



HAL
open science

Propositions pour améliorer l'équipement biométrique du détective écologique - Application à la modélisation de la relation entre gestion forestière et biodiversité

Frédéric Gosselin

► To cite this version:

Frédéric Gosselin. Propositions pour améliorer l'équipement biométrique du détective écologique - Application à la modélisation de la relation entre gestion forestière et biodiversité. Sciences de l'environnement. HDR Spécialité Biologie, Université Pierre et Marie Curie Paris, 2011. tel-02594486

HAL Id: tel-02594486

<https://hal.inrae.fr/tel-02594486v1>

Submitted on 15 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Pierre et Marie Curie
Mémoire d'Habilitation à Diriger des Recherches

Frédéric GOSSELIN

**Propositions pour améliorer l'équipement
biométrique du détective écologique**
*Application à la modélisation de la relation entre
gestion forestière et biodiversité*

soutenu le 11 janvier 2011, devant le jury composé de :

M. Hervé JACTEL	Directeur de Recherche INRA	Rapporteur
M. Eric PARENT	Enseignant-Chercheur PR / ICPEF	Rapporteur
M. Robert BARBAULT	Professeur Paris 6	Examineur
M. Jean-Dominique LEBRETON	Directeur de Recherche CNRS	Examineur
M. Claude MILLIER	Retraité du corps des Directeurs de Recherche INRA	Examineur

Remerciements

Le travail autour de ce rapport a été rendu possible par la décharge de certaines responsabilités – d'animation scientifique notamment – qui m'a permis de passer davantage de temps à l'encadrement des étudiants et au projet scientifique. Je remercie donc mes supérieurs hiérarchiques locaux et nationaux qui ont accepté ce contrat, ainsi que mes collègues du Cemagref et particulièrement Laurent Bergès et Frédéric Archaux qui ont pris ma suite en tant qu'animateurs d'équipe. Je remercie aussi ceux qui m'ont plus que secondé pour l'animation des projets, en particulier Yoan Paillet pour le projet Gestion-Naturalité-Biodiversité, et tous ceux qui ont aidé à leur bonne réalisation : Richard Chevalier, Yann Dumas, Carl Moliard, Guillaume Legoff, Marc Roussel, Gérard Goujon, Benoît Nusillard, Jean-François Abgrall, Hilaire Martin, Françoise Dumas, Gilles Bergevin, Philippe Guillemard, Vickie Moore, sans oublier ceux déjà cités par ailleurs.

Faisant le point sur le travail réalisé depuis mon arrivée au Cemagref, mes pensées vont aussi aux collègues qui nous ont quittés, durant leur activité ou juste après – Claude-Bernard Malphettes, Jacques Lhonoré, Jean Mosnier, Antoine Gama, Gérard Pinon, Jean-Claude Rameau. A tous, un grand merci pour votre implication et votre professionnalisme.

Les points de vue proposés dans ce mémoire sont le fruit de discussions avec les étudiants que j'ai eu la chance d'encadrer et avec des collègues. Pêle-mêle me reviennent notamment à l'esprit :

– les questions d'Elfie Perdereau, en stage Master2, sur les p-valeurs de qualité d'ajustement, sources du travail résumé dans la partie 4.2 ;

– la suggestion de Christophe Bouget, à l'issue de son DEA, de passer aux modèles mixtes : ce point n'est pas traité directement dans ce mémoire mais émergera inévitablement dans les années à venir ;

– le travail sur les modèles linéaires et non-linéaires mixtes gaussiens, avec notamment Emmanuelle Dauffy-Richard, Ghislain Vieilledent et Julien Delcher. Ce travail n'a pas été repris en tant que tel dans ce rapport puisque je suis maintenant convaincu par l'intérêt d'une modélisation plus rigoureuse sur le plan de la qualité probabiliste (cf. 4.3). Les nombreuses discussions autour des modèles mixtes – et de leurs impasses – ont clairement été à la base des développements probabilistes de ce rapport de HDR. Il en va de même des échanges avec Stéphane Barbier sur les modèles Bayésiens de données de comptage ;

– la réflexion épistémologique menée par Emmanuelle Dauffy-Richard dans sa thèse m'a incité à poursuivre le travail entrepris alors par une réflexion sur les modèles (cf. parties 2 et 3) ;

– enfin, les nombreuses discussions écologiques avec les étudiants que j'ai encadrés, en particulier Emmanuelle Dauffy-Richard, Christophe Bouget et Stéphane Barbier, m'ont été fort utiles même si elles trouvent peu d'échos dans ce rapport davantage orienté vers les méthodes.

Toute ma reconnaissance va aussi à l'équipe de Jean-Dominique Lebreton au CEFÉ, et plus largement au CEFÉ, à Jean Jacod et à Claude Millier : il est clair dans mon esprit que mon goût pour la recherche s'est ancré avec eux, pendant mes années de thèse.

Un grand merci aussi à Marc Lavielle et à Eric Parent pour m'avoir accueilli dans leurs réseaux ou groupes de travail (respectivement MONOLIX et APPLIBUGS/Rochebrune) : tous les deux m'ont beaucoup apporté pour ce rapport de HDR (je pense notamment aux parties 3, 4.2 et 4.3).

Enfin, mes plus grands remerciements vont à ma famille : Marion, mon épouse, pour son amour et son soutien ; mes enfants, Jean, Eliane et François, pour leur gaieté et leur amour. Merci pour votre patience quand j'avais l'esprit encore un peu trop rempli d'équations ou d'hypothèses... Un grand merci aussi à mes parents qui m'ont donné, par leur exemple, le goût des études et de la recherche appliquée.

De manière plus large, merci à tous ceux qui autour de moi se donnent pour rendre au quotidien la vie plus belle – que ce soit parmi mes collègues du Cemagref et des autres organismes sur le domaine des Barres, ou parmi mes amis et connaissances à Nogent et au-delà.

Je remercie enfin les relecteurs des premières versions de ce rapport : Marion Gosselin, Jean-Dominique Lebreton, Christophe Bouget, Frédéric Archaux et Thomas Cordonnier.

1. Introduction

Une partie non négligeable des défis que nos sociétés contemporaines doivent affronter met en question la viabilité des activités de l'homme "moderne" (cf. Bourg, 1996 pour le contenu de ce mot), dans les domaines tant économique et social qu'environnemental. Si l'époque moderne s'est bâtie sur l'hypothèse d'une nature illimitée non affectée par les activités humaines, nous remettons aujourd'hui en cause cette hypothèse et sommes amenés à inclure la nature dans nos choix de société, notamment parce que la technologie et la science nous fournissent une puissance capable de perturber les cycles régulateurs de la biosphère (Bourg, 1996, Badré et Décamps, 2005). Certes, des incertitudes demeurent quant à la réalité de ces menaces : ces problèmes sont scientifiquement complexes et la perception des menaces est pour partie le fruit de processus sociaux (e.g. Roqueplo, 1995) marqués par une conscience contemporaine particulièrement attachée à la protection de la nature (Bourg, 1996). Pourtant, ces défis ne sont pas que sociaux : ils sont plutôt caractérisés par une dialectique entre la dimension sociale et la dimension "physique", "naturelle" ou "objective". Il serait alors dangereux de ne pas garder contact avec la dimension "objective" de ces problèmes, sous prétexte qu'ils comprennent une composante sociale forte (Roqueplo, 1995).

Ainsi, de nombreuses politiques tentent d'enrayer la "6^{ème} crise d'extinction" et l'érosion de la biodiversité qui pourrait entraîner une formidable homogénéisation du monde vivant, et par là d'autres conséquences non forcément toutes identifiées. Or, ces politiques sont très souvent à la fois orphelines et peu productrices de données pertinentes renseignant sur l'état de la biodiversité (Yoccoz *et al.*, 2001, Nichols et Williams, 2006, Pereira et Cooper, 2006, Boutin *et al.*, 2009), hormis pour les espèces menacées d'extinction à travers l'UICN – pour les taxons "supérieurs". Les phases de construction et d'évaluation de ces politiques s'en trouvent appauvries, surtout en comparaison avec des domaines riches en données comme les domaines climatiques et météorologiques. Pourtant, les quelques cas où des données sont venues enrichir la gestion ou la politique publique de conservation de la biodiversité devraient nous entraîner dans cette direction. Par exemple, en Australie, des données de grande envergure pour de nombreuses espèces ont été analysées avec des modèles statistiques paramétriques ou semi-paramétriques pour ensuite choisir des réserves (par Ferrier *et al.*, 2002). Les auteurs ont montré tout l'intérêt de passer par une modélisation espèce par espèce pour choisir des réserves contenant le plus d'espèces plutôt que par des modèles de type de forêt – approche habitat – ou des modèles multivariés. Un autre exemple où les données et

leur interprétation statistique – démographiques cette fois – ont joué un rôle important dans le choix d'une politique et son évaluation concerne la gestion forestière mise en place dans l'aire de répartition de la chouette tachetée nordique (cf. Gosselin, 2009).

Le "détective écologique" (Hilborn et Mangel, 1997) manque ainsi d'"indices", particulièrement quand il est engagé à l'interface entre politique et recherche, pour choisir ou pour évaluer des politiques publiques de conservation par exemple. Mais ce n'est pas uniquement de données dont il a besoin. Il lui faut aussi des outils affûtés pour analyser ces données. C'est dans cette direction que se déploient les recherches que je présente ici : au cours de mes recherches sur le lien entre gestion forestière et biodiversité, j'ai été amené à améliorer progressivement les outils statistiques à la disposition du détective écologique, dans trois directions. Premièrement, j'insiste sur l'étude de la magnitude des effets plutôt que de leurs significativité statistique (partie 4.1), avec des modèles paramétriques Bayésiens (partie 3). Ce premier développement n'est pas nouveau en écologie, même s'il est peu mis en pratique. Il devrait pourtant permettre un meilleur lien avec les gestionnaires et les décideurs. Mon second apport consiste à proposer d'utiliser de nouvelles distributions de probabilité de comptages permettant notamment de modéliser des données sous-dispersées par rapport à la loi de Poisson en même temps que des données sur-dispersées (partie 4.3). A ma connaissance, ce point est nouveau en écologie ; il ne l'est pas en statistiques appliquées, où de nombreuses distributions de ce type ont déjà été proposées. Je montre néanmoins que la plupart de ces distributions ont des limites pratiques importantes rendant leur utilisation problématique en régression par exemple. La troisième avancée proposée ici est un nouvel outil de critique de modèle statistique – ou d'évaluation de la qualité d'ajustement (la p-valeur postérieure échantillonnée ; partie 4.2). Cet outil est nouveau en statistique et en écologie. Il a l'avantage d'être assez universel au sein des modèles statistiques paramétriques. On peut ainsi jauger l'ajustement du modèle aux données sous des angles très variés. J'éclaire ici ces travaux d'amélioration statistique à la lumière d'une réflexion épistémologique sur la manière de concevoir et d'utiliser les modèles, et sur la place relative des mathématiques et du domaine d'application – en l'occurrence l'écologie – dans ce processus (partie 2). J'ai inclus dans cette réflexion la perspective de Duhem-Quine, peu présente en écologie.

Mon propos m'amènera donc dans un premier temps à poser quelques jalons sur la place des modèles dans les sphères académique et décisionnelle (partie 2). J'expliquerai ensuite les intérêts et limites des modèles statistiques paramétriques Bayésiens dans ces deux sphères (partie 3). Puis je présenterai les trois principales améliorations statistiques de ce rapport

(partie 4). Dans ce rapport de HDR, j'ai privilégié la partie méthodologique de mon travail de recherche. Les autres parties de mon activité, qui traitent de questions écologiques ou des relations entre recherche écologique et gestion/décision, sont réintégrées lors de l'évocation des pistes de recherche à moyen terme (partie 5).

2. Modèles et méthode : Quelques points de repère épistémologiques

Avant d'aller plus loin sur le sujet, faisons le point sur ce que contient la trousse à outil du détective écologique. Bien entendu, je ne prétends pas à l'exhaustivité du contenu mais souhaite présenter d'un point de vue épistémologique deux outils centraux : les modèles ; et les méthodes d'utilisation des modèles en écologie.

2.1. Modèles : les intermédiaires

Le terme de modèle a plusieurs significations qu'il est peut-être vain de vouloir réduire (Le Moigne, 1989, Legay, 1997 pp. 19-23, Huber, 2002a). Néanmoins, en insistant à la suite de Legay (1997) sur le caractère instrumental des modèles – et en espérant ne pas trop diverger sur le plan sémantique, nous entendrons ici par modèle une représentation ou imitation de la réalité, représentation dans notre cas mathématique.

2.1.1. Le modèle comme représentation

Un modèle ne vise pas à reproduire parfaitement la réalité. Il en est une représentation (Legay, 1997, Bouleau, 1999, Beven, 2002, Huber, 2002a). La phrase de Paul Valéry illustre bien la philosophie de la modélisation : "ce qui est simple est toujours faux. Ce qui ne l'est pas est inutilisable". C'est dans cet entre deux entre faux et inutilisable, dans l'écart entre modèle et réalité, et dans la manière d'appréhender cet écart que se noue le problème de la modélisation (Legay, 1997). Nous reviendrons sur ce point par la suite.

On conçoit souvent un modèle comme la mise en exergue de certains phénomènes ou mécanismes, qui sont "modélisés". Pourtant, au moins dans une partie des cas, c'est peut-être aussi sur le contraire que s'appuie la modélisation, i.e. sur la "négligence" de certains phénomènes (Nouvel, 2002). Le modèle étant une représentation de la réalité associée à des

phénomènes ou structures mis en exergue ou au contraire négligés, un modèle est donc forcément subjectif, affilié au point de vue du concepteur (Legay, 1997, Bouleau, 1999¹).

2.1.2. Composantes du processus de modélisation

Si la modélisation est probablement un "art" qu'il n'est pas si simple de maîtriser et pour lequel bien des outils restent à développer (Bouleau, 1999), les grandes phases de développement d'un modèle sont *a priori* connues. La modélisation comprend plusieurs composantes, soit temporelles soit méthodologiques :

– les **intrants du modèle**. Ce sont les éléments qui ne sont pas forcément dans le modèle lui-même mais qui ont conditionné son développement. Parmi eux, il y a des intrants scientifiques et techniques : les données et connaissances disponibles dans la discipline d'application, mais aussi en mathématiques appliquées – par exemple, le "réservoir de modèles" mentionné par Lehmann (1990) en parlant des approches de la modélisation de Fisher et de Neyman. Un autre intrant d'importance est – si elle existe formellement – la demande du modèle, aboutissant au cahier des charges du modèle et notamment à la définition de ses objectifs ;

– la définition de la **structure du modèle et de certains de ses paramètres**. Il faut d'abord définir la nature des composantes du modèle – par exemple la nature des données modélisées, leur transformation ou pas. C'est un choix important dans le cas de la biodiversité, et il faut probablement le faire en fonction des objectifs du modèle. Par exemple, si un modèle veut être utilisé pour prédire une extinction, mieux vaut travailler sur des données discrètes incluant des zéros (cf. Gosselin, 1997). En statistiques paramétriques et en probabilité, en plus de la nature des données modélisées, il faut souvent leur attribuer une famille de distributions probabilistes. La conjonction entre le choix de la nature des données et le choix des distributions n'est pas anodine : elle définit des familles de modèles non réductibles les unes aux autres (Bouleau, 1999, pp. 314-315). D'autres choix à cette étape concernent les relations entre les composantes du modèle : s'agit-il de relations linéaires, exponentielles, logarithmiques... ? Ce choix est aussi d'importance en pratique.

– le **traitement qui sera fait du modèle par le modélisateur** (statistique, mathématique, décisionnel...). Les mathématiques peuvent fournir des outils pour analyser

¹ "Finalement, qu'est-ce qu'un modèle? A cette question je propose de répondre un simulacre utile. Il me semble crucial de mettre en avant que les modèles sont des représentations partisans, au bon sens du terme, c'est-à-dire qu'au sein de multiples possibilités d'expression et de représentation, ils sont le choix d'un parti." (Bouleau, 1999, p.301)

plus efficacement les modèles : connaître les propriétés mathématiques du modèle, par exemple asymptotiques, indépendantes des conditions initiales permet souvent d'économiser du temps et de l'énergie dans l'analyse des résultats (Bouleau, 1999) ; un des objectifs de ma thèse était de montrer l'existence de ce type de convergences (Gosselin, 1997). C'est aussi ici que pourra s'effectuer une partie de la validation du modèle statistique, qui consiste à tester sa cohérence interne ou externe, par exemple la cohérence des données avec les différentes hypothèses distributionnelles faites. C'est aussi bien sûr le moment où on extrait les résultats du modèle, et où on met en forme l'éventuelle interface du modèle avec le monde extérieur ;

– **éventuellement, l'utilisation qui sera faite du modèle ou de ses résultats par d'autres utilisateurs** que le modélisateur. Qui utilisera le modèle ? Avec quelle compétence et quelle capacité à analyser, se former ? Ces questions sont très importantes à avoir en tête lors du développement du modèle. "Avant de s'engager dans une modélisation l'ingénieur doit impérativement se demander à qui va servir le modèle, s'il s'agit d'un organisme, quels sont ses buts, ses moyens d'actions et ses capacités à faire des mesures. Ce dernier point est primordial. ... Les modèles les plus grossiers sont simples à valider alors que les modèles fins nécessitent tant de mesures pour préciser les fonctions inconnues qu'en pratique leur validation est très mauvaise. Or leur logique déductive et les phénomènes qu'ils font apparaître dépendent de ces choix" (Bouleau, 1999, p. 272). Les critères suivant lesquels se fait la validation du modèle, la qualification de sa qualité, ne semblent pas encore stabilisés d'après Bouleau (1999, p. 279, p. 284)². Les outils d'évaluation de la qualité de l'ajustement dont il sera question plus loin (cf. 4.2) peuvent aider à améliorer cette pratique, pour ce qui est des modèles statistiques.

Sur tous ces points, il est important de connaître le point de vue adopté lors de la construction du modèle (Legay, 1997, Bouleau, 1999). En effet, ce n'est pas parce qu'il emprunte le langage mathématique – *a priori* gage de rigueur – qu'un modèle est neutre : la modélisation utilise extrêmement souvent les mathématiques comme "langage entre acteurs sociaux exactement au même titre que le langage ordinaire, aussi tendancieux, aussi trompeur parfois, mais aussi utile, indispensable concrètement" (Bouleau, 1999, p. 268). Ce constat, assez partagé dans la communauté scientifique, n'empêche pas que la société accorde

² "Il apparaît surtout que chacun [des modèles] ne saurait être critiqué en dehors de l'utilisation sociale réelle du modèle. Nous sommes dans une période transitoire où les critères de jugement de la qualité de ce nouveau travail de l'ingénieur sont en cours d'émergence"

davantage de crédit aux modèles qui utilisent le formalisme mathématique – confusion fréquente entre outils technologiques et outils scientifiques (cf. Encadré).

Les connaissances non mathématiques injectées comme intrants dans le modèle sont en général cruciales. Si certains vont jusqu'à identifier modélisation et mise en équation, donc modélisation et mathématiques appliquées (anonyme, 1997), il s'agit selon moi d'une vision réductrice du modèle. Je pense en effet que le modèle doit être considéré aussi en fonction du corpus de connaissances dans la discipline où intervient le modèle (anonyme, 1997) ainsi que de ses objectifs – scientifiques ou décisionnels. C'est en ce sens que la modélisation est plutôt une activité d'ingénieur ou de chercheur non mathématicien – éventuellement en lien avec des spécialistes d'autres disciplines – que de mathématicien appliqué.

Ecologie et mathématiques, bidisciplinarité et indisciplinarité

Je réfléchis ici à la bonne prise en compte des "énoncés" mutuels des deux disciplines – mathématiques et écologie, dans le cadre de ce que j'appelle une "bidisciplinarité active". De manière générale, les écologues recherchent parfois un peu trop de "sens", de contenu dans les mathématiques. Pire, ils peuvent être amenés comme dans d'autres disciplines à utiliser des résultats mathématiques comme arguments d'autorité faisant taire le débat ou l'esprit critique (Lévy-Leblond, 1977, Bouleau, 1999³, Morad, 2004). Or, je partage le point de vue de Poincaré (1905) selon lequel les **mathématiques permettent avant tout de fournir aux autres disciplines des "cadres vides" extrêmement utiles**, mais dont le contenu est, dans notre cas, biologique ou écologique⁴ (cf. Bouleau, 1999, références dans Morad, 2004

³ " [Le langage de l'ingénieur] perd par rapport au langage de la géométrie l'universalité, mais peut-être, peut-il faire émerger un nouvel esprit de finesse et une réconciliation de l'abstrait et du concret? Cela ne sera possible que si l'on dénonce avec la plus grande vigueur l'emploi de mathématiques savantes aux seules fins d'impressionner par l'apparence de la scientificité... Il s'agit là, pour le moins, de graves fautes de style, ces modélisations utilisent les mathématiques comme les médecins de Molière utilisaient le latin, elles sont très dommageables à la communication, comme tout ce qui détériore le langage". (Bouleau, 1999, p.307).

"Il ne faudrait pas croire que la présence de mathématique dans le modèle ait une quelconque vertu purificatrice à cet égard. Au contraire, les mathématiques sont souvent une bienséance, une politesse, qui peut cacher n'importe quoi" (Bouleau, 1999, p. 293).

⁴ "L'analyse mathématique, dont l'étude de ces cadres vides est l'objet principal, n'est-elle donc qu'un vain jeu de l'esprit? Elle ne peut donner au physicien qu'un langage commode; n'est-ce pas là un médiocre service, dont on aurait pu se passer à la rigueur; et même, n'est-il pas à craindre que ce langage artificiel ne soit un voile interposé entre la réalité et l'œil du physicien? Loin de là, sans ce langage, la plupart des analogies intimes des choses nous seraient demeurées à jamais inconnues; et nous aurions toujours ignoré l'harmonie interne du monde, qui est, nous le verrons, la seule véritable réalité objective. " (Poincaré, 1905, p.22)

"Les mathématiques interviennent comme langue interdisciplinaire. C'est la seule qui puisse jouer ce rôle, car c'est la seule qui ne soit pas pieds et poings liés à une sémantique" (Bouleau, 1999, p. 324 + p. 329).

pour des points de vue similaires) . Les mathématiques sont ainsi un langage utile aux autres sciences plus qu'une science en elle-même – si l'on retient comme critères de la science la confrontation à la réalité; la réfutabilité par la réalité...⁵

Premier cas de figure : Les mathématiques pourront par exemple montrer l'indépendance asymptotique d'un objet mathématique par rapport aux conditions initiales, ou l'existence d'un objet donné, et ce sous certaines hypothèses, mais qui restent assez générales d'un point de vue biologique ou écologique. C'est ce que j'ai fait de faire durant ma thèse, en démontrant les propriétés asymptotiques d'une large classe de modèles d'extinction d'espèces, sous la direction de Jean-Dominique Lebreton, biométricien, et Jean Jacod, mathématicien (Gosselin, 2001a).

Second cas de figure : à l'inverse, certains développements mathématiques peuvent être faits sous des hypothèses assez contraignantes, et l'écologue autant que le mathématicien doit vérifier avec soin l'adéquation des hypothèses avec le cas étudié. C'est ce que j'ai montré sur deux exemples durant ma thèse (Gosselin, 1998, Gosselin, 1999). Notons que le fait de ne pas vérifier les hypothèses – ou de ne pas réfléchir à leur signification dans la discipline d'application – est présenté comme une des explications des dernières crises financières que nous avons vécues (Bouchaud, 2008).

Ces deux cas de figure – développer des résultats mathématiques sous contrainte d'utilité pratique et vérifier les hypothèses mathématiques avant d'appliquer les résultats mathématiques – me semblent devoir être deux des briques d'une bi-disciplinarité active.

On pourrait, à la suite de Jean-Marie Legay (1986), aller plus loin, en abordant la notion d'indisciplinarité, qui s'apparente dans une version "sage" au développement d'écotones entre disciplines – Legay parle de zones d'intersection –, et dans une version plus audacieuse – et souvent moins acceptée – à un travail scientifique qui est tout simplement en dehors des

"Que deviennent les mathématiques alors ? Essentiellement un exemple de perfection du langage qui laisse complètement ouverte leur référence sémantique... Les mathématiques nous donnent l'exemple d'un langage précis grâce auquel les mathématiciens s'entendent." (Bouleau, 1999, p.168)

⁵ "Qu'une structure mathématique ait valeur universelle, comme telle, en ce qui concerne la réalité empirique ne va pas de soi. Ce ne sont d'ailleurs pas les mathématiciens qui le revendiquent, mais plutôt une certaine métaphysique tirée d'une interprétation des mathématiques qui met de toute évidence en jeu la conception même de la réalité empirique" (Mathiot, 2002, p.230).

disciplines. Je n'ai pas l'expérience de Jean-Marie Legay sur ces questions, et si ses analyses me semblent tout à fait intéressantes, j'insisterai sur un point qui m'est apparu souvent durant ma thèse : il y a une opposition entre une bi-disciplinarité active – telle que présentée ci-dessus..., marquée par un respect mutuel des deux disciplines – et une bi-disciplinarité à sens unique, où souvent des résultats mathématiques existants – valables sous des hypothèses simplificatrices fortes – sont appliqués sans beaucoup de précautions à des cas concrets. Cette vision est partagée par North et Jeffers (1990, p.530 ⁶) et – si je l'interprète bien, par Austin (1979) ⁷.

2.2. Méthodes : méthode des modèles

En préambule, signalons que des modèles ou ensemble de méthodes peuvent être utilisés par des praticiens (ingénieurs, techniciens...) en dehors de leur domaine de validité et sans qu'on sache clairement l'importance de la déviation entre modèle et réalité : c'est ce que Bouleau (1999) qualifie de "sciencette", corpus de connaissances et de méthodes bien installées dans un domaine, facilement enseignable et utile aux praticiens – du moins tant que des démentis cinglants ne la contredisent pas. Il en va probablement de même pour certaines méthodes utilisées par les scientifiques : consensuelles à un moment donné mais pas forcément bien justifiées, elles sont utilisées ... tant qu'elles ne sont pas remises en cause. Je m'efforcerai ici de préciser les fondements de ce que pourrait être des éléments de méthode des modèles plus rationnels.

2.2.1. Place des modèles dans le processus scientifique

Souvent, le modèle est un objet intermédiaire et central dans le processus scientifique : il "s'interpose entre les phénomènes et l'interprétation que la science en donne" (anonyme, 1997). Dans le cadre de cette confrontation entre phénomènes et interprétation de la science, on peut se demander quelles sont les différences entre les notions de modèle, de théorie et d'hypothèse. Plusieurs points de vue existent sur la question.

⁶ "Slobodkin (1975) suggested that mathematicians should learn biology and only then create new and appropriate mathematics, rather than fitting poorly understood biology into mathematical moulds which were originally created for hydrodynamics, thermodynamics, economics, etc."

⁷ "Whatever happens, I hope ecologists and mathematicians will manage to avoid making errors of the third kind – "solving the wrong problem". This type of error has been prevalent in statistical ecology, particularly in ordination methodology"

A partir de Nichols (2001) et de Giere (2004), je retiendrai quatre principaux concepts opérationnels : les "principes", sur la base desquels peuvent être construits des "théories" ou des "hypothèses"⁸, lesquelles constituent⁹, ou aident à construire des "modèles" ; les "prédictions"¹⁰, issues des modèles, sont alors confrontées à la réalité sous forme de tests (cf. Figure 1). Les principes sont des énoncés généraux qui servent à construire ces modèles. Les principes ne sont en général pas testés en tant que tels mais sont comme des ferments ou des fondations aidant à créer des modèles. Ils forment ce que Peters (1991, p. 21) appelle une soupe préscientifique qui sera une source d'inspiration utile aux scientifiques. La définition du modèle est variable mais garde le trait constant d'après Giere (2004) d'une représentation de la réalité (cf. aussi 2.1). Cette présentation est compatible avec les positions épistémologiques classiques de la science, notamment celle de Popper (cf. Hilborn et Mangel, 1997)¹¹. Dans cette manière de voir les modèles, ce sont les prédictions qui sont confrontées aux données, et non les modèles directement (cf. aussi Levins in Hilborn et Mangel, 1997), ni les hypothèses ou théories, et encore moins les principes. Classiquement, ce qui fait retenir un modèle ou pas – ainsi que les théories et hypothèses qui l'ont nourri – c'est la confrontation des prédictions issues du modèle aux données, i.e. son caractère prédictif (Peters, 1991).

Il existe néanmoins d'autres manières de voir les modèles, notamment parmi les écologues théoriciens (Loehle, 1983, Caswell, 1988, Odenbaugh, 2005), qui m'amènent à proposer une modification du schéma de Giere (2004) (cf. Figure 1) : à côté des *modèles prédictifs* – cf. ci-dessus – existent des *modèles* que nous appelleront "*théoriques*" qui ne répondent pas à la même logique de confrontation aux données (Loehle, 1983, Nichols, 2001). Les fonctions des modèles théoriques sont de mieux comprendre les théories (Grimm, 1994), de les "tester" d'un point de vue purement logique ou théorique (Goldstein, 1977), et à terme,

⁸ qui d'après Nichols (2001) sont des "histoires" ou explications sur comment un système fonctionne ou répond à des actions de gestion ... Une théorie étant une hypothèse qui a survécu à plusieurs confrontations à la réalité.

⁹ cf. aussi Bouleau (1999, p.314) sur la proximité entre les notions de théorie et de modèle. Néanmoins, ce qui les distingue d'après Lewontin in Nichols (2001), c'est que: "the scientific model can be viewed as the 'calculating engine' ... needed to project the consequences of a scientific hypothesis". A la suite de nombreux auteurs, Nichols (2001) insiste par ailleurs sur l'importance du lien entre hypothèse formulée a priori et modèle portant cette hypothèse. C'est un point de la méthode de travail que j'ai adoptée (Gosselin et Gosselin, 2004) ainsi que les étudiants que j'ai eu la chance d'encadrer (Richard 2004, Bouget 2004, Delcher 2005, Perdereau 2006 et Barbier 2007 ; cf. aussi Barbier *et al.*, 2009). Sur ce sujet, je recommande notamment la partie épistémologique de la thèse d'Emmanuelle Richard (2004).

¹⁰ Giere (2004) parle pour ce niveau d'"hypothèses" ou de "généralisations", mais je préfère ici le terme de prédiction car (i) la prédiction peut être assez spécifique et pas forcément générale ; et (ii) la prédiction se situe après le modèle – elle en est issue – là où l'hypothèse est préalable au modèle dans mon schéma (et par exemple dans la vision de Nichols, 2001, qui me semble assez dominante en écologie).

¹¹ mais pourrait être contredite par l'école de Duhem-Quine (Morad, 2004). Nous ne rentrerons néanmoins pas ici dans cette discussion faute de place.

éventuellement, d'inspirer de nouveaux modèles prédictifs (Conroy, 1993) – un peu à la façon des "réservoirs de modèles" (Lehmann, 1990) faisant partie des intrants de la modélisation. Caswell (1988, p. 38) identifie de nombreuses autres fonctions de ces modèles, qui peuvent se résumer comme constituant la "science du possible" (Lewontin *in* Nichols, 2001) ¹².

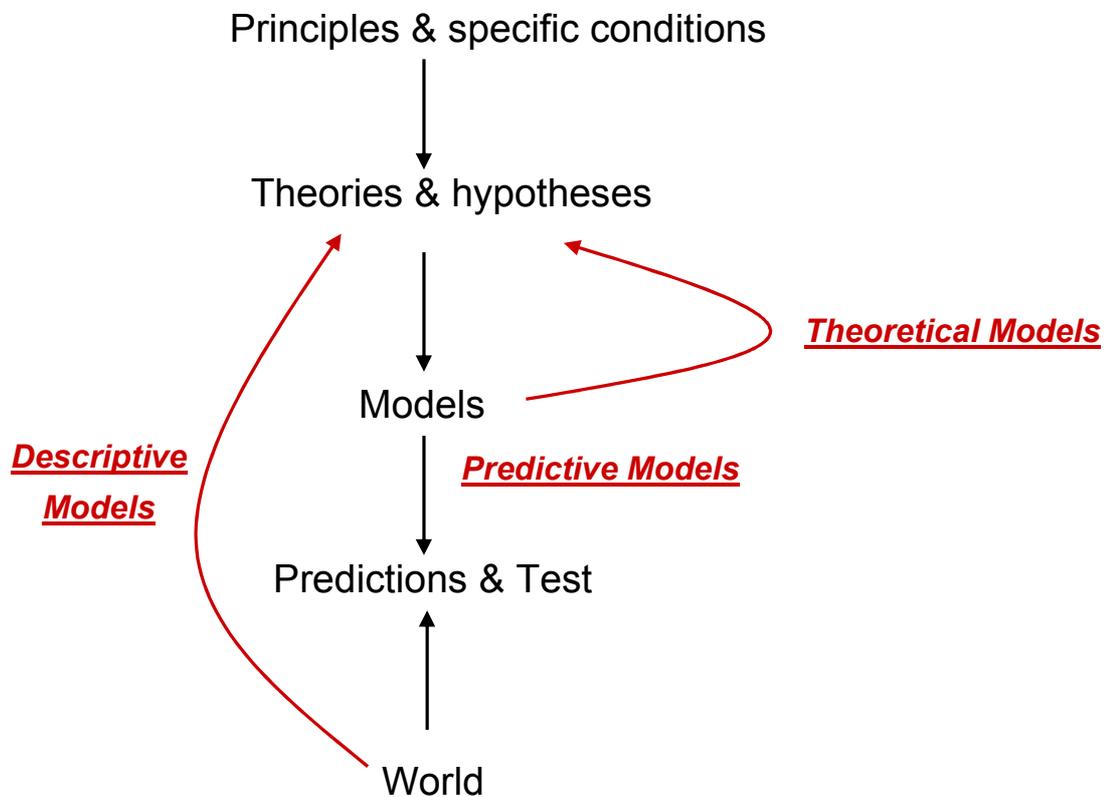


Figure 1. Représentation des modèles et de leur lien avec les théories et le monde réel selon le schéma de Giere (2004; en noir) modifié (en rouge), dans laquelle trois types de modèles sont distingués : les modèles théoriques (cf. Loehle, 1983) et les modèles descriptifs en plus des modèles prédictifs. Ce schéma s'applique avant tout à des modèles à visée scientifique dans la sphère académique ¹³.

Mais ce n'est pas tout : en plus de ces deux types de modèles – prédictifs et théoriques –, des *modèles descriptifs* peuvent être utilisés pour résumer ou synthétiser les données (cf. Figure 1). Ces modèles suivent majoritairement une logique inductive là où les modèles théoriques ou prédictifs utilisent au moins en partie une logique hypothético-déductive. De plus les modèles théoriques et certains modèles prédictifs sont majoritairement à visée explicative, c'est-à-dire essaient d'identifier des mécanismes ou causes de ce qui est observé. Les modèles descriptifs et les modèles explicatifs ont tous les deux un rôle à jouer dans les

¹² Les modèles théoriques font donc partie du monde n°3 selon Popper, monde à la fois abstrait mais ayant son existence propre à côté du monde "réel" (n°1) et du monde constitué par notre intersubjectivité (n°2; Stamos, 2007). Ils semblent en ceci faire partie du même monde que les mathématiques par exemple, et fonctionnent le plus souvent suivant une logique hypothético-déductive, très proche de celle des mathématiques.

¹³ Cette articulation forte entre modèle et hypothèse n'est toutefois pas le fait uniquement du monde académique : on peut la retrouver – ou regretter de ne pas la retrouver ! – dans le monde de la gestion ou de l'évaluation des politiques publiques (Gosselin, 2009)...

sphères scientifiques et décisionnelles (Lehmann, 1990). Les deux ne me semblent pas à opposer mais plutôt à comparer ou à associer (Peters, 1991). L'approche explicative s'avèrera probablement plus féconde sur le long terme que l'approche descriptive. Néanmoins, plusieurs remarques – de bon sens, d'avantage d'ingénieur que de chercheur – modulent cette assertion :

(i) appliquer une approche explicative à tous les niveaux hiérarchiques des systèmes aboutit à une approche très précise, réductionniste. Pourtant, une approche plus synthétique est aussi à considérer, probablement dans les mêmes théories/ hypothèses que l'approche réductionniste. Par exemple c'est aussi parce que Newton s'est arrêté sur les bons niveaux "hiérarchiques" que sont l'accélération et la masse qu'il a pu trouver ses lois. C'est à ces niveaux qu'émergent les propriétés intéressantes du point de vue de ce qu'il étudiait. Cela fait écho à la notion de propriété émergente de la systémique (e.g. Bertrand-Krajewski, 2007). Autre exemple : en dynamique des populations, les mathématiques nous ont appris que le niveau populationnel pouvait être résumé par quelques paramètres nous évitant – en première approche – de devoir suivre le destin précis de chaque individu pour étudier la dynamique de la population ;

(ii) ce qui apparaît explicatif à certains ou à certains niveaux de discours peut être descriptif à un autre niveau. Reprenons les lois de la gravitation de Newton. Elles peuvent apparaître explicatives, car elles ont une bonne capacité prédictive. Mais elles sont aussi descriptives car nous ne connaissons pas – que je sache – les causes profondes de ces relations.

Pourtant, une partie de l'écologie semble marquée par une préférence nette pour l'explicatif par rapport à l'empirique – ou descriptif. Lors d'un test d'une théorie par rapport aux données, certains écologues théoriciens privilégient en effet les théories (formulation extrême : Fagerström, 1987 ; plus intermédiaires : Levins, 1966, Loehle, 1983, Clark et Gelfand, 2006, Cox, 1990) :

"A theory may fail miserably to fit the data, yet be compelling in its explanatory power, elegance, completeness, logic, etc. We may decide that the data are wrong."
(Loehle, 1983),

ou encore :

"The validation of a model is not that it is "true" but that it generates good testable hypotheses relevant to important problems. A model may be discarded in favor of a more powerful one, but it usually is simply outgrown when the live issues are not any longer those for which it was designed." (Levins, 1966).

Ces points de vue très en faveur des théories et modèles théoriques – y compris à l'issue de la confrontation aux données – sont peut-être favorisés par un terrain épistémologique trouble, correspondant à :

- une très grande hétérogénéité des points de vue sur ce qu'est la validation d'un modèle en écologie (partie 2 de Rykiel, 1996) ou sur ce qu'est un modèle (Nichols, 2001) ;
- une grande influence des thèses Poppériennes en écologie, selon lesquelles la marque de fabrique d'une théorie scientifique est son caractère explicatif, mécaniste et hypothético-déductif. Pourtant, Popper lui-même reconnaît que ce point de vue, non testable, est "métaphysique" et reconnaît que la capacité prédictive des théories permet de distinguer plus clairement science et non-science (Peters, 1991, p. 106¹⁴).

Cette vision très théorique des modèles écologiques par certains écologues se voit remise en cause par d'une part les critiques de Peters (1991) à l'encontre du développement des tautologies en écologie et d'autre part les réflexions de Bouchaud (2008) sur l'état actuel de l'économie :

"Classical economics is built on very strong assumptions that quickly become axioms: the rationality of economic agents, the invisible hand and market efficiency, etc. An economist once told me, to my bewilderment: These concepts are so strong that they supersede any empirical observation... Physicists, on the other hand, have learned to be suspicious of axioms and models. If empirical observation is incompatible with the model, the model must be trashed or amended, even if it is conceptually beautiful or mathematically convenient. So many accepted ideas have been proven wrong in the history of physics that physicists have grown to be critical and queasy about their own models" (Bouchaud, 2008).

L'écologie, "canal théorique", ne court-elle pas le même risque que l'économie ? En écologie, ne devrait-on pas envier la physique pour son sens empirique – pouvant aller jusqu'à un certain scepticisme devant les développements théoriques selon Bouchaud (2008; cf. aussi Peters, 1991, pp. 108-109) – en même temps que pour ses prodigieux développements mathématiques et théoriques ? Ne vaut-il pas mieux adopter un point de vue davantage diversifié sur les théories, acceptant qu'il y ait des théories empiriques ou descriptives aux côtés des théories explicatives, comme Peters (1991, pp. 108-109) le propose ?

¹⁴ Sinon, "theories can not only 'be right for the wrong reason' (Dayton 1973) because successful prediction does not guarantee truth content, but if predictive success is not paramount, they can also be wrong for the right reason. This view debilitates the criterion of predictive power and allows other criteria to confound the selective process by which science winnows theory and grows" (Peters, 1991, p.108).

2.2.2. Différence de traitement des modèles dans la sphère académique et dans la sphère décisionnelle : le gouffre, le "meilleur en stock" et l'écart

La manière de traiter l'approximation que représente la modélisation varie suivant que le modèle est utilisé dans la sphère académique ou dans la sphère décisionnelle. Elle varie aussi au sein de la sphère académique.

La première vision est celle du gouffre ; elle est associée à la conception Poppérienne de la science. La distance qui sépare le modèle de la réalité est conçue comme un gouffre, qui est le fondement de la méthode des modèles (Legay, 1997, pp. 72 et 75¹⁵). La notion de gouffre provient du fait que tout modèle, théorie ou loi, n'étant vérifié que sur un ensemble de données fini, ayant leurs limites en termes de précisions, est amené à être contredit, affiné, amélioré avec davantage de données (Poincaré, 1905, p.171¹⁶, Caswell, 1976, Loehle, 1983). Le scientifique qui utilise cette méthode des modèles aura donc à construire de nouveaux modèles et hypothèses sur la base des expériences "cruciales" passées et à les confronter à des données de manière rigoureuse – par exemple à travers de nouvelles expériences "cruciales".

La seconde vision est celle de l'écart. Même si ce terme peut être utilisé pour décrire la même réalité que celle décrite par le gouffre Poppérien, cette seconde vision dédramatise la situation et cherche à mieux connaître et circonscrire cet écart. La citation suivante de Loehle (1983) est assez éclairante sur la différence entre le gouffre et l'écart:

"Note, that if our model fits the data 99% of the time, but behaves in a bizarre fashion 1% of the time, we would have to work to improve it if it were a theory, though it might be perfectly adequate as a calculation tool".

Dans certaines conditions, l'ingénieur ou le décideur sera tout à fait satisfait d'un tel modèle alors que le scientifique aura tendance à le rejeter et à en rechercher un nouveau. En effet, quid si un modèle approche la réalité de manière satisfaisante dans un domaine de validité bien défini? Quid si le domaine de validité n'est pas si clairement établi, mais le "coût" des 1% d'erreur est faible? Considérons à nouveau la loi de la gravitation selon Newton. Doit-elle

¹⁵ "Pour moi, le grain de sable se situe principalement dans ce qu'on appelle en abrégé "l'écart entre le modèle et la réalité", c'est-à-dire entre les prévisions autorisées par le modèle et l'observation directe de la réalité. Cet écart est le fondement même de la méthode des modèles. C'est lui qu'on attend et qui va faire progresser nos connaissances, puisqu'il va remettre en cause les mécanismes ou les hypothèses que nous avons imaginés. C'est parfois une surprise. Tout l'art du chercheur va consister à détecter cet écart (grâce à des expériences judicieuses), à en évaluer l'importance, et à le réduire en modifiant le modèle d'une manière ou d'une autre" (Legay, 1997, p.75).

¹⁶ "Si nous envisageons une loi particulière quelconque, nous pouvons être certains d'avance qu'elle ne peut être qu'approximative. Elle est, en effet, déduite de vérifications expérimentales et ces vérifications n'étaient et ne pouvaient être qu'approchées. On doit toujours s'attendre à ce que des mesures plus précises nous obligent à ajouter de nouveaux termes à nos formules; c'est ce qui est arrivé par exemple pour la loi de Mariotte."

être rejetée au motif qu'elle est dépassée dans certaines conditions par la théorie de la relativité généralisée ? Le scientifique répondra oui ; l'ingénieur, non dans la plupart des cas pratiques – en fonction notamment des "coûts" associés aux erreurs. Acceptons donc cette différence de perspective, qui se retrouve pour partie au sein du monde académique : constatant que les modèles auront du mal à tester les mécanismes sous-jacents à des observations dans des champs disciplinaires complexes comme les sciences de la vie et les sciences sociales, Costanza (2003) propose ainsi de les juger davantage sur leur qualité et leur utilité plutôt que leur caractère explicatif – i.e. pour tester des hypothèses sur les mécanismes sous-jacents (cf. aussi Bouleau, 1999, notamment p.315-316). Dans cette perspective, le lien entre "théories et hypothèses" et "modèle" (cf. Figure 1) pourra souvent se trouver atténué (cf. aussi 2.2.3). Selon cette vision de l'écart, avec tout modèle on devrait avoir un critère pour spécifier si le modèle est une approximation correcte de la réalité ou non, éventuellement dans certaines conditions – définissant ainsi son domaine de validité (Caswell, 1976, Innis *et al.*, 1977 in Loehle, 1983, Huber, 2002a).

Ces deux premières visions – le gouffre et l'écart – peuvent être implémentées assez facilement dans le cadre de la Figure 1, en laissant de côté au moins pour l'écart les modèles théoriques. Une troisième vision existe, celle du "meilleur modèle en stock" (cf. Figure 2). C'est une vision académique, qui peut être considérée comme une conséquence de la conception de Lakatos (1978) sur les programmes de recherche scientifiques. Elle est proche de la vision pluraliste du processus scientifique défendue il y a plus d'un siècle par Chamberlin (1965) ainsi que de la méthode des modèles pour l'ingénieur développée par Bouleau (pp.115-116, 172, 316-318 ; cf. aussi Goldstein, 1977) : elle vise à développer plusieurs modèles en même temps et à comparer leurs capacités respectives à expliquer les données : est alors gardé – au moins temporairement – le ou les meilleurs des modèles testés¹⁷, même s'ils sont d'assez mauvaises représentations de la réalité (par rapport à la vision "écart") ou même s'il existe des éléments pour les remettre en cause (vision "gouffre"). Pour les modèles d'ingénieur ou décisionnels, l'idée serait aussi non pas d'essayer de les critiquer par l'approche "falsificatrice" de Popper, mais plutôt de les contre-expertiser, par exemple en développant d'autres modèles, ayant d'autres bases (Bouleau, 1999). C'est une approche moins ontologique et plus relativiste que l'approche Poppérienne du gouffre. En écologie, la science s'est parfois développée suivant une seule hypothèse ou théorie, allant à l'encontre de la

¹⁷ "Ecologists, like many others, do not reject theories for the futile reason that they are wrong; theories are retained until better ones emerge... Picasso defined art as the "lie that helps us to see the truth". I believe it is the same with scientific theories: a lie that helps us will survive; a truth see will be forgotten" (Fagerström, 1987).

méthode des hypothèses multiples : ce fut le cas dans les années 1960 et 1970 avec l'hypothèse du rôle central de la compétition interspécifique en écologie des communautés (cf. par exemple le résumé fait par Gotelli et Graves, 1996, pp.8-9).

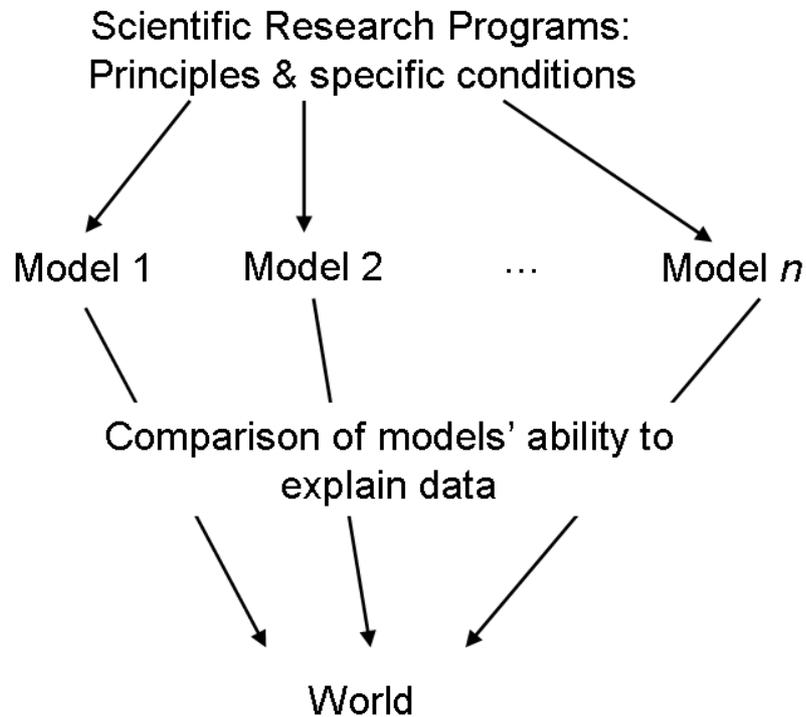


Figure 2. Représentation schématique de la conséquence des positions de Lakatos (1978) en termes de méthode des modèles : "le meilleur modèle en stock" parmi divers modèles alternatifs est retenu.

2.2.3. Une vision encore plus "relativiste" des modèles : la thèse de Duhem-Quine

La vision du "meilleur modèle en stock" de Lakatos met en avant une méthode des modèles éminemment pluraliste et relativiste. Une autre école insiste sur le caractère relatif de la modélisation et plus généralement du test des hypothèses et théories : l'école Quinienne, que l'on peut aussi associer à Duhem (cf. Sawyer *et al.*, 1997, Bouleau, 1999 pour plus de détails). Les difficultés à interpréter les résultats d'un modèle ou d'une expérience ont plusieurs sources :

- d'abord, Quine insiste sur le fait que différentes théories ou hypothèses peuvent être équivalentes du point de vue de l'observation sans l'être sur le plan des mécanismes explicatifs. Cela rejoint la réflexion souvent faite en écologie qu'on a du mal à inférer les mécanismes à partir des patrons (Rykiel, 1996, Costanza, 2003). Cette remarque semble davantage concerner l'approche Lakatosienne de la science et de la modélisation – difficulté à

comparer des modèles équivalents sur le plan des observations – que l'approche Poppérienne, fondée sur le caractère temporaire des théories et leur réfutabilité plutôt que sur leur validation ;

– ensuite – et cela concerne d'abord l'approche Poppérienne des sciences –, la confrontation d'un modèle avec la réalité n'a pas un lien si net avec la falsification ou non de la théorie utilisée pour construire le modèle. En effet, les modèles et la collecte des données ne sont en toute généralité pas construits à partir de cette seule théorie mais nécessitent des hypothèses auxiliaires pour prendre forme. De sorte que ce n'est pas uniquement la théorie principale qui est testée à travers un modèle ou une expérience cruciale, mais l'ensemble (théorie principale+hypothèses auxiliaires ; cf. Figure 3). Ainsi, il se peut que la théorie principale soit valable mais que ce soient les hypothèses auxiliaires qui soient contredites par les observations. Selon Lakatos (1978), lors de la confrontation aux données, ce sont souvent d'abord les hypothèses auxiliaires périphériques qui sont remises en cause, préservant ainsi le cœur de la théorie. Celle-ci se déploie alors avec d'autres hypothèses auxiliaires, jusqu'à ce qu'elle soit dépassée par une autre théorie en termes de capacité prédictive nouvelle. Ce constat *peut* expliquer la position de certains écologues – notamment théoriciens – qui accordent presque davantage de crédit aux théories qu'aux données (cf. 2.2.1). Néanmoins, la généralisation de cette conception pose problème car elle conduirait à remettre en cause le caractère prédictif des théories comme critère scientifique central (cf. fin de la partie 2.2.1). Les différents types d'hypothèses auxiliaires comprennent les conditions spécifiques d'application du modèle, l'approche méthodologique utilisée (expérimentale, observationnelle, type d'échantillonnage), le type de modèle mathématique utilisé et ses ingrédients, ainsi que les quantités analysées (cf. Figure 3 ; cf. aussi la partie 2.1.2). Sur ce dernier point, Bouleau (1999) donne l'exemple de la modélisation statistique du comportement d'un fleuve en cas d'inondation : pour ce faire, on peut choisir de modéliser la hauteur d'eau, le débit du fleuve, le logarithme de la hauteur d'eau ou celui du débit du fleuve. Ces différentes modélisations pourront donner des résultats différents, notamment pour les valeurs extrêmes, sans qu'on puisse trancher entre ces différents modèles, surtout en l'absence de données suffisamment nombreuses sur les valeurs extrêmes... Les descriptions des modèles faites par un écologue (Austin, 2007)¹⁸ et par un statisticien (McCullagh, 2002) entrent aussi en résonance avec le point de vue Quinien des modèles.

¹⁸ Selon Austin (2007), les composantes de modèles de biodiversité ou de répartition d'espèces – associés à des paradigmes...- incluent: (i) un modèle écologique incluant la théorie écologique utilisée ou supposée ; (ii) un modèle de données précisant la nature des données utilisées et les méthodes de collecte des données mais

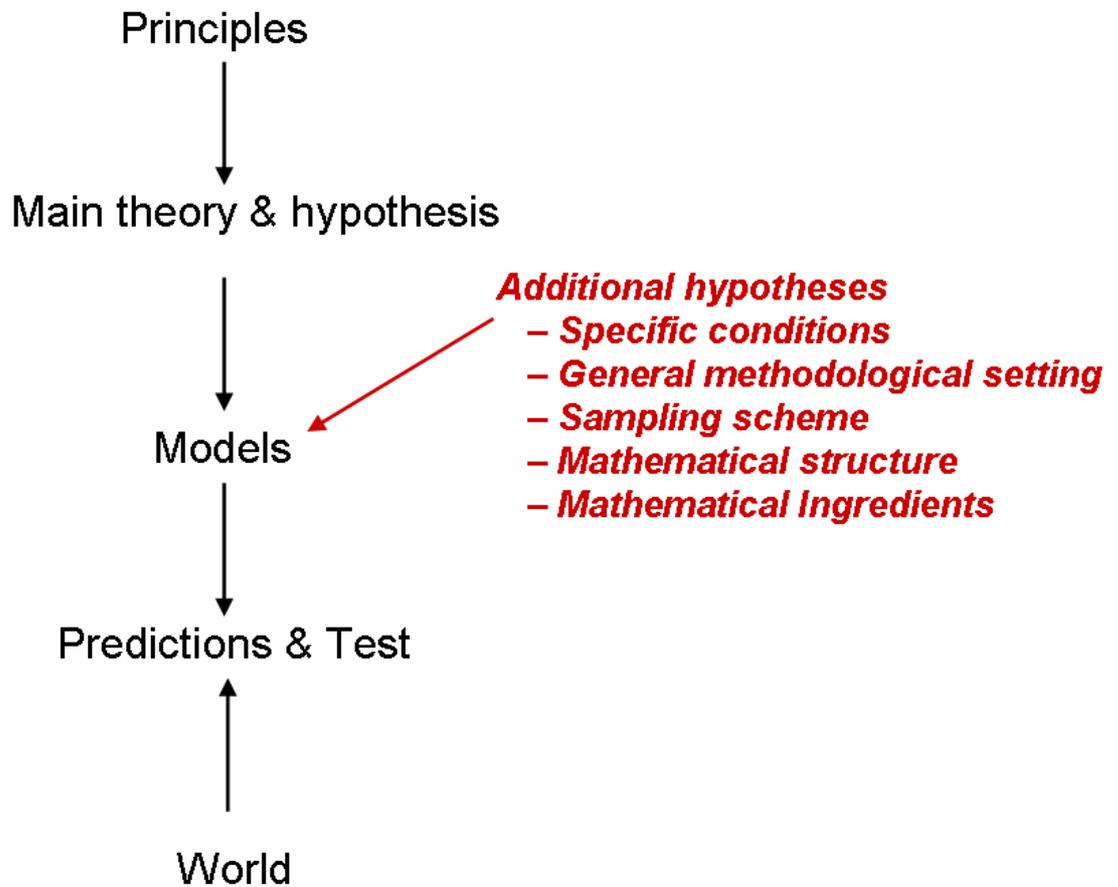


Figure 3. Représentation schématique de la vision de la modélisation selon la thèse de Duhem-Quine.

Dans la Figure 3, j'ai descendu les "conditions spécifiques" de validité d'une théorie du niveau des principes à celui des hypothèses auxiliaires. C'est une place somme toute assez logique pour cet élément important en écologie mais assez peu intégré aux discours sur les modèles en écologie : la notion de domaine de validité (anonyme, 1997). Reprenant les objectifs des lois physiques et chimiques, ainsi que l'avis de Popper (Stamos, 2007), certains écologues semblent rechercher des modèles, théories ou lois qui soient universels (comme le propose Loehle, 1983). Pour ma part, il me semble plus raisonnable de tester des modèles associés à des domaines de validité écologiques (Prisley et Mortimer, 2004), à cause de la relative complexité des systèmes écologiques et de l'importance de phénomènes de type aléatoires ou historiques associés... Un exemple d'une telle hypothèse, "contextualisée", est celle développée par Bersier et Meyer (1994) à la suite de Rotenberry (1985) : elle stipule que, sur un cycle de succession entier, c'est la structure du peuplement arboré davantage que

aussi souvent associés à un ensemble de problématiques restreint (ex: contrôle climatique de la distribution des espèces) et (iii) un modèle statistique comprenant les méthodes et théories statistiques appliquées.

sa composition en essences qui influence la composition en espèces d'oiseaux, et que c'est l'inverse si on se restreint aux peuplements arborés adultes. En corollaire, le manque d'attention porté au domaine de validité lorsqu'on teste un modèle en tant que généralité peut faire que ces tests ont lieu dans des conditions très restreintes qui posent problème pour l'applicabilité de leurs résultats dans la pratique (Romesburg, 1991, Peters, 1991, Bunnell et Huggard, 1999).

Bien entendu, cette présentation est assez caricaturale. Nombre d'hypothèses ou de théories écologiques intègrent en effet des conditions de contexte dans leur énoncé même, réduisant ainsi la part des hypothèses auxiliaires. Il n'en reste pas moins qu'une partie de l'écologie semble envier à la physique ses théories et lois générales. Enfin, notons que si certains économistes ou philosophes de l'économie par exemple ont réfléchi aux liens entre leur discipline et la thèse de Duhem-Quine, je ne connais aucun travail qui en fasse de même pour l'écologie. C'est probablement un point à explorer.

En résumé, contrairement à l'approche Poppérienne de la science, fondée sur la réfutabilité des théories scientifiques par les expériences, la vision Quinienne insiste au contraire sur l'indétermination des théories par les faits. D'après Bouleau (1999, p.329), "les théories quiniennes sont le cas général et il n'y a pas d'autre moyen de les réfuter que d'en construire d'autres à côté".

Je retiens de ces différentes perspectives sur le processus de modélisation non pas une méthode unique mais des éléments de méthode que je développerai pour certains dans les parties à venir. De l'approche "gouffre" ou Poppérienne, je retiens essentiellement le travail "hypothético-déductif" – mais aussi inductif – sur la théorie et les hypothèses de façon à les faire mûrir et à aborder la confrontation aux données plus sereinement. De l'approche Lakatosienne, je garde l'idée de travailler sur plusieurs modèles et hypothèses alternatifs en même temps. Ce sont des points de méthode développés dans Richard (2004) et Gosselin et Gosselin (2004), sur lesquels je ne reviendrai pas par la suite. De l'approche écart, je retiens l'intérêt d'estimer la magnitude des effets et son incertitude (cf. 4.1). Enfin, le point de vue de Quine-Duhem insiste selon moi avec raison sur l'importance des hypothèses auxiliaires lors de la confrontation du modèle aux données ; c'est ce qui m'a motivé pour travailler davantage sur la qualité "probabiliste" des modèles statistiques, en proposant à la fois des outils pouvant aider au diagnostic de cette qualité (4.2), et des constituants probabilistes a priori plus souples (4.3). Tout cela ne fait pas une méthode des modèles, j'en conviens, mais davantage des repères méthodologiques que je pense utiles pour l'utilisation des modèles.

3. Plaidoyer pour les modèles statistiques paramétriques Bayésiens

Cette partie est assez orthogonale à la partie précédente. J'y présente les intérêts que je vois à l'utilisation des modèles paramétriques Bayésiens en écologie. La jonction avec la partie précédente se fera en conclusion, une fois mon argumentaire posé. Je reconnais néanmoins que la rencontre entre méthodes Bayésiennes et épistémologie est beaucoup plus riche que ce que je serai amené à traiter ici – faute de place et de temps ¹⁹.

Dans ce qui suit, je m'intéresserai presque exclusivement aux modèles permettant d'estimer les paramètres d'une population statistique ou les paramètres de modèles statistiques à partir de données réelles. Nous entendrons par modèle statistique un modèle de nature probabiliste où les observations et éventuellement certaines variables sont supposées être des réalisations de variables aléatoires (Cox, 1990, Yoccoz, 1999). Ce sont la plupart du temps des modèles qui comportent une dimension descriptive (cf. Figure 1) – par l'estimation de paramètres du modèle à partir des données – mais aussi assez souvent une dimension prédictive – par le test d'hypothèses définies a priori.

Dans un premier temps, j'expliquerai pourquoi j'ai opté pour les modèles statistiques plutôt que des méthodes statistiques sans modèle pour aborder les diverses questions et politiques relatives à la biodiversité. Ensuite, j'explicitai les atouts – et les quelques faiblesses – d'une approche par modèle statistique paramétrique dans le cadre Bayésien. Faute de place, je ne traiterai pas de questions connexes pourtant importantes, telles que l'importance des approches statistiques, stochastiques ou probabilistes en écologie (cf. Yoccoz, 1994, p. 58, Gosselin, 1997, notamment partie I).

Pourquoi s'orienter vers des modèles statistiques, et lesquels ? Dans cette partie, je fais le point sur les raisons qui m'ont poussé à travailler sur les modèles statistiques paramétriques Bayésiens. En trois temps.

Dans un premier temps, pourquoi des *modèles statistiques*, tout simplement ? Il existe en effet des méthodes statistiques qui sont fondées uniquement sur les propriétés du plan d'échantillonnage et qui ne font aucune hypothèse probabiliste quant à la population étudiée (e.g. de Gruijter et ter Braak, 1990, Gregoire, 1998). Ces méthodes permettent d'estimer les paramètres d'une population statistique, paramètres qui sont de première

¹⁹ aboutissant à la création de ce que certains appellent une épistémologie Bayésienne (cf. par exemple Dorling, 1979, Hellman, 1997, Crupi *et al.*, 2007, Ullman *et al.*, Fitelson, 2010, Henderson *et al.*, 2010)

importance pour la sphère de la décision. Ce sont souvent ces méthodes statistiques qui sont utilisées par les inventaires statistiques pour analyser leurs données. Les méthodes statistiques basées sur des modèles peuvent néanmoins avoir un avantage dans ce contexte, par exemple s'il y a des variations de détectabilité (Buckland *et al.*, 2000) dans la population échantillonnée. Je reconnais que les deux approches ont leurs avantages et leurs inconvénients. J'ai néanmoins choisi l'approche fondée sur les modèles car elle me semble plus souple, par exemple pour pouvoir incorporer des fonctions de lien non-linéaires, en même temps que des dépendances spatiales... Le coût à payer, bien entendu, est de porter une attention particulière à la validation des hypothèses faites lors de la construction des modèles et à leur signification écologique.

Dans un deuxième temps, pourquoi des modèles statistiques *paramétriques* ? Il existe en effet de nombreuses méthodes non-paramétriques, qui permettent de faire des tests statistiques sans passer par l'estimation de paramètres de modèles statistiques. Mais les arguments suivants plaident en faveur des modèles paramétriques :

(i) contrairement à ce qu'on croit, les modèles non-paramétriques font aussi des hypothèses (par exemple sur l'indépendance des observations ; Richard, 2004), qui peuvent rendre les méthodes non-paramétriques moins robustes que leurs équivalents paramétriques ;

(ii) la modélisation dans des dispositifs complexes (avec par exemple hétérogénéité des variations, dépendances entre les données,...) est beaucoup plus difficile en modèle non paramétrique que paramétrique (cf. aussi Laara, 2009) ;

(iii) les modèles paramétriques permettent d'estimer la magnitude des effets et le niveau d'incertitude associé, sur des échelles interprétables, là où les méthodes non-paramétriques ne donneront en général que des probabilités de significativité du test (en désaccord avec certaines recommandations de biométriciens ; cf. aussi partie 4.1). Ainsi, si le but d'un modèle est de rentrer dans une comparaison d'analyses de données dans différentes situations (Yoccoz, 1999), on aura tout intérêt à utiliser une forme de modèle "canonique", et à travailler sur les paramètres estimés, plutôt qu'à travailler sur des probabilités de test, des R^2 ou autres quantités dépendantes du bruit résiduel ou du nombre de relevés ;

(iv) les modèles paramétriques permettent de comparer des modèles incluant des hypothèses écologiques adverses (Richard, 2004).

Bien entendu, cela n'interdit pas d'incorporer des éléments de méthodes non-paramétriques ou semi-paramétriques dans des modèles paramétriques, par exemple pour estimer les effets de paramètres de "nuisance" (covariables, effets aléatoires...).

Enfin, pourquoi donner priorité aux modèles statistiques paramétriques *Bayésiens* ? Un détour par l'histoire des sciences (cf. http://en.wikipedia.org/wiki/Bayesian_probability) nous rappelle que les statistiques ont été Bayésiennes (Laplace, Bayes) avant d'être fréquentistes, et que les deux branches des statistiques se sont développées de manière assez indépendante, sans qu'on sache clairement aujourd'hui hiérarchiser les deux en toute généralité (Bayarri et Berger, 2004, Little, 2006). Voici néanmoins une proposition de liste des arguments en faveur du Bayésien (B) et en sa défaveur (F comme fréquentiste) :

(B1) & (F1) le premier élément est à la fois un avantage et un inconvénient de l'approche Bayésienne. Dans le cadre fréquentiste classique, les distributions de probabilité apparaissent le plus souvent au niveau des variables à expliquer ou de leurs résidus. La caractéristique du cadre Bayésien est de mettre des **distributions de probabilité aussi sur les paramètres du modèle** – par exemple, la moyenne et la variance d'une loi normale. C'est pour cela qu'on qualifie les modèles Bayésiens de complètement probabilistes. La distribution de probabilité des paramètres change quand on observe des données : avant de les observer, nous avons affaire à la distribution *a priori* des paramètres ; après les avoir observées, elle devient la distribution *a posteriori* des paramètres. On obtient cette mise à jour des distributions de probabilité des paramètres par le théorème de Bayes, qui fonctionne comme un véritable "processeur d'information" (Rivot, 2003). Les opposants à l'approche Bayésienne crient au scandale que constitue l'injection d'autres chose que les données – la distribution *a priori* des paramètres – dans l'analyse, arguant du caractère subjectif de cette manière de faire (Dennis, 2004, Lele *et al.*, 2007). Même les distributions *a priori* peu informatives peuvent poser problème (Van Dongen, 2006, Lele et Dennis, 2009) : en effet, une distribution plate ou presque plate – i.e. équiprobable – pour un paramètre donné ne l'est plus pour des transformations non-linéaires de ce paramètre. Les partisans du Bayésien y voient au contraire soit une procédure assez peu influencée par les distributions *a priori* (statistiques Bayésiennes "objectives", avec des distributions *a priori* adaptées), soit une procédure prenant en compte une vraie information *a priori* (statistiques Bayésiennes "subjectives"), que ce soit par exemple du savoir d'expert ou des connaissances scientifiques préalables. Les partisans du Bayésien trouvent donc en cette procédure une manière cohérente de mettre à jour nos connaissances explicites au fur et à mesure que l'information arrive, au cours du temps ou à partir de différentes sources (Huber, 1997, Cressie *et al.*, 2009, http://en.wikipedia.org/wiki/Bayesian_probability). Certains Bayésiens proposent même de dépasser cette opposition entre subjectif et objectif, pour reconnaître que dans tout modèle ou

méthode statistique il y a du subjectif ou des connaissances préalables – y compris dans le fréquentiste (Lecoutre *et al.*, 2001) –, et que la question est plutôt de connaître la validité des choix effectués et la sensibilité des résultats par rapport à ces choix (Cressie *et al.*, 2009) ;

(B2) **l'intervalle de crédibilité**, notion Bayésienne correspondant à l'intervalle de confiance en fréquentiste, a une interprétation plus directe et moins alambiquée que ce dernier : un intervalle de crédibilité de 95% d'un paramètre est bien un intervalle qui contient 95% de la distribution a posteriori du paramètre – donc 95% des valeurs du paramètre sous la distribution *a posteriori*... (Lecoutre, 2005, Cressie *et al.*, 2009) ;

(B3) et (F2) les **algorithmes d'estimation des paramètres des modèles Bayésiens** – et notamment les chaînes de Markov de Monte Carlo (MCMC) – sont clairement responsables pour partie du succès actuel des modèles Bayésiens. Comme je l'ai déjà dit, ils permettent de prendre en compte des modèles beaucoup plus complexes que ce que nous pouvions faire jusqu'ici en fréquentiste (Clark, 2005, Cressie *et al.*, 2009). Les outils Bayésiens permettent de décomposer ces modèles complexes en briques de modèles simples en utilisant des conditionnements probabilistes. Ceci permet par exemple de prendre en compte les multiples sources d'incertitude, y compris sur les mesures des variables explicatives – ce qu'à ma connaissance peu d'outils fréquentistes permettent, et dans peu de conditions. Cependant, le développement d'algorithmes qui convergent correctement et facilement n'est pas encore complètement stabilisé, de sorte que leur mise en place demande souvent du temps et de l'expérience²⁰ ;

(B4) un avantage certain des méthodes Bayésiennes sur la plupart des méthodes fréquentistes est que l'estimation des paramètres n'y **repose pas sur des résultats asymptotiques sur les distributions des estimateurs**. Souvent, pour des modèles assez simples, on note assez peu de différences entre fréquentiste et Bayésien. Néanmoins, un cas typique où le Bayésien permet de sortir de l'ornière dans laquelle les résultats asymptotiques peuvent mettre les modèles fréquentistes concerne les régressions logistiques – en modèle linéaire généralisé –, avec par exemple une variable catégorielle dont un des niveaux correspond à des données uniformément nulles (absence de l'espèce). Dans ce cas, à cause de l'hypothèse de normalité – donc de symétrie – de la distribution de l'estimateur issue des résultats asymptotiques, les erreurs types des estimateurs fréquentistes peuvent très facilement approcher des valeurs énormes – faisant penser qu'il n'y a pas assez d'information pour

²⁰ "There is much judgment involved in constructing an MCMC algorithm that burns in quickly and yields stable samples from the posterior distribution" (Cressie *et al.*, 2009).

conclure sur la valeur de l'estimateur – alors que les distributions postérieures Bayésiennes sont plus informatives dans une direction – mais asymétriques ;

(B5) les méthodes Bayésiennes permettent de mettre à jour les distributions de probabilité de l'ensemble des paramètres en **prenant en compte simultanément tous les niveaux d'incertitude** (Cressie *et al.*, 2009; Uriarte et Yackulic, 2009; Valpine, 2009) tandis que les estimations de certains paramètres fréquentistes ne peuvent se faire pour une seule valeur "optimale" d'un autre paramètre (par exemple une variance), sans prendre en compte sa variabilité ;

(F3) bien que de nombreux auteurs annoncent la supériorité du Bayésien pour la **critique de modèles** ("p-valeurs", test d'ajustement) (e.g. Uriarte et Yackulic, 2009), l'outil classique de critique de modèle Bayésien – la p-valeur postérieure prédictive – est plus mauvais que l'outil utilisant le maximum de vraisemblance, fréquentiste (e.g. Valpine, 2009 ; cf. partie 4.2). Il existe bien des outils Bayésiens de substitution, mais ils sont compliqués à utiliser en toute généralité. De manière plus générale, la culture de l'estimation robuste semble manquer aux Bayésiens ²¹ ;

(B6) & (F4) les méthodes Bayésiennes respectent le **principe de Vraisemblance**, qui stipule que toute l'information provenant d'un jeu de données est contenue dans la fonction de vraisemblance. Le Bayésien partage cette propriété avec les approches non-Bayésiennes fondées sur la vraisemblance (cf. Royall, 2004), mais pas avec beaucoup de méthodes fréquentistes (Little, 2006). Si cela est plutôt un avantage du Bayésien, ce principe est quand même sujet à débat (anonyme, 2007). En lien avec cette polémique, une autre différence entre le Bayésien et le fréquentiste est la possibilité de regarder les analyses de données pendant qu'on collecte les données en Bayésien et pas en fréquentiste (anonyme, 2007).

(B7) d'après différentes études, l'utilisation de méthodes Bayésiennes pour **mettre à jour les connaissances** est plus efficace pour apprendre que l'utilisation de méthodes non-formalisées (http://en.wikipedia.org/wiki/Bayesian_probability) ;

(B8) les méthodes Bayésiennes offrent aussi l'avantage de pouvoir **comparer des modèles** de manière théoriquement plus rigoureuse (Little, 2006) – au moins sur le principe, car les problèmes numériques peuvent être importants – que les méthodes fréquentistes, et de

²¹ "The term "robust" was coined by Bayesians (Box and Andersen, 1955). It is puzzling that Bayesian statistics never managed to assimilate the modern robustness concept, but remained stuck with antiquated parametric supermodels. *Ad hoc* supermodels and priors chosen by intuition (personal beliefs) or convenience (conjugate priors) do not guarantee robustness, for that one needs some theory. Compare already Hampel (1973)" (Huber, 1997).

permettre des estimations multi-modèles de manière assez naturelle, intégrant l'incertitude sur le meilleur modèle dans l'estimation ;

(B9) enfin, les méthodes Bayésiennes semblent plus naturellement utilisables pour des **analyses de risques** en science de la décision (Wade, 2000, Rivot, 2003).

Au bilan, les méthodes Bayésiennes semblent ainsi meilleures que leurs équivalents fréquentistes pour l'estimation et l'inférence sous un modèle donné (Little, 2006). La situation semble être l'inverse pour les techniques de critiques de modèle. C'est pourquoi Little (2006) a proposé de mixer les deux approches, programme que j'ai fait mien pour le développement de nouvelles techniques de critique de modèles (cf. partie 4.2).

Quelques exemples d'utilisation de modèles statistiques Bayésiens en écologie

– le Bayésien peut être utilisé pour combiner des sources d'informations de différents types (par exemple, deux manières de modéliser une probabilité de présence d'espèce : Bayliss *et al.*, 2005 ; ou couplage entre modélisation statistique des processus démographiques et des données d'abondance : Fonnesbeck et Conroy, 2004; Chaloupka et Balazs, 2007) ;

– pour l'analyse des essais en plein champ en agriculture – un domaine dominé jusqu'ici par les statistiques fréquentistes –, le Bayésien offre l'avantage de simplifier l'interprétation des données (Besag et Higdon, 1999) ;

– des modèles bayésiens sont utilisés pour analyser des données de répartition de l'abondance entre espèces (proches des analyses de diversité et d'équitabilité ; Etienne et Olf, 2005, Golicher *et al.*, 2006) ;

– pour modéliser plusieurs espèces en même temps en partageant certains effets entre espèces, des modèles Bayésiens peuvent être utilisés au même titre que les modèles fréquentistes (Richard, 2004, Warton et Hudson, 2004, Gelfand *et al.*, 2005, Ferrier et Guisan, 2006, Kéry et Royle, 2008) ;

– des modèles de type "structural equation modeling", faisant apparaître différents chemins explicatifs, sont utilisés pour étudier l'effet de la biodiversité vs d'autres forces sur la productivité des écosystèmes (Grace *et al.*, 2007, Grace et Bollen, 2008, Vacher *et al.*, 2008). Ces modèles sont pour l'instant essentiellement développés dans le cadre fréquentiste mais pourraient aussi profiter du cadre Bayésien (Lee, 2007) ;

– des réseaux Bayésiens sont utilisés pour modéliser des règles de décision et savoir d'expert (Marcot *et al.*, 2006).

D'un certain point de vue, l'approche Bayésienne hiérarchique telle qu'elle est en train de se déployer semble contredire la méthode de recherche de Levins (1966) sur deux points – pourvu qu'on accepte de mettre dans le même paquet modèles statistiques et modèles dynamiques, ceux auxquels pensaient Levins :

– Levins était gêné dans l'utilisation des modèles complexes par le manque de solution sous forme explicite de séries d'équations aux dérivées partielles nombreuses (cf. aussi Odenbaugh, 2006). Clairement, les outils Bayésiens hiérarchiques ont pour partie résolu ce problème en remplaçant les solutions explicites par des approximations basées sur des outils comme le MCMC ;

– en second lieu, comme le souligne Odenbaugh (2006), les statistiques Bayésiennes par leur principe de conditionnalité permettent d'appréhender de manière compréhensible des modèles statistiques aux paramètres multiples : elles cassent en quelque sorte le lien implicite chez Levins (1966) entre modèle compliqué et modèle ininterprétable.

Par contre, ce qui reste de l'héritage de Levins, et qu'a redéveloppé Bouleau (1999) par exemple, c'est la nécessité de ne pas être prisonnier du modèle développé, et d'adopter une approche mixte écologique-modélisation, fondée sur plusieurs modèles ou de tester assez complètement les hypothèses faites dans les modèles (cf. partie 4).

Quel lien y a-t-il finalement entre la modélisation statistique Bayésienne et les différentes visions de la science et de la modélisation exposées dans la partie 2 ? Ce qui semble clair c'est que les modèles Bayésiens permettent de faire vivre la confrontation entre modèles promue par Lakatos (Hilborn et Mangel, 1997). Par différents types d'outils, ils devraient aussi permettre de traiter le modèle comme écart dont on souhaite maîtriser l'importance par rapport à la réalité (cf. suite).

4. Vers des outils améliorés pour le détective écologique et certains de ses collègues

4.1. Pour populariser l'étude de la magnitude des effets

La partie statistique du test d'hypothèses en écologie a historiquement reposé sur une vision Popérienne du développement de la science, par le test d'hypothèses nulles ponctuelles : en utilisant cette méthode, le but est en effet de montrer l'existence de relations entre variables ou de patrons en essayant de montrer qu'une hypothèse nulle ponctuelle – impliquant la nullité d'un certain paramètre, $\theta = 0$ – est très peu probable (Lecoutre *et al.*, 2001) – donc en cherchant à la falsifier. Sous la condition que l'hypothèse nulle est vraie, on calcule la probabilité qu'une fonction des données observées et/ou du paramètre "optimum" estimé à partir des données – souvent correspondant au maximum de vraisemblance – soit inférieure à la même fonction de données et des paramètres simulés sous l'hypothèse nulle ponctuelle. La probabilité est appelée p-valeur ou valeur p. Si la p-valeur est faible, les données sont surprenantes si l'hypothèse nulle est vraie et :

– l'hypothèse nulle sera donc tenue pour peu probable, si on se trouve dans le cadre néoFisherien (Hurlbert et Lombardi, 2009) ;

– l'hypothèse nulle sera rejetée – et le test est déclaré significatif – si elle est en-dessous d'un seuil α fixé a priori, si on se trouve dans le cadre paléoFisherien (Hurlbert et Lombardi, 2009).

Cette approche a néanmoins des défauts désormais bien reconnus :

(i) confusion fréquente entre absence de significativité et acceptation de l'hypothèse nulle, et entre le complémentaire de la p-valeur et la probabilité que l'hypothèse alternative soit vraie (Lecoutre *et al.*, 2001) ;

(ii) confusion fréquente entre le niveau de significativité et l'intérêt scientifique du résultat (Lecoutre *et al.*, 2001). Or un effet peut être très significatif sans être fort pour autant ;

(iii) interprétation très variable d'un scientifique à l'autre du même résultat d'un test de ce type (Lecoutre *et al.*, 2001). A titre d'exemple, en cas de test non significatif, une bonne partie des scientifiques remettent en question la qualité de l'expérience²² alors que d'autres tirent des conclusions statistiques (Lecoutre *et al.*, 2001) ;

²² comportement qui n'était pas observé si le test était significatif – une manifestation d'un biais psychologique très probablement (Lecoutre *et al.*, 2001).

(iv) pour une même valeur du paramètre sous-jacent θ non-nulle, la p-valeur du test diminuera vers la limite 0 quand le nombre d'observations tend vers l'infini et variera dans le même sens que le niveau de variance dans le jeu de données. On peut ainsi aboutir au problème du "N obèse" (Hurlbert et Lombardi, 2009), où, à partir d'un certain nombre d'observations, l'hypothèse nulle aura une p-valeur très faible.

Une autre école en écologie – qu'on peut baptiser "vraisemblantiste" car elle est fondée essentiellement sur le calcul de la vraisemblance des modèles – promeut plutôt l'utilisation de méthodes de comparaisons de modèles, dans une approche générale très proche du point de vue de Lakatos (cf. partie 2.2.2). L'objectif ici est de comparer l'accord de différents modèles explicatifs avec les données et de ne garder que le ou les meilleurs. Des outils simples basés sur des critères d'informations (d'Akaike, de Bayes...) sont utilisés à cet effet. On oppose souvent ces outils aux tests d'hypothèses nulles. Pourtant, il faut rappeler qu'ils partagent pour partie les mêmes limites que les tests d'hypothèses nulles ponctuelles : ainsi, pour des modèles très simples de régression sur variables explicatives orthogonales, les résultats de ces méthodes sont, pour certains critères d'information, asymptotiquement équivalents à des tests F des effets associés à des valeurs d' α précises (McQuarrie et Tsai, 1999) ; de sorte que là encore, au moins pour le cas de comparaisons de modèles emboîtés, le modèle le plus compliqué sera retenu pourvu qu'on ait suffisamment de données observées. Autrement dit, les méthodes "vraisemblantistes" souffrent donc aussi du problème du N obèse et il est possible que les limites des tests d'hypothèses nulles ponctuelles – sur lesquelles les vraisemblantistes insistent beaucoup (<http://www.indiana.edu/~stigtsts/>) – aboutissent à scier la branche sur laquelle les vraisemblantistes sont perchés.

Comment traiter ce problème d'obésité, docteur ? Une première méthode souvent employée en écologie consiste à raisonner sur des mesures de variance ou déviance relative expliquée par l'effet étudié. Le résultat ne dépend pas de N – le problème du N obèse est donc traité –, mais souffre par contre d'une dépendance de la variance du jeu de données : un même θ induira un R^2 différent suivant le niveau de bruit dans les données par exemple.

Une seconde parade serait de coupler des analyses de puissance à des analyses de données. Néanmoins, le consensus actuel semble être d'utiliser les analyses de puissance avant de récolter les données et non au moment de leur analyse (Lecoutre *et al.*, 2001, Hoenig et Heisey, 2001).

Je pense que la meilleure méthode à notre disposition pour aborder le test d'une hypothèse écologique consiste à **remplacer les hypothèses ponctuelles par des hypothèses**

"intervalles" (Camp *et al.*, 2008, Dixon et Pechmann, 2005)²³. L'hypothèse nulle ponctuelle $\theta = 0$, comparée à son alternative $\theta \neq 0$, est remplacée par une hypothèse intervalle $\theta \in [-\varepsilon_1; \varepsilon_2]$, comparée aux deux alternatives : $\theta \in]\varepsilon_2; \infty[$, et $\theta \in]-\infty; -\varepsilon_1[$. Très souvent, nous aurons $\varepsilon_1 = \varepsilon_2 > 0$, sans que ce soit une nécessité. Ce sont les probabilités *a posteriori* de ces hypothèses²⁴ qui sont comparées ; les résultats sont en général assez proches – et le sont asymptotiquement – que l'on utilise des méthodes Bayésiennes ou des méthodes "vraisemblantistes" pour estimer le modèle et les probabilités des différentes hypothèses (Edwards, 1996, Camp *et al.*, 2008). Le premier avantage, intuitif, de cette approche est de comparer des choses comparables : un intervalle à des intervalles et non une valeur ponctuelle à un intervalle. En conséquence, quand on accumulera suffisamment de données, on ne sera pas forcé de conclure au rejet de l'hypothèse nulle – syndrome du N obèse dans le cas de l'hypothèse nulle ponctuelle – mais on conclura soit que l'effet est négligeable – l'hypothèse nulle intervalle est acceptée –, soit que l'effet est non négligeable et négatif, soit qu'il est non négligeable et positif. Le second avantage est d'intégrer dans l'hypothèse nulle d'absence d'effet la possibilité de petits écarts à la valeur nulle, par exemple à cause d'erreurs systématiques de mesures, possibles dans la plupart des cas. Le troisième avantage des hypothèses intervalles est qu'elles rapprochent les notions d'hypothèse statistique et d'hypothèse biologique (voire décisionnelle) (Lukacs *et al.*, 2007 vs Hurlbert et Lombardi, 2009) : en effet, les nombres $\varepsilon_1, \varepsilon_2$ seront très souvent choisis de sorte que si $\theta \in [-\varepsilon_1; \varepsilon_2]$, l'effet est déclaré biologiquement négligeable (Dixon et Pechmann, 2005, Camp *et al.*, 2008). Par exemple, Dixon et Pechmann (2005) ont choisi $\varepsilon_1 = \varepsilon_2 = 0.0346$ pour des modèles exponentiels analysant la tendance temporelle de l'abondance de diverses espèces. Ces seuils correspondent à un doublement ou une division par deux de l'abondance en 20 ans, ce qui correspond peu ou prou à certains des critères utilisés par l'UICN pour classer les espèces dans les catégories de menace d'extinction. Au lieu d'un seul couple de seuils de négligeabilité $\varepsilon_1, \varepsilon_2$, on peut en outre en définir plusieurs (Walshe *et al.*, 2007; Barbier *et al.*, 2009). Quatrième avantage de cette méthode : on prête attention à la magnitude et au signe des effets, là où une approche basée uniquement sur du test d'hypothèse nulle ponctuelle ou du R^2 peut facilement l'ignorer (e.g. Yates, 1951) – comme c'est assez souvent le cas en écologie. Or la

²³ et non pas par des intervalles de confiance (Di Stefano, 2004, Walshe *et al.*, 2007), qui partagent avec le test "classique" d'hypothèses nulles ponctuelles l'emploi d'un seuil arbitraire de significativité (Hurlbert et Lombardi, 2009).

²⁴ plus exactement que le paramètre appartienne aux intervalles associés à ces hypothèses.

science est intéressée au premier chef par la magnitude des effets (Simberloff, 1990; Cohen, 1990).

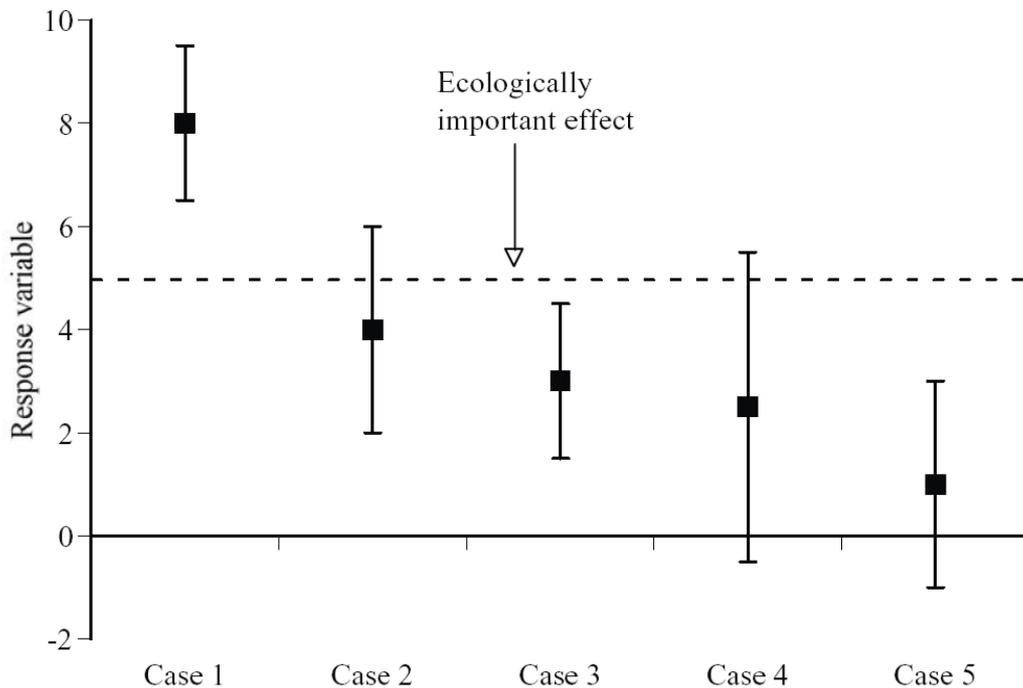


Figure 4. Représentation simplifiée des différents cas de figure de croisement possible entre hypothèses intervalles et hypothèse nulle ponctuelle (figure issue de Di Stefano, 2004²⁵). La figure n'aborde pas le cas de l'intervalle non-négligeable négatif et adopte une approche par intervalle de confiance là où nous proposons plutôt une approche "probabilité postérieure d'être dans les intervalles". Dans le cas 1, le paramètre est jugé comme significativement différent de 0 et non-négligeable et positif. Dans le cas 2, le paramètre est indéterminé du point de vue de sa négligeabilité alors qu'il est statistiquement significativement différent de 0. Dans le cas 4, le paramètre reste indéterminé du point de vue de sa négligeabilité mais n'est plus significativement différent de 0. Enfin, dans les cas 3 et 5, le paramètre sera jugé négligeable et, respectivement, significativement différent de 0 et non significativement différent de 0.

Dans le cas précédent, trois alternatives étaient distinguées dans l'analyse : l'effet est négligeable ; l'effet est non-négligeable négatif ; l'effet est non-négligeable positif. Ceci étant, dans certains cas – par exemple les catégories de menace d'extinction de l'UICN, on cherchera juste à distinguer les cas où l'effet est négatif et non-négligeable (i.e. le taxon est menacé) du reste du monde (i.e. le taxon n'est pas menacé, correspondant à un effet négligeable ou non-négligeable et positif) : on rejoint alors les approches plus classiques basées sur des tests t unilatéraux (Hurlbert et Lombardi, 2009), mais avec des intervalles $]-\infty; -\varepsilon_1]$ et $]-\varepsilon_1; \infty[$, et non forcément $]-\infty; 0]$ et $]0; \infty[$. Mais nous sommes alors bien plus proche sur le fond d'un test impliquant deux intervalles que d'un test d'hypothèse nulle ponctuelle (Hurlbert et Lombardi, 2009). Une alternative appauvrie – car ne donnant pas d'information sur

²⁵ "Reprinted from Forest Ecology and Management, vol. 187, J. Di Stefano, A confidence interval approach to data analysis, pp.173-183, Copyright (2004), with permission from Elsevier."

l'estimateur du paramètre – serait enfin de comparer des modèles basés sur des hypothèses nulles ponctuelles avec chacun des valeurs différentes des paramètres (cf. Hurlbert et Lombardi, 2009; Anderson *et al.*, 2000; exemple donné par Lewin-Koh *et al.*, 2004).

Dans Barbier *et al.* (2009), nous avons adopté une méthode d'interprétation des données utilisant des hypothèses intervalles. Je reprends le cas de l'analyse de la richesse spécifique des groupes écologiques de la flore du sous-bois relevée sur 400 m². Le modèle est ici multiplicatif : on modélise le facteur par lequel on multiplie la moyenne prédite quand on augmente la variable écologique d'une quantité jugée importante. Nous avons ici considéré des ajouts de 5 m²/ha en surface terrière (soit totale, soit par catégories d'essence) : ce chiffre est proche de l'écart-type de surface terrière dans notre jeu de données et correspond à une variation à la fois réaliste et significative – sur le plan de la gestion – du paramètre écologique. L'idée est de calculer la probabilité que le paramètre multiplicatif soit dans chacun des intervalles suivants :

- (i) le paramètre est "biologiquement" négligeable;
- (ii) le paramètre est non-négligeable et inférieur à 1 ;
- (iii) le paramètre est non-négligeable et supérieur à 1.

Si le paramètre ne se retrouve pas avec une probabilité suffisante dans un de ces trois intervalles, on ne peut pas conclure quant à sa négligeabilité. Nous avons considéré comme intervalle de négligeabilité biologique $[\exp(-0.2); \exp(0.2)]$ correspondant à une évolution de richesse spécifique comprise entre -19% et +22% pour la variation de surface terrière ci-dessus. Les résultats (Tableau 1) indiquent :

- des bryophytes forestières subissant un impact négligeable de la surface terrière du chêne et des essences de demi-ombre ;
- des herbacées dont on ne peut dire si elles sont impactées de façon non-négligeable ou non par la surface terrière du chêne, alors que les ligneuses "subissent un impact" qui est négligeable ;
- des ligneuses forestières subissant un impact négligeable de la surface terrière des essences de demi-ombre, des herbacées forestières et ligneuses péri-forestières dont on ne sait pas si l'impact est négligeable ou non et enfin des herbacées péri-forestières et non-forestières pour lesquelles l'impact est non-négligeable négatif avec une probabilité de 95%.

Pour tous les cas où nous ne pouvons pas conclure, il faut soit nous résigner à rester indécis, soit acquérir plus de données pour resserrer l'intervalle de crédibilité et arriver à une conclusion.

A signaler un cas intéressant dans le Tableau 1 : les ligneuses péri-forestières qui ont une réponse à la fois statistiquement significative et négligeable à la surface terrière du chêne.

Variable écologique	Bryophytes		Herbacées		Ligneuses et semi- ligneuses	
			Péri-	Non-	Péri-	
	Forestières	Forestières	Forestières	forestières	Forestières	Forestières
	0.99 ⁰	1.06	0.85	0.81	1.06 ⁰	0.89 ^{*,0}
G.Qu	[0.93;1.06]	[0.91;1.22]	[0.62;1.14]	[0.58;1.08]	[0.93;1.20]	[0.81;0.98]
	1.03 ⁰	0.83 ^{**}	0.62 ^{***,-}	0.40 ^{***,-}	0.93 ⁰	0.80 ^{***}
G.Un	[0.99;1.07]	[0.75;0.92]	[0.48;0.79]	[0.29;0.53]	[0.85;1.01]	[0.75;0.85]

Tableau 1. Analyse du facteur multiplicatif sur la richesse spécifique des différents groupes écologiques lors d'un ajout de 5m²/ha de surface terrière du chêne et du châtaigner (G. Qu) et des essences de demi-ombre (essentiellement charme et tilleul ; G.Un). Les chiffres correspondent à la moyenne du facteur (1^{ère} ligne) et à son intervalle de crédibilité (2^{nde} ligne). On indique par des sigles en exposant à la fois une éventuelle différence significative bilatérale de l'estimateur par rapport à la valeur 1 (* = p<0.001, ** = p<0.01, * = p<0.05) et son caractère négligeable avec une probabilité de 95% ("0") ou son caractère non-négligeable et négatif avec une probabilité de 95% (-). Il n'y a pas de cas non-négligeable et positif.**

Les considérations précédentes m'amènent à une cartographie du pluralisme statistique paramétrique assez orthogonale à celle proposée par Hurlbert et Lombardi (2009, pp. 340-341) qui oppose les méthodes qui estiment la plausibilité absolue d'une hypothèse nulle ponctuelle aux méthodes "vraisemblantistes" qui estiment la plausibilité relative entre plusieurs modèles. Je propose plutôt d'opposer d'un côté les méthodes qui comparent la "plausibilité" d'une hypothèse nulle ponctuelle à son hypothèse alternative intervalle et de l'autre côté des méthodes qui comparent les probabilités *a posteriori* que le paramètre d'intérêt soit dans différents intervalles.

Ce travail sur la magnitude des effets peut être poursuivi de manière très simple dans le cadre Bayésien par une analyse de risque pour la prise de décision (Wade, 2000, Rivot, 2003). Il suffit simplement d'intégrer une fonction d'utilité – ou à l'inverse une fonction de risque – associée à chaque prise de décision (par exemple mettre en place un plan de restauration, augmenter la surface terrière d'un peuplement,...), sur la distribution de probabilité postérieure pour avoir un risque moyen associé à chaque type de décision. On peut par exemple choisir la décision correspondant au risque moyen le plus faible.

Cette vision du test d'hypothèse "intervalle", en soi et par rapport aux hypothèses ponctuelles ou aux comparaisons de modèles, peut être mise en relation avec les différentes visions de la modélisation présentées dans la partie 2.2.2 :

- classiquement, dans le cas de modèles statistiques, la vision "gouffre" de la modélisation utilisera le test d'hypothèse nulle ponctuelle, avec les risques associés ²⁶ ;
- la vision "écart" mettra davantage l'accent sur l'estimation des effets et de leur intervalle de confiance et sur les hypothèses intervalles ;
- alors que l'approche relativiste multi-modèle de Lakatos privilégiera des méthodes de comparaison de modèles.

4.2. De nouveaux tests d'ajustement du modèle aux données

Nous avons vu dans la partie 2.2.3 comment Quine et Duhem insistaient sur l'importance des hypothèses auxiliaires dans la construction des modèles et leur confrontation aux données. Ces auteurs insistent notamment sur la difficulté à savoir si l'on teste l'hypothèse principale ou une des hypothèses auxiliaires. En statistiques, les hypothèses auxiliaires constituent la partie de la description statistique qui n'est pas en question au moment du test de l'hypothèse principale. Néanmoins, pour s'assurer que la confrontation du modèle aux données concerne avant tout l'hypothèse principale, encore faut-il jauger avec des outils adéquats la qualité du modèle portant l'hypothèse. De manière assez étonnante, certains statisticiens, faisant écho à des propos d'écologues cités plus haut (Fagerström, 1987), démontent par avance l'intérêt de ce genre de méthodes (Clark et Gelfand, 2006 ²⁷) et, dans certains cas, semblent donner le modèle par avance vainqueur dans sa confrontation aux données et considérer que les détails probabilistes des modèles statistiques ont peu d'importance. C'est cette attitude que Peters (1991) avait en son temps dénoncée comme non-scientifique. Pour ma part, je prends le point de vue inverse, selon lequel la qualité probabiliste des modèles statistiques – i.e. l'adéquation des structures probabilistes auxiliaires du modèle statistique aux données – est plus importante qu'on ne le pense généralement, notamment pour des modèles statistiques qui sont de plus en plus compliqués et qu'on ne sait

²⁶ "There has not been a single date in the history of the law of gravitation when a modern significance test would not have rejected all laws and left us with no law" (Jeffreys, 1961 in Kass et Wasserman, 1996). Mais le processus n'en demeure pas moins utile : "A rejected model might be useful for many purposes (e.g., Newton's law) while its rejection encourages us to continue searching for an improved model (e.g., Einstein's relativity)." (Gelman *et al.*, 1996b; cf. also Rubin, 1996, Huber, 2002a).

²⁷ "Model adequacy is an inherently informal exercise as it is essentially impossible to 'calibrate' how well a complex process model 'fits' the data. Indeed, we might believe more in the model specification than in the data that we obtain. Especially in ecological models, data are notoriously noisy, signals might be weak and confidence in the specification might far exceed the information in the data" (p. 379)

plus forcément bien critiquer (O'Hagan, 2003²⁸). C'est une pratique assez développée en économétrie (Sawyer *et al.*, 1997), aussi défendue par certains écologues (Yoccoz, 1994, p.63, Guthery *et al.*, 2005 ; cf. aussi partie 4.3) ; Hilborn et Mangel (1997) en ont même fait un des quatre outils principaux du détective en écologie²⁹. Les statisticiens ne sont pas en reste, à commencer par Fisher, qui a donné ses lettres de noblesse à la critique des modèles par des tests d'adéquation du calage (ou "tests of significance" ; Anscombe, 1963, Christensen, 2005), à la suite duquel de nombreux statisticiens contemporains gardent une place de choix à la critique des modèles dans les développements des modèles statistiques : Box (1976 et 1980) en fait ainsi une des deux étapes principales du développement des modèles statistiques, Cox (1997) un des six outils principaux et McCullagh et Nelder (1989) un des trois principes de l'art de la modélisation. Cette pratique n'implique qu'un seul modèle et a une logique proche de celle de la démonstration par l'absurde : conditionnellement au modèle, nous évaluons dans quelle mesure les données contredisent le modèle. Cette approche se distingue ainsi du test d'une hypothèse nulle ponctuelle contre son alternative, ou de deux hypothèses intervalles mutuellement incompatibles, chères à l'approche de Neyman-Pearson (cf. Christensen, 2005). La critique des modèles est le prix à payer quand on utilise des méthodes statistiques basées sur des modèles. Et ce, d'autant plus dans le cas des statistiques Bayésiennes qui peuvent sembler trop dépendantes du modèle spécifié et intrinsèquement incapables de prendre du recul pour critiquer le modèle choisi³⁰. Le raisonnement des tests d'adéquation ou p-valeurs est le suivant (d'après Christensen, 2005) :

(i) on définit le modèle probabiliste des données, qui peut provenir en partie du modèle statistique ;

(ii) les données multidimensionnelles observées sont ramenées à une seule dimension grâce à une fonction, la statistique de test. Suivant la statistique de test choisie, on pourra critiquer différents aspects – ou hypothèses auxiliaires– du modèle, comme les distributions de probabilité utilisées, les fonctions de lien (cf. Figure 5) ;

(iii) on simule ou on calcule la distribution de la statistique de test sous l'hypothèse que le modèle probabiliste est bon ;

²⁸ "The genesis of this work lies in my growing unease that the power of HSSS [Highly Structured Stochastic Systems] modelling with MCMC computational tools was tempting us to build models that we did not know how to criticize effectively"

²⁹ Le statisticien Anscombe (1963) prend aussi l'image du détective – la même que celle de mon titre de rapport – pour donner une image familière du contexte dans lequel ces p-valeurs peuvent être utilisées.

³⁰ "The philosophical problem of Bayesian statistics is that it is congenitally unable to separate the model, the underlying true situation, and the statistical procedure. It acts as if the model were exactly true, and it then uses the corresponding optimal procedure. A fundamentalist Bayesian, for whom probabilities exist only in the mind, will not be able to see that there is a problem [...]; it takes a pragmatist like George Box to be aware of it." (Huber, 2009)

(iv) la p-valeur est la probabilité d'observer quelque chose d'au moins aussi étrange – par exemple aussi grand ou plus grand du point de vue de la statistique de test – que ce qu'on a observé, toujours sous le modèle probabiliste "testé".

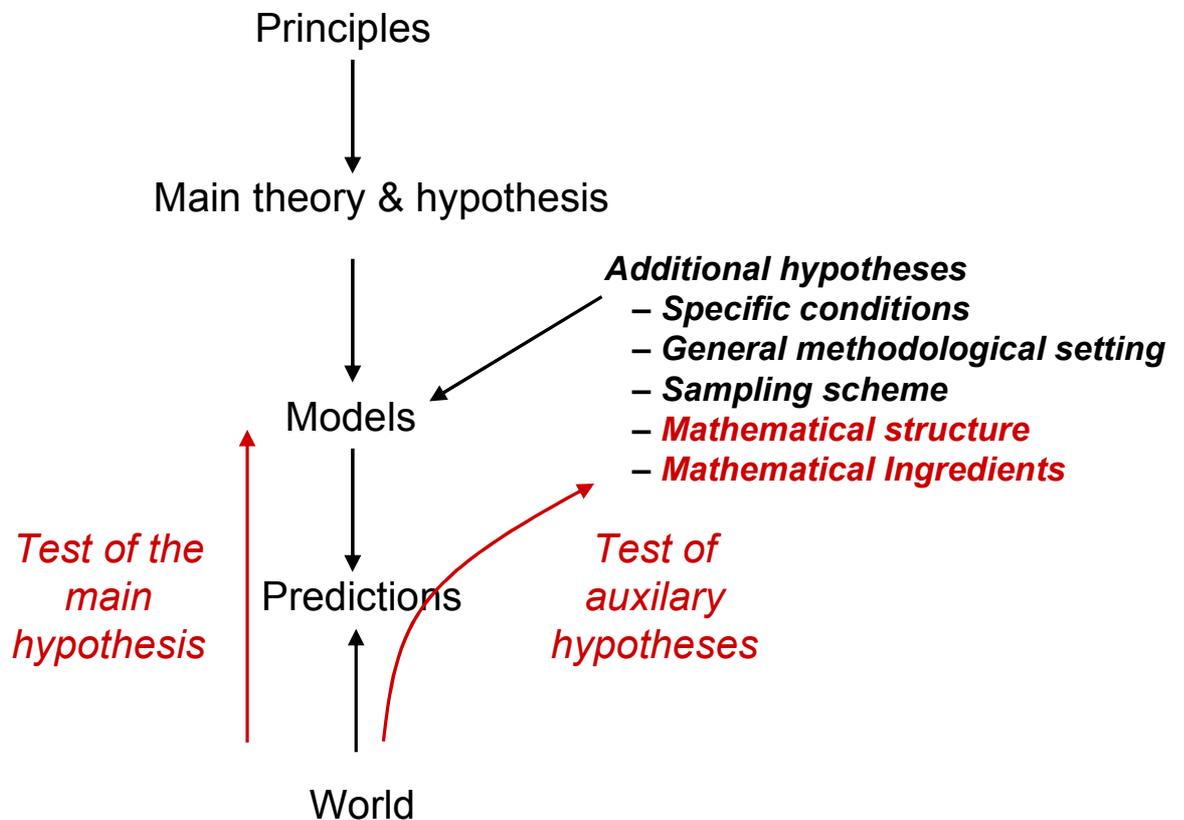


Figure 5. Représentation schématique des différents types de tests en statistique paramétrique : le test de l'hypothèse principale a lieu suite au processus d'estimation des paramètres du modèle à partir des données (éventuellement en comparant des modèles). Il est néanmoins dépendant selon la thèse de Duhem-Quine d'hypothèses auxiliaires, certaines concernant la structure du modèle statistique lui-même : d'autres types de tests cherchent donc à "critiquer" des composantes du modèle statistique lui-même.

Beaucoup de statisticiens utilisant les tests d'ajustement souhaitent ne pas employer le terme de "validation", mais plutôt ceux de "critique" (O'Hagan, 2003), "vérification" (O'Hagan, 2003, Bertrand-Krajewski, 2007), "évaluation" (Beck, 2002, Mentré et Escolano, 2006, Prisley et Mortimer, 2004). On pourrait aussi parler de corroboration de modèle pour faire le parallèle avec le terme employé dans l'approche Poppérienne (Soares *et al.*, 1995 in Prisley et Mortimer, 2004). Ce choix semble tout à fait justifié eu égard à la signification plus précise du terme de "validation" ³¹.

³¹ La validation vise à démontrer que le modèle dans un certain domaine d'applicabilité possède un niveau de précision satisfaisant par rapport à ce qu'on veut faire du modèle (Rykiel, 1996). Huber (2002a) propose de qualifier le modèle ayant passé la phase de validation de modèle "adéquate".

Que les structures "en dur" des modèles statistiques – comme les fonctions de lien plus ou moins non-linéaires, ou les structures de corrélation ou au contraire d'indépendance des variables aléatoires – soient importantes ne fait guère de doute. Par contre, que des structures plus molles comme des distributions de probabilités puissent avoir une influence parfois importante sur les résultats est moins présent dans les esprits des utilisateurs des statistiques. Je donnerai des exemples de ce point dans la partie suivante. Toujours est-il que cette pratique de la critique du modèle est, au moins dans son stade actuel, assez proche de la vision "gouffre" du modèle : le but est d'accroître la crédibilité d'un modèle en le soumettant à des critiques/tests poussés, ou de l'invalider (Holling 1978 in Rykiel, 1996, p.231). La démarche n'est toutefois pas vraiment stabilisée et laisse beaucoup de degrés de liberté à l'utilisateur (McCullagh et Nelder, 1989, Bouleau, 1999). Tout au plus peut-on définir quelques types de vérification fréquemment utilisés en écologie (Araújo et Guisan, 2006, p.1680) : (i) adéquation de la distribution des résidus, et notamment de leur variance ; (ii) adéquation de la fonction de lien ; (iii) recherche de valeurs aberrantes ; (iv) recherche de non-indépendance des données non prise en compte dans le modèle. Certains auteurs préconisent d'utiliser ces techniques sur des données autres que celles qui ont servi à estimer les coefficients du modèles (Goodall, 1972 et autres références dans Rykiel, 1996, Araújo *et al.*, 2005, Araújo et Guisan, 2006, Botkin *et al.*, 2007) – la pratique inverse de "resubstitution" des données étant tenue pour incorrecte. On peut aussi se poser des questions sur quelques pratiques fréquentes en écologie, qui de mon point de vue relèvent de la confusion, comme l'utilisation du pourcentage de variance expliquée (R^2 , Guisan et Thuiller, 2005) ou de la méthode de la surface sous la courbe ROC (AUC, Lobo *et al.*, 2008) comme méthodes de test de l'adéquation du modèle aux données.

4.2.1. De nouveaux tests d'ajustement à la distribution de référence connue

J'ai longtemps pratiqué le type de vérification décrit ci-dessus en utilisant l'estimateur du maximum de vraisemblance (cf. par exemple Pinheiro et Bates, 2000 et Harrell, 2001)³², avant d'opter pour des outils bayésiens du même acabit (p-valeur postérieure prédictive ;

³² développé et pratiqué par exemple dans le cadre des thèses d'Emmanuelle Richard (2004) et Christophe bouget (2004).

Gelman *et al.*, 1996a, Gelman *et al.*, 2004)³³. Ce n'est que récemment que j'ai réalisé que toutes ces méthodes étaient difficilement interprétables, simplement parce qu'on ne sait pas à quelle distribution de probabilité comparer les résultats : si j'observe une p-valeur de 0.4, est-ce surprenant ou pas ? A priori, si j'ai en tête la distribution uniforme sur [0;1], cette valeur n'est pas du tout surprenante. Par contre, si j'utilise la p-valeur postérieure prédictive comme simulée par exemple par Sinharay et Stern (2003) sur le modèle normal hiérarchique avec la fonction statistique moyenne des effets aléatoires, 95% des valeurs de la p-valeur devraient se trouver dans l'intervalle [0.47;0.53]. Dans ces conditions, la valeur de 0.4 est très surprenante... Le problème est que l'on ne connaît la distribution cible que quand on connaît le vrai modèle qui a généré les données – ce qui n'est pas le cas en général en statistiques appliquées –, dans certaines conditions, ou au prix de calculs très longs (Hjort *et al.*, 2006).

Certes, il existe des p-valeurs correctes pour certains types de distribution (Marchetti et Mudholkar, 2002) – et notamment pour les distributions ayant des statistiques ancillaires, c'est-à-dire qui ne dépendent pas des paramètres du modèle. Ce que je cherche néanmoins ici ce sont des p-valeurs dont la distribution de références soit connue et qui (i) s'appliquent dans n'importe quel contexte probabiliste ; (ii) puissent jauger des caractéristiques variables – i.e. différents points de vue – des modèles et (iii) puissent se mettre en œuvre facilement, sans besoin de calculs numériques supplémentaires lourds.

C'est ce genre de p-valeurs que j'ai proposé dans Gosselin (2010b) : ce sont ce que j'ai appelé des p-valeurs postérieures échantillonnées (en anglais, "sampled posterior p-value"). Leur calcul est très simple et ressemble grandement à ce qui est fait en fréquentiste : il consiste à générer des données virtuelles à partir du modèle considéré non pas à la valeur des paramètres correspondant au maximum de vraisemblance (Pierce et Schafer, 1986, Pinheiro et Bates, 2000, Harrell, 2001) mais à une valeur aléatoire des paramètres, tirée au sort dans la distribution postérieure à l'observation des données (cf. Figure 6).

³³ développé et pratiqué par exemple dans le cadre du master d'Elfie Perdereau (2006).

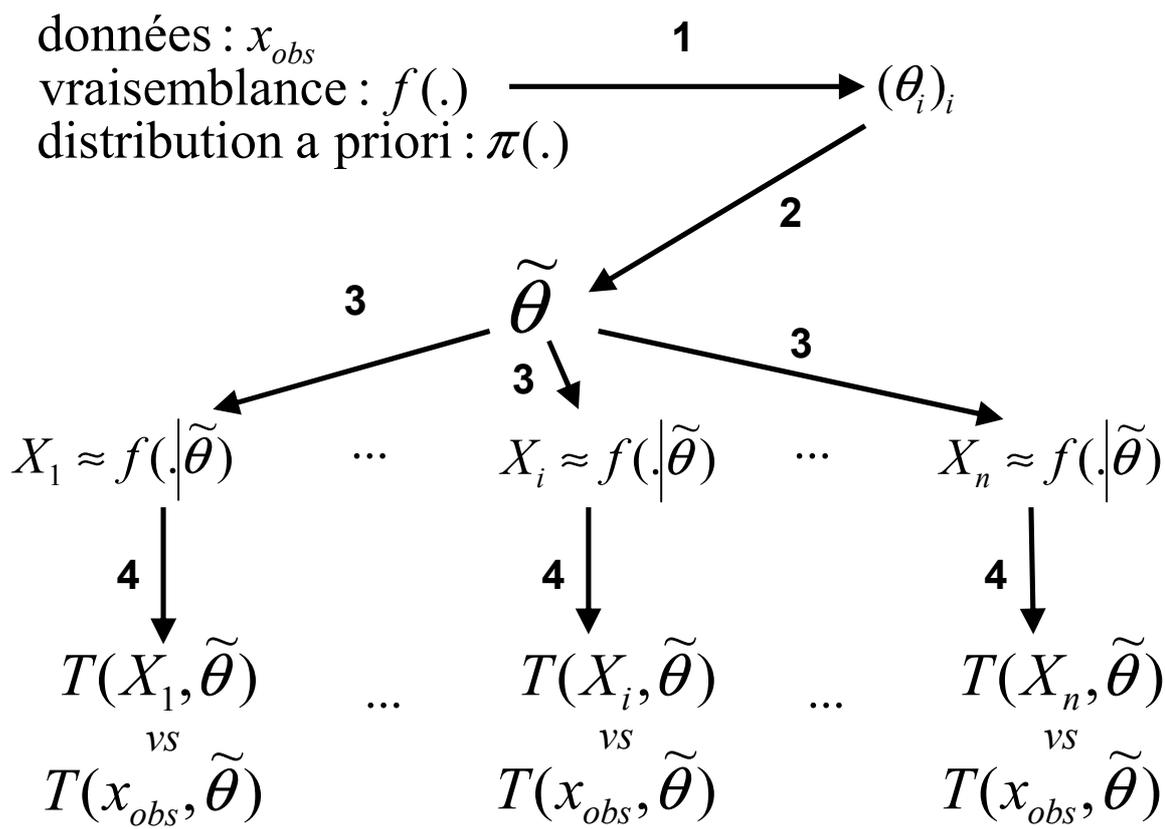


Figure 6. Représentation graphique du fonctionnement de la p-valeur postérieure échantillonnée. Etape 1 : estimation de la distribution des paramètres θ postérieure à l'observation des données ; Etape 2 : tirage au sort d'une valeur dans cette distribution ; Etape 3 : tirage au sort de données virtuelles à partir de cette valeur $\tilde{\theta}$ et du modèle ; Etape 4 : calcul de la fonction de discrédance pour chaque jeu de données virtuelles et comparaison à la discrédance observée (incluant $\tilde{\theta}$) pour former la p-valeur.

Les p-valeurs postérieures échantillonnées peuvent s'appliquer à la fois à des distributions continues et discrètes. De plus, une version normalisée de cette p-valeur devrait permettre des comparaisons plus faciles entre contextes différents. Les revers de la médaille de cette p-valeur sont (i) qu'elle n'a une distribution approximativement uniforme que pour de grands échantillons (i.e. le résultat est "asymptotique" – comme pour pratiquement tous les résultats connus sur le sujet) et (ii) qu'elle est aléatoire car elle dépend du tirage au sort d'une seule valeur du paramètre du modèle statistique dans la distribution de probabilité *a posteriori* – d'où son nom.

Néanmoins, ces limites me semblent mineures eu égard à la généralité et à la facilité d'utilisation de l'outil, qui donne plus de puissance statistique à l'utilisateur pour détecter des incohérences entre le modèle statistique et les données. A titre d'exemple, si les 100 effets aléatoires du modèle de Sinharay et Stern (2003) ne sont pas distribués normalement mais exponentiellement et qu'on applique un modèle où les effets aléatoires sont normaux avec la fonction de test statistique qui consiste à calculer le maximum des observations, seuls 2.5% de

p-valeurs seront plus grandes que 0.91 avec la p-valeur postérieure prédictive alors que 50% de p-valeurs seront plus grandes que 0.989 avec la p-valeur postérieure échantillonnée (Gosselin, In Prep.). On a donc plus de chances de détecter cette incohérence entre le modèle et les données avec la nouvelle p-valeur qu'avec l'ancienne. Un autre exemple pour des modèles plus simples, non hiérarchiques montre un avantage moins spectaculaire à la p-valeur postérieure échantillonnée (Figure 7). Cela devrait être un phénomène général car la p-valeur postérieure prédictive est connue pour être conservative (Robins *et al.*, 2000).

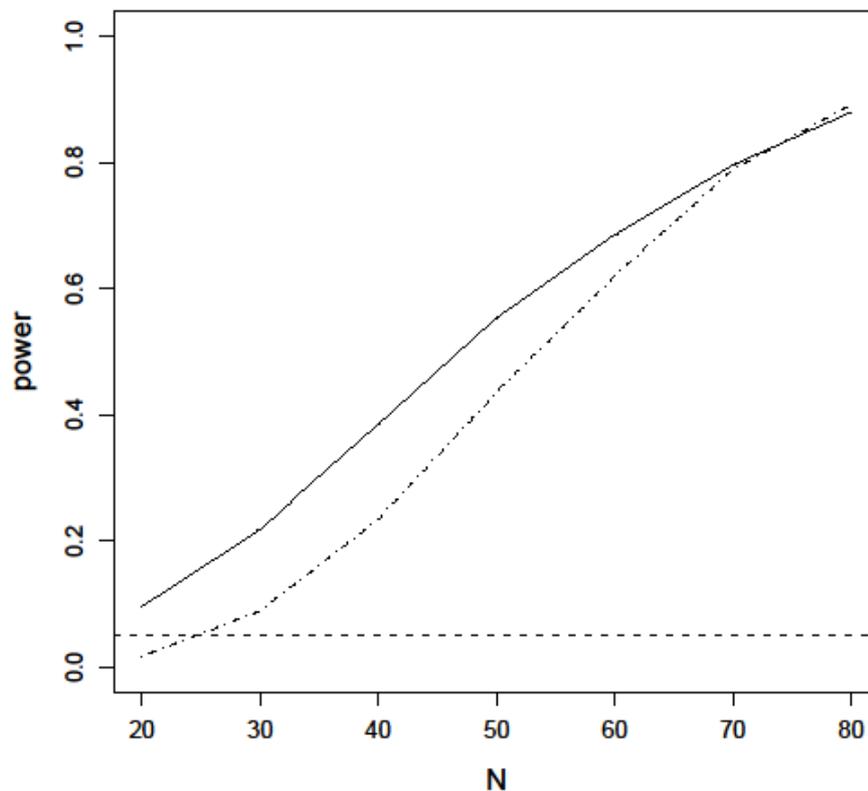


Figure 7. Puissance de la p-valeur postérieure échantillonnée (ligne continue) et de la p-valeur postérieure prédictive (ligne en pointillés alternés) à détecter une différence significative avec une probabilité inférieure à 5% entre les données observées simulées à partir d'une distribution de Polya avec $n_{\max} = 5$ et la distribution de Poisson utilisée pour analyser les données, le tout avec une taille d'échantillon N variant entre 20 et 80 et la fonction de test statistique "maximum" (Gosselin, 2010b).

4.2.2. Discussion sur les tests d'ajustement

Comme l'avaient proposé différents statisticiens (Box, 1980, Little, 2006), la critique des modèles Bayésiens emprunte des approches chères aux statisticiens fréquentistes. En effet, le point de vue utilisé pour développer la p-valeur postérieure échantillonnée incorpore un point de vue fréquentiste par l'étude de la distribution de la p-valeur quand on réplique un

grand nombre de fois les jeux de données et les analyses, sous condition d'un certain modèle. On peut néanmoins défendre aussi le point de vue inverse suivant lequel le développement d'une p-valeur purement fréquentiste ne fonctionne pas : en effet – jusqu'à preuve du contraire – il a été nécessaire d'introduire dans la démonstration de Gosselinb (2010b) une distribution *a priori* sur les paramètres, une caractéristique des statistiques Bayésiennes. Nous sommes donc conduits provisoirement à faire l'hypothèse qu'un contexte Bayésien calibré (pour reprendre les termes de Little, 2006) est le plus fertile pour développer des p-valeurs générales avec des propriétés connues – ainsi que d'autres pans des statistiques modernes (Bayarri et Berger, 2004).

On pourrait néanmoins rester dans la pratique purement Bayésienne du test d'adéquation qui consiste à comparer un modèle à tester à un modèle plus général, faisant moins d'hypothèses (Box, 1980) ; cela revient à préférer une approche Lakatosienne de la critique des modèles – fondée sur la comparaison de différents modèles aux hypothèses auxiliaires différentes – à une approche Poppérienne, cherchant à critiquer le modèle utilisé en soi. Soulignons, à la suite de Rubin (1996) et Huber (2002b) les limites de la première approche : nous avons aussi besoin en science de savoir si un modèle fonctionne en soi et pas uniquement s'il est le meilleur modèle parmi ceux qui ont été envisagés³⁴.

D'autres questions peuvent se poser par rapport à ces p-valeurs : finalement, comment doit-on les utiliser ? Que faire si une p-valeur est très surprenante ? Comment ces p-valeurs s'intègrent-elles dans la méthode des modèles ? Les réponses à ces questions dépendent grandement du contexte, ainsi que des visions des modèles qui sont adoptées – notamment entre gouffre et écart.

En premier lieu, signalons que les résultats proposés permettent de calculer des p-valeurs sur les données qui ont été utilisées pour estimer les paramètres du modèle statistique. Ils permettent ainsi d'utiliser les données deux fois, en faisant de la "re-substitution", contrairement aux recommandations habituelles en écologie (cf. références citées ci-dessus). Ce faisant, on teste bien l'adéquation des données aux hypothèses auxiliaires du modèle, et non la validité du modèle dans son ensemble pour d'autres données qui n'ont pas servi à l'estimation des paramètres – ce qu'on appelle la transportabilité modèle (Vaughan et

³⁴ [purely GOF Bayesian methods] “are based on the unwarranted presumption that by throwing in a few additional parameters one can obtain a perfectly fitting model. But how and where to insert those additional parameters often is far from obvious [...]. Remember that Kepler rejected the epicyclic Ptolemaic/Copernican models because he could not obtain an adequate fit within that class” ([Huber, 2002 #110233]).

Ormerod, 2005). Ce dernier problème est différent, mais pourrait aussi être traité avec des p-valeurs, mais plutôt par des tests d'ajustement externes.

L'adéquation du modèle statistique aux données n'est pas le seul critère de qualité des modèles. D'autres critères incluent : le lien avec les connaissances existantes (Cox, 2004, Prisley et Mortimer, 2004) ; le comportement "logique" ou mathématique ou crédible du modèle (Prisley et Mortimer, 2004, Caswell, 1976, Bayarri et Berger, 1999) ; l'utilisation qui sera faite du modèle et les critères associés (interprétabilité des paramètres...; Cox, 1990, Rykiel, 1996; Bouleau, 1999). Dans de nombreux cas, on pourra essayer de modifier certaines hypothèses auxiliaires du modèle pour remédier au problème. Cette solution ne sera toutefois pas toujours possible, et ce qui sera fait en cas de p-valeur surprenante est donc à adapter à chaque cas et est à la charge des disciplines selon le point de vue de Fisher (in Lehmann, 1990). Le point de vue de Cox (1990) me semble bien résumer cette tension ³⁵:

"Empirical models that do not fit will normally be replaced by models that do fit, but for substantive models in whose basic soundness there is considerable confidence, it may be more fruitful to specify qualitatively the nature of any failure and the broad interpretation to be given to the departures".

La définition de critères de validation du modèle en lien avec son utilisation est apparemment un processus non stabilisé (Bouleau, 1999, p.284, Rykiel, 1996) et les techniques proposées ci-dessus ne permettent pas vraiment de résoudre ce problème pratique. Tout au plus permettent-elles d'envisager de très nombreuses fonctions de statistiques servant à critiquer le modèle. Les p-valeurs développées permettent simplement de quantifier le degré de surprise d'un aspect des données par rapport au modèle, et non de révéler l'impact d'un mauvais choix dans le modèle pour l'utilisation du modèle (le "how wrong do we go when using this model for this data set" de Molenaar, 2004). Pour aborder cette question, Draper (1996) insiste quant à lui sur l'importance de mieux intégrer les notions d'utilité et la pratique de l'analyse de sensibilité dans la pratique des tests d'ajustement: on ne peut probablement pas en rester au niveau p-valeur, à moins de se rattacher à une notion assez vague de crédibilité (cf. Rykiel, 1996). A la place, il faut reconnaître le problème de la magnitude des "discrepancies" (Draper, 1996, Hill, 1996, Kass et Wasserman, 1996 ; cf. aussi partie précédente), et dans ces conditions les p-valeurs sont utiles au mieux comme des outils de diagnostics préliminaires (Hill, 1996). On en arrive alors à des outils considérant le modèle

³⁵ cette vision s'oppose pour partie au point de vue plus simple de Loehle, 1983 (citant Caswell, 1976) : "The problem is that goodness of fit is necessary but not sufficient as a criterion for the evaluation of a theory."

comme un écart à la réalité et non comme séparé par un gouffre de la réalité (cf. partie 2.2.2). En parallèle, l'utilisation de plusieurs modèles au lieu d'un seul pourrait être fort utile (Levins 1966, Bouleau 1999, Molenaar, 2004, Odenbaugh, 2006), un modèle validé n'étant jamais qu'une approche parmi d'autres (Bouleau, 2002, p. 113-114).

4.3. Distributions probabilistes pour données de comptage : ne pas négliger le versant sous-dispersé

Le choix d'une famille de distributions (de probabilité) appropriée pourrait être l'étape la plus ardue de l'analyse statistique (Cox, 1990). Pourtant, nous n'avons pas toujours beaucoup de choix en termes de distributions de probabilité. Ainsi, dans les modèles hiérarchiques, les distributions des effets aléatoires sont très souvent normales – en tout cas dans les procédures prêtes à l'emploi. De même, pour les données de comptage, la distribution de Poisson fait référence, avec les deux extensions que sont la distribution négative binomiale et les méthodes basées sur de la quasi-vraisemblance de type Poisson (Bolker *et al.*, 2009). Il existe aussi des modèles plus élaborés, incluant soit de l'inflation de zéros (i.e. une probabilité supplémentaire d'observer 0 par rapport aux distributions classiques) soit la probabilité de détection des espèces. Ces deux domaines sont très actifs actuellement en écologie, à juste titre. Rappelons que la loi de Poisson fait l'hypothèse très contraignante que la moyenne est égale à la variance. L'utiliser est un peu équivalent à utiliser la loi normale réduite (i.e. avec variance 1) pour analyser n'importe quel jeu de données continues. La distribution négative binomiale permet quant à elle de modéliser d'une certaine façon de la sur-dispersion, c'est-à-dire des cas où la variance est strictement supérieure à la moyenne.

En quoi cela pose-t-il un problème de ne pas utiliser la bonne distribution de probabilité dans le modèle statistique, notamment pour des données de comptage ? Il y a quelques années, nous portions une attention plus grande à l'homogénéité des variances qu'à la distribution des résidus (à la suite de Mitchell, 1977 *in* White et Bennetts, 1996). Nous avons donc mis l'accent sur la modélisation de l'hétéroscédasticité dans le cadre de modèles par ailleurs gaussiens³⁶. Cette approche semblait cohérente avec les simulations de White et Bennetts (1996) et McArdle et Anderson (2004), montrant une relative robustesse des modèles gaussiens utilisés avec des distributions de comptage qui pouvaient être sur-dispersées.

³⁶ par exemple dans le cadre de la thèse d'Emmanuelle Richard (2004) et du master de Julien Delcher (2005).

Pourtant, j'ai été amené à prêter davantage attention à la qualité de la distribution de probabilités employée, à la fois à partir de ma pratique – que je ne peux développer ici faute de place – et de la bibliographie. White et Bennetts (1996) ont ainsi montré à partir de simulations combien l'utilisation de la distribution de Poisson ou de la quasi-vraisemblance de type Poisson donnait des estimateurs ayant une mauvaise erreur-type sur des données simulées avec une distribution négative binomiale. McArdle et Anderson (2004) ont pour leur part montré que la stratégie de transformation des données en $\log(x+1)$ est assez catastrophique quand on effectue une analyse de variance des données simulées par une loi négative binomiale de moyenne fixée mais d'indice de dispersion variable. Enfin, Kikvidze et Moya-Larano (2008) ont montré par simulation des intervalles de confiance mauvais dans le cas d'ANOVA et ses équivalents non-paramétriques (y compris par randomisation) quand deux des problèmes suivants étaient présents en même temps : distribution asymétrique (ici log-normale) ; inégalité de variances entre traitements ; inégalité de taille d'échantillons entre traitements (cf. aussi Zimmerman, 2006).

Mon expérience et ces quelques exemples m'ont donc amené à réfléchir à des extensions des distributions de données de comptage³⁷. En effet, les méthodes présentées ci-dessus – hormis celle basée sur la quasi-vraisemblance – ne permettent pas de prendre en compte une éventuelle sous-dispersion des données de comptage. Il semble en effet tenu pour acquis en écologie que le "minimum" de variabilité pour des données de comptage soit la distribution de Poisson – pour laquelle variance=moyenne. En témoignent l'accent mis sur la sur-dispersion dans des articles méthodologiques (Anderson *et al.*, 1994, Bolker *et al.*, 2009), pouvant impliquer la non-considération délibérée de la possibilité de sous-dispersion (Burnham et Anderson, 2002, p.69). Or, des processus écologiques peuvent rendre les processus ponctuels plus réguliers, moins variables que les processus ponctuels Poissoniens. Un exemple tout simple correspond à l'analyse de la richesse spécifique d'un groupe mélangeant des espèces très fréquentes avec un très grand nombre d'espèces peu fréquentes (cf. Gosselin, 2010d). Nous avons observé ce cas de figure pour la richesse spécifique de certains groupes écologiques en Brie Francilienne (Barbier *et al.*, 2009, Gosselin, 2010d). J'ai donc effectué une recherche bibliographique dans d'autres domaines que l'écologie à la recherche de distributions de données de comptage sous-dispersées. J'ai été assez surpris d'en trouver en grand nombre (Gosselin, 2010c). Néanmoins, leur utilisation était

³⁷ travail commencé dans le cadre du master d'Elfie Perdereau (2006), poursuivi dans le cadre de la thèse de Stéphane Barbier (2007) et amplifié depuis.

systématiquement limitée par un problème (lien entre paramètres et moments de la distribution mal maîtrisé) ou un autre (la distribution de probabilités ne somme pas à 1) – sauf pour deux d'entre elles, qui sont des modifications de la loi de Poisson généralisée. Ces limites des modèles existants m'ont incité à construire – un peu au forceps – des distributions de comptage sous-dispersées ne présentant pas ces problèmes (Gosselin, 2010d). Ces distributions sont construites de façon à respecter au mieux les deux premiers moments de la distribution : la moyenne, qui comprend en général les variables explicatives ; et l'indice de dispersion, qui sera souvent constant mais pourra aussi être une fonction de variables. L'indice de dispersion voulu ne peut être atteint que s'il existe des distributions de données de comptage permettant cet indice de dispersion pour la moyenne modélisée. Par exemple, viser un indice de dispersion de 0.1 pour une moyenne de 0.5 est impossible à atteindre pour des données discrètes prenant des valeurs entières. Pour réussir à modéliser la sous-dispersion la plus forte avec des données discrètes, j'ai construit les distributions de probabilités comme étant des mélanges entre une distribution de Bernoulli chargeant les deux entiers les plus proches de la moyenne, et une autre distribution de probabilités (entre un mélange de distributions de Polya – la plus souple en termes de sous-dispersion –, les deux formes de distribution de Poisson généralisées issues de Gosselin, 2010c, et la distribution de Poisson). Le paramètre réglant le mélange est choisi de façon à être au plus près de l'indice de dispersion recherché. J'aboutis ainsi à des distributions qui sont certes un peu compliquées, mais qui ont les avantages suivants :

- (i) elles sont de vraies distributions de probabilité ;
- (ii) elles permettent de bien contrôler les moments à partir des paramètres du modèle ;
- (iii) elles sont des fonctions continues des paramètres ;
- (iv) elles tendent vers la distribution de Poisson quand l'indice de dispersion tend vers 1 ; ces distributions ont ainsi l'avantage de pouvoir être incluses dans une famille de distributions comportant la distribution de Poisson et la distribution négative binomiale couvrant toute la gamme possible de la sous-dispersion la plus faible à la sur-dispersion la plus forte.

J'ai utilisé ces distributions dans un cadre Bayésien sur les données de richesse spécifique mentionnées ci-dessus ; clairement, à l'échelle de 400 m², pour trois groupes écologiques sur six, les diagnostics des modèles étaient meilleurs que pour des modèles uniquement sur-dispersés, et les paramètres estimés gagnaient en précision – avec une baisse parfois de moitié de l'erreur-type ; aux échelles inférieures (4 et 100 m², modélisés ensemble), tous les groupes écologiques étaient sous-dispersés ; les niveaux de sous-dispersion étaient parfois élevés

($ID < 0.2$ pour les deux groupes ligneux à l'échelle de 100 m^2), cas dans lesquels les nouvelles distributions proposées montraient un avantage certain par rapport à celles dérivées de la littérature dans Gosselin (2010c).

L'utilisation de cette famille de distribution de manière plus générique devrait donc permettre de gagner en précision pour les cas où des distributions sous-dispersées sont rencontrées. Je ne sais pas à ce stade si cette situation est fréquente ; en guise d'exemple, signalons que la Figure 3 de Warton (2005), contient apparemment deux cas de situations sous-dispersées (sur 17).

5. Pistes de recherche à moyen terme

J'ai insisté dans ce mémoire sur l'utilité de modèles statistiques reliant la biodiversité aux pressions pesant sur elle. C'est dans cette thématique que se déploie mon projet de recherches principal sur 4 ans, en me restreignant à la biodiversité forestière. Des hypothèses écologiques ou à l'interface écologie/gestion ont déjà été définies dans les projets antérieurs, en lien avec les étudiants encadrés. Des pistes de développements méthodologiques ont aussi été identifiées, comme la modélisation de la biodiversité aux niveaux concomitants espèce et groupe écologique, dans le cadre de modèles paramétriques de l'abondance des espèces, comme proposé par Richard (2004) et Gelfand et al. (2005). Une partie de ces travaux n'a pas encore été publiée, notamment parce que le comportement des modèles statistiques utilisés ne nous donnait pas satisfaction. Mon projet de recherche est fondé sur l'hypothèse que ces comportements erratiques sont associés à une mauvaise adéquation entre le modèle statistique et les données – ce que j'appelle ici une mauvaise qualité probabiliste du modèle statistique. Le projet de recherche vise donc à améliorer cette qualité, pour tester les hypothèses écologiques sur des fondations statistiques plus saines.

Ce projet comprendra trois domaines principaux. D'abord, un travail pourra avoir lieu sur les indices d'équitabilité (suite à Gosselin, 2001b et Gosselin, 2006) ainsi que sur l'interface entre recherche en écologie et gestion ou applications (suite à Gosselin, 2008, Gosselin, 2009, Cordonnier et Gosselin, 2009 ; cf. aussi 6.3). Suite à l'analyse du cas de la chouette tachetée (Gosselin, 2009), on pourra être amené à réfléchir au rôle des modèles et des hypothèses (multiples) dans le cadre de ces interfaces : dans quelles conditions un vrai travail sur les hypothèses et modèles – plus abouti que dans le cas de la chouette tachetée – peut-il être organisé à ces interfaces ? De manière moins ambitieuse, comment nourrir cette interface : à partir de concepts, de méta-analyses, de savoir d'expert, de suivis (cf. un début de

réflexion dans Gosselin, 2010a) ? Enfin, peut-on relier l'évaluation des politiques publiques par des systèmes d'indicateurs avec la méthode des modèles (Gosselin et Barbier, 2005 vs Bockstaller et Girardin, 2003)?

Ensuite, le projet comprend le développement et le test d'**outils pour jauger la qualité probabiliste de modèles statistiques**. Il s'agit ici de continuer le travail présenté ci-dessus (4.2), en l'étendant aux modèles hiérarchiques ainsi qu'aux tests d'ajustement externes des modèles (i.e. à la comparaison des résultats prédits par le modèle à des données non utilisées pour estimer les paramètres du modèle statistique). Je prévois aussi de comparer la puissance statistique de ces méthodes sur des jeux de données simulés. Enfin, il s'agit de populariser ces développements en écologie, par exemple en les confrontant avec les outils principalement utilisés dans différents champs de l'écologie pour remplir des fonctions proches (AUC par exemple ; cf. partie 4.2).

Enfin, il s'agira de proposer des **structures probabilistes plus variées** (comme dans la partie 4.3) et de jauger – par simulation ou par analyse de données réelles – l'impact de mauvais choix distributionnels ou de fonctions de lien sur les résultats des analyses statistiques. Mon hypothèse est que dans certains cas de figure des mauvais choix distributionnels peuvent avoir des conséquences statistiques importantes, par exemple en termes de biais ou de précision. Ce volet du projet de recherche pourrait faire l'objet **d'un travail de thèse (cf. ci-dessous)**, qui devrait permettre d'améliorer la modélisation de la biodiversité aux niveaux concomitants espèce et groupe écologique. A titre d'exemple, des premières comparaisons par simulation entre modèles Bayésiens basés sur les distributions sous-dispersées présentées dans la partie 4.3 et analyses par quasi-vraisemblance de type Poisson, sur des données simulées avec une distribution de comptage sous-dispersée montrent la supériorité du modèle Bayésien sur le modèle par maximum de quasi-vraisemblance dans certaines situations (cf. Figure 8). Il semble d'ailleurs que la défaillance de la méthode par quasi-vraisemblance dans ces situations ne concerne pas que les distributions sous-dispersées.

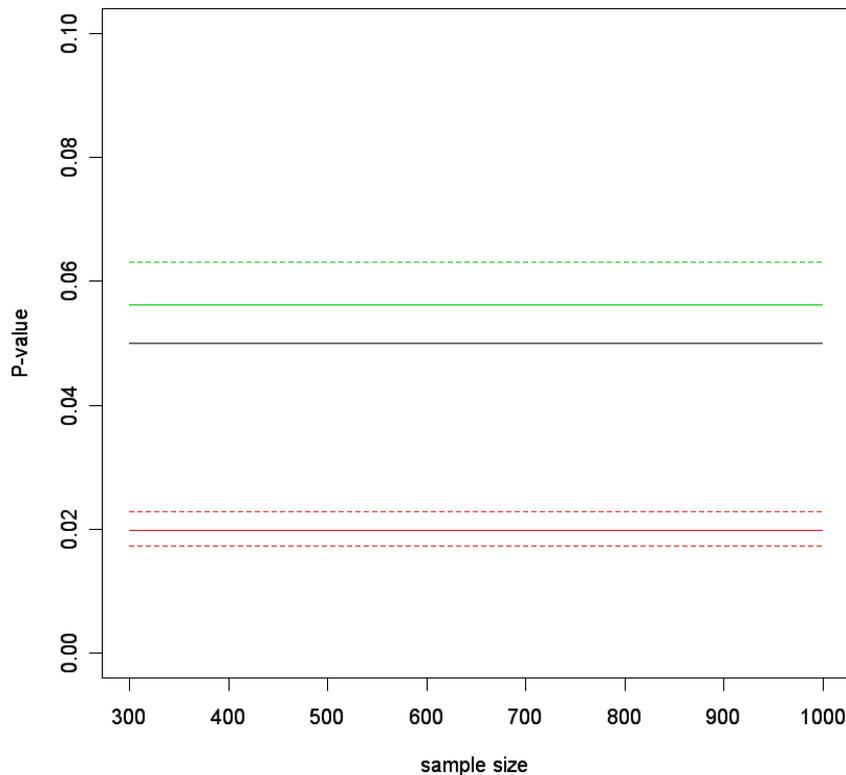


Figure 8. P-valeur – et son intervalle de confiance de 95% – issue des modèles Bayésiens (en vert) et de maximum de quasi-vraisemblance de type Poisson (en rouge) de détecter au seuil de 5% l'effet d'une covariable n'ayant pas d'effet, dans les simulations, sur les observations. La simulation impliquait une distribution sous-dispersée à indice de dispersion asymptotique de 0.5 et des valeurs moyennes de 0.05 (avec probabilité 0.5) et 6 (avec probabilité 0.6) (Gosselin, In Prep b). La valeur normalement attendue – 5%- est représentée par la ligne noire.

L'objectif du travail de thèse serait de proposer une nouvelle palette d'outils statistiques pour l'analyse de ce type de données, dans un cadre probabiliste "vrai" (en excluant des outils non-probabilistes comme les quasi-vraisemblances) en se concentrant sur trois points :

Axe 1 – les distributions de probabilité des variables de biodiversité ;

Axe 2 – les distributions de probabilité des effets aléatoires, notamment spatiaux ;

Axe 3 – différentes formes de fonctions non-linéaires reliant prédicteurs et données de biodiversité.

Ces différents travaux devraient notamment nous amener à renouveler le consensus écologique en matière d'utilisation de modèles hiérarchiques (Bolker *et al.*, 2009).

Le cœur de la thèse sera d'étudier ces trois points pour le cas des données de comptage (ligne orange foncé du tableau). Suivant l'avancée du projet, les deux derniers points pourront être abordés dans le cas de données en présence/absence (cases orange clair du tableau).

	Axe 1 – Distribution de probabilité biodiversité	Axe 2 – Distribution de probabilité effets aléatoires	Axe 3 – Fonctions de lien
Données de comptage	*	*	*
Données en présence-absence	A priori, la loi de Bernoulli devrait suffire	*	*

Néanmoins, une autre articulation pourrait voir le jour, en fonction notamment des projets déposés par l'équipe, dans lequel la thèse serait davantage axée sur les fonctions de lien (orange foncé), après ou parallèlement à un ou plusieurs Masters axés sur les distributions des effets aléatoires, y compris la modélisation de la dépendance spatiale entre données (approche géostatistique ou autre ; orange clair) :

	Axe 1 – Distribution de probabilité biodiversité	Axe 2 – Distribution de probabilité effets aléatoires	Axe 3 – Fonctions de lien
Données de comptage		*	*
Données en présence-absence	A priori, la loi de Bernoulli devrait suffire	*	*

Nous utiliserons ces outils statistiques rénovés sur les données de projets passés, pour tester les hypothèses écologiques de ces projets. Par ailleurs, les développements écologiques futurs concerneront aussi la thématique du lien entre naturalité, intensité de gestion et biodiversité, dans le cadre du projet GNB (Gestion, Naturalité et Biodiversité, initié avec l'ONF et RNF, comparant la biodiversité de réserves intégrales à celle de forêts gérées). Plus généralement, les outils statistiques seront utilisés pour jauger le caractère indicateur de différentes variables, notamment liées à la gestion forestière : pour chaque indicateur potentiel, il s'agira de définir les niveaux taxonomiques "bien" et "fortement" indiqués en même temps que les domaines écologiques dans lesquels ces relations ont lieu et que le niveau taxonomique le plus pertinent – espèces ou groupes écologiques (Richard, 2004, Gelfand *et al.*, 2005). Pour chaque espèce ou groupe écologique, il s'agira de développer des modèles statistiques incluant des indicateurs potentiels mais aussi des covariables (par exemple climatiques) dans le cadre de modèles non-linéaires renouvelés (de type "Structural Equation Model" non-linéaires ou modèles non-linéaires plus simples). Des formes de modèles non-linéaires plus explicites devraient permettre un meilleur lien avec les théories écologiques que des modèles avec des composantes non-paramétriques ou semi-

paramétriques (Austin, 2007). Enfin, on examinera en quoi les outils statistiques développés pourraient améliorer l'analyse des données de suivi de biodiversité.

6. Epilogue

Pour finir, je voudrais insister sur trois contradictions qui ont traversé mon discours – ici ou dans des travaux précédents –, et qui pourraient se révéler soit des impasses soit des tensions sources de nouvelles solutions – et donc des pistes de recherche à plus long terme.

6.1. *Magnitude de l'effet versus significativité statistique*

Si on voulait résumer mon propos dans la partie 4, on pourrait le décrire comme l'application de l'estimation de la magnitude des effets quand on estime les paramètres d'un modèle statistique, puis l'application de la significativité statistique pour critiquer le modèle et éventuellement le changer. Comme je l'ai déjà dit, cette vision ne "colle" pas. Il faudra donc travailler pour recoller les deux morceaux du puzzle, peut-être dans le cadre de la théorie de la décision, soit en développant une vision plus large de la notion de modèle (cf. par exemple McCullagh, 2002), soit en améliorant nos connaissances sur les impacts en termes d'estimateurs de mauvaises hypothèses probabilistes de modèles – c'est ce que je me propose de faire dans le moyen terme ; cf. ci-dessus –, soit en adoptant une vision clairement multi-modèles (et multi types de modèles), comme promue par Legay (1997) et Bouleau (1999).

6.2. *L'indisciplinarité, ses différentes formes et ses limites*

J'ai par ailleurs évoqué (voir encadré dans la partie 2.1.2) l'importance de l'indisciplinarité pour le travail de recherche dans les domaines environnementaux, point de vue défendu par Legay (1997)³⁸. Je pense avoir développé mon activité de recherche dans une perspective plus limitée de bi-disciplinarité active : en thèse, ce sont les prises en compte des contraintes biologiques et écologiques qui m'ont entraîné à rechercher des résultats mathématiques qui soient valables sous des hypothèses mathématiques qui correspondent à

³⁸ "La conclusion majeure qui accompagne depuis plusieurs années ces transformations est que les disciplines traditionnelles sont appelées à plus de modestie et doivent être considérées comme des points de vue au moins en partie arbitraires et souvent temporaires dans l'examen de ces objets. Il n'y a plus pour le chercheur pratiquant ni privilège d'une discipline, ni hiérarchie entre disciplines (4), mais plutôt organisation de ces disciplines, de leurs techniques et de leurs pratiques, en vue d'atteindre un objectif, selon la nature de cet objectif, et celle de l'objet bien entendu, cette organisation peut être différente."

des conditions acceptables sur le plan écologique ; dans mon travail actuel, mon activité bi-disciplinaire se développe à deux interfaces :

– biométrie/écologie, par la critique des outils statistiques existants et la proposition de nouveaux outils, pour des raisons issues de la pratique ;

– écologie/"gestion", par la discussion de l'interaction entre écologie académique et différents champs de la gestion (ingénierie écologique, conservation d'espèces et gestion forestière).

L'indisciplinarité a néanmoins des limites : en premier lieu, il faut signaler un certain malaise identitaire de l'indiscipliné à une époque d'affichage ou d'ancrage disciplinaire de plus en plus fort ; en second lieu, on ne peut – il me semble – être indiscipliné que dans quelques disciplines : de ce point de vue, l'indisciplinarité forte doit s'organiser et se construire collectivement : par exemple, pour proposer les analyses ou modèles des mêmes réalités à partir de différents points de vue, comme proposé par Legay (1997) et Bouleau (1999).

6.3. Expertise scientifique, connaissance scientifique et décision

Signalons enfin la tension qui existe entre la logique de connaissance scientifique et la logique de l'action – par exemple politique. Poincaré (1905) a bien mis en évidence cette tension qui peut apparaître comme une contradiction entre science et action. A première vue et à court terme, la recherche peut ainsi apparaître comme déstabilisante et non souhaitable pour l'action. Pourtant, Poincaré (1905) justifie cette recherche non pas uniquement par son utilité mais par un aspect éthique voire esthétique : elle est d'après lui consubstantielle à notre nature humaine. Il concluait :

"Aussi l'homme ne peut être heureux par la science, mais aujourd'hui il peut bien moins encore être heureux sans elle".

Un siècle plus tard, maintiendrait-on cette position, malgré les nombreuses vicissitudes que le XXème siècle scientifique ou technologique nous a apportées ? Toujours est-il que cette tension perdure, que la recherche de la vérité fait toujours partie de la condition humaine, et que même si la science a peut-être perdu de son aura par son rôle dans certaines affaires, bon nombre des progrès de notre civilisation lui sont dus. Ce qui pose davantage question c'est l'articulation entre la science et l'action, l'indépendance ou non du chercheur scientifique – pourtant au cœur de l'éthique scientifique selon Poincaré (1905) – et les différents types de recherche définis durant ces dernières décennies, de manière générale (recherche de mode 2

vs de mode 1 ; Gibbons *et al.*, 1994) ou en écologie (rechercher de catégories 1, 2, 3...; Underwood, 1995). Ainsi, une tension persiste – et persistera – pour les scientifiques entre leur travail académique et leur position d'experts ou d'interprètes (Roqueplo, 1999, Latour, 2004). Les scientifiques spécialistes de la biodiversité jouent ainsi à la fois un rôle de détective écologique – selon les mots de Hilborn et Mangel, 1997 (1997) – dans leur fonction de développement de connaissances sur des cas précis et d'interprète ou de porte-parole de celle-ci, dans les cas où les scientifiques sont questionnés, le plus souvent en tant qu'experts, à l'interface entre science et politique (comme par exemple sur le cas de la chouette tachetée : Gosselin, 2009). Il me semble important que ces deux facettes du travail du scientifique soient davantage connues et reconnues du politique³⁹ – pourquoi pas dans le cadre de processus politiques renouvelés (Roqueplo, 1999, Latour, 2004, Gosselin, 2009) – et qu'en même temps les scientifiques en écologie ne nourrissent pas l'interface entre science et action essentiellement avec des idées générales ou des concepts à la mode (Gosselin, 2010a).

7. Références bibliographiques citées

- anonyme, 1997, Modèle, *in* Serres M. et Farouki N. (Eds), *Le Trésor. Dictionnaire des Sciences*, p. 599-601.
- anonyme, 2007, *Likelihood principle*, http://en.wikipedia.org/wiki/Likelihood_principle#Arguments_for_and_against_the_likelihood_principle .
- Anderson D.R., Burnham K.P. et Thompson W.L., 2000, Null hypothesis testing: Problems, prevalence, and an alternative, *Journal of Wildlife Management*, 64, 4, p. 912-923.
- Anderson D.R., Burnham K.P. et White G.C., 1994, AIC model selection in overdispersed capture-recapture data, *Ecology*, 75, 6, p. 1780-1793.
- Anscombe F.J., 1963, Tests of goodness of fit, *Journal of the Royal Statistical Society, Series B*, 25, p. 81-94.
- Araújo M.B. et Guisan A., 2006, Five (or so) challenges for species distribution modelling, *Journal of Biogeography*, 33, 10, p. 1677-1688.
- Araújo M.B., Pearson R.G., Thuiller W. et Erhard M., 2005, Validation of species-climate impact models under climate change, *Global Change Biology*, 11, 9, p. 1504-1513.
- Austin M.P., 1979, Current approaches to the non-linearity problems in vegetation analysis, *in* Patil G.P. et Rosenzweig M.L. (Eds), *Contemporary Quantitative Ecology and Related Econometrics*, Fairland, MD, International Co-operative Publishing House, p. 197-210.
- Austin M., 2007, Species distribution models and ecological theory: A critical assessment and some possible new approaches, *Ecological Modelling*, 200, 1-2, p. 1-19.
- Badré M. et Décamps H., 2005, Michel Badré : la forêt au rythme des sciences et de la société, *Natures Sciences Sociétés*, 13, p. 428-436.

³⁹ ce qui n'a pas l'air d'être toujours le cas actuellement comme en témoigne la place assez maigre des scientifiques au Grenelle de l'environnement.

- Barbier S., Balandier P. et Gosselin F., 2009, Influence of several tree traits on rainfall partitioning in temperate and boreal forests: a review, *Annals of Forest Science*, 66, 602, .
- Barbier S., Chevalier R., Loussot P., Bergès L. et Gosselin F., 2009, Improving biodiversity indicators of sustainable forest management: tree genus abundance rather than tree genus richness and dominance for understory vegetation in French lowland oak hornbeam forests, *Forest Ecology and Management*, 258, Supplement 1, p. S176-S186.
- Barbier S., 2007, *Influence de la diversité, de la composition et de l'abondance des essences forestières sur la diversité floristique des forêts tempérées*, Ecole Doctorale "Sciences et Technologies", Ph.D thesis, Orléans, 2007 .
- Bayarri M.J. et Berger J.O., 1999, Quantifying surprise in the data and model verification, in Bernardo J.M., Berger J.O., Dawid A.P. et Smith A.F.M. (Eds), *Bayesian Statistics 6*, Oxford, Oxford University Press, p. 53-82.
- Bayarri M.J. et Berger J.O., 2004, The interplay of Bayesian and frequentist analysis, *Statistical Science*, 19, 1, p. 58-80.
- Bayliss J.L., Simonite V. et Thompson S., 2005, The use of probabilistic habitat suitability models for biodiversity action planning, *Agriculture, Ecosystems and Environment*, 108, 3, p. 228-250.
- Beck B., 2002, Model evaluation and performance, in El-Shaarawi A.H. et Piegorsch W.W. (Eds), *Encyclopedia of Environmetrics*, Chichester, John Wiley & Sons, p. 1275-1279.
- Bersier L.F. et Meyer D.R., 1994, Bird assemblages in mosaic forests: the relative importance of vegetation structure and floristic composition along the successional gradient, *Acta Oecologica*, 15, 5, p. 561-576.
- Bertrand-Krajewski J.L., 2007, Stormwater pollutant loads modelling: Epistemological aspects and case studies on the influence of field data sets on calibration and verification, *Water Science and Technology*, 55, 4, p. 1-17.
- Besag J. et Higdon D., 1999, Bayesian analysis of agricultural field experiments, *Journal of the Royal Statistical Society Series B-statistical Methodology*, 61, 4, p. 691-717.
- Beven K., 2002, Towards a coherent philosophy for modelling the environment, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 458, 2026, p. 2465-2484.
- Bockstaller C. et Girardin P., 2003, How to validate environmental indicators, *Agricultural Systems*, 76, 2, p. 639-653.
- Bolker B.M., Brooks M.E., Clark C.J., Geange S.W., Poulsen J.R., Stevens M.H. et White J.-S., 2009, Generalized linear mixed models: a practical guide for ecology and evolution, *Trends in Ecology and Evolution*, 24, 3, p. 127-135.
- Botkin D.B., Saxe H., Araujo M.B., Betts R., Bradshaw R.H., Cedhagen T., Chesson P., Dawson T.P., Ettlerson J.R., Faith D.P., Ferrier S., Guisan A., Hansen A.S., Hilbert D.W., Loehle C., Margules C., New M., Sobel M.J. et S., 2007, Forecasting the effects of global warming on biodiversity, *BioScience*, 57, 3, p. 227-236.
- Bouchaud J.P., 2008, Economics needs a scientific revolution, *Nature*, 455, 7217, p. 1181-.
- Bouleau N., 1999, *Philosophies des mathématiques et de la modélisation. Du chercheur à l'ingénieur*, Paris, France, L'Harmattan, 363 p.
- Bouleau N., 2002, La modélisation et les sciences de l'ingénieur, in Nouvel P. (Eds), *Enquête sur le concept de modèle*, Paris (France), Presses Universitaires de France, p. 101-119.
- Bourg D., 1996, Modernité et appartenance à la nature, *Esprit*, Juin 1996, p. 55-80.
- Boutin S., Haughland D., Schieck J., Herbers J. et Bayne E., 2009, A new approach to forest biodiversity monitoring in Canada, *Forest Ecology and Management*, 258, Supplement 1, p. S168-S175.

- Box G.E.P., 1976, Science and statistics, *Journal of the American Statistical Association*, 71, 356, p. 791-799.
- Box G.E.P., 1980, Sampling and Bayes' inference in scientific modelling and robustness, *Journal of the Royal Statistical Society, Series A*, 143, 4, p. 383-430.
- Buckland S.T., Goudie I.B.J. et Borchers D.L., 2000, Wildlife population assessment : past developments and future directions, *Biometrics*, 56, 1, p. 1-12.
- Bunnell F.L. et Huggard D.J., 1999, Biodiversity across spatial and temporal scales: problems and opportunities, *Forest Ecology and Management*, 115, 2-3, p. 113-126.
- Burnham K.P. et Anderson D.R., 2002, *Model selection and multimodel inference: a practical information theoretic approach*, New York, USA, Springer-Verlag.
- Camp R.J., Seavy N.E., Gorresen P.M. et Reynolds M.H., 2008, A statistical test to show negligible trend: Comment, *Ecology*, 89, 5, p. 1469-1472.
- Caswell H., 1976, The validation problem, in Patten B.C. (Eds), *Systems analysis and simulation in Ecology*, New York, Academic Press, p. 313-325.
- Caswell H., 1988, Theory and models in ecology: a different perspective, *Ecological Modelling*, 43, 1-2, p. 33-44.
- Chaloupka M. et Balazs G., 2007, Using Bayesian state-space modelling to assess the recovery and harvest potential of the Hawaiian green sea turtle stock, *Ecological Modelling*, 205, 1-2, p. 93-109.
- Chamberlin T., 1965, The method of multiple working hypotheses, *Science*, 148, p. 754-759.
- Christensen R., 2005, Testing Fisher, Neyman, Pearson, and Bayes, *The American Statistician*, 59, 2, p. 121-126.
- Clark J.S. et Gelfand A.E., 2006, A future for models and data in environmental science, *Trends in Ecology & Evolution*, 21, 7, p. 375-380.
- Clark J.S., 2005, Why environmental scientists are becoming Bayesians, *Ecology letters*, 8, 1, p. 2-14.
- Cohen J., 1990, Things I have learned (so far), *American Psychologist*, 45, p. 1304-1312.
- Conroy M.J., 1993, The use of models in natural resources management: prediction, not prescription, *Transactions of the North American Wildlife and Natural Resources Conference*, 58, p. 509-519.
- Cordonnier T. et Gosselin F., 2009, La gestion forestière adaptative: intégrer l'acquisition de connaissances parmi les objectifs de gestion, *Revue Forestière Française*, 61, 2, p. 131-144.
- Costanza R., 2003, A vision of the future of science: Reintegrating the study of humans and the rest of nature, *Futures*, 35, 6, p. 651-671.
- Cox D.R., 1990, Role of models in statistical analysis, *Statistical Science*, 5, 2, p. 169-174.
- Cox D.R., 1997, The current position of statistics: A personal view, *International Statistical Review*, 65, 3, p. 261-290.
- Cox D.R., 2004, Commentary on: Statistical distances as loss functions in assessing model adequacy, in Taper M.L. et Lele S.R. (Eds), *The nature of scientific evidence: statistical, philosophical, and empirical considerations*, Chicago, Illinois, University of Chicago Press, p. 478-480.
- Cressie N., Calder C.A., Clark J.S., Ver Hoef J.M. et Wikle C.K., 2009, Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling, *Ecological Applications*, 19, 3, p. 553-570.
- Crupi V., Tentori K. et Gonzalez M., 2007, On Bayesian measures of evidential support: Theoretical and empirical issues, *Philosophy of Science*, 74, 2, p. 229-252.
- de Gruijter J.J. et ter Braak C.J., 1990, Model-free estimation from spatial samples: A reappraisal of classical sampling theory, *Mathematical Geology*, 22, 4, p. 407-415.

- Delcher J., 2005, *Impacts de la taille des coupes et de la composition en essences sur la biodiversité interspécifique de Coléoptères Carabiques en forêt de Brie (77)*, Cemagref, Master 2- Sciences de l'Univers, Environnement, Ecologie- sp Ecologie, biodiversité, évolution, Paris, 2005 .
- Dennis B., 2004, Statistics and the Scientific method in ecology, in Taper M.L. et Lele S.R. (Eds), *The nature of scientific evidence: statistical, philosophical, and empirical considerations*, Chicago, Illinois, University of Chicago Press, p. 327-378.
- Di Stefano J., 2004, A confidence interval approach to data analysis, *Forest Ecology and Management*, 187, 2-3, p. 173-183.
- Dixon P.M. et Pechmann J.H.K., 2005, A statistical test to show negligible trend, *Ecology*, 86, 7, p. 1751-1756.
- Dorling J., 1979, Bayesian personalism, the methodology of scientific research programmes, and Duhem's problem, *Studies in History and Philosophy of Science*, 10, 3, p. 177-187.
- Draper D., 1996, Comment: Utility, sensitivity analysis, and cross-validation in Bayesian model checking, discussion of A. Gelman, X.-L. Meng, H. Stern, Posterior predictive assessment of model fitness via realized discrepancies, *Statistica Sinica*, 6, 4, p. 760-767.
- Edwards A.W.F., 1992, *Likelihood*, Baltimore, Maryland, USA, Johns Hopkins University Press.
- Edwards D., 1996, Comment: The first data analysis should be journalistic, *Ecological Applications*, 6, 4, p. 1090-1094.
- Etienne R.S. et Olf H., 2005, Confronting different models of community structure to species-abundance data: a Bayesian model comparison, *Ecology Letters*, 8, 5, p. 493-504.
- Fagerström T., 1987, On theory, data and mathematics in ecology, *Oikos*, 50, 2, p. 258-261.
- Ferrier S. et Guisan A., 2006, Spatial modelling of biodiversity at the community level, *Journal of Applied Ecology*, 43, 3, p. 393-404.
- Ferrier S., Watson G., Pearce J. et Drielsma M., 2002, Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling, *Biodiversity and Conservation*, 11, 12, p. 2275-2307.
- Fitelson B., 2010, Pollock on probability in epistemology, *Philosophical Studies*, 148, 3, p. 455-465.
- Fonnesbeck C.J. et Conroy M.J., 2004, Application of integrated Bayesian modeling and Markov chain Monte Carlo methods to the conservation of a harvested species, *Animal Biodiversity and Conservation*, 27, 1, p. 267-281.
- Gelfand A.E., Schmidt A.M., Wu S., Silander J.A. et Latimer A., 2005, Modelling species diversity through species level hierarchical modelling, *Journal of the Royal Statistical Society, Series C*, 54, 1, p. 1-20.
- Gelman A., Carlin J.B., Stern H.S. et Rubin D.B., 2004, *Bayesian Data Analysis*, Boca Raton, Chapman & Hall, 668 p.
- Gelman A., Meng X.L. et Stern H., 1996a, Posterior predictive assessment of model fitness via realized discrepancies, *Statistica Sinica*, 6, 4, p. 733-760.
- Gelman A., Meng X.L. et Stern H., 1996b, Rejoinder on Posterior predictive assessment of model fitness via realized discrepancies, *Statistica Sinica*, 6, 4, p. 796-807.
- Gibbons M., Limoges C., Nowotny H., Schwartzman S., Scott P. et Trow M., 1994, *The new production of knowledge: the dynamics of science and research in contemporary societies*, London, Sage.
- Giere R.N., 2004, How models are used to represent reality, *Philosophy of Science*, 71, 5, p. 742-752.

- Goldstein R.A., 1977, Reality and models: difficulties associated with applying general ecological models to specific situations, *in* Levin S.A. (Eds), *Mathematical models in biological discovery*, New York, NY, Springer, p. 206-215.
- Golicher D.J., O'Hara R.B., Ruiz-Montoya L. et Cayuela L., 2006, Lifting a veil on diversity: A Bayesian approach to fitting relative-abundance models, *Ecological Applications*, 16, 1, p. 202-212.
- Goodall D.W., 1972, Building and testing ecosystem models, *in* Jeffers J.N.J. (Eds), *Mathematical models in ecology*, Oxford, Blackwell, p. 173-194.
- Gosselin F. et Barbier S., 2005, Indicateurs, gradients écologiques et sélection de modèles. Cas de la composition et de la richesse en essences comme indicateurs de biodiversité floristique., *in* Barbault R. et Le Duc J.-P. (Eds), *Biodiversité. Science et gouvernance*, Unesco, p. 1-3.
- Gosselin F. et Gosselin M., 2004, Analyser les variations de biodiversité : outils et méthodes, *in* Gosselin M. et Larroussinie O. (Eds), *Biodiversité et gestion forestière : connaître pour préserver - synthèse bibliographique*, Antony, Coédition GIP Ecofor - Cemagref Editions, p. 58-99.
- Gosselin F., 1997, *Modèles stochastiques d'extinction de population: propriétés mathématiques et leurs applications*, Laboratoire de Probabilités, PhD, Paris, 1997 .
- Gosselin F., 1998, Reconciling theoretical approaches to stochastic patch-occupancy metapopulation models, *Bulletin of Mathematical Biology*, 60, 5, p. 955-971.
- Gosselin F., 1999, Test of mathematical assumptions behind the 'incidence function' estimation process of metapopulations' dynamic parameters, *Mathematical Biosciences*, 159, 1, p. 21-32.
- Gosselin F., 2001a, Asymptotic behavior of absorbing Markov chains conditional on nonabsorption for applications in conservation biology, *Annals of Applied Probability*, 11, 1, p. 261-284.
- Gosselin F., 2001b, Lorenz partial order: the best known logical framework to define evenness indices, *Community Ecology*, 2, 2, p. 197-207.
- Gosselin F., 2006, An assessment of the dependence of evenness indices on species richness, *Journal of Theoretical Biology*, 242, 3, p. 591-597.
- Gosselin F., 2008, Redefining ecological engineering to promote its integration with sustainable development and tighten its links with the whole of ecology, *Ecological Engineering*, 32, 3, p. 199-205.
- Gosselin F., 2009, Management on the basis of the best scientific data or integration of ecological research within management? Lessons learned from the northern spotted owl saga on the connection between research and management in conservation biology, *Biodiversity and Conservation*, 18, 4, p. 777-793.
- Gosselin F., 2010a, From ecology to ecological engineering: mainly through theory and concepts?, Soumis.
- Gosselin F., 2010b, Improving calibrated Bayesian goodness-of-fit methods: sampled posterior p-values as simple and general p-values that allow double use of the data, Soumis.
- Gosselin F., 2010c, Probability Distributions to Model Underdispersed Count Data: Nearly all Existing Solutions Have Serious Limitations in a Bayesian Regression Context, Soumis.
- Gosselin F., 2010d, Probability distributions to model under-dispersed count data in a Bayesian regression context: the case for mixtures of distributions involving the Bernoulli and Pòlya distributions, Soumis.
- Gotelli N.J. et Graves G.R., 1996, *Null models in ecology*, Washington, D.C., USA, Smithsonian Institution Press, 368 p.

- Grace J.B., Anderson T.M., Smith M.D., Seabloom E., Andelman S.J., Meche G., Weiher E., Allain L.K., Jutila H., Sankaran M., Knops J., Ritchie M. et Willig M.R., 2007, Does species diversity limit productivity in natural grassland communities?, *Ecology Letters*, 10, 8, p. 680-689.
- Grace J.B. et Bollen K.A., 2008, Representing general theoretical concepts in structural equation models: the role of composite variables, *Environmental and Ecological Statistics*, 15, 2, p. 191-213.
- Gregoire T.G., 1998, Design-based and model-based inference in survey sampling: Appreciating the difference, *Canadian Journal of Forest Research*, 28, 10, p. 1429-1447.
- Grimm V., 1994, Mathematical models and understanding in ecology, *Ecological Modelling*, 75-76, C, p. 641-651.
- Guisan A. et Thuiller W., 2005, Predicting species distribution: offering more than simple habitat models, *Ecology Letters*, 8, 9, p. 993-1009.
- Guthery F.S., Brennan L.A., Peterson M.J. et Lusk J.J., 2005, Information theory in wildlife science: Critique and viewpoint, *Journal of Wildlife Management*, 69, 2, p. 457-465.
- Harrell F.E., 2001, *Regression Modeling Strategies, With Applications to Linear Models, Logistic Regression, and Survival Analysis*, New York, USA, Springer, xxiii + 568 p.
- Hellman G., 1997, Bayes and Beyond, *Philosophy of Science*, 64, 2, p. 191-221.
- Henderson L., Goodman N.D., Tenenbaum J.B. et Woodward J.F., 2010, The structure and dynamics of scientific theories: A hierarchical bayesian perspective, *Philosophy of Science*, 77, 2, p. 172-200.
- Hilborn R. et Mangel M., 1997, *The ecological detective: confronting models with data*, Princeton (NJ), Princeton University Press, xvii+315 p.
- Hill B.M., 1996, Comment: discussion of A. Gelman, X.-L. Meng, H. Stern, Posterior predictive assessment of model fitness via realized discrepancies, *Statistica Sinica*, 6, 4, p. 767-773.
- Hjort N.L., Dahl F.A. et Hognadottir G., 2006, Post-processing posterior predictive p values, *Journal of the American Statistical Association*, 101, 475, p. 1157-1174.
- Hoening J.M. et Heisey D.M., 2001, The abuse of power: The pervasive fallacy of power calculations for data analysis, *The American Statistician*, 55, 1, p. 19-24.
- Huber P.J., 1997, Robustness: where are we now?, *IMS Lecture Notes–Monograph Series*, 31, p. 487-498.
- Huber P.J., 2002a, Approximate models, in Huber-Carol C., Balakrishnan N., Nikulin M.S. et Mesbah M. (Eds), *Goodness-of-fit tests and model validity*, Boston, Birkhäuser, p. 25-41.
- Huber P.J., 2002b, [Uncertainty, Policy Analysis and Statistics]: Comment, *Statistical Science*, 2, 3, p. 279-281.
- Huber P.J., 2009, On the non-optimality of optimal procedures, *IMS Lecture Notes–Monograph Series*, 57, p. 31-46.
- Hurlbert S.H. et Lombardi C.M., 2009, Final collapse of the Neyman-Pearson decision theoretic framework and rise of the neoFisherian, *Annales Zoologici Fennici*, 46, 5, p. 311-349.
- Innis G.S., Schlesinger S. et Sylvester, R. J., 1977, Model certification--Varying views from different specialties, in anonymous (Eds), *Summer Computer Simulation Conference Proceedings*, Montvale, NJ, AFIPS Press, p. 695-699.
- Jeffreys H., 1961, *Theory of probability*, London, Oxford University Press.
- Kass R.E. et Wasserman L., 1996, Comment on A. Gelman, X.-L. Meng, H. Stern, Posterior predictive assessment of model fitness via realized discrepancies, *Statistica Sinica*, 6, 4, p. 774-779.

- Kéry M. et Royle J.A., 2008, Hierarchical Bayes estimation of species richness and occupancy in spatially replicated surveys, *Journal of Applied Ecology*, 45, 2, p. 589-598.
- Kikvidze Z. et Moya-Larano J., 2008, Unexpected failures of recommended tests in basic statistical analyses of ecological data, *Web Ecology*, 8, p. 67-73.
- Laara E., 2009, Statistics: reasoning on uncertainty, and the insignificance of testing null, *Annales Zoologici Fennici*, 46, 2, p. 138-157.
- Lakatos I., 1978, *The methodology of scientific research programmes*, New York, Cambridge University Press.
- Latour B., 2004, *Politiques de la nature. Comment faire entrer les sciences en démocratie*, La Découverte, 382 p.
- Le Moigne J.-L., 1989, Qu'est-ce qu'un modèle?, *Confrontations Psychiatriques*, 30, p. 11-36.
- Lecoutre B., Lecoutre M.P. et Poitevineau J., 2001, Uses, abuses and misuses of significance tests in the scientific community: Won't the Bayesian choice be unavoidable?, *International Statistical Review*, 69, 3, p. 399-417.
- Lecoutre B., 2005, Et si vous étiez un bayésien qui s'ignore?, *Modulad*, 32, p. 92-105.
- Lee S.Y., 2007, *Structural Equation Modelling: A Bayesian Approach*, Wiley, 458 p.
- Legay J.M., 1986, Quelques réflexions à propos d'écologie. Défense de l'indisciplinarité, *Acta Oecologica, Oecologia Generalis*, 7, 4, p. 391-398.
- Legay J.-M., 1997, *L'expérience et le modèle. Un discours sur la méthode*, Paris, INRA Editions, 111 p.
- Lehmann E.L., 1990, Model specification: the views of Fisher and Neyman, and later developments, *Statistical Science*, 5, 2, p. 160-168.
- Lele S.R., Dennis B. et Lutscher F., 2007, Data cloning: Easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods, *Ecology Letters*, 10, 7, p. 551-563.
- Lele S.R. et Dennis B., 2009, Bayesian methods for hierarchical models: Are ecologists making a Faustian bargain?, *Ecological Applications*, 19, 3, p. 581-584.
- Levins R., 1966, The strategy of model building in population ecology, *American Scientist*, 54, 4, p. 421-431.
- Lévy-Leblond J.M., 1977, Des mathématiques catastrophiques, *Critique*, 33, 359, p. 430-441.
- Lewin-Koh N., Taper M.L. et Lele S.R., 2004, A brief tour of statistical concepts, in Taper M.L. et Lele S.R. (Eds), *The nature of scientific evidence: statistical, philosophical, and empirical considerations*, Chicago, Illinois, University of Chicago Press, p. 3-16.
- Little R.J., 2006, Calibrated Bayes: A Bayes/frequentist roadmap, *American Statistician*, 60, 3, p. 213-223.
- Lobo J.M., Jiménez-Valverde A. et Real R., 2008, AUC: a misleading measure of the performance of predictive distribution models, *Global Ecology Biogeography*, 17, 2, p. 145-151.
- Loehle C., 1983, Evaluation of theories and calculation tools in ecology, *Ecological Modelling*, 19, 4, p. 239-247.
- Lukacs P.M., Thompson W.L., Kendall W.L., Gould W.R., Doherty Jr. P.F., Burnham K.P. et Anderson D.R., 2007, Concerns regarding a call for pluralism of information theory and hypothesis testing, *Journal of Applied Ecology*, 44, 2, p. 456-460.
- Marchetti C.E. et Mudholkar G.S., 2002, Characterization theorems and goodness-of-fit tests, in Huber-Carol C., Balakrishnan N., Nikulin M.S. et Mesbah M. (Eds), *Goodness-of-fit tests and model validity*, Boston, Birkhäuser, p. 125-142.
- Marcot B.G., Hohenlohe P.A., Morey S., Holmes R., Molina R., Turley M.C., Huff M.H. et Laurence J.A., 2006, Characterizing species at risk II: Using Bayesian belief networks as decision support tools to determine species conservation categories under the Northwest Forest Plan, *Ecology and Society*, 11, 2, .

- Mathiot J., 2002, La légitimité paradoxale des modèles, in Nouvel P. (Eds), *Enquête sur le concept de modèle*, Paris (France), Presses Universitaires de France, p. 223-236.
- McArdle B.H. et Anderson M.J., 2004, Variance heterogeneity, transformations, and models of species abundance: a cautionary tale, *Canadian Journal of Fisheries and Aquatic Sciences*, 61, 7, p. 1294-1302.
- McCullagh P. et Nelder J.A., 1989, *Generalized linear methods*, London, Chapman.
- McCullagh P., 2002, What is a statistical model? (with discussion), *Annals of Statistics*, 30, 5, p. 1225-1310.
- McQuarrie A. et Tsai C.L., 1999, Model selection in orthogonal regression, *Statistics & Probability Letters*, 45, 4, p. 341-349.
- Mentré F. et Escolano S., 2006, Prediction discrepancies for the evaluation of nonlinear mixed-effects models, *Journal of Pharmacokinetics and Pharmacodynamics*, 33, 3, p. 345-367.
- Mitchell J.H., 1977, *A robustness study of the analysis of variance on negative binomial data*, Seattle, WA, USA, 1977 .
- Molenaar I.W., 2004, About handy, handmade and handsome models, *Statistica Neerlandica*, 58, 1, p. 1-20.
- Morad M., 2004, Talking hypothetically: the Duhem-Quine thesis, multiple hypotheses and the demise of hypothetico-deductivism, *Geoforum*, 35, 6, p. 661-668.
- Nichols J.D. et Williams B.K., 2006, Monitoring for conservation, *Trends in Ecology and Evolution*, 21, 12, p. 668-673.
- Nichols J.D., 2001, Using models in the conduct of science and management of natural resources, in Shenk T.M. et Franklin A.B. (Eds), *Modeling in natural resource management. Development, interpretation, and application*, Washington, D. C., Island Press, p. 11-34.
- North P.M. et Jeffers J.N.R., 1990, Modelling: a basis for management or an illusion?, in Spellerberg I.F., Goldsmith F.B. et Morris M.G. (Eds), *The scientific management of temperate communities for conservation*, London, UK, Blackwell, p. 523-541.
- Nouvel P., 2002, Modèles et métaphores, in Nouvel P. (Eds), *Enquête sur le concept de modèle*, Paris (France), Presses Universitaires de France, p. 189-202.
- Odenbaugh J., 2005, Idealized, inaccurate but successful: A pragmatic approach to evaluating models in theoretical ecology, *Biology and Philosophy*, 20, 2-3, p. 231-255.
- Odenbaugh J., 2006, The strategy of "The strategy of model building in population biology", *Biology and Philosophy*, 21, 5, p. 607-621.
- O'Hagan A., 2003, HSSS model criticism, in Green P.J., Hjort N.L. et Richardson S.T. (Eds), *Highly Structured Stochastic Systems*, Oxford University Press, p. 423-444.
- Perdereau E., 2006, *Impact de la modalité d'exploitation et des habitats sur les Coléoptères carabiques dans les parcelles touchées par les chablis de 1999 en Brie (77)*, Mention Ecosystèmes Terrestres et action de l'homme, Master II, Orléans, 2006 .
- Pereira H.M. et Cooper H.D., 2006, Towards the global monitoring of biodiversity change, *Trends in Ecology & Evolution*, 21, 3, p. 123-129.
- Peters R.H., 1991, *A critique for ecology*, Cambridge, Cambridge University Press, 366 p.
- Pierce D. et Schafer D., 1986, Residuals in generalized linear models, *Journal of the American Statistical Association*, 81, 396, p. 977-986.
- Pinheiro J.C. et Bates D.M., 2000, *Mixed-effects models in S and S-PLUS*, New York, Springer, 528 p.
- Poincaré H., 1905, *La valeur de la science*, Paris, Flammarion.
- Prisley S.P. et Mortimer M.J., 2004, A synthesis of literature on evaluation of models for policy applications, with implications for forest carbon accounting, *Forest Ecology and Management*, 198, 1-3, p. 89-103.

- Richard E., 2004, *Réponse des communautés de carabiques à la conversion en futaie régulière de chêne : aspects écologiques et méthodologiques*, Sciences de l'Environnement, Thèse de Doctorat, Paris, 2004 .
- Rivot E., 2003, *Investigations bayésiennes de la dynamique des populations de saumon atlantique (Salmo salmar l.) : des observations de terrain à la construction du modèle statistique pour apprendre et gérer*, spécilité halieutique, PhD, 2003 .
- Robins J.M., van der Vaart A. et Ventura V., 2000, Asymptotic distribution of P values in composite null models, *Journal of the American Statistical Association*, 95, 452, p. 1143-1156.
- Romesburg H.C., 1991, On improving the natural resources and environmental sciences, *Journal of Wildlife Management*, 55, 4, p. 744-756.
- Roqueplo P., 1995, La forêt menacée: le regard du sociologue, in Meiller D. et Vannier P. (Eds), *La forêt. Les savoirs et le citoyen. Regards croisés sur les acteurs, les pratiques et les représentations*, Chalon-sur-Saône, France, ANCR, p. 85-88.
- Roqueplo P., 1999, *Entre savoir et décision, l'expertise scientifique*, Paris, INRA Editions, 112 p.
- Rotenberry J.T., 1985, The role of habitat in avian community composition: physiognomy or floristics?, *Oecologia*, 67, p. 213-217.
- Royall R., 2004, The likelihood paradigm for statistical evidence, in Taper M.L. et Lele S.R. (Eds), *The nature of scientific evidence: statistical, philosophical, and empirical considerations*, Chicago, Illinois, Univeristy of Chicago Press, p. 119-152.
- Rubin D.B., 1996, Comment: On posterior predictive p-values, discussion of A. Gelman, X.-L. Meng, H. Stern, Posterior predictive assessment of model fitness via realized discrepancies, *Statistica Sinica*, 6, 4, p. 787-792.
- Rykiel E.J.J., 1996, Testing ecological models: the meaning of validation, *Ecological Modelling*, 90, 3, p. 229-244.
- Sawyer K.R., Beed C. et Sankey H., 1997, Underdetermination in economics. The Duhem-Quine thesis, *Economics and Philosophy*, 13, 1, p. 1-23.
- Simberloff D., 1990, Hypotheses, errors and statistical assumptions, *Herpetologica*, 46, 3, p. 351-357.
- Sinharay S. et Stern H.S., 2003, Posterior predictive model checking in hierarchical models, *Journal of Statistical Planning and Inference*, 111, 1-2, p. 209-221.
- Soares P., Tomé M., Skovsgaard J.P. et Vanclay J.K., 1995, Evaluating a growth model for forest management using continuous forest inventory data, *Forest Ecology and Management*, 71, 3, p. 251-265.
- Stamos D.N., 2007, Popper, laws, and the exclusion of biology from genuine science, *Acta Biotheoretica*, 55, 4, p. 357-375.
- Ullman T.D., Goodman N.D. et Tenenbaum J.B., Theory acquisition as stochastic search, in anonymous (Eds), *Proceedings of the Thirty-Second Annual Conference of the Cognitive Science Society*, .
- Underwood A.J., 1995, Ecological research and (and research into) environmental management, *Ecological Applications*, 5, 1, p. 232-247.
- Uriarte M. et Yackulic C.B., 2009, Preaching to the unconverted, *Ecological Applications*, 19, 3, p. 592-596.
- Vacher C., Vile D., Helion E., Piou D. et Desprez-Loustau M.L., 2008, Distribution of parasitic fungal species richness: Influence of climate versus host species diversity, *Diversity and Distributions*, 14, 5, p. 786-798.
- Valpine P.D., 2009, Shared challenges and common ground for Bayesian and classical analysis of hierarchical statistical models, *Ecological Applications*, 19, 3, p. 584-588.

- Van Dongen S., 2006, Prior specification in Bayesian statistics: Three cautionary tales, *Journal of Theoretical Biology*, 242, 1, p. 90-100.
- Vaughan I.P. et Ormerod S.J., 2005, The continuing challenges of testing species distribution models, *Journal of Applied Ecology*, 42, 4, p. 720-730.
- Wade P.R., 2000, Bayesian methods in conservation biology, *Conservation biology*, 14, 5, p. 1308-1316.
- Walshe T., Wintle B., Fidler F. et Burgman M., 2007, Use of confidence intervals to demonstrate performance against forest management standards, *Forest Ecology and Management*, 247, 1-3, p. 237-245.
- Warton D.I. et Hudson H.M., 2004, A manova statistic is just as powerful as distance-based statistics, for multivariate abundances, *Ecology*, 85, 3, p. 858-874.
- Warton D.I., 2005, Many zeros does not mean zero inflation: Comparing the goodness-of-fit of parametric models to multivariate abundance data, *Environmetrics*, 16, 3, p. 275-289.
- White G.C. et Bennetts R.E., 1996, Analysis of frequency count data using the negative binomial distribution, *Ecology*, 77, 8, p. 2549-2557.
- Yates R., 1951, The influence of Statistical methods for research workers on the development of the science of statistics, *Journal of the American Statistical Association*, 46, 253, p. 19-34.
- Yoccoz N.G., Nichols J.D. et Boulinier T., 2001, Monitoring of biological diversity in space and time, *Trends in Ecology and Evolution*, 16, 8, p. 446-453.
- Yoccoz N.G., 1994, Rapport d'Habilitation à Diriger des Recherches, 1994 .
- Yoccoz N.G., 1999, Évolution de l'utilisation des statistiques : quelques réflexions sur le rôle des modèles, *Nature, Sciences, Sociétés*, 7, 4, p. 14-18.
- Zimmerman D.W., 2006, Two separate effects of variance heterogeneity on the validity and power of significance tests of location, *Statistical Methodology*, 3, 4, p. 351-374.

Annexe 1.

**Eléments de bilan d'activité en
lien avec la HDR**

Curriculum Vitae

Nom : Gosselin

Prénom : Frédéric

Né le : 24 avril 1968 au Mans (72)

Nationalité : Française

Etat civil : Marié, trois enfants

Etudes :

Institution : Université Pierre et Marie Curie, Paris
Date : 12/1993 à 06/1997
Diplôme : Doctorat en Probabilités et Applications

Institution : Université Pierre et Marie Curie, Paris
Date : 09/1992 à 09/1993
Diplôme : DEA en Probabilités et Applications

Institution : Ecole Nationale du Génie Rural, des Eaux et des Forêts
Date : 09/1991 à 12/1993
Diplôme : Ingénieur du Génie Rural, des Eaux et des Forêts

Institution : Ecole Polytechnique
Date : 09/1988 à 07/1991
Diplôme : Ingénieur diplômé de l'Ecole Polytechnique

Associations professionnelles :

Society for Conservation Biology (SCB)

IUFRO – The Global Network for Forest Science Cooperation (via le Cemagref)

Société Française de Statistiques

Amicale des Ingénieurs du GREF (AIGREF)

Expérience professionnelle :

Date	04/1997 à aujourd'hui
Lieu	Nogent-sur-Vernisson
Institution	Cemagref
Poste	Ingénieur-chercheur en écologie forestière
Description	<p>Animateur du thème de recherche Biodiversité et gestion des forêts de plaine de 1998 à fin 2003</p> <p>Montage de projets</p> <p>Enseignement et encadrement d'étudiants</p>

Date	12/1993 à 04/1997
Lieu	Montpellier
Institution	CEFE (CNRS)
Poste	Ingénieur GREF en FCPR
Description	Thèse

Expérience d'enseignement :

Etablissement	Période	Durée	Matière enseignée
Master EBE - Université Paris 6	2004 2005 2006 2007 2008	3h	Introduction à l'ingénierie écologique
ENGREF	2004 2005 2006 2007 2008 2009	3h	Introduction à l'ingénierie écologique
Formation des Ingénieurs Forestiers	2005 2006	3h	Biodiversité et gestion forestière
Master Ecosystèmes Terrestres et Actions de l'Homme – Université d'Orléans	2005 2006 2007 2008	3h	Principes de gestion des écosystèmes terrestres
CFPPA Le Chesnoy-Les Barres (formation continue)	2000 2001 2002 2003 2004	4 à 10h	Biodiversité et gestion forestière
DEA ETES - MNHN	1999 2000 2001	3h	Biodiversité et gestion forestière

Formation à et pour la recherche

Encadrement d'étudiants en master/DEA/stages ingénieur

Timstet C., 1995, *Modélisation analytique de la dynamique du mutualisme figuiers-pollinisateurs*, Palaiseau (France), Ecole Polytechnique (co-encadrée avec M.-C. Anstett).

Bouget C., 2000, *Echantillonnage de la diversité des Coléoptères Carabidae - Limites du piège à fosse en relation avec le milieu forestier*, Rapport technique de DEA, ENESAD Dijon, Université Claude Bernard Lyon I, Cemagref Nogent-sur-Vernisson, 28 p.

Vieilledent G., 2001, *Etude statistique de l'influence des paramètres sylvicoles sur la stabilité au vent des peuplements forestiers à base de chênes en forêt de Brie*, Stage d'Agro, Rennes, ENSAR, 120 p.

Delcher J., 2005, *Impacts de la taille des coupes et de la composition en essences sur la biodiversité interspécifique de Coléoptères Carabiques en forêt de Brie (77)*, Cemagref, Master 2- Sciences de l'Univers, Environnement, Ecologie- sp Ecologie, biodiversité, évolution, Paris, Université Pierre et Marie Curie, 54 p.

Perdureau E., 2006, *Impact de la modalité d'exploitation et des habitats sur les Coléoptères carabiques dans les parcelles touchées par les chablis de 1999 en Brie (77)*, Mention Ecosystèmes Terrestres et action de l'homme, Master II, Orléans, Université d'Orléans, 31 p.

Dallari R., 2007, *Synthèse bibliographique sur les indicateurs directs de biodiversité et leur application dans un suivi forestier*, Master report, Université de Paris 6, 66 p.

Vuidot A., 2009, *Contribution à l'amélioration du protocole de suivi d'espaces naturels protégés (SENP) en vue de l'étude de la biodiversité : Relevé des microhabitats dans les arbres*, Cemagref, Master 2, Université Paris Est, 75 p. (stage co-encadré avec Y. paillet)

Co-encadrement d'étudiants en thèse

Bouget C., 2004, *Chablis et diversité des coléoptères en forêt feuillue de plaine : impact à court terme de la trouée, de sa surface et de son contexte paysager*, Ecologie, Thèse de Doctorat, Paris, Muséum National d'Histoire Naturelle, 452 p.

Richard E., 2004, *Réponse des communautés de carabiques à la coupe de régénération et au cycle sylvicole, lors de la conversion vers la futaie régulière de chêne : lien avec les caractéristiques écologiques des espèces*, Sciences de l'Environnement, Thèse de Doctorat, Paris, ENGREF, 406 p.

Barbier S., 2007, *Influence de la diversité, de la composition et de l'abondance des essences forestières sur la diversité floristique des forêts tempérées*, Ecole Doctorale "Sciences et Technologies", Thèse de Doctorat, Orléans, Université d'Orléans, 273 p.

Liste des principales publications

Publications dans des revues scientifiques à comité de lecture indexées par l'ISI

16. Elek Z., Dauffy-Richard et **Gosselin F.**, 2010, Carabid species responses to hybrid poplar plantation in floodplains in France, *Forest Ecology and Management*, 260, 9, p. 1446-1455.
15. Vuidot A., Paillet Y., Archaux F. et **Gosselin F.**, Acceptée, Influence of tree characteristics and forest management on tree microhabitats in France, *Biological Conservation*.
14. Barbier S., Chevalier R., Loussot P., Bergès L. et **Gosselin F.**, 2009, Improving biodiversity indicators of sustainable forest management: tree genus abundance rather than tree genus richness and dominance for understory vegetation in French lowland oak hornbeam forests, *Forest Ecology and Management*, 258, Supplement 1, p. S176-S186.
13. **Gosselin F.**, 2009, Management on the basis of the best scientific data or integration of ecological research within management? Lessons learned from the northern spotted owl saga on the connection between research and management in conservation biology, *Biodiversity and Conservation*, 18, 4, p. 777-793.
12. **Gosselin F.** et Lebreton J.-D., 2009, Asymptotic properties of infinite Leslie matrices, *Journal of Theoretical Biology*, 256, 2, p. 157-163.
11. Barbier S., Balandier P. et **Gosselin F.**, 2009, Influence of several tree traits on rainfall partitioning in temperate and boreal forests: a review, *Annals of Forest Science*, 66, 602, .
10. **Gosselin F.**, 2008, Redefining ecological engineering to promote its integration with sustainable development and tighten its links with the whole of ecology, *Ecological Engineering*, 32, 3, p. 199-205.
9. Barbier S., **Gosselin F.** et Balandier P., 2008, Influence of tree species on understory vegetation diversity and mechanisms involved - a critical review for temperate and boreal forests, *Forest Ecology and Management*, 254, 1, p. 1-15.
8. Lebreton J.-D., **Gosselin F.** et Niel C., 2007, Extinction and viability of populations: Paradigms and concepts of extinction models, *Ecoscience*, 14, 4, p. 472-481.
7. **Gosselin F.**, 2006, An assessment of the dependence of evenness indices on species richness, *Journal of Theoretical Biology*, 242, 3, p. 591-597.
6. Archaux F., **Gosselin F.**, Bergès L. et Chevalier R., 2006, Effects of sampling time, quadrat richness and observer on exhaustiveness of plant censuses, *Journal of Vegetation Science*, 17, p. 299-306.
5. **Gosselin F.**, 2001, Lorenz partial order: the best known logical framework to define evenness indices, *Community Ecology*, 2, 2, p. 197-207.
4. **Gosselin F.**, 2001, Asymptotic behavior of absorbing Markov chains conditional on nonabsorption for applications in conservation biology, *Annals of Applied Probability*, 11, 1, p. 261-284.
3. **Gosselin F.**, 1999, Test of mathematical assumptions behind the 'incidence function' estimation process of metapopulations' dynamic parameters, *Mathematical Biosciences*, 159, 1, p. 21-32.
2. **Gosselin F.**, 1998, Reconciling theoretical approaches to stochastic patch-occupancy metapopulation models, *Bulletin of Mathematical Biology*, 60, 5, p. 955-971.
1. **Gosselin F.**, 1996, Extinction in a simple source/sink system: Application of new mathematical results, *Acta Oecologica International Journal of Ecology*, 17, 6, p. 563-584.

Publications dans d'autres revues scientifiques

Gosselin F., 1999, What kinds of stability and equilibrium for "discrete" demographic models? Mathematical results and their implications, *Ecologie*, 30, 4, p. 241-245.

Publications scientifiques soumises

Gosselin F., 2010, Improving calibrated Bayesian goodness-of-fit methods: sampled posterior p-values as simple and general p-values that allow double use of the data, Soumis.

Gosselin F., 2010, Probability distributions to model count data underdispersed relative to the Poisson distribution: the case for the Bernoulli/double Pôlya mixture, Soumis.

Gosselin F., 2010, Probability Distributions to Model Underdispersed Count Data: Nearly all Existing Solutions Have Serious Limitations in a Bayesian Regression Context, Soumis.

Gosselin, F. 2010. From ecology to ecological engineering: mainly through theory and concepts? Soumis

Publications dans des ouvrages en anglais (autres qu'actes de colloques)

Gosselin F. et Lebreton J.D., 2000, The potential of branching processes as a modeling tool for conservation biology, in Ferson S. et Burgman M.A. (Eds), *Quantitative methods for conservation biology*, New York, Springer, p. 199-225.

Publications dans des actes de colloques

Gosselin F., Archaux F. et Gosselin M., 2008, Suivre la biodiversité en forêt : Pourquoi ? Quoi? Comment ?, in Landmann G. et Landeau S. (Eds), *De l'observation des écosystèmes forestiers à l'information sur la forêt*, Quae, SympoScience, p. 26-32.

Gosselin F., 2007, Eléments bibliographiques à propos du lien entre bois mort et biodiversité, in Jaulin S. et al. e. (Eds), *Les vieux arbres et la conservation de la biodiversité*, Perpignan, OPIE-LR éditeur, p. 107-110.

Gosselin F. et Barbier S., 2005, Indicateurs, gradients écologiques et sélection de modèles. Cas de la composition et de la richesse en essences comme indicateurs de biodiversité floristique., in Barbault R. et Le Duc J.-P. (Eds), *Biodiversité. Science et gouvernance*, Unesco, p. 1-3.

Bouget C. et Gosselin F., 2005, Windthrow gaps as dead wood islands for saproxylic beetles in managed deciduous French forests, in Barclay M. et Telnov D. (Eds), *Latvijas entomologs (Proceedings of the 3rd Symposium and Workshop on the Conservation of Saproxylic Beetles)*, Acta Coleopterologica Latvica, p. 16-25.

Richard E. et Gosselin F., 2005, Ecological traits: surrogates for short-term carabid response to regeneration felling of mature oak stands?, in Serrano J.G. et Ruiz C. (Eds), *XII European Carabidologists Meeting, "Ground beetles as a key group for biodiversity conservation studies in Europe"*, Murcia, Spain, Nausicaä Edicion Electronica, S.L., p. 85-92.

Richard E., Gosselin F. et Lhonoré J., 2004, Short-term and mid-term response of ground beetle communities (*Coleoptera, Carabidae*) to disturbance by regeneration felling, in Honnay O., Bossuyt B., Verheyen K. et Hermy M. (Eds), *Forest Biodiversity:*

Lessons from history for conservation, Oxon, UK, IUFRO Res. Ser. CAB International, p. 179-192.

- Gosselin F., 2004, Évaluation de la biodiversité forestière en Brie : influence du type de peuplement, in Millier C., Barre V. et Landeau S. (Eds), *Restitution des résultats du programme de recherche Biodiversité et Gestion Forestière*, Paris, GIP-Ecofor, p. 124-131.
- Gosselin F., 2004, Enjeux et stratégie : l'avis du scientifique, in anonymous (Eds), *8ème forum des gestionnaires "Les enjeux écologiques dans la gestion forestière"*, Paris., p. 57-60.

Publications techniques dans des revues avec comité de lecture (certaines indexées au WOS ou à SCOPUS)

18. Cordonnier T. et Gosselin F., 2009, La gestion forestière adaptative: intégrer l'acquisition de connaissances parmi les objectifs de gestion, *Revue Forestière Française*, 61, 2, p. 131-144.
17. Cordonnier T., Gosselin F., Bouget C., Brézard J.-M. et Allain R., 2009, Gestion adaptative ou gestion expérimentale du bois mort, des vieux arbres et des arbres à cavités: exercice de prospective, *RenDez-Vous techniques*, 25-26, p. 34-37.
16. Gosselin F., Bergès L., Bouget C., Perdereau E., Thuault F., Dumas Y., Goujon G., Moliard C. et Legoff G., 2008, Réponse de la biodiversité aux chablis en Brie: influence de l'exploitation et de la taille des trouées, *Forêt-Entreprise*, 183, p. 28-32.
15. Barbier S., Gosselin F. et Balandier P., 2008, Le mélange d'essences est-il favorable à la diversité végétale en forêt ?, *Revue Forestière Française*, 60, 2, p. 159-167.
14. Gosselin F. et Gosselin M., 2008, Pour une amélioration des indicateurs et suivis de biodiversité forestière, *Ingénieries-EAT*, 55-56, p. 113-120.
13. Gosselin F., 2007, L'Europe au milieu du gué, *Les Dossiers de La Recherche*, 28, p. 83-83.
12. Bouget C., Gosselin F., Goujon G. et Moliard C., 2007, L'après-tempête en forêts feuillues de plaine : contribution des trouées de chablis à la diversité des coléoptères saproxyliques, *Bourgogne Nature*, 5, p. 176-182.
11. Bouget C., Goujon G., Moliard C. et Gosselin F., 2005, Contribution à la connaissance des Coléoptères des forêts feuillues de Brie (Seine-et-Marne). Partie II : des Cucujoidea aux Curculionoidea, *L'Entomologiste*, 61, 1, p. 17-32.
10. Bouget C., Goujon G., Moliard C. et Gosselin F., 2004, Contribution à la connaissance des Coléoptères des forêts feuillues de Brie (Seine-et-Marne). Partie I : des Caraboidea aux Cleroidea, *L'Entomologiste*, 60, 6, p. 261-275.
9. Gosselin F., 2004, Intégrer recherche scientifique en écologie et gestion dans le cadre de l'ingénierie écologique : intérêts et limites, *Ingénieries EAT*, n° spécial 2004, p. 113-120.
8. Gosselin F., 2004, Pour une définition de l'ingénierie écologique plus intégrée avec le développement durable et avec l'écologie, *Ingénieries EAT*, n° spécial 2004, p. 139-147.
7. Gosselin F., Bouget C. et Nageleisen L., 2004, Bois mort et biodiversité : synthèse scientifique pour une réflexion sur la gestion du bois mort, *Forêt Entreprise*, 155, p. 27-29.
6. Gosselin M., Gosselin F. et Bergès L., 2003, Le point de vue du scientifique. Dossier "Des indicateurs fiables pour une gestion forestière durable", *Forêt Entreprise*, 150, p. 37-39.

5. Gosselin F. et Bouget C., 2003, L'évolution des pratiques d'exploitation forestière pourrait bénéficier à "la" biodiversité: réflexions scientifiques autour du guide de reconstitution de l'ONF, suite à la tempête, *Ingénieries*, 35, p. 61-73.
4. Bergès L., Gosselin M., Gosselin F., Dumas Y. et Laroussinie O., 2002, Prise en compte de la biodiversité dans la gestion forestière : éléments de méthode, *Ingénieries - EAT*, n° spécial, p. 45-55.
3. Chevalier R., Bergès L., Gilbert J.M. et Gosselin F., 2001, Station forestière et végétation spontanée d'accompagnement : perspectives pour la diversification des jeunes pineraies de l'Orléanais, *Revue Forestière Française*, 53, 2, p. 151-170.
2. Le Tacon F., Selosse M.A. et Gosselin F., 2001, Biodiversité, fonctionnement des écosystèmes et gestion forestière. Deuxième partie : interventions sylvicoles et biodiversité, *Revue Forestière Française*, 53, 1, p. 55-80.
1. Le Tacon F., Selosse M.A. et Gosselin F., 2000, Biodiversité, fonctionnement des écosystèmes et gestion forestière. Première partie, *Revue Forestière Française*, 52, 6, p. 477-496.

Ouvrages collectifs

Landmann G., Gosselin F. et Bonhême I., 2009, *Bio2, Biomasse et Biodiversité forestières. Augmentation de l'utilisation de la biomasse forestière : implications pour la biodiversité et les ressources naturelles*, Paris, MEEDDM-ecofor,, 210p.

Publications dans des ouvrages en français

- Gosselin F., Bouget C., Gosselin M., Chauvin C. et Landmann G., 2009, L'état et les enjeux de biodiversité forestière en France, *in* Landmann G., Gosselin F. et Bonhême I. (Eds), *Bio2 - Biomasse et Biodiversité Forestière - Augmentation de l'utilisation de la biomasse forestière: implications pour la biodiversité et les ressources naturelles*, Paris, GIP Ecofor, MEEDM, p. 63-69.
- Landmann G. et Gosselin F., 2009, Utilisation de la biomasse forestière, biodiversité et ressources naturelles: synthèse et pistes d'approfondissement, *in* Landmann G., Gosselin F. et Bonhême I. (Eds), *Bio2 - Biomasse et Biodiversité Forestière - Augmentation de l'utilisation de la biomasse forestière: implications pour la biodiversité et les ressources naturelles*, Paris, GIP Ecofor, MEEDM, p. 177-191.
- Bouget C., Gosselin M., Gosselin F. et Bergès, 2009, Conséquences de l'augmentation des prélèvements de biomasse ligneuse pour la biodiversité forestière : chap. 7, *in* G. L., Gosselin F. et Bonhême I. (Eds), *Bio2 - Biomasse et Biodiversité Forestière - Augmentation de l'utilisation de la biomasse forestière: implications pour la biodiversité et les ressources naturelles*, Paris, GIP Ecofor, MEEDM, p. 73-87.
- Bouget C. et Gosselin F., 2005, Distribution spatiale du bois mort : enjeux pour la conservation des espèces cavicoles et saproxyliques, *in* Vallauri D., André J., Dodelin B., Eynard-Machet R. et Rambaud D. (Eds), *Bois mort et à cavités : une clé pour la biodiversité des forêts*, Paris, Lavoisier Tec et Doc, p. 107-113.
- Gosselin F. et Gosselin M., 2004, Analyser les variations de biodiversité : outils et méthodes, *in* Gosselin M. et Laroussinie O. (Eds), *Biodiversité et gestion forestière : connaître pour préserver - synthèse bibliographique*, Antony, Coédition GIP Ecofor - Cemagref Editions, p. 58-99.
- Gosselin F., 2004, Imiter la nature, hâter son œuvre ? Quelques réflexions sur les éléments et stades tronqués par la sylviculture, *in* Gosselin M. et Laroussinie O. (Eds), *Gestion*

Forestière et Biodiversité : connaître pour préserver - synthèse bibliographique, Antony, Coédition GIP Ecofor - Cemagref Editions, p. 217-256.

Gosselin F., 2004, Influence de la composition et de la richesse spécifique du peuplement arboré sur la biodiversité, *in* Gosselin M. et Laroussinie O. (Eds), *Biodiversité et gestion forestière : connaître pour préserver*, Antony, Coédition GIP Ecofor - Cemagref Editions, p. 127-148.

Gosselin M., Laroussinie O., Gosselin F. et Bergès L., 2004, Biodiversité et gestion forestière : pour une interaction croissante entre gestion et recherche, *in* Gosselin M. et Laroussinie O. (Eds), *Biodiversité et Gestion Forestière : connaître pour préserver - synthèse bibliographique*, Antony, Coédition GIP Ecofor - Cemagref Editions, p. 41-56.

Communications orales à congrès "technique" – sans publication écrite

Bouget C., Gosselin F., Gosselin M. et Jactel H., 2009, *Intensification des prélèvements de biomasse et préservation de la biodiversité forestière*, Ateliers REGEFOR La forêt face aux défis énergétiques, Champenoux, 08-10/06/2009 .

Rapports de stage ou de diplômes

Gosselin F., 1997, *Modèles stochastiques d'extinction de population: propriétés mathématiques et leurs applications*, Laboratoire de Probabilités, PhD, Paris, Université Paris 6, 297 p.

Gosselin F., 1996, *A journey into the country of spotted owls*, Paris, ENGREF, 258p.

Gosselin F., 1993, *A propos de l'existence de la distribution quasi-stationnaire de processus de ramification densité-dépendants en temps discret*, DEA, Paris, Université Paris 6, 64 p.

Annexe 2.

**Manuscrits soumis en lien
avec la HDR**

Université Pierre et Marie Curie

Mémoire d'Habilitation à Diriger des Recherches

Frédéric GOSSELIN

2011

Propositions pour améliorer l'équipement biométrique du détective écologique Application à la modélisation de la relation entre gestion forestière et biodiversité

Résumé. L'utilisation des mathématiques en écologie s'est considérablement développée durant la seconde moitié du vingtième siècle, souvent en prenant comme exemple la grande sœur que sont les sciences physiques. Pourtant, ce développement n'a pas été sans difficultés. Nous souhaitons dans ce mémoire faire le bilan de notre expérience de travail à cette interface, pour ensuite proposer quelques développements dans le domaine de la modélisation statistique des pressions pesant sur "la" biodiversité – avec un clin d'œil à l'ouvrage de Hilborn & Mangel (1997) expliquant le titre.

Dans un premier temps, nous rappelons et de discutons les différents rôles joués par les modèles en science – et notamment en écologie – en insistant notamment sur les différences de perspectives entre le modèle "gouffre", le modèle "écart" et le "meilleur modèle en stock". Cette première partie est aussi l'occasion d'insister sur l'importance de la bonne prise en compte des "énoncés" mutuels des deux disciplines – mathématiques et écologie, dans le cadre d'une "bidisciplinarité active". De manière générale, les écologues recherchent parfois un peu trop de "sens", de contenu dans les mathématiques, alors que nous partageons le point de vue de Poincaré (1905) selon lequel les mathématiques permettent avant tout de fournir aux autres disciplines des "cadres vides" extrêmement utiles, mais dont le contenu est, dans notre cas, biologique ou écologique.

Dans la seconde partie du mémoire, nous réfléchissons aux conditions d'une meilleure modélisation statistique des pressions pesant sur des parties de la biodiversité. Nous précisons pourquoi les modèles statistiques nous semblent importants d'un point de vue appliqué. Nous y défendons une approche paramétrique et Bayésienne de la modélisation, mettant davantage l'accent sur l'estimation des effets et sur le test d'hypothèses intervalles plutôt que sur les classiques tests d'hypothèses nulles "ponctuelles". Néanmoins, les outils à la disposition de l'écologue pour développer des modèles statistiques Bayésiens sont pour partie encore limités. Outre les problèmes numériques, le développement d'outils d'évaluation de la qualité probabiliste des modèles – un des outils principaux du détective écologique selon Hilborn & Mangel (1997) – n'est pas encore stabilisé. Nous présentons quelques résultats qui permettent d'améliorer ces outils, et proposons des familles de distributions de probabilité pour données de comptage plus souples que celles utilisées en écologie aujourd'hui.

Références citées:

Hilborn R. et Mangel M., 1997, *The ecological detective: confronting models with data*, Princeton University Press, xvii+315 p.

Poincaré H., 1905, *La valeur de la science*, Paris, Flammarion.