

Prédiction de la répartition des espèces de poissons dans le réseau hydrographique de la Seine : comparaison des performances de différents modèles existants

Jérôme Belliard^{1*}, Léna Marollier¹, Amandine Zahm¹, Evelyne Tales¹

¹Irstea UR HBAN, 1 rue Pierre-Gilles de Gennes - CS 10030, 92761 Antony

jerome.belliard@irstea.fr

* personne à contacter

Résumé

Les modèles de distribution d'espèces sont des outils numériques qui permettent de prédire la présence/absence d'une espèce en fonction des conditions d'environnement. Nous avons testé, dans le contexte du bassin de la Seine, la capacité prédictive de trois séries de modèles de distribution d'espèces de poissons développés précédemment et qui ont été calibrés sur des domaines géographiques variés : bassin de la Seine (modèles PIREN), France entière (modèles IPR) et Europe (modèles IPR+). Par ailleurs les trois séries de modèles différaient selon qu'ils prenaient en compte, via les variables prédictives, l'influence explicite des activités et aménagements anthropiques (Modèles PIREN) ou non (modèles IPR et IPR+). Ces tests ont été menés sur un jeu de données de 269 sites distribués sur l'ensemble du bassin sur lesquels des pêches électriques ont permis de qualifier la présence/absence réelle des espèces de poissons.

Pour les modèles PIREN et IPR, on observe une concordance étroite entre le nombre d'occurrences observées et le nombre d'occurrences prédites pour l'ensemble des 269 sites test. Les modèles IPR+ tendent au contraire à sous-estimer l'occurrence des espèces notamment les plus abondantes.

Différents indicateurs évaluant la performance de prédictions à l'échelle du site (AUC, indice de Kappa...) ont été calculés. Ils montrent que les capacités prédictives des modèles peuvent varier fortement d'une espèce à l'autre. Les modèles se révèlent tout à fait satisfaisant pour 6 à 9 espèces (selon le critère d'évaluation et la série de modèles considérés) sur les 19 espèces testées. À l'inverse, les capacités prédictives des modèles se révèlent médiocres voire mauvaises pour 6 espèces. Les indicateurs de qualité des modèles semblent légèrement meilleurs pour les modèles PIREN sans que cette différence soit statistiquement significative.

Des tests complémentaires ont montré que le risque de non-détection des espèces dans notre jeu de données test (espèce présente sur le site mais non capturée lors de l'échantillonnage) était susceptible d'affecter nos résultats.

Par ailleurs contrairement à nos hypothèses de départ, nous n'avons pas détecté de biais dans la prédiction des modèles pour les cours d'eau de la craie sur lesquels nous attendions une surreprésentation des espèces d'eau froide.

1 Introduction

Les modèles de distribution d'espèces (SDM – species distribution models) (appelés aussi parfois « modèles de niche » ou « modèles d'habitat ») sont des outils numériques qui mettent en relation les conditions d'environnement avec la présence/absence d'une espèce (voire pour certains d'entre eux de son abondance) (Guisan & Zimmerman, 2000; Elith & Leathwick, 2009). Leur développement et leur usage ont connu un essor

spectaculaire au cours des 25 dernières années (Guisan & Thuiller, 2005). Des recommandations méthodologiques touchant notamment au choix des outils de modélisation, à l'identification des variables prédictives les plus pertinentes et à l'évaluation des performances des prédictions, au regard des objectifs assignés aux modèles se sont progressivement structurées (Austin, 2007).

L'utilisation des modèles de distribution d'espèces s'est notamment imposée autour d'objectifs aussi divers que (1) identifier et quantifier l'influence des différents facteurs qui contrôlent la répartition des espèces, (2) déterminer les territoires les plus propices à la conservation d'une espèce, (3) prédire la distribution spatiale d'une espèce sur une zone géographique peu échantillonnée, (4) reconstituer la composition historique des assemblages d'espèces à partir de la connaissance de l'environnement passé, ou (5) prévoir les conséquences de changements environnementaux sur la répartition des espèces. Ce dernier point a généré une littérature considérable en lien notamment avec la question des impacts à venir du changement climatique sur la biodiversité (Pearson & Dawson, 2003).

Depuis une quinzaine d'années, de nombreux travaux ont donné lieu à la mise au point de modèles de distribution pour les poissons des cours d'eau européens pour des contextes territoriaux variés allant de régions relativement restreintes (voir par exemple Radinger et al. (2015) pour le bassin de la rivière Treene [760 km² - nord de l'Allemagne]) à l'ensemble de l'Europe (Logez et al., 2012). Ces modèles reposent sur des méthodes de modélisation très variées : GLM [Generalized Linear Models] (Oberdorff et al., 2001), GAM [Generalized Additive Models] (Buisson et al., 2008), BRT [Boosted Regression Trees] (Radinger et al., 2015), voire combinaison de méthodes (Buisson et al., 2010). Potentiellement, ces modèles peuvent s'appliquer au bassin de la Seine dans la mesure où ils concernent des espèces présentes sur ce territoire. Cependant, la plupart des publications ne fournissent pas les éléments de base (paramètres des modèles / jeux de données d'origine) qui permettraient de mettre en œuvre les modèles développés.

Le travail mené en 2016 a consisté à tester et comparer 3 séries de modèles déjà développés et calibrés sur des entités spatiales d'étendue variable (bassin de la Seine, France, Europe) afin d'évaluer leurs capacités à prédire la répartition actuelle des espèces dans le réseau hydrographique de la Seine. Outre l'évaluation des performances des modèles et leur comparaison, nous nous sommes également attachés à considérer certaines questions susceptibles d'affecter la qualité des modèles et notamment (1) les biais potentiels générés par les problèmes d'échantillonnage (non détection des espèces), (2) la question de la non prise en compte de la température de l'eau.

2 Matériel et Méthode

Les modèles disponibles

Le travail de comparaison a porté sur 3 séries de modèles : (1) des modèles développés précédemment dans le cadre des travaux du PIREN-Seine (Belliard et al., 2010), (2) des modèles développés par Oberdorff et al. (2001) utilisés dans le calcul de l'indice poisson rivière (IPR), et (3) des modèles développés par Logez et al. (2012), disponibles avec le calcul de l'indice poisson rivière révisé (IPR+). Dans la suite, et par souci de simplification, nous désignerons ces 3 ensembles par les appellations suivantes : « modèles PIREN », « modèles IPR » et « modèles IPR+ ».

Ces 3 séries de modèles reposent sur les mêmes approches statistiques, à savoir des modèles de régression logistique qui permettent de relier la probabilité de présence d'une espèce aux caractéristiques environnementales du cours d'eau et de son bassin versant. En revanche, ils diffèrent par un certain nombre d'éléments susceptibles de conditionner leurs capacités prédictives et leur domaine de mise en œuvre :

- L'emprise spatiale des données ayant servi à caler les modèles varie fortement : elle concerne le bassin de la Seine pour les modèles PIREN, la France entière pour les modèles IPR et 14 pays européens pour les modèles IPR+
- La nature des sites de calibration des modèles varie également selon les cas : les modèles IPR et IPR+ ont été calés à partir de sites sélectionnés pour leur relativement faible impact anthropique ; à l'inverse le jeu de calibration des modèles PIREN intègre à la fois des sites peu anthropisés et des sites fortement influencés par les activités humaines.

- Les variables environnementales utilisées comme variables d'entrée des modèles varient selon les trois ensembles (Tableau I). En particulier, les modèles PIREN se distinguent des modèles IPR et IPR+ par le fait qu'ils intègrent des variables étroitement dépendantes des activités humaines (variables de qualité d'eau, navigation, proportion de plans d'eau sur le bassin).

Tableau 1 : Liste des variables environnementales utilisées comme variables d'entrée des différents ensembles de modèles. *b* : variable intégrée directement dans le modèle. *c* : variable combinée avec d'autres dans un paramètre plus global utilisé en entrée du modèle¹.

		PIREN	IPR	IPR+
Taille et géométrie du lit	Surface du bassin versant	b	c	b,c
	Distance à la source		c	
	Pente	b	c	c
	Largeur		c	
	Profondeur		c	
Contexte hydro-climatique	Altitude		b	
	Précipitations moyennes annuelles sur le bassin			c
	Température moyenne annuelle sur le bassin			c
	Température moyenne de Juillet (station)	b,c	c	b,c
	Température moyenne de Janvier (station)	c	c	c
Qualité de l'eau	Oxygène	b		
	Ammonium	b		
	Chlorophylle			
Aménagement du lit et du bassin	Navigation	b		
	Densité de plans d'eau	b		
Biogéographie	Bassin		c	

Compte tenu de ces particularités, on peut donc supposer que (1) les modèles IPR, et a fortiori les modèles IPR+, en intégrant des espaces géographiques et donc des gradients environnementaux plus étendus, intègrent mieux l'ensemble de la niche écologique des espèces que les modèles PIREN, bâtis sur un domaine beaucoup plus restreint. (2) À l'inverse, les modèles PIREN pourraient se révéler plus performant pour rendre compte de particularités écologiques locales « diluées » lorsque des échelles spatiales plus grandes sont considérées. Ceci semble plus particulièrement vrai pour certains taxons (comme par exemple le chabot ou le goujon) montrant une forte diversification génétique géographique (pouvant conduire à la création de plusieurs espèces sœurs) (Kottelat & Freyhof, 2007) et susceptible d'être accompagnée par des variations géographiques dans les préférences écologiques. Par ailleurs, la nature des jeux de données de calibration des modèles et les variables prédictives utilisées suggèrent que (3) les modèles PIREN seront mieux à même d'intégrer et de rendre compte des effets de l'anthropisation sur la répartition des espèces que les modèles IPR et IPR+.

¹ À titre d'exemple, dans les modèles IPR, la surface de bassin versant et la distance à la source sont combinées au sein d'un même paramètre indicateur de la taille du cours d'eau ; c'est ce paramètre qui est intégré comme variable d'entrée.

Tableau 2 : Liste des espèces modélisées pour les 3 séries de modèles.

	PIREN	IPR	IPR+	Espèces communes aux 3 séries	Présence sur la Seine	Introduite
Ablette	*	*	*	*	*	
Anguille	*	*	*	*	*	
Barbeau	*	*	*	*	*	
Barbeau méridional		*				
Blageon		*	*		*	I
Bouvière	*	*	*	*	*	
Brème bordelière	*	*			*	
Brème commune	*				*	
Brochet	*	*	*	*	*	
Carassins (sp)		*			*	I
Carpe		*			*	I
Chabot	*	*	*	*	*	
Chevesne	*	*	*	*	*	
Epinoche	*	*	*	*	*	
Epinochette	*	*	*	*	*	
Gardon	*	*	*	*	*	
Goujon	*	*	*	*	*	
Grémille	*	*			*	I
Hotu	*	*	*	*	*	I
Loche franche	*	*	*	*	*	
Lote		*	*		*	
Lamproie de	*	*	*	*	*	
Ombre commun		*	*		*	I
Perche	*	*	*	*	*	
Perche soleil	*	*			*	I
Poisson-chat		*			*	I
Rotengle	*	*			*	
Sandre		*			*	I
Saumon		*	*			
Spirlin	*	*	*	*	*	
Tanche	*	*			*	
Toxostome		*			?	I
Truite	*	*	*	*	*	
Vairon	*	*	*	*	*	
Vandoise	*	*	*	*	*	
	25	34	23	19		

Le nombre d'espèces modélisées varie selon les séries de modèles entre 23 (IPR+) et 34 (IPR²), 19 espèces étant communes aux 3 séries. C'est sur cet ensemble de 19 espèces que porteront en priorité les comparaisons et évaluations présentées dans ce rapport.

Le jeu de données test

Les modèles PIREN, IPR et IPR+ de distributions d'espèces ont été testés sur un jeu de données de 269 stations réparties sur l'ensemble du bassin de la Seine et rendant compte de la diversité des cours d'eau (surface de bassin versant allant de 6 à 64685 km²) (Figure 1). Pour ces 269 stations, des données sur la distribution actuelle des différentes espèces ont été obtenues grâce à des échantillonnages par pêche électrique (selon les sites, les échantillonnages disponibles ont été effectués entre 2006 et 2011). Une espèce est considérée comme présente sur un site lorsqu'elle a été capturée et, inversement est considérée comme

² A noter que pour les modèles IPR, les espèces brèmes commune et brème bordelière ont été regroupées sous un même taxon qui a fait l'objet du travail de modélisation.

absente dans le cas où elle n'a pas été capturée. Il convient néanmoins de retenir que, en pratiquant ainsi, on tend à sous évaluer la présence réelle des espèces en raison de risques de non détection (l'espèce est présente mais non capturée lors de l'échantillonnage). Potentiellement, ce biais est plus particulièrement sensible pour les espèces benthiques (plus difficilement capturées) et pour les plus grands cours d'eau pour lesquels seule une portion limitée de la station est réellement échantillonnée (contrairement aux petits cours d'eau pour lesquels l'intégralité de la station est prospectée).

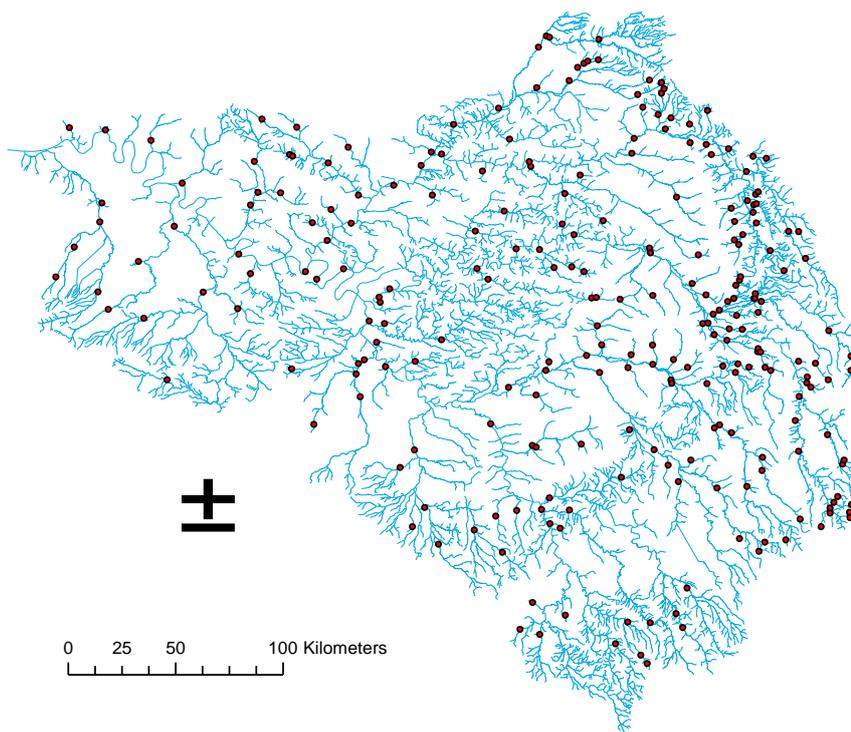


Figure 1 : Localisation des 269 stations utilisées pour tester les modèles de distribution d'espèces.

Pour chacune des 269 stations test, les différentes variables nécessaires au calcul des modèles de distribution d'espèces ont été évaluées soit par mesures directes sur le terrain lors des opérations d'échantillonnages, soit à partir d'outils SIG, soit à partir de modélisations. Les variables climatiques (températures et précipitations), ont été dérivées des données SAFRAN de Météo France. Les variables de qualité d'eau ont été calculées à partir de simulations effectuées à partir du modèle Sénèque en se basant sur l'année hydrologique 2000 et en considérant les moyennes des concentrations simulées pour les décades de juillet (voir Belliard et al. (2010) pour plus de détail). La mise en œuvre des modèles IPR nécessite la connaissance de la largeur et de la profondeur du cours d'eau. Or, ces informations n'étaient pas disponibles pour l'intégralité des 269 stations test. Pour pallier cette lacune, nous avons établi des relations statistiques empiriques entre la surface de bassin versant d'une part et la largeur ou la profondeur du cours d'eau d'autre part, à partir d'un jeu de données complémentaires de sites sur lesquelles ces variables avaient été mesurées sur le terrain (971 mesures disponibles). Ces relations permettent d'expliquer une part importante des variations observées ($r^2 = 0,72$ et $0,88$ respectivement pour la profondeur et la largeur). Nous avons ainsi utilisé ces relations pour estimer ces deux variables largeur et profondeur sur nos 269 sites test. Il est à noter que le fait d'avoir utilisé des estimations des largeurs et profondeurs plutôt que des mesures réelles est susceptible de réduire les performances des modèles qui utilisent ces variables (modèles IPR).

Évaluation de la performance des modèles

L'évaluation de la performance des modèles est basée sur la confrontation entre prédiction des modèles. Il existe différents indicateurs qui permettent de juger des performances des prédictions des modèles, la plupart étant basés sur la constitution de la matrice de confusion (Figure 2). Dans la mesure où les modèles de distribution d'espèces prédisent une probabilité de présence (et non une présence ou une absence), la construction d'une matrice de confusion suppose la définition préalable d'un seuil à partir duquel on considérera que la valeur de probabilité correspond à une présence ou à une absence. Intuitivement, un seuil de 0.5 semble le choix le plus logique pour distinguer les présences des absences, mais cette valeur ne correspond pas nécessairement au seuil permettant d'optimiser les prédictions. On verra par la suite comment ce seuil a été ajusté à chaque cas.

	Présence prédite	Absence prédite
Présence observée	Nombre de vrais positifs (VP)	Nombre de faux négatifs (FN)
Absence observée	Nombre de faux positifs (FP)	Nombre de vrais négatifs (VN)

Figure 2 : Matrice de confusion permettant de classer les cas en fonction de la présence/absence observée et de la présence/absence prédite.

Une fois le seuil de décision choisi et la matrice de confusion construite, le paramètre le plus classique pour juger de l'efficacité des modèles consiste à calculer le taux de bonnes prédictions :

$$\frac{(VP + VN)}{(VP + FN + FP + VN)}$$

Au-delà, la capacité prédictive d'un modèle peut être mesurée par sa spécificité (taux de bonnes prédictions des présences) :

$$\frac{VP}{(VP + FN)}$$

et sa sensibilité (taux de bonnes prédictions des absences) :

$$\frac{VN}{(FP + VN)}$$

Il est à noter que pour chaque série de modèles et chaque espèce, la valeur seuil de probabilité pour la distinction entre présences et absences a été choisie de façon à maximiser la somme de la spécificité et de la sensibilité.

L'indice de Kappa est un indice synthétique plus complexe qui consiste à mesurer l'amélioration des prédictions apportée par le modèle par rapport à une prédiction au hasard :

$$\frac{P_o - P_e}{1 - P_e}$$

où P_o : proportion de concordance observée entre prédictions et observations

et P_e : proportion de concordance entre prédictions et observations due au hasard

Potentiellement, l'indice de Kappa varie de 0 à 1. On considère en général que l'efficacité du modèle devient acceptable pour des valeurs de Kappa supérieures à 0,4.

Un autre indicateur pour évaluer la performance des modèles est basé sur la construction de la courbe ROC (Receiver operating characteristic curve) qui représente la variation de la sensibilité en fonction de (1-spécificité) pour des valeurs seuil variant progressivement entre 0 et 1 (Figure 3). L'aire sous la courbe (AUC = area under the curve) est un indicateur de la qualité du modèle et varie potentiellement entre 0,5 (le modèle n'apporte aucune information sur la présence absence) et 1 (le modèle permet de distinguer parfaitement les absences des présences). On considère en général que les capacités prédictives d'un modèle sont bonnes pour des valeurs d'AUC supérieures à 0.8.

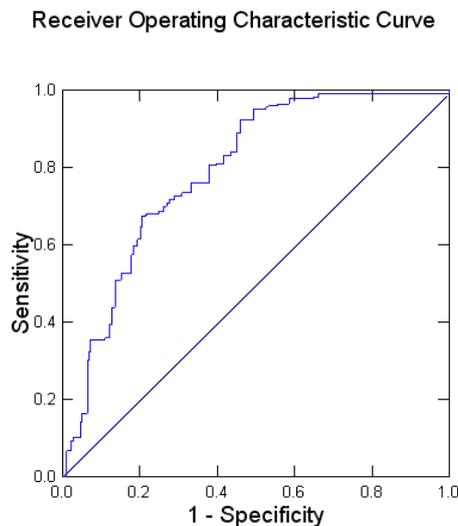


Figure 3 : Exemple de courbe ROC représentant les variations de sensibilité en fonction de (1- spécificité). Les performances prédictives du modèle sont d'autant plus élevées que la courbe s'écarte de la diagonale.

Pour comparer les modèles entre eux nous avons calculé l'ensemble des indicateurs évoqués ci-dessus.

Une autre manière, plus globale d'apprécier la pertinence d'un (ou plusieurs) modèle(s) de distribution d'espèce consiste à confronter le nombre d'occurrences (présences) observées (OCC_{Obs}) et le nombre d'occurrences prédites (OCC_{Pred}) pour un ensemble de sites. Pour un modèle donné (donc une espèce donnée), et un ensemble de sites, le nombre d'occurrences prédites peut être évalué de la manière suivante :

$$OCC_{Pred} = \sum_j P_{ij}$$

Avec P_{ij} correspondant à la probabilité d'observer l'espèce i sur le site j .

Pour une série de modèles portant sur différentes espèces, il est possible d'examiner la relation entre OCC_{Obs} et OCC_{Pred} . Si cette série de modèles fournit des prédictions correctes et non biaisées, on s'attend à ce que les variations entre relation OCC_{Obs} et OCC_{Pred} suivent une relation linéaire de pente 1 et d'intercept (ordonnée à l'origine) 0. Nous avons examiné ces relations entre occurrences observées et occurrences prédites pour les 3 séries de modèles (PIREN, IPR et IPR+) soit pour l'ensemble des sites et l'ensemble des 19 espèces communes, soit sur des sous-ensembles de sites et le cas échéant des sous-ensembles d'espèces.

Ainsi, pour évaluer dans quelle mesure l'échantillonnage (et donc les problèmes éventuels de non détection- des espèces) était susceptible d'affecter la réponse des modèles, nous avons effectué ce type d'analyse en distinguant les sites sur lesquels l'échantillonnage est effectué de manière complète (l'intégralité de la station est échantillonnée par pêche électrique) des sites sur lesquels l'échantillonnage est effectué de manière

partielle (seule une portion, souvent réduite, de la station est échantillonnée). De même, pour évaluer si le comportement thermique des cours d'eau était susceptible d'affecter la réponse des modèles, nous avons effectué ce type d'analyse en distinguant les sites sur substrat géologique crayeux (pour lesquels, globalement, la température de l'eau estivale est très sensiblement plus fraîche que celle de l'air), des sites localisés sur d'autres substrats géologiques (pour lesquels, globalement, la température de l'eau suit plus étroitement celle de l'air). Pour ces dernières analyses, les espèces d'eau froide (température de pont < 14°C) ont été séparées des espèces d'eau plus chaude.

3 Résultats

Évaluation globale des modèles

Les différents indicateurs de qualité d'ajustement des modèles pour les 3 séries de modèles et les différentes espèces sont présentés en détail au tableau 3. Ces résultats montrent que la qualité des prédictions est susceptible de varier fortement d'une espèce à l'autre pour une même série de modèles, et d'un modèle à l'autre pour une espèce donnée.

Selon les espèces, le taux global de bonnes prédictions varie de 42% à 82 % pour les modèles PIREN, de 51% à 80% pour les modèles IPR et de 35% à 80% pour les modèles IPR+. Sur la base des valeurs AUC, les capacités prédictives peuvent être considérées comme bonnes ($AUC > 0.8$) pour 6 modèles PIREN et 5 modèles IPR ou IPR+. Sur base des valeurs de Kappa, 9 modèles IPR peuvent être considérés comme acceptables ($Kappa > 0.4$) contre 8 pour les modèles PIREN et 7 pour les modèles IPR+. Pour une série de modèles donnée, on observe une corrélation forte entre valeur d'AUC et d'indice de Kappa indiquant une cohérence globale entre ces 2 indicateurs de performance des modèles. Cette cohérence souffre néanmoins quelques exceptions (par exemple pour le modèle PIREN du barbeau, la valeur d'AUC est forte [0.87] suggérant une bonne performance du modèle alors que l'indice de Kappa est très faible [0.07]). Ces discordances parfois importantes peuvent conduire à s'interroger sur la méthode employée pour optimiser le seuil de détection entre présence et absence.

Tableau 3 : Résumé des performances des modèles de prédiction de la présence/absence des espèces sur la base de différents critères. N : nombre de sites où l'espèce a été observée, Seuil : seuil de probabilité optimisé pour la distinction entre présence et absence, TBC : taux global de bonne classification, Sens : sensibilité, Spec : spécificité, AUC : aire sous la courbe ROC, Kappa : indice de Kappa. Les nombres en rouge indiquent pour chaque espèce, le modèle apparaissant le plus performant sur la base de l'indice de Kappa.

Espèce	Code	N	PIREN						IPR						IPR+					
			Seuil	TBC	Sens	Spec	AUC	Kappa	Seuil	TBC	Sens	Spec	AUC	Kappa	Seuil	TBC	Sens	Spec	AUC	Kappa
Ablette	ABL	44	0.187	0.822	0.795	0.827	0.841	0.488	0.092	0.781	0.886	0.760	0.836	0.446	0.185	0.784	0.818	0.778	0.846	0.431
Anguille	ANG	94	0.364	0.773	0.798	0.760	0.842	0.528	0.178	0.732	0.745	0.726	0.800	0.444	0.225	0.766	0.777	0.760	0.859	0.510
Barbeau	BAF	49	0.008	0.361	0.918	0.236	0.871	0.067	0.501	0.803	0.816	0.800	0.854	0.482	0.190	0.803	0.735	0.818	0.801	0.455
Bouvière	BOU	22	0.026	0.532	0.818	0.506	0.718	0.094	0.059	0.781	0.818	0.777	0.841	0.290	0.075	0.639	0.909	0.615	0.824	0.179
Brochet	BRO	74	0.354	0.721	0.581	0.774	0.682	0.337	0.247	0.691	0.703	0.687	0.716	0.335	0.185	0.569	0.878	0.451	0.704	0.234
Chabot	CHA	207	0.682	0.714	0.783	0.484	0.630	0.248	0.525	0.725	0.865	0.258	0.555	0.137	0.435	0.606	0.565	0.742	0.649	0.217
Chevesne	CHE	151	0.556	0.736	0.609	0.898	0.816	0.486	0.598	0.729	0.675	0.797	0.766	0.461	0.435	0.721	0.603	0.873	0.789	0.457
Epinoche	EPI	45	0.262	0.639	0.578	0.652	0.627	0.151	0.226	0.691	0.556	0.719	0.667	0.199	0.095	0.346	0.844	0.246	0.487	0.037
Epinochette	EPT	86	0.252	0.647	0.663	0.639	0.685	0.271	0.238	0.569	0.663	0.525	0.625	0.159	0.225	0.654	0.314	0.814	0.555	0.139
Gardon	GAR	134	0.658	0.691	0.604	0.778	0.733	0.382	0.560	0.695	0.597	0.793	0.715	0.390	0.485	0.691	0.597	0.785	0.735	0.382
Goujon	GOU	142	0.405	0.669	0.796	0.528	0.714	0.328	0.760	0.647	0.437	0.882	0.702	0.310	0.535	0.691	0.613	0.780	0.742	0.388
Hotu	HOT	37	0.065	0.762	0.838	0.750	0.867	0.370	0.181	0.777	0.865	0.763	0.876	0.400	0.210	0.818	0.784	0.823	0.851	0.442
Loche franche	LOF	184	0.392	0.781	0.902	0.518	0.774	0.453	0.858	0.773	0.821	0.671	0.781	0.484	0.435	0.680	0.788	0.447	0.611	0.241
Lamproie de planer	LPP	103	0.177	0.483	0.971	0.181	0.529	0.122	0.197	0.513	0.922	0.259	0.587	0.150	0.115	0.494	0.932	0.223	0.551	0.126
Perche	PER	111	0.368	0.751	0.694	0.791	0.773	0.485	0.400	0.751	0.622	0.842	0.769	0.474	0.435	0.743	0.541	0.886	0.760	0.446
Spiralin	SPI	33	0.039	0.706	0.848	0.686	0.855	0.282	0.086	0.654	0.818	0.631	0.750	0.219	0.135	0.695	0.758	0.686	0.761	0.239
Truite	TRF	152	0.628	0.710	0.717	0.701	0.757	0.415	0.310	0.725	0.921	0.470	0.725	0.411	0.725	0.721	0.914	0.470	0.758	0.404
Vairon	VAI	145	0.518	0.740	0.766	0.710	0.795	0.476	0.685	0.680	0.814	0.524	0.686	0.344	0.485	0.632	0.662	0.597	0.628	0.259
Vandoise	VAN	88	0.280	0.755	0.614	0.823	0.771	0.439	0.293	0.706	0.784	0.669	0.772	0.404	0.285	0.647	0.841	0.552	0.708	0.329

On constate par ailleurs qu'il existe une concordance étroite entre les performances des modèles par espèce entre les différentes séries. Ainsi, on observe une corrélation forte entre les séries de modèles prises 2 à 2 à la fois pour les valeurs d'AUC (corrélation de Pearson entre 0.74 et 0.81) et les valeurs d'indice de Kappa (0.47 et 0.80). Ceci suggère que globalement, une espèce bien prédite par un des modèles est en général également bien prédite par les 2 autres modèles, et qu'à l'inverse, une espèce mal prédite par un des modèles est en général également mal prédite par les 2 autres modèles.

Ainsi, les espèces comme l'ablette, l'anguille, le chevesne, la perche ou la truite sont des espèces pour lesquelles la présence/absence est plutôt bien prédite (indice de Kappa > 0.4) quels que soient les modèles considérés (PIREN, IPR ou IPR+). À l'inverse pour la bouvière, le chabot, l'épinoche, l'épinochette, la lamproie de planer et le spirilin, les modèles se révèlent systématiquement non performants (indice de Kappa < 0.3). Ces espèces mal prédites regroupent à la fois des espèces plutôt rares à distribution fragmentée (spirilin et bouvière) et des espèces plus fréquentes mais connues pour présenter des difficultés d'échantillonnage (chabot, épinoche, épinochette, lamproie de planer).

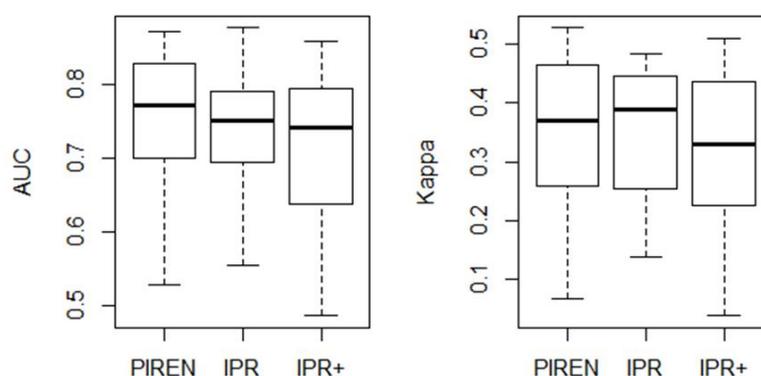


Figure 4 : Distribution des valeurs d'AUC et d'indice de Kappa pour les modèles PIREN, IPR et IPR+.

On constate que les modèles PIREN présentent les valeurs d'indice de Kappa les plus fortes pour 11 espèces suivis des modèles IPR pour 6 espèces et des modèles IPR+ pour 2 espèces seulement (Tableau 3). Toutefois la distribution des valeurs d'AUC et d'indice de Kappa révèlent des différences très ténues entre les 3 séries de modèles qui ne se révèlent pas statistiquement significatives (tests t sur séries de valeurs appariées, $p > 0.05$) (Figure 4).

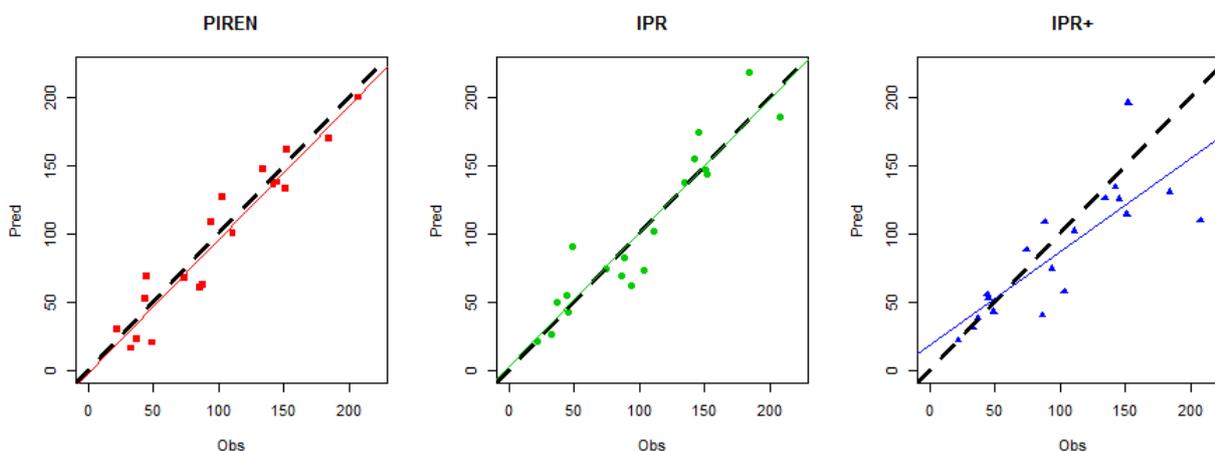


Figure 5 : Relations entre le nombre d'occurrences observées et le nombre d'occurrences prédites sur l'ensemble de 269 stations pour les 19 espèces selon les 3 séries de modèles (PIREN, IPR et IPR+). La droite pointillée noire (droite 1:1) correspond à l'égalité entre observation et prédiction. La droite colorée correspond à l'ajustement linéaire entre occurrences observées et occurrence prédite.

Une autre manière, plus globale, d'évaluer les modèles, consiste à confronter, pour l'ensemble du jeu de données test, le nombre de sites où la présence (ou l'absence) d'une espèce est prédite par rapport au nombre de présences (absences) réellement observées (Figure 5 et tableau 4). Globalement, on constate que pour les modèles PIREN et IPR, la relation entre valeurs observées et valeurs prédites suit la droite (1:1) suggérant que ces modèles permettent de prédire sans biais l'occurrence des espèces dans le réseau hydrographique de la Seine. L'ajustement de cette relation est cependant légèrement meilleur dans le cas des modèles PIREN que dans le cas des modèles IPR (r^2 de 0.91 et 0.87 respectivement). En revanche, pour les modèles IPR+, l'ajustement de la relation entre valeurs observées et valeurs prédites est nettement moins étroit ($r^2 = 0.64$). Par ailleurs dans cette relation linéaire, l'intercept s'écarte significativement de 0 et la pente s'écarte significativement de 1 témoignant de l'existence d'un biais (tendance à la sous-évaluation des espèces les plus communes).

Tableau 4 : Régressions linéaires entre le nombre d'occurrences prédites et le nombre d'occurrences observées sur l'ensemble de 269 stations pour les 19 espèces en fonction des 3 séries de modèles (PIREN, IPR et IPR+). Les chiffres entre parenthèses correspondent aux erreurs standards des estimations des intercepts et des pentes des différentes relations.

	r^2	intercept	pente
PIREN	0.91	-2.36 (\pm 8.23)	0.98 (\pm 0.07)
IPR	0.87	2.40 (\pm 9.99)	0.98 (\pm 0.09)
IPR+	0.64	18.58 (\pm 13.52)	0.68 (\pm 0.12)

Mode d'échantillonnage et effet potentiel de la non-détection des espèces

Afin d'évaluer dans quelles mesures les performances des modèles testés étaient susceptibles d'être affectées par la non-détection des espèces dans les jeux de données test, nous avons examiné, pour les 3 séries de modèles (PIREN, IPR et IPR+), la concordance entre l'occurrence observée des espèces et l'occurrence prédite, en distinguant les sites ayant fait l'objet d'un échantillonnage complet ($n = 186$) d'une part (pour lesquels on suppose a priori que les problèmes de non détection d'espèces sont les plus réduits) et les sites ayant fait l'objet d'un échantillonnage partiel ($n = 83$) d'autre part (pour lesquels on suppose a priori que les problèmes de non détection d'espèces sont les plus importants) (Tableau 5 et figure 6). Notre hypothèse est que pour les sites échantillonnés de manière partielle, on constate une tendance à la sous-représentation globale des espèces (l'occurrence prédite tend à être plus élevée que l'occurrence observée), ou que, a minima, les liens entre occurrences prédites et occurrences observées soient affaiblis (moindre cohérence des prédictions des modèles).

Tableau 5 : Régressions linéaires entre le nombre d'occurrences prédites et le nombre d'occurrences observées sur les 186 sites échantillonnés de manière complète (Compt) et les 83 sites échantillonnés de manière partielle (Part) pour les 19 espèces en fonction des 3 séries de modèles (PIREN, IPR et IPR+). Les chiffres entre parenthèses correspondent aux erreurs standards des estimations des intercepts et des pentes des différentes relations.

	Échantillonnage	r^2	intercept	penne
PIREN	Compt	0.91	-5.14 (± 5.22)	1.03 (± 0.07)
	Part	0.83	1.97 (± 4.11)	0.92 (± 0.10)
IPR	Compt	0.94	-4.64 (± 4.90)	1.08 (± 0.07)
	Part	0.65	4.95 (± 6.34)	0.88 (± 0.15)
IPR+	Compt	0.72	6.42 (± 7.42)	0.70 (± 0.10)
	Part	0.58	7.45 (± 6.53)	0.78 (± 0.15)

Les résultats obtenus sur les modèles IPR+ confirment les analyses globales précédentes mettant en lumière une discordance entre les occurrences prédites et les occurrences observées (sous-évaluation des espèces les plus fréquentes). Pour les modèles PIREN et IPR, on observe au contraire que les points représentant les occurrences prédites en fonction des occurrences observées s'ajustent autour de la droite 1:1 aussi bien pour les sites échantillonnés de manière complète que pour ceux échantillonnés de manière partielle (intercepts et pentes des régressions linéaires cohérentes avec les valeurs 0 et 1 respectivement). On constate néanmoins que pour ces deux séries de modèles, la dispersion des points est plus importante pour le jeu de données de sites échantillonnés de manière partielle que pour les autres sites (r^2 passant de 0.83 à 0.91 pour les modèles PIREN et de 0.65 à 0.94 pour les modèles IPR). Ceci souligne la moindre concordance entre observations et prédictions lorsque les risques de non-détection d'espèces sont plus importants.

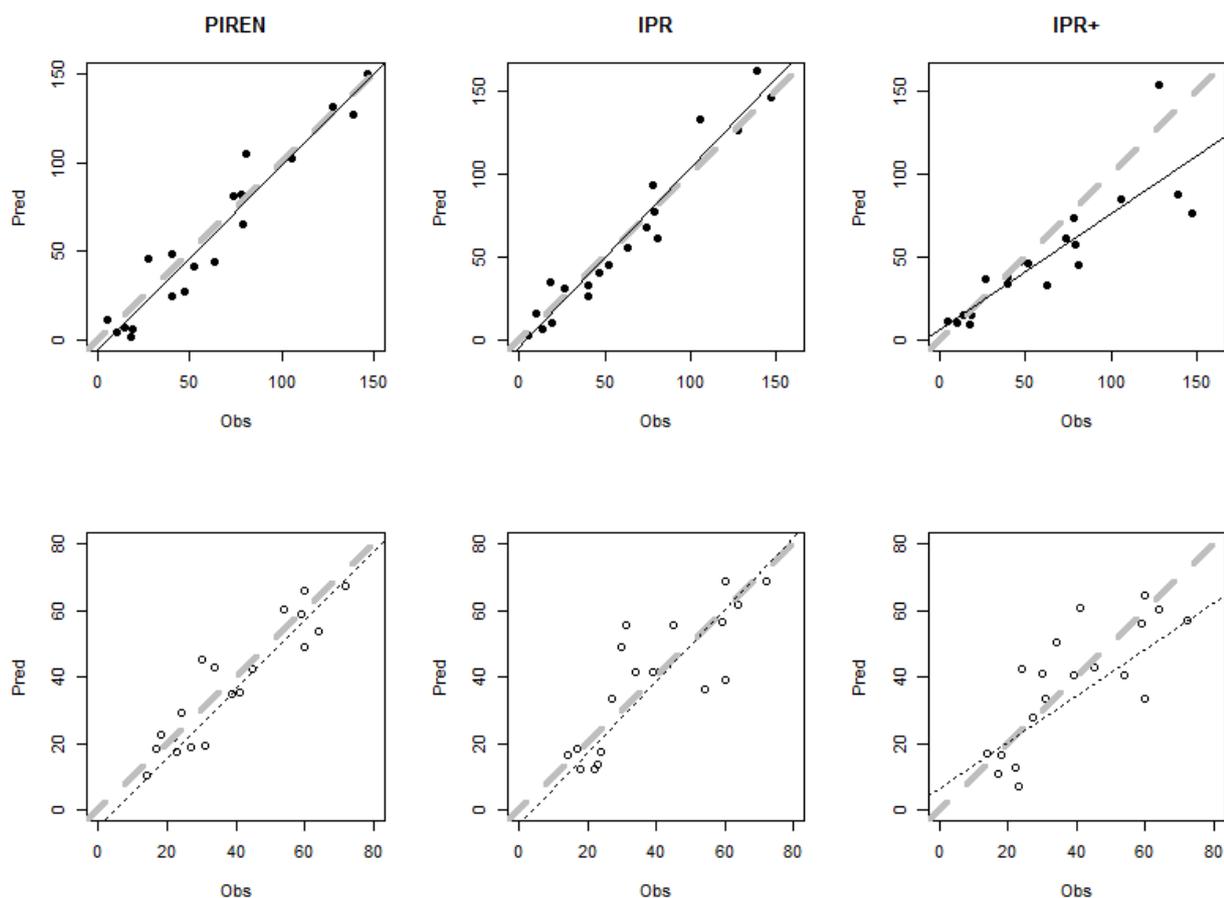


Figure 6 : Relations entre le nombre d'occurrences observées et le nombre d'occurrences prédites des 19 espèces selon les 3 séries de modèles (PIREN, IPR et IPR+). Le cas des sites échantillonnés de manière complète (points noirs ; droite noire pour l'ajustement linéaire entre points) et celui des sites échantillonnés de manière partielle (points blancs ; droite pointillée pour l'ajustement linéaire entre points) ont été distingués. La droite hachurée grise (droite 1:1) correspond à l'égalité entre observation et prédiction.

Influence potentielle de la température de l'eau

Les 3 séries de modèles intègrent des variables thermiques comme variables prédictives de la distribution des espèces de poissons. Par commodité (disponibilité des données), le choix s'est porté sur des données de température de l'air (moyenne de juillet et de janvier) considérées comme des bons proxys de la température de l'eau. Cependant selon le substrat géologique, le comportement thermique des cours d'eau peut présenter des variations importantes, notamment pour des cours d'eau de nappe (particulièrement pour les cours d'eau de la craie) pour lesquels la température de l'eau s'écarte très sensiblement des températures de l'air (température plus élevées en hiver et plus fraîches en été) (Belliard et al., 2010). Ces comportements thermiques particuliers de certains cours d'eau, non intégrés par les données de températures de l'air sont donc susceptibles de générer des biais dans la prédiction des distributions des espèces.

Pour évaluer ce problème, nous avons examiné la cohérence entre occurrences observées et prédites pour les cours d'eau situés sur substrat géologique crayeux (situations où la discordance entre température de l'eau et température de l'air est maximum) d'une part et les cours d'eau sur d'autres substrats géologiques (situations où cette discordance est potentiellement plus atténuée) (Tableau 6 et figure 7). Notre hypothèse était que pour les cours d'eau de la craie, à température de l'eau plus fraîche durant l'été, les espèces d'eau froide seraient surreprésentées (occurrences prédites < occurrences observées) et qu'à l'inverse, les espèces d'eau

chaudeseraient sous-représentées (occurrences prédites > occurrences observées). Les analyses sur les relations entre occurrences observées et prédites ont donc été menées en distinguant les espèces d'eau froide des espèces d'eau plus chaude. Nous avons effectué la distinction entre les deux groupes d'espèces sur la base de leur température de ponte, en considérant comme espèces d'eau froide celles se reproduisant à une température inférieure à 14°C (voir Poulet et al. (2011) pour les données utilisées).

Tableau 5 : Régressions linéaires entre le nombre d'occurrences prédites et le nombre d'occurrences observées sur les 100 sites localisés sur substrat crayeux et les 169 sites échantillonnés sur d'autres substrats géologiques en fonction des 3 séries de modèles (PIREN, IPR et IPR+). Les analyses ont été menées séparément pour les espèces d'eau froides (F) et les espèces d'eau chaude (C). Les chiffres entre parenthèses correspondent aux erreurs standards des estimations des intercepts et des pentes des différentes relations.

	Géologie	Type d'espèce	r ²	intercept	pente
PIREN	Craie	F	0.91	-5.28 (± 5.65)	1.05 (± 0.12)
	Craie	C	0.86	2.99 (± 4.80)	0.90 (± 0.12)
	Autre	F	0.90	-6.08 (± 9.53)	1.02 (± 0.12)
	Autre	C	0.89	0.43 (± 6.84)	0.99 (± 0.12)
IPR	Craie	F	0.82	-4.60 (± 8.32)	1.05 (± 0.17)
	Craie	C	0.86	1.55 (± 5.28)	1.00 (± 0.13)
	Autre	F	0.90	0.82 (± 9.11)	0.96 (± 0.11)
	Autre	C	0.85	2.78 (± 8.38)	1.02 (± 0.14)
IPR+	Craie	F	0.23	12.63 (± 14.27)	0.55 (± 0.30)
	Craie	C	0.94	3.58 (± 2.59)	0.77 (± 0.06)
	Autre	F	0.38	18.53 (± 18.80)	0.58 (± 0.24)
	Autre	C	0.93	6.48 (± 4.41)	0.80 (± 0.07)

En dehors des modèles IPR+ (pour lesquels on retrouve le problème d'ajustement entre nombre d'occurrences observées et prédites), on constate une concordance étroite entre l'occurrence observée des espèces et l'occurrence prédite, aussi bien sur les sites de la craie que pour les autres sites, à la fois pour les modèles PIREN et les modèles IPR. Cette concordance s'observe aussi bien pour les espèces d'eau chaude que celles d'eau froide. Ainsi, contrairement à nos hypothèses de départ, on ne constate ni de surreprésentation des espèces d'eau froide, ni une sous-représentation des espèces d'eau chaude par rapport aux prédictions pour les cours d'eau de la craie.

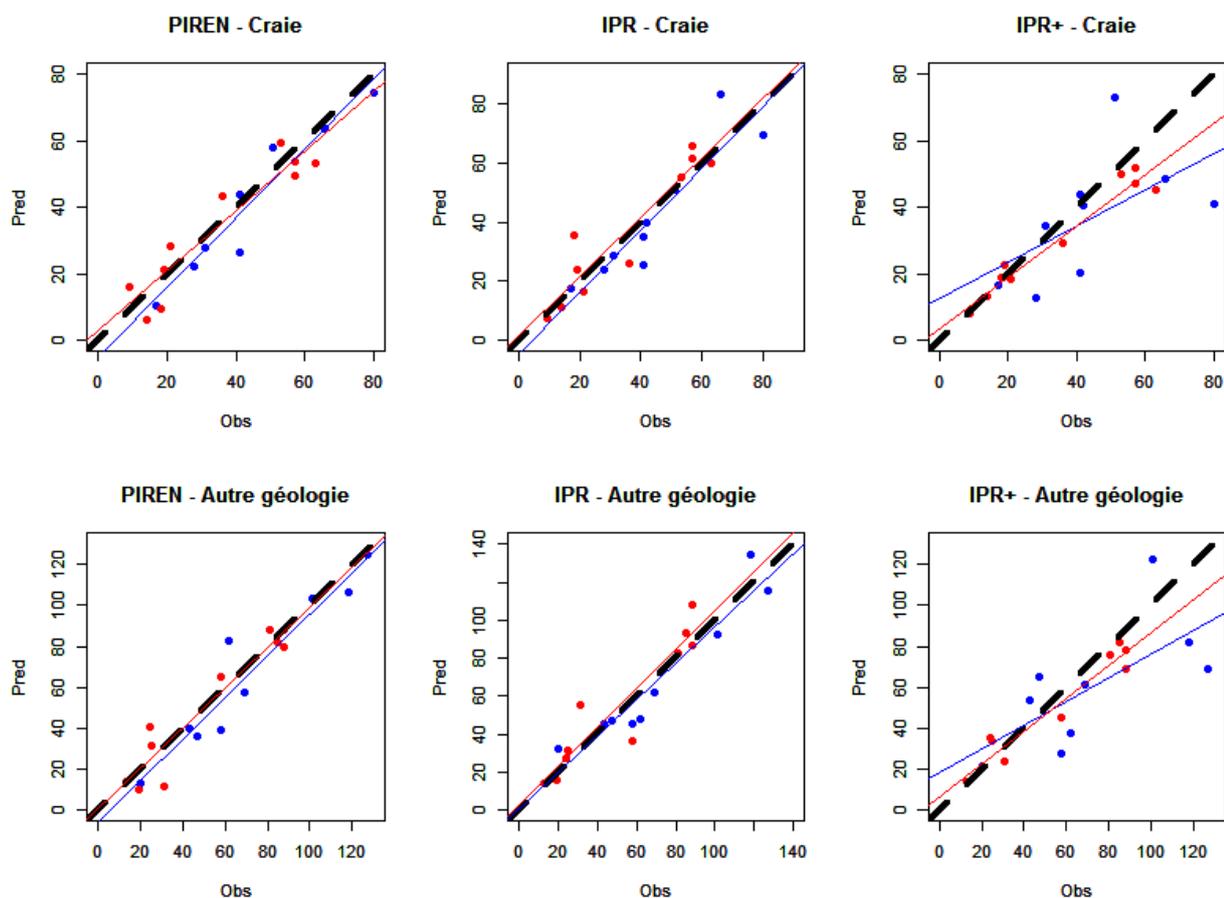


Figure7: Relations entre le nombre d’occurrences observées et le nombre d’occurrences prédites selon les 3 séries de modèles (PIREN, IPR et IPR+) en fonction du type de substrat géologique. Les espèces d’eau froide (points bleus ; droite bleue pour l’ajustement linéaire entre points) et les espèces d’eau chaude (points rouges ; droite rouge pour l’ajustement linéaire entre points) ont été distinguées. La droite hachurée noire (droite 1:1) correspond à l’égalité entre observation et prédiction.

4 Discussion et perspectives

Dans ce travail, nous avons utilisé deux types d’approches complémentaires pour évaluer la qualité des modèles de distribution d’espèces. Le premier, le plus global, consiste à examiner, pour un ensemble de sites, la concordance entre le nombre d’occurrences observées et le nombre d’occurrences prédites (estimé par la somme des probabilités d’occurrence prédites par le modèle sur les différents sites). Le second, plus précis, repose sur différents indicateurs mesurant la capacité d’un modèle à prédire correctement la présence et/ou l’absence d’une espèce à l’échelle d’un site.

Le premier type d’approche montre qu’à l’échelle d’un ensemble de sites, les modèles PIREN et IPR permettent de prédire correctement le nombre d’occurrences des 19 espèces examinées (concordance étroite entre nombre d’occurrences observées et nombre d’occurrences prédites – $r^2 = 0.91$ et 0.87 respectivement). Pour les modèles IPR+, le lien entre nombre d’occurrences observées et nombre d’occurrences prédites est beaucoup plus brouillé ($r^2 = 0.64$) et l’occurrence prédite des espèces tend à être sous-évaluée en particulier pour les espèces les plus fréquentes sur le bassin. Ce résultat suggère que les modèles PIREN et IPR calés sur des jeux de données plus étroitement lié au bassin de la Seine fournissent des prédictions plus cohérentes que les modèles IPR+ construit à une échelle spatiale beaucoup plus vaste.

Le second type d'approche, fondé sur l'examen d'indicateurs de performance de prédictions à l'échelle du site (AUC, indice de Kappa...) permet de compléter et de nuancer sensiblement ces premières conclusions. On constate notamment que la qualité des modèles varie fortement d'une espèce à l'autre. Ainsi, si pour certaines espèces, la capacité prédictive des 3 séries de modèles est globalement satisfaisante, leurs performances sont en revanche franchement médiocres voire mauvaises pour plusieurs espèces. Ces espèces pour lesquelles la présence/absence est mal prédite regroupent à la fois des espèces rares à distribution fragmentée (spirin, bouvière) et des espèces plus fréquentes mais connues pour leur difficulté d'échantillonnage (épinuche, épinochette, lamproie, chabot). Globalement, toutes espèces confondues, les valeurs des indicateurs de performances des modèles testés ici sont plutôt faibles en comparaison à celles observées dans la littérature pour d'autres zones géographiques et/ou d'autres modèles évalués. Cette situation est sans doute liée dans une large mesure à la relative homogénéité des conditions environnementales qui caractérise les cours d'eau du bassin de la Seine.

Les indicateurs de performance des modèles (AUC et Kappa notamment) tendent à être globalement légèrement meilleurs pour les modèles PIREN en comparaison avec les modèles IPR et IPR+ mais cette différence ne revêt pas de caractère significatif du point de vue statistique. Ainsi, même si les modèles PIREN intègrent des variables prédictives liés aux pressions anthropiques, contrairement aux autres modèles disponibles, cela ne se traduit pas par un gain sensible en termes de performance prédictive.

Les analyses comparant les sites en fonction de la nature, complète ou partielle de l'échantillonnage, suggèrent que la non-détection probable de certaines espèces, dans le jeu de données ayant servi à évaluer les modèles, conduirait à affaiblir la cohérence entre observations et prédictions. Ainsi une partie des erreurs des prédictions des modèles seraient en fait le résultat d'un artefact d'échantillonnage propre au jeu de données test, et donc sans lien avec le « réalisme écologique » des modèles.

Les analyses comparant les sites en fonction de la nature du substrat géologique (en lien avec son influence supposée sur le régime thermique des cours d'eau) suggèrent que la prise en compte de la température de l'air au lieu de la température de l'eau dans les variables prédictives ne conduit pas à un biais systématique des prédictions de présences/absences. Cette situation peut sembler paradoxale au regard des données de terrain qui révèlent clairement que les cours d'eau de la craie, qui présentent des faibles températures estivales, abritent essentiellement des espèces d'eau froide et peu ou pas d'espèces d'eau chaude contrairement à d'autres types de cours d'eau de taille équivalente (Belliard et al., 1997). Une des hypothèses pouvant expliquer un tel résultat serait que le découpage que nous avons pratiqué n'est pas optimal (les cours d'eau « autre substrat » intègrent des cours d'eau sur calcaire avec des influences de nappe potentiellement fortes ; les grands cours présentant des bassins géologiquement hétérogènes ont été inclus dans les jeux de données et sont susceptibles de brouiller les tendances). Une autre hypothèse également plausible serait que l'éventuel effet « température de l'eau » pourrait être intégré de manière indirecte via les autres variables prédictives utilisées dans les modèles.

Ce travail fournit une première évaluation de la performance des modèles de distributions d'espèces de poissons applicables sur le bassin de la Seine. Cette évaluation concerne la capacité des modèles à prédire la distribution spatiale actuelle des espèces dans le réseau hydrographique. Elle pourrait être complétée, à terme, par une évaluation de leur capacité à prédire des modifications temporelles des distributions d'espèces (ou des modifications temporelles de la composition des communautés à l'échelle d'un site). Sur ce point, et compte tenu des variables mises en jeu, on sait néanmoins d'ores et déjà que les modèles IPR et IPR+ ne peuvent potentiellement rendre compte que de changements temporels liés à des modifications hydro-climatiques alors que les modèles PIREN pourraient également intégrer des changements liés à l'évolution de la qualité d'eau ou de certains aménagements sur le bassin.

Pour l'avenir plusieurs pistes d'amélioration peuvent être envisagées pour construire et proposer des modèles prédictifs plus performants :

- Choix des méthodes de modélisation : les outils que nous avons testés sont basés sur des GLM. Or, ce type de modèle n'est pas nécessairement le plus performant et de nombreux travaux montrent qu'il existe d'autres méthodes de modélisation (ou combinaisons de méthodes) qui permettent d'améliorer les pouvoirs prédictifs (Buisson et al., 2010 ; Leclère et al., 2011). De telles approches pourraient assez aisément être testées sur le bassin de la Seine à partir des jeux de données déjà disponibles.
- Variables prédictives : un choix plus raisonné et optimisé des variables prédictives pouvant être intégrées dans les modèles devrait également conduire à une amélioration sensible des performances des modèles. Au-delà des performances prédictives proprement dites, une connexion plus directe avec les autres modèles du PIREN-Seine, à partir de l'intégration de certaines de leurs prédictions comme variables d'entrée des modèles poissons (par ex. sur les aspects hydrologie, température et qualité physico-chimique de l'eau) nous semble une piste prometteuse.
- Prise en compte de l'autocorrélation spatiale : La prise en compte des phénomènes d'autocorrélation spatiale au niveau des variables environnementales et de la distribution des espèces constitue une voie potentielle d'amélioration des modèles (Segurado et al., 2006).
- Prise en compte de l'effet des barrages et seuils : Les seuils et barrages peuvent avoir un double effet sur la répartition des espèces aquatiques : (1) localement (notamment en amont, dans la retenue) ils transforment radicalement les conditions d'environnements et donc les habitats disponibles pour les espèces ; (2) plus globalement ils induisent des discontinuités au sein du réseau hydrographique qui altèrent les processus de dispersion des espèces (Van Looy et al., 2014). À cet égard, leur prise en compte dans les modèles est susceptible d'améliorer sensiblement leur pertinence. Il existe maintenant des bases de données relativement exhaustives et géo référencées qui permettent d'envisager leur prise en compte à l'échelle du bassin de la Seine.

5 Références bibliographiques

- Austin, M. 2007. Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling***1–2**: 1–19.
- Belliard, J., Boët, P. & Tales E. 1997. Regional and longitudinal patterns of fish community structure in the Seine River basin, France. *Environmental Biology of Fishes***50**: 133–147.
- Belliard, J., Gorges, G., Hette, N., Le Pichon, C., Tales, E. & Zahm, A. 2010. La répartition des peuplements de poissons dans le bassin de la Seine: modèles à large échelle et focus sur les petits bassins versants franciliens. Programme PIREN-Seine: Peuplements de poissons dans le bassin de la Seine, 52p.
- Buisson, L., Blanc, L. & Grenouillet, G. 2008. Modelling stream fish species distribution in a river network: the relative effects of temperature versus physical factors. *Ecology of Freshwater Fish***17**: 244–257.
- Buisson, L., Thuiller, W., Casajus, N, Lek, S. & Grenouillet, G. 2010. Uncertainty in ensemble forecasting of species distribution. *Global Change Biology***16**: 1145–1157.
- Elith, J. & Leathwick, J.R. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics***40**: 677–697.
- Guisan, A. & Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters***8**: 993–1009.
- Guisan, A. & Zimmerman, N.E. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling***135**: 147–186.
- Kottelat, M. & Freyhof, J. 2007. Handbook of European freshwater fishes. Publications Kottelat, Cornol and Freyhof, Berlin. 646 pp.
- Leclère, J., Oberdorff, T., Belliard, J. & Leprieux, F. 2011. A comparison of modeling techniques to predict juvenile 0+ fish species occurrences in a large river system. *Ecological Informatics***6**: 276–285.
- Logez, M., Bady, P. & Pont D. 2012. Modelling the habitat requirement of riverine fish species at the European scale: sensitivity to temperature and precipitation and associated uncertainty. *Ecology of Freshwater Fish* **21**: 266–282.
- Oberdorff, T., Pont, D., Hugueny, B. & Chessel, D. 2001. A probabilistic model characterizing fish assemblages of French rivers: A framework for environmental assessment. *Freshwater Biology***46**: 399–415.
- 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.
- Poulet, N., Beaulaton, L. & Dembski, S. 2011. Time trends in fish populations in metropolitan France: insights from national monitoring data. *Journal of Fish Biology***79**: 1436–1452.
- Radinger, J., Wolter, C. & Kail, J. 2015. Spatial scaling of environmental variables improves species-habitat models of fishes in a small, sand-bed lowland river. *PLoS ONE***10**: e0142813.
- Segurado, P., Araujo, M.B. & Kunin, W.E. 2006. Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology***43**: 433–444.
- Van Looy, K., Tormos, T. & Souchon, Y. 2014. Disentangling dam impacts in river networks. *Ecological Indicators***37**: 10–20.