'espèces spécialistes moins compétitives et agressives (Waide *et al.* 1999 ; Mittelbach *et al.*

2001 ; Moss *et al.* 2009). Compte tenu des nombreuses interactions biotiques au sein des réseaux trophiques, les impacts du changement global sur la biodiversité pourraient être encore plus importants que ceux projetés par les modèles. Les poissons, situés au sommet de ces réseaux trophiques, pourraient subir l'accumulation de l'ensemble des perturbations affectant les niveaux trophiques inférieurs.

2.2 Atténuations possibles des impacts du changement climatique ?

Si le fonctionnement des écosystèmes aquatiques semble exposé à d'incontournables bouleversements dans le futur, les différents scénarios climatiques et d'occupation des sols testés au cours de cette thèse laissent entrevoir des atténuations possibles à ce changement. Globalement, les différences entre les trois scénarios climatiques ne se distinguent pas avant la deuxième moitié du 21^{ième} siècle. Au-delà, l'intensité des changements se poursuit de manière plus marquée dans le cas du scénario climatique le plus pessimiste (A2), que dans le cas des deux autres scénarios les plus optimistes, A1B et surtout B1. Les différences dans ces changements pourraient être tout particulièrement perceptibles au niveau de l'hydrologie où le scénario A2 pourrait augmenter drastiquement la sévérité des étiages en automne. Ces différences entre scénarios climatiques, seulement perceptibles à long terme, soulignent que de concrètes actions environnementales doivent considérer le temps de réponse relativement long du climat et de ses processus rétroactifs (Cox *et al.*, 2000 ; Beaumont *et al.*, 2008), avant de pouvoir mesurer des changements significatifs.

Les projections hydro-chimiques ont mis en évidence l'influence prépondérante de l'occupation des sols sur les changements relatifs en nitrates, particulièrement dans les zones montagneuses et faiblement agricoles à l'heure actuelle. Une augmentation progressive des zones agricoles jusqu'en 2100 pourraient entraîner un doublement des concentrations actuelles en azote. A l'inverse, un scénario agro-pastoral favorisant l'expansion des prairies et des forêts pourrait limiter considérablement les apports azotés entrant dans le système, augmenter les processus de dénitrification des sols, et conduire à la stabilité voire à la diminution relative des concentrations actuelles en nitrates. Bien que les scénarios d'occupation des sols testés dans notre étude soient fictifs et ne prennent absolument pas en compte le réel développement social et économique des régions, les différences entre scénarios sont telles que ces résultats méritent entière considération. Toute modification des pratiques agricoles pouvant mener à une diminution des apports en fertilisants azotés serait susceptible de réduire significativement le lessivage des sols et, par conséquent, de réduire les

concentrations en azote dans les rivières. Selon Ormerod (2009), favoriser le développement des zones ripariennes, c'est-à-dire les zones recouvertes de végétation longeant le cours d'eau, pourraient également aider à réguler les flux de matières (ex. sédiments) et la température de l'eau et contribuer ainsi à la rétention des nutriments, notamment à celle des nitrates.

Un dernier élément pouvant atténuer les effets du changement climatique, qui a été négligé dans cette thèse, concerne la capacité de résilience des écosystèmes aquatiques (Poff 2002). Au travers de la diversité des organismes biologiques qui les constituent, les écosystèmes sont 'vivants' et possèdent donc une certaine capacité à évoluer naturellement dans un environnement changeant. En supposant que les changements globaux soient suffisamment progressifs dans le temps, il n'est pas impossible que certaines espèces puissent développer des réponses physiologiques ou comportementales leur permettant de s'adapter aux modifications environnementales. De telles modifications comportementales d'ordre phénologique (ex. période de reproduction ou de floraison plus précoce) ont déjà été mises en évidence chez les amphibiens (Beebee 1995), les oiseaux (Dunn & Winkler 1999) et les plantes (Bradley *et al.* 1999), mais jamais chez les poissons à notre connaissance. Une des difficultés majeures restent donc d'appréhender cette capacité d'adaptation des espèces sur une période de temps relativement courte (les 100 prochaines années) alors que les processus évolutifs se produisent généralement sur des échelles de temps beaucoup plus longues (Wrona *et al.* 2006).

CONCLUSIONS ET PERSPECTIVES

1.1 Synthèse des résultats

L'approche de modélisation développée aux cours de cette thèse présente l'intérêt majeur de favoriser les relations multidisciplinaires entre les sciences du climat, de l'hydrologie et de l'écologie en contribuant notamment à : (i) une meilleure compréhension des connexions entre le climat à large échelle et la variabilité hydrologique régionale ou locale ; (ii) fournir des projections hydro-climatiques robustes en entrée des modèles d'impacts écologiques, en améliorant notamment la qualité du signal saisonnier dans les projections hydrologiques ; (iii) exploiter la dimension interannuelle du signal de changement climatique afin de quantifier la variabilité des changements écologiques de manière spatialement et temporellement explicite.

Les résultats principaux issus des projections hydro-écologiques futures suggèrent des conséquences importantes des changements globaux sur la biodiversité des poissons d'eau douce ainsi que sur la dynamique hydro-chimique des nitrates dans le bassin Adour-Garonne. L'intensité des perturbations pourrait être hétérogène dans l'espace, en particulier le long du gradient amont-aval des rivières, et dans le temps, avec une modification importante de la dynamique saisonnière de l'hydrologie et des nitrates.

Les zones en amont des bassins versants pourraient être les plus vulnérables, en étant notamment exposées à de sévères périodes d'étiages qui favoriseraient l'augmentation des températures et des concentrations en nitrates ainsi qu'une perte de biodiversité (diminution de la richesse spécifique). Il apparait donc urgent de mettre en place des plans de conservation pour ces zones particulières afin qu'elles ne soient pas dégradées par d'autres facteurs anthropiques et restent favorables à la survie d'espèces vulnérables comme la truite.

Selon un scénario optimiste de réduction des gaz à effet de serre dans le futur ainsi qu'une modification des pratiques agricoles vers une expansion des zones pastorales et boisées ou une diminution de l'épandage des fertilisants azotés, l'intensité des impacts du changement global sur les écosystèmes aquatiques pourrait être atténuée. En revanche, ces atténuations ne pourraient être réellement perceptibles qu'à partir de la deuxième moitié du 21^{ième} siècle.

1.2 Vers une modélisation statistico-dynamique plus réaliste

Les différents résultats de ma thèse laissent entrevoir des perspectives de recherche très prometteuses. Par exemple, dans le cadre des projections hydro-biologiques, les modèles statiques de distribution d'espèces pourraient être orientés vers une modélisation dynamique décrivant les mécanismes écologiques de manière spatialement et temporellement explicite (Dormann 2007 ; Barnard & Thuiller 2008 ; Williams *et al.* 2008 ; Zurell *et al.* 2009). Dans le cas des poissons, ce modèle dynamique pourrait inclure la variabilité interannuelle et saisonnière des projections hydro-climatiques afin de simuler la dynamique spatiale (capacité de migration) et démographique des espèces (mortalité, fécondité, interactions biologiques). Les projections pourraient être ensuite comparées entre modèles statiques et dynamiques (e.g. Morin & Thuiller 2009) afin d'identifier le rôle respectif des processus climatiques et écologiques dans la structuration et la distribution des populations.

Alors que l'approche de modélisation statique développée au cours de ma thèse suppose que la capacité de dispersion des espèces est illimitée (e.g. une espèce est présente si les conditions hydro-climatiques lui sont favorables), leur distribution réelle peut dépendre à la

fois de leur physionomie (e.g. taille) ou de leur comportement (e.g. espèces migratrices) ainsi que de la connectivité des réseaux hydrographiques. La capacité de dispersion des espèces commence tout juste à être intégrée dans certains modèles bioclimatiques (e.g. Morin *et al.* (2007) et Thuiller *et al.* (2008) chez les végétaux, Zurell *et al.* (2009) chez les papillons). Chez les espèces de poisson d'eau douce, le taux de dispersion pourrait être estimé à partir des vitesses de colonisation des habitats suite à des (ré)-introductions ou des patrons de recolonisation depuis la dernière glaciation (Durand *et al.* 2000, 2003 ; Griffiths 2006). Quant à la connectivité des réseaux hydrographiques, comme indicateur de la capacité physique des espèces à coloniser dans le futur un site actuellement inoccupé, elle pourrait être modélisée à travers une mesure de distance entre les sites, géographique (e.g. Euclidienne) et/ou hydrologique (e.g. tenant compte des connections et de la sinuosité des rivières) (e.g. Peterson *et al.* 2007). Il pourrait être également judicieux de prendre en compte le nombre d'obstacles qui fragmentent la continuité amont-aval des rivières (e.g. Barrages) et entravent ainsi la libre circulation des espèces (Lassalle *et al.* 2009). Ainsi, combiner le taux de dispersion des espèces, la connectivité des réseaux hydrographiques ainsi que le nombre d'obstacles pourrait permettre d'estimer de manière robuste la capacité réelle de dispersion des espèces.

Au cours de cette thèse, l'approche de modélisation statique des espèces a fait également abstraction des interactions biotiques, comme la compétition inter- et intra-spécifiques pour les ressources, qui sont pourtant susceptibles de favoriser le développement futur des espèces les plus compétitrices. Bien que Araujo & Luoto (2007) aient montré que l'influence des interactions biotiques pouvait être significative sur le résultat des projections d'espèces futures, peu d'études les ont jusqu'à maintenant intégrées dans des modèles bioclimatiques. De même, à notre connaissance très peu de modèles bioclimatiques ont explicitement considéré les interactions biotiques au sein des réseaux trophiques, à l'instar de Zurell *et al.* (2009) qui ont modélisé la dynamique d'une espèce de papillon en interaction avec celle de son parasite. Par ailleurs, l'influence des interactions biotiques sur les projections futures pourrait être accentuée par l'introduction et l'expansion d'espèces exotiques et invasives, pouvant mener à l'extirpation de certaines espèces de poisson natives et modifier profondément le fonctionnement des réseaux trophiques (e.g. Mercado-Silva *et al.* 2006; Olden *et al*. 2006 ; Rahel & Olden 2008 ; Leprieur *et al.* 2009). C'est ainsi que Jackson & Mandrak (2002) ont mis en évidence que l'augmentation de l'aire de distribution du black-bass à petite bouche en réponse aux modifications du climat futur pourrait provoquer l'extirpation de plus de 25000 populations de quatre espèces de cyprinidés au Canada.

Les interactions entre la qualité de l'eau, c'est-à-dire l'habitat chimique, et le fonctionnement des réseaux trophiques ont été discutées dans la section Discussion. A notre connaissance, aucune étude n'a explicitement intégré l'habitat chimique dans des projections futures de distribution de poisson ou autres organismes biologiques. Une perspective incontournable serait donc d'intégrer l'effet du climat futur sur l'habitat chimique (e.g. eutrophisation), thermique et hydrologique afin d'évaluer leurs répercussions sur les organismes biologiques. Une approche bayésienne pourrait être tout particulièrement adaptée à ce genre de problème, en modélisant de manière conditionnelle et hiérarchique les interactions possibles entre espèces ou entre l'hydrologie et la chimie de l'eau.

Enfin, plusieurs éléments pourraient contribuer à une meilleure gestion futures des ressources hydriques et de la biodiversité aquatique. Tout d'abord, une meilleure quantification des sources d'incertitude dans les projections hydro-écologiques pourrait se faire en développant d'avantage des approches d'ensemble (Araujo & New 2007), en utilisant notamment plusieurs méthodes statistiques pour la construction des modèles de downscaling et de distribution d'espèces. Cette nécessité est renforcée par plusieurs études ayant souligné l'influence du choix de la méthode statistique sur la distribution future des espèces (Lawler *et al.* 2006 ; Buisson *et al.* 2009). D'autre part, il serait utile d'étendre les notions de biodiversité à d'autres groupes taxonomiques (ex. macrophytes, invertébrés, diatomées), ce qui favoriserait une meilleure compréhension des perturbations potentielles des changements globaux sur les écosystèmes aquatiques. La prise en compte d'autres critères de biodiversité, comme la diversité fonctionnelle à partir des traits biologiques des espèces (e.g. taille des organismes, date de ponte, régime alimentaire, etc.), pourrait également permettre de mieux appréhender les bouleversements fonctionnels des écosystèmes en réponse aux changements globaux.

RÉFÉRENCES

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T. & Siebert, S. (2003). Development and testing of the WaterGAP 2 global model of water use and availability/Développement et évaluation du modèle global WaterGAP 2 d'utilisation et de disponibilité de l'eau. *Hydrological Sciences Journal / Journal des Sciences Hydrologiques*, 48, 317-337.

Araujo, M.B. & Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, 16, 743-753.

Araujo, M.B. & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22, 42-47.

Barnard, P. & Thuiller, W. (2008). Introduction. Global change and biodiversity: future challenges. *Biology Letters*, 4, 553-555.

Beaumont, L.J., Hughes, L. & Pitman, A.J. (2008). Why is the choice of future climate scenarios for species distribution modelling important? *Ecology Letters*, 11, 1135-1146.

Beebee, T.J.C. (1995). Amphibian breeding and climate. *Nature*, 374, 219-220.

Beven, K. & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249, 11-29.

Boe, J., Terray, L., Martin, E. & Habets, F. (2009). Projected changes in components of the hydrological cycle in French river basins during the 21st century. *Water Resources Research*, 45, W08426.

Bradley, N.L., Leopold, A.C., Ross, J. & Huffaker, W. (1999). Phenological changes reflect climate change in Wisconsin. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 9701-9704.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32.

Broennimann, O., Thuiller, W., Hughes, G., Midgley, G.F., Alkemade, J.M.R. & Guisan, A. (2006). Do geographic distribution, niche property and life form explain plants' vulnerability to global change? *Global Change Biology*, 12, 1079-1093.

Buisson, L. & Grenouillet, G. (2009). Contrasted impacts of climate change on stream fish assemblages along an environmental gradient. *Diversity and Distributions*, 15, 613-626.

Buisson, L., Thuiller, W., Casajus, N., Lek, S. & Grenouillet, G. (2009). Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, 9999.

Buisson, L., Thuiller, W., Lek, S., Lim, P. & Grenouillet, G. (2008). Climate change hastens the turnover of stream fish assemblages. *Global Change Biology*, 14, 2232-2248.

Caballero, Y., Voirin-Morel, S., Habets, F., Noilhan, J., LeMoigne, P., Lehenaff, A. & Boone, A. (2007). Hydrological sensitivity of the Adour-Garonne river basin to climate change. *Water Resources Research*, 43, W07448.

Cannon, A.J. & Whitfield, P.H. (2002). Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models. *Journal of Hydrology*, 259, 136-151.

Cattanéo, F. (2005). Does hydrology constrain the structure of fish assemblages in French streams? Local scale analysis. *Archiv für Hydrobiologie*, 164, 345-365.

Chevan, A. & Sutherland, M. (1991). Hierarchical Partitioning. *American Statistician*, 45, 90-96.

Chu, C., Mandrak, N.E. & Minns, C.K. (2005). Potential impacts of climate change on the distributions of several common and rare freshwater fishes in Canada. *Diversity and Distributions*, 11, 299-310.

Cox, P.M., Betts, R.A., Jones, C.D., Spall, S.A. & Totterdell, I.J. (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, 408, 184-187.

Croisé, L., Ulriche, E., Duplat, P. & Jaquet, O. (2002). Renecofor – Deux approches indépendantes pour l'estimation et la cartographie des dépôts atmosphériques totaux hors couvert forestier sur le territoire français. In. Office Nationale des Forêts Paris.

Daufresne, M. & Boet, P. (2007). Climate change impacts on structure and diversity of fish communities in rivers. *Global Change Biology*, 13, 2467-2478.

Daufresne, M., Lengfellner, K. & Sommer, U. (2009). Global warming benefits the small in aquatic ecosystems. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 12788-12793.

De'ath, G. (2007). Boosted trees for ecological modeling and prediction. *Ecology*, 88, 243-251.

De'ath, G. & Fabricius, K.E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81, 3178-3192.

Deque, M. (2007). Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global and Planetary Change*, 57, 16-26.

Doll, P., Kaspar, F. & Lehner, B. (2003). A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology*, 270, 105-134.

Dormann, C.F. (2007). Promising the future? Global change projections of species distributions. *Basic and Applied Ecology*, 8, 387-397.

Dunn, P.O. & Winkler, D.W. (1999). Climate change has affected the breeding date of tree swallows throughout North America. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 266, 2487-2490.

Durand, J.D., Bianco, P.G., Laroche, J. & Gilles, A. (2003). Insight into the origin of endemic Mediterranean ichthyofauna: Phylogeography of Chondrostoma genus (Teleostei, Cyprinidae). *Journal of Heredity*, 94, 315-328.

Durand, J.D., Unlu, E., Doadrio, I., Pipoyan, S. & Templeton, A.R. (2000). Origin, radiation, dispersion and allopatric hybridization in the chub Leuciscus cephalus. *Proceedings Of The Royal Society Of London Series B-Biological Sciences*, 267, 1687-1697.

Elith, J., Leathwick, J.R. & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77, 802-813.

Fowler, H.J., Blenkinsop, S. & Tebaldi, C. (2007). Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 27, 1547-1578.

Ghosh, S. & Mujumdar, P.P. (2008). Statistical downscaling of GCM simulations to streamflow using relevance vector machine. *Advances in Water Resources*, 31, 132.

Griffiths, D. (2006). Pattern and process in the ecological biogeography of European freshwater fish. *Journal of Animal Ecology*, 75, 734-751.

Guisan, A. & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8, 993-1009.

Guisan, A. & Zimmermann, N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147-186.

Hastie, T. & Tibshirani, R. (1990). *Generalized Additive Models*. Chapman & Hall/CRC.

Heino, J., Virkkala, R. & Toivonen, H. (2009). Climate change and freshwater biodiversity: detected patterns, future trends and adaptations in northern regions. *Biological Reviews*, 84, 39-54.

Hutchinson, G.E. (1957). Population studies - Animal ecology and demography - Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 415-427.

Jaccard, P. (1901). Etude comparative de la distribution florale dans une portion des Alpes et des Jura.[Comparative study of the distribution of flora in a region of the Alps and the Jura] Bull. *Soc. Vaudoise Sci. Nat*, 37, 547–549.

Jackson, D.A. & Mandrak, N.E. (2002). Changing fish biodiversity: predicting the loss of cyprind biodiversity due to global climate change. In. American Fisheries Society, pp. 89-98.

Jurasinski, G. & Kreyling, J. (2007). Upward shift of alpine plants increases floristic similarity of mountain summits. *J. Veg. Sci.*, 18, 711-718.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. & Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77, 437-471.

Kingston, D.G., Lawler, D.M. & McGregor, G.R. (2006). Linkages between atmospheric circulation, climate and streamflow in the northern North Atlantic: research prospects. *Progress in Physical Geography*, 30, 143-174.

Kingston, D.G., McGregor, G.R., Hannah, D.M. & Lawler, D.M. (2007). Large-scale climatic controls on new England river flow. *Journal of Hydrometeorology*, 8, 367-379.

Lamouroux, N. & Cattanéo, F. (2006). Fish assemblages and stream hydraulics: consistent relations across spatial scales and regions. *River Research and Applications*, 22, 727-737.

Lassalle, G., Crouzet, P. & Rochard, E. (2009). Modelling the current distribution of European diadromous fishes: an approach integrating regional anthropogenic pressures. *Freshwater Biology*, 54, 587-606.

Lawler, J.J., White, D., Neilson, R.P. & Blaustein, A.R. (2006). Predicting climate-induced range shifts: model differences and model reliability. *Global Change Biology*, 12, 1568-1584.

Lek, S. & Guegan, J.F. (1999). Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling*, 120, 65-73.

Leprieur, F., Beauchard, O., Hugueny, B., Grenouillet, G. & Brosse, S. (2008). Null model of biotic homogenization: a test with the European freshwater fish fauna. *Diversity and Distributions*, 14, 291-300.

Leprieur, F., Brosse, S., Garcia-Berthou, E., Oberdorff, T., Olden, J.D. & Townsend, C.R. (2009). Scientific uncertainty and the assessment of risks posed by non-native freshwater fishes. *Fish And Fisheries*, 10, 88-97.

Lindstrom, G., Johansson, B., Persson, M., Gardelin, M. & Bergstrom, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, 201, 272-288.

Matulla, C., Schmutz, S., Melcher, A., Gerersdorfer, T. & Haas, P. (2007). Assessing the impact of a downscaled climate change simulation on the fish fauna in an Inner-Alpine River. *International Journal of Biometeorology*, 52, 127-137.

McCullagh, P. (1984). Generalized Linear-Models. *European Journal of Operational Research*, 16, 285-292.

McKinney, M.L. (2004). Measuring floristic homogenization by non-native plants in North America. *Global Ecology and Biogeography*, 13, 47-53.

Mercado-Silva, N., Olden, J.D., Maxted, J.T., Hrabik, T.R. & Zanden, M.J.V. (2006). Forecasting the spread of invasive rainbow smelt in the Laurentian Great Lakes region of North America. *Conservation Biology*, 20, 1740-1749.

Michelangeli, P.A., Vrac, M. & Loukos, H. (2009). Probabilistic downscaling approaches: Application to wind cumulative distribution functions. *Geophysical Research Letters*, 36, L11708.

Mittelbach, G.G., Steiner, C.F., Scheiner, S.M., Gross, K.L., Reynolds, H.L., Waide, R.B., Willig, M.R., Dodson, S.I. & Gough, L. (2001). What is the observed relationship between species richness and productivity? *Ecology*, 82, 2381-2396.

Mohseni, O., Stefan, H.G. & Eaton, J.G. (2003). Global warming and potential changes in fish habitat in US streams. *Climatic change*, 59, 389-409.

Morin, X., Augspurger, C. & Chuine, I. (2007). Process-based modeling of species' distributions: What limits temperate tree species' range boundaries? *Ecology*, 88, 2280-2291.

Morin, X. & Thuiller, W. (2009). Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology*, 90, 1301-1313.

Moss, B., Hering, D., Green, A.J., Aidoud, A., Becares, E., Beklioglu, M., Bennion, H., Boix, D., Brucet, S., Carvalho, L., Clement, B., Et Al, Moss, B., Hering, D., Green, A.J., Aidoud, A., Becares, E., Beklioglu, M., Bennion, H., Boix, D., Brucet, S., Carvalho, L. & Clement, B. (2009). Climate change and the future of freshwater biodiversity in Europe: a primer for policy-makers. *Freshwater Reviews*, 2, 103-130.

Noss, R.F. (1990). Indicators for monitoring biodiversity - a hierarchical approach. *Conservation Biology*, 4, 355-364.

Olden, J.D. (2006). Biotic homogenization: a new research agenda for conservation biogeography. *Journal of Biogeography*, 33, 2027-2039.

Olden, J.D., Kennard, M.J. & Pusey, B.J. (2008). Species invasions and the changing biogeography of Australian freshwater fishes. *Global Ecology and Biogeography*, 17, 25-37.

Olden, J.D., Poff, N.L.R. & Bestgen, K.R. (2006). Life-history strategies predict fish invasions and extirpations in the Colorado River Basin. *Ecological Monographs*, 76, 25-40.

Olden, J.D. & Rooney, T.P. (2006). On defining and quantifying biotic homogenization. *Global Ecology and Biogeography*, 15, 113-120.

Ormerod, S.J. (2009). Climate change, river conservation and the adaptation challenge. *Aquatic Conservation-Marine and Freshwater Ecosystems*, 19, 609-613.

Pachauri, R.K. & Reisinger, A. (2007). *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland.

Palmer, M., Lettenmaier, D., Poff, N., Postel, S., Richter, B. & Warner, R. (2009). Climate change and river ecosystems: protection and adaptation options. *Environmental Management*, 44, 1053-1068.

Parmesan, C. & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421, 37-42.

Pearson, R.G. & Dawson, T.P. (2003). Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, 12, 361-371.

Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sanchez-Cordero, V., Soberon, J., Buddemeier, R.H. & Stockwell, D.R.B. (2002). Future projections for Mexican faunas under global climate change scenarios. *Nature*, 416, 626-629.

Peterson, E.E., Theobald, D.M. & Hoef, J.M.V. (2007). Geostatistical modelling on stream networks: developing valid covariance matrices based on hydrologic distance and stream flow. *Freshwater Biology*, 52, 267-279.

Phillips, I.D., McGregor, G.R., Wilson, C.J., Bower, D. & Hannah, D.M. (2003). Regional climate and atmospheric circulation controls on the discharge of two British rivers, 1974–97. *Theoretical and Applied Climatology*, 76, 141-164.

Poff, N.L. (2002). Ecological response to and management of increased flooding caused by climate change. *Philosophical Transactions of the Royal Society of London Series A-Mathematical physical and engineering sciences*, 360, 1497-1510.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E. & Stromberg, J.C. (1997). The natural flow regime. *Bioscience*, 47, 769-784.

R Development Core team (2009). R: A language and environment for statistical computing. In: (ed. Computing, RFfS) Vienna, Austria.

Rahel, F.J. (2000). Homogenization of fish faunas across the United States. *Science*, 288, 854-856.

Rahel, F.J. & Olden, J.D. (2008). Assessing the effects of climate change on aquatic invasive species. *Conservation Biology*, 22, 521-533.

Reed, R.D. & Marks, R.J. (1998). *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks*. MIT Press Cambridge, MA, USA.

Root, T.L., Price, J.T., Hall, K.R., Schneider, S.H., Rosenzweig, C. & Pounds, J.A. (2003). Fingerprints of global warming on wild animals and plants. *Nature*, 421, 57-60.

Rumelhart, D.E., Hintont, G.E. & Williams, R.J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533-536.

Sax, D.F. & Gaines, S.D. (2003). Species diversity: from global decreases to local increases. *Trends in Ecology & Evolution*, 18, 561-566.

Sharma, S., Jackson, D.A., Minns, C.K. & Shuter, B.J. (2007). Will northern fish populations be in hot water because of climate change? *Global Change Biology*, 13, 2052-2064.

Statzner, B., Gore, J.A. & Resh, V.H. (1988). Hydraulic stream ecology - observed patterns and potential applications. *Journal of the North American Benthological Society*, 7, 307-360.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L. & Williams, S.E. (2004). Extinction risk from climate change. *Nature*, 427, 145-148.

Thuiller, W., Albert, C., Araujo, M.B., Berry, P.M., Cabeza, M., Guisan, A., Hickler, T., Midgely, G.F., Paterson, J., Schurr, F.M., Sykes, M.T. & Zimmermann, N.E. (2008). Predicting global change impacts on plant species' distributions: future challenges. *Perspectives in Plant Ecology Evolution and Systematics*, 9, 137-152.

Thuiller, W., Araujo, M.B., Pearson, R.G., Whittaker, R.J., Brotons, L. & Lavorel, S. (2004). Biodiversity conservation - Uncertainty in predictions of extinction risk. *Nature*, 430.

Thuiller, W., Lavorel, S., Araujo, M.B., Sykes, M.T. & Prentice, I.C. (2005). Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 8245-8250.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. & Cushing, C.E. (1980). River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Sciences*, 37, 130-137.

Vrac, M., Stein, M.L., Hayhoe, K. & Liang, X.Z. (2007). A general method for validating statistical downscaling methods under future climate change. *Geophysical Research Letters*, 34.

Wade, A.J., Durand, P., Beaujouan, V., Wessel, W.W., Raat, K.J., Whitehead, P.G., Butterfield, D., Rankinen, K. & Lepisto, A. (2002). A nitrogen model for European catchments: INCA, new model structure and equations. *Hydrology and Earth System Sciences*, 6, 559-582.

Waide, R.B., Willig, M.R., Steiner, C.F., Mittelbach, G., Gough, L., Dodson, S.I., Juday, G.P. & Parmenter, R. (1999). The relationship between productivity and species richness. *Annual Review of Ecology and Systematics*, 30, 257-300.

Walther, G.R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.M., Hoegh-Guldberg, O. & Bairlein, F. (2002). Ecological responses to recent climate change. *Nature*, 416, 389-395.

Whitehead, P.G., Wilby, R.L., Battarbee, R.W., Kernan, M. & Wade, A.J. (2009). A review of the potential impacts of climate change on surface water quality. *Hydrological Sciences Journal / Journal des Sciences Hydrologiques*, 54, 101-123.

Whitehead, P.G., Wilson, E.J., Butterfield, D. & Seed, K. (1998). A semi-distributed integrated flow and nitrogen model for multiple source assessment in catchments (INCA): Part II - application to large river basins in south Wales and eastern England. *Science of the Total Environment*, 210, 559-583.

Wilby, R.L., Dawson, C.W. & Barrow, E.M. (2002). sdsm-a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling and Software*, 17, 145-157.

Williams, S.E., Shoo, L.P., Isaac, J.L., Hoffmann, A.A. & Langham, G. (2008). Towards an integrated framework for assessing the vulnerability of species to climate change. *Plos Biology*, 6, 2621-2626.

Wood, S.N. (2008). Fast stable direct fitting and smoothness selection for generalized additive models. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 70, 495-518.

Wrona, F.J., Prowse, T.D., Reist, J.D., Hobbie, J.E., Levesque, L.M.J. & Vincent, W.F. (2006). Climate change effects on aquatic biota, ecosystem structure and function. *Ambio*, 35, 359-369.

Xenopoulos, M.A., Lodge, D.M., Alcamo, J., Marker, M., Schulze, K. & Van Vuuren, D.P. (2005). Scenarios of freshwater fish extinctions from climate change and water withdrawal. *Global Change Biology*, 11, 1557-1564.

Xu, C.Y. (1999). From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. *Progress in Physical Geography*, 23, 229-249.

Zurell, D., Jeltsch, F., Dormann, C.F. & Schroder, B. (2009). Static species distribution models in dynamically changing systems: how good can predictions really be? *Ecography*, 32, 733-744.

ARTICLE N°1

Modeling the Stream Water Nitrate Dynamics in a 60,000-km^2 European Catchment, the Garonne, Southwest France.

Tisseuil, C., Wade, A.J., Tudesque, L. and Lek, S. (2008).

J Environ Qual. 37: 2155-2169

Modeling the Stream Water Nitrate Dynamics in a 60,000-km^2 European Catchment, the Garonne, Southwest France

Clément Tisseuil* CNRS- Université Paul Sabatier
Andrew J. Wade University of Reading
Loïc Tudesque and Sovan Lek CNRS- Université Paul Sabatier

The spatial and temporal dynamics in the stream water NO_3–N concentrations in a major European river-system, the Garonne (62,700 km^2), are described and related to variations in climate, land management, and effluent point-sources using multivariate statistics. Building on this, the Hydrologiska Byråns Vattenbalansavdelning (HBV) rainfall-runoff model and the Integrated Catchment Model of Nitrogen (INCA-N) are applied to simulate the observed flow and N dynamics. This is done to help us to understand which factors and processes control the flow and N dynamics in different climate zones and to assess the relative inputs from diffuse and point sources across the catchment. This is the first application of the linked HBV and INCA-N models to a major European river system commensurate with the largest basins to be managed under the Water Framework Directive. The simulations suggest that in the lowlands, seasonal patterns in the stream water NO_3–N concentrations emerge and are dominated by diffuse agricultural inputs, with an estimated 75% of the river load in the lowlands derived from arable farming. The results confirm earlier European catchment studies. Namely, current semi-distributed catchment-scale dynamic models, which integrate variations in land cover, climate, and a simple representation of the terrestrial and in-stream N cycle, are able to simulate seasonal NO_3–N patterns at large spatial (>300 km^2) and temporal (≥ monthly) scales using available national datasets.

Published in J. Environ. Qual. 37:2155–2169 (2008).
doi:10.2134/jeq2007.0507
Received 24 Sept. 2007.
*Corresponding author (tisseuil@cict.fr).

677 S. Segoe Rd., Madison, WI 53711 USA

The over-enrichment of fresh, transitional, and marine waters with nitrogen (N) can lead to the problems associated with eutrophication, such as changes in species composition of aquatic plants and nuisance algal blooms (James et al., 2005; Barker et al., in press). The main sources of N in lowland catchments are fertilizer from farming and domestic and industrial effluents, and the main sources in upland areas are from atmospheric deposition (Skeffington and Wilson, 1988). Given the diverse nature of the N problem, integrated catchment-scale modeling approaches can be used to help quantify the relative inputs of N from different sources. Also, they help in the design of strategies to remediate nutrient inputs set against a background of expected land-use and climate change because they represent the integration of key N source areas, pathways, and transformations (Ruiz et al., 2002; Wasson et al., 2003; European Parliament, 2005; Langan et al., 1997). However, the application of dynamic nutrient models is difficult because of an inability to scale measurements of N concentrations and mass from a point in space and time to a value representative of an area that is required by such models. This inability leads to problems of uncertainty in model simulations and forecasts (Beven, 1993).

As part of the European Water Framework Directive, River Basin Management Plans will be created for large river systems (European Parliament, 2005); the range of areas for the Water Framework Directive Pilot River Basins is 1200 to 37,000 km^2. Therefore, it is important to understand the key factors and processes controlling N dynamics in large catchments and to determine how well models, typically developed for smaller research catchments, perform at larger spatial and temporal scales when using available national datasets describing the climate, hydrology, water quality, and catchment characteristics such as land use and fertilizer practice. Testament to the need for such research is the large number of recent and ongoing research projects that address this topic: ELOISE (Cornell et al., 2004), CHESS (Boorman, 2003), DYNAMO (Ferrier, 1998), EUROHARP (Van Liedekerke et al., 2003), and EURO-LIMPACS (Wade et

C. Tisseuil, L. Tudesque, and S. Lek, Laboratoire Evolution et Diversité Biologique (EDB) UMR 5174, CNRS- Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 4– France. A.J. Wade, Aquatic Environments Research Centre, School of Human and Environmental Sciences, Univ. of Reading, RG6 6AB, UK.

Abbreviations: AEAG, Agence de l'Eau Adour Garonne; FYM, farm yard manure; HBV, Hydrologiska Byråns Vattenbalansavdelning; HER, hydrological effective rainfall; INCA-N, Integrated Catchment Model of Nitrogen; PCA, principal component analysis; RDA, redundancy analysis; SMD, soil moisture deficit.

al., 2004). Semi-distributed models, which make a compromise between data availability and space and time resolution, are often used to simulate the flow and nutrient dynamics in catchments (Müller-Wohlfeil, 2002). Despite the diversity in available semi-distributed dynamic N models (Arheimer and Olsson, 2003), few examples of applications to catchments larger than 1000 km^2 are reported in the literature. Such models tend to be applied in small research catchments where typically there is a substantial database describing the hydrology, water chemistry, soil, and land-use. At the daily time step, the semi-distributed Hydrologiska Byråns Vattenbalansavdelning (HBV)-N (Arheimer and Wittgren, 1994), INCA-N (Whitehead et al., 1998; Wade et al., 2002), and the SWAT model (Soil and Water Assessment Tool) (Arnold et al., 1998) are among the few nutrient models that have been applied to catchments larger than 1000 km^2.

The aim of this study is to assess the ability of the coupled HBV and INCA-N models to simulate the flow and stream water nitrate (NO_3–N) concentrations observed in the Garonne river in southwest France, an example of a large heterogeneous catchment (62,700 km^2) that incorporates different climatic and land use zones. The simulations of the N budget are used to investigate the relative input of N from point and diffuse sources in each of the climate zones. INCA-N has been used to simulate in-stream N dynamics in a broad range of ecosystem types, ranging from intensively farmed systems in northwest Europe to forest systems in Brazil (Whitehead et al., 1998; Wade et al., 2001, 2002, 2004; Neal et al., 2002), but no single application has been made to such a large river system covering such a diverse range of climate and land cover types. The Garonne is a large gravel-bed river important to the regional economy because of agriculture, viniculture, tourism, conservation, and navigation. Algal blooms and oxygen depressions occur in the middle and lower reaches during summer low-flow periods. The application of HBV and INCA-N builds on an initial assessment of the stream water NO_3–N dynamics done in this study using multivariate statistical techniques. Previous studies suggest that two factors control the stream water NO_3–N dynamics: land use and a downstream transition from nival to pluvial regime (Probst, 1985; Etchanchu, 1998). The predominant form of N in the Garonne is NO_3, and this form of N is assessed in this work as a starting point for model testing and load evaluation. The null hypothesis tested is "Current semi-distributed models cannot be used with readily available national data sets to simulate the observed stream water NO_3–N dynamics in a major European catchment." The objectives of this study were (i) to review the stream water NO_3–N concentration data collected between 1991 and 2005 to determine the dynamics and possible casual factors for the observed spatial and temporal patterns using principle component analysis (PCA) and redundancy analysis (RDA); (ii) to apply the HBV and INCA-N models to assess the capability to represent the observed NO_3–N dynamics in the largest system to which the linked models have ever been applied for calibration (1996–2005) and test (1991–1995) periods; and (iii) to evaluate the annual mean load of NO_3–N (kg N ha^{-1} yr^{-1}) exported from different land use types and to compare the results with previous investigations of the catchment to provide another assessment of model behavior and also to assess the dominant sources of NO_3–N pollution within the Garonne basin.

Study Area

The Garonne is the principal catchment of southwest France and is the country's third longest river. With a length of about 640 km from its source in the Pyrenean massif in Spain (from 1870 m altitude) to its mouth in the Atlantic Ocean (near Bordeaux in France) (Fig. 1), it covers an area of approximately 60,000 km^2. The Garonne is eutrophic in the middle and lower catchment from Toulouse to Bordeaux, where summer algal blooms and oxygen depressions regularly occur. The mean slope decreases from 3.9‰ in the Pyrenees to 0.25‰ in the lowland plain.

The climate and hydrology of the catchment are largely influenced by orographic factors. The Pyrenees dominate the Garonne upstream of Toulouse where the hydrology is influenced by snow, and typically snow melt leads to high spring flows. Downstream of Toulouse, the flow in the Garonne is derived from precipitation over the Central Massif and the Pyrenees. In the lower reaches of the Garonne, western winds from the Atlantic Ocean cause high precipitation and cool temperatures. The Mediterranean climate is less of an influence on the hydrology but is manifested by hot and dry southeastern winds around Toulouse, such as the Föhn type, which are typified by infrequent but intense summer rainfall. The flow of the catchment is regulated by approximately 210 dams; most of them are used to generate hydroelectricity. Of these dams, 40% are located upstream of Toulouse, and 12 cross the main channel of the Garonne. At the daily time step, dams are responsible for significant fluctuations in the discharge (Sauvage et al., 2003).

Agriculture occupies approximately 60% of the total catchment area. Of this, approximately 50% is under cereal production: 60% maize, 30% durum and common wheat, and 10% oilseed (sunflower, colza, and soya). The remaining agricultural area, classified as "other agriculture," is dominated by vineyards and fruit trees. Agriculture dominates the middle and lower catchment, whereas woodland and grassland dominate the headwaters. In parallel to the downstream increase of agriculture, the urban area also increases; the towns and cities of Toulouse, Agen, Tonneins, and Marmande are located alongside the main river channel. Commensurate with this increase in urban area, the number of urban and industrial effluent inputs also increases downstream.

Data Resource

Daily time series of precipitation, mean air temperature, and potential evapotranspiration from 133 climate stations maintained by METEOFRANCE were available for this study. There are seven flow gauging stations along the main channel of the Garonne providing continuous (15-min) discharge data from 1991 to 2005 (Fig. 1). These gauges are maintained by the Direction Régionale de l'Environnement. Stream water NO_3–N and NH_4 concentrations were measured at monthly intervals by the Agence de l'Eau Adour Garonne (AEAG) at 16 sites along the Garonne from 1991 to 2005 (Fig. 1). The analy-

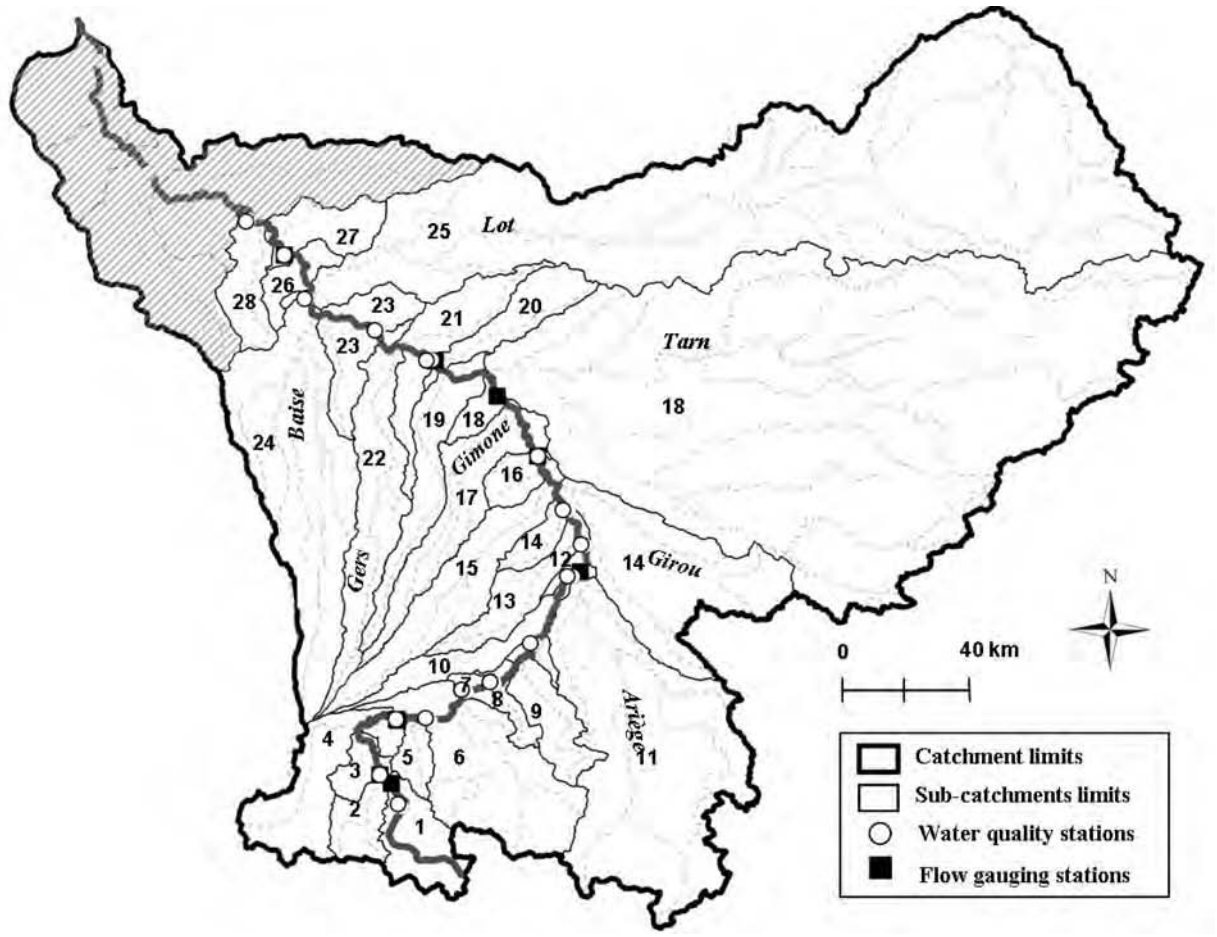


Fig. 1. Garonne catchment boundaries and reach/subcatchment structure defined for the INCA application (1991–2005).

sis of the water samples was done according to the guidelines of the French standards association. The detection limit for the stream water samples varied during the measurement period. The detection thresholds were 0.23 mg N L^{-1} for NO_3–N and 0.07 mg N L^{-1} for NH_4–N between 1999 and 2005, whereas the detection limits were 0.05 and 0.01 mg N L^{-1} for NO_3–N and NH_4–N before 1999. Daily ground water chemistry data, monitored at monthly intervals, were obtained from the ADES database (Bureau des Ressources Géologiques et Minières), and these data were used to estimate the initial ground water N concentrations for the INCA-N calibration.

Land-use data were available from the Corine Landcover map (Institut Français de l'Environnement), and statistics describing the land cover where obtained from the AGRESTE database (Ministère de l'Agriculture et de la Pêche). Data describing fertilizer amounts and application periods and crop growth periods were provided by the Chambre d'Agriculture and the ARVALIS institute. Inorganic fertilizers and farm yard manure (FYM) are used in the Garonne, although FYM is estimated to contribute less than 5% of the total annual fertilizer load (Rabaud and Cesses, 2004). Total NO_3–N and NH_4–N bulk atmospheric deposition (mg N L^{-1}) were taken from the Renecofor maps (Croisé et al., 2002).

The annual arithmetic mean of effluent flows and NO_3–N and NH_4–N concentrations in urban sewage were obtained from the AEAG (Tables 1 and 2). Missing annual mean flows and concentrations were estimated from the long-term mean over the period of record from 1991 to 2005. Industrial effluents were not considered in this study because, according to the AEAG, they represent only 5% of the total N effluent input; the influence of industrial effluents at the annual timestep was assumed negligible relative to urban sewage.

Materials and Methods

Preliminary Exploration of the Spatial and Temporal Dynamics of Flow and NO_3

For the purposes of the following statistical analysis and application of the INCA-N model (Whitehead et al., 1998; Wade et al., 2002), 28 reaches were defined from the headwater source to the beginning of the estuary. The reach boundaries were defined according to the locations of water sample sites, discharge gauges, effluent inputs, and confluence points (Fig. 1 and Table 2). Reach lengths and the associated subcatchment area were determined using ARCVIEW 9.1 (ESRI, Meudon, France) and digital elevation data. Each subcatchment was at least 300 km^2 (Table 2).

A principal component analysis (PCA) was done using monthly mean flows from 1991 to 2005 as columns in the

Table 1. Summary of data used in INCA modeling of the river Garonne from 1991 to 2005.

Data	Description	Source of data
Stream water NO_3 and NH_4 concentrations, mg N L^{-1}	Spot samples from 16 sites along the stem of the river Garonne; monthly sampling from 1996 to 2005	Agence de l'eau Adour Garonne
Effluent NO_3 and NH_4 concentrations (mg N L^{-1}) and flow, m^3 s^{-1}	Theoretical annual concentrations and spot flows samples to calculate mean daily flow and concentrations through the years	Agence de l'eau Adour Garonne
River flows, m^3 s^{-1}	Daily flows from seven gauging stations on the main stem of the river Garonne (1996–2005)	DIREN
Rainfall, temperature, and evapotranspiration, mm	Daily measurements on 18 stations localized on the catchment	METEOFRANCE
Hydrological effective rainfall and soil moisture deficit, m	Daily estimations derived from the HBV hydrological model for seven subcatchment groups; necessitating daily temperature (°C), rainfall data, actual flow, altitude, and land use area for each group	HBV (Bergström, 1992)
Base flow index; a and b parameters from the velocity/flow relation $V = aQ^b$	Derived from flow gauging stations and extended to eight groups of subcatchments, assumed environmentally homogeneous	DIREN
Fertilizer practice: application, kg N ha^{-1} yr^{-1}	Annual rate of fertilizer applications according to the variety of crops (cereals, oilseed) and regional practices	AGRESTE statistics (DRAF)
Wet and dry NH_4 and NO_3 atmospheric depositions, mg N L^{-1}	Annual mean of total NH_4 and NO_3 deposition from digitalized map; equitable sharing between wet and dry depositions from the total depositions	RENECOFOR NETWORK (Croisé et al., 2002)
Ground water NO_3 and NH_4 concentrations, mg N L^{-1}	Spot samples from 21 stations across the catchment issued from the ADES database	Bureau des Ressources Géologiques et Minières
Land use, km^2	corine landcover map	Institut Français de l'Environnement
Fertilizer practice: timing	Start day and period of fertilizer applications for crops constituting cereals and oilseed categories	ARVALIS

analysis, and the analysis was centered by site. For each site, the mean of the monthly mean flows was subtracted from each of the monthly mean flows so that the centered distribution had a mean of 0. As such, the pattern of the monthly mean flows rather than the absolute flows was compared between sites. The spatial patterns in the stream water NO_3–N and NH_4–N concentrations were analyzed using a redundancy analysis (RDA), also called a principal component analysis with instrumental variables. This was done to determine the direct influence of land use management on in-stream inorganic N concentrations. For this analysis, the observed stream water NO_3–N and NH_4–N concentrations were set as two quantitative dependant variables, and seven explanatory variables characterized the percentage of the six land use types and the annual fertilizer rate (kg N ha^{-1} yr^{-1}) within each of the 28 subcatchments. The results from the PCA and the RDA were considered together by delimiting three groups of subcatchments, representative of different climate types, to help identify the seasonal dynamics in the stream water flow and N concentrations. The seasonal analysis was done for each of the three groups with another RDA, using the 12 months of the year as dummy predictor variables and flow and NO_3–N concentrations as the two predictands. All statistical analyses were done using the R software using ade4 package (available at www.r-project.org).

INCA-N Set-up

The structure of INCA-N is described in detail in Wade et al. (2002) and shown in Fig. 2. Briefly, the model is semi-distributed so that spatial variations in land use and management can be taken into account. The flow and nutrient fluxes from different land-use classes and subcatchment boundaries are modeled simultaneously at a daily step, and the information is fed sequentially into a multi-reach river model. The input fluxes taken into account are atmospheric deposition of NH_4–N and NO_3–N (wet and dry), NH_4–N and NO_3–N fertilizer applications, mineralization of organic matter (to form NH_4–N) and nitrification (to form NO_3–N), and N fixation by plants. From these are subtracted various output fluxes, such as plant uptake, NH_4–N immobilization, and NO_3–N denitrification, before the amount available for stream output is calculated. The model also accounts for stocks of NH_4–N and NO_3–N in the soil, ground water pools, and stream reaches. The model was applied using the data resource described in the previous section and summarized in Table 1. The model was calibrated for the period 1996 to 2005 and tested for an independent period (1991–1995). INCA-N version 1.11 was used in this study. For each of the 28 reaches, INCA-N simulates the flow and in-stream NO_3–N and NH_4–N concentrations and quantifies the N processes and fluxes into the soil and ground water within each subcatchment and in-stream.

Based on the Corine Landcover and the French agricultural statistics, six land classes were defined: cereal, oilseed, other agriculture, urban, grassland, and woodland. The area of each land class area per subcatchment was calculated using ARCVIEW 9.1 (ESRI) (Table 2). It was assumed that FYM, together with the waste from grazing animals, is added to the unlimited pool of organic N available for mineralization. Due to a lack of a method to readily distinguish between the wet and dry components of atmospheric N deposition, the bulk deposition mass was split equally into the wet and dry forms. The base flow index is used as a measure of the base flow characteristics of catchments (Gustard et al., 1987). It provides a systematic way of assessing the proportion of base flow in the total runoff of a catchment (Table 2).

To account for the spatial variability in the climate, the 28 subcatchments were grouped into seven climatic regions that represented a transitional shift from a nival to pluvial regime (Table 2). INCA-N requires an estimate of actual precipita-

tion, hydrological effective rainfall (HER), soil moisture deficit (SMD), and air temperature for each climate region. For each of the seven subcatchment groups, daily observed precipitation and air temperature data were used directly as input into INCA-N; HER and SMD were derived from the rainfall-runoff hydrological HBV model (Bergström, 1992). The HBV model has been used successfully in past INCA-N applications. The model is capable of simulating snow accumulation and melt as well as HER and SMD (Kaste and Skjelkvale, 2002). The HBV model was set up for each of the seven climate data sets with daily rainfall, temperature, and in-stream flow time series as input variables and was calibrated from 1991 to 2005.

INCA-N Calibration and Validation

Calibration followed the manual procedure proposed by Butterfield et al. (2006). Briefly, because the simulated N concentrations in the land and in-stream components of INCA-N depend on water volumes, the hydrology of the terrestrial and aquatic components of the model was calibrated first. The terrestrial N processes, such as nitrification, denitrification, and mineralization, that directly affect NH_4–N or NO_3–N concentrations were then adjusted, as were those parameters relating to the physiological characteristics of plants (i.e., growth period and uptake rate). The loads associated with the land-based N processes were kept within values reported in the literature (Table 3; from Butterfield et al., 2006). This is a use of "soft-data" as described by Rankinen et al. (2006). The in-stream biological processes were then adjusted. The in-stream nitrification rate was modeled as being higher in the upper reaches because of the observed higher oxidation potential. The oxygen concentrations measured by AEAG are approximately 9 mg O_2 L^{-1} above

Table 2. Reach details for the application of INCA to the Garonne.

Reach ID	Reach length	Base Flow Index (ø)	Catchment area	Urban	Cereals	Oilseeds	Other culture	Grassland	Woodland	Gauging stations- DIREN name/code	Water quality station (WA name/code)	# dams†	# urban effluents‡	Climatic set§
	m		km²				%							
1	48,400	0.66	320	0	2	1	2	25	70	St Béat (O0010040)	Pont du Roi (5184000)	2	0	A
2	3800	0.64	563	2	0	0	0	44	53	Chaum (O0050010)	Chaum (5183900)	11	1	B
3	20,300	0.64	464	2	4	1	5	22	66				1	C
4	17,700	0.68	1271	3	7	1	9	30	50	St Gaudens (O0200040)	Valentine (5181800)	20	2	C
5	11,900	0.68	455	2	7	3	6	19	63		Labarthe (5181000)	1	2	D
6	17,000	0.68	2048	2	4	2	14	26	52		Boussens (5177600)	11	3	D
7	14,500	0.68	328	7	17	8	15	19	34		Cazeres (5177000)	1	2	D
8	9200	0.68	400	1	11	5	23	21	38			1	1	D
9	7000	0.68	778	1	9	4	24	25	36		Marquefavre (5175800)	2	1	D
10	35,400	0.68	808	5	28	13	25	17	12		Pinsaguel (5174000)	1	8	D
11	900	0.64	4336	2	7	3	29	21	38	Portet (O1900010)		33	0	D
12	14,300	0.64	316	63	12	6	11	2	7		St (Pierre (5163000)	1	3	E
13	12,200	0.64	793	16	29	14	25	8	9		Gagnac (5161000)	6	4	E
14	9200	0.64	2485	8	31	14	33	5	9			12	1	E
15	5000	0.64	1383	1	35	16	29	10	10			1	1	E
16	12,100	0.60	511	2	34	14	36	2	11	Verdun (O2620010)	Verdun (5154500)	1	1	E
17	24,500	0.60	1285	1	37	15	34	6	8			3	2	F
18	13,200	0.60	15746	2	11	3	34	17	33			56	4	F
19	17,000	0.60	1073	1	38	15	34	4	8			2	3	F
20	2900	0.58	794	1	21	7	41	10	21	Lamagistere (O6140010)	Lamagistere (5117000)		1	F
21	17,300	0.58	1053	1	33	12	37	8	10			2	6	G
22	8000	0.58	1533	3	35	14	28	11	9		Aqueduc (5112000)	1	1	G
23	22,200	0.58	924	3	34	11	35	11	7			2	5	G
24	7000	0.58	3170	1	29	11	32	8	19		St (Leger (5104000)	6	0	G
25	4300	0.58	11788	1	6	1	22	30	40			31	1	G
26	16,300	0.59	511	2	30	9	31	4	24	Tonneins (O9000010)	Mas d'Agenais (5083580)		1	G
27	20,600	0.59	754	3	37	11	38	2	9			1	3	G
28	10,100	0.59	733	1	13	4	16	8	58		Couthure (5081000)		2	G

† Number of counted dams across subcatchments.

‡ Number of point urban effluents.

§ Climatic set groups for INCA and HBV application to represent a transitional shift from a nival to pluvial hydrological regime.

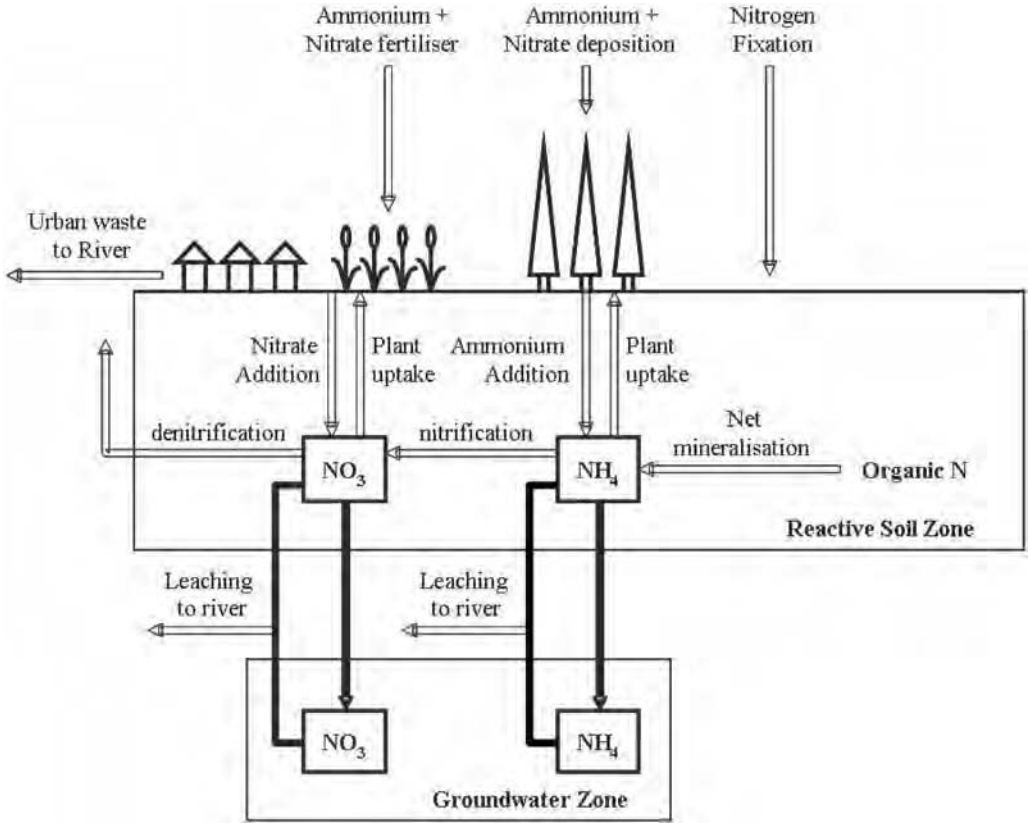


Fig. 2. INCA N process schematic (adapted from Wade et al., 2002)

reach 4 and 5 mg O_2 L^{-1} downstream of Toulouse (reach 12). The in-stream denitrification rate was increased downstream from the headwaters to the freshwater limit because moving downstream, the river becomes deeper, slower, and less oxygenated. Following the recommendation by McIntyre et al. (2005), all the model parameters were considered in a final calibration step to best match the simulated flow and stream water NO_3–N and NH_4–N concentrations to those observed. The goodness-of-fit for each subcatchment was estimated graphically and by two statistics, the coefficient of efficiency (*E*) (Nash and Sutcliffe, 1970) and the R^2, given as:

$$E = 1 - \frac{\sum_{t=1}^{T}(O_t - P_t)^2}{\sum_{t=1}^{T}(O_t - \bar{O})^2} \quad [1]$$

$$R = \left\{ \frac{\sum_{t=1}^{T}(O_t - \bar{O})(P_t - \bar{P})}{\left[\sum_{t=1}^{T}(O_t - \bar{O})^2\right]^{0.5}\left[\sum_{t=1}^{T}(P_t - \bar{P})^2\right]^{0.5}} \right\}^2 \quad [2]$$

where *O* and *P* represent observed and predicted values from the initial, *t* to the final, *T* daily time step in the formula. $\bar{O}$ and $\bar{P}$ are the mean of the observations and of the predictions, respectively. R^2 is the square of the Pearson's product moment correlation coefficient and describes the proportion of the total variance in the observed data that can be explained by the model. It ranges from 0.0 to 1.0 (perfect model). *E* has been widely used to evaluate the performance of hydrologic models, which ranges from minus infinity to 1.0 (perfect model). If the square of the differences between the model simulations and the observations is as large as the variability in the observed data, then $E = 0.0$, and if it exceeds it, then $E < 0.0$. Thus, a value of zero for *E* indicates that the observed mean is as good a predictor as the model, whereas negative values indicate that the observed mean is a better predictor than the model (Legates and McCabe, 1999). However, because R^2 and E are functions of the squared difference between the observations and simulations, they are sensitive to extreme rare values. The coefficient of efficiency is affected by bias in the model predictions. Namely, if a model produces the correct pattern but all the values are mean shifted, then the value of *E* will be low. As such, R^2 and *E* are typically used to assess the model fit to the observed flow dynamics, but only R^2 is used to compare the simulated and observed solute dynamics. In this study, R^2 and *E* are used to provide a rigorous assessment of model flow and nitrate behavior during calibration and testing.

Due to an inability to scale point flux measurements to be representative of an area, it is difficult to identify the optimum parameter set, and different parameter combinations can give equally acceptable results (Beven, 1993; Durand, 2004). All calibrated parameter values have limited physical

Table 3. Simulated N loads and process from the INCA-N calibration in comparison with "soft data" measurements (Rankinen et al., 2006) taken from Butterfield et al. (2006).

Process	Land cover types	Simulations	Vegetation/ecosystem type	Range	Literature source
		kg N ha^{-1} yr^{-1}		kg N ha^{-1} yr^{-1}	
N total load					
	woodland	8			
	grassland	8			
	cereals	173			
	oculture	54			
	oilseed	56			
	urban	20			
N fixation					
	woodland	0			
	grassland	0			
	cereals	0			
	oculture				
	oilseed	0			
N retention					
	woodland	93			
	grassland	63			
	cereals	120			
	oculture	58			
	OILSEED	50			
N uptake					
	woodland	136	deciduous forest (range of sites)	72–153	Melillo (1981)
	grassland	92	heather moorland	42	Miller (1981)
			unimproved grassland (Snowdonia)	162	Heal and Perkins (1978)
	cereals	153	winter wheat	95	Miller (1981)
	oculture	138	crop	200	Powlson (1993)
	oilseed	144	range of crops grown in USA, from wheat, silage corn and fertilized grass hay	100–350	Brady and Weil (1996)
Denitrification					
	woodland	4	coniferous forests across Europe (NITREX sites), temperate coniferous forests (range of soil types)	<0.01–4	Reynolds et al. (1998)
	grassland	2	unimproved grassland (N Wales)	1	Emmett et al. (1997)
			unimproved grassland (grass-clover/herballey)	3.4–4.4	Ruz-Jerez et al. (1994)
	cereals	69	Rothamsted experimental plot, fertilized	10–50	Powlson (1993)
	oculture	36	restricted drainage, large amounts of applied fertilizer	30–60	Brady and Weil (1996)
	oilseed	37			
Mineralization					
	woodland	123	sitka spruce forest, N Wales (net mineralization, forest floor)	10–292	Emmett et al. (1997)
			typical deciduous forest	25–149	Melillo (1981)
	grassland	82	Dutch heathland	44–126	Vuuren et al. (1992)
	cereals	103	accumulation during late summer, autumn, and early winter	30–100	Powlson (1993)
	oculture	71	net mineralized, Jealott's Hill Research Station, crop of winter wheat	171	Rowell (1994)
	oilseed	74			
Nitrification					
	woodland	14	Welsh spruce forests	15	Stevens et al. (1994)
			Dutch coniferous forests	1–35	Tietema (1993)
	grassland	9	Dutch heatland	3–54	Vuuren et al. (1992)
	cereals	43	arable fertilized	10–50	Powlson (1993)
	oculture	35			
	oilseed	50			
Inorganic N leaching					
	woodland	13	Welsh spruce forests	0–30	Emmett et al. (1993)
	grassland	13	Welsh moorland	1.8–5.3	Stevens et al. (1994)
	cereals	31	fields growing continuous arable crops for many years	15–65	Rowell (1994)
	oculture	19	(fertilized with 96–192 kg N ha^{-1} yr^{-1})	40–41	Addiscott and Powlson (1989)
	oilseed	21	Rothamsted experimental fields	20–100	Powlson (1993)

meaning because they do not relate to quantities that can be measured directly. Despite this, the model application is useful because it provides a methodology for testing concepts of catchment functioning and, once calibrated and tested satisfactorily, the exploration of future management and climate scenarios.

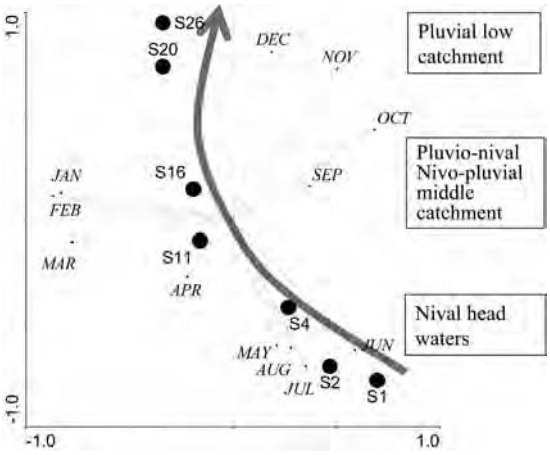


Fig. 3. Standardized principal components analysis, centered by subcatchment (S), of the monthly mean flow on the Garonne between 1991 and 2005, characterizing the transition from nival to pluvial hydrological regime across the catchment. Axis 1 and 2 explain 37.3 and 27.9% of the total variation, respectively. Subcatchments are plotted as passive variables and grouped according to their hydrological regime: nival (S1, S2, S4), nivo-pluvial (S11, S16), and pluvial (S20, S26).

Results

Flow and NO_3–N Dynamics

Spatial Factors Controlling the Observed Flow and Stream Water Nitrogen Concentrations

The first two components of the PCA explain 65% of the total variability in the monthly mean flows calculated for the outlet of each of the seven subcatchment groups (Fig. 3). In Fig. 3, the length of the thin PCA arrows represents the magnitude of the mean flow for a given month, and the direction indicates the degree of correlation with the other monthly mean flows; the higher the correlation between two variables, the smaller the angle between the arrows. The first axis explains 37% of the variation and shows a clear seasonal flow pattern. The monthly mean flows for January, February, March, and April plot at the left of the axis and are separated from the other months, which plot to the right. The second axis, which explains 28% of the variation, represents the upstream-downstream gradient characterizing a nival to pluvial transition in the hydrological regime, as shown by the hard-drawn arrow. The combination of the two axes in the analysis reveals the relationship between hydrological patterns according to location: The nival regime occurs in Pyrenean headwaters (S1, S2, S4) where the hydrograph has a double peak; one peak occurs typically in the late autumn and early winter (October to December) in response to rainfall, and a second peak occurs during the snowmelt period typically from March to June. During January and February, the precipitation falls predominantly as snow. The regulation of flows by dams may also reduce the mean monthly flows during the winter and spring periods. These double hydrographs are shown and discussed further in the section "Spatial and Temporal Dynamics within the Geographical Zones." Further downstream, the monthly mean flows in the Garonne represent a nivo-pluvial/pluvio-nival system (S11 and S16), which still exhibits a double peak in the annual hydrograph and a transition to a pluvial hydrological regime (S20 and S26) where the double-peak is evident but less pronounced than in the upper reaches. In the subcatchments S11 to S26, the highest monthly mean flows occur in November and December in response to rainfall mainly over the Central Massif. The curved arrow drawn in bold in Fig. 3 illustrates the spatial transition from a nival to pluvial regime from the top to bottom of the catchment.

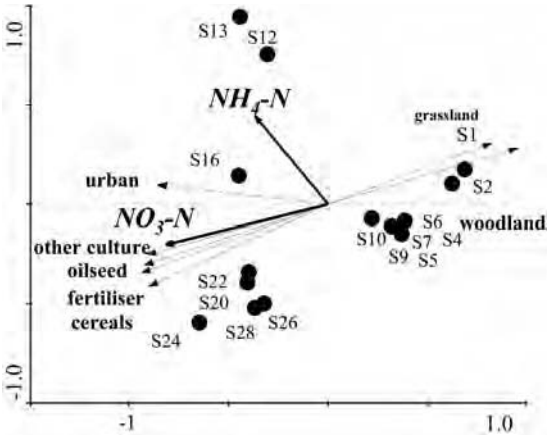


Fig. 4. Redundancy analysis explaining the spatial variation in NO_3–N and NH_4–N concentrations variables according to land use percentage and annual fertilization rate per subcatchment (S). Fifty-three percent of the total variation is explained, and the subcatchments, plotted as passive variables, are grouped according to their physical and land use characteristics: woodland (S1 to S10), woodland to agricultural transition (S12, S13, S16), and agricultural (S20, S22, S24, S26, S28).

The RDA results (Fig. 4) describe the influence of land use and fertilizer loads (kg N km^{-2} yr^{-1}) on the stream water NO_3–N and NH_4–N concentrations. The first axis explains 39% of the total variation in the stream water NO_3–N and NH_4–N concentrations. Figure 4 shows a positive correlation between agricultural land (which is more abundant in the middle and lower reaches of the Garonne [S20–S28]), annual fertilizer amounts, and high (>5 mg N L^{-1}) in-stream NO_3–N concentrations. The downstream increase of NH_4–N concentrations from approximately 0.05 to 0.6 mg N L^{-1} is correlated to the increase of urbanization and arable agriculture. The headwaters of the Garonne catchment, mostly occupied by woodland and forest, have the lowest in-stream NO_3–N concentrations of approximately 0.1 mg N L^{-1}.

The RDA results are in agreement with the PCA examinations of monthly mean flows, which identified three hydrological regimes. The RDA results show that the catchment can be split into three geographical zones: (i) the woody headwater catchment (S1–S10) with a nival regime, (ii) the agroforestry middle catchment between S11 to S17 with a nival to pluvial transition regime, and (iii) the agricultural lowland catchment (S18–S28) with a predominantly pluvial regime.

Temporal Factors Controlling the Observed Flow and Stream Water Nitrogen Concentrations

The subcatchment monthly mean flows and stream water NO_3–N and NH_4–N concentrations were grouped according to the three geographical zones identified in the preceding PCA and

RDA. Using these three groups, another RDA was done to examine the correlations between month and the observed monthly mean flows and stream water NO_3–N concentrations (Fig. 5). In the headwaters, month of the year explains 22% of the variation in mean monthly flow and NO_3–N concentration. There is a negative correlation between monthly mean flow and NO_3–N concentrations shown by an opposition in the arrows representing flow and NO_3–N (Fig. 5a); as monthly mean flow increases, the mean monthly stream water NO_3–N concentration decreases. However, inspection of the observed hydrograph and chemograph shows a very complex pattern (Fig. 6). Between 1995 and 2000, during spring snowmelt when the flows are at a peak, the stream water NO_3–N is diluted, and the concentrations remain low during the summer months but increase during the winter high flows. During the period 1991 to 1995, stream water nitrate concentrations of approximately 0.7 mg N L^{-1} occurred during the summer and autumn months, whereas during 2000 to 2005 the concentrations showed no clear pattern, with the concentrations ranging between 0.2 and 1.2 mg N L^{-1}. In the middle reaches (S11–S17), 31% of the variation is explained by month. Between 1995 and 2005, high stream water NO_3–N concentrations coincide with autumn and winter high flows, and low concentrations coincide with spring and early summer high flows due to snowmelt and subsequent summer low flows (Fig. 5b and 7). This indicates a possible flushing of NO_3–N by elevated flow conditions during autumn and winter, dilution during snowmelt periods, and possible instream biological activity during summer. Before 1995, the observed stream water NO_3–N concentrations do not show a clear relationship with flow (Fig. 7). In the lower reaches (S18–S28), the seasonal relationship between flow and observed stream water NO_3–N concentrations is positively correlated. In these reaches, the maximum monthly mean NO_3–N concentrations coincide with the maximum flows in winter and spring (Fig. 5c and 8). The seasonal RDA investigating the relationship between months, monthly mean flow, and stream water NH_4–N concentrations did not show any significant correlations.

INCA-N Simulation Results

Spatial and Temporal Dynamics within the Three Geographical Zones

The three greatest changes in the annual flow along the main channel of the Garonne correspond to confluences with three major tributaries: the Ariège (reach 11), the Tarn (reach 18), and the Lot (reach 23). These three confluences mark the lower boundaries of the three zones defined by the PCA and RDA analyses (Fig. 9a). INCA-N tends to overestimate the annual mean flow in the lower reaches. This occurs due to a tendency to overestimate flow peaks and the falling limb of the hydrograph. The simulated annual mean NO_3–N concentrations compares well with observations (Fig. 9b). The highest concentrations are observed immediately downstream of the confluence with the Tarn (Fig. 9b), which is a large tributary dominated by agriculture (reach 18; Table 2).

The hydrological transition from nival to pluvial regime across the three geographical zones is simulated well during the calibration of the model (1996–2005) according to R^2

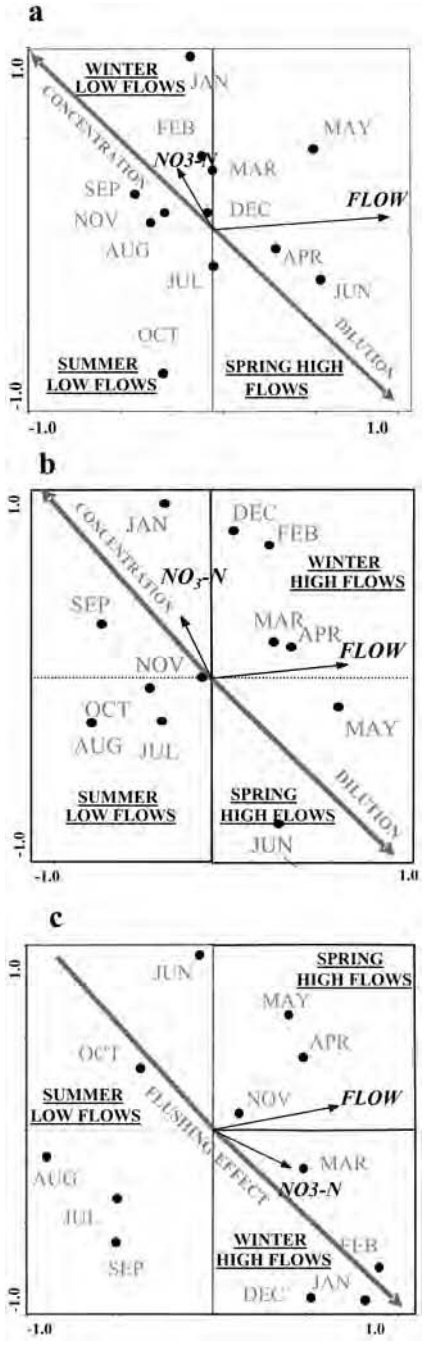


Fig. 5. Redundancy analysis characterizing the influence of season on flow and nitrate dynamics in (a) the woodland and nival headwater catchments (21%); (b) the woodland-agriculture, nivo-pluvial transitions of the mid-catchments (31%); and (c) the agricultural and pluvial catchments of the lower reaches (35.5%).

statistics. The R^2 and E coefficients for flow have mean values of 0.69 and –0.39 in the upper basin, 0.64 and –0.39 in the middle Garonne, and 0.75 and 0.12 in the lower zone, respectively (Table 4). For model validation (1991–1995), flow statistics are similar to those obtained during calibration. For flow, R^2 has mean values of 0.62, 0.69, and 0.66 for the upper, mid, and lower reaches, respectively, during validation (Table 4). The cause of negative E values is an overestimation of observed extreme high flows; the E statistics increase when

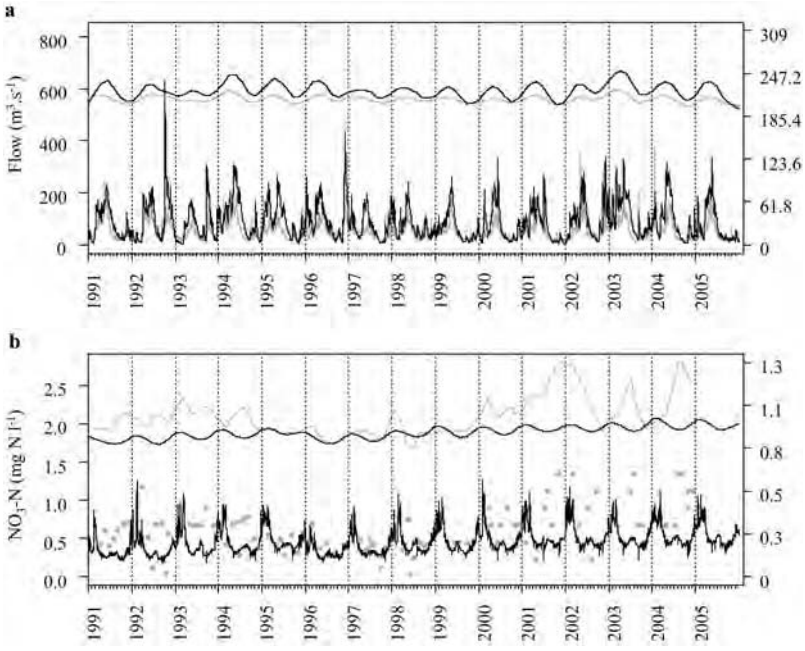


Fig. 6. The observed and simulated (a) hydrograph and (b) stream water nitrate concentrations in the Pyrenean headwaters (reach 4) of the Garonne for the period 1991 and 2005. The observed flows and stream water nitrate concentrations are plotted as a gray line and gray points, respectively. The simulated flows and stream water nitrate concentrations are plotted as black lines. The lines above the hydrograph show a smoothed curve (lowess), where the gray and black lines relate to the observed and simulated data, respectively. The scale for the lowess curve is shown on the right side of the diagram.

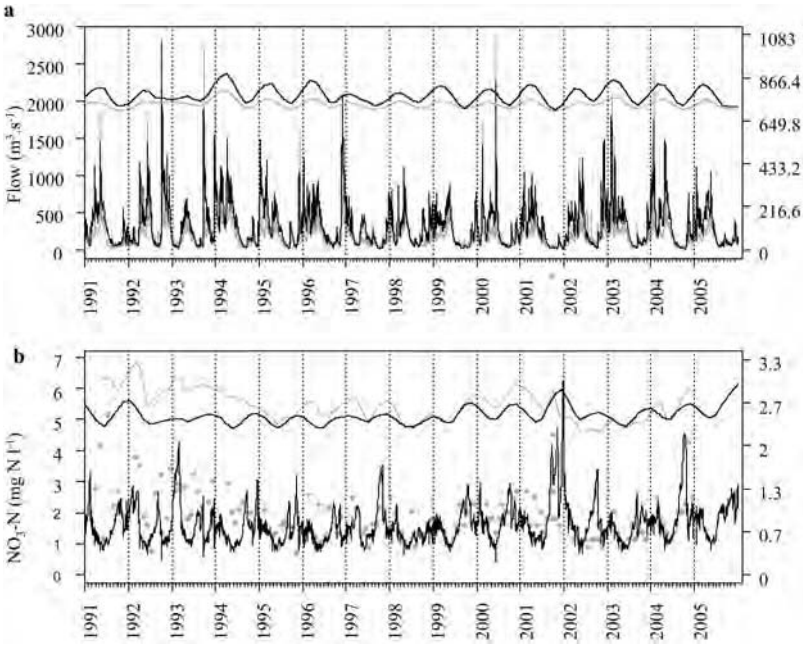


Fig. 7. The observed and simulated (a) hydrograph and (b) stream water nitrate concentrations in a mid-reach (reach 16) between Toulouse and Agen of the Garonne for the period 1991 and 2005. The observed flows and stream water nitrate concentrations are plotted as a gray line and gray points, respectively. The simulated flows and stream water nitrate concentrations are plotted as black lines. The lines above the hydrograph show a smoothed curve (lowess), where the gray and black lines relate to the observed and simulated data, respectively. The scale for the lowess curve is shown on the right side of the diagram.

the flow data are transformed by a log function, which diminishes the effect of extreme flows. The unknown influence of the operation of the dams also causes uncertainty in the flow simulations. Despite this, in the headwaters the R^2 coefficients for flow are good (R^2 ~0.61; Table 4), and the hydrograph shows the influence of the spring and early summer snow-melt and lower winter flows due to more precipitation falling as snow compared with the spring and autumn months. The good simulation demonstrates that HBV and INCA-N can reproduce the nival regime (Fig. 6a). The interannual trend for flow dynamics, illustrated by a smoothing curve (lowess) plotted above the hydrographs in Fig. 6, shows a decrease of discharge with time, which may be explained by a regional decrease in precipitation or by a reduced ground water input to the river due to reduced recharge or increased abstraction or more offtake from dam reservoirs. In the middle Garonne, daily simulated flows fit the observations well (R^2 ~0.66; Table 4), and the hydrograph successfully reproduces the seasonal nivo-pluvial dynamics as high flows are well simulated in winter and spring (Fig. 7a). In the lower Garonne (reach 26), the R^2 for flow is approximately 0.72 (Table 4), suggesting an improvement in the ability of linked HBV-INCA-N to simulate pluvial hydrological systems compared with nival systems (Fig. 8a).

To assess model behavior for NO_3–N, only the R^2 statistics were considered because E is more affected than R^2 by bias in the model predictions. The results in Table 4 suggest that the model fit is equally good in the headwaters, mid-reaches, and lowlands because the means of the R^2 values for calibration and validation, respectively, are 0.17 and 0.02 in the headwaters, 0.14 and 0.02 in the mid-reaches, and 0.18 and 0.21 in the lowlands. However, results show that the model failed to reproduce the daily variability in NO_3–N patterns because the R^2 values are globally poor. For each geographical zone (Fig. 6b, 7b, 8b), the simulated chemograph reproduces well the observed seasonal patterns described with the PCA and RDA in the middle and lower reaches; namely, the maximum NO_3–N concentrations, which are observed to occur during high flows. In the headwaters draining more extensive land use types, such as grassland and woodland, the stream water NO_3–N concentrations were observed below 0.23 N L^{-1} before 1999 (Fig. 6b). From 1999, Fig. 6b seems to show an increase in the stream water NO_3–N concentrations, which coincides with the change in the analytic methodology in 1999; the detection limit of 0.23 mg N L^{-1} after 1999 increased the minimum

detectable NO_3–N concentration and will have caused an upward shift in the lowest stream water NO_3–N concentrations measured. The observed NO_3–N concentrations also show a larger range from 2000 to 2005 compared with 1991 to 1999 (Fig. 6b). Because of the change in analytical method, it is difficult to determine if this increase in the observed stream water NO_3–N concentrations is the result of an environmental change or an artifact of the change in methodology. To determine if the increase in the detection limit after 1999 affects the goodness-of-fit statistics, the R^2 statistics were calculated before and after the analytic change for two 3-yr periods, from 1996 to 1998 (M9698) and from 1999 to 2001 (M9901) (Table 4). Table 4 indicates that, for reach 4, the model fit is better before (R^2 = 0.38) the analytical change compared with afterward (R^2 = 0.11) due to the model underestimating the observed stream water NO_3–N concentrations during the period 1999 to 2001. The R^2 values for the other reaches show no overall improvement in the goodness-of-fit for the period 1996 to 1998 compared with 1999 to 2001, and therefore the increased detection limit does not seem to have worsened model performance when measured using the R^2 statistics overall, although some worsening may be apparent in reach 4.

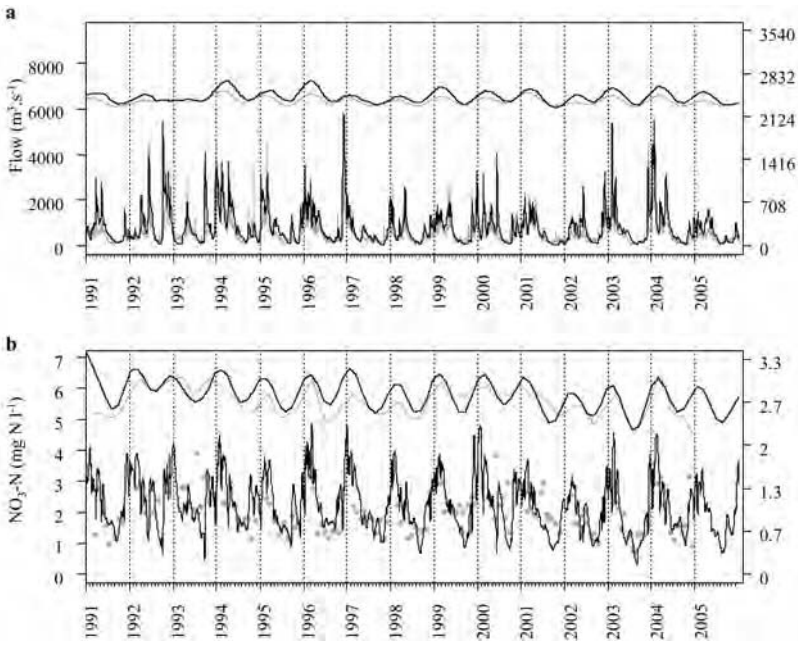


Fig. 8. The observed and simulated (a) hydrograph and (b) stream water nitrate concentrations in a mid-reach (reach 26) between Agen and Bordeaux of the Garonne for the period 1991 and 2005. The observed flows and stream water nitrate concentrations are plotted as a gray line and gray points, respectively. The simulated flows and stream water nitrate concentrations are plotted as black lines. The lines above the hydrograph show a smoothed curve (lowess), where the gray and black lines relate to the observed and simulated data, respectively. The scale for the lowess curve is shown on the right side of the diagram.

Considering the Garonne as a whole, arable lands (cereals + oilseed + "other") contribute approximately 75% of the annual mean NO_3–N load simulated at the catchment outlet. The total NO_3–N output load is estimated as 150 kt N yr^{-1}, which is approximately 2.8 t N km^{-2} yr^{-1} (Fig. 10a and 10b). This result is within the range of N fluxes estimated for other European rivers dominated by agriculture: 2 and 1.35 t N km^{-2} yr^{-1} on the Dender and Enza river (Boorman, 2003) and 2.5 and 3 t N km^{-2} yr^{-1} on the Rhine and Scheldt rivers (Garnier et al., 2002). In the upper reaches of the Garonne, woodland and grassland contribute nearly 60% of the total NO_3–N budget of the river (Fig. 10a). In the Pyrenean headwaters (reach 4), arable lands cover approximately 10% of the subcatchment area but contribute to nearly 60% of the annual NO_3–N load (Table 2; Fig. 10a). In the mid-reaches of the Garonne from reach 12 to reach 16, the sewage inputs from Toulouse (reach 12) enhance the total annual N load into the river, contributing 32% of the annual load (Fig. 10a). Downstream of Toulouse, the contribution of urban effluent declines as a percentage of the total load, whereas the load from arable lands increases (Fig. 10a). The Tarn tributary (reach 18) contributes 25% (40 kt N yr^{-1}) of the annual N load to the lower Garonne, and nearly 80% of this amount is derived from arable farming (Fig. 10a).

The modeled annual loads (kg N ha^{-1} yr^{-1}) associated with the NO_3–N and NH_4–N inputs, processes, and outputs under different land management units are shown in Table 3. Cereals, oilseed, and "other agriculture" receive the largest total NO_3–N load because of fertilizer inputs and exhibit the greatest simulated export of NO_3–N (31, 21, and 19 kg N ha^{-1} yr^{-1}, respectively). The modeled processes involving the largest transfers of N were NO_3–N plant uptake (to crops), organic matter mineralization, and NH_4–N nitrification. The load estimates for the different land use types were able to be constrained within the ranges published in the literature (Table 3; from Butterfield et al., 2006). Woodland and grassland exhibit the lowest total NO_3–N export to the river (8 kg N ha^{-1} yr^{-1}).

Discussion

This discussion focuses on three items: (i) the factors and processes controlling the observed patterns in the flow and NO_3–N concentrations at the monitoring sites along the Garonne, (ii) the emergent properties controlling the behavior of the system and how these can be used for modeled projections of future flows and

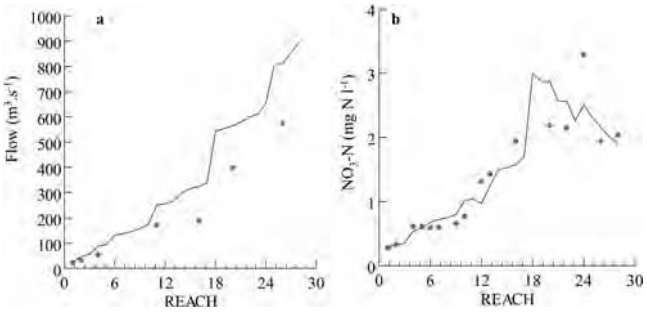


Fig. 9. Spatial trend in the INCA simulations (line) and observation (points) on the Garonne between 1996 and 2005 for (a) annual mean flow and (b) stream water annual mean NO_3–N concentrations.

Table 4. Coefficients of determination (R^2 and E) (Nash and Sutcliffe, 1970) for flow and stream water NO_3–N concentrations for calibration (cal, 1996–2005) and validation (val, 1991–1995) periods, before suspected analytic changes in NO_3–N measurements for upper reaches, 1996 and 1998 (M9698), and after changes, 1999 to 2001 (M9901).

		NO_3–N						Flow			
		R^2						R^2		E	
Zone	Reach	cal	val	M9698	M9901	Zone	Reach	cal	val	cal	val
High	1	0.14***	0.02	0.12	0.32***	High	1	0.67***	0.59***	0.60	0.37
	2	0.09***	0.01	0.03	0.03		2	0.69***	0.61***	−0.66	−1.00
	4	0.24***	0	0.38***	0.11		4	0.71***	0.64***	−1.10	−0.94
	5	0.25***	0	0.17*	0.16*	Middle	11	0.66***	0.69***	−0.13	−0.04
	6	0.18***	0.03	0.18*	0.13*		16	0.61***	0.68***	−0.65	−0.18
	7	0.16***	0.02	0.2*	0.04	Low	20	0.73***	0.65***	0.12	0.19
	9	0.28***	0.09*	0.2*	0.36***		26	0.77***	0.68***	0.12	0.22
	10	0.03*	0.02	0	0.03						
Middle	12	0.1***	0.01	0.05	0.08						
	13	0.02	0	0.01	0						
	16	0.28***	0.04	0.27***	0.36***						
Low	20	0.24***	0.25***	0.42***	0.03						
	22	0.03	0.14**	0.21*	0.07						
	24	0.27***	0.32***	0.21	0.12						
	26	0.22***	0.15**	0.27***	0.02						
	28	0.12***	0.18***	0.04	0.06						

* Significant at $p < 0.05$.

** Significant at $p < 0.01$.

*** Significant at $p < 0.005$.

stream water NO_3–N concentrations, and (iii) the ability of the linked HBV-INCA-N models using national datasets to explain the observed patterns and represent the emergent properties.

The factors and processes controlling the stream water flow and NO_3–N patterns were investigated using principle components analysis and redundancy analysis. The results showed the importance of flow controls on the stream water NO_3–N dynamics in three distinct geographical zones: the nival and woodland dominated headwaters; the nival to pluvial and woodland to agriculture transition, which characterizes the middle reaches; and the pluvial and agriculture dominated lowlands. From 1995 to 2000, in the headwaters there was an increase in stream water NO_3–N concentration as flows increased during late autumn or early winter, perhaps indicating the transport of nitrate from the soil to the stream water, and then a general dilution of the stream water NO_3–N concentrations with increasing flow during the late spring and early summer melt periods; a monthly sampling frequency was too infrequent to identify the effects of the elution of NO_3–N during the first period of melt. The simulations of the headwaters were characterized by low N

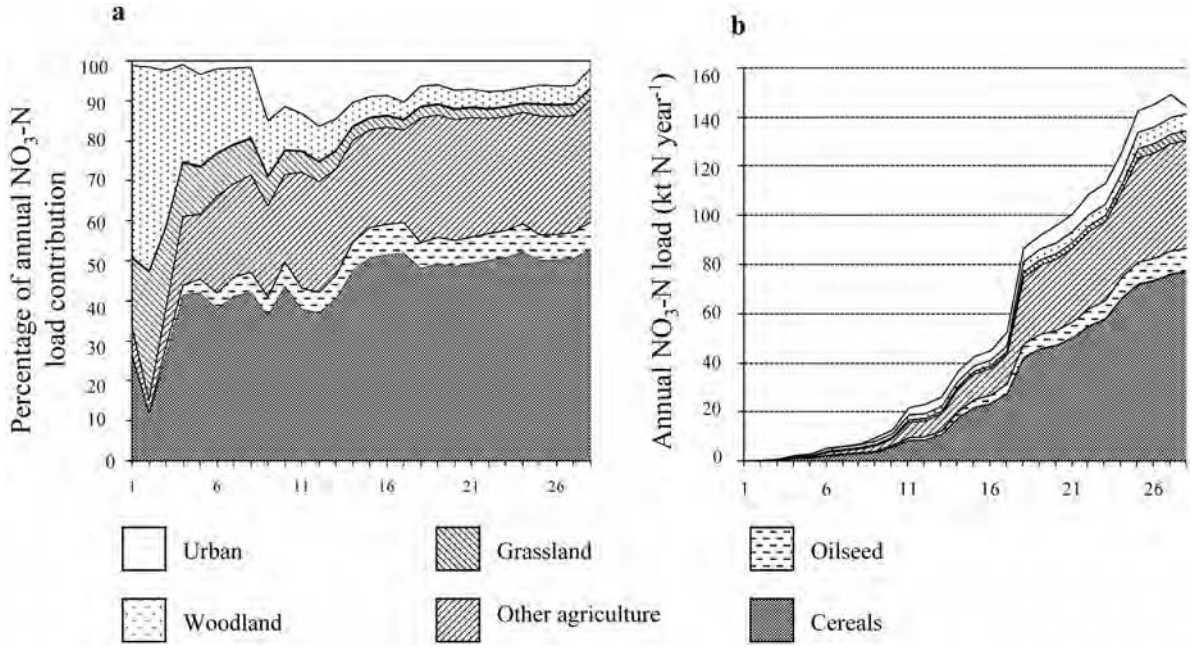


Fig. 10. The contribution of different land classes to the cumulative NO_3–N load simulated along the river Garonne expressed as (a) a percentage and (b) a load (kt N yr^{-1}).

inputs from fertilizers and atmospheric deposition, high retention of N by trees, and low in-stream productivity. Observations suggest that the headwater streams are typically oligotrophic (i.e., the stream water NO_3–N concentrations are low at approximately 0.02–1.0 mg N L^{-1}). In the lower reaches, NO_3–N is typically added as fertilizer during the winter and spring months, and therefore the stream water NO_3–N concentrations, which range from 0.08 to 5 mg N L^{-1}, are higher than those in the headwaters. The fertilizer additions coincide with the period of the year when the soil is at its wettest, and therefore mineralization and soil NO_3–N concentrations are typically at an annual maximum; soil and in-stream plant uptake and denitrification are also at a minimum. Thus, the factors and processes integrate to cause an increase in the observed stream water NO_3–N concentrations during the wetter months of December, January, and February relative to the rest of the year when flows and N inputs are lower and when there is greater terrestrial and aquatic plant uptake. In the middle reaches, the dynamics of the stream water NO_3–N concentrations suggest a transition between the upper and lower reaches. High stream water NO_3–N concentrations during autumn and winter high-flow periods indicate a possible flushing of NO_3–N from fertilizers, although dilution during snowmelt periods is evident, as is possible instream biological activity during summer, giving low stream water NO_3–N concentrations of approximately 1 mg N L^{-1}. This conceptual model of NO_3–N source and delivery controlled by hydrology and land use is supported by the work of Etchanchu (1998) and Probst (1985). In these two studies it was identified that, in the lower reaches of the Garonne, the diffuse source contribution of NO_3–N increased with flow due to a flush of NO_3–N from surface and subsurface pathways. The elevated summer stream water NO_3–N concentrations of in the upper and middle reaches between 1991 and 1995 (Fig. 6 and 7) may result from point source inputs that have now been reduced in terms of their N input to the main channel or removed. Further work is required to investigate the number of point source inputs and their impact in the upper and middle reaches before 1995.

The analysis of the factors and processes controlling the hydrology and the stream water NO_3–N concentrations show that, when considering seasonal (monthly) variations in the stream water flows and NO_3–N, two key catchment characteristics emerge that explain the observed patterns. The first is the seasonal precipitation pattern. The flows in the Garonne are correlated to the rainfall input and snowmelt. The second key characteristic is land management. At the seasonal time step in the subcatchments defined for the application of the statistical analyses and the INCA-N application, there was a strong relationship between the percentage of agricultural and urban land and in-stream NO_3–N concentration. Such a relationship has been identified in other studies of smaller rivers, across regions, and across Europe (Edwards et al., 1990; Neal et al., 2002; Davies and Neal, 2004). Land use and management are correlated to monthly stream water NO_3–N concentrations because of the seasonal variations in fertilizer and manure applications and because of the seasonal variation in the flow available to wash in excess fertilizer and manure and potential to dilute effluent inputs (Neal et al., 2006).

Given the two emergent catchment characteristics that control stream water NO_3–N patterns in large catchments at the seasonal time-scale, the modeling approaches based on that incorporate these factors work well. The results suggest that in all reaches, the two coupled semi-distributed models, INCA-N and HBV (which together include a simple representation of the N cycle; flow pathways through the soil and ground water and snow-pack dynamics; and N stores in the soil, ground water, and in-stream and which use readily available national datasets describing the hydrology and land management) can reproduce the observed seasonal patterns in flow and stream water NO_3–N concentrations. Thus, the null hypothesis defined in the Introduction was rejected. This type of modeling approach seems to be a pragmatic way to help understand how key factors (e.g., hydrological and N inputs) and processes (e.g., routing along flow pathways) are integrated in large river-systems. This approach can also be used to apportion the sources of N that contribute to the in-stream load and provide a load estimate at the catchment outlet. The daily estimates of flow and NO_3–N concentrations could be used as input to an estuarine model, as demonstrated in an application of HBV and INCA to the Birkenes river in Norway (Kaste et al., 2006). Water quality data would need to be more typical of research catchments monitored at the highest frequency possible to help identify the model structure and parameters (Kirchner, 2006). The modeled nitrate export rates are within the ranges estimated for other European rivers, although the stream water nitrate concentrations in the Garonne are low at less than 5 mg N L^{-1}. This raises the question about the importance of phosphorus in the eutrophication problems associated with the mid- and lower reaches of the Garonne.

Uncertainties in the model input data, structure, and parameterization remain. In this model application, it was assumed that industrial effluents were less important than urban effluent. Given the goodness-of-fit in the lower reaches, this assumption seems reasonable, but when considering daily dynamics, effluent inputs of an episodic nature must be considered. For example, the high NO_3–N concentrations observed at Toulouse (reach 12) in September 2001, which occurred after the release of N into the Garonne from a factory, demonstrate the potential effects of industrial discharges (Fig. 7). High-frequency water quality data would allow the impacts on the stream water quality of such inputs to be characterized. However, modeling these events and other short-term episodic events is difficult in a deterministic way because the causes of elevated steam water NO_3–N concentrations are not always known. An alternative approach might be to represent short-term events within a stochastic component embedded within a deterministic, physically based model or to use a Monte-Carlo–based approach with model inputs, such as urban and industrial effluents or atmospheric inputs, defined as input distributions. The next phase of this work will use uncertainty techniques such as those demonstrated by McIntyre et al. (2005), Rankinen et al. (2006), and Futter et al. (2007) to explore how the INCA-N model outputs are affected by the variability of the model parameters and input.

In terms of reducing the NO_3–N load in the Garonne, the easiest input to reduce would be the urban effluent from Toulouse.

Reduction of the inputs from arable land would be more difficult but may be achieved through a reduction in the amount of fertilizer and manure applied. Preliminary model simulations suggest that any reduction may take at least 10 yr to have an effect on the ground water and stream water concentrations. Other monitoring and modeling studies have shown that although fertilizer reduction reduces leaching at the bottom of the soil profile, the NO_3–N stored in the ground water can confound recovery in the river (Ruiz et al., 2002; Jackson et al., 2007). Further work is required to assess how changes in the flow and stream water NO_3–N concentrations affect stream water biological communities in the Garonne basin, which are thought to be sensitive to stream water NO_3–N concentrations and to flow, sediment, and light conditions.

Conclusions

The results of this study demonstrate that in large river systems dominated by agricultural inputs, simple sinusoidal patterns emerge in the monthly stream water NO_3–N concentrations due to climate and agricultural inputs. In the agricultural, low-, and mid-reaches of the Garonne, the NO_3–N concentrations exhibit a seasonal pattern with a peak concentration coinciding with fertilizer applications and a minimum concentration in summer coinciding with terrestrial and aquatic uptake in plant biomass. In the upper reaches, the climate controls on flow are the most important in determining the stream water nitrate patterns. There are no discernable stream water nitrate dynamics that can be attributed to point sources except for an extreme event where there was an uncontrolled release of high nitrate water into the Garonne from a factory in Toulouse in September 2001, although further work is required to investigate if elevated stream water nitrate concentrations observed before 1995 are the result of effluent discharges. The linked HBV-INCA-N model was able to simulate the stream water nitrate response to climate and agricultural transitions down the Garonne and thereby provided a useful tool to simulate agricultural and urban catchments at a scale commensurate with the largest river-systems in Europe.

Acknowledgments

This work was done as part of the EU FP6 Integrated Project "EURO-LIMPACS" (GOCE-CT-2003-505540). We thank the following, who provided data and helped to set up model: J-P Serrano, G.Navarro, J-L Scharffe, S. Gomes, L.Verdié, and Paschini M. (AEAG) for water quality and effluent data and their advice; Meteofrance for daily climate data; J-C Teurlay and C.Villa (DRAF) for land management data; A. Probst (LMTG, CNRS) for N deposition data; C. Delmas and M. Douard (Direction Régionale de l'Environnement) for industrial effluent data; J-L Le Rohellec (Direction Régionale de l'Environnement) for flow data.; P. Durand (SAS, INRA) for advice on the INCA-N application; P. Gate and C. Lesoudère (ARVALIS) for crop growth data; J-L Probst (ENSAT) and S. Sauvage (ECOLAB) for their valuable knowledge of the Garonne system; the Bureau des Ressources Géologiques et Minières for ground water data; and B. Jackson from Imperial College, London for Fig. 2.

References

Addiscott, T., and D. Powlson. 1989. Laying the ground rules for nitrate. New Sci. 122:28–29.

Arheimer, B., and J. Olsson. 2003. Integration and coupling of hydrological models with water quality models. Applications in Europe. No. 49. Report of the Swedish Meteorological and Hydrological Inst., Norrkoping, Sweden.

Arheimer, B., and H.B. Wittgren. 1994. Modelling the effects of wetlands on regional nitrogen transport. Ambio 23:378–386.

Arnold, J., G.R. Srinivasan, R.S. Muttiah, and J.R. Williams. 1998. Large area hydrologic modelling and assessment. Part I: Model development. J. Am. Water Resour. Assoc. 34:73–89.

Barker, T., K. Hatton, M. O'Connor, L. Connor, and B. Moss. Effects of nitrate load on submerged plant biomass and species richness: Results of a mesocosm experiment. Freshwater Biol. (in press).

Bergström, S. 1992. The HBV model: Its structure and applications. No. 4. Report of the Swedish Meteorological and Hydrological Inst., Norrkoping, Sweden

Beven, K.J. 1993. Prophecy, reality, and uncertainty in distributed hydrological modelling. Adv. Water Resour. 16:41–51.

Boorman, D.B. 2003. Climate, hydrochemistry, and economics of surface-water systems (CHESS): Adding a European dimension to the catchment modelling experience developed under LOIS. Sci. Total Environ. 411:314–316.

Brady, N.C., and R.R. Weil. 1996. The nature and properties of soils. Prentice Hall, Upper Saddle River, NJ.

Butterfield, D., A.J. Wade, and P.G. Whitehead. 2006. INCA-N v1.9 user guide. Univ. of Reading, UK.

Cornell, S., N. Jackson, D. Hadley, K. Turner, and D. Burgess. 2004. Land-ocean interactions and climate change: Insights from the ELOISE projects. SEC-Tyndall HQ, UK.

Croisé, L., E. Ulriche, P. Duplat, and O. Jaquet. 2002. Renecofor: Deux approches indépendantes pour l'estimation et la cartographie des dépôts atmosphériques totaux hors couvert forestier sur le territoire français. Office Nationale des Forêts, Département Recherche et Développement, Paris, France.

Davies, H., and C. Neal. 2004. GIS-based methodologies for assessing nitrate, nitrite and ammonium distributions across a major UK basin, the Humber. Hydrol. Earth Syst. Sci. 8:823–833.

Durand, P. 2004. Simulating nitrogen budgets in complex farming systems using INCA: Calibration and scenario analyses for the Kervidy catchment (W. France). Hydrol. Earth Syst. Sci. 8:793–802.

Edwards, A.C., K. Pugh, G. Wright, A.H. Sinclair, and G.A. Reaves. 1990. Nitrate status of two major rivers in NE Scotland with respect to land use and fertiliser additions. Chem. Ecol. 4:97–107.

Emmett, B.A., B.J. Cosby, R.C. Ferrier, A. Jenkins, A. Tietema, and R.F. Wright. 1997. Modelling the ecosystem effects of nitrogen deposition: Simulation of nitrogen saturation in a Sitka spruce forest, Aber, Wales, UK. Biogeochemistry 38:129–148.

Etchanchu, D. 1998. Géochimie des eaux de la Garonne: Transferts de matières dissoutes et particulaires vers l'océan Atlantique. Ph.D. thesis. Université Paul Sabatier de Toulouse, France.

European Parliament. 2005. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Overall approach to the classification of ecological status and ecological potential. Office for Official Publications of the European Communities, Luxembourg, LU.

Ferrier, R.C. 1998. The DYNAMO project: An introduction: Dynamic models to predict and scale up the impact of environmental change on biogeochemical cycling. Hydrol. Earth Syst. Sci. 2:375–383.

Futter, M.N., D. Butterfield, B.J. Cosby, P.J. Dillon, A.J. Wade, and P.G. Whitehead. 2007. Modeling the mechanisms that control in-stream dissolved organic carbon dynamics in upland and forested catchments. Water Resour. Res. 43:W02424, doi:10.1029/2006WR004960.

Garnier, J., G. Billen, E. Hannon, S. Fonbonne, Y. Videnina, and M. Soulie. 2002. Modelling the transfer and retention of nutrients in the drainage network of the Danube River. Estuarine Coastal Shelf Sci. 54:285.

Gustard, A., D.C.W. Marshall, and M.F. Sutcliffe. 1987. Low flow estimation in Scotland. Institute of Hydrology Report 101. Inst. of Hydrology, Wallingford, UK.

Heal, O.W., and D.F. Perkins. 1978. Production ecology of British Moors and Montane Grasslands. Springer Verlag, Berlin, Germany.

Jackson, B.M., H.S. Wheater, A.J. Wade, D. Butterfield, S.A. Mathias,

A.M. Ireson, A.P. Butler, N. McIntyre, and P.G. Whitehead. 2007. Catchment-scale modelling of flow and nutrient transport in the Chalk unsaturated zone. Ecol. Modell. 209:41–52.

James, C., J. Fisher, V. Russell, S. Collings, and B. Moss. 2005. Nitrate availability and hydrophyte species richness in shallow lakes. Freshwater Biol. 50:1049–1063.

Kaste, Ø., and B.L. Skjelkvale. 2002. Nitrogen dynamics in runoff from two small heathland catchments representing opposite extremes with respect to climate and N deposition in Norway. Hydrol. Earth Syst. Sci. 6:351–362.

Kaste, Ø., R.F. Wright, L.J. Barkved, B. Bjerkeng, T. Engen-Skaugen, J. Magnusson, and N.R. Sælthun. 2006. Linked models to assess the impacts of climate change on nitrogen in a Norwegian river basin and fjord system. Sci. Total Environ. 365:3–14.

Kirchner, J.W. 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. Water Resour. Res. 42, W03S04.

Langan, S.J., A.J. Wade, R. Smart, A.C. Edwards, C. Soulsby, M.F. Billet, H.P. Jarvie, M.S. Cresser, R. Owen, and R.C. Ferrier. 1997. The prediction and management of water quality in a relative unpolluted major Scottish catchment: Current issues and experimental approaches. Sci. Total Environ. 194/195:419–435.

Legates, D.R., and G.J. McCabe, Jr. 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resour. Res. 35:233–241.

McIntyre, N., B. Jackson, A.J. Wade, D. Butterfield, and H.S. Wheater. 2005. Sensitivity analysis of a catchment-scale nitrogen model. J. Hydrol. 315:71–92.

Müller-Wohlfeil, D.I. 2002. Model frameworks for the calculation of annual runoff and nitrogen emissions from Danish catchments. National Environmental Research Inst., Silkeborg, Denmark.

Melillo, J.M. 1981. N-cycling in deciduous forests. *In* F.E. Clark and T. Rosswall (ed.) Terrestrial nitrogen cycles. Ecol. Bull. 33:427–442.

Miller, H.G. 1981. Nutrient cycles in forest plantations, their change with age, and the consequence for fertilizer practice. p. 187–200. In: Australian forest nutrition workshop: Productivity in perpetuity. Proceedings, Canberra, Australia.

Nash, J.E., and J.V. Sutcliffe. 1970. River flow forecasting through conceptual models: I. A discussion of principles. J. Hydrol. 10:282–290.

Neal, C., H.P. Jarvie, M. Neal, L. Hill, and H. Wickham. 2006. Nitrate concentrations in the river waters of the upper Thames and its tributaries. Sci. Total Environ. 365:15–32.

Neal, C., P.G. Whitehead, and N. Flynn. 2002. INCA: Summary and conclusions. Hydrol. Earth Syst. Sci. 6:607–616.

Powlson, D.S. 1993. Understanding the soil nitrogen cycle. Soil Use Manage. 9:86–94.

Probst, J.L. 1985. Nitrogen and phosphorus exportation in the Garonne basin (France). J. Hydrol. 76:281–305.

Rabaud, V., and M. Cesses. 2004. AGRESTE Chiffres et Données Agriculture: Enquête sur les pratiques culturales en 2001. Ministère de l'Agriculture et de la Pêche, Paris, France.

Rankinen, K., T. Karvonen, and D. Butterfield. 2006. An application of the GLUE methodology for estimating the parameters of the INCA-N model. Sci. Total Environ. 365:123–139.

Reynolds, B., E.J. Wilson, and B.A. Emmett. 1998. Evaluating critical loads of nutrient nitrogen and acidity for terrestrial systems using ecosystem-scale experiments (NITREX). For. Ecol. Manage. 101:81–94.

Rowell, D.L. 1994. Soil science: Methods and applications: Longman Scientific & Technical, Essex, England.

Ruiz, L., S. Abiven, P. Durand, C. Martin, F. Vertès, and V. Beaujouan. 2002. Effect on nitrate concentration in stream water of agricultural practices in small catchments in Brittany: I. Annual nitrogen budgets. Hydrol. Earth Syst. Sci. 6:497–506.

Ruz-Jerez, B.E., R.E. White, and P.R. Ball. 1994. Long-term measurement of denitrification in three contrasting pastures grazed by sheep. Soil Biol. Biochem. 26:29–39.

Sauvage, S., S. Teissier, P. Vervier, T. Améziane, F. Garabetian, F. Delmas, and B. Caussade. 2003. A numerical tool to integrate bio-physical diversity of a large regulated river: Hydro-biogeochemical bases; the case of the Garonne River (France). River Res. Applic. 19:181–198.

Skeffington, R., and E.J. Wilson. 1988. Excess nitrogen deposition: Issues for consideration. Environ. Pollut. 54:159–184.

Stevens, P.A., D.A. Norris, T.H. Sparks, and A.L. Hodgson. 1994. The impact of atmospheric inputs on throughfall soil and streamwater interactions for different aged forest and moorland catchments in Wales. Water Air Soil Pollut. 73:297–317.

Tietema, A. 1993. Mass loss and nitrogen dynamics in decomposing litter of five forest ecosystems in relation to increased nitrogen deposition. Biogeochemistry 20:45–62.

Van Liedekerke, M., F. Bouraoui, P. Panagos, and A. Nogueira. 2003. The EUROHARP data management system. Prog. Water Resour. 7:431–440.

Vuuren, M.M., E. Aerts, F. Berendse, and W. Visser. 1992. Nitrogen mineralization in heathland ecosystems dominated by different plant species. Biogeochemistry 16:151–166.

Wade, A.J., P. Durand, V. Beaujouan, W. Wessel, K.J. Raat, P.G. Whitehead, D. Butterfield, K. Rankinen, and A. Lepisto. 2002. Towards a generic nitrogen model of European ecosystems: INCA, new model structure and equations. Hydrol. Earth Syst. Sci. 6:559–582.

Wade, A.J., C. Neal, D. Butterfield, and M.N. Futter. 2004. Assessing nitrogen dynamics in Europeean ecosystems, integrating measurement and modeling: Conclusions. Hydrol. Earth Syst. Sci. 8:846–857.

Wade, A.J., C. Soulsby, S.J. Langan, P.G. Whitehead, A.C. Edwards, D. Butterfield, R.P. Smart, Y. Cook, and R.P. Owen. 2001. Modelling instream nitrogen variability in the Dee catchment, NE Scotland. Sci. Total Environ. 265:229–252.

Wasson, J.G., M.H. Tusseau-Vuillemin, V. Andréassian, C. Perrin, J.B. Faure, O. Barreteau, M. Bousquet, and B. Chastan. 2003. What kind of water models are needed for the implementation of the European Water Framework Directive? Examples from France. Int. J. River Basin Manage. 1:1–11.

Whitehead, P.G., E.J. Wilson, and D. Butterfield. 1998. A semi-distributed nitrogen model for multiple source assessments in catchments (INCA): Part 1-model structure and process equations. Sci. Total Environ. 210/211:547–558.

ARTICLE N°2

Statistical downscaling of river flow.

Tisseuil C., Vrac M., Wade AJ., Lek S. (2009)

En révision dans Journal of Hydrology

STATISTICAL DOWNSCALING OF RIVER FLOWS

Clement Tisseuil[a], Mathieu Vrac[b], Sovan Lek[a], Andrew J Wade[c]

[a]Université de Toulouse, UMR CNRS-UPS 5174, Evolution et Diversité Biologique (EDB), 118 route de Narbonne, 31062 Toulouse cedex 4 – France

[b]Laboratoire des Sciences du Climat et de l'Environnement (LSCE-IPSL) CNRS/CEA/UVSQ, Centre d'étude de Saclay, Orme des Merisiers, Bat. 701 91191 Gif-sur-Yvette, France

[c]Aquatic Environments Research Centre, School of Human and Environmental Sciences, University of Reading, RG6 6AB, UK

Abstract

An extensive statistical 'downscaling' study is done to relate large-scale climate information from a general circulation model (GCM) to local-scale river flows in SW France for 51 gauging stations ranging from nival (snow-dominated) to pluvial (rainfall-dominated) river systems. This study helps to select the appropriate statistical method at a given spatial and temporal scale to downscale hydrology for future climate change impact assessment of hydrological resources. The four proposed statistical downscaling models use large-scale predictors (derived from climate model outputs or reanalysis data) that characterize precipitation and evaporation processes in the hydrological cycle to estimate summary flow statistics. The four statistical models used are generalized linear (GLM) and additive (GAM) models, aggregated boosted trees (ABT) and multi-layer perceptron neural networks (ANN). These four models were each applied at two different spatial scales, namely at that of a single flow-gauging station (local downscaling) and that of a group of flow-gauging stations having the same hydrological behaviour (regional downscaling). For each statistical model and each spatial resolution, three temporal resolutions were considered, namely the daily mean flows, the summary statistics of fortnightly flows and a daily 'integrated approach'. The results show that flow sensitivity to atmospheric factors is significantly different between nival and pluvial hydrological systems which are mainly influenced, respectively, by shortwave solar radiations

and atmospheric temperature. The non-linear models (i.e. GAM, ABT and ANN) performed better than the linear GLM when simulating fortnightly flow percentiles. The aggregated boosted trees method showed higher and less variable R^2 values to downscale the hydrological variability in both nival and pluvial regimes. Based on GCM cnrm-cm3 and scenarios A2 and A1B, future relative changes of fortnightly median flows were projected based on the regional downscaling approach. The results suggest a global decrease of flow in both pluvial and nival regimes, especially in spring, summer and autumn, whatever the considered scenario. The discussion considers the performance of each statistical method for downscaling flow at different spatial and temporal scales as well as the relationship between atmospheric processes and flow variability.

Keywords: Hydrological regimes; evaporation; precipitation; generalized linear models; generalized additive models; boosted trees; neural networks.

1.Introduction

Climate change is expected to adversely impact water resources, water quality and the freshwater ecology and therefore methods are required to quantify the likely impacts to develop mitigation and adaptation strategies (Whitehead et al, 2009). Such quantification requires an ability to forecast river flow based on the projected changes in climate to assess changes in flow-pathways, pollutant source area, dilution and residence times, all of which affect the water quality and the aquatic ecosystem. Classically, future climate change is modelled under several hypothetical scenarios using General Circulation Models (GCM) which are mechanistic models built to physically represent the main atmospheric processes. However, GCM remain relatively coarse in resolution (approximately 2.5° x 2.5°, i.e. about 250 km x 250 km) and are unable to resolve sub-grid scale features such as topography, clouds and land use. This represents a considerable problem for the impact assessment of climate change on hydrological dynamics in river-systems. Thus, considerable efforts in the climate community have focused on the development of techniques, the so called ‘downscaling’ step, to bridge the gap between large- and local-scale climate data. To date, impact studies of climate change on hydrology involve a two-step approach: (i) GCM outputs are used to generate local climate conditions such as precipitation and temperature, which is known as ‘downscaling’, then (ii) these downscaled local climate data are used as input to a hydrological model to project the hydrological changes according to future climate. Fowler et

al. (2007) made a comparative review of downscaling models applied to hydrological studies, which are usually separated into either dynamical or statistical approaches. Dynamical downscaling is performed through Regional Climate Models (RCMs) which physically simulate the smaller-scale dynamical processes that control climate at the regional level down to 5 km x 5 km. GCM outputs are used to define the boundary conditions of Regional Climate Models. However, RCMs are computationally expensive in the production of the regional simulations. As such, it is currently possible to apply RCMs to limited periods and regions only.

This study relies on statistical downscaling models (SDMs). Based on observed data, SDMs define relationships between the large-scale variable fields, derived either from climate model outputs or observations, and local-scale surface conditions. The large-scale variable fields from General Circulation Models or reanalysis data (the predictors) are chosen such that they are strongly related to the local-scale conditions of interest (the predictands or response variable). The relationships can then be used to estimate changes in river-flow, or other local hydrological measures such as precipitation or air temperature, based on future projections from global or regional climate models. SDMs are generally separated into three types of approach which can be combined: regression models, weather typing schemes and weather generators (Vrac and Naveau, 2007a). Multiple linear models, in the regression-based approach are the most applied in downscaling, for example the well known SDSM tool (Wilby et al., 2002). These assume a linear relationship between large-scale atmospheric predictors and the response variable. However, several studies have shown that taking into account non-linearity between predictors and the predictand in statistical downscaling can improve the goodness-of-fit (Huth et al., 2008) including polynomial regression (Hewitson, 1994), recursive partitioning tree (Schnur and Lettenmaier, 1998), nearest neighbour (Zorita and von Storch, 1999), artificial neural networks (Harpham and Wilby, 2005; Khan et al., 2006) or generalized additive models (Vrac et al., 2007a; Salameh et al., 2009).

The two-step modelling framework, linking GCM outputs to a hydrological model, is usually constrained in space by the domain of calibration of the hydrological model. Furthermore the data requirement for setting the hydrological model parameters may be large, both for conceptual and fully distributed hydrological models (Arheimer and Wittgren, 1994; Eckhardt et al., 2005; Thompson et al., 2004; Habets et al., 2008). One possibility to increase the spatial extent of forecasting river flow at large spatial scales in response to climate change is to develop SDMs able to simulate instream flows directly from GCM atmospheric

variables. Seeking a direct association between river flows and GCM outputs may be relevant to facilitate the generalization and extrapolation of river flow simulations over large spatial scales. In the past, such a direct link has been criticized by some authors because of an over-simplification of the hydrological cycle through a lack of consideration of water stores and transfers within the soils and groundwater of a catchment (Xu, 1999), previous poor performances of SDMs linking directly GCM to flow (Wilby et al., 1999) or simply GCM outputs are deemed inappropriate as direct predictors of river flows (Prudhomme et al., 2002). Furthermore, the direct downscaling to streamflow from GCM atmospheric variables generally do not take into account other important factors affecting the streamflow variability such as the land use and soil cover, assuming deterministically that those factors do change with time.

However during the last decade, the relationship between GCM large-scale atmospheric variables and instream flows has been better described. Kingston et al. (2006) made a useful synthesis of recent integrated hydrological-climate research regarding the links between large-scale atmospheric circulation patterns (e.g., characterizing the North Atlantic Oscillation – NAO), regional climate and streamflow variations in the northern North Atlantic region over the last century and especially the last 50 years. Surprisingly, few studies have investigated such a link between atmospheric circulation patterns and flow in a purely predictive way, e.g. through downscaling applications. Examples include Cannon and Whitfield (2002) who applied an ensemble neural network downscaling approach to 21 watersheds in British Columbia; Ghosh and Mujumdar (2008) who simulated the streamflow of an Indian river for the monsoon period using a relevance vector machine; Landman et al. (2001) who downscaled the seasonal streamflow at the inlets of twelve dams in South Africa from predicted monthly-mean sea-surface temperature fields; Phillips et al. (2003) who used atmospheric circulation patterns and regional climate predictors to generate mean monthly flows in two British rivers; Déry and Wood (2004) who have shown that the recent variability in Hudson Bay river was significantly explained by the Arctic Oscillation over the last decades; Lawler et al. (2003) who investigated the influence of changes in atmospheric circulation and regional climate variability on river flows and suspended sediment fluxes in southern Iceland; and Ye et al. (2004) who used combinations of climate and atmospheric variables to explain from about 31% to 55% of the variance of the annual total discharges of three Siberian rivers.

In this study, various direct downscaling strategies linking flows to GCM outputs are investigated to estimate the flows measured at 51 hydrological gauging stations located in southwest France, representative of a transition from nival (snow-dominated) to pluvial (rainfall-dominated) hydrological conditions. Reanalysis data from the National Centers for Environmental Prediction and the National Center for Atmospheric Research. (NCEP/NCAR; Kalnay et al., 1996) are used as large-scale atmospheric predictors to calibrate the models and validate the approaches. The focus of this study will address the three following questions:

(1) Which spatial or temporal scale resolution and statistical methods could be the most relevant to downscale the streamflow variability from GCM outputs? As such, the statistical downscaling framework is built upon an extensive comparative approach which has three aspects (Fig. 1, Table 1). Four linear or non-linear statistical methods are applied at two different spatial scales, either to individual stations or regionally to a group of stations, according to three temporal resolutions varying from daily to fortnightly time resolutions.

(2) Can the relationship between climate processes and the hydrological variability be modelled by the downscaling framework according to different hydrological systems? As such, a wide set of NCEP/NCAR atmospheric variables are tested as potential predictors for flows and an extensive sensitivity analysis is performed to quantify the relationship between flows and atmospheric predictors according a range of hydrological regimes from nival to pluvial.

(3) As a synthesis of this work, is the proposed downscaling framework relevant for future climate change impacts studies? As an illustration, future seasonal changes in flows are projected and discussed according to nival and pluvial regimes over the region, using one GCM (cnrm-cm3) and two scenarios (A2, A1B).

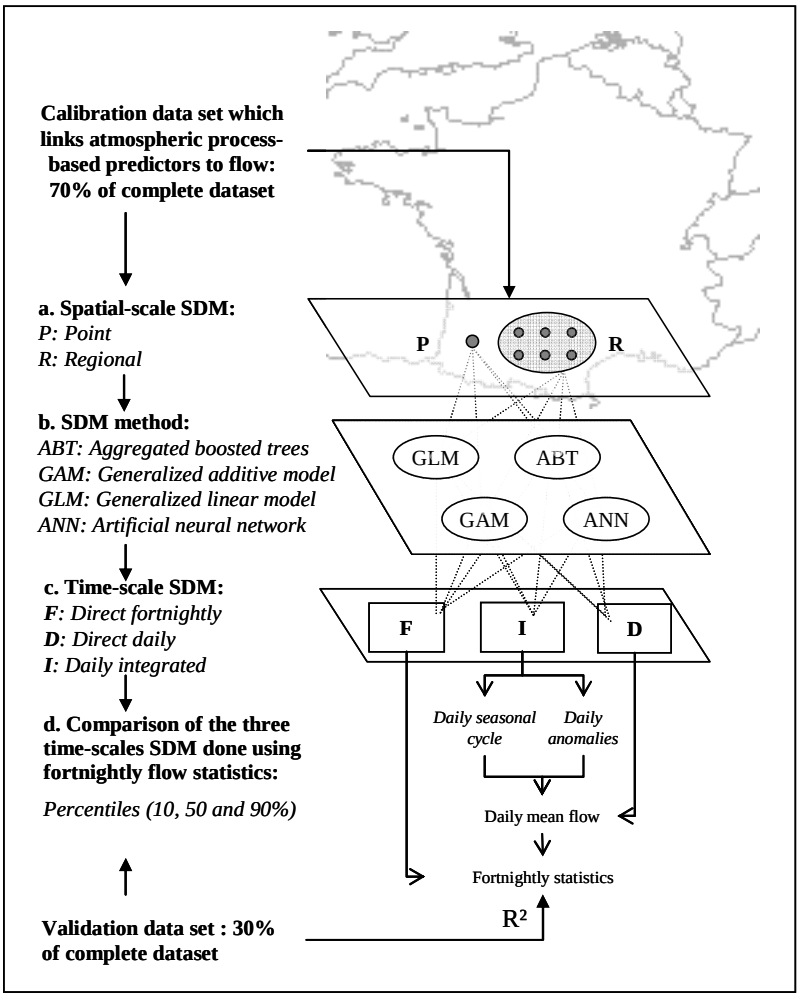


Fig. 1: Statistical downscaling framework. Four different statistical downscaling methods were calibrated using 70% of complete dataset which linked synthesized atmospheric predictors, derived from NCEP/NCAR reanalysis data, to observed flows summarised at different three time-scales and point and regional spatial scales. Testing was done using the remaining 30% of the dataset.

Table 1. Abbreviations

Full name	Abbreviation
Statistical downscaling model	SDM
Generalized linear model	GLM
Generalized additive model	GAM
Aggregated boosted tree	ABT
Artificial neural network	ANN
General circulation model	GCM
Classification and regression trees	CART
Hierarchical ascending clustering	HAC
Principal component analysis	PCA

2.Study area and data resource

Mean daily streamflow data for 51 stations located in the south west of France were obtained from the Hydro2 database maintained by the Ministère de l'Ecologie et du Développement Durable (http://www.hydro.eaufrance.fr/; Table 2; Fig.2). Three criteria were employed to determine the stations to be selected: (1) a continuous record spanning at least 15 years and starting after 1945; (2) inclusion of a large range of hydrological conditions over the region; (3) gauging stations close to water chemistry and biological sampling points and therefore of use to investigate the interactions between hydrology, water chemistry and/or biological communities in future studies. In general, the daily flow data from the 51 stations were available from 1968 to 1999.

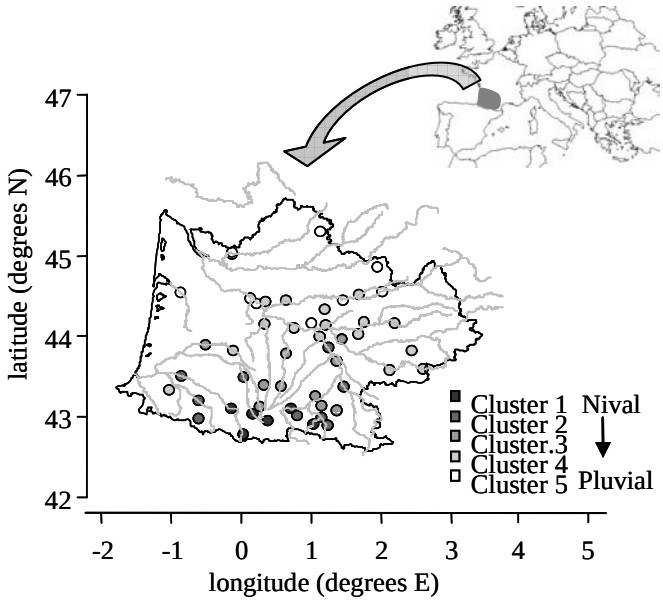


Fig. 2: The locations of the 51 hydrological gauging stations in the Adour-Garonne river-system (SW France). The grey-scale colours represent the hydrological transition from nival (cluster 1) to pluvial (cluster 5) hydrological regimes. Hydrological clusters were identified using HAC.

Table 2. Description of the 51 hydrological gauging stations located SW France with their hydrological regime scales from nival (1) to pluvial (5).

Station ID	Station name	Catchment area (km²)	Longitude (degrees E)	Latitude (degrees N)	Years	Hydrological regimes
O0174010	La Neste d'Aure à Sarrancolin	606	0.38	42.955	1961-1999	1
O0200020	La Garonne	2230	0.707	43.098	1984-1999	1
O0234020	Le Ger à Aspet	95	0.795	43.021	1983-1999	2
O0384010	L'Arac à Soulan	169	1.232	42.899	1962-1999	2
O0444010	Le Lez aux Bordes-sur-Lez	212	1.029	42.903	1971-1999	1
O0502520	Le Salat à Saint-Lizier	1154	1.141	42.991	1974-1999	2
O0624010	Le Volp à Montberaud	91	1.142	43.145	1968-1999	3
O0744030	L'Arize au Mas-d'Azil	218	1.361	43.083	1974-1999	3
O0964030	La Louge au Fousseret	272	1.06	43.267	1970-1999	3
O1712510	L'Ariège à Auterive	3450	1.467	43.369	1966-1999	2
O2034010	L'Aussonnelle à Seilh	192	1.356	43.692	1968-1999	3
O2620010	La Garonne à Verdun-sur-Garonne	13730	1.242	43.855	1972-1999	2
O2883310	La Gimone à Garganvillar	827	1.111	43.998	1965-1999	4
O4142510	L'Agout à Anglès	364	2.596	43.595	1972-1999	4
O4544020	Le Sor à Cambounet-sur-le-Sor	372	2.115	43.577	1977-1999	4
O4704030	Le Dadou à Paulinet	72	2.441	43.822	1968-1999	4
O4984320	Le Tescou à Saint-Nauphary	287	1.432	43.966	1974-1999	3
O5534010	Le Lézert à Saint-Julien-du-Puy	222	2.196	44.162	1968-1999	4
O5685010	La Bonnette à Saint-Antonin-Noble-Val	179	1.748	44.172	1968-1999	4
O5754020	La Vère à Bruniquel	311	1.673	44.024	1971-1999	4
O5964020	Le Lemboulas à Lafrançaise	403	1.203	44.137	1968-1999	4
O6125010	La Petite Barguelonne à Montcuq	62	1.191	44.334	1971-1999	4
O6134010	La Barguelonne à Valence	477	0.998	44.17	1968-1999	5
O6164310	L'Auroue à Caudecoste	196	0.756	44.107	1968-1999	4
O6212530	Le Gers à Panassac	159	0.568	43.383	1965-1999	3
O6312520	Le Gers à Montestruc-sur-Gers	678	0.64	43.791	1965-1999	3
O6692910	La Baïse à Nérac	1327	0.335	44.148	1965-1999	4
O6804630	L'Osse à Castex	10.2	0.324	43.399	1965-1999	3
O7971510	Le Lot à Faycelles	6840	2.016	44.557	1979-1999	5
O8133520	Le Célé à Orniac	1194	1.679	44.52	1971-1999	4
O8231510	Le Lot à Cahors	9170	1.446	44.449	1960-1999	5
O8584010	La Lède à Casseneuil	411	0.634	44.446	1970-1999	4
O9000010	La Garonne à Tonneins	51500	0.222	44.412	1989-1999	5
O9034010	Le Tolzac à Varès	255	0.353	44.433	1970-1999	4
O9134010	L'Avance à Montpouillan	405	0.137	44.464	1968-1999	5
P2054010	La Bave à Frayssinhes	183	1.948	44.858	1961-1999	5
P6342510	L'Auvézère à Cherveix-Cubas	586	1.127	45.298	1966-1999	5
P7261510	L'Isle à Abzac	3752	-0.126	45.022	1972-1999	5
P8462510	La Dronne à Coutras	2816	-0.132	45.042	1967-1999	5

Q0100010	L'Adour	272	0.164	43.037	1940-1999	1
Q0280030	L'Adour à Estirac	906	0.029	43.498	1968-1999	2
Q0522520	L'Arros à Gourgue	173	0.259	43.132	1968-1999	3
Q2062510	Le Midour à Laujuzan	256	-0.117	43.821	1966-1999	4
Q2192510	Le Midou à Mont-de-Marsan	800	-0.502	43.892	1967-1998	3
Q4124010	Le Gave d'Héas à Gèdre	84	0.022	42.787	1948-1995	1
Q4801010	Le Gave de Pau à Saint-Pé-de-Bigorre	1120	-0.143	43.103	1955-1999	1
Q5501010	Le Gave de Pau à Bérenx	2575	-0.853	43.509	1940-1999	2
Q6332510	Le Gave d'Aspe à Bedous	425	-0.604	42.981	1948-1999	2
Q7002910	Le Gave d'Oloron à Oloron-Sainte-Marie	1085	-0.608	43.199	1940-1999	2
Q8032510	La Bidouze à Aïcirits-Camou-Suhast	246	-1.028	43.334	1969-1999	4
S2242510	L'Eyre à Salles	1650	-0.872	44.548	1967-1999	5

NCEP/NCAR reanalysis data were used to model the river flows at the 51 gauging stations. NCEP/NCAR reanalysis data are atmospheric model outputs derived from the assimilation of surface observation stations, upper-air stations and satellite-observing platforms with long records starting in 1948 and continuing to present day. These data are typically viewed as ‘observed’ large-scale data on a regular grid with a spatial resolution of approximately 2.5° x 2.5° (250 km x 250 km). To improve the understanding between atmospheric conditions and flows, 27 atmospheric variables were tested here as potential explanatory variables. These variables included long wave and short wave radiation fluxes, cloud cover, land skin temperature, latent and sensible heat fluxes at surface. The full list is given in Table 3. As this study was built upon a climate change perspective, NCEP/NCAR variables were carefully selected as readily-available GCM outputs (available online at https://esg.llnl.gov:8443/index.jsp) so that these outputs could be used in further studies to generate the flow response to projected climate change. Each NCEP/NCAR variable was interpolated to each of the 51 hydrological stations locations using bilinear interpolation. For a given station, the interpolated data result from the weighted average of the data of the nearest points located on the regular grid. Then each interpolated NCEP/NCAR variable was normalized so that its mean was zero and its variance was 1.

Table 3. Description of the NCEP/NCAR reanalysis predictors used into the downscaling framework for river flow simulation, with their acronyms and correspondence with Global Circulation Models outputs.

NCEP names	NCEP short names	Pressure levels (hPa)	Units	Corresponding monthly GCM output	Corresponding daily GCM output
Mean daily air temperature	air	500, 850,1000	K	ta	ta
Mean daily convective precipitation rate at surface	cprat		kg m^{-2} s^{-1}	prc	
Mean daily clear sky downward longwave flux at surface	csdlf		W m^{-2}	rldscs	
Mean daily clear sky downward solar flux at surface	csdsf		W m^{-2}	rsdscs	
Mean daily clear sky upward solar flux at surface	csusf		W m^{-2}	rsuscs	
Mean daily downward longwave radiation flux at surface	dlwrf		W m^{-2}	rlds	rlds
Mean daily downward solar radiation flux at surface	dswrf		W m^{-2}	rsds	rsds
Mean daily geopotential height	hgt	500, 850, 1000	m	zg	zg
Mean daily upward longwave radiation flux at surface	ulwrf		W m^{-2}	rlus	rlus
Mean daily precipitation rate at surface	prate		kg m^{-2} s^{-1}	pr	pr
Mean daily surface pressure	pres		Pa	ps	ps
Mean daily relative humidity	rhum	500, 850, 1000	%	hur	
Mean daily upward solar radiation flux at surface	uswrf		W m^{-2}	rsus	rsus
Mean daily specific humidity	shum	500, 850, 1000	kg kg^{-1}	hus	hus
Mean daily SST/land skin Temperature	skt		K	ts	
Mean daily sea level pressure	slp		Pa	psl	psl
Mean daily total cloud cover	tcdc		%	clt	
Mean daily latent heat net flux at surface	lhtfl		W m^{-2}	hfls	hfls
Mean daily sensible heat net flux at surface	shtfl		W m^{-2}	hfss	hfls

3.Method

The statistical downscaling framework may be summarized in four steps (Fig. 1, Table 1). At step 1, information from the 27 NCEP/NCAR variables was first synthesised into five process-based predictors to be more readily interpreted, namely precipitation, temperature, pressure, radiation and heat flux (see Section 3.1; Fig.3). At step 2, these process-based predictors were used in the statistical downscaling framework (SDM; see Section 3.2, Fig. 1) to simulate river flow according to two spatial resolutions, namely at a single flow-gauging station or a group of flow-gauging stations having the same hydrological behaviour (Fig. 1a).

For each spatial resolution, four statistical models (Fig. 1b) including generalized linear models (GLM), generalized additive models (GAM), aggregated boosted trees (ABT) and artificial neural networks (ANN) were each applied to three temporal resolutions, namely daily mean flow, fortnightly-derived flows statistics (percentiles 10, 50 and 90%) and a daily integrated approach (Fig.1c). This daily integrated approach separates the daily flow downscaling process into the downscaling of the daily seasonal cycle, which is defined as the mean flow for each day of the year over the calibration period, and the downscaling of the corresponding daily anomalies which are the values resulting from the subtraction of the daily seasonal cycle from the daily flow data. Performances of the different SDMs are compared between observed and downscaled flow statistics calculated at the fortnightly time scale for each station (Fig. 1d). At step 3, a sensitivity analysis was performed based on the regional downscaling approach to quantify and describe the relationship between river flow and the five process-based atmospheric variables, according to the hydrological regions and the five statistical methods used (see Section 3.3). At step 4, future relative changes of seasonal flow were projected to assess the potential impact of climate change on nival and pluvial systems according to different time periods and future scenarios (see Section 3.4).

3.1.Deriving process-based NCEP/NCAR predictors

The approach was based on a regional, process-based representation of atmospheric variables, which aimed at synthesizing the initial 27 NCEP/NCAR atmospheric variables into a limited number of moderately correlated, physically meaningful, predictors for the downscaling of flows (Fig. 3). With such a representation, correlations between predictors were reduced, so that their relationship with the flow variability could be quantified with more robustness than if using the 27 highly correlated NCEP/NCAR predictors directly. In practice, co-linearity would not impact the performances of the downscaling process; however, the individual contribution of predictors to the flow variance explained, as well as the coefficients estimates in downscaling models, could change erratically. Furthermore, limiting the number of atmospheric predictors reduces the computation time for downscaling models. The method to derive the process-based factors is based on two steps:

(1) A hierarchical ascending cluster analysis (HAC) with Ward criterion was applied to the Euclidean distance matrix of the 27 normalized mean monthly NCEP/NCAR atmospheric variables (Ward, 1963). HAC has been applied in several climate studies, such as Vrac et al. (2007b) who categorized the regional climate conditions in the state of Illinois, USA, in terms

of circulation and precipitation atmospheric patterns. By applying HAC in our study, the atmospheric variables which have the most similar "behaviours" have been grouped together within five homogeneous clusters related to precipitation, temperature, pressure, shortwave radiation and heat flux processes (Fig. 3a). The relevance of selecting five clusters was assessed using the silhouette information (SI) calculated for each variable, ranging from 0 to 1 for badly to perfectly clustered variables (Rousseeuw, 1987). In this study, the 27 variables were correctly placed within the five clusters (SI >0.5). Furthermore, the five clusters represented physically meaningful information on well identified atmospheric processes.

(2) A principal component analysis (PCA) was applied to each of the five groups of variables to derive a physically meaningful and synthetic description of the given process. The first PC of each group, containing more than 80% of the total variance, was retained as predictor into the downscaling (Fig. 3b). The pairwise Pearson correlation between the first PC of each group was ensured not to exceed 0.7.

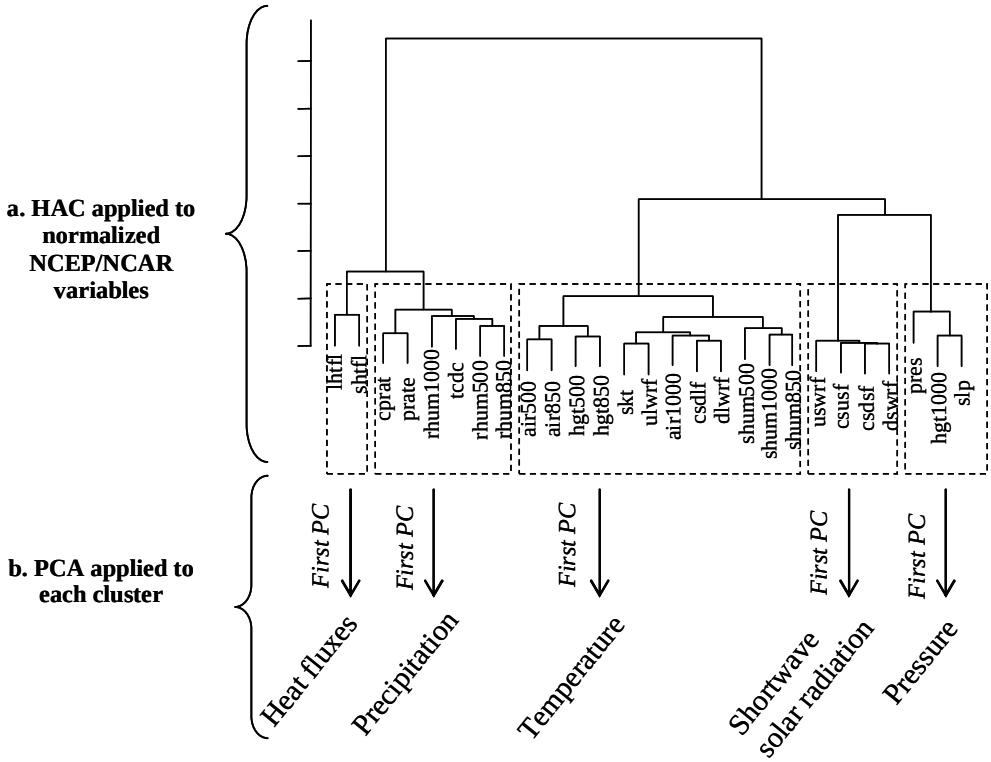


Fig. 3. Atmospheric predictors, namely heat flux, precipitation, temperature, shortwave solar radiation and pressure fields were derived from the 27 normalized NCEP/NCAR atmospheric variables. The atmospheric variables were first clustered (**a**) using hierarchical ascending analysis with Ward criterion (HAC), then process-based predictors were synthesized into the first component of a principal component analysis (PCA) applied to each cluster (**b**). The 27 variables are explained in Table 3.

3.2.Statistical downscaling framework

Prior to the downscaling process, flow data were first standardized per station. For a given station, the annual mean flow was subtracted from the time series of daily flows and the result divided by the standard deviation of the daily flow time series. This was done to make the dimension of flow values comparable between stations. The standardised data were then transformed using box-cox power transformations to make the shape of the distribution as Gaussian as possible, so that the GLM and GAM assumption of normality was valid (Box and Cox, 1964). The whole analysis was made with the R statistical software and supporting routines that have been compiled into the DS package for R, available on request.

3.2.1.Point (P) and regional (R) downscaling

Point downscaling refers to the calibration of a statistical model to each of the 51 gauging stations. Regional downscaling, in this study, refers to the calibration of a statistical model to a group of gauging stations representative of a hydrological regime. These regimes were previously identified via HAC method with Ward criterion to group the 51 gauging stations into 5 homogeneous and well identified hydrological regimes ranging from nival to pluvial (Fig. 2). The 5 selected clusters were assumed to be the optimal number of clusters for the present analysis in comparison to a larger or smaller number of clusters. Thus all the stations from the same hydrological regime have the same calibrated model. HAC was applied to the Euclidean distance matrix of stations based on their standardized monthly flow percentiles (10, 50 and 90%). Note that HAC was performed based on monthly flow percentiles only, and not other basin characteristics.

3.2.2.Daily (D) vs. fortnightly (F) direct downscaling vs. daily integrated downscaling (I)

The comparative downscaling framework includes three different time scale strategies (Fig. 1c). In this study, SDM aims at relating directly the daily mean (D) and fortnightly mean (F) atmospheric predictors, respectively to the daily mean flow and fortnightly flow statistics which were the fortnightly percentiles 10, 50, and 90%. Such indices have been applied in downscaling context to improve percentiles estimates, especially extremes (Dibike and Coulibaly, 2006). The fortnightly scale was preferred to monthly scale to increase the number of sampling units and improve the statistical inference.

The daily 'integrated' SDM (I) was based on two separate downscaling steps from the initial daily time series of flow. Firstly, the downscaling of the daily seasonal cycle was done; secondly, the downscaling of the corresponding daily anomalies. Finally, the downscaled daily seasonal cycle and anomalies are summed afterwards to complete the daily integrated approach. As such, downscaling the seasonal cycle aims at modelling the flow seasonality while downscaling the anomalies aims at modelling the variation around the daily seasonal cycle. A review of the literature suggests that such an approach has not been tried previously.

3.2.3.Statistical models

For each of the six SDM spatial/temporal combinations examining point and regional down-scaling at each of the daily, fortnightly and 'integrated' timescales, GLM, GAM, ABT and ANN statistical methods were also compared for each of the six combinations (Fig. 1b).

3.2.3.1 Generalized Linear and Generalized Additive Models

Generalized Linear Models (GLM) are a flexible generalization of ordinary least squares regression, unifying various other statistical models, including linear, logistic and Poisson regression under one framework (McCullagh, 1984). In GLM, each outcome of the response variable *Y* (i.e., flow) is assumed to be generated from a particular distribution function in the exponential family that includes the normal, binomial and Poisson distributions. Flow data were assumed to be normally distributed after box-cox transformation. The mean of the distribution, μ, depends on the predictor variables *X*, namely the NCEP/NCAR predictors. The model was defined as:

$$g(E(Y \mid X)) = \beta X + \alpha \qquad (1)$$

where $E(Y|X)$ is the expected value of Y conditionally on X; β and α corresponds respectively to a vector of unknown parameters to be estimated and the intercept; g is the function relating the predictors to the flow variable. The g function is called the "link" function and can take many shapes (determined by the user) in order to make applicable the right parts of Eqs. (1). Indeed, according to the distribution family of Y, the link function g has to be changed. In the present study, the flow variability to downscale are assumed to be Gaussian distributed and then E(Y|X) is directly related the right parts of Eqs. (1) (see Hastie and Tibshirani, 1990 for technical and theoretical details). Hence, g is taken as the identity function.

Generalized Additive Models (GAM) have been developed for extending properties of GLM to non-linear relationships between *X* and *Y* through additive properties (Hastie and Tibshirani, 1990). GAM fits the conditional expectation of *Y* given *X*, as the sum of *m* spline functions f_i of some or all of the covariates (Wood, 2008), where *m* is the dimension of X:

$$g(E(Y \mid X)) = \sum_{i=1}^{m} f_i(x_i) + \theta_0 \quad (2)$$

As for GLM, GAM specifies a distribution for the response variable. The functions f_i can be parametric or non-parametric, thus providing the potential for non-linear fits to the data which GLM does not allow. In this study, the spline functions, f_i, are defined as natural cubic splines, namely splines constructed of piecewise third-order polynomials with continuity conditions expressed until second derivatives (Hastie and Tibshirani, 1990). θ_0 is a constant to be estimated and *g* was defined as the identity function.

3.2.3.2 Feedforward artificial neural network

A multi-layer perceptron feedforward artificial neural network (ANN) was used in this study. This type of neural network is extremely flexible and has been applied to a wide variety of hydrological and climate situations (Reed and Marks, 1998). In this study the artificial neural network was trained using a back-propagation algorithm (Rumelhart et al.,1986). The architecture of the neural network used was three layers of neurons: the input layer, the hidden layer and the output layer. Every neuron of a layer was connected with every neuron of the previous layer by weight links that were modified during successive iterations. The value of the output from each neuron was calculated using the tanh sigmoid transfer function [$f(x) = 1/(1+e^{-x})$]. The backpropagation algorithm adjusted the connection weights according to the back propagated error computed between the observed and the estimated results. This is a supervised training procedure that attempts to minimize the error between the desired and the predicted output (Lek and Guégan, 2000). The output Y from the neural network was given by:

$$Y = \sum_j \tanh\left(\sum_i x_i w_{i,j}^1 + b_j^1\right) w_j^2 + b^2 \quad (3)$$

where x_i represents the i^{th} input predictors, $w_{i,j}^1$ and w_i^2 are the hidden input and output layer weights, and $b_{i,j}^1$ and b^2 the hidden input and output layer biases. Here, j = 4 internal

nodes were chosen for the single-hidden layer by comparing the downscaling performances with a different number of nodes whose range was defined using the empirical formula (Huang and Foo, 2002):

$$\sqrt{2i} + o < j < 2i + 1 \quad (4)$$

Where i is the number of input nodes corresponding to the number of atmospheric predictors (i.e. in our case the five process-based predictors), o is the number of output nodes (i.e. in this study $o = 1$).

3.2.3.3 Aggregated boosted regression trees (ABT)

There was no evidence in the readily-accessible literature that boosted trees have been used in downscaling studies. Friedman et al. (2000) and Hastie et al. (2001) introduced the technique for use in applied statistics, especially in ecological applications. Boosted trees are based on a compilation of classification and regression tree (CART) models. CART models explain variation of a single response variable by repeatedly splitting the data into more homogeneous groups, using combinations of explanatory variables that may be categorical and/or numeric. Each group is characterized by a typical value of the response variable, the number of observations in the group and the values of the explanatory variables that define it (De'ath and Fabricius 2000).

The aim of boosted trees is to improve the performance of a single CART model by fitting m models, in our case 1000 models, where each successive CART is built for the prediction residuals of the preceding tree (Elith et al., 2008). Considering a loss function that represents the loss in predictive performance (e.g. deviance explained) between two models, boosting is a numerical optimization technique that minimizes the value of the loss function by adding, at each new step, a new CART that best reduces the loss function (Elith et al. 2008). To limit the over-fitting of the boosted trees caused by the construction of too many CART models, each new CART is grown on a randomized subset of the dataset. Then, the optimal number of trees is automatically selected, after the 1000 generated CART in our study, so that that the loss in predictive performance calculated on the remaining subset of the dataset was minimized (De'Ath, 2007).

Aggregated boosted trees (ABT) are themselves an extension of boosted trees. Aggregated boosted trees comprise a collection of boosted trees generated on a cross-validation subset, which reduce the prediction error relative to a single boosted tree (De'ath, 2007).

3.2.4.Validation and evaluation of model performances

The same validation procedure was applied to all downscaling schemes. Observations were chosen from the whole sample to form the training dataset (the first 70% of each time series), and the remaining observations (i.e. corresponding to the last 30% of each time series) were retained as the validation dataset (Fig. 1d). Hence, validation and training datasets are temporally independent. For the comparison between the different spatial and time scales downscaling models, performance was evaluated using the coefficient of determination, R^2, calculated by station for each fortnightly statistics (percentiles 10, 50 and 90% of flow) between observations (O) and simulations (S) from year i to n, through:

$$R^2 = \left\{ \frac{\sum_{i=1}^{n}(O_i - \bar{O})(S_i - \bar{S})}{\left[\sum_{i=1}^{n}(O_i - \bar{O})^2\right]^{0.5}\left[\sum_{i=1}^{n}(S_i - \bar{S})^2\right]^{0.5}} \right\}^2 \qquad (5)$$

R^2 values range from 0 (poor model) to 1 (perfect model). Statistical downscaling models with R^2 values above 0.5 will be interpreted here as good models, showing that 50% of the flow variability is explained by the atmospheric predictors (Fig. 1d).

3.3.Sensitivity of downscaled flows to atmospheric predictors

Based on the regional fortnightly downscaling approach, a sensitivity analysis was performed to quantify the contribution from each of the five process-based predictors to the explained variance of the river flow, according to the different hydrological regions and the four statistical methods, namely GLM, GAM, ANN and ABT. Since the core from the four statistical methods is based upon different algorithm, the sensitivity approach developed here to quantify the influence of predictors to the flow variability was specific to each statistical method. However, to make comparable the results between the four statistical methods, the percentage contribution of each predictor to the flow variance explained (i.e. R^2) is scaled so that the sum adds to 100, with higher numbers indicating stronger contribution to the response (Elith et al. 2008).

3.3.1.Sensitivity measure

For GAM and GLM, the sensitivity of flow variability to the atmospheric predictors was estimated via the Fisher-Snedecor statistic, F, calculated for each of the predictors. Typically in GLM and GAM framework, the F statistic is the ratio of the explained variability by a given predictor (as calculated by the R^2 coefficient of determination) and the unexplained variability (as calculated by $1-R^2$), divided by the corresponding degree of freedom (Lomax, 2007). Thus, the larger the F statistic, the more important is the predictor to flow variance explained.

For ANN, the influence predictor to the flow variability was evaluated via the method of partial derivates (Dimopoulos et al., 1995; Gevrey et al., 2003). With the method of partial derivates, the sum of square derivatives value was obtained per input variable and allowed a classification of the input variables according to their increasing contribution to the output variable (i.e. river flows) in the model. The input variable with the highest sum of square derivatives value was the variable most influencing the output variable.

For ABT, the flow sensitivity to each atmospheric predictor was assessed using the method described by Friedman (2001). The contribution of predictors is based on the number of times a predictor is selected for splitting during the boosting process, weighted by the squared improvement (i.e. the loss in predictive performance) to the model as a result of each of those splits, and averaged over all models.

3.3.2.Multivariate Analysis of Variance

Each downscaling model was performed 500 times using flow datasets of size 500 ($m = 500$), randomly drawn from the training dataset and representing approximately 25% of it. A Multivariate Analysis of Variance (Manova) was applied to test if the relative contribution of the five atmospheric predictors ($a=5$) was significantly different between each statistical model ($s=4$) and between each hydrological regime ($h=5$). Manova is a direct extension of anova where the two tested variables of interest are not tested on a single continuous variable but on the distance matrix. Here, the Euclidean distance matrix was calculated from the $i \times a$ matrix of predictors contribution, where $i = m \times s \times h$.

3.4.Future projections

Based on the regional fortnightly downscaling approach, future projections of median flow conditions were performed to illustrate the ability of using the downscaling framework

for future climate change impact studies. The future projections were based on the GCM cnrm-cm3 from Meteo-France according to two scenarios from the IPCC (Pachauri and Reisinger, 2007), namely scenarios A2 and B1. Three time periods, namely 2025-2050, 2050-2075 and 2075-2100 were investigated and the relative changes of flow (RC) were calculated seasonally for each station to highlight the contrasted changes between nival and pluvial regimes according to the two scenarios. RC was calculated as difference between future projected and observed (1970-2000) flow condition, divided by the observed condition. For example, a relative change of +0.20 indicates a future flow increase of 20%. The future flow projections were made in three steps:

The GCM atmospheric variables for the two future scenarios were standardized according to their control period, i.e. under the scenario '20c3m' which represents a simulation of the GCM over 1970-2000 based on historical trends. This was done to remove the potential bias in the mean and the standard deviation of GCM atmospheric variables over the period 1970-2000.

As many hydrological change impact studies (e.g. Hay et al. 2000), the delta method was applied to each of the 21 atmospheric variables by adding the change in climate to an observational database to represent the future climate. More specifically for a given station and a given month, the delta method was calculated as the mean difference between the observations, i.e. the averaged NCEP/NCAR conditions over 1970-2000, and the averaged GCM projections over a given future time period. Then the observations and the estimated mean difference were summed afterwards to recombine a future fortnightly times series of atmospheric variables.

The future fortnightly times series of the 21 atmospheric variables were then projected onto the first principal component axis from their respective group of atmospheric variables (See Section 3.1) to derive the four atmospheric predictors for the downscaling.

4.Results

The Hierarchical Ascending Cluster analysis applied to our 51 stations produced five hydrological regimes, ranging from nival to pluvial systems (Fig. 2). The nival regime characterizes stations mostly located in the headwaters of the Pyrenees (six stations) with the annual peak of flows generally occurring during the spring snowmelt. Conversely, the pluvial regime characterizes lowland stations (10 stations), influenced by heavy winter rainfall in the Massif Central leading to maximum annual flows in winter. Transitional nival to pluvial

regimes are observed for intermediate stations collecting water both from Pyrenees and Central Massif (Fig. 2). The seasonal Pearson correlations between the observed flows and the corresponding process-based predictors, namely precipitation, temperature, solar radiations, heat fluxes and pressure PC (Fig. 3), show some seasonal correlations according to nival or pluvial regimes (Fig. 4). Temperature and shortwave solar radiations correlated with observed flows show the largest seasonal variability in the correlations (Fig. 4a, b). While the correlation between flow and temperature is globally negative in summer and autumn as well as weak in winter for both nival and pluvial regimes, the temperature in spring correlates flow negatively in pluvial systems and positively in nival ones (Fig. 4a). The seasonal correlation of flow with the shortwave solar radiations exhibits the same trends than those observed with the temperature, excepted in summer where the correlation between flow and the shortwave radiations remains positive for both pluvial and nival regimes (Fig. 4.b). The correlation between precipitation and flows is globally positive throughout the year, approximately R=0.4 (Fig. 4c). Heat fluxes and pressure predictors do not show strong seasonal correlations with flows, although flow correlation to the pressure PC averaged -0.2, nor major differences between nival and pluvial regimes (Fig. 4d, e).

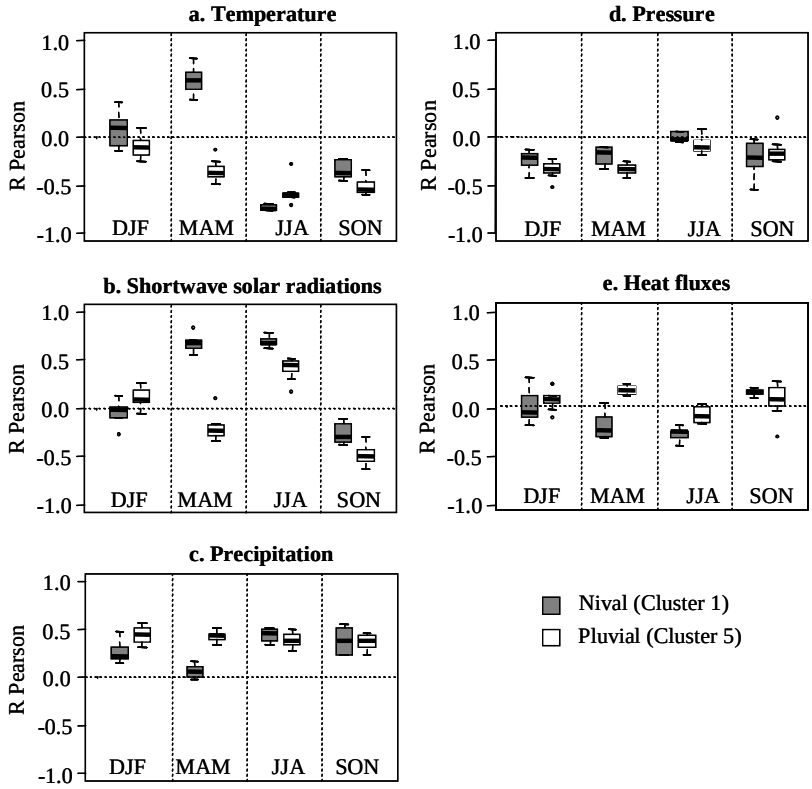


Fig. 4. Seasonal Pearson R correlation coefficients between flow and the five derived atmospheric predictors, as described in Fig.3, according to nival (dark grey) and and pluvial (white) systems: **(a)** temperature, **(b)** shortwave solar radiation, **(c)** precipitation, **(d)** pressure and (e) heat fluxes.

The mean percentage contribution and standard deviation from the five process-based predictors to the flow variance explained was estimated for each statistical model (i.e. GLM, GAM, ABT, ANN) and per hydrological regime using the daily regional downscaling from 500 samples (Fig. 5). The Manova results show that the contribution of the atmospheric predictors was significantly different between the four statistical methods (Manova, $p<0.001$) and between the five hydrological regimes (Manova, $p<0.001$). Nival regimes are mainly driven by solar radiation fluxes whereas temperature is the key-process involved in pluvial regimes (Fig. 5a). Aggregated boosted trees seem to be more stable than other methods since the percentage of contribution calculated for each predictor show less variability then the one estimated from the GLM, GAM and ANN, as shown by the smaller amplitude in the boxplot (Fig. 5b). GLM, GAM and ANN emphasise the importance of temperature and solar radiation principal components to explain the flow variance. However, temperature and solar radiation remain the two most important factors for both statistical models (Fig. 5b).

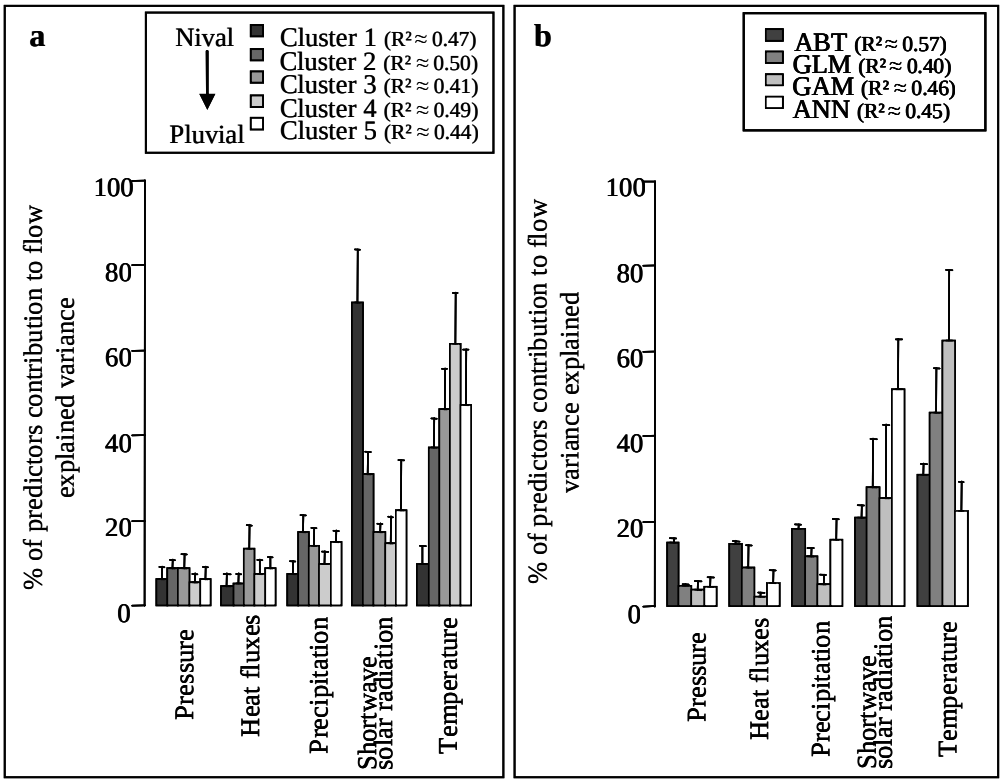


Fig. 5. Results of the sensitivity analysis showing the percentage contribution of the five atmospheric predictors to the explained flow variability, according to hydrological regimes (**a**) from nival (black) to pluvial (white); statistical downscaling models (**b**) ABT= aggregated boosted trees, GAM=generalized additive model, GLM=generalized linear model, ANN=artificial neural network.

Model performances (i.e. R^2 calculated between fortnightly observed and simulated flow statistics) were compared according to each spatial/time scale combination, as well as according to the four statistical models and the five hydrological regimes. The results are presented in Fig. 6. Mean R^2 performances for aggregated boosted trees (ABT) are significantly better than those of the GLM (paired t-test, $p<0.001$), GAM (paired t-test, $p<0.001$) and ANN (paired t-test, $p<0.001$), while GLM shows significantly lower performances (Fig. 6a; $R^2_{ABT}=0.49$, $R^2_{GAM}= 0.44$, $R^2_{ANN}=0.44$, $R^2_{GLM}=0.40$. When averaging results from all methods, a slight decrease in high flow percentiles estimates is observed ($R^2_{p10}=0.48$, $R^2_{p50}= 0.47$, $R^2_{p90}=0.41$). Overall, fortnightly downscaling (F) slightly outperforms daily downscaling (D) and daily downscaling with integrated seasonal cycle and anomalies (I) ($R^2_F=0.47$, $R^2_D=0.43$, $R^2_I=0.43$). Additional results from the daily integrated downscaling (not presented here) show its good performance in downscaling the seasonal cycle, but its lack of efficiency to simulate the daily anomalies. Point downscaling performs significantly better than the regional one as $R^2_{Point} =0.51$ and $R^2_{Regional}=0.46$ (paired t-test, $p <0.001$) and it is significantly better for modelling high fortnightly flow percentiles (paired t-test, $p <0.001$). The mean performance of downscaling models is lower in nival ($R^2_{cluster\ 1}=0.41$) than in pluvial ($R^2_{cluster\ 5}=0.45$) regimes, especially for high flow percentiles estimates (Fig. 6c; unpaired t-test, $p <001$). Globally for the three percentiles, fortnightly flows is better simulated by the downscaling models in summer ($R^2_{JJA} = 0.28$) than in winter ($R^2_{DJF} = 0.11$), spring ($R^2_{MAM} = 0.16$) and autumn ($R^2_{SON} = 0.19$) (Fig. 6d).

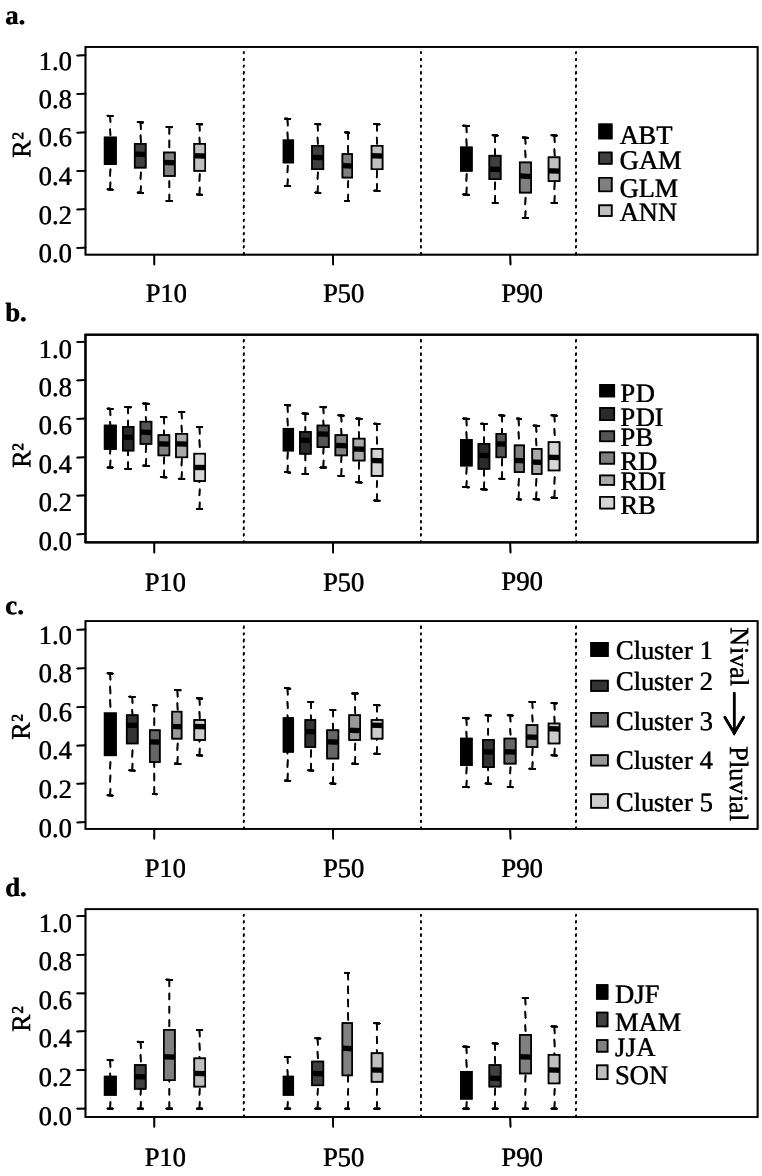


Fig. 6. Boxplot representing the variability in the performance of statistical downscaling models, as the variance explained (R^2) in modelling three fortnightly percentiles of river flow, namely percentiles 10% (P10), 10% (P50) and 90% (P90). Comparison is made between: **(a)** statistical downscaling models (ABT=aggregated boosted trees; GAM=generalized additive model; GLM=generalized linear model; ANN=artificial neural network); **(b)** downscaling approaches (PD=point daily downscaling; PI= point daily downscaling with integrated season and anomalies; PF=point fortnightly downscaling; RD=regional daily downscaling; RI= regional daily downscaling with integrated season and anomalies; RF=regional fortnightly downscaling); **(c)** hydrological regimes ranging from nival (cluster 1; dark grey) to pluvial (cluster 5; white); **(d)** seasons, namely winter (DJF), spring (MAM), summer (JJA) and autumn (SON).

Future projections in median flow conditions were performed based on the regional bimonthly downscaling approach and the ABT statistical method, according to two scenarios and analysed for three periods, namely 2025-2050, 2050-2075 and 2075-2100 (Fig. 7). Globally, the median flow conditions decrease in both nival (-17%) and pluvial (-15%) systems (Fig. 7a, b). In nival systems (Fig. 7a), this decrease is more particularly severe in

spring (RC_{MAM} =-40%) and autumn (RC_{SON} =-24%) than in winter (RC_{DJF} = -7%) and summer (RC_{JJA} =-7%). The future relative change of flows in nival systems is not significantly different between the A2 and A1B scenarios (paired t-test, p=0.32) nor between the different periods (one-way Anova, p=0.45). In pluvial systems (Fig. 7b), flows could globally increase in winter (RC_{DJF} = +20%) while decreasing during the other seasons (RC_{MAM} = -30%, RC_{JJA} = -32% and RC_{SON} = -26%). The relative changes in pluvial systems are relatively the same according to the A2 and A1B scenarios, excepted in spring where flows decrease dramatically under the A2 scenario (Fig. 7b; RC_{MAM} = -50%). Globally for both scenarios, the relative changes of flows in pluvial systems are significantly different between the three periods in winter only (Fig. 7b; one-way Anova; p<001).

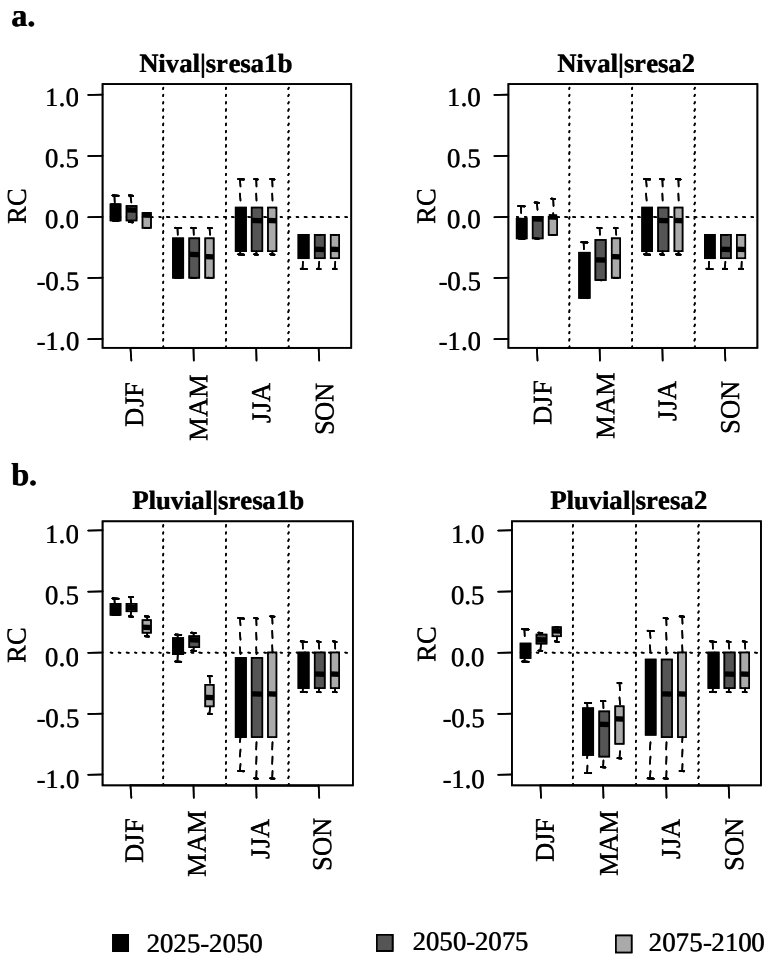


Fig. 7. Future relative changes (RC) in seasonal flow conditions projected for nival **(a)** and pluvial **(b)** regimes. Relative changes are highlighted for three periods, namely 2025-2050 (black), 2050-2075 (grey) and 2075-2100 (light grey) according to scenarios A2 and A1B.

5.Discussion

The discussion will address the three main questions mentioned in the Introduction. Firstly, the technical aspects related to the different downscaling strategies will be discussed to highlight their main strengths and limits as well as some possibilities of improvements. Secondly, the reliability of the downscaling framework will be discussed in regards to the physical meaning linking atmospheric factors to streamflow variability according to nival and pluvial hydrological systems. Thirdly, future flow projections in nival and pluvial systems will be analysed to illustrate the applicability of the downscaling framework for future climate change impact studies.

5.1.Comparison between the different statistical downscaling strategies

In this study, a direct statistical downscaling approach from GCM to streamflow variability was experimented, which is less commonly applied than the approach involving an intermediate hydrological model between GCM and streamflow to reproduce the hydrological cycle (Fowler et al., 2007). While a direct downscaling approach may allow the assessment of the relationship between flow and atmospheric process over large spatial scales more easily than if using an intermediate hydrological model, some limits should be considered. Particularly, the direct downscaling approach was developed from a deterministic point of view by assuming that the variability of streamflow was influenced by climate factors only. Thus the direct downscaling approach developed in this study do not explicitly take into account for some physical factors, such as the land use and soil cover, which interact with climate and influence flow pathways (e.g.. interception, infiltration and groundwater processes) and may vary under future climate. In this context, using a hydrological model that classically integrates those physical factors within a delimited structure of the river catchment (e.g. HBV; Lindstrom et al., 1997) may provide a more realistic projection of the potential future hydrological conditions than the use of a direct downscaling approach. However, by comparing different statistical downscaling approaches according to different spatial/time scale strategies and statistical models, our study has revealed three key results encouraging further developments for the use of direct statistical downscaling approaches to assess the potential impact of climate change on hydrological resources.

Firstly, the downscaling performances using the regional approach did not deteriorate too much the quality of the projected fortnightly statistics in comparison to the point downscaling. This makes the regional approach very attractive from a technical point of view as well as for

the understanding of large scale hydro-climatic processes. Technically, the regional approach is 10 times faster to compute than the local one, approximately 30 minutes on a regular computer to calibrate the four statistical methods. Furthermore, the regional approach has shown to summarize satisfactorily the key relationship between climate and streamflow variability according to the different hydrological systems ranging from nival to pluvial. These two specificities make the regional approach of particular interest to extend feasibly the downscaling framework of streamflow across Europe. Finally, a few additional features could be added to the regional downscaling approach to improve the regional flow projections from GCM outputs, such as integrating the land cover, geology and soil covers to better identify hydrological region. The spatial autocorrelation between hydrological sites could be also integrated into a statistical downscaling framework of streamflow, which has never been done to our knowledge, for example to help projecting the flow variability from atmospheric process to ungauged hydrological stations.

Secondly, the fortnightly flow percentiles downscaling recorded better performances than daily and daily integrated downscaling, especially for high flow percentiles. These results are in agreement with studies focused on the downscaling of extreme climate events which highlighted good performances when downscaling seasonal extreme indices derived from daily climate data (Moberg and Jones, 2005, Hanson et al., 2007). The daily direct and daily integrated downscaling simulations were shown to reproduce accurately the daily flow seasonal cycle across the study area but failed to simulate the magnitude of high flow events. The difficulty for those two daily approaches to simulate high flow events may come from the statistical inability of models to relate high flow events to climate processes. Since high flow events may result from local climate processes and controlling processes such as localised convective precipitation or orographically, enhanced precipitation and thus the simulation of extreme floods from large-scale atmospheric conditions may not be satisfactorily simulated. The proposed downscaling of daily anomalies suffered the incapacity of the models to take into account for the seasonal variation in the relationship between the daily mean atmospheric processes and the daily anomalies. Thus, the downscaling of daily anomalies could be improved by possibly adding a seasonal signal (e.g. sin and cosin values related to the different months) to relate the daily mean atmospheric processes to the daily anomalies at a given season; or, even more simply, by conditioning the downscaling model per season. Moreover, the downscaling of daily anomalies could also take advantage of the extreme value theory (Coles, 2001; Katz et al., 2002; Vrac and Naveau, 2007) to improve high percentiles

simulations. Recent studies have also characterized drought and floods at the daily time scale in relation to circulation patterns using fuzzy coding (e.g., Bardossy et al., 1995; Samaniego and Bardossy, 2007).

Secondly, the non-linear statistical models such as aggregated boosted tree, generalized additive models and artificial neural networks performed better than the generalized linear models to project the hydrological variability from atmospheric processes. Some similar results have been highlighted by Cannon and Whitfield (2002) and Ghosh and Mujumdar (2008) who respectively applied an ensemble of neural networks and support vectors machines to forecast streamflow from atmospheric processes. Although all three non-linear statistical models performed comparably in our study, the best performance was obtained for the aggregated boosted trees models. To our knowledge, this study is the first application of the aggregated boosted tree method for climate downscaling studies. However, earlier studies from Elith et al. (2008) and De'ath (2007) in ecology confirmed the relatively higher predictive power of boosted trees than that of other statistical methods. Anyway, since none statistical method may definitely assumed to be the best one, especially for climate change impact studies, it would worth to take into the uncertainty in downscaling projections from different statistical methods.

5.2. Relationship between atmospheric factors and streamflow variability

The hydrological response in catchments results from the complex interactions between hydro-climatic conditions, for example rainfall intensity and duration and the condition of soil moisture preceding a rainfall event, and the physical characteristics of the catchment, namely the land cover, the morphology of the river network and the soil characteristics. The hydrological cycle may be viewed as a balance between the evaporation and precipitation processes which drive the dynamics of water and the active flow pathways regulating the soil moisture, the infiltration, groundwater recharge and surface runoff (Sun and Pinker, 2004; Li et al., 2007).

Atmospheric processes are generally related to river flows through atmospheric weather regimes (Kingston et al., 2006). Atmospheric weather regimes characterize the large spatial scale structure of a given atmospheric variable, often geopotential height, sea level pressure or specific humidity at different atmospheric levels, which are then used to relate flow dynamics. This was done by Kingston et al. (2006) in Britain; Stewart et al. (2005) and Molnár and Ramírez (2001) in north-western New Mexico; Anctil and Coulibaly (2004) and Déry and

Wood (2004) in Canada; Krepper et al. (2003) in Uruguay; Lawler et al. (2003) in south-west Iceland; Struglia et al. (2004) across the Mediterranean region; and Ye et al. (2004) in Siberia.

In this study, a simplified representation of the relationship between atmospheric fields and flow generation was developed throughout five synthetic regional atmospheric factors derived from clustering and principal component analysis. Those five factors were related to precipitation, pressure, temperature, shortwave solar radiation and heat flux and they may show different or combined effect on the hydrological cycle. For example, evaporation mainly depends on the energy available in the system (e.g. heat fluxes, temperature, shortwave radiations) as well as the capacity of the air to store water (e.g. the pressure of water saturation in the air influence the air relative humidity). Similarly, precipitation results from a change in temperature and/or pressure, conditioned by a sufficient air relative humidity (Hufty, 2001). The sensitivity analysis of flow to those five atmospheric predictors revealed that pluvial and nival systems were mostly driven by temperature and shortwave solar radiation, i.e. by evaporation processes, more than by precipitation. Such results are in agreement with those of Phillips et al. (2003), who highlighted the main influence of regional temperature on flow in two pluvial rivers in southern Britain. Furthermore, the influence of temperature and shortwave radiation on streamflow variability showed some differences between nival and pluvial regimes.

In pluvial regimes, precipitation tends to fall as rain all year and the air temperature is negatively correlated to flow all year. That is, an increase in air temperature tends to actively increase the evaporation process and reduce the soil moisture, as shown by the negative correlation between the mean air temperature and flow in summer. In winter, the evaporation is reduced while the frequency and the intensity of precipitation increases, which leads to a saturation of the soil and higher groundwater levels. Thus, rainfall in winter is likely to contribute directly to a rising flow when the catchment is saturated, as shown by the positive correlation between the mean precipitation and flow in pluvial catchments.

Conversely in nival catchments, winter precipitation are generally stored as snow until spring, which do not contribute to soil moisture saturation and do not consequently lead a rising flow, as confirmed by a very weak are correlation between mean precipitation and flow in winter (Fig. 4c). From spring, the rising shortwave solar radiations and temperature triggers snowmelt and typically generates a flow increase in nival systems, which may continue until summer. Shortwave solar radiations remain positively correlated to flow in spring and summer (Fig. 4b), possibly indicating a stronger control than temperature on the snowmelt

process, as confirmed by some recent studies on snowmelt runoff modelling (Li and Williams, 2008).

Globally, high flows were less well simulated in nival systems than in pluvial ones. Although, the shortwave radiation and temperature were shown to be important processes for trigging the snowmelt from spring to summer, the prediction of high flows from snowmelt remains very difficult. This may be due to an inability for the downscaling models to capture the subtleties of snow-pack accumulation over the winter, ripening and melt.

5.3.Future hydrological projections in nival and pluvial systems

The suitability of the downscaling framework for future climate change impact studies was illustrated using a single statistical method, namely the aggregative boosted trees, and the regional approach to highlight how the nival and pluvial systems may respond to future climate change over the region. The interpretation of these future projections should be considered carefully since only one GCM model was used to characterize the future climate. Furthermore, the relevance of hydrological projections could be also criticized by the delta method used to derive the future atmospheric predictors for the regional downscaling. A major disadvantage of the delta approach is that representation of extremes from future climate scenarios effectively gets filtered out in the transfer process. The extremes resulting from this approach are simply the extremes from present climate observations that have either been enhanced or dampened according to the delta factors (Graham et al., 2007).

Globally, streamflow could decrease in both nival and pluvial systems over the region of study. In nival systems, the decrease of flow could be particularly important in spring while the precipitation and temperature increases could lead to the snow cover storage reduction and to an earlier melt (Caballero et al., 2007). In pluvial regimes, the rising precipitation in winter could be related to the dramatic increase of streamflow in winter. These results are in agreement with Caballero et al. (2007) who assessed the potential future changes of flows based on the mechanistic hydrological model, SAFRAN-ISBA-MODCOU (SIM), applied to the Adour Garonne basin. However, recent applications of the SIM models over the same region highlighted a global diminution of precipitation all over the year leading to likely the same global diminution of flows all over the year (Boé et al., 2009).

6.Conclusion and wider perspective

To our knowledge, this study is one of the first one to compare extensively a number of statistical downscaling approach to project the hydrological variability directly from GCM atmospheric processes for a wide range of hydrological conditions. A first important result showed the ability of the downscaling modelling framework to highlight the contrasted dynamics of streamflow variability in nival and pluvial systems in response to key atmospheric processes. The results also emphasised the particular interest of using a regional approach to downscale directly the hydrological variability from GCM, for three reasons at least: (i) the capacity to capture the key relationship between the atmospheric and hydrological variability within each hydrological system; (ii) the possibility to extend feasibly the downscaling approach to higher spatial scales such as Europe; (iii) the possibility to improve the approach by taking into account for the spatial autocorrelation between sites or adding physical information to better help identifying hydrological regions or projecting hydrological changes at ungauged sites. This study was also the first application of the aggregated boosted trees method in statistical downscaling studies of hydro-climatology. That is, the aggregated boosted trees appeared to be the most efficient and stable method for modelling river flows in this case study, in comparison to others methods such as generalized linear models, generalized additive models and neural networks.

The main objective of this study was essentially to build and validate a downscaling framework of river flow directly from GCM outputs, to be used for future climate change impact studies. Thus, results from the projected future changes in the hydrology between nival and pluvial regimes were preliminary; however they were sufficiently encouraging to further development in the downscaling of river flow. For example, an ensemble method could be developed to downscale seasonal forecasts or future hydrological changes in different hydrological systems, by using several GCM, downscaling methods and different scenarios. Although this type of ensemble procedure has already been applied in several future hydrological studies based on an hydrological model to make the connection between downscaled climate conditions to river streamflow (Graham et al., 2007; Boé et al., 2009; Hangemann et al., 2009; Kay et al., 2009; Tapiador et al., 2009), to our knowledge it has never been applied to direct statistical downscaling framework of river flow from GCM. Further investigations are also under progress to build an integrated model chain linking the directly downscaled hydro-climatic conditions from GCM to some ecological models e.g. to

project the potential impact of future hydro-climatic changes on the river ecosystem, from the nutrient loads to the structure of hydro-biological organisms.

7.Acknowledgements

This work was done as part of the EU FP6 Integrated Project “EURO-LIMPACS” (GOCE-CT-2003-505540) and GIS-REGYNA project. We thank J-L Le Rohellec (DIREN) for providing flow data; and National Oceanic and Atmospheric Administration (NOAA) for NCEP/NCAR reanalysis. We also thank the two anonymous referees for their insightful comments on an earlier draft of this manuscript.

8.References

Anctil, F. and Coulibaly, P., 2004. Wavelet Analysis of the Interannual Variability in Southern Québec Streamflow. Journal of Climate, 17(1): 163-173.

Arheimer, B. and Wittgren, H.B., 1994. Modeling The Effects Of Wetlands On Regional Nitrogen Transport. Ambio, 23(6): 378-386.

Bardossy, A., Duckstein, L. and Bogardi, I., 1995. Fuzzy Rule-Based Classification Of Atmospheric Circulation Patterns. International Journal Of Climatology, 15(10): 1087-1097.

Boe, J., Terray, L., Martin, E. and Habets, F., 2009. Projected changes in components of the hydrological cycle in French river basins during the 21st century. Water Resources Research, 45.

Box, G.E.P. and Cox, D.R., 1964. An analysis of transformations. Journal of the Royal Statistical Society. Series B (Methodological): 211-252.

Caballero, Y., Voirin-Morel, S., Habets, F., Noilhan, J., LeMoigne, P., Lehenaff, A. and Boone, A., 2007. Hydrological sensitivity of the Adour-Garonne river basin to climate change. Water Resources Research, 43(7).

Cannon, A.J. and Whitfield, P.H., 2002. Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models. Journal Of Hydrology, 259(1-4): 136-151.

Coles, S. (2001), An Introduction to Statistical Modeling of Extreme Values, Springer, London.

De'ath, G. and Fabricius, K.E., 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology, 81(11): 3178-3192.

De'ath, G., 2007. Boosted trees for ecological modeling and prediction. Ecology, 88(1): 243-251.

Déry, S.J. and Wood, E.F., 2004. Teleconnection between the Arctic Oscillation and Hudson Bay river discharge. Geophys. Res. Lett, 31: 18.

Dibike, Y.B. and Coulibaly, P., 2006. Temporal neural networks for downscaling climate variability and extremes. Neural Networks, 19(2): 135-144.

Dimopoulos, Y., Bourret, P. and Lek, S., 1995. Use of some sensitivity criteria for choosing networks with good generalization ability. Neural Processing Letters, 2(6): 1.

Hagemann, S., Göttel, H., Jacob, D., Lorenz, P. and Roeckner, E., 2009. Improved regional scale processes reflected in projected hydrological changes over large European catchments. Climate Dynamics, 32(6): 767.

Hay, L.E., Wilby, R.J.L. and Leavesley, G.H., 2000. A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States. Journal Of The American Water Resources Association, 36(2): 387-397.

Hufty, A., 2001. Introduction à la climatologie. De Boeck University, Bruxelles, 541 pp

Eckhardt, K., Fohrer, N. and Frede, H.G., 2005. Automatic model calibration. Hydrological Processes, 19(3): 651-658.

Elith, J., Leathwick, J.R. and Hastie, T., 2008. A working guide to boosted regression trees. Journal Of Animal Ecology, 77(4): 802-813.

Fowler, H.J., Blenkinsop, S. and Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. International Journal Of Climatology, 27(12): 1547-1578.

Friedman, J., Hastie, T. and Tibshirani, R., 2000. Additive logistic regression: a statistical view of boosting. Ann. Statist, 28(2): 337-407.

Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Statist, 29(5): 1189-1232.

Gevrey, M., Dimopoulos, L. and Lek, S., 2003. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecological Modelling, 160(3): 249-264.

Ghosh, S. and Mujumdar, P.P., 2008. Statistical downscaling of GCM simulations to streamflow using relevance vector machine. Advances in Water Resources, 31(1): 132.

Graham, L.P., Hagemann, S., Jaun, S. and Beniston, M., 2007. On interpreting hydrological change from regional climate models. Climatic Change, 81: 97-122.

Habets, F., Boone, A., Champeaux, J.L., Etchevers, P., Franchisteguy, L., Leblois, E., Ledoux, E., Le Moigne, P., Martin, E., Morel, S., Noilhan, J., Segui, P.Q., Rousset-Regimbeau, F. and Viennot, P., 2008. The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France. Journal Of Geophysical Research-Atmospheres, 113(D6).

Hanson, C.E., Palutikof, J.P., Livermore, M.T.J., Barring, L., Bindi, M., Corte-Real, J., Durao, R., Giannakopoulos, C., Good, P., Holt, T., Kundzewicz, Z., Leckebusch, G.C., Moriondo, M., Radziejewski, M., Santos, J., Schlyter, P., Schwarb, M., Stjernquist, I. and Ulbrich, U., 2007. Modelling the impact of climate extremes: an overview of the MICE project. Climatic Change, 81: 163-177.

Harpham, C. and Wilby, R.L., 2005. Multi-site downscaling of heavy daily precipitation occurrence and amounts. Journal Of Hydrology, 312(1-4): 235-255.

Hastie, T.J. and Tibshirani, R.J., 1990. Generalized Additive Models, volume 43 of Monographs on Statistics and Applied Probability. Chapman & Hall, 9: 41.

Hastie, T.J., Tibshirani, R.J. and Friedman, J.H., 2001. The Elements of Statistical Learning. Springer-Verlag, New York.

Hewitson, B., 1994. Regional Climates in the GISS General Circulation Model: Surface Air Temperature. Journal of Climate, 7(2): 283-303.

Huang, W. and Foo, S., 2002. Neural network modeling of salinity variation in Apalachicola River. Water Research, 36(1): 356-362.

Huth, R., Kliegrova, S. and Metelka, L., 2008. Non-linearity in statistical downscaling: does it bring an improvement for daily temperature in Europe? International Journal Of Climatology, 28(4): 465-477.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W.Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, Roy Jenne, Dennis Joseph (1996). "The NCEP/NCAR 40-Year Reanalysis Project". Bulletin of the American Meteorological Society 77 (3): 437–471.

Katz, R.W., Parlange, M.B. and Naveau, P., 2002. Statistics of extremes in hydrology. Advances In Water Resources, 25(8-12): 1287-1304.

Kay, A.L., Davies, H.N., Bell, V.A. and Jones, R.G., 2009. Comparison of uncertainty sources for climate change impacts: flood frequency in England. Climatic Change, 92(1-2): 41-63.

Khan, M.S., Coulibaly, P. and Dibike, Y., 2006. Uncertainty analysis of statistical downscaling methods. Journal Of Hydrology, 319(1-4): 357-382.

Kingston, D.G., Lawler, D.M. and McGregor, G.R., 2006. Linkages between atmospheric circulation, climate and streamflow in the northern North Atlantic: research prospects. Progress in Physical Geography, 30(2): 143-174.

Krepper, C.M., Garcia, N.O. and Jones, P.D., 2003. Interannual variability in the Uruguay river basin. International Journal of Climatology, 23(1): 103-115.

Landman, W.A., Mason, S.J., Tyson, P.D. and Tennant, W.J., 2001. Statistical downscaling of GCM simulations to streamflow. Journal Of Hydrology, 252(1-4): 221-236.

Lawler, D.M., McGregor, G.R. and Phillips, I.D., 2003. Influence of atmospheric circulation changes and regional climate variability on river flow and suspended sediment fluxes in southern Iceland. Hydrological Processes, 17(16): 3195-3223.

Lek, S. and Guégan, J.F., 2000. Artificial neuronal networks: application to ecology and evolution. Springer Berlin.

Li, H., Robock, A. and Wild, M., 2007. Evaluation of Intergovernmental Panel on Climate Change Fourth Assessment soil moisture simulations for the second half of the twentieth century. J. Geophys. Res, 112.

Li, X.G. and Williams, M.W., 2008. Snowmelt runoff modelling in an arid mountain watershed, Tarim Basin, China. Hydrological Processes, 22(19): 3931-3940.

Lindstrom, G., Johansson, B., Persson, M., Gardelin, M. and Bergstrom, S., 1997. Development and test of the distributed HBV-96 hydrological model. Journal Of Hydrology, 201(1-4): 272-288.

Lomax, R.G., 2007. An Introduction to Statistical Concepts. Lawrence Erlbaum.

McCullagh, P., 1984. Generalized Linear-Models. European Journal Of Operational Research, 16(3): 285-292.

Moberg, A. and Jones, P.D., 2005. Trends in Indices for Extremes in Daily Temperature and Precipitation in Central and Western Europe, 1901-99. International Journal of Climatology, 25(9): 1149-1171.

Molnár, P. and Ramírez, J.A., 2001. Recent Trends in Precipitation and Streamflow in the Rio Puerco Basin. Journal of Climate, 14(10): 2317-2328.

Pachauri, R.K. and Reisinger, A., 2007. Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland.

Phillips, I.D., McGregor, G.R., Wilson, C.J., Bower, D. and Hannah, D.M., 2003. Regional climate and atmospheric circulation controls on the discharge of two British rivers, 1974–97. Theoretical and Applied Climatology, 76(3): 141-164.

Prudhomme, C., Reynard, N. and Crooks, S., 2002. Downscaling of global climate models for flood frequency analysis: where are we now? Hydrological Processes, 16(6): 1137-1150.

Reed, R.D. and Marks, R.J., 1998. Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks. MIT Press Cambridge, MA, USA.

Rousseeuw, P.J., 1987. Silhouettes - A Graphical Aid To The Interpretation And Validation Of Cluster-Analysis. Journal Of Computational And Applied Mathematics, 20: 53-65.

Rumelhart, D.E., Hintont, G.E. and Williams, R.J., 1986. Learning representations by back-propagating errors. Nature, 323(6088): 533-536.

Salameh, T., Drobinski, P., Vrac, M. and Naveau, P., 2009. Statistical downscaling of near-surface wind over complex terrain in southern france. Meteorology and Atmospheric Physics, 103(1): 253-265.

Samaniego, L. and Bardossy, A., 2007. Relating macroclimatic circulation patterns with characteristics of floods and droughts at the mesoscale. Journal Of Hydrology, 335(1-2): 109-123.

Schnur, R. and Lettenmaier, D.P., 1998. A case study of statistical downscaling in Australia using weather classification by recursive partitioning. Journal Of Hydrology, 213(1-4): 362-379.

Stewart, I.T., Cayan, D.R. and Dettinger, M.D., 2005. Changes toward Earlier Streamflow Timing across Western North America. Journal of Climate, 18(8): 1136-1155.

Struglia, M.V., Mariotti, A. and Filograsso, A., 2004. River Discharge into the Mediterranean Sea: Climatology and Aspects of the Observed Variability. Journal of Climate, 17(24): 4740-4751.

Sun, D. and Pinker, R.T., 2004. Case study of soil moisture effect on land surface temperature retrieval. Geoscience and Remote Sensing Letters, IEEE, 1(2): 127-130.

Tapiador, F.J., Sanchez, E. and Romera, R., 2009. Exploiting an ensemble of regional climate models to provide robust estimates of projected changes in monthly temperature and precipitation probability distribution functions. Tellus Series A-Dynamic Meteorology And Oceanography, 61(1): 57-71.

Thompson, J.R., Sørenson, H.R., Gavin, H. and Refsgaard, A., 2004. Application of the coupled MIKE SHE/MIKE 11 modelling system to a lowland wet grassland in southeast England. Journal of Hydrology, 293(1-4): 151.

Vrac, M. and Naveau, P., 2007. Stochastic downscaling of precipitation: From dry events to heavy rainfalls. Water Resources Research, 43(7).

Vrac, M., Marbaix, P., Paillard, D. and Naveau, P., 2007a. Non-linear statistical downscaling of present and LGM precipitation and temperatures over Europe, Clim. Past, 3: 669-682.

Vrac, M., Stein, M. and Hayhoe, K., 2007b. Statistical downscaling of precipitation through nonhomogeneous stochastic weather typing. Climate Research, 34(3): 169-184.

Ward J., 1963. Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236–244.

Whitehead, P.G., Wilby, R.L., Battarbee, R.W., Kernan, M. and Wade, A.J., 2009. A review of the potential impacts of climate change on surface water quality. Hydrological Sciences Journal, 54(1): 101-123.

Wilby, R.L., Hay, L.E. and Leavesley, G.H., 1999. A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado. Journal of Hydrology, 225(1-2): 67.

Wilby, R.L., Dawson, C.W. and Barrow, E.M., 2002. sdsm-a decision support tool for the assessment of regional climate change impacts. Environmental Modelling and Software, 17(2): 145-157.

Wood, S.N., 2008. Fast stable direct fitting and smoothness selection for generalized additive models. Journal Of The Royal Statistical Society Series B-Statistical Methodology, 70: 495-518.

Xu, C.Y., 1999. From GCM to river flow: a review of downscaling methods and hydrologic modelling approaches. Progress in Physical Geography, 23(2): 229-249.

Ye, H., Yang, D., Zhang, T., Zhang, X., Ladochy, S. and Ellison, M., 2004. The Impact of Climatic Conditions on Seasonal River Discharges in Siberia. Journal of Hydrometeorology, 5(2): 286-295.

Zorita, E. and von Storch, H., 1999. The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. Journal Of Climate, 12(8): 2474-2489.

ARTICLE N°3

Validating a hydro-ecological model to project fish community structure from general circulation models using downscaling techniques.

Tisseuil C., Vrac M, Wade AJ, Grenouillet G, Gevrey M, Lek S

En préparation

VALIDATING A HYDRO-ECOLOGICAL MODEL TO PROJECT FISH COMMUNITY STRUCTURE FROM GENERAL CIRCULATION MODELS USING DOWNSCALING TECHNIQUES

Tisseuil C.[a], Vrac M.[b], Wade A.J.[c], Grenouillet G.[a], Gevrey M.[a], Lek S.[a]

[a]Laboratoire Evolution et Diversité Biologique (EDB) UMR 5174, CNRS - Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 4 – France

[b]Laboratoire des Sciences du Climat et de l'Environnement (LSCE-IPSL) CNRS/CEA/UVSQ, Centre d'étude de Saclay, Orme des Merisiers, Bat. 701 91191 Gif-sur-Yvette, France

[c]Aquatic Environments Research Centre, School of Human and Environmental Sciences, University of Reading, RG6 6AB, UK

Abstract

To understand how projected climate change will impact the freshwater ecology it is important to determine the inter-relationships between climate and hydrology and the response of the aquatic ecology to changes in habitat and food-web structure. This understanding is required to develop informed management plans regarding the use of water resources whilst protecting the ecological services of surface waters. As part of this research effort, a hydro-climatic-ecological (HCE) model-chain was developed for south-west France to test hypotheses regarding how the climate controls fish communities through invoked changes in the regional hydrology and temperature. The hydro-climatic modelling was calibrated using the reanalysis data from the National Centre for Environmental Prediction and the National Centre for Atmospheric Research (NCEP/NCAR) and five general circulation models (GCMs) were selected to project the hydro-climatic conditions under the control period i.e. from 1970 to 2000. The downscaled outputs of the GCMs showed good overall ability to model the observed seasonal hydrological and temperature variability. Coupled to fish-specific distribution models, the downscaled hydro-climatic projections were able to represent satisfactorily the observed occurrence for the 13 most prevalent fish species

over the region. The HCE model was validated based on historical data, confirming its suitability for future climate change impact studies. It is envisaged that this work will form the basis for the quantification of how fish community structures will change under future climate projections.

Keywords:

GCM, hydrology, boosted tree, hydro-ecology, regional climate, stream fish.

Introduction

Modelling the impact of climate change on freshwater ecosystems is a major challenge for scientists worldwide. The most recent report of the Intergovernmental Panel of Climate Change (IPCC; Pachauri & Reisinger, 2007) provides evidence that the on-going climate change modelled by GCMs will affect natural ecosystems across the world. Specifically, in terms of water quality and freshwater ecology, air temperature increases could accelerate the acidification of streams and negatively affect the recovery process of acidified lakes increase levels of nutrients entering the river system and alter the annual hydrological cycle (Schindler, 1997; Whitehead et al., 2009). In terms of freshwater biodiversity, declining river flow rates are shown to be a major cause of species loss through the impact on breeding seasons for fish and on post-spawning recruitment (Jackson, 1989; Humphries & Lake, 2000; Postel & Richter, 2003). Global warming is expected to shift cold-water species towards higher latitudes and altitudes by exceeding temperature preferences and tolerance limits (Rahel et al., 1996; O' Brien et al., 2000; Reid et al., 2001; Hari et al., 2006). Such temperature increases could have a ruinous effect on species presently found in mountainous headwaters or in high-altitude lakes which would not be able to migrate whilst species in the downstream sections of rivers are expected to expand their range of distribution (Buisson et al., 2008).

Modelling the local freshwater biological community response to global change usually requires consideration of multiple spatial and temporal scales (Heino et al., 2009). Such communities are not solely a product of local environmental filters such as hydrological variability (Cattanéo et al., 2002) and water chemistry, but they also have imprints of factors associated with larger spatial and temporal scales, such as mean annual air temperature and topography (Poff et al., 1997; Heino et al., 2009). From an evolutionary perspective, the pattern of spatial and temporal variations in habitat influences the relative success of a species in a particular environmental setting (Poff et al., 1997). While several climate change impact studies on freshwater biodiversity have emerged in the last decade, most of them have

focused on determining the future thermal habitat suitability for biological communities, for example for fish in North America (Minns & Moore, 1995; Eaton & Sheller, 1996; Magnusson et al., 1997; Jackson & Mandrak, 2002; Mohseni et al., 2003; Chu et al., 2005; Sharma et al., 2007) and in Europe (Buisson et al., 2008; Lassalle et al., 2009). However, to our knowledge few studies have explored the future hydrological habitat suitability for biological communities (Xenopoulos et al., 2005), and even more rare are studies that integrating both future hydrological and thermal habitat suitability (Matulla et al., 2007). An explanation for this knowledge gap is the difficulty of applying hydrological models to multiple sites across a region, such as the Garonne in south-west France which covers approximately 60,000 km2, where many tributaries may be ungauged and due to the heterogeneity of catchment soils, geology, vegetation and local climate conditions there is uncertainty regarding the transfer of modelled outcomes from gauged to ungauged sub-catchments. Lane (2008) suggests that when modelling the impact of hydrological changes on an organism there is a need to know what matters most in terms of the hydrology. Furthermore, linking the outputs from a hydrological model to a model-based representation of the ecology remains a major challenge, in particular the specification of the thresholds in the ecological response to changes in flow remains a key research topic (Poff et al., 1996; Cattanéo et al., 2002; Cattanéo et al., 2005).

This aim of this study is to address the identified knowledge gap through the development of a hydro-climatic-ecological (HCE) model-chain to help understand how climate change will affect the hydrological variability across a region and the subsequent consequences for fish biodiversity. The development of the HCE model-chain represents one of the first models to make a link between climate, hydrology and the freshwater ecological response at the regional (116,000 km^2) scale. The study has three main objectives: (1) to select a reliable dataset of large-scale atmospheric fields from climate re-analysis and the ouputs of General Circulation Models to use as input to the HCE model; (2) to calibrate independently the statistical hydro-climatic downscaling and fish-specific distribution models; (3) to compare five GCMs outputs to project the hydro-climatic variability under the control period, specified as the last 30 years of observed records where available, and (4) to assess the HCE model goodness-of-fit to model the historical hydro-climatic variability and fish species distribution over the region. It is envisaged that this model development will allow the subsequent exploration of the impacts of climate change on the distribution of freshwater fish communities across large (> 100,000 km^2) regions.

Materials and methods

Study area

The study area is the Adour–Garonne drainage basin in south-western France. This hydrographic network comprises 120,000 km of flowing waters draining a total area of 116000 km^2. Six hydrographic sub-basins (Adour, Charente, Dordogne, Garonne, Lot, Tarn-Aveyron) form this large watershed which covers 20% of France. The development of the HCE model used data collected from 50 flow gauges which characterised a wide range of hydrological (snow to rainfall dominated regimes) and climatic conditions (of mountainous, continental or oceanic influence) (Fig. 1, 2).

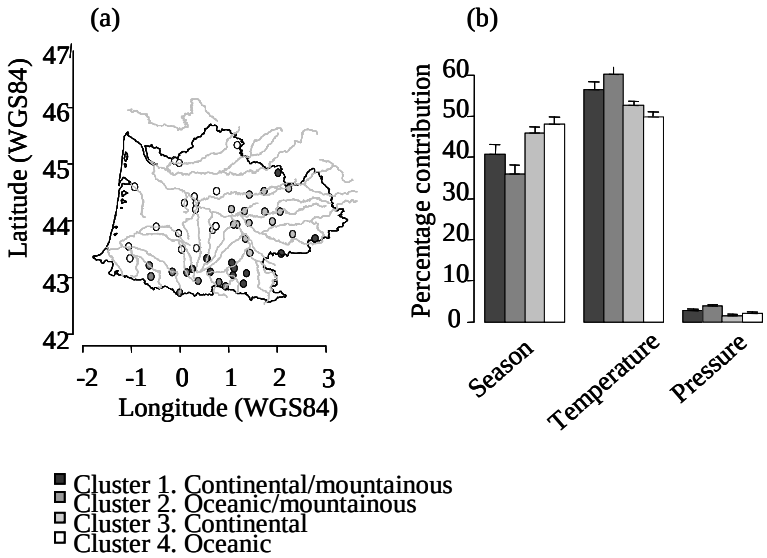


Fig. 1. Location of the four homogeneous regions for temperature identified using Hierarchical Ascending Clustering (a). For each region, the relative contribution to the regional temperature variability explained by the atmospheric temperature, pressure as well as seasons, was derived from the regional downscaling (b).

Data

Daily mean flow data (m^3s^{-1}) were collated from the Hydro2 database maintained by the Ministère de l'Ecologie et du Développement Durable (http://www.hydro.eaufrance.fr/) for the period 1970 to 2000. This period was defined as the control period for the hydro-climatic modelling. Hydrological stations were located less than 20 km upstream from the fish sampled sites to assume hydrological data were indicative of the hydrological conditions at the sites where the fish species characterisation was surveyed.

Daily time series of temperature (°C) were interpolated using kriging at the 50 fish monitoring sites over the control period using 160 local daily climate stations provided by Météo-France, located over the region. The kriging model was set with an exponential covariance function. Namely, the correlation between sites was assumed to be an exponential function of their Euclidean distance based on longitude, latitude and altitude.

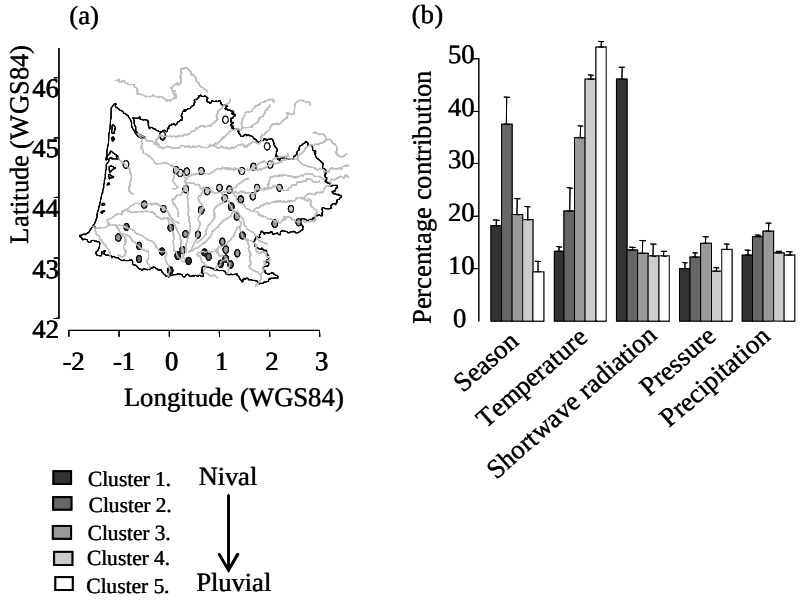


Fig. 2. Location of the five hydrological regions ranging from nival (cluster 1) to pluvial (cluster 5) systems and identified using Hierarchical Ascending Clustering (a). For each hydrological region, the relative contribution to the local flow variability explained by the atmopsheric temperature, shortwave solar radiation, pressure and precipitation as well as seasons, was derived from the regional downscaling (b).

For each fish monitoring site, the monthly low, median and high hydro-climatic conditions were characterized by the monthly flow and the 10, 50 and 90% temperature percentiles (noted P10, P50 and P90), derived over the control period from the daily flow and temperature time series. Annual fish survey data were extracted from the Office National de l'Eau et des Milieux Aquatiques (ONEMA) ranging from 1992 to 2000. Fish occurrence data

(i.e. presence-absence) were used for the 13 most prevalent species which were present in more than 30% of sites over the period of survey (Table 1). Geomorphological data collated from the ONEMA database described the physical catchments characteristics at the 50 sites such as their distance from the river source (km), their drainage area (km²), longitude (degree EW) and latitude (degree NS), altitude (m), slope (%), river width (m) and depth (m).

Table 1. Prevalence of 13 studied species over the region of study

Species name	Common name	Code	Prevalence
Perca fluviatilis	Perch	Pef	0.30
Chondrostoma toxostoma	Soufie	Cht	0.30
Leuciscus leuciscus	Dace	Lel	0.36
Lepomis gibbosus	Pumpkinseed	Leg	0.36
Salmo trutta fario	Brown trout	Sat	0.49
Anguilla anguilla	European eel	Ana	0.51
Alburnus alburnus	Bleak	Ala	0.54
Barbatula barbatula	Stone loach	Bab	0.55
Barbus barbus	Barbel	Bar	0.59
Rutilus rutilus	Roach	Rur	0.62
Phoxinus phoxinus	Minnow	Php	0.63
Leuciscus cephalus	Chub	Lec	0.69
Gobio gobio	Gudgeon	Gog	0.77

Reanalysis data from the National Centre for Environmental Prediction and the National Centre for Atmospheric Research (NCEP/NCAR; Kalnay et al., 1996) were used over the control period to calibrate the HCE model. Reanalysis data are considered as large spatial scale records of atmospheric variables of approximately 2.5° x 2.5° spatial scale resolution derived from the assimilation of surface observation stations, upper-air stations and satellite-observing platforms with long records. Eleven GCMs were tested to validate the HCE model projections under the control period, downloaded online from the IPCC website at https://esg.llnl.gov:8443/index.jsp (Table 2).

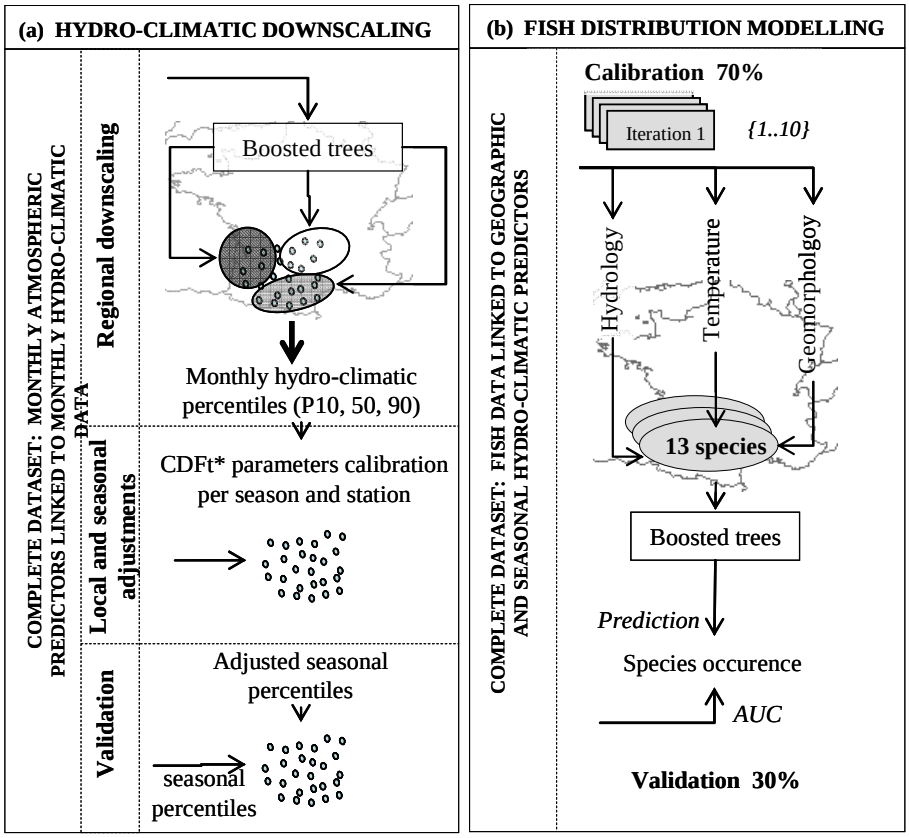
Table 2. Acceptable atmospheric variables among 11 tested GCMs. In bold, the best set of GCM and variable selected for the downscaling.

Atmospheric variables			GCM availability										
Name	Code	Unit	cccma_cgcm3_1	**cnrm_cm3**	csiro_mk3_5	**gfdl_cm2_0**	**gfdl_cm2_1**	giss_model_e_r	inmcm3_0	ipsl_cm4	**miroc3_2_medres**	mpi_echam5	**mri_cgcm2_3_2a**
Mean daily air temperature at 1000 hPa	air.1000	K	×	×		×	×	×	×		×		×
Mean daily air temperature 2 meters above surface	air.2m	K	×	×		×	×	×	×		×	×	×
Mean daily air temperature at 500 hPa	air.500	K	×	×		×	×	×	×		×		×
Mean daily air temperature at 850 hPa	air.850	K	×	×		×	×	×	×		×		×
Mean daily convective precipitation rate at surface	cprat	kg m^{-2} s^{-1}			×	×	×	×			×	×	×
Mean daily clear sky downward longwave flux at surface	csdlf	W m^{-2}	×	×		×	×	×	×		×		×
Mean daily clear sky upward solar flux at surface	csusf	W m^{-2}		×		×	×	×				×	
Mean daily downward longwave radiation flux at surface	dlwrf	W m^{-2}	×	×	×	×	×	×	×		×	×	×
Mean daily downward solar radiation flux at surface	dswrf	W m^{-2}		×							×		
Mean daily geopotential height at 1000 hPa	hgt.1000	m				×							
Mean daily geopotential height at 500 hPa	hgt.500	m	×			×	×	×	×		×		
Mean daily geopotential height at 850 hPa	hgt.850	m	×			×	×	×	×	×	×		
Mean daily precipitation rate at surface	prate	kg m^{-2} s^{-1}	×	×	×	×	×	×	×	×	×	×	×
Mean daily surface pressure	pres	Pa	×	×	×	×	×				×	×	×
Mean daily relative humidity at 1000 hPa	rhum.1000	%	×			×	×						
Mean daily relative humidity at 500 hPa	rhum.500	%	×			×		×			×		×
Mean daily relative humidity at 850 hPa	rhum.850	%	×			×	×	×			×		×
Mean daily specific humidity at 1000 hPa	shum.1000	kg kg^{-1}					×		×		×		×
Mean daily specific humidity at 1000 hPa	shum.500	kg kg^{-1}					×		×	×	×		
Mean daily specific humidity at 1000 hPa	shum.850	kg kg^{-1}	×				×		×		×		
Mean daily SST/land skin Temperature	skt	K		×		×	×	×	×		×		×
Mean daily sea level pressure	slp	Pa			×	×							
Mean daily total cloud cover	tcdc	%	×							×	×	×	×
Mean daily upward longwave radiation flux at surface	ulwrf	W m^{-2}		×		×	×		×		×		×
Mean daily upward solar radiation flux at surface	uswrf	W m^{-2}	×	×		×	×	×			×	×	×

Overall, 21 atmospheric variables both from reanalysis and GCMs database were tested as relevant predictors to drive the hydro-climatic modelling (Table 2). These atmospheric variables were related to long wave and short wave radiation fluxes, cloud cover, land skin temperature, latent and sensible heat fluxes at surface. The overall atmospheric variables were interpolated at the 50 sites of study using bilinear interpolation, and then standardized using the mean and standard deviation).

Conceptual framework

The conceptual framework of the HCE model was based on two separated statistical downscaling models which simulated respectively the seasonal hydrological and temperature variability (Fig. 3a). Those two statistical downscaling models were then coupled to fish-specific statistical distribution models to calculate the probability of occurrence over the region for the 13 fish species (Fig. 3b). The hydrological and temperature variability within the year is of particular importance to complete the biological cycle for most fish species over the region. Thus three seasons were defined according to Cattanéo et al. (2001), for the adjustment of downscaled hydro-climatic outputs as well as the definition of predictors for fish models: (i) the winter season, from October to February, commonly defined as a period of low activity for fish; (ii) the spawning season, from March to June, encompassing the major part of the reproduction time of most fish (except for the brown trout), although some species can extend their spawning activity beyond this limit, until mid-summer; (iii) the growth period, from July to September, during which fish actively feed.



* Cumulative Distribution Function transformation

Fig. 3. Structure of the hydro-climatic-ecological (HCE) model-chain built upon two downscaling models to model respectively the hydrological and temperature variability at the 50 local sites of study (a), coupled to a distribution models to simulate the fish occurrence for 13 species (b).

The core of the HCE model was built upon the boosted trees (BT) statistical model. The calibration and validation of HCE was done in three steps. Firstly, an optimal set of large-scale atmospheric predictors was selected to drive the hydro-climatic downscaling models, representative of key hydro-climatic atmospheric processes accurately modelled by a maximum number of GCMs. This was done to limit some erratic sources of uncertainty in hydro-climatic projections which may be due to the inability of some GCMs to accurately model certain atmospheric fields. Secondly, the hydro-climatic downscaling and the fish distribution models were calibrated independently from each other based on historical observations. Thirdly, the selected GCMs outputs were used as input predictors to drive the HCE model and both hydro-climatic and fish projections under the control period were compared to the current historical records.

Boosted trees (BT)

Boosted trees are based on a compilation of classification and regression tree (CART) models. CART models (Breiman et al., 1984) explain variations of a single response variable by repeatedly splitting the data into more homogeneous groups, using combinations of explanatory variables that may be categorical and/or numeric. Each group is characterized by a typical value of the response variable, the number of observations in the group and the values of the explanatory variables (De'ath & Fabricius, 2000). The aim of boosted trees is to improve the performance of a single CART model by fitting several CART models, in this study 1000 models. Each successive CART model was built for the prediction residuals of the preceding tree, each time based on a randomized subset from the original database, here 70%. Such a randomization in the boosting algorithm makes each boosted trees model run unique, which may help to assess the uncertainty in predictions if performing different model runs and thus, improve the robustness of results (Elith et al., 2008).

The relative importance of each predictor was assessed using the method developed by Friedman et al. (2001). This was done to better understand the relationship between the atmospheric process and the regional hydro-climatic variability as well as between environmental descriptors of the fish habitat. The method is based on the number of times a predictor is selected for splitting, weighted by the squared improvement (i.e. the loss in predictive performance) to the model as a result of each of those splits, and averaged over all CART models. The relative importance of each predictor is scaled so that the sum adds to 100, with higher numbers indicating stronger contribution to the response (Elith et al., 2008).

Pre-selection of atmospheric variables and GCMs

A pre-selection of atmospheric variables and GCMs was performed in three steps to get the most reliable large scale atmospheric variables to drive the hydro-climatic downscaling process. At step 1, four well-identified key atmospheric hydro-climatic processes related to precipitation, temperature, solar radiations and pressure were highlighted by clustering the 21 standardized NCEP/NCAR atmospheric fields using hierarchical ascending clustering (HAC) with Ward criterion and the Euclidean distance (Fig. 4).

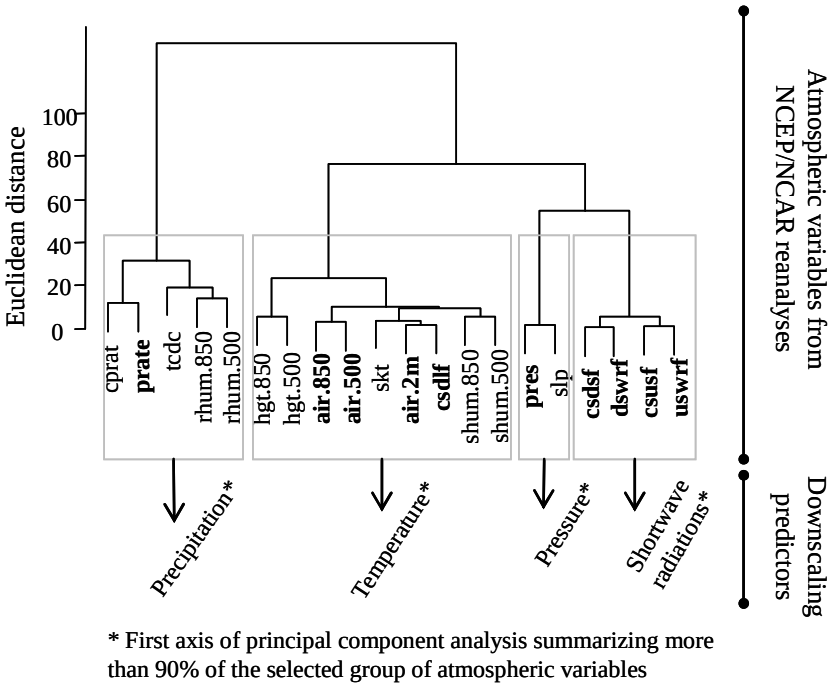


Fig. 4. Hierarchical Ascending Clustering (HAC) of atmospheric the 21 NCEP/NCAR variables highlighting key hydro-climatic atmospheric processes such as temperature, shortwave radiations and precipitations pressure. Pre-selected variable in bold (Table 2) were used to apply Principal Component Analysis (PCA) within each cluster. The fisrt PC of each group was used as predictors into the downscaling framework (Fig. 3a).

At step 2, each GCM standardized variable was compared to the equivalent for NCEP/NCAR and the tested variable was assumed ‘acceptable’ if its inter-annual variability was accurately modelled by the GCMs, following two criteria:(i) the null hypothesis of equality of the two empirical distributions over the control period was accepted using the Cramer von Mises test (Anderson, 1962) at the 5% level of confidence; (ii) for each of the three biological seasons, namely winter, spawning and growth, the absolute difference of the means of the two datasets (GCM and NCEP/NCAR) was lower than half a standard deviation

(i.e., lower than 0.5 since the data are standardized with unit variance) for variables in the temperature and short-wave radiations groups, and lower than three-quarters of a standard deviation (i.e., lower than 0.75) for variables in pressure and precipitation groups. The selection threshold was particularly higher in the case precipitation groups of atmospheric variables since it was verified that the GCM performed worse in modelling the water cycle than radiation processes (Xu et al., 1999).

At step 3, for each GCM, all possible combinations between the acceptable variables were generated, containing at least one variable of each hydro-climatic process identified at step 1. For a given combination, the number of relevant GCMs was recorded and a score was calculated as the absolute difference between the mean of NCEP/NCAR and the GCMs (step 2), averaged over the whole variables in the combination, and divided by the sum of the total number of variables and GCMs related to this combination.

The best combination (i.e. that with the lowest score) maximized the number of variables and GCMs while minimizing the error in the inter-annual variability of GCMs. Thus five GCMs and 11 NCEP/NCAR variables, distributed in the four atmospheric hydro-climatic processes identified in step 1, were selected and are listed in bold in Table 2.

Hydro-climatic downscaling model

The hydro-climatic downscaling aimed at projecting the local flows and temperature percentiles (P10, P50 and P90) at the 50 sites at a monthly time step, from the 11 pre-selected atmospheric variables. The hydro-climatic downscaling process separated the hydrological downscaling from that for temperature, although the overall procedure was globally the same. The procedure was summarized into the five following steps that are more detailed in the next paragraphs (Fig.3a): (i) the 11 pre-selected atmospheric variables were synthesized into four atmospheric predictors related to precipitation, temperature, shortwave radiation and pressure; (ii) the 50 sites of study were grouped into five hydrological and four temperature regions based on clustering techniques; (iii) for each region a statistical model was built using the boosted tree model to relate the regional hydro-climatic variability to the four large-scale atmospheric predictors; (iv) the resulting regional projections were finally refined to each gauging station individually using the cumulative distribution function transformation (CDFt; Michelangeli et al., 2009); (v) the overall downscaling model was based on a cross-validation process to project the local hydro-climatic variability.

Large scale atmospheric predictors

For each of the four identified clusters related to the temperature, precipitation, shortwave radiation and pressure atmospheric processes, a principal component analysis (PCA) was performed onto the matrix defined by the corresponding NCEP/NCAR standardized variables of this cluster (Fig. 4). The first axis of PCA was retained as a synthetic descriptor of the process of interest, summarizing more than 90% of the variance, and was used as predictor into the downscaling framework. Such representation of predictors had the main advantage of summarizing the space of atmospheric fields in a limited number of physically meaningful predictors. This representation also reduces the collinearity between each pair of predictors (Pearson cross-correlations below 0.7) so that their influence on the local hydro-climatic variability may be quantified with more confidence throughout the statistical downscaling modelling framework.

Hydro-climatic regions

The 50 sites of study were grouped into four temperature (Fig. 1a) and five hydrological (Fig. 2a) regions using HAC with Ward criterion (Ward, 1963) and the Euclidean distance, based on standardized monthly P10, P50 and P90 of stations. The number of hydrological and temperature regions was determined qualitatively according their meaningful physical and/or geographical interpretation. The four climate groups highlighted different climate influence, from continental/mountainous (clusters 1), oceanic/mountainous (cluster 2), continental (cluster 3) to oceanic (cluster 4) influence (Fig. 1a). The five hydrological regions ranging from cluster 1 to cluster 5 characterized a nival (snow-dominated) to pluvial (rainfall-dominated) hydrological gradient (Fig. 2a).

Standardization of the monthly flow and temperature percentiles was done by subtracting the median and dividing by the standard deviation of each station, so that the dimension of the monthly statistics was comparable between stations. Standardization was based on the median rather than on the mean to better represent the statistical mode of the distribution, especially in the case of a skewed distribution.

Regional downscaling

For each of the five hydrological and four temperature regions, a single boosted tree model was built for each standardized monthly flow or temperature percentiles (P10, P50, P90). That is, 27 different boosted tree models were calibrated (i.e. 3 percentiles × 5 hydrological regions + 3 percentiles × 4 temperature regions). The four atmospheric predictors were used to

downscale the hydrology while only the temperature and pressure predictors were used to downscale the temperature, as classically done in most downscaling studies (Wilby et al., 1999).

Furthermore, rather than calibrating separated downscaling models per season or per month, a single model was developed by including the sine and cosine values of the 12 months of the year, as two additional monthly predictors for the hydro-climatic downscaling. Throughout additional results, not presented here, including those two monthly predictors had to major advantages: (i) to allow the downscaling models to take into account the period of the year where the downscaling has to be performed; (ii) to reduce the variability in seasonal projections and improve the regional model fitting to observations.

Local and seasonal adjustment of regional downscaling projections

The hydro-climatic projections from the 27 regional downscaling models were then adjusted seasonally to each individual station. This was done for two major reasons: (i) the regional projections give a baseline in the hydro-climatic processes occurring at each stations but they generally underestimate the extremes of individual station (i.e. the tails in the probability distribution of the local hydro-climatic processes are underestimated); (ii) the pre-selection step of the atmospheric 11 variables had revealed that the 5 GCMs were particularly biased during the winter season and for precipitation related fields, which may be explained by the well known weakness of GCM to accurately model the water cycle.

The local and seasonal adjustment was done using the “Cumulative Distribution Function - transform” approach (CDF-t; Michelangeli et al., 2009) which is an extension of the more commonly applied quantile-quantile approach (Déqué, 2007). CDFt is a mathematical transformation which was used to transpose the probability distribution of the regional downscaled projection to that of observations. More particularly, CDFt was applied per biological season and per station in the aim to optimize the quality of projections to be used as predictors into the fish models.

Cross-validation of models and hydro-climatic projections

The hydro-climatic downscaling models were validated using a cross-validation procedure based on three temporally independent periods of approximately 10 years over the control period, denoted *a*, *b*, *c*. These three periods were successively used to calibrate the regional downscaling model (e.g. period *a* in the first instance), calibrate the CDF-t approach (e.g. on the downscaled data for period *b* from the downscaling model calibrated on *a*), and validate

the adjusted downscaled results (e.g., *c*), so that six combinations of downscaling models were generated, namely *abc*, *acb*, *bac*, *bca*, *cab*, *cba*. This sampling design was used to account for the problem of non-stationarity in the temperature and flow time series over the control period. Specifically, the probability distribution of the temperature and flow time series was not constant over the three periods, presumably due to some natural cyclic variations in the temperature and flow signal. Hydro-climatic projections from the six validation periods were then monthly averaged for each year over the control period to estimate the goodness-of-fit of models. The hydro-climatic projections were done successively for NCEP/NCAR reanalyses and for the five selected GCMs from the downscaling models calibrated from NCEP/NCAR data.

Fish species distribution model

Seasonal hydro-climatic and geomorphologic predictors

For each year where data were collected at each fish survey site (i.e., 50 sites × for approximately 6 years of sampling), 24 seasonal hydro-climatic predictors for fish models were derived from the three monthly hydro-climatic percentiles (P10, P50, P90), for each biological season (winter, spawning, growth), separately for flows and temperature data . The seasonal flow percentiles were divided by the median discharge of each site, computed for the entire period of the flow record, to highlight the magnitude of flows related to the overall median conditions (Cattanéo, 2005). The overall hydro-climatic variability was defined as the difference between the P90 and P10 values for both seasonal flows and temperature. This difference characterised the amplitude in the shift between the low and high hydrological or thermal conditions.

Two geomorphological indices were derived from the two first axis of another PCA applied to the standardized variables related to catchment characteristics. These characteristics were: the distance from the source, catchment size, longitude and latitude, altitude, slope, river width and depth. Distance of the sites from the source and the catchment size were first box-cox transformed (Box & Cox, 1964) to make the shape of their distribution as Gaussian as possible. The first axis, explaining 62 % of the variance, characterized the position of the sites along longitudinal gradient whereas the second axis, explaining 16 % of the variance, described a SW–NE gradient.

Bootstrap calibration of species-specific models

Thirteen species-specific Boosted Tree models were built, relating the current fish species occurrence at each annual site to the seasonal hydro-climatic and geomorphological predictors. A binomial distribution of errors was assumed and the probability of species occurrence was related to the predictors via a logistic link function. Seventy percent of the dataset were randomly selected to calibrate the models and the probability of occurrence was simulated on the remaining 30% of validation dataset. The whole procedure of calibration and validation was randomly repeated 10 times and the simulated results for the validation period were then averaged to give consistency and robustness in the results (Fig. 3b).

Validation of the hydro-climatic-ecological model chain

The downscaled monthly times series for the 24 seasonal hydro-climatic projections from the five GCMs and NCEP/NCAR data were derived as predictors into the calibrated fish-specific BT models to project the fish species occurrence over the control period. Thus, the validation of the HCE model was performed by evaluating the quality in both downscaled hydro-climatic projections (as relevant predictors for the fish-specific models) and the resulting fish projections.

Each of the 24 seasonal hydro-climatic projections was averaged per site over the control period according to the downscaled GCM or NCEP/NCAR data. Similarly, the projected probability of species occurrence from the five GCM and NCEP/NCAR was averaged per site over the period of fish record, individually for each species.

Both statistical tests described in the following were performed to assess the HCE ability to fit the observed spatial variability in the hydro-climatic conditions and fish occurrence. The statistical significance of the different tests was evaluated by 1000 permutations under 5% level of confidence, by randomly permuting sites as a way to test the spatial consistency between the observed and simulated results.

Validation of hydro-climatic downscaling models

For each of the 24 seasonal hydro-climatic projections, the Spearman rank correlation co-efficient, ρ was calculated between the downscaled and the observed hydro-climatic data and the downscaled and observed data were considered as significantly dependent/correlated if rejecting the null hypothesis of independence. A Kendall test was done to test the null

hypothesis that the ρ correlations for the 24 hydro-climatic projections were significantly of similar quality according to the five GCMs and NCEP/NCAR downscaling models.

For each GCM, the overall quality of the 24 hydro-climatic projections was tested using the Mantel r correlation test as a measure of the correlation between two Euclidean dissimilarity matrices (Legendre & Legendre, 1998), in our case the two matrix of observed versus downscaled hydro-climatic predictors. The Mantel r test was performed by considering successively the hydrological and temperature sets of predictors alone (i.e. the 12 hydrological predictors first, then the 12 temperature ones), then the overall sets of 24 hydro-climatic predictors. The downscaled hydro-climatic projections were significantly and spatially correlated to observations if rejecting the null hypothesis of independence between the two matrices (i.e. if $p<0.05$).

Validation of fish models

The quality of each fish projection from each GCM or NCEP/NCAR database-driven was evaluated using the area under the curve (AUC) method of a receiver operating characteristic (ROC) plot (Fielding & Bell, 1997; Pearce & Ferrier, 2000). The AUC score was calculated from the observed occurrences and the projected probability of occurrence. Then for each GCM and NCEP/NCAR database-driven, the projected presence-absence of species was derived by maximizing the number of true presences and true absences of species.

The AUC score ranges between 0 and 1 with a value of 0.5 for models that do not discriminate better than chance, and 1 for a 'perfect' model (Swets, 1988). AUC scores were tested by permutations and they were considered as significant if rejecting the null hypothesis of spatial independence between the observed and projected fish occurrence. A Kendall test was performed to test the null hypothesis that the AUC scores of the 13 species were significantly consensual between GCM or NCEP/NCAR models-driven and observations.

For each GCM and NCEP/NCAR, the Mantel r correlation test was performed to test if the projected fish assemblages were spatially consistent with the observations. The projected presence-absence for the 13 species was combined as a matrix of 13 columns and the Jaccard dissimilarity matrix of sites was calculated (Jaccard, 1901). The expected spatial correlation between the observed and the projected matrix of fish assemblages was assumed if rejecting the null hypothesis of spatial independence (i.e. $p<0.05$).

Results

Hydro-climatic downscaling models

The percentage contribution of atmospheric predictors to the regional temperature variability, calculated through the calibration of the regional downscaling models, was comparable between the four temperature regions (Fig. 1b). The atmospheric temperature and the seasonal information contributed approximately to 50 ± 4 % and 37 ± 5 % of the local temperature variability (Fig. 1b).

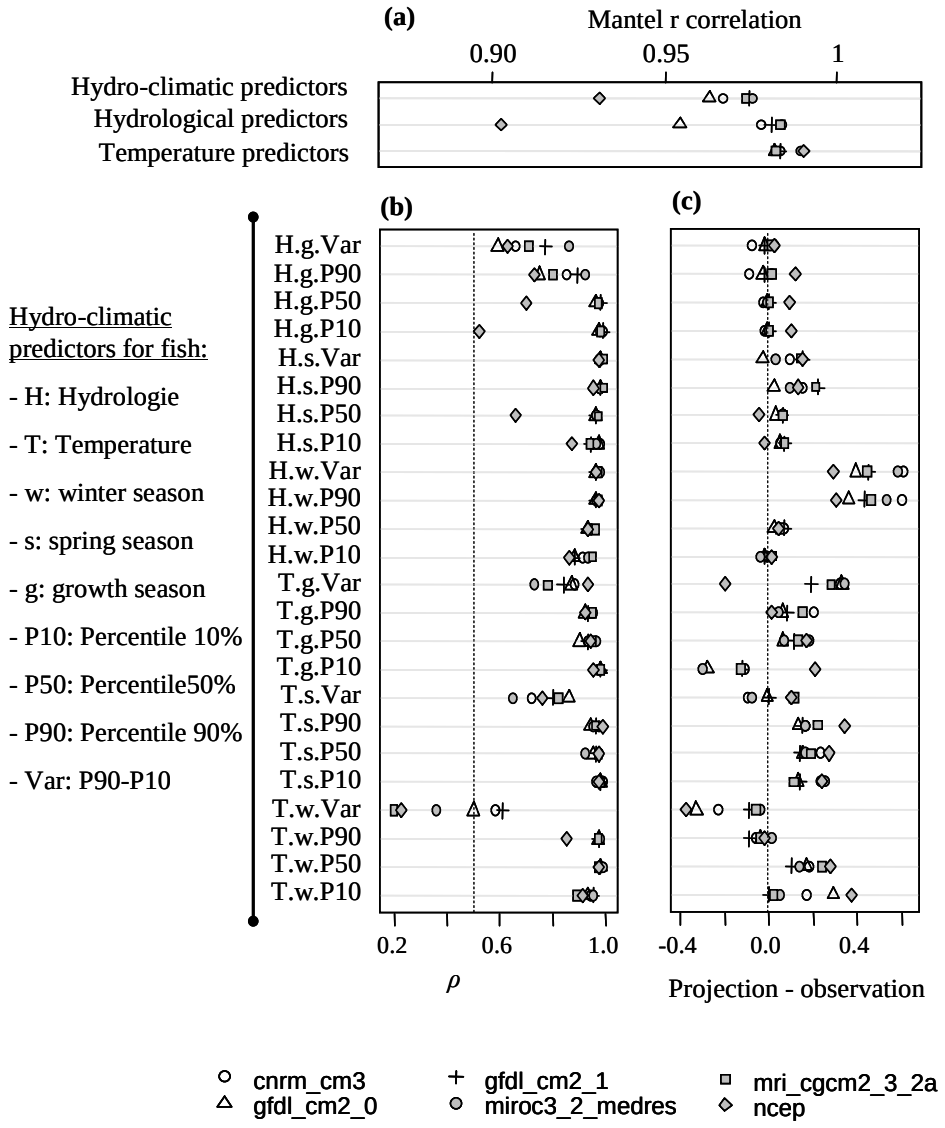


Fig. 5. Quality of the downscaled hydro-climatic projections under the control period (1970-2000) according to the downscaling models driven by NCEP/NCAR reanalysis or GCMs: Mantel r correlation between the Euclidean dissimilarity matrices of observed and projected results, separately for the hydrology, temperature and both set of predictors (a); Spearman rank correlations (b) and difference between individual downscaled and observed predictors (c).

The contribution of the four atmospheric predictors to the monthly flow variability was contrasted between nival and pluvial regimes (Fig. 2b). Whereas the contribution of shortwave solar radiations gradually decreased from 50 ± 2 % to 13 ± 2 % in pluvial systems (Fig. 2b; cluster 4-5), that of atmospheric temperature gradually increased from 13 ± 3 % to 45 ± 2% (Fig. 2b; cluster 1-2). The contributions of precipitation and pressure atmospheric predictors remained more stable between hydrological regimes, ranging between 12± 4 % and 11± 4 %, while the monthly signal contributed to 18± 3 % and until 35± 6 % in transitional nival to pluvial regimes (Fig. 2b; cluster 3).

Globally, both hydrological and climatic variability was satisfactorily well modelled by the five GCM- and reanalyses-driven model as the Mantel *r* correlation with the observations was higher than 0.9 (Fig. 5a, $p < 0.01$). The hydrological downscaling models showed somewhat lower performances when driven by NCEP/NCAR than by GCM predictors since the Mantel *r* correlation respectively averaged at 0.90 and at 0.96. However, such a range of correlation values should be considered as very comparable (Fig. 5a).

Individually, each downscaled hydro-climatic variable was significantly well spatially correlated to the observations since, on average, ρ was higher than 0.7 (Fig. 5b; $p < 0.01$). During winter, downscaling models of temperature performed not as good as during the other biological seasons as the ρ correlation value was lower than 0.5. However, globally the whole hydro-climatic projections correlated well with the observations and were not statistically different between the GCM- or NCEP/NCAR-driven downscaling models (Kendall, $p > 0.45$).

The difference in the means between each individual downscaled and observed predictors showed that the downscaling models slightly overestimated both the seasonal hydrological and the temperature during all seasons, although this overestimation was reasonable and averaged approximately at +0.15 for flow, and at +0.2 °C for temperature (Fig. 5c).

Prediction of fish species occurrence

Globally, the seasonal temperature and the geography explained the main part of the fish species occurrence, respectively 40 ± 4 % and 34 ± 4 %, while the seasonal hydrological variability explained approximately 26 ± 2 %. More particularly, the median conditions of temperature during the spawning season, as well as the longitudinal gradient, explained respectively 10 and 24% of species distribution (Fig. 6).

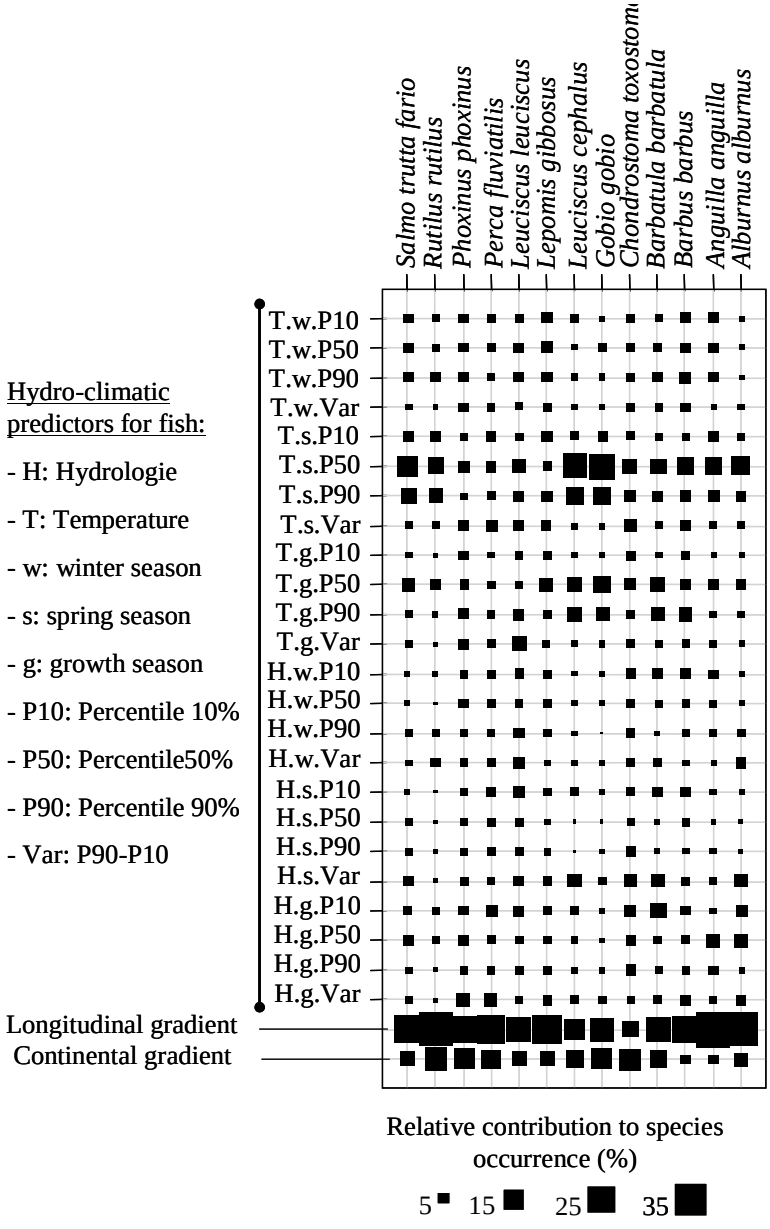


Fig. 6. Relative contribution of the 24 hydro-climatic and two geographic predictors to the 13 fish species occurrence simulated by the fish-specific boosted tree models.

The overall spatial structure in fish assemblages was consistently modelled by both GCM- and reanalyses-driven models as the Mantel correlation r was statistically significant and averaged at 0.6 for all the models (Fig. 7a; $p < 0.05$) while the highest correlations were shown by the models driven by observed hydro-climatic predictors.

AUC scores were higher than 0.7 for all the 13 fish models when driven by the five GCMs or NCEP/NCAR datasets, and averaged at 0.86 ± 0.08 (Fig. 7b; $p<0.01$), showing the overall good performance of the HCE models to simulate the fish species distribution over the region. Globally, fish models driven by observed predictors comparatively recorded better AUC scores than others models, in average 0.86 (driven by observations) and 0.83 (driven by downscaled projections). However, the AUC scores of species were not significantly different between GCM- and observation-driven models (Kendall, p=0.34).

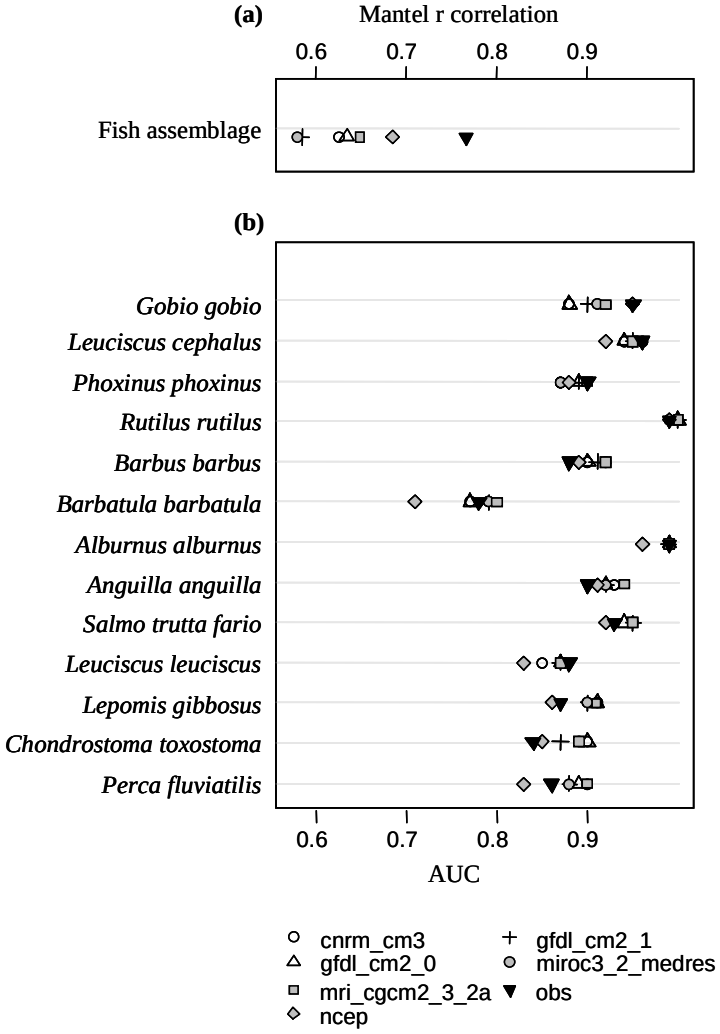


Fig. 7. Quality of spatial fish projections according to models driven by observed or downscaled GCM and reanalysis hydro-climatic predictors: Mantel r correlation between projected and observed Jaccard dissimilarity matrix of species occurrence (a); AUC scores for individual species (b).

Discussion

Consistency of hydro-climatic projections

Selecting an adequate set of large-scale atmospheric variables is of major importance in downscaling applications since the relationships between the large and small spatial scale hydro-climatic processes have to be physically meaningful (Wilby et al., 1999). More especially for future climate change impact studies, the selected atmospheric variables also require to be satisfactorily modelled by GCMs under the control period to be credibly used for future climate change projections. Whereas this first validation step is rarely done in most climate change impact studies of species distribution, a first key result of this study highlighted the inability of more than half of the 11 tested GCMs to represent the key seasonal features of the atmospheric circulation over the past 30 years in the region of study. Thus, before projecting future climate change, careful selection of GCMs and their atmospheric variables is required to reduce uncertainty and increase confidence in the simulation results.

The regional component of the statistical downscaling method presented here has shown the good ability of the boosted trees models to capture the relationship between the large-scale atmospheric processes and the local hydro-climatic variability. As found in other downscaling studies (Wilby et al., 2002), the regional temperature variability was shown to be mainly driven by the large-scale atmospheric temperature variability. In this study, this relationship was observed in the overall temperature regions. Conversely, the relationship between atmospheric predictors and flows showed some contrasted patterns between hydrological regions. The shortwave solar radiations showed their strongest control in the hydrology of nival systems (snow-dominated), by presumably triggering the snowmelt process as discussed in some recent snowmelt modelling studies (Li & Williams, 2008). In pluvial regimes (rainfall dominated), the atmospheric temperature was the key driver of the hydrological process, by possibly regulating the soil moisture and flow pathways throughout the evaporation.

The CDFt approach was then used to adjust the regional hydro-climatic projections to each station, which showed the good ability of the five tested GCMs to model the spatial and seasonal variability of the observed low, median and high hydro-climatic conditions. The hydrological projections were likely better modelled by the GCMs than by the NCEP/NCAR reanalyses, although not significantly different. This may result from the combination of two

factors, namely the overall lower variability in the climate signal of GCM fields than in the observations, as well as the non-stationarity detected in the flow and temperature times series over the control period. Thus, the downscaled NCEP/NCAR reanalyses could lead to some erratic overestimations in the projections throughout the CDF-t processing, as shown for the hydrological modelling in winter by inflating the probability density of downscaled flow.

Validation of the simulations of fish species assemblages

This is one of the first studies to combine both the hydrological and climatic temporal variability, as well as the geographical characteristics of the catchments, as environmental predictors of fish species occurrence at the regional scale. The 13 studied species were shown to have different sensitivity to the three environmental predictors, which confirms the interest of using specific-fish models to represent species assemblages, as confirmed by Pont et al. (2005), rather than using direct predictive techniques of fish assemblages such as canonical correspondence analysis. The stream gradient and thermal conditions globally explained more than 60% of most fish species occurrence, which is in agreement with both the zonation concept of Huet (1959) and thermal niche preferences (Magnuson et al., 1979) that are known to be the two factors that best explain fish species distribution (Matthews, 1998). The hydro-climatic conditions were shown to be critical during the spawning and growth seasons for most species, which is in accordance with the life history of most species (Mills & Mann, 1985; Daufresne et al., 2004). Cattaneo (2005) highlighted the importance of high flows magnitude during these two seasons to promote the recruitment of some species guilds. During the growth period of summer, the lower summer flows in combination with the rising temperature generally lead to oxygen depletion in freshwaters that may particularly impact the survival rate for the young-of-the-year fish populations (Gibson et al., 2005). During the spawning season, the magnitude and variability of high flows may disturb the fish spawning in several ways according to Cattaneo (2005). For those species that spawn in one batch before April and thus have a low number of reproductive cycles, brood hiders, with a medium fecundity, e.g., dace (*Leuciscus leuciscus*), stone loach (*Barbatula barbatula*), sculpin (*Cottus gobio*), high flows may directly affect the early-life stages. On the opposite, high spawning flows may favour another groups of species which are later and fractional spawners, highly fecund, lithophilic, small-egg depositors, non brood-hiders and that have a high number of reproductive cycles, e.g., chub (*Leuciscus cephalus*), minnow (*Phoxinus phoxinus*) and barbel (*Barbus barbus*).

Although the goodness of projected fish results could be slightly overestimated due to pseudo replication problems in the sampling design which could lead to underestimated standard errors and an inflated Type I error rate (Hurlbert, 1984; Millar & Anderson, 2004), the spatial distribution of the 13 species was significantly well projected by the different selected GCMs under the control period. When analysing the whole projected species altogether, Mantel test highlighted that the overall spatial structure in fish assemblages was significantly well projected by the five GCMs. Thus the 'predict first, assemble later' strategy used to assess the goodness-of-fit in projected fish species assemblages was validated. This is in agreement with Buisson et al. (2008) who employed a similar strategy to model the thermal habitat suitability for several fish species over France. Our projected fish results were not significantly different according to the five downscaled GCMs, which validate the whole HCE model-chain as reliable to couple the downscaled hydro-climatic conditions to the fish-specific distribution models.

Conclusion

The validation of the HCE model is a baseline for further investigations, especially to assess the future integrated impact of climate change on the hydro-climatic conditions and fish species distributions according to different greenhouse gas emission scenarios. In this context, the HCE model has several strengths: (i) the model is spatially and temporally explicit, which may help to give an overview about the potential spatial and temporal future dynamics in the hydro-climatic conditions and fish species distributions, which has rarely been investigated to our knowledge; (ii) although the model was applied to a 116 000 km² area, the diversity in the hydro-climatic conditions and fish assemblages over the region should help addressing and corroborating several hypotheses such as the global shift of fish species toward higher elevations and the upstream (Matulla et al., 2007; Buisson et al., 2008); (iii) the core of the model is non-parametric and thus very flexible, which makes the model concretely extendable to higher spatial scales (e.g. from national to continental scale), and the possibility to integrate some other ecological modelling components into the model chain such as water chemistry or land cover change.

Acknowledgments

This study is funded by the Eurolimpacs European project (contract number GOEC-CT-2003-505540) and partly supported by the GIS REGYNA project. We are indebted to the Office National de l'Eau et des Milieux Aquatiques (ONEMA) for providing fish data, the Ministère de l'Environnemenent et du Développement Durable (MEDD) for flow data, Météo-France for climate data, the National Oceanic and Atmospheric Administration (NOAA) for NCEP/NCAR reanalyses and the Intergovernmental Panel on Climate Change (IPCC) for downloadable GCM data at https://esg.llnl.gov:8443/index.jsp. We are grateful to Bertrand Urroz for helping us collating and formatting GCMs data.

References

Alcamo J, Döll P, Henrichs T, Kaspar F, Lehner B, Rösch T, Siebert S (2003) Development and testing of the WaterGAP 2 global model of water use and availability/Développement et évaluation du modèle global WaterGAP 2 d'utilisation et de disponibilité de l'eau. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques,* **48,** 317-337.

Anderson TW (1962) On the distribution of the two-sample Cramer-von Mises criterion. *The Annals of Mathematical Statistics,* **33,** 1148-1159.

Araujo MB, Pearson RG, Thuiller W, Erhard M (2005) Validation of species-climate impact models under climate change. *Global Change Biology,* **11,** 1504-1513.

Box GEP, Cox DR (1964) An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological),* **26,** 211-252.

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth. *Inc., Belmont, CA,* **358**.

Buisson L, Thuiller W, Lek S, Lim PUY, Grenouillet G (2008) Climate change hastens the turnover of stream fish assemblages. *Global Change Biology,* **14,** 2232-2248.

Cattanéo F (2005) Does hydrology constrain the structure of fish assemblages in French streams? Local scale analysis. *Archiv für Hydrobiologie,* **164,** 345-365.

Cattanéo F, Lamouroux N, Breil P, Capra H (2002) The influence of hydrological and biotic processes on brown trout (Salmo trutta) population dynamics. *Canadian Journal Of Fisheries And Aquatic Sciences,* **59,** 12-22.

Chu C, Mandrak NE, Minns CK (2005) Potential impacts of climate change on the distributions of several common and rare freshwater fishes in Canada. *Diversity & Distributions,* **11,** 299-310.

Daufresne M, Roger MC, Capra H, Lamouroux N (2004) Long-term changes within the invertebrate and fish communities of the Upper Rhône River: effects of climatic factors. *Global Change Biology,* **10,** 124-140.

De'ath G, Fabricius KE (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology,* **81,** 3178-3192.

Eaton JG, Scheller RM (1996) Effects of climate warming on fish thermal habitat in streams of the United States. *Limnology And Oceanography***,** 1109-1115.

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. *Journal Of Animal Ecology,* **77,** 802-813.

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation,* **24,** 38-49.

Fowler HJ, Blenkinsop S, Tebaldi C (2007) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal Of Climatology,* **27,** 1547-1578.

Friedman JH (2001) Greedy function approximation: A gradient boosting machine. *Ann. Statist,* **29,** 1189-1232.

Gibson CA, Meyer JL, Poff NL, Hay LE, Georgakakos A (2005) Flow regime alterations under changing climate in two river basins: Implications for freshwater ecosystems. *River Research and Applications,* **21,** 849-864.

Hari RE, Livingstone DM, Siber R, Burkhardt-Holm P, Guttinger H (2006) Consequences of climatic change for water temperature and brown trout populations in Alpine rivers and streams. *Global Change Biology,* **12,** 10-26.

Humphries P, Lake PS (2000) Fish larvae and the management of regulated rivers. *Regulated Rivers: Research & Management,* **16**.

Hurlbert SH (1984) Pseudoreplication and the design of ecological field experiments. *Ecological monographs,* **54,** 187-211.

Jaccard P (1901) Etude comparative de la distribution florale dans une portion des Alpes et des Jura.[Comparative study of the distribution of flora in a region of the Alps and the Jura] Bull. *Soc. Vaudoise Sci. Nat,* **37,** 547–549.

Jackson DA, Mandrak NE (2002), pp. 89-98. American Fisheries Society.

Jackson PBN (1989) Prediction of regulation effects on natural biological rhythms in south-central African freshwater fish. *Regulated Rivers: Research & Management,* **3,** 205-220.

Jani Heino RVHT (2009) Climate change and freshwater biodiversity: detected patterns, future trends and adaptations in northern regions. *Biological Reviews,* **84,** 39-54.

Jenkins M (2003) Prospects for Biodiversity. *Science,* **302,** 1175-1177.

Kalnay E, Kanamitsu M, Kistler R, *et al.* (1996) The NCEP/NCAR 40-year reanalysis project. *Bulletin Of The American Meteorological Society,* **77,** 437-471.

Lane SN (2008) What makes a fish (hydrologically) happy? A case for inverse modelling. *Hydrological Processes,* **22,** 4493 - 4495.

Lassalle G, Rochard E (2009) Impact of twenty-first century climate change on diadromous fish spread over Europe, North Africa and the Middle East. *Global Change Biology,* **15,** 1072-1089.

Legendre L (1998) *Numerical ecology.* Elsevier Science, Amsterdam.

Li XG, Williams MW (2008) Snowmelt runoff modelling in an arid mountain watershed, Tarim Basin, China. *Hydrological Processes,* **22,** 3931-3940.

Lindstrom G, Johansson B, Persson M, Gardelin M, Bergstrom S (1997) Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology,* **201,** 272-288.

Magnuson JJ, Crowder LB, Medvick PA (1979) Temperature as an ecological resource. *Integrative and Comparative Biology,* **19,** 331-343.

Magnuson JJ, Webster KE, Assel RA, *et al.* (1997) Potential effects of climate changes on aquatic systems: Laurentian Great Lakes and Precambrian shield region. *Hydrological Processes,* **11,** 825-871.

Matthews WJ (1998) *Patterns in freshwater fish ecology.* Chapman and Hall, New York.

Matulla C, Schmutz S, Melcher A, Gerersdorfer T, Haas P (2007) Assessing the impact of a downscaled climate change simulation on the fish fauna in an Inner-Alpine River. *International Journal Of Biometeorology,* **52,** 127-137.

Michelangeli PA, Vrac M, Loukos H (2009) Probabilistic downscaling approaches: Application to wind cumulative distribution functions. *Geophysical Research Letters,* **36**.

Millar RB, Anderson MJ (2004) Remedies for pseudoreplication. *Fisheries Research,* **70,** 397-407.

Mills CA, Mann RHK (1985) Environmentally-induced fluctuations in year-class strength and their implications for management. *Journal Of Fish Biology,* **27,** 209-226.

Nations U (2003) Water for People–Water for Life–The United Nations World Water Development Report. *Co-published with Berghahn Books, UK http://www. unesco. org/water/wwap/wwdr*.

O'Brien CM, Fox CJ, Planque B, Casey J (2000) Fisheries: climate variability and North Sea cod. *Nature,* **404,** 142.

Olden JD, Poff NL (2003) Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River Research and Applications,* **19,** 101-121.

Pachauri RK, Reisinger A (2007) *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Geneva, Switzerland.

Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling,* **133,** 225-245.

Poff NL, Allan JD, Bain MB, *et al.* (1997) The natural flow regime. *Bioscience,* **47,** 769-784.

Pont D, Hugueny B, Oberdorff T (2005) Modelling habitat requirement of European fishes: do species have similar responses to local and regional environmental constraints? *Canadian Journal Of Fisheries And Aquatic Sciences,* **62,** 163-173.

Postel S, Richter BD (2003) *Rivers for life: managing water for people and nature.* Island Press, Washington DC.

Rahel FJ, Keleher CJ, Anderson JL (1996) Potential habitat loss and population fragmentation for cold water fish in the North Platte River drainage of the Rocky Mountains: response to climate warming. *Limnology And Oceanography,* **41,** 1116-1123.

Reid PC, Borges MF, Svendsen E (2001) A regime shift in the North Sea circa 1988 linked to changes in the North Sea horse mackerel fishery. *Fisheries Research,* **50,** 163-171.

Schindler DW (1997) Widespread effects of climatic warming on freshwater ecosystems in North America. *Hydrological Processes,* **11**.

Sharma S, Jackson DA, Minns CK, Shuter BJ (2007) Will northern fish populations be in hot water because of climate change? *Global Change Biology,* **13,** 2052-2064.

Swets JA (1988) Measuring The Accuracy Of Diagnostic Systems. *Science,* **240,** 1285-1293.

Thuiller W, Lavorel S, Sykes MT, Araujo MB (2006) Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Diversity And Distributions,* **12,** 49-60.

Vorosmarty CJ, Green P, Salisbury J, Lammers RB (2000) Global Water Resources: Vulnerability from Climate Change and Population Growth. *Science,* **289,** 284-288.

Ward JH (1963) Hierarchical grouping to optimize an objective function. *Journal Of The American Statistical Association,* **58,** 236-244.

Whitehead PG, Wilby RL, Battarbee RW, Kernan M, Wade AJ (2009) A review of the potential impacts of climate change on surface water quality. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques,* **54,** 101-123.

Wilby RL, Dawson CW, Barrow EM (2002) sdsm-a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling and Software,* **17,** 145-157.

Wilby RL, Hay LE, Leavesley GH (1999) A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado. *Journal of Hydrology,* **225,** 67-91.

Xenopoulos MA, Lodge DM, Alcamo J, Marker M, Schulze K, Van Vuuren DP (2005) Scenarios of freshwater fish extinctions from climate change and water withdrawal. *Global Change Biology,* **11,** 1557-1564.

Xu C-y (1999a) From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. *Progress in Physical Geography,* **23,** 229-249.

Xu CY (1999b) Climate change and hydrologic models: A review of existing gaps and recent research developments. *Water Resources Management,* **13,** 369-382.

ARTICLE N°4

Spatio-temporal impacts of climate change on biodiversity: strengthen the link between downscaling and bioclimatic models.

Tisseuil C., Vrac M, Wade AJ, Grenouillet G, Gevrey M, Lek S

(En préparation)

SPATIO-TEMPORAL IMPACTS OF CLIMATE CHANGE ON BIODIVERSITY: STRENGTHEN THE LINK BETWEEN DOWNSCALING AND BIOCLIMATIC MODELS

Tisseuil C.[a], Vrac M.[c], Grenouillet G., Wade A.J.[b], Gevrey M.[a], Lek S.[a]

[a]Laboratoire Evolution et Diversité Biologique (EDB) UMR 5174, CNRS - Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 4 – France

[b]Aquatic Environments Research Centre, School of Human and Environmental Sciences, University of Reading, RG6 6AB, UK

[c]Laboratoire des Sciences du Climat et de l'Environnement (LSCE-IPSL) CNRS/CEA/UVSQ, Centre d'étude de Saclay, Orme des Merisiers, Bat. 701 91191 Gif-sur-Yvette, France

ABSTRACT

INTRODUCTION

Whereas the temporal variability of climate change may increase the probability of population extinction (Thuiller et al. 2008), this dimension has rarely been considered in most bioclimatic studies explicitly, but see Morin et al. (2007) and Zurell *et al.* (2009) based on dynamic modelling. Instead, most bioclimatic studies usually draw a picture of the potential spatial changes in biodiversity over few mean time period, classically 2050, 2080 or 2100 (e.g., Thuiller 2004; Araujo *et al.* 2006; Tuck *et al.* 2006; Mika *et al.* 2008; Buisson *et al.* 2009). However, better integrating both the temporal and spatial dimensions of climate change into bioclimatic models could better help anticipate the strength and kinetics of global change impacts on biodiversity structure and functioning. More particularly for management purposes, this could help identify core areas within a species' range (see Osborne & Suarez-Seoane 2007) and thus core areas for nature conservation (Zurell *et al.* 2009). Furthermore, improving the spatial and temporal resolution of climate change projections (e.g. inter-annual and seasonal climate variability) could also help better integrate some dynamical processes within bioclimatic models (e.g. dispersal or phonological processes such as flowering, leafing

or fruiting). However, the use of spatially and temporally high resolution climate projections into bioclimatic models is facing several difficulties. Firstly because General Circulation Models (GCM), which are currently the best tool we have for simulating future climate, are too coarse in spatial resolution (approximately 250 km × 250 km) to be directly used as input into bioclimatic models (Beaumont *et al.* 2008). Secondly, whereas GCMs projections are relatively consensual at the global or continental scale between or within (different runs) models, they may be importantly divergent when focussing at lower spatial and temporal scale (Beaumont *et al.* 2007).

In this context, GCM projections have to be necessarily downscaled at higher spatial and/or temporal scale resolution to provide suitable climate predictors for most impact models, e.g. by taking into account for the regional features of climate variability (e.g. topography or land cover). Although downscaling techniques have been increasingly developed for hydro-climatic impact studies during last ten years, their use for bioclimatic studies has paid much less attention, but see Beaumont *et al.* (2008) who warn the use of climate scenarios for species distribution modelling. Fowler *et al.* (2007) made a review of the recent advances in downscaling techniques that may be classified into dynamical and statistical ones. The dynamical downscaling usually involves a regional climate model (RCM) nested within a global climate model (GCM) at a lower scale than that of GCM, classically 50 km × 50 km or less. Therefore, RCM can realistically simulate regional climate features including orographic precipitation, extreme climate events and regional scale climate anomalies. However, RCM is relatively more computer intensive than statistical downscaling and the variability in internal parameterizations of RCM provides considerable uncertainty (Fowler *et al.* 2007). Statistical downscaling models are generally separated into three types of approach which can be combined: regression models, weather typing schemes and weather generators (Fowler *et al.* 2007). Multiple linear models, in the regression-based approach are the most applied in downscaling, for example the well known SDSM tool (Wilby *et al.* 2002). These assume a linear relationship between large-scale atmospheric predictors and the response variable. However, several studies have shown that taking into account non-linearity between predictors and the predictand in statistical downscaling can improve the goodness-of-fit (Huth *et al.* 2008) including polynomial regression (Hewitson 1994), recursive partitioning tree (Schnur & Lettenmaier 1998), nearest neighbour (Zorita & von Storch 1999), artificial neural networks (Harpham & Wilby 2005; Khan *et al.* 2006) or generalized additive models (Vrac *et al.* 2007; Salameh *et al.* 2009).

This study introduces a readable statistical downscaling framework of GCMs to derive suitable climatic predictors for bioclimatic models. Conceptually, the future climate modelled by GCM is downscaled to fulfil the spatial and/or temporal requirements for bioclimatic models. Here, the concept was experimented by downscaling the potential future hydro-climatic conditions over South-West France to project the potential future distribution of 13 fish species. An ensemble procedure was performed to track and disentangle different sources of uncertainty in future projections, basing on five GCM, three scenarios and different realization of a single bioclimatic model. This study aims at promoting the link between downscaling and bioclimatic modelling to go ahead in the comprehension of future climate change impacts on biodiversity. Accordingly, the objectives of this study are two-fold:

(i) Disentangling the different sources of uncertainty in projected fish communities (e.g. from GCM, scenarios and different runs of bioclimatic models) and characterizing the spatio-temporal patterns of uncertainty in future projections.

(ii) Exploring the spatio-temporal kinetics of future climate changes impacts on biodiversity, in terms of richness, similarity in species composition between sites and types of species assemblages.

MATERIAL AND METHODS

Study area and data requirements

The study area is the Adour–Garonne drainage basin in south-western France including 120,000 km of flowing waters draining a total area of 116000 km^2. 50 sites characterized by a wide range of hydrological (snow to rainfall dominated regimes) and climatic conditions (of mountainous, continental or oceanic influence) were ensured to include both hydrological, climate and fish data for a mean period of approximately six years, ranging from 1992 to 2000. Daily mean flow data (m^3s^{-1}) were collated from the Hydro2 database maintained by the Ministère de l'Ecologie et du Développement Durable (http://www.hydro.eaufrance.fr/) from 1970 to 2000. Daily time series of temperature (° C) were interpolated by ordinary kriging at the 50 fish sites over the control period based on an exponential covariance distance matrix of 160 local daily climate stations provided by Météo-France. Fish occurrence data (i.e., presence-absence) collated from the Office National de l'Eau et des Milieux Aquatiques (ONEMA), were used for the 13 most prevalent species which were present in more than 30% of sites over the period of survey (Table 1). Geomorphological data were also collated from

the ONEMA and described the physical catchments such as their distance from the river source (km), their drainage area (km²), longitude (degree EW) and latitude (degree NS), altitude (m), slope (%), river width (m) and depth (m).

Table 1: Description and prevalence of studied species

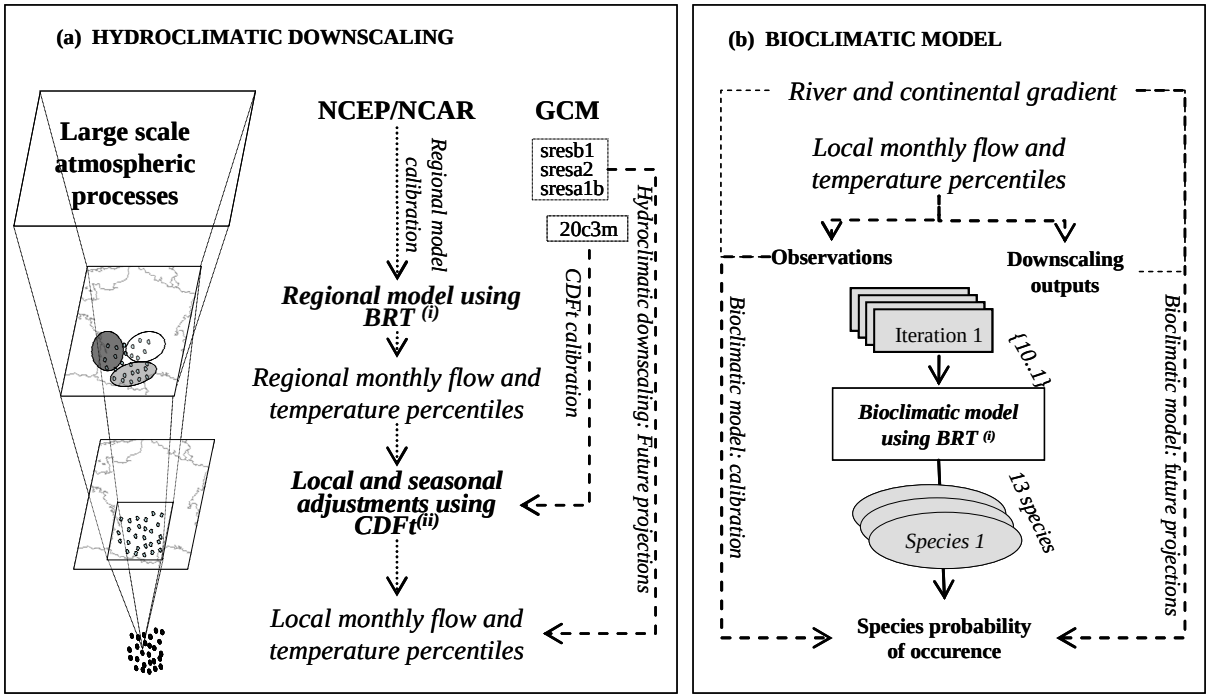
Species name	Common name	Code	Prevalence
Perca fluviatilis	Perch	Pef	0.30
Chondrostoma toxostoma	Soufie	Cht	0.30
Leuciscus leuciscus	Dace	Lel	0.36
Lepomis gibbosus	Pumpkinseed	Leg	0.36
Salmo trutta fario	Brown trout	Sat	0.49
Anguilla anguilla	European eel	Ana	0.51
Alburnus alburnus	Bleak	Ala	0.54
Barbatula barbatula	Stone loach	Bab	0.55
Barbus barbus	Barbel	Bar	0.59
Rutilus rutilus	Roach	Rur	0.62
Phoxinus phoxinus	Minnow	Php	0.63
Leuciscus cephalus	Chub	Lec	0.69
Gobio gobio	Gudgeon	Gog	0.77

Eleven atmospheric variables were used to characterize large spatial-scale atmospheric processes controlling the regional hydro-climatic variability. Monthly reanalysis data from 1970 to 2000 were collated from National Centers for Environmental Prediction and the National Center for Atmospheric Research (NCEP/NCAR; Kalnay *et al.* 1996), to characterize the long-term observations for the 11 atmospheric variables over the region of study. Their equivalent monthly data from five GCMs were downloaded from the Intergovernmental Panel on Climate Change (IPCC) website at https://esg.llnl.gov:8443/index.jsp. Thus NCEP/NCAR reanalysis were used as observed predictors to calibrate the statistical downscaling framework, while GCM data were used as predictors to project the future regional hydro-climatic conditions.

Four scenarios were considered for each GCM to highlight the historical (20c3m) and future potential scenarios, namely A2, A1B, B1, as reported from the Special Report on Emission Scenarios (SRES; Pachauri & Reisinger 2007). Under scenario 20c3m, the five GCM are running with increasing greenhouse gases emissions as observed through the 20th century. The A2 scenario is based on a very heterogeneous world with continuously increasing global population and regionally oriented economic growth that is more fragmented and slower than in other storylines. The A1B storyline is based on a future world of very rapid economic growth, global population that peaks in mid-century and declines

thereafter, and rapid introduction of new and more efficient technologies, with the development balanced across energy sources. The B1 scenario is based on a convergent world with the same global population as in the A1B storyline but with rapid changes in economic structures toward a service and information economy, with reductions in material intensity, and the introduction of clean and resource-efficient technologies.

Both NCEP/NCAR and GCM atmospheric variables were interpolated at the 50 sites using bilinear interpolation, and standardized for GCM based on scenario 20c3m. The 11 selected atmospheric variables then were synthesized into four atmospheric processes related to pressure, temperature, precipitation and shortwave radiation, which characterized key atmospheric controls on the regional hydrology and temperature. This was done by grouping the 11 atmospheric variables within the four atmospheric processes using Hierarchical Ascending Clustering (HAC). Then for each group of variables related to a given atmospheric process, the first Principal Component Analysis (PCs) axis was extracted, which summarised more than 90% of the variance for each process.



(i): boosted regression trees

(ii): Cumulative distribution function transformation

Fig. 1. Structure of the hydro-climatic-ecological (HCE) model-chain built upon downscaling models to model the hydro-climatic variability at the 50 local sites of study (a), coupled to a distribution models to simulate the fish occurrence for 13 species (b).

Hydro-climatic statistical downscaling

The hydro-climatic statistical downscaling was developed to provide reliable predictors for the bioclimatic models (Fig. 1a). Thus, the projected low, median and high flow and temperature conditions, characterized by percentiles 10%, 50% and 90%, were optimized for each of the three most important season in the life cycle of most streamwater fish species: (i) the winter season, from October to February, commonly defined as a period of low activity for fish; (ii) the spawning season, from March to June, encompasses the major part of the reproduction time of most fish (except for the brown trout); (iii) the growth period, from July to September, during which fish actively feed. The downscaling process was performed separately for the temperature and hydrology and was based on two successive modelling components. The regional component project the regional hydro-climatic variability from the four atmospheric processes using the boosted regression trees (Elith *et al.* 2008). Then the local component adjusted the regional projections, seasonally to each of the 50 sites of study.

The regional hydro-climatic conditions were defined as the clustering of the 50 sites within five hydrological and four temperature regions using HAC with Ward criterion. Thus a single regional downscaling model was built for each of the three percentiles according of each of the five hydrological and four temperature regions i.e. overall 27 boosted regression tree models were calibrated. The four atmospheric PCs were used to downscale the hydrological variability, whereas the temperature and pressure PCs were used only to project the monthly local temperature. Two additional predictors related to the sin and cosin values of the 12 month were included as predictors for both temperature and hydrology models. They were shown to improve the seasonal stability of the downscaled projections. The local component of the downscaling model corrected the statistical bias in the regional projections to each station individually using the “Cumulative Distribution Function - transform” approach (CDFt, Michelangeli *et al.* 2009). With CDFt, the probability distribution of downscaled projections was adjusted to that of observations, individually for each quartile of each biological season of each station.

Future hydro-climatic projections from the five GCMs were performed based on a cross-validation procedure using two temporally independent periods from the control period of approximately 15 years, denoted *a* and *b*. Those two periods were successively used to calibrate the regional downscaling from NCEP/NCAR reanalyses (e.g. period *a* from 1970 to 1985) and setting the CDF-t parameters (e.g. period *b* ranging from 1986 to 2000) from the GCMs projected on a. Consequently two combinations of downscaling models were

generated, (i.e., *ab* and *ba*) to project future monthly hydro-climatic conditions, which were averaged afterward to build a single monthly time series from 2005 to 2100. The hydro-climatic downscaling models projections were validated in Tisseuil *et al.* (in prep) based on observed data records.

Bioclimatic models of fish species distribution

The bioclimatic model was used to project the future distribution of each of the 13 fish species from the downscaled hydro-climatic projections (Fig. 1b). For each biological season the monthly hydro-climatic percentiles 10, 50 and 90% (P10, P50 and P90) were averaged over the corresponding season. The seasonal hydro-climatic variability (VAR) was defined as the difference between the seasonal P90 and P10 hydro-climatic conditions, characterizing the amplitude between low and high hydro-climatic conditions. The seasonal flow conditions were divided by the median discharge of each site, computed for the entire period of observed flow records (Cattanéo 2005), to highlight the magnitude related to the overall median flow conditions. Two geomorphological predictors were derived as the first two axis of another PCA applied onto the standardized variables related to catchment characteristics, namely the distance from the source, catchment size, longitude and latitude, altitude, slope, river width and depth. The first PCA axis (63 % of the total variance) characterized the position of the 50 sites along the river gradient whereas the second one (16 % of the total variance) described a SW–NE continental gradient.

Thirteen species-specific boosted regression trees were calibrated from observations, to model the fish species probability of occurrence at each annual site (approximately 50 sites × 6 years) from the two geomorphological predictors, i.e. the river and continental gradient, and the 24 seasonal hydro-climatic predictors, i.e. for both hydrology and temperature, four statistics (P10, P50, P90 and VAR) were considered for each of the three biological seasons (winter, spawning, growth). A binomial distribution of errors was assumed and the probability of species occurrence was related to the predictors via a logistic link function. The model was calibrated using randomly sampled 50% of the full reanalysis dataset, while the remaining 50% data were used during the boosting process to estimate the loss deviance and thus the optimal number of trees to grow. This calibration step was repeated 10 times to introduce randomness in the simulations and future projections were performed using the downscaled hydro-climatic conditions as predictors into the bioclimatic models. The bioclimatic model

projections driven by the downscaled hydro-climatic conditions were validated in Tisseuil *et al.* (in prep) based on observed data records.

Changes in fish biodiversity

From the probability of occurrence projected by a given GCM under the scenario 20c3m, the projected occurrence of a given species (i.e. presence-absence) was derived by identifying the optimal threshold which maximized the number of true presences and absences from the observed data, using a receiver operating characteristic (ROC) plot (Fielding & Bell 1997; Pearce & Ferrier 2000). For each projected species over scenario 20c3m, the errors of commission (falsely predicted presence) and of omission (falsely predicted absence) were likely the same for all species and did not exceed 20%. Furthermore, the residuals from each model were not significantly spatially correlated (Mantel test, $p>0.25$), so that the consistency for the 'predict first, assemble later' strategy used to assess the potential future change on fish assemblages was strengthened.

The spatial and temporal changes in fish biodiversity were highlighted at each site and year using three indices: species richness (α-diversity), similarity in species composition between sites (β-diversity) and fish structure (species composition). The α-diversity was calculated as the number of species whereas the β-diversity was estimated as the mean similarity in species composition (i.e. presence or absence of the 13 species) of site *i* with the *n-i* others sites (n = 50 sites). Based on the dissimilarity measure of Jaccard (1901), the similarity between sites ranged between 0 and 1, respectively from sites poorly to highly similar in species composition.

A redundancy analysis (RDA) was performed to describe how species were structured along the river gradient according to years. RDA likely extends the properties of multiple regression to a matrix of multiple response variables (Y), explained by a matrix of predictors (X). In this study Y was characterized, in columns, by the projected probability of occurrence for the 13 species, whereas X was defined by the position of sites along the river gradient as well as the different years as two quantitative predictors. The first axe of the RDA, namely the weighted averaged score, was extracted and highlighted more than 60% of the fish community structure constrained by the river gradient and the inter-annual variability.

Disentangling the variability in projected fish biodiversity indices

The projected variability of each biodiversity index (i.e. α-diversity, β-diversity and species composition) was disentangled according to five factors: the river gradient, the inter-annual signal of global change, the five GCM, the three future scenarios and the ten bioclimatic model runs. Hierarchical partitioning was applied to evaluate the independent and joint contribution from each of these five factors (predictors) to the projected variance (adjusted R^2) of each biodiversity index (response). Hierarchical partitioning is classically built upon generalized linear models, by considering linear combinations between the predictors and the response (Chevan & Sutherland 1991). In this study, hierarchical partitioning was extended to generalized additive models to take into account for non-linear relationships between the response (e.g. α-diversity) and the spatial (i.e. the river gradient) and temporal (i.e. inter-annual signal of global change) factors.

For each biodiversity index, the spatio-temporal variation between the different GCM, scenarios and bioclimatic model runs was evaluated using the coefficient of variation (CV). Let us consider the spatio-temporal variation between GCM in the projected α-diversity projected. At each annual site, the projected α-diversity was first averaged according to each GCM then the CV was calculated as the ratio between the standard deviation and the averaged projected α-diversity. CV is dimensionless, however the higher CV, the higher variation between GCM.

Smoothing spatio-temporal patterns

The spatio-temporal patterns of each projected biodiversity index, as well as their respective sources of variation related to GCM, scenarios and boosted regression tree models runs, were highlighted using generalized additive models. That is, the river gradient and the inter-annual variations were used as two predictors into the GAM to regress the process of interest (e.g. α-diversity) using thin plate regression splines of low dimension, to smooth global spatio-temporal patterns in the process. Additionally, the CDFt method was applied to adjust the probability distribution of smoothed results to that of the process, especially to intensify the spatio-temporal contrasts in the process.

RESULTS

Variability in projected streamwater fish biodiversity

Changes in streamwater biodiversity were essentially explained by spatio-temporal patterns since the independent contribution from the river gradient and the inter-annual signal of global change respectively approximated 60% and 2% of the total explained variance (Fig. 2). The variability related to the five GCM, the three future scenarios and the ten bioclimatic model runs explained less than 2% of the total variability in the projections. Additional results, not presented here, showed that the projected variability explained by each factor was likely independent from each others since their joint contribution was lower than 1%.

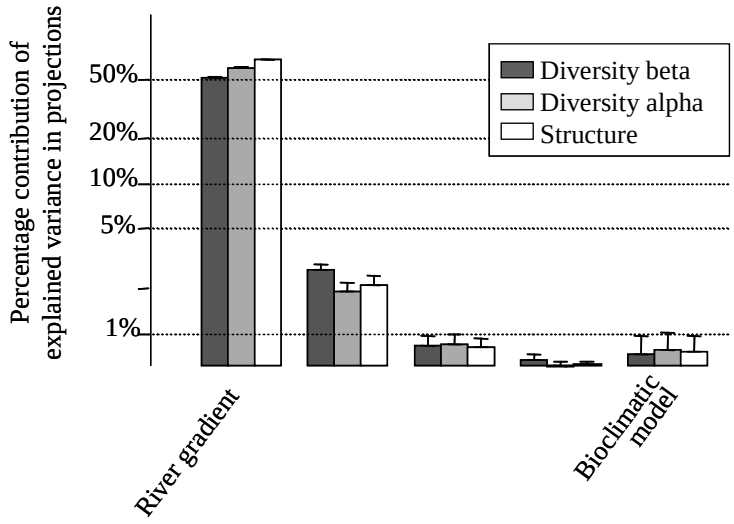


Fig. 2. Individual percentage contribution in projected future fish β-diversity (dark grey), α-diversity (light grey) and structure (white) explained by the river gradient, the inter-annual variability from 2005 to 2100, five general circulation models (GCM), three greenhouse gaze emission scenarios (SRES) and ten bioclimatic model runs based on boosted regression tree.

The overall spatial and temporal patterns of uncertainty in the projected α-diversity (Fig. 3a), β-diversity (Fig. 3b) and community structure (Fig. 3c) were investigated. Uncertainty in the projected α-diversity and community structure was consistently explained by the river gradient and inter-annual trends, approximately 30% of the explained CV variability (Fig. 3a, c). More specifically, uncertainty was likely higher in the middle river gradient and progressively decreased with time, excepted for that from α-diversity which seemed to increase during the second mid-century (Fig. 3a). Uncertainty in the projected β-diversity did not exhibit particular spatio-temporal patterns as the river gradient and the inter-annual variability approximately explained only 6% of CV variability (Fig. 3b).

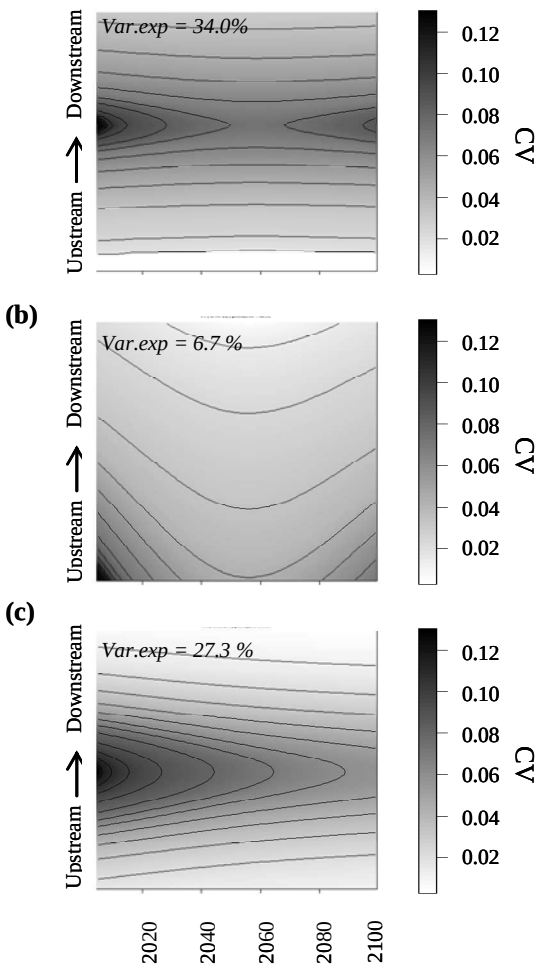


Fig. 3. Spatial (river gradient in y-axis) and temporal (inter-annual variability from 2005 to 2100 in x-axis) patterns of uncertainty in the projected streamwater fish α-diversity (a) , β-diversity (b), and structure (c). Measure of uncertainty was based on the coefficient of variation (CV). The percentage variance in uncertainty explained by the spatial and temporal dimensions (Var.exp) is shown at the top-left of each representation.

Changes in projected streamwater fish biodiversity

The spatio-temporal changes in projected α-diversity, β-diversity and fish structure showed similar patterns (Fig. 3). Globally, the potential number of species was likely to increase all along the river gradient over the 21st century, of approximately 12% in 2045-2055 and 25% in 2090-2100 (Fig. 3a; red colour gradient). This increasing α-diversity over 2015-2100 was more important in the upstream than in the downstream, in average respectively +22 and +5% (Fig. 3a). Similarly, the spatio-temporal trends in the projected β-diversity showed that fish composition was homogenizing over the region as the similarity between sites was continuously increasing with time (Fig. 3b). In comparison to the averaged conditions between 2005 and 2015, the global β-diversity was likely to increase of approximately 6% in 2045-2055 and 11 % in 2090-2100. This tendency of homogenisation was particularly highlighted in the downstream (+10%) whereas fish communities in the upstream were likely to diversify (-2%) (Fig. 3b). Future hydrological and thermal habitats suitability was more favourable to the expansion of warm species over the major part of the river gradient, more particularly toward the upper gradient (Fig. 3c, from green to orange colour gradient).

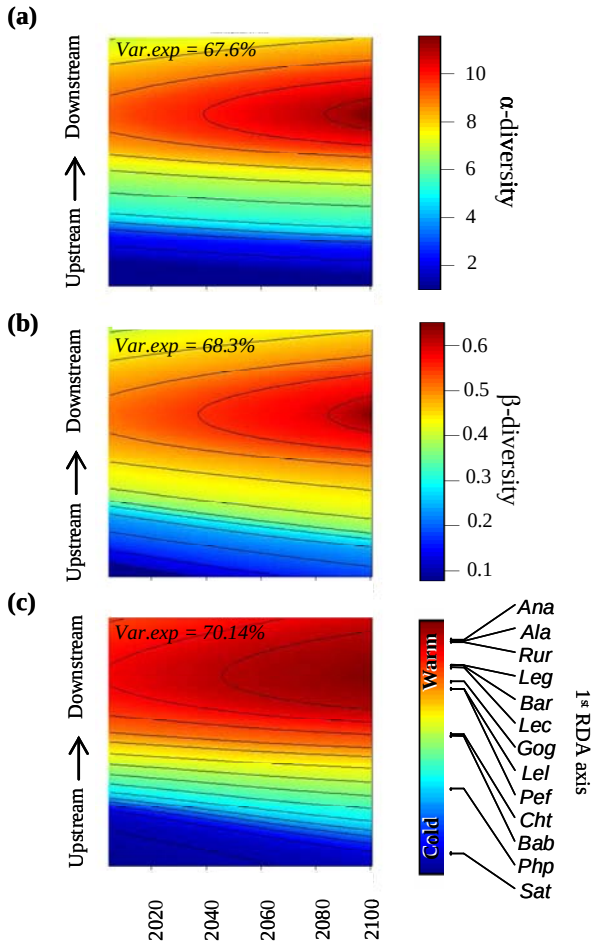


Fig. 4. Spatial (river gradient in y-axis) and temporal (inter-annual variability from 2005 to 2100 in x-axis) patterns in future projected streamwater fish α-diversity (a) , β-diversity (b), and structure (c). The percentage variance in the projections explained by the spatial and temporal dimensions (Var.exp) is shown at the top-left of each representation.

DISCUSSION

REFERENCES

Araujo, M.B., Thuiller, W. & Pearson, R.G. (2006). Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography*, 33, 1712-1728.

Beaumont, L.J., Hughes, L. & Pitman, A.J. (2008). Why is the choice of future climate scenarios for species distribution modelling important? *Ecology Letters*, 11, 1135-1146.

Beaumont, L.J., Pitman, A.J., Poulsen, M. & Hughes, L. (2007). Where will species go? Incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology*, 13, 1368-1385.

Buisson, L., Thuiller, W., Casajus, N., Lek, S. & Grenouillet, G. (2009). Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, 9999.

Cattanéo, F. (2005). Does hydrology constrain the structure of fish assemblages in French streams? Local scale analysis. *Archiv für Hydrobiologie*, 164, 345-365.

Chevan, A. & Sutherland, M. (1991). Hierarchical Partitioning. *American Statistician*, 45, 90-96.

Elith, J., Leathwick, J.R. & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77, 802-813.

Fielding, A.H. & Bell, J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24, 38-49.

Fowler, H.J., Blenkinsop, S. & Tebaldi, C. (2007). Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 27, 1547-1578.

Harpham, C. & Wilby, R.L. (2005). Multi-site downscaling of heavy daily precipitation occurrence and amounts. *Journal of Hydrology*, 312, 235-255.

Hewitson, B. (1994). Regional Climates in the GISS General Circulation Model: Surface Air Temperature. *Journal of Climate*, 7, 283-303.

Huth, R., Kliegrova, S. & Metelka, L. (2008). Non-linearity in statistical downscaling: does it bring an improvement for daily temperature in Europe? *International Journal of Climatology*, 28, 465-477.

Jaccard, P. (1901). Etude comparative de la distribution florale dans une portion des Alpes et des Jura.[Comparative study of the distribution of flora in a region of the Alps and the Jura] Bull. *Soc. Vaudoise Sci. Nat*, 37, 547–549.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. & Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77, 437-471.

Khan, M.S., Coulibaly, P. & Dibike, Y. (2006). Uncertainty analysis of statistical downscaling methods using Canadian Global Climate Model predictors. *Hydrological Processes*, 20, 3085-3104.

Michelangeli, P.A., Vrac, M. & Loukos, H. (2009). Probabilistic downscaling approaches: Application to wind cumulative distribution functions. *Geophysical Research Letters*, 36.

Mika, A.M., Weiss, R.M., Olfert, O., Hallett, R.H. & Newman, J.A. (2008). Will climate change be beneficial or detrimental to the invasive swede midge in North America? Contrasting predictions using climate projections from different general circulation models. *Global Change Biology*, 14, 1721-1733.

Morin, X., Augspurger, C. & Chuine, I. (2007). Process-based modeling of species' distributions: What limits temperate tree species' range boundaries? *Ecology*, 88, 2280-2291.

Osborne, P.E. & Suarez-Seoane, S. (2007). Identifying core areas in a species' range using temporal suitability analysis: An example using little bustards Tetrax tetrax L. in spain. *Biodiversity and Conservation*, 16, 3505-3518.

Pachauri, R.K. & Reisinger, A. (2007). *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland.

Pearce, J. & Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133, 225-245.

Salameh, T., Drobinski, P., Vrac, M. & Naveau, P. (2009). Statistical downscaling of near-surface wind over complex terrain in southern france. *Meteorology and Atmospheric Physics*, 103, 253-265.

Schnur, R. & Lettenmaier, D.P. (1998). A case study of statistical downscaling in Australia using weather classification by recursive partitioning. *Journal of Hydrology*, 213, 362-379.

Thuiller, W. (2004). Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, 10, 2020-2027.

Thuiller, W. (2007). Biodiversity - Climate change and the ecologist. *Nature*, 448, 550-552.

Tuck, G., Glendining, M.J., Smith, P., House, J.I. & Wattenbach, M. (2006). The potential distribution of bioenergy crops in Europe under present and future climate. *Biomass & Bioenergy*, 30, 183-197.

Vrac, M., Marbaix, P., Paillard, D. & Naveau, P. (2007). Non-linear statistical downscaling of present and LGM precipitation and temperatures over Europe, Clim. *Past*, 3, 669-682.

Wilby, R.L., Dawson, C.W. & Barrow, E.M. (2002). sdsm-a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling and Software*, 17, 145-157.

Zorita, E. & von Storch, H. (1999). The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate*, 12, 2474-2489.

Zurell, D., Jeltsch, F., Dormann, C.F. & Schroder, B. (2009). Static species distribution models in dynamically changing systems: how good can predictions really be? *Ecography*, 32, 733-744.

MODELLING THE IMPACT OF CLIMATE CHANGE ON FRESHWATER ECOSYSTEMS USING DOWNSCALING APPROACHES

SUMMARY:

This thesis aimed at assessing the impact of global change on freshwater ecosystems during the 21st century in the Adour Garonne area (SW France). A downscaling approach was developed linking techniques from climate, hydro-chemical and ecological sciences. The main results suggest an increase of high flows in winter as well as more severe low flows in summer. Nitrogen concentrations and thermophile fish species distribution may also increase. Reducing green house gas emissions and modifying agricultural practices (e.g reducing nitrate fertilizers) could reduce the intensity of ecological disturbances. This study is an original contribution to the management of future hydrological and ecological resources.

KEYWORDS: climate change, species assemblages, species distribution, environmental gradients, uncertainty, statistical modelling, mechanistic modelling, ecological niche, stream-water fish, future projections, nitrates, hydrological regimes, downscaling, spatiao-temporal variability, climate scenarios.

AUTEUR : Clément Tisseuil

TITRE : Modéliser l'impact du changement climatique sur les écosystèmes aquatiques par approche de downscaling.

DIRECTEURS DE THÈSE : Sovan Lek, Andrew J. Wade, Mathieu Vrac

LIEU ET DATE DE SOUTENANCE : Toulouse, vendredi 4 Décembre 2009

RÉSUMÉ :
L'objectif de ma thèse était d'évaluer l'impact du changement global sur les écosystèmes aquatiques au cours du 21ème siècle, dans le bassin Adour Garonne (S-O France). Une approche de « downscaling » a été développée à l'interface entre les sciences du climat, de l'hydro-chimie et de l'écologie. Les résultats suggèrent une augmentation globale des débits hivernaux et une diminution des débits d'étiage. Les concentrations en nitrate ainsi que la distribution des espèces de poisson thermophiles pourraient également augmenter. Toutefois, des scénarios de diminution des gaz à effet de serre ainsi qu'une modification des pratiques agricoles (ex. diminution des fertilisants azotés) pourraient limiter l'intensité des perturbations écologiques. Cette thèse offre une contribution originale, notamment pour la gestion future des ressources hydriques et écologiques.

TITRE et résumé en anglais au recto de la dernière page

MOTS CLÉS : assemblages d'espèces, changement climatique, distribution d'espèces, gradients environnementaux, incertitudes, modélisation statistique, modélisation mécanistique, niche écologique, poissons d'eau douce, projections futures, nitrates, régime hydrologique, régionalisation, downscaling, variabilité spatio-temporelle, scénarios climatiques.

DISCIPLINE : Ecologie

ADRESSE DU LABORATOIRE DE RATTACHEMENT : Laboratoire Evolution & Diversité Biologique Bâtiment 4R3 Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 4, France.