

PRÉDICTION DES RISQUES DE COLONISATION DE LA GAMBUSIE *GAMBUSIA AFFINIS* (POECILIIDAE) À L'ÉCHELLE MONDIALE

Duron Quiterie

Master 1 de Biologie des Organismes, des Populations et des Ecosystèmes.
UFR Sciences de la Vie et de l'Environnement. Université de Rennes 1.



GECCO, Laboratoire Evolution & Diversité Biologique (UMR-CNRS 5174)
Bâtiment 4R3. Université Paul Sabatier Toulouse III
118, route de Narbonne
31062 Toulouse cedex 4, France

Stage du 2 avril au 13 juin 2007. Soutenu à Rennes le 15 juin 2007.
Sous la direction de **Fabien Leprieur** (*Doctorant, UPS Toulouse 3*)
Et **Sébastien Brosse** (*Maître de Conférences, UPS Toulouse 3*)

Tuteur : **Pascal Laffaille** (*Maître de Conférences, Université de Rennes 1*)

REMERCIEMENTS

Je tiens à remercier Fabien Leprieur et Sébastien Brosse pour m'avoir accueillie au sein de leur équipe et pour leur encadrement exemplaire! Je les remercie pour leur disponibilité, leur écoute constante et leurs conseils toujours judicieux. Nos nombreuses discussions m'ont permis d'avancer dans mon travail et dans mon rapport.

Merci Fab pour m'avoir aidée à emprunter le dût et long chemin menant à la compréhension des modèles statistiques et pour m'avoir supportée ! Dommage, ça y est c'est la fin... alors maintenant « un suppôt, la cassette de Winnie l'ourson et au lit » (Leprieur 1992) !

Bon courage pour ta thèse ! Big Up !!

Seb, je te souhaite de bons voyages en Guyane ou ailleurs !

J'aimerais aussi remercier Gaël pour son aide précieuse en statistiques, Loïc pour les cartes, Géraldine pour la relecture de mon rapport et Karl pour m'avoir trouvé un logement à quelques encablures du labo.

Merci à Clément, Laetitia, Leslie et Simon... Bonne continuation !!

Enfin, je tiens à tous vous féliciter pour votre bonne humeur, votre formidable entraide et votre concertation continuelle.

PURA VIDA

SOMMAIRE

INTRODUCTION.....	1
MATERIEL ET METHODES.....	3
1) <i>Caractéristiques du modèle biologique.....</i>	3
2) <i>Présentation de la base de données</i>	4
3) <i>Identification de la niche écologique de la gambusie.....</i>	4
4) <i>Mise en place d'un modèle prédictif.....</i>	5
a) <u>Présentation des modèles</u>	
- Modèles linéaires généralisés (GLMs)	
- Modèles additifs généralisés (GAMs)	
b) <u>Construction des modèles</u>	
c) <u>Evaluation de la qualité des modèles</u>	
d) <u>Comparaison des prédictions données par les GLMs et les GAMs</u>	
- Comparaison de la qualité des modèles	
- Comparaison de la prédiction des sites	
5) <i>Impact potentiel de la gambusie sur les espèces des bassins prédits comme colonisables.....</i>	7

RESULTATS.....8

1) Identification de la niche écologique de la gambusie.....8

2) Modèles prédictifs.....9

a) Evaluation des modèles

b) Comparaison des modèles

c) Prédictions obtenues par les GLMs et les GAMs

3) Impact potentiel de la gambusie sur les espèces des bassins prédits comme colonisables.....10

DISCUSSION.....11

1) Importance de la niche réalisée exotique11

2) Prédiction des risques de colonisation.....12

3) Mesures de conservation.....13

4) Conclusions et perspectives.....15

BIBLIOGRAPHIE.....16

Annexe 1 : Présentation de la structure d'accueil.....21

Annexe 2 : Bilan Personnel du Stage.....22

INTRODUCTION

Les espèces invasives sont aujourd'hui considérées comme étant la deuxième cause de perte de biodiversité (e.g. Simberloff *et al.* 1996, Vitousek *et al.* 1997, Mack *et al.* 2000) et comme responsables de pertes économiques importantes (Leung *et al.* 2002). Elles font partie des principaux agents responsables des changements globaux (Vitousek *et al.* 1996). Ainsi, les espèces exotiques peuvent parfois conduire à l'élimination des espèces natives par des processus de prédation et de compétition ou par détérioration de leurs habitats, par transmission de maladies ou encore par hybridation avec ces dernières (Lockwood *et al.* 2007). Leur impact sur les populations natives se répercute également à des échelles plus larges, puisqu'elles peuvent modifier la structure des communautés et le fonctionnement des écosystèmes (Mack *et al.* 2000, Simon & Townsend 2003).

Les milieux d'eau douce constituent des écosystèmes enclins à des altérations et des dégradations fréquentes dues aux usages par l'homme et sont, par conséquent, sensibles aux invasions par des espèces opportunistes (Moyle & Light 1996). En général, les invasions résultent de facteurs anthropiques; ainsi, beaucoup d'espèces de poissons ont été introduites au cours du XXème siècle à des fins économiques, ornementales ou sanitaires (Moyle & Cech 2004, Zambrano *et al.* 2006). Cependant, leur présence affecte les écosystèmes natifs (Mack *et al.* 2000) et accélère l'homogénéisation biotique du milieu, c'est-à-dire l'augmentation de la similarité faunistique entre bassins, souvent due à une diminution d'espèces rares au profit d'espèces communes (Mc Kinney & Lockwood 1999, Rahel 2002).

Une fois établies, les espèces exotiques sont difficilement éradiquées et les méthodes de contrôle sont très coûteuses (Mack *et al.* 2000, Leung *et al.* 2002, Zambrano *et al.* 2006). De ce fait, prédire les risques de colonisation de ces espèces est essentiel afin de permettre un enrayment rapide des invasions grâce à des mesures de prévention et de surveillance (Moyle & Light 1996, Zambrano *et al.* 2006). La mise en place de méthodes de gestion ne peut se faire qu'après avoir déterminé d'une part les sites à risque, et en particulier les zones d'invasion potentielle (Peterson & Vieglais 2001, Thuiller 2003), et d'autre part l'impact potentiel des espèces invasives sur les assemblages natifs (Kolar & Lodge 2001, Sakai *et al.* 2001).

Au cours de cette étude, nous nous intéresserons au potentiel d'invasion de la gambusie *Gambusia affinis*. Ce poisson, originaire du sud-est des Etats-Unis, a été introduit dans plus de 50 pays (Garcia Berthou *et al.* 2005) afin de contrôler les populations de moustiques vecteurs du paludisme (Krumholz 1948) et constitue probablement aujourd'hui le poisson d'eau douce le plus répandu à l'échelle mondiale (Krumholz 1948, Unmack 1998). Il fait partie des 100 espèces exotiques retenues par le GISP (Global Invasive Species Programme

<http://www.issg.org/database/>). En effet, parmi les 67 introductions reportées par Garcia Berthou (2005), la gambusie s'est établie avec succès dans 96,8% des cas. En outre, son efficacité de prédation des moustiques est très controversée (Hoy *et al.* 1972, Baker *et al.* 2004) et des impacts négatifs sur les populations natives de différents taxons ont pu être démontrés (zooplancton (Hurlbert *et al.* 1972), amphibiens (e.g. Lawler *et al.* 1999, Kats & Ferrer 2003), poissons (e.g. Mills *et al.* 2004)). Enfin, la gambusie perturbe indirectement les écosystèmes et serait un générateur d'eutrophisation (Hurlbert *et al.* 1972, Margaritora *et al.* 2001). Ces constatations montrent l'importance de pouvoir prévoir les risques de colonisation de cette espèce pour limiter son expansion.

La prédiction des risques d'invasion peut être effectuée grâce à des méthodes statistiques prédictives se basant sur la niche écologique réalisée de l'espèce, celle-ci correspondant à l'ensemble des conditions abiotiques et biotiques permettant la survie de l'espèce (Hutchinson 1957). Différentes méthodes sont utilisées tels que les modèles linéaires généralisés (GLM; Mc Cullagh & Nelder 1989), les modèles additifs généralisés (GAM; Hastie & Tibshirani 1995), les arbres de classification (CART; De'ath & Fabricius 2000), les réseaux de neurones artificiels (ANN; Rumelhart *et al.* 1986) et les algorithmes génétiques (GARP; Stockwell & Peters 1999). Cependant, chaque méthode présente des avantages et des biais qui lui sont propres et ce n'est que récemment que certains auteurs se sont risqués à la comparaison de ces méthodes (Manel *et al.* 1999, Olden & Jackson 2002, Thuiller *et al.* 2003). Ainsi, comme le suggèrent Araujo & Guisan (2006), il est important d'étudier différents modèles sur un même jeu de données afin de connaître leurs avantages et inconvénients et de fournir des informations fiables sur les risques d'invasion.

Le but de notre étude est de prédire les risques de colonisation de *G. affinis* à l'aide de différents outils de modélisation pour identifier les bassins susceptibles de l'accueillir à l'échelle mondiale. Nous analyserons alors les impacts négatifs potentiels de la gambusie sur les communautés natives de ces bassins. Dans un premier temps, une représentation de la niche native et de la niche exotique de la gambusie en fonction des conditions climatiques est effectuée afin de déterminer les principales variables intervenant dans la construction de la niche et d'apprécier un éventuel élargissement de celle-ci. Les résultats obtenus nous permettent de décider si les modèles peuvent être établis à partir de l'aire native ou doivent prendre en compte l'ensemble des sites où l'espèce est présente. Dans un deuxième temps, deux types de modèles prédictifs (GLMs et GAMs) sont construits dans le but d'identifier au mieux les zones d'invasion potentielle. Enfin, les risques de perturbations des communautés natives suite à l'invasion de la gambusie ont été évalués et discutés.

MATERIEL ET METHODES

1) Caractéristiques du modèle biologique

G. affinis est un poisson d'eau douce de la famille des Poeciliidae (Ordre des Cyprinodontiformes). Cette espèce présente une biologie et une morphologie très similaires à celle de *G. holbrooki*. Les deux formes sont souvent confondues et parfois considérées comme deux sous-espèces : *G. affinis affinis* et *G. affinis holbrooki* (Pyke 2005). Cependant, nos données se réfèrent uniquement aux occurrences de *G. affinis*.

La gambusie est une espèce vivipare de petite taille de 1 à 5 cm de longueur totale. Elle présente une maturité sexuelle précoce, une croissance rapide, un taux de reproduction élevé et une durée de vie courte. Omnivore et opportuniste, elle se nourrit principalement en surface d'arthropodes, mollusques, œufs, petits poissons mais aussi d'algues et autres particules végétales. Son mode de prédation implique principalement la vue c'est pourquoi les milieux turbides lui sont peu favorables (Pyke 2005).

De plus, *G. affinis* a été observée dans des milieux présentant des caractéristiques environnementales très différentes. Elle peut survivre sous des conditions extrêmes: températures fortes ou faibles (valeurs variant suivant les auteurs, températures proche de zéro à 40°C), taux d'oxygène faible (jusqu'à 0,28 mg/l ; Pyke 2005), salinité variable (0 ‰ à 30‰ ; Nordlie & Mirandi 1996). Cependant, elle semble préférer des températures fortes (31°C en moyenne; Winkler 1979), des eaux peu profondes (8-15cm ; Pyke 2005), des milieux lenticules et riches en végétation (Casterlin & Reynolds 1977). Elle n'est retrouvée en général qu'à des altitudes inférieures à 1000m (Alcaraz 2006). De part cette forte capacité d'adaptation et le nombre important de ses introductions, *G. affinis* est présente à travers le monde entier (Fig.1) et la probabilité de colonisation de nouveaux sites serait donc importante.

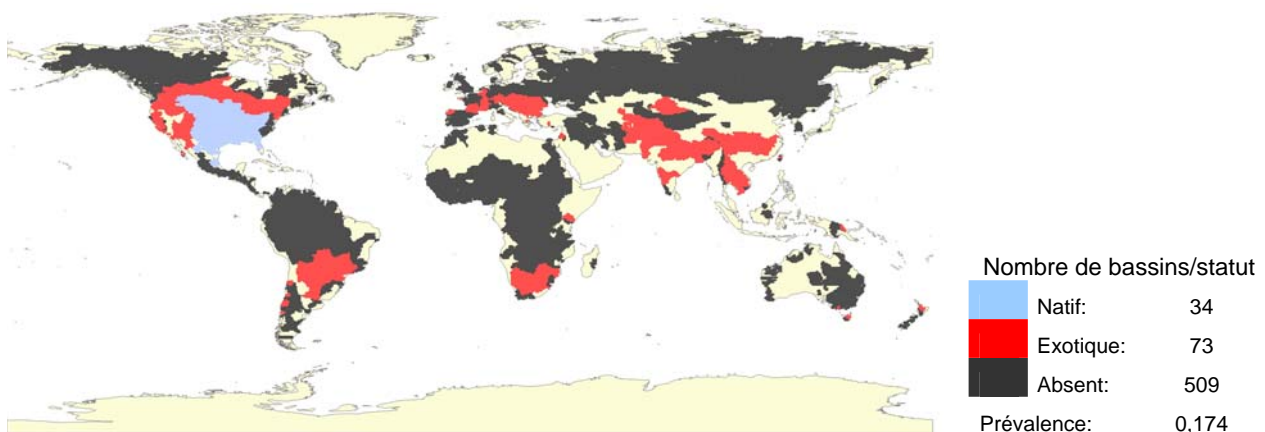


Figure 1 : Distribution de la gambusie (*G. affinis*) à travers le monde, aire native, exotique et absences connues dans 616 bassins versants.

2) Présentation de la base de données

Les données utilisées lors de notre étude proviennent de la base de données de l'ANR « Freshwater Fish Diversity » regroupant à la fois la liste faunistique des poissons et les caractéristiques climatiques et physiques de 1050 bassins hydrographiques à l'échelle mondiale. Dans notre cas, seuls les bassins pour lesquels toutes les variables environnementales listées dans la **Table 1** sont disponibles, ont été retenus ; limitant l'étude à 616 bassins versants répartis sur six zones biogéographiques (Néarctique: 132, Paléarctique: 210, Afrotropicale: 72, Australienne: 98, Néotropicale: 71, Orientale: 33).

Pour chaque bassin, la composition spécifique est connue ainsi que le statut de chaque espèce (native ou exotique). Dans cette étude, 17 variables environnementales et climatiques permettent de caractériser chaque bassin. Ces dernières ont été calculées avec un Système d'Information Géographique (**Tab.1**).

Table 1: Variables climatiques et physiques caractérisant les 616 bassins : abréviations, significations, unités, moyennes et erreurs standards (moy \pm e.s), valeurs minimums (min) et maximums (max).

Variable	Signification	Unité	moy \pm e.s.	min	max
Range	Altitude maximale - altitude minimale du bassin	semi-quantitatif	3.95 \pm 0,06	0	8
Alt_Moy	Altitude moyenne du bassin	m	573 \pm 21.7	8	4062
PET	Evapotranspiration potentielle	mm/an	69.9 \pm 1.23	18.6	149
AET	Evapotranspiration réelle	mm/an	50.5 \pm 1.06	5.11	126
NPP	Production primaire nette	kg/C/m ² /an	0.54 \pm 0.01	0	1.23
EnsAn	Ensoleillement/an	kWh/m ²	610 \pm 7.07	243	954
Tmin	Température minimale au cours de l'année	°C	5.06 \pm 0.58	-41.8	25.6
Tmax	Température maximale au cours de l'année	°C	21,0 \pm 0.25	0.2	32.8
Tmoy	Moyenne des températures mensuelles	°C	13.1 \pm 0.39	-16.2	28.2
TCV	Coeff. de variation des t° intermensuel	sans unité	1.02 \pm 0.07	0.01	11.8
T_Ampl	Amplitude des températures au cours du mois	°C	10.4 \pm 0.12	1.16	16.7
NBJP	Nombre de jours de pluie/An	jour	120 \pm 2.06	11	241
Pmin	Précipitation minimale au cours de l'année	mm/mois	33.2 \pm 1.27	0	192
Pmax	Précipitation maximale au cours de l'année	mm/mois	167.1 \pm 5.44	14	941
Pmoy	Moyenne des précipitations mensuelles	mm/mois	1040 \pm 28.4	51.6	3568
PCV	Coeff. de variation des préc° intermensuel	sans unité	0.55 \pm 0.01	0.07	1.67
Surface	Surface totale du bassin	km ²	134875 \pm 18897	6	6145186

3) Identification de la niche écologique de la gambusie

Une Analyse en Composantes Principales (ACP) normée est effectuée sur les 616 bassins et les variables climatiques. Cette ACP permet d'ordonner les bassins les uns par rapport aux autres en fonction de leurs caractéristiques climatiques. Les bassins où la gambusie est native puis ceux où elle a été introduite sont regroupés sur le plan formé par les deux premiers axes de l'ACP. Le but est i) de visualiser la niche écologique réalisée de l'espèce en zone native et exotique ; ii) d'identifier les principales variables climatiques qui déterminent la

distribution spatiale de la gambusie en zone native et exotique. Enfin, une analyse interclasses est effectuée afin de tester si la différence entre niche exotique et niche native diffère significativement de la différence due au hasard (test de Monte Carlo, 999 permutations, H_0 : la différence entre les deux niches n'est pas significativement différente de celle attendue par le hasard, revue dans Broenniman *et al.* 2007). Les analyses ont été réalisées en utilisant la librairie 'ade4' du logiciel R (version 2.4.1).

4) Mise en place d'un modèle prédictif

Deux méthodes de modélisation sont utilisées afin de comparer leur capacité de prédiction dans le but de sélectionner le modèle fournissant les meilleurs résultats: une méthode linéaire, les Modèles Linéaires Généralisés (GLMs) et une méthode non linéaire, les Modèles Additifs Généralisés (GAMs).

a) Présentation des modèles

- **Modèles linéaires généralisés (GLMs)** (Mc Cullagh & Nelder 1989)

Les GLMs nous ont permis de modéliser une relation linéaire entre les variables climatiques et physiques (variables indépendantes) et l'occurrence de la gambusie (variable dépendante) grâce à des relations de type binomiale et des liens logit. Les GLMs étant basés sur l'étude de relations linéaires, les variables ne suivant pas une loi normale sont transformées grâce à un logarithme de base 10. Cette procédure a été appliquée pour les variables suivantes : Alt_Moy, PET, AET, NPP, TCV, Pmin, Pmax, Pmoy, PCV, Surface.

-**Modèles additifs généralisés (GAMs)** (Hastie & Tibshirani 1990)

Les GAMs sont construits de la même manière que les GLMs. A la différence de ces derniers, ils permettent de s'affranchir de la linéarisation des variables (méthode non-paramétrique). En écologie, les relations entre variables étant le plus souvent de forme complexe, les GAMs assurent un meilleur ajustement aux données (Chambers & Hastie 1997, Yee & Mitchell 1991). Les librairies 'gam' et 'splines' du logiciel R sont utilisées. Les variables sont préalablement lissées par la fonction spline.

b) Construction des modèles

Dans un premier temps, le modèle est construit à partir de l'ensemble des données. Il est ensuite testé sur ces mêmes données afin d'avoir une première estimation de la qualité du modèle (voir méthodes d'évaluation de qualité en 4c). Cette méthode, dite de « Resubstitution », a tendance à surestimer les capacités de prédiction du modèle, à cause de la non indépendance des données d'apprentissage et de test du modèle (Fielding & Bell 1997). Pour cette raison, si le modèle obtenu par Resubstitution est considéré de bonne qualité, les

données doivent être partagées entre un jeu d'apprentissage et un jeu test indépendants, dans le but de déterminer la capacité de généralisation du modèle. Dans notre cas, nous utiliserons la méthode du Leave-One-Out (LOO), qui consiste à établir le modèle à partir de (n-1) données et de le tester sur la donnée restante autant de fois que de données existantes. Cette méthode permet d'obtenir l'indépendance entre le jeu d'apprentissage et le jeu test tout en conservant le maximum d'information pour la construction du modèle (revu dans Fielding & Bell 1997). Enfin, un essai d'optimisation du modèle est effectué par sélection des variables grâce à la fonction 'step' (disponible au sein du logiciel R) basé sur le critère d'information de Akaike (AIC ; Quinn & Keough 2002). Après cette sélection, les méthodes de « Resubstitution » et de « LOO » sont à nouveau utilisées.

Pour chaque bassin, les modèles permettent de déterminer une probabilité de présence de l'espèce (valeur entre 0 et 1).

c) Evaluation de la qualité des modèles

Pour chaque modèle, le pouvoir prédictif est évalué en utilisant les « Receiver Operating Characteristic Curves » (ROC Curves). Ces courbes représentent la proportion de Vraies Présences (Sensitivité) en fonction des Fausses Présences (1-Spécificité) (Hanley & McNeil 1982). L'avantage de cette méthode réside en l'absence de dépendance au seuil déterminant l'occurrence de l'espèce (Manel *et al.* 2001). L'aire sous la courbe ROC (ou AUC) constitue un bon indicateur de la qualité du modèle: plus l'AUC augmente (plus elle tend vers 1), plus la précision du modèle est importante. En général, le modèle est considéré comme bon lorsque $AUC > 0,8$; en revanche, si $AUC < 0,5$, le modèle prédit moins bien que le hasard (Hanley & Mc Neil 1982).

De plus, les courbes ROC permettent de calculer le seuil optimal considéré comme étant celui pour laquelle la somme de la spécificité et de la sensibilité est maximale (Manel *et al.* 2001). Ce seuil définit la probabilité limite à partir de laquelle l'espèce est présente et permet ainsi de transformer les données en présence/absence. La matrice de confusion est alors générée grâce au seuil obtenu (**Tab.2**).

Table 2: Matrice de confusion théorique

		Observé	
		Présence	Absence
Prédit	Présence	a (VP)	b (FP)
	Absence	c (FA)	d (VA)

a = VP (Vraie Présence)
b = FP (Fausse Présence)
c = FA (Fausse Absence)
d = VA (Vraie Absence)

D'autres estimateurs de la qualité du modèle, dépendants du seuil, peuvent alors être calculés à partir de la matrice de confusion:

- le taux de bonne prédiction $(a+d)/N^*$: proportion de Vraie Présence et de Vraie Absence sur le nombre total de prédictions.

- la sensibilité $a/(a+c)$: proportion de Vraie Présence sur l'ensemble des présences observées

- la spécificité $d/(b+d)$: proportion de Vraie Absence sur l'ensemble des absences observées

- l'indice Kappa $[(a+d)-(((a+c)(a+b)+(b+d)(c+d))/N^*)]/[N^*-(((a+c)(a+b)+(b+d)(c+d))/N^*)]$:

proportion pondérée de bonnes prédictions. N^* : nombre total de sites. Si $K < 0,4$ le modèle est mauvais; bon lorsque $0,4 < K < 0,75$ et excellent quand $K > 0,75$ (Fielding & Bell 1997)

d) Comparaison des prédictions données par les GLMs et les GAMs

- Comparaison de la qualité des modèles

Les AUC des courbes ROC du GAM et du GLM générés après LOO sont comparées par une méthode de ré-échantillonnage (« bootstrapping »). Pour chaque modèle, les probabilités d'occurrence de la gambusie pour 500 bassins (parmi les 616) sont tirées aléatoirement et leurs AUC relatives sont calculées. Cet échantillonnage est répété 1000 fois. Une différence significative entre modèles est détectée si plus de 5% des AUC d'un modèle sont supérieures à l'autre.

- Comparaison de la prédiction des sites

Les prédictions de colonisation des bassins données par les GLMs et les GAMs sont illustrées sous forme de cartes mondiales réalisées sous ArcView vers. 9.0. Afin de vérifier la concordance entre résultats des GLMs et des GAMs, une matrice de confusion entre valeurs prédites par les deux méthodes est alors mise en place afin d'effectuer un test du χ^2 d'homogénéité avec correction de Yates. De plus, des tests de corrélation de Pearson entre les probabilités d'occurrence obtenues à partir des GLMs et des GAMs, puis entre valeurs observées et probabilités prédites par les GLMs puis par les GAMs sont effectués.

5) *Impact potentiel de la gambusie sur les espèces natives des bassins colonisables*

Après établissement des prédictions, il est intéressant de savoir si la colonisation de la gambusie présente un danger pour les communautés natives. Pour cela, un examen de la composition spécifique des bassins est requis. Seuls ceux prédits comme colonisables à la fois par les GLMs et les GAMs sont pris en compte car l'exactitude de la prédiction augmente avec la concordance des modèles (Araújo & New 2006). Tout d'abord, les bassins présentant des espèces endémiques sont analysés : le statut IUCN (disponible sur <http://www.iucnredlist.org/>) ainsi que les principaux traits biologiques et écologiques (taille, lieu de vie, alimentation, comportement) des endémiques sont notés. Ensuite, toutes les espèces, endémiques ou non, considérées en danger critique (CR), en danger (EN) et vulnérables (VU) sont dénombrées.

RESULTATS

1) Identification de la Niche écologique de la gambusie

L'ACP permet d'identifier deux axes principaux de variables climatiques. Le premier axe, d'inertie 41,6%, est principalement défini par l'évapotranspiration potentielle et réelle (PET: -0,938 ; AET: -0,724) et les températures (Tmin: -0,927; Tmoy: -0,957 et Tmax: -0,845). Le deuxième axe, d'inertie 33%, est principalement corrélé aux précipitations (NBJP: -0,866; Pmin: -0,786; Pmoy: -0,737) à l'ensoleillement (EnsAn: 0,712), à la productivité primaire nette (NPP: -0,792) ainsi qu'aux amplitudes de température (T_Ampl: 0,755). Le cercle de corrélation (Fig.2A) montre les contributions absolues des variables aux axes. La niche réalisée native est significativement différente de la niche réalisée exotique (Fig.2B ; Test de Monte-Carlo : $P=0,001$) même s'il existe une superposition des deux statuts. Un élargissement de la niche selon l'axe 1 et l'axe 2 est observé, ceci vers des zones à températures et précipitations plus faibles et des variations de ces dernières plus importantes. Compte tenu de ce résultat, les modèles prédictifs doivent à la fois se baser sur la niche native et la niche exotique.

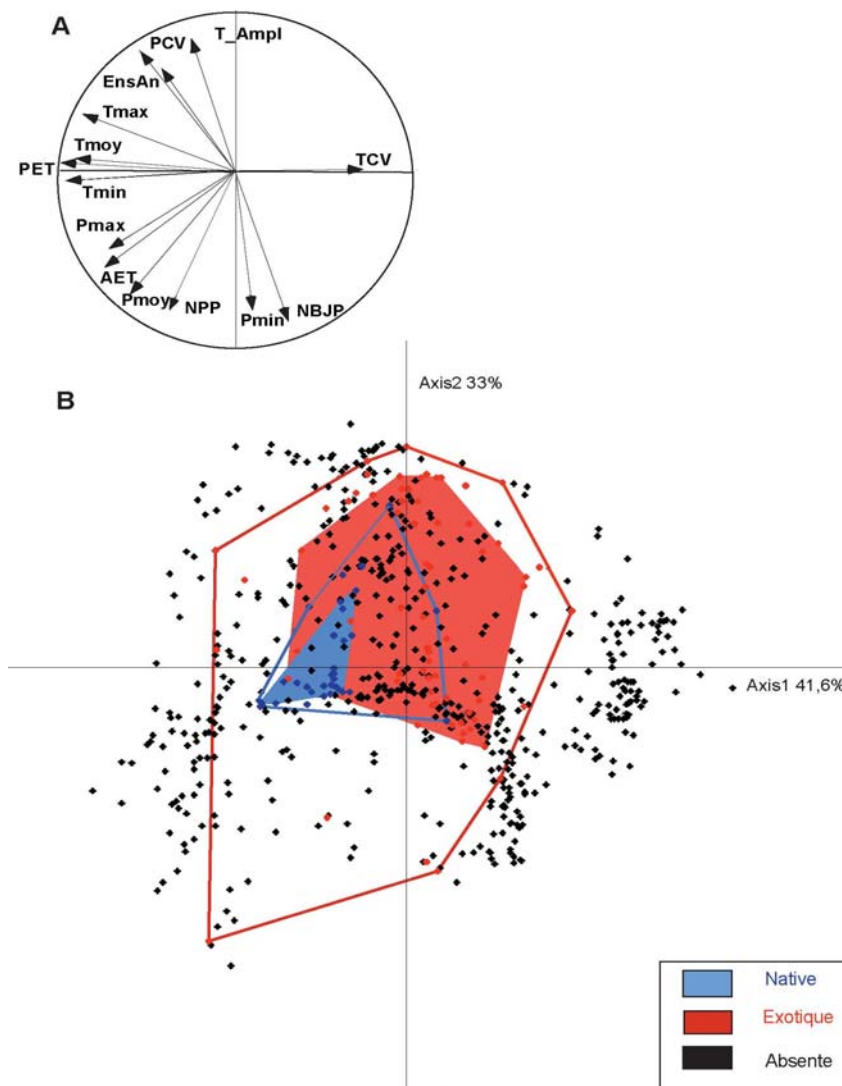


Figure 2 : Variables climatiques définissant la niche de *G. affinis* et identification de la niche réalisée native et de la niche réalisée exotique.

A) Cercle de corrélation

B) Superposition des sites sur le plan factoriel des variables. Les sites sont représentés par des points de couleurs différentes suivant le statut de l'espèce (natif, exotique ou absente). Pour les statuts natifs et exotiques, les zones coloriées regroupent 75% des sites partageant le plus de similitudes environnementales et les traits colorés regroupent 100% des sites.

2) Modèles prédictifs

a) Evaluation des modèles

Pris individuellement, les modèles sont de bonne qualité avec une faible diminution du taux de bon classement entre Resubstitution et LOO, et un AUC élevé en LOO ($AUC \geq 0,75$). Par contre, l'indice Kappa passe de 0,36 à 0,30 pour le GLM et de 0,65 à 0,40 pour le GAM. De plus, le nombre de Fausse Absence (FA) reste sensiblement le même pour le GLM alors que pour le GAM cette dernière est doublée. Après sélection des variables, les modèles sont toujours de bonne qualité et le nombre de Fausse Absence est encore diminué (Tab.3). Les variables sélectionnées pour le GLM : Tmoy, NBJP, Pmoy, Alt_Moy, Tmax, AET, TCV, Pmax, Range, EnsAn sont différentes de celles sélectionnées pour le GAM: s(Range, 2) + Alt_Moy + AET + NPP + EnsAn + s(Tmin, 2) + s(Tmax, 2) + s(Tmoy, 2) + s(NBJP, 2) + s(Pmin, 2) + Pmax + Pmoy + PCV.

Table 3: Comparaison du pouvoir de prédiction des GLMs et des GAMs obtenus après Resubstitution et Leave-One-Out sans et avec sélection des variables (non/oui)

MODELE	GLM				GAM			
	Resubstitution		Leave-One-Out		Resubstitution		Leave-One-Out	
Méthode	non	oui	non	oui	non	oui	non	oui
<i>Sélection des variables</i>	<i>non</i>	<i>oui</i>	<i>non</i>	<i>oui</i>	<i>non</i>	<i>oui</i>	<i>non</i>	<i>oui</i>
Déviance expliquée	0,24	0,24	-	-	0,53	0,35	-	-
AIC	467	457	-	-	418	409	-	-
AUC	0,83	0,83	0,79	0,80	0,95	0,88	0,85	0,85
Spéc+Sens	1,50	1,51	1,45	1,49	1,77	1,64	1,56	1,59
Seuils	0,19	0,16	0,16	0,15	0,21	0,24	0,13	0,18
VP	81	87	80	86	96	85	84	88
FP	129	157	154	165	63	81	121	122
FA	26	20	27	21	11	22	23	19
VA	380	352	355	344	446	428	388	387
Taux Bon classement	0,75	0,71	0,71	0,70	0,88	0,83	0,77	0,77
Sensitivité	0,76	0,81	0,75	0,80	0,90	0,79	0,79	0,82
Spécificité	0,75	0,69	0,70	0,68	0,88	0,84	0,76	0,76
Indice Kappa	0,36	0,34	0,30	0,31	0,65	0,52	0,40	0,42

b) Comparaison des modèles

Malgré les bons résultats obtenus après sélection des variables, dans un souci de comparaison, les modèles de LOO sans sélection sont pris en compte. Après test de randomisation, les AUC des GAMs sont apparues significativement supérieures aux AUC des GLMs ($P < 0,001$). Les occurrences prédites

Table 4: Matrice de confusion GLM/GAM

Modèle		GAM	
		Présence	Absence
GLM	Présence	152	82
	Absence	53	329

par les deux méthodes sont significativement différentes ($\chi^2 = 168$, ddl= 1, $P < 0,001$, Tab.4). Le coefficient de

corrélation entre les deux modèles est de $R=0.526$. La corrélation entre occurrences prédites par les GAMs et occurrences observées est supérieure à celle des GLMs. ($R=0,440 > R=0,347$).

c) Prédiction obtenues par les GLMs et les GAMs (Tab.3, Fig3, Fig.4)

Les GLMs prédisent 154 bassins colonisables (FP) et les GAMs, 121. Des populations de *Gambusia holbrooki* sont présentes dans 43 des sites prédits par les GLM, dans 51 pour les GAMs et dans 33 bassins prédits à la fois par les deux méthodes. Les vraies présences (VP), les vraies absences (VA) ainsi que les fausses présences (FP) et fausses absences (FA) obtenues par les deux modèles sont représentées sur les figures 3 et 4 (voir aussi Tab.3).

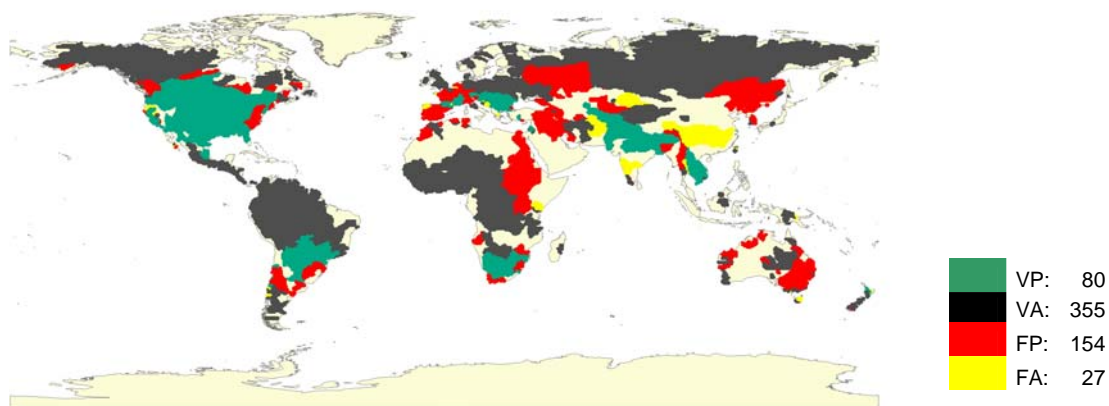


Figure 3 : Prédiction des risques de colonisation de la gambusie par les GLMs (FP: bassins colonisables)

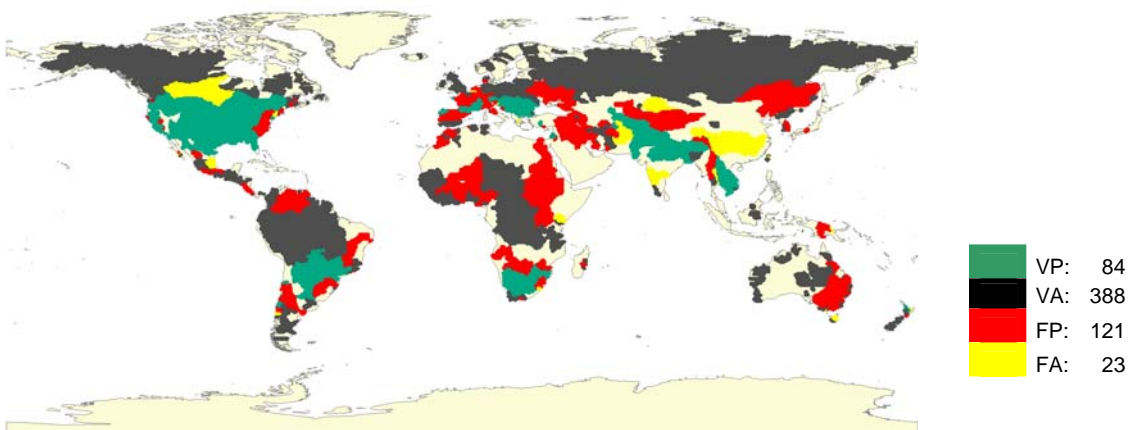


Figure 4 : Prédiction des risques de colonisation de la gambusie par les GAMs (FP: bassins colonisables)

3) Impact potentiel de la gambusie sur les espèces des bassins prédits comme colonisables

Sur les 77 bassins prédits en commun entre GLMs et GAMs, 16 bassins présentent des espèces endémiques. Parmi celles-ci, deux espèces endémiques sont sur la liste rouge de l'IUCN et une seule possède des traits correspondants à des espèces pouvant être perturbées par la gambusie: *Notropis mekistocholas* (Cyprinidae).

Sur les 77 bassins avec espèces endémiques ou non, 47 abritent des espèces présentes sur la liste rouge de l'IUCN soit un total de 126 espèces. Après examen de leurs traits, 18 se trouveraient potentiellement menacées dans le cas d'une introduction de la gambusie, ceci pour 14 bassins. Sur ces 14 bassins, 11 ont déjà été colonisés par *G.holbrooki* (Tab.5).

Table 5 : Bassins et espèces (endémiques ou non) potentiellement menacés par l'introduction de la gambusie. Nom et situation géographique du bassin. Présence de *G.holbrooki* (*G. hol* : oui/non). Espèces en danger/bassin. Pour chaque espèce, sont indiqués le statut IUCN (CR: en danger critique, EN: en danger, VU: vulnérable) et l'existence de bibliographie démontrant un impact négatif de la gambusie sur l'espèce considérée (B).

Bassin	Pays	<i>G.hol</i>	Espèces en danger (famille, statut, bibliographie)
Cape Fear	Etats-Unis	oui	<i>Notropis mekistocholas</i> (Cyprinidae,CR,B)
Clarence	Australie	oui	<i>Craterocephalus fluviatilis</i> (Atherinidae,EN,B), <i>Nannoperca oxleyana</i> (Percichthyidae,EN)
Douro	Espagne	oui	<i>Rutilus lemmingii</i> , <i>R. macrolepidotus</i> (Cyprinidae,VU) ; <i>Cobitis calderoni</i> (Cobitidae,VU)
Ebro	Espagne	oui	<i>Cobitis calderoni</i> (Cobitidae,VU), <i>Valencia hispanica</i> (Valenciidae,EN,B)
Incomati	Af. du Sud	non	<i>Opsaridium peringueyi</i> (Cyprinidae,VU)
Jucar	Espagne	oui	<i>Barbus guiraonis</i> (Cyprinidae,VU); <i>Valencia hispanica</i> (Valenciidae,EN,B)
Keurbooms	Af. du Sud	non	<i>Pseudobarbus tenuis</i> (Cyprinidae,EN)
Mary	Australie	oui	<i>Nannoperca oxleyana</i> (Percichthyidae, EN)
Murray Darling	Australie	oui	<i>Craterocephalus fluviatilis</i> , <i>C.arniculus</i> (Atherinidae,EN,VU,B); <i>Galaxias fuscus</i> , <i>G. rostratus</i> (Galaxiidae,CR,VU,B); <i>Nannoperca oxleyana</i> (Percichthyidae,EN)
Peedee	Etats-Unis	oui	<i>Fundulus waccamensis</i> (Fundulidae,VU)
Po	Italie	oui	<i>Knipowitschia punctatissima</i> (Gobiidae, VU)
Roanoke	Etats-Unis	oui	<i>Fundulus waccamensis</i> (Fundulidae,VU)
Tagliamento	Italie	non	<i>Knipowitschia punctatissima</i> (Gobiidae, VU)
Tagus	Portugal	oui	<i>Cobitis calderoni</i> (Cobitidae, VU); <i>Barbus comizo</i> , <i>B.microcephalus</i> ; <i>Rutilus lemmingii</i> , <i>R. macrolepidotus</i> (Cyprinidae, VU)

DISCUSSION

1) Importance de la niche réalisée exotique

La niche réalisée exotique de *G. affinis* est apparue comme significativement différente de sa niche réalisée native (Fig.2). Plus étendue et répartie sur six zones biogéographiques, elle comprend des zones à températures et précipitations plus faibles avec des variations plus fortes de ces dernières. Ce poisson étant une espèce euryèce (Pyke 2005) et possédant des traits biologiques favorables aux invasions (Sakai *et al.* 2001), il n'est pas étonnant de le retrouver au sein d'un large éventail de conditions. Il semble cependant limité par des précipitations fortes et des températures trop faibles rendant les milieux tropicaux ou de latitudes élevés peu propices à son implantation.

La niche exotique de la gambusie doit être prise en compte dans la construction des modèles prédictifs afin d'approximer au mieux sa niche potentielle puisque ici la niche native

ne comprend qu'une petite portion des conditions climatiques favorables à l'implantation de l'espèce. Jusqu'alors, la plupart des modèles prédictifs sont générés à partir de la niche native (e.g : Mercado-Silva *et al.* 2006). Cette méthode basée sur la théorie du conservatisme de la niche (Wiens & Graham 2005) suppose que la niche réalisée native regroupe toutes les conditions écologiques permettant la présence de l'espèce et soit donc bien représentative de sa niche potentielle. Or, la niche potentielle d'une espèce est généralement plus large que sa niche réalisée puisque l'aire de distribution de la niche est souvent limitée par des contraintes historiques et écologiques (Broennimann *et al.* 2007). En effet, des barrières géographiques ou des relations interspécifiques (prédation et compétition) existant au sein de l'aire native peuvent limiter localement la dispersion de l'espèce considérée.

Les études comparant les prédictions effectuées à partir de l'aire native avec celles effectuées à partir de l'aire exotique sont peu nombreuses et très récentes. Ainsi Loo *et al.* (2007) ont montré que les prédictions réalisées à partir de l'aire native d'une espèce d'escargot aquatique étaient de moindre qualité que celles effectuées à partir de l'aire exotique.

Nous recommandons donc de comparer la niche réalisée native avec la niche réalisée exotique afin de représenter au mieux la niche potentielle de l'espèce et ce, surtout lorsque l'espèce est connue comme résistante à une grande variété de conditions.

2) Prédiction des risques de colonisation

Les GLMs et les GAMs ont permis d'établir avec succès les risques de colonisation de *Gambusia affinis* à partir de sa distribution mondiale actuelle. En effet, les Vraies Présences (sensitivité) étant bien prédites et le nombre de Fausses Absences étant faibles (**Tab.3**), les Fausses Présences représentent une estimation raisonnable des zones d'invasion potentielle de la gambusie. Les Fausses Présences peuvent, d'une part, s'expliquer par une absence d'introduction de l'espèce. D'autre part, elles peuvent correspondre à des échecs d'introductions lorsque l'espèce introduite est confrontée (i) à de nouveaux prédateurs et compétiteurs (résistance biotique, Sakai *et al.* 2001) ou (ii) à des conditions environnementales défavorables (résistance environnementale, Leprieur *et al.* 2006). La prise en compte de variables intervenant à l'échelle locale semble être un bon moyen pour diminuer les erreurs de prédictions (Ficetola *et al.* 2007). Néanmoins, les informations à l'échelle locale (station au sein d'un bassin) sont difficilement accessibles à l'échelle mondiale.

La comparaison des fausses présences avec les zones accueillant une espèce sœur permet de vérifier la capacité d'accueil de l'espèce considérée (Peterson *et al.* 1999). *G.*

holbrooki nécessitant des conditions écologiques similaires (Pyke 2005), la possibilité d'établissement de *G. affinis* peut ainsi être évaluée. Cependant, nous pouvons nous demander si la présence d'une espèce sœur ne pourrait pas être à l'origine d'une résistance biotique interdisant l'implantation de *G. affinis*.

La comparaison des modèles montre que les GAMs donnent de meilleurs résultats que les GLMs (AUC supérieurs). Ceci peut s'expliquer par l'absence de relations linéaires entre l'occurrence de la gambusie et certaines variables. De plus, les sites prédits comme colonisables par les GLMs et les GAMs ne sont pas toujours les mêmes ($P < 0,001$). Ceci montre que l'utilisation simultanée de plusieurs méthodes est un moyen d'améliorer la robustesse des prédictions (Thuiller 2003). Par exemple, il serait possible de déterminer une probabilité d'invasion de chaque bassin en prenant en compte la concordance des résultats de chaque modèle (Araújo & New 2006). Parmi les autres méthodes couramment utilisées, nous avons testé les arbres de classification (CART). Les résultats obtenus n'ont pas été intégrés à cette étude car ces modèles ne permettent pas d'utiliser la procédure de LOO. La procédure utilisée est de type « K fold 10 », c'est-à-dire que le jeu d'apprentissage est construit sur 90% des données et le jeu test constitue 10% des données. Une modification du programme serait donc requise afin de pouvoir comparer leur performance aux GAMs et GLMs générés après LOO. Nous avons également testé la performance des réseaux de neurones (ANN). Les premiers essais ont montré que ces techniques ont un fort pouvoir de prédiction, mais que la procédure LOO limite le pouvoir de généralisation (les fausses présences sont pratiquement absentes). Il serait donc nécessaire de modifier soit l'architecture du modèle, soit la méthode de validation pour optimiser la capacité de prédiction des invasions.

3) Mesures de conservation

Dans notre analyse, seule l'influence potentielle de la gambusie sur les espèces natives listées par l'IUCN a été prise en compte. Parmi ces espèces, 18 ont de fortes chances d'interagir directement avec la gambusie, puisqu'elles présentent des traits biologiques et écologiques redondants avec ceux de l'espèce invasive (e.g : petite taille, prédateur de macroinvertébrés et de poissons, préférendum pour des eaux peu profondes et calmes...). Cette superposition de niche accroît donc la probabilité d'impact négatif de la gambusie, mais ne permet pas de quantifier son impact. Pour cela, il serait nécessaire de mettre en place des approches expérimentales dans le cadre de programmes locaux de gestion et de conservation.

Il serait également important, dans les études futures, d'élargir cette approche à des espèces autres qu'endémiques ou non listées dans la base IUCN. En effet, de nombreuses espèces non endémiques au sens strict (i.e. présentes dans un seul bassin versant) ne sont représentées que par de très petites populations mais la majorité ne figure pas dans la liste rouge de l'IUCN. Ainsi, par exemple, Alcaraz (2006) rapporte que *Aphanius baeticus*, *A. iberus* et *A. fasciatus*, bien que non répertoriées dans la base IUCN, sont menacées par *G. holbrooki*. De plus, des espèces non listées mais présentant des traits biologiques et écologiques très proches de la gambusie pourraient devenir menacées après l'introduction de cette dernière.

L'impact d'une espèce introduite peut se manifester directement par compétition ou prédation, mais peut aussi affecter indirectement le fonctionnement d'un écosystème. Ainsi, *Gambusia* exerce une pression de prédation importante sur les invertébrés aquatiques (e.g: diminution du nombre de libellules en Nouvelle-Zélande et Italie, Cachini *et al.* 2003), sur les œufs et alevins de poissons (e.g : *Galaxias maculatus*, Baker *et al.* 2004; *Heterandria formosa*, Belk & Lydeard 1994) ainsi que sur les œufs et larves d'amphibiens (e.g: Hamer *et al.* 2002). De plus, elle manifeste des comportements d'agressivité (morsures des nageoires) envers des poissons adultes de taille parfois supérieure (e.g : Laha & Mattingly 2007). Elle peut également constituer un compétiteur trophique pour les espèces natives comme l'ont démontré expérimentalement Caiola & De Sostoa (2005) pour *Valencia hispanica* et *Aphanius iberus*. Enfin, en se nourrissant de zooplancton, elle participe indirectement à l'eutrophisation du milieu en entraînant un contrôle « top-down » de la chaîne trophique (Hurlbert *et al.* 1972).

Gambusia est donc capable d'affecter les écosystèmes de différentes manières même si l'importance de son impact serait fonction de différents facteurs. Ainsi, selon Moyle (1986), son impact dépendrait de la complexité du milieu (augmentation du nombre de niches disponibles), des dimensions des niches des espèces natives et de la sensibilité de ces dernières aux perturbations. De même, Laha & Mattingly (2006, 2007) montrent que ses effets négatifs seraient atténués lorsque les températures sont faibles ceci permettant une coexistence avec les espèces natives. Néanmoins, lorsque la gambusie est établie au sein d'un bassin, empêcher sa dispersion semble difficile car celle-ci peut se déplacer dans 3mm d'eau (Alemadi & Jenkins 2007). De plus, selon Rehage & Sih (2004), *G. affinis* est l'espèce de gambusie la plus enclin à une dispersion rapide de part son fort comportement exploratoire du milieu.

En conséquence, les 77 bassins prédits à la fois colonisables par les GLMs et les GAMs par la gambusie doivent bénéficier d'une attention particulière. Comme le souligne Ruesink *et al.* (1995), tant qu'aucune étude ne démontre l'absence d'impacts négatifs sur les écosystèmes natifs, il est recommandé d'interdire l'introduction d'espèces exotiques. En général, les espèces utilisées en lutte biologique ont des effets non désirés sur les écosystèmes (Simberloff & Stiling 1996). Compte tenu de l'efficacité de certaines espèces natives contre les moustiques, un effort doit être accompli afin de les identifier et de préférer leur utilisation à celle de la gambusie (Lloyd *et al.* 1986).

d) Conclusion et perspectives

Notre étude se place au cœur des préoccupations et nécessités actuelles en offrant des lignes directrices pour la prédiction des risques de colonisation d'espèces invasives à l'échelle mondiale. D'abord, une bonne identification de la niche potentielle de l'espèce doit être réalisée avant toute tentative de modélisation des capacités d'invasion d'une espèce donnée. Ensuite, il faut préférer l'utilisation simultanée de plusieurs méthodes afin d'augmenter la robustesse des prédictions d'invasion. Enfin, lorsque cela est possible, la distribution d'une espèce « sœur » peut être comparée aux zones prédites comme colonisables par les modèles, constituant ainsi un moyen de valider la capacité de prédiction du modèle.

Les analyses à l'échelle mondiale sont encore peu nombreuses de part la forte dispersion de l'information (Peterson & Vieglais 2001). La plupart se sont limitées à des petites étendues et beaucoup portent sur les Etats-Unis (e.g : California, Moyle & Marchetti 2006; Région des Grands Lacs, Mercado-Silva *et al.* 2006; Amérique du Nord, Chen *et al.* 2007). Or, les prédictions à l'échelle mondiale constituent un bon outil permettant d'identifier globalement les bassins à risque avant de se focaliser sur les impacts de l'espèce introduite à une échelle plus fine.

Cette méthode peut être étendue à l'étude de l'évolution de la distribution des espèces en conséquence des changements climatiques actuels. En effet, les modèles construits ici pourront être utilisés avec les prédictions des données climatiques futures. Ceci sera l'objet de différents travaux dans le cadre de plusieurs programmes de recherche.

BIBLIOGRAPHIE

- Alcaraz CC. 2006. Ecological interactions between an invasive fish (*Gambusia holbrooki*) and native cyprinodonts : the role of salinity. Th. Ecologie. Soutenu le 20/06/06 à Girona. Université de Girona. 164p.
- Alemadi SD & Jenkins DG. 2007. Behavioral constraints for the spread of the eastern mosquitofish, *Gambusia holbrooki* (Poeciliidae). *Biological Invasions*. [en ligne]. Disponible sur <<http://www.springerlink.com/content/?k=alemadi>> (consultée le 04/05/07)
- Araújo MB & Guisan A. 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography* **33**: 1677-1688
- Araújo MB & New M. 2006. Ensemble forecasting of species distributions. *Trends in Ecology and Evolution* **22**(1): 42-47
- Baker C, Rowe D, Smith J. 2004. *Gambusia*-a biodiversity threat? *Water & Atmosphere* **12**(2): 24-25
- Belk MC & Lydeard C. 1994. Effect of *Gambusia holbrooki* on a Similar-Sized, Syntopic Poeciliid, *Heterandria formosa*: Competitor or Predator? *Copeia* **2**: 296-302
- Broennimann O, Treier UA, Müller-Schärer H, Thuiller W., Peterson AT, Guisan A. Evidence of Climatic Niche Shift during Biological Invasions. *Ecology letters* (in press)
- Carchini G, Di Domenico M, Pacione T, Solimini A, Tanzilli C. 2003. Species distribution and habitat features in lentic Odonata. *Italian Journal of Zoology* **70**: 39-46.
- Caiola N & De Sostoa A. 2005. Possible reasons for the decline of two native toothcarps in the Iberian Peninsula: evidence of competition with the introduced Eastern mosquitofish. *Journal of Applied Ichthyology* **21**: 358-363
- Casterlin ME & Reynolds WW. 1977. Aspects of habitat selection in the mosquitofish *Gambusia affinis*. *Hydrobiologia* **55**(2):125-127
- Chambers JM & Hastie TJ. 1997. Statistical models in S. Chapman & Hall, London,UK.595pp
- Chen P, Wiley EO, Mac Nysset KM. 2007. Ecological niche modelling as a predictive tool: silver and bighead carps in North America. *Biological Invasions* **9**: 43-51
- De'Ath G & Fabricius KE. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* **81**(11): 3178-3192
- Ficetola GF, Thuiller W, Miaud C. 2007. Prediction and validation of the potential global distribution of a problematic alien invasive species—the American bullfrog. *Diversity and distributions* (in press)
- Fielding AH & Bell JF. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* **24** (1): 38-49

- Garcia-Berthou E, Alcaraz C, Pou-Rovira Q, Zamora L, Coenders G, Feo C. 2005. Introduction pathways and establishment rates of invasive aquatic species in Europe. *Canadian Journal of Fisheries and Aquatic Sciences* **62**: 453-463
- Hamer AJ, Lane SJ, Mahony MJ. 2002. The role of introduced mosquitofish (*Gambusia holbrooki*) in excluding the native green and golden bell frog (*Litoria aurea*) from original habitats in south-eastern Australia. *Oecologia* **132**: 445-452
- Hanley JA & McNeil BJ. 1982. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology* **143**(1): 29-36.
- Hastie TJ & Tibshirani R. 1990. Generalized additive models. *Monographs on Statistics and Applied Probabilities* 43. Chapman & Hall, London,UK. [en ligne]. Disponible sur <<http://books.google.com/>> (consultée le 01/05/07)
- Hoy JB, Kauffman EE, O'Berg AG. 1972. A large-scale test of *Gambusia affinis* and chlorpyrifos for mosquito control. *Mosquito News* **32**:161-169
- Hurlbert SH, Zedler J, Fairbanks D. 1972. Ecosystem alteration by mosquitofish (*Gambusia affinis*) predation. *Science* **175**: 639-641
- Hutchinson GE. 1957. Concluding remarks. *Cold Springs Harbor Symposium on Quantitative Biology* **22**: 423-427
- Kats LB & Ferrer RP. 2003. Alien predators and amphibian declines: review of two decades of science and the transition to conservation. *Diversity & Distributions* **9**: 99-110
- Kolar CS & Lodge DM. Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution* **16**(4): 199-204
- Krumholz LA. 1948. Reproduction in the western mosquitofish, *Gambusia affinis affinis* (Baird & Girard), and its use in mosquito control. *Ecological monographs* **18**: 1-43
- Laha M & Mattingly HT. 2006. Identifying environmental conditions to promote species coexistence: an example with the native Barrens topminnow and invasive western mosquitofish. *Biological Invasions* **8**: 719-725
- Laha M & Mattingly HT. 2007. Ex situ evaluation of impacts of invasive mosquitofish on the imperilled Barrens topminnow. *Environmental Biology of Fishes* **78**(1): 1-11
- Lawler SP, Dritz D, Strange T, Holyoak M. 1999. Effects of introduced mosquitofish and bullfrogs on the threatened California red-legged frog. *Conservation Biology* **13**: 613-622
- Leprieur F, Hickey MA, Arbuckle CJ, Closs GP, Brosse S, Townsend CR. 2006. Hydrological disturbance benefits a native fish at the expense of an exotic fish. *Journal of Applied Ecology* **43**: 930-939
- Leung B, Lodge DM, Finnoff D, Shogren JF, Lewis MA, Lamberti G. 2002. An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceeding of the Royal Society of London B*. 269: 2407-2413

- Lloyd LN, Arlington AH, Milton DA. 1986. The mosquitofish—a valuable mosquito-control agent or a pest? *In: Kitching R. L. ed. The ecology of exotic animals and plants: some Australian case histories.* Australia, John Wiley and Sons.
- Lockwood JL, Hoopes MF, Marchetti MP. 2007. *Invasion Ecology.* Blackwell Scientific Press. UK. 304p
- Loo SE, Mac Nally R, Lake PS. 2007. Forecasting New Zealand mudsnail invasion range: model comparison using native and invaded ranges. *Ecological Applications* **17**(1):181-189
- Lowe S, Browne M, Boudjelas S, De Poorter M. 2000. 100 of the World's Worst Invasive Alien Species A selection from the Global Invasive Species Database. ISSG, SSC and IUCN. 12p. **[en ligne]**. Disponible sur <<http://www.issg.org/booklet.pdf>> (consultée le 18/05/07)
- Mac Kinney ML & Lockwood JL. 1999. Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends in Ecology and Evolution* **14**(11): 450-453.
- Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M, Bazzaz FA. 2000. Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications* **10**: 689-710
- Manel S, Dias JM, Buckton ST, Ormerod SJ. 1999. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *Journal of Applied Ecology* **36**: 734-747
- Manel S, Williams HC, Ormerod SJ. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* **38**: 921-931
- Margaritora FG, Ferrara O, Vagaggini D. 2001. Predatory impact of the mosquitofish (*Gambusia holbrooki* Girard) on zooplanktonic populations in a pond at Tenuta di Castelporziano (Rome, Central Italy). *Journal of Limnology* **60**: 189-193.
- Mercado-Silva N, Olden JD, Maxted JT, Hrabik TR, Vander Zanden MJ. 2006. Forecasting the Spread of Invasive Rainbow Smelt in the Laurentian Great Lakes Region of North America. *Conservation Biology* **20**(6): 1740-1749
- Mills MD, Rader RB, Belk MC. 2004. Complex interactions between native and invasive fish: the simultaneous effects of multiple negative interactions. *Oecologia* **141**: 713-721.
- Moyle PB, Li HW, Barton BA. 1986. The Frankenstein effect: impact of introduced fishes on native fishes in North America. *In: Stroud RH. Ed. Fish culture in fisheries management.* Bethesda, MD, American Fisheries Society.
- Moyle PB & Light T. 1996. Biological invasions of freshwater: empirical rules and assembly theory. *Biological conservation* **78**: 149-161
- Moyle PB & Cech JJ. 2004. *Fishes: An Introduction to Ichthyology.* 5th edition. Prentice Hall. U.S.A. 612p

- Moyle PB & Marchetti MP. 2006. Predicting Invasion Success: Freshwater Fishes in California as a Model. *BioScience* **56**(6): 1-10
- Nordlie FG & Mirandi A. 1996. Salinity relationships in a freshwater population of eastern mosquitofish. *Journal of Fish Biology* **49**: 1226-1232
- Olden JD & Jackson DA. 2002. A comparison of statistical approaches for modelling fish species distributions. *Freshwater Biology* **47**:1976-1995
- Peterson AT, Soberon J, Sanchez-Cordero V. 1999. Conservatism of Ecological Niches in Evolutionary Time. *Science* **285**: 1265-1267
- Peterson AT & Vieglais DA. 2001. Predicting Species Invasions Using Ecological Niche Modeling: New approaches from Bioinformatics Attack a Pressing Problem. *Bioscience* **51**(5): 363-371
- Pyke GH. 2005. A review of the biology of *Gambusia affinis* and *G. holbrooki*. *Reviews in Fish Biology and Fisheries* **15**: 339-365
- Quinn GP & Keough MJ. 2002. Experimental Design and Data Analysis for Biologists. Cambridge University Press 537p
- Rahel FJ. 2002. Homogenization of freshwater faunas. *Annual Review of Ecology Evolution and Systematics* **33**: 291-315
- Rehage JS & Sih A. 2004. Dispersal behaviour, boldness, and the link to invasiveness: a comparison of four *Gambusia* species. *Biological invasions* **6**: 379-391
- Ruesink JL, Parker IM, Groom MJ, Kareiva PM. 1995. Reducing the Risks of Nonindigenous Species Introductions. *BioScience* **45**(7): 465-477
- Rumelhart DE, Hinton GE, Williams RJ. 1986. Learning internal representations by back-propagating errors. *Nature* **323**: 533-536
- Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, Weller SG. 2001. The population biology of invasive species. *Annual Review of Ecology Evolution and Systematics* **32**: 305-332
- Simberloff D. 1996. Impacts of introduced species in the United States. *Consequences: The Nature and Implications of Environmental Change* **2**:13-22.
- Simberloff D & Stiling P. 1996. How Risky is Biological Control? *Ecology* **77**(7): 1965-1974
- Simon KS & Townsend CR. 2003. Impacts of freshwater invaders at different levels of ecological organisation, with emphasis on salmonids and ecosystem consequences. *Freshwater Biology* **48**: 982-994
- Stockwell DRB & Peters D. 1999. The GARP modelling system: problems and solutions to automated spatial predictions. *International Journal of Geographical Information Science* **13**: 143-158

- Thuiller W, Araujo MB, Lavorel S. 2003. Generalized models vs. classification tree analysis: Predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science* **14**: 669-680
- Thuiller W. Impact des changements globaux sur la biodiversité en Europe: projections et incertitudes. Th : Biologie des Populations et Ecologie. Soutenue le 15 Décembre 2003 à Montpellier. Université de Montpellier II. 105 p.
- Unmack P. 1998. Gambusia control homepage. [en ligne]. Disponible sur <<http://www.gambusia.net/>> (consultée le 04/04/07)
- Vitousek PM, D'Antonio CM, Loope LL, Westbrooks R. 1996. Biological invasions as global environmental change *American Scientist* **84** : 218-228
- Vitousek PM, Mooney HA, Lubchenco J, Melillo JM. 1997. Human Domination of Earth's Ecosystems. *Science* **277**: 494-499
- Wiens JJ & Graham CH. 2005. Niche Conservatism : Integrating Evolution, Ecology, and Conservation Biology. *Annual Review of Ecology Evolution and Systematics* **36**: 519-539
- Winkler P. 1979. Thermal Preference of *Gambusia affinis affinis* as Determined under Field and Laboratory Conditions. *Copeia* **1979**(1): 60-64
- Yee TW & Mitchell ND. 1991. Generalized Additive Models in Plant Ecology. *Journal of Vegetation Science* **2**(5): 587-602
- Zambrano L, Martinez-Meyer E, Menezes N, Peterson AT. 2006. Invasive potential of common carp (*Cyprinus carpio*) and Nile tilapia (*Oreochromis niloticus*) in American freshwater systems. *Canadian Journal of Fisheries and Aquatic Sciences* **63**: 1903-1910

Annexe I : Présentation de la structure d'accueil

Groupe d'Ecologie Comportementale et Structure des Communautés (GECCO)
Laboratoire Evolution & Diversité Biologique (UMR-CNRS 5174)
Bâtiment 4R3. Université Paul Sabatier Toulouse III
118, route de Narbonne
31062 Toulouse cedex 4, France.

Le laboratoire Evolution et Diversité Biologique (EDB) travaille sur des sujets à l'interface entre l'écologie et l'évolution. Différentes approches sont utilisées: biologie moléculaire, génétique des populations, macroécologie (modélisation mathématique), étude des interactions durables (symbiose et coévolution)... Pour cela, des études théoriques, expérimentales et de terrain sont effectuées.

EDB compte environ 40 chercheurs et enseignants chercheurs. Ces derniers sont membres de l'IFR 40 "Agrobiosciences, Interactions et Biodiversité", et de la Génopôle Toulouse Midi-Pyrénées. Ils participent au master 2 Recherche "Biologie Evolutive et Ecologie" et à l'Ecole Doctorale SEVAB (Sciences Ecologiques Vétérinaires Agronomiques Bioingénieries).

Le laboratoire s'articule autour de trois groupes de recherche :

- Génétique et Evolution de la Biodiversité. *Responsable : Christophe Thébaud.*
- Coévolution et Interaction. *Responsable : Monique Gardes.*
- Groupe d'Ecologie Comportementale et Structure des Communautés (GECCO)

Responsable : Philipp Heeb.

Groupe d'Ecologie Comportementale et Structure des Communautés (GECCO)

Ce groupe de recherche étudie le rôle des facteurs écologiques qui influencent les comportements individuels, la dynamique des populations, la spéciation et la structure des communautés. Plusieurs approches sont utilisées : modélisation et expérimentation en laboratoire sur différents modèles biologiques (oiseaux-reptiles-poissons-insectes). Cependant, une part importante de la recherche est consacrée à l'étude des systèmes aquatiques continentaux. Il faut noter également que ce groupe possède une haute qualification en biostatistiques et modélisation.

Les thématiques de recherche abordées sont :

- Interactions durables et co-évolution
- Ecologie comportementale
- Dynamique des populations
- Structure des communautés et macroécologie

Annexe II : Bilan personnel du stage

Le sujet de mon stage m'a permis d'acquérir des connaissances et des compétences nouvelles :

- techniques statistiques non abordées au cours de mon cursus universitaire
- utilisation du logiciel R et apprentissage de la programmation
- recherche de sources bibliographiques pertinentes
- aptitude accrue de compréhension de la littérature scientifique
- connaissances théoriques en biologie et écologie des espèces invasives
- rédaction d'un rapport scientifiquement rigoureux

Ensuite, ce stage a également été l'occasion pour moi de :

- intégrer une équipe de recherche et connaître les différents axes de recherche existants au sein du laboratoire.
- assister à des conférences (en anglais) au sein du laboratoire
- participer à des campagnes de terrain (ne faisant pas partie du stage lui-même) dans le cadre de problématiques abordées par des chercheurs du laboratoire
 - Impact des changements climatiques sur la reproduction des poissons
 - Impact de la fragmentation sur les écosystèmes aquatiques

Techniques utilisées lors du terrain: pêche électrique, identification et mesures d'espèces de poissons.

Mon stage a fait l'objet d'un prolongement d'un mois et demi (jusqu'à fin juillet 2007) afin de :

- approfondir le sujet
- compléter la base de données de l'ANR « Freshwater Fish Diversity »

Bilan

Ce stage était mon premier stage dans un laboratoire de recherche. J'avais choisi volontairement un sujet de stage abordant la problématique des invasions biologiques. Celui-ci m'ayant beaucoup plu, il confirme mon envie d'intégrer une équipe de recherche travaillant sur ce thème plus tard.

Prédiction des risques de colonisation de la gambusie *Gambusia affinis* (Poeciliidae) à l'échelle mondiale

Résumé

Les hydrosystèmes continentaux sont aujourd'hui menacés par l'invasion d'espèces exotiques. La gambusie (*Gambusia affinis*) a été introduite dans le monde entier et est susceptible de perturber de nombreux écosystèmes. L'identification des bassins colonisables par ce poisson est donc nécessaire afin de limiter sa dispersion. La niche écologique de la gambusie a ici été déterminée afin d'identifier les bassins colonisables et d'estimer son impact potentiel sur les assemblages natifs. La niche réalisée a été identifiée par une Analyse en Composantes Principales (ACP) en incluant aire native et zone d'invasion actuelle de l'espèce. La niche native étant significativement différente de la niche exotique, les bassins potentiellement colonisables ont été identifiés à partir de l'ensemble de la niche réalisée. Les prédictions ont été obtenues par des modèles linéaires généralisés (GLMs) et des modèles additifs généralisés (GAMs). Les deux modèles prédisent respectivement 154 et 121 bassins colonisables. GLMs et GAMs sont de bonne qualité ($AUC > 0,75$) bien qu'ils ne prédisent pas toujours les mêmes bassins. L'étude des traits biologiques des espèces natives des bassins colonisables a permis de lister 18 espèces dans 13 bassins pouvant être menacées par la gambusie. Cependant, les bassins colonisables devront faire l'objet d'études à échelle plus fine afin d'estimer les véritables dangers de l'introduction de la gambusie. Les résultats obtenus suggèrent l'importance de considérer la niche réalisée exotique dans la définition de la niche potentielle de l'espèce. Nous recommandons également l'emploi de plusieurs modèles et indices de qualité ainsi qu'une vérification des prédictions par différentes stratégies.

Mots clefs : invasions biologiques, modèles linéaires généralisés, modèles additifs généralisés, bassins hydrographiques, niche écologique

Global scale prediction of mosquitofish *Gambusia affinis* (Poeciliidae) colonization risks

Abstract

Freshwater ecosystems are threatened by exotic species invasion. Mosquitofish (*Gambusia affinis*) was introduced all over the world and is susceptible to disturb ecosystems integrity. Identifying basins that may support mosquitofish populations is necessary to control its dispersion. The ecological niche of *Gambusia* was defined by Principal Components Analysis (PCA) by including its native and actual invaded range. Because the native and exotic ranges were significantly different, sustainable basins for *Gambusia* were identified from the entire realized niche. Predictions were derived from Generalized Linear Models (GLMs) and Generalized Additive Models (GAMs). These models respectively predict 154 and 121 sustainable basins. GLMs and GAMs have good predictive power ($AUC > 0.75$) although the predicted basins are not always the same according to the modelling method. Among basins with a high potential to be invaded by *Gambusia*, 18 species in 14 different basins may be threatened. However, finer-scale studies will be necessary to forecast the real impact of *Gambusia* in these basins. Our results underlines the importance to consider the exotic realized niche together with native niche in the aim to accurately approach the definition of the potential niche of the species. We also recommend using several methods, quality indices and to verify the predictions by comparing different estimators of the model predictions.

Keywords: biological invasions, Generalized Linear Models, Generalized Additive Models, river basins, ecological niche