THESE DE DOCTORAT DE
L'UNIVERSITE PIERRE ET MARIE CURIE

Spécialité
Ecologie Marine

Présentée par
Christophe Loots

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Contrôles de la distribution spatiale de l'habitat de reproduction chez les populations de poissons

Approche multi-modèles appliquée à la plie et au merlan de Mer du Nord et à l'anchois du Golfe de Gascogne

soutenue le 18 novembre 2009 devant le jury composé de :

Philippe KOUBBI	Directeur de thèse
Guy DUHAMEL	Rapporteur
Anthony LEHMANN	Rapporteur
Paul NIVAL	Examinateur
André CARPENTIER	Examinateur
Sandrine VAZ	Responsable scientifique
Benjamin PLANQUE	Responsable scientifique

Résumé

La reproduction est une phase importante du cycle de vie des poissons car elle assure le renouvellement du nombre d'individus dans la population. La détermination des facteurs qui influencent sa dynamique spatiale est donc un point capital car elle affecte le succès de la reproduction.

Une approche multi-modèles a été développée afin d'identifier les facteurs contrôlant la distribution spatiale des adultes reproducteurs pour la plie et le merlan de Mer du Nord et pour l'anchois du Golfe de Gascogne. Les modèles additifs généralisés ont été utilisés pour relier la présence-absence des individus et leurs abondances aux différentes hypothèses de contrôle externes (position géographique et environnement) et internes à la population (dépendance spatiale, taille, structure en âge et mémoire). Les modèles construits ont été classés en fonction de leur aptitude à refléter les distributions observées. Les hypothèses de contrôle importantes ont alors été identifiées à partir des meilleurs modèles.

L'habitat de reproduction montre une structuration très importante dans l'espace. L'approche multi-modèles révèle (1) l'influence de l'environnement sur l'occupation des zones de reproduction et (2) le degré d'abondance avec lequel ces zones sont occupées est déterminé par des facteurs internes à la population. La persistance dans le temps de la structuration spatiale de la reproduction est facilitée par un attachement des individus à leur site de reproduction. Celui ci est plus fort pour des bentho-démersaux comme la plie et le merlan que pour un pélagique comme l'anchois.

Mots clés : reproduction, habitat, hypothèses de contrôle, approche multi-modèles, plie, merlan, anchois

Abstract

Control of the spatial distribution of the spawning habitat of fish populations.

A multi-model approach applied to North Sea plaice and whiting
and Bay of Biscay anchovy

Spawning is a crucial phase of fish life cycle as it ensures the renewal of individuals in the population. The identification of factors that influence its spatial dynamic is therefore a key feature as it will affect the reproduction success of fish population.

A multi-model approach has been developed in order to identify factors that control the spatial distribution of spawning adults of plaice and whiting in the North Sea and of anchovy in the Bay of Biscay. Generalised additive models have been used to relate presence-absence of adults and their abundances to several hypotheses of control external (geographical position and environment) and internal to the population (spatial dependency, size, age structure and memory). Models were then ranked according to their ability to reflect the observed distribution. Relevant hypotheses of control were then identified from the best models.

The spawning habitat depicts a strong spatial structure. The multi-model approach highlights that (1) environmental factors influence the occupation of spawning areas and (2) population internal factors determine the degree of abundance by which spawning areas are occupied. Temporal persistence of spawning distribution is ensured by attachment of adults to their spawning site. This attachment is stronger for bentho-demersal fish like plaice and whiting than for pelagic fish like anchovy.

Keywords : spawning, habitat, hypotheses of control, multi-model approach, plaice, whiting, anchovy

Remerciements

Ce travail de thèse vient clore une aventure qui même si elle n’a pas commencé « il y a bien longtemps, dans une galaxie lointaine, très lointaine…», remonte déjà à bientôt dix ans. A cette époque, un certain Philippe Koubbi accepte de me prendre en stage d’été pour me faire découvrir le monde de l’ichtyologie sous la coupe d’Alain Grioche. Notre collaboration se poursuit au sein de son LIMUL (Laboratoire d’Ichtyoécologie marine de l’Université du Littoral), grâce à un stage en maîtrise sur le zooplancton arctique sous la direction de Elvire Antajan et Jean-Michel Brylinski, puis en DEA sous la direction conjointe avec Guy Duhamel. Mon hivernage en Antarctique lui laisse alors le temps de trouver un financement et de monter un sujet de thèse pendant que de mon coté je contribue à son programme de recherche polaire ICOTA (Ichtyologie Côtière en Terre Adélie). De retour d’hivernage, les portes de l’Ifremer de Boulogne s’ouvrent à moi pour trois ans et c’est là qu’une nouvelle page commence…

Mes premiers remerciements vont tout d’abord aux trois personnes qui m’ont encadré durant ces trois années, et en particulier à Sandrine Vaz et Benjamin Planque, pour m’avoir confié sans me connaître le sujet de thèse qu’ils avaient écrit.

Philippe, merci de m’avoir offert la possibilité de découvrir et de vivre pendant 14 mois dans ce monde merveilleux qu’est l’Antarctique. Merci également d’avoir accepté d’encadrer ce travail de thèse et de l’avoir si bien fait pendant ces trois années. C’est grâce à tes qualités humaines indéniables ainsi qu’à tes remarques, parfois épuisantes, mais toujours percutantes et justifiées que j’ai pu mener à bien ce travail. Merci pour ta disponibilité à toute épreuve, surtout dans les moments difficiles, qui furent rares, mais où il fait toujours bon de se sentir soutenu. Merci également d’avoir continué à m’encadrer comme tu l’as fait après ton départ pour le grand sud et d’avoir fait l’effort de remonter me voir à chaque fois que cela a été nécessaire. J’espère que tu trouveras dans ces quelques pages la récompense de l’aide, du soutien et des encouragements que tu m’as apportés à chaque instant durant toutes ces années. J’espère également que l’avenir nous donnera l’occasion de continuer à travailler ensemble et nous permettra d’entretenir cette relation d’amitié qui s’est développée au fils des années.

Sandrine, merci à toi de m’avoir encadré et supporté au quotidien durant ces trois années. J’espère qu’elles ne t’auront pas trop démoraliser à superviser de futurs étudiants en thèse. Merci pour ta disponibilité et le pragmatisme qui te caractérisent et qui ont fait que la thèse s’est déroulée dans les meilleures conditions possibles. Merci de m’avoir transmis une partie de ton immense savoir-faire sur les analyses multivariées et géostatistiques. Je n’oublie pas non plus nos discussions toujours sérieuses mais souvent teintées d’humour que nous avons eues quant aux méthodes utilisées et aux résultats obtenus. J’espère que ce travail de thèse a répondu à tes attentes vis à vis du sujet et que nous aurons l’occasion de travailler de nouveau ensemble, qui sait peut être grâce à un éventuel post-doc.

Merci à toi Benjamin d’avoir accepté de confier à un inculte en statistique comme moi la lourde responsabilité de développer un tel sujet. Merci pour ton aide, ton enthousiasme, ton flot d’idées, tes remarques et ta bonne humeur qui ont contribué à la réussite de cette thèse. Merci également d’avoir accepté de continuer à me superviser après ton départ pour le grand nord norvégien. La distance n’a en rien empêché les longues conversations et débats téléphoniques que nous avons eus à propos de l’autocorrélation spatiale, de la manière de combiner les modèles et bien d’autre chose encore. J’espère que ce travail t’auras permis d’avancer dans ta propre recherche même si je sais qu’il nous reste encore beaucoup à faire.

Je tiens ensuite à remercier les personnes qui ont accepté de lire et de critiquer ce travail de thèse.

Merci à messieurs Guy Duhamel, professeur au Muséum National d’Histoire Naturelle de Paris et Anthony Lehmann, maître d’enseignement et de recherche à l’université de Genève, d’avoir accepté d’être rapporteurs de ce travail de thèse. J’espère qu’ils auront eu plaisir à lire ce manuscrit et qu’ils y auront trouvé quelques idées qui pourront leur être utiles. Merci également à Paul nival, professeur émérite à l’université Pierre et Marie Curie d’avoir accepté d’examiner ce travail. J’espère qu’il trouvera là le fruit des connaissances, du savoir faire et du perfectionnisme qu’il m’a transmis lorsque j’étais son étudiant dans l’option paramétrisation et modélisation du DEA Océanologie Biologique et Environnement Marin de Paris VI. Merci à André Carpentier, chef du département halieutique du centre Ifremer Manche-Mer du Nord d’avoir également accepté le rôle d’examinateur. Merci à toi, André, de m’avoir accueilli au sein de ton équipe, tout d’abord en ta qualité de chef du laboratoire Ressources Halieutiques de Boulogne-sur-Mer puis par la suite en tant que chef du département halieutique. Merci pour la motivation dont tu fais preuve, que tu sais communiquer aux personnes qui t’entourent et qui fait que cela a toujours été un plaisir de discuter avec toi le matin en arrivant ou le soir avant de partir.

Merci à Olivier Le Pape et Pierre Petitgas d’avoir accepté de faire partie de mon comité de thèse. Je les remercie pour leurs conseils avisés et leurs critiques constructives qui m’ont fait réfléchir et permis d’avancer. C’est aussi grâce à eux que ce travail est ce qu’il est aujourd’hui. Merci plus particulièrement à toi Pierre, de m’avoir prêté une partie de tes données sur l’anchois et de m’avoir permis de jouer avec. Merci également de m’avoir intégré au projet RECLAIM et de m’avoir prêté de l’argent pour mes missions à chaque fois que je te l’ai demandé. J’en profite également pour remercier Martin, Huret, Pierre Beillois et Jacques Massé avec qui j’ai eu l’occasion de collaborer et de discuter lors de mes quelques déplacement à l’Ifremer de Nantes.

Je tiens également à remercier Marc Morel et Dominique Godefroy, tous deux directeurs du centre Manche-Mer du Nord de Boulogne pour m’avoir accueilli et permis de réaliser ma thèse au sein de leur locaux. Merci également à Jean-Paul Delpech, actuel chef du laboratoire Ressources Halieutiques, de m’avoir offert les conditions que tout étudiant rêverait d’avoir pour terminer sa thèse.

Je remercie également les deux financeurs de la thèse, d’une part la région Nord Pas de Calais et d’autre part l’Europe à travers le projet Européen Reclaim, sans qui ce travail de thèse n’aurait jamais pu avoir lieu.

Merci aux amis pour la sympathie et le soutien dont ils ont fait preuve à mon égard durant ces années. Merci à toi Elvire pour m’avoir transmis le goût de la recherche et la passion pour les copépodes. Merci à Caro, Emilie, Didier, Dorothée, Carole et Eric pour leur amitié. Merci également à toutes les personnes de l’Ifremer de Boulogne qu’il m’a été donné de rencontrer et avec qui j’ai eu le plaisir de travailler. Merci à Yves pour les nounours du jeudi, à Caro pour les sucettes du mardi et à Jean Paul pour les carambars.

Merci également aux personnes avec qui j’ai été amené à partager un bureau. J’espère ne pas avoir été trop insupportable. Merci à Youen et Emilie pour leur bonne humeur lorsque nous partagions la bibliothèque lors de ma première année de thèse. Un merci tout particulier à mes collègues du bassin d’essai, Grégory, Benoit, Jean-Valéry et Christèle, qui m’ont ouvert leur porte lors de ma deuxième année et qui j'espère garderont un bon souvenir de moi. Enfin, merci à Caro, Stéphanie, Juliette et Djiby d'avoir fait l'effort de rigoler à mes blagues lorsque nous partagions la même pièce durant ma dernière année.

Je voudrais remercier mes parents, sans qui je n'en serai pas là aujourd'hui. Qu'ils trouvent dans ce manuscrit le témoignage de la reconnaissance et de l'amour que je leur porte. Merci d'avoir respecté mon choix de vouloir partir pendant 14 mois en Antarctique et de ce fait de m'être éloigné de vous de plusieurs milliers de kilomètres. Merci de m'avoir soutenu, parfois de m'avoir 'supporté' dans les moments difficiles, mais surtout d'avoir toujours cru en moi durant toutes ces années. Merci également de m'avoir inculqué les valeurs du courage, du travail, de la persévérance, de la curiosité et de l'envie de découvrir et d'apprendre. Ce sont ces valeurs qui font de moi ce que je suis aujourd'hui et qui m'ont permis de mener à bien ce projet de thèse qui me tenait tant à cœur. J'espère qu'à présent j'aurai plus de temps à vous consacrer et que nous pourrons rattraper le temps perdu. Je vous aime fort.

Enfin, je voudrais remercier une personne chère à mes yeux qu'il m'a été donné de rencontrer au début de ma thèse. Elodie, merci à toi pour avoir débarqué dans ma vie. Tu as pris le temps de me connaître et tu as accepté de me supporter quotidiennement dans les circonstances exceptionnelles d'une thèse où règnent souvent la mauvaise humeur, la fatigue, le stress et l'indisponibilité. Une relation qui naît et grandit dans de telles conditions ne peut que prospérer dans les conditions normales de la vie de tous les jours. Il ne tient qu'à nous maintenant de faire qu'il en soit ainsi. Merci de m'aimer comme tu m'aimes, merci d'être là, merci d'être toi.

Que la force soit avec vous!

Pour Pierre, Jeanne et Pascal, qui m’ont quitté trop tôt.
J’espère que vous me regardez et que vous êtes fiers de moi.
Je ne vous oublierai jamais.

Tables des matières

Liste des figures

Liste des tables

Liste des abréviations

AIC : Akaike Information Criterion

DDHS : Density Dependent Habitat Selection

GAMs : Generalised Additive Models

GIS : Geographic Information System

IBM : Individual Based Modelling

IBTS : International Bottom Trawl Surveys

ICES : International Council for the Exploration of the Sea

IFD : Ideal Free Distribution

LRT : Likelihood Ratio Test

NHST : Null Hypothesis Statistical Testing

PCA : Principal Component Analysis

PCNM : Principal Coordinates of Neighbour Matrices

PELGAS : Pelagique Gascogne

RDA : Redundancy Analysis

RMSE : Root Mean Squared Error

PDMs : Population Distribution Models

SDMs : Species Distribution Models

SSB : Spawning Stock Biomass

“…there can never be a correct model. There may be a best model…”

The Ecological Detective. Confronting models with data.

Hilborn et Mangel, 1997.

Introduction

Les populations de poissons ne sont pas réparties de manière aléatoire mais montrent des distributions qui sont organisées dans l'espace et dans le temps. La connaissance des mécanismes qui régissent ces distributions est donc importante d'un point de vue fondamental et appliqué car cela permet une meilleure utilisation, gestion et conservation lorsque les espèces étudiées sont des ressources exploitées. Pourquoi les poissons se trouvent là où ils sont? ; pourquoi se reproduisent-ils parfois dans certaines zones et vont se nourrir ailleurs? ; pourquoi font-ils parfois de longues migrations pour revenir à leur lieu de reproduction? ; sont autant de questions qui ont motivé le travail de thèse présenté ici.

Organisation spatiale du cycle de vie chez les poissons

Le cycle de vie d'un organisme est l'ensemble des phases nécessaires à son ontogénie, c'est à dire à son développement de sa naissance jusqu'à sa mort. Chez les poissons, on distingue trois phases principales ayant des durées différentes. La phase larvaire qui est généralement la plus courte (de quelques jours à quelques mois); elle est représentée par des stades planctoniques ou ichthyoplancton (œufs, larves et post-larves) dont la morphologie est très différente des adultes. La phase juvénile dont la durée (de quelques mois à quelques années) varie en fonction de la durée de vie des espèces, est représentée par les individus immatures sexuellement. La phase adulte est la phase la plus longue, de quelques années pour les petits pélagiques à plusieurs dizaines d'années pour les espèces benthiques, et est représentée par les individus capables de se reproduire (matures).

Le cycle de vie des poissons possède souvent une structuration spatiale des différentes phases qui est très importante. Les zones où sont rencontrées les adultes sont distinctes de celles où se trouvent les juvéniles et de celles occupées par les œufs et les larves. Le déplacement entre les différentes zones est assuré par des migrations actives ou passives de la part des individus. Ce schéma de distribution et de migrations des différents stades dans l'espace est décrit dans le 'triangle de migration' de Harden Jones (1968) qui a été repris et complété plus tard (fig. 1, Gibson, 1997). La reproduction a lieu au niveau des frayères, généralement situées au large, et où se rencontrent les adultes. Les œufs (lorsque ceux-ci sont pélagiques) et les larves dérivent alors suivant l'hydrodynamisme de la zone jusqu'aux nourriceries où les juvéniles vont pouvoir se développer. Lorsque les juvéniles deviennent capables de se reproduire, ils entrent dans la population adulte à travers le processus de recrutement. Une fois la reproduction achevée, les adultes regagnent leur zone de nutrition où ils vont emmagasiner l'énergie nécessaire pour se reproduire l'année suivante.

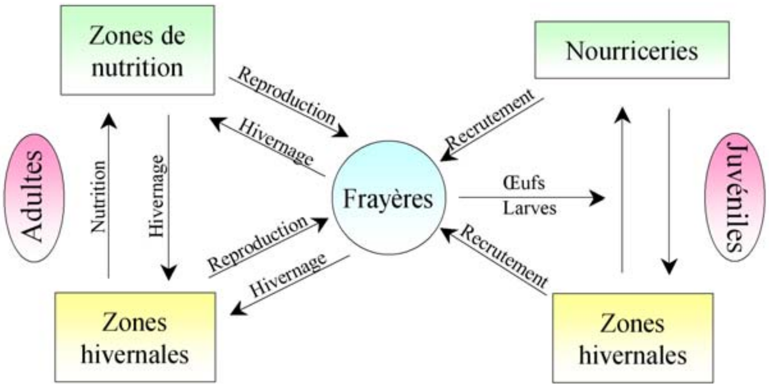


Figure 1. Cycle de vie et schéma de migration des poissons (d'après Gibson, 1997). Ce schéma reprend le triangle de migrations de Harden Jones (1968) et le complète en tenant compte des variations possibles qui ont été observées pour différentes espèces.

La réussite de chaque phase est contrôlée par un ensemble de facteurs différents qui interagissent entre eux. Une zone qui cumulera un nombre important de fonctions (reproduction, croissance, nutrition…) devra regrouper un nombre de caractéristiques favorables qui sera d'autant plus élevé. Elle représentera alors un compromis global entre plusieurs fonctions sans pour autant être adaptée de manière optimale à chacune d'entre elle (Boulinier *et al.*, 2005). Dans ce contexte, la ségrégation spatiale du cycle de vie des poissons est donc avantageuse car elle évite le cumul des fonctions au niveau d'une même zone et permet ainsi une occupation optimale de chaque zone pour la fonction pour laquelle elle est dédiée. Les mécanismes qui déterminent l'accomplissement du cycle de vie chez les poissons sont repris dans la notion d'habitats essentiels comme l'ensemble des facteurs favorables à la survie, la croissance et à la reproduction d'une espèce (Benaka, 1999). L'identification de ces différents facteurs est donc importante si l'on veut comprendre ce qui détermine la distribution spatiale des poissons, depuis les différents stades de vie pris individuellement, jusqu'au cycle de vie dans sa globalité.

Dynamique spatiale de la reproduction

La reproduction est une phase cruciale du cycle de vie car elle permet d'assurer la survie d'une espèce grâce à l'apport de nouveaux individus. Cet apport permet le renouvellement et le maintien du niveau d'abondance des populations en compensant les pertes engendrées par la mortalité naturelle subie par les individus. D'un point de vue spatial, les frayères occupent une place importante car elles tiennent lieu de carrefour par lequel les individus des différents stades seront amenés à passer (fig. 1). Bellier *et al.* (2007a) émet l'hypothèse que la configuration de la distribution spatiale de la reproduction va affecter le recrutement qui en retour va jouer sur la distribution de la reproduction en modifiant la démographie de la population. Le succès du recrutement dépend en partie du taux de survie des œufs et des larves qui lui même est influencé par les conditions environnementales rencontrées (Borja *et al.*, 1998; Borja *et al.*, 2008). La localisation des frayères et la connaissance des facteurs qui régissent leur distribution est donc un point essentiel si l'on veut comprendre la dynamique spatiale et temporelle des populations de poissons.

L'étude des frayères se fait classiquement à l'aide des œufs (ICES, 2005; Martin *et al.*, 2007; van der Molen *et al.*, 2007) dont la distribution spatiale permet de localiser les habitats récurrent, occasionnel et non favorable (Bellier *et al.*, 2007a). L'habitat récurrent correspond aux zones où la ponte a lieu chaque année. L'habitat optionnel correspond aux zones où la ponte n'a lieu que certaines années. L'habitat non favorable est l'ensemble des zones qui sont non favorables à la ponte. Les facteurs qui contraignent la distribution spatiale des œufs sont déterminés par la modélisation de l'habitat de ponte qui vise à mettre en relation des distributions observées avec des facteurs physiques de l'environnement. Trois types d'habitats de ponte sont alors définis: l'habitat potentiel, l'habitat réalisé et l'habitat efficace (fig. 2).

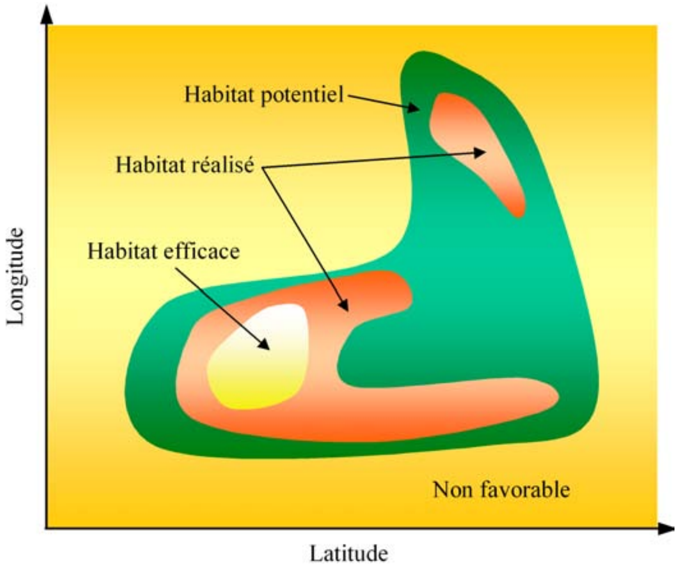


Figure 2. Représentation schématique des trois types d'habitats de ponte: potentiel, réalisé et efficace. L'habitat potentiel correspond aux zones favorables à la ponte. L'habitat réalisé correspond aux zones de l'habitat potentiel où la ponte a effectivement lieu à un moment donné. L'habitat efficace correspond aux zones de l'habitat réalisé qui ont donné lieu à un recrutement positif pour la population. A l'extérieur, les zones ne sont pas favorables à la ponte. D'après Planque *et al.* (2007).

Cependant, l'environnement physique ne peut expliquer à lui seul les distributions observées. L'occupation de l'habitat de reproduction résulte de l'influence d'un nombre important de contraintes qui interviennent tout au long du processus de sélection de l'habitat (fig. 3). Ces contraintes sont de natures différentes et sont liées soit à l'environnement externe des individus, soit aux individus eux même, qu'ils appartiennent à la population étudiée ou à d'autres populations. Ces aspects sont repris dans les notions d'habitat potentiel et réalisé. Ainsi, l'habitat potentiel est l'ensemble des zones regroupant les conditions environnementales qui sont favorables à la ponte tandis que l'habitat réalisé représente la partie de l'habitat potentiel qui est effectivement utilisé par les individus compte tenu des caractéristiques démographiques de la population (taille, structure en taille, structure en âge…).

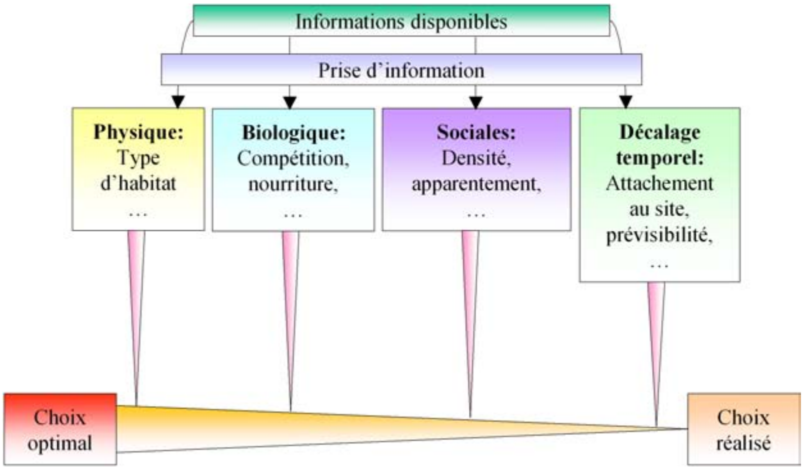


Figure 3. Schéma du processus de sélection de l'habitat. Il retrace l'influence conjointe des différents facteurs et montre le lien entre le choix optimal pour les individus et le choix qu'ils effectuent réellement compte tenu des différentes contraintes qui agissent lors de la sélection. D'après Boulinier *et al.* (2005).

La reproduction montre des variations dans sa réalisation selon plusieurs échelles spatio-temporelles (Froese et Pauly, 2009). A l'échelle de l'espèce, on distingue ainsi les espèces sémelpares, qui ne se reproduisent qu'une fois (p. ex. l'anguille, *Anguilla anguilla*), des espèces itéropares, qui sont capables de se reproduire plusieurs fois au cours de leur vie. Certaines ont une période de reproduction restreinte tandis que d'autres ont une période de reproduction beaucoup plus étalée dans le temps. A une échelle spatiale moins large, les différentes populations d'une même espèce peuvent se reproduire à des périodes différentes compte tenu de leur localisation géographique (p. ex. selon un gradient latitudinal). Enfin, les individus d'une même population peuvent se reproduire une ou plusieurs fois au cours de la période de reproduction. La capacité à se reproduire varie selon l'âge et aura lieu d'autant plus tôt dans la vie des individus que ces derniers auront une vie courte. Par exemple, pour des espèces de petits pélagiques comme l'anchois (*Engraulis sp.*), la sardine (*Sardina sp.*) et la sardinelle (*Sardinella sp.*), la reproduction a lieu dès la première année. Pour des espèces démersales comme le merlan (*Merlangius merlangus*) et le rouget barbet (*Mullus surmuletus*), la reproduction a lieu dès la deuxième année. Pour des espèces comme la morue (*Gadus* *morhua*), la plie (*Pleuronectes platessa*) et la sole (*Solea solea*), la reproduction a lieu au bout de quelques années. En revanche, pour des espèces longévives comme le sébaste (*Sebastes mentella*), certaines espèces de poissons plats comme le flétan noir (*Hippoglossus hippoglossus*) et les espèces de fonds comme le grenadier (*Coryphanoides rupestris*), les individus ne sont matures et capables de se reproduire que très tardivement (à partir de douze ans pour le sébaste).

Quels facteurs déterminent l'habitat de reproduction?

Dans le cadre de la thèse, il s'agit d'étudier l'habitat de reproduction des poissons afin de caractériser et identifier les facteurs qui en contrôlent la distribution spatiale. Nous préférons le terme 'd'habitat de reproduction' car c'est la distribution des adultes prêts à se reproduire qui a été utilisée et non pas la distribution des oeufs. La méthodologie illustrée en fig. 4 a été utilisée comme fil conducteur tout au long de la thèse.

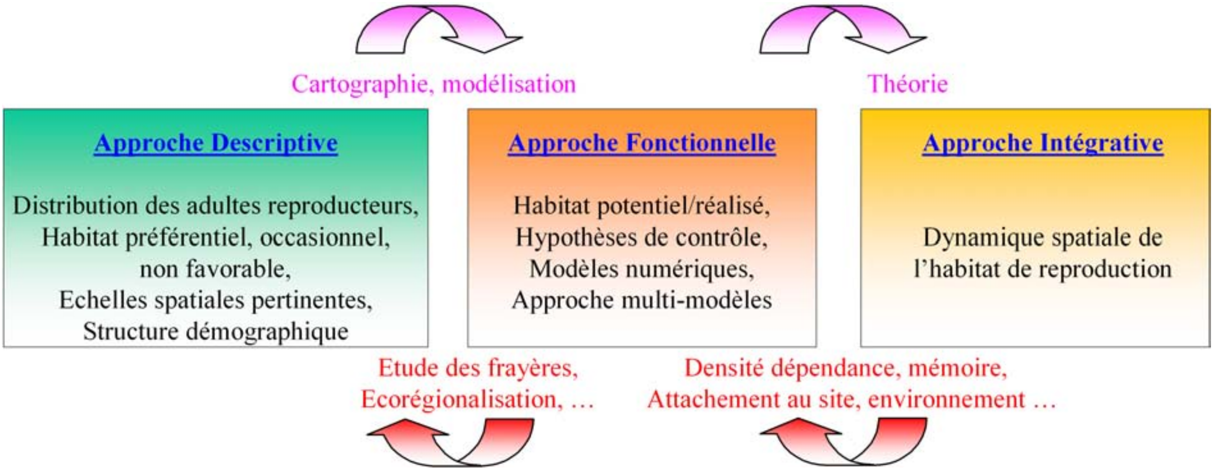


Figure 4. Schéma méthodologique de l'approche utilisée pour décrire l'habitat de reproduction des populations de poissons. La distribution des adultes reproducteurs est cartographiée afin de localiser les habitats préférentiels, optionnels et non favorables. Ici, le terme 'préférentiel' est préféré au terme 'récurrent' car ce sont des abondances qui sont cartographiées. Ces abondances sont ensuite modélisées selon diverses hypothèses de contrôle afin d'étudier les habitat potentiel et réalisé. Enfin, les résultats obtenus par les deux approches sont discutés d'un point de vue théorique dans le contexte de l'habitat de reproduction.

Elle est structurée en trois approches complémentaires: une approche descriptive, une approche fonctionnelle et une approche intégrative. Dans l'approche descriptive, la distribution spatiale de l'habitat de reproduction est décrite grâce à la cartographie des distributions observées qui sont ensuite utilisées dans des analyses multivariées pour en explorer la variabilité spatio-temporelle. Dans l'approche fonctionnelle, des modèles numériques sont développés afin de simuler les distributions spatiales observées empiriquement et déterminer le rôle des différentes contraintes qui peuvent s'exercer sur les individus au moment de se reproduire. L'approche intégrative vise à synthétiser les résultats obtenus dans les approches descriptive et fonctionnelle afin de définir et caractériser d'un point de vue théorique les mécanismes qui contrôlent l'habitat de reproduction.

Trois populations cibles ont été choisies comme cas d'étude pour étudier la distribution spatiale de l'habitat de reproduction: la plie et le merlan de mer du Nord et l'anchois du Golfe de Gascogne. Ces trois espèces possèdent des traits de vie et des modalités de reproduction différents. La plie de mer du Nord (*P. platessa*) est un poisson plat benthique qui se reproduit en hiver. Le merlan de mer du Nord (*M. merlangus*) est une espèce démersale qui se reproduit à la fin de l'hiver et au début du printemps. L'anchois du Golfe de Gascogne (*Engraulis encrasicolus*) est un petit pélagique qui se reproduit à la fin du printemps et au début de l'été pendant le pic de production planctonique. La plie et le merlan sont des espèces à vie longue (plusieurs dizaines d'années) ne se reproduisant qu'une fois par an (itéropares annuelles) sur une période plus ou moins étalée, tandis que l'anchois est une espèce à vie courte (3-4 ans) qui se reproduit plusieurs fois en une saison. Ces trois espèces ont en commun l'émission d'œufs qui sont pélagiques et qui sont donc fortement soumis aux caractéristiques hydrologiques des masses d'eaux dans lesquelles ils sont émis et vont dériver.

Plan de la thèse

Le manuscrit se présente comme une thèse rédigée d'après des articles qui ont été soumis ou acceptés pendant la période de doctorat dans des revues internationales avec comité de lecture. Les articles ne sont pas intégrés directement mais les différentes parties ont été replacées au sein de chaque chapitre de manière à rendre le document cohérent et éviter les répétitions quant aux données et méthodes utilisées. Cinq articles ont été rédigés au total; trois sont soumis, un est sous presse et un est en préparation:

(1) B. Planque, C. Loots, P. Petitgas, U. Lindstrom, S. Vaz
Modelling the spatial distribution of fish populations using a statistical multi-model approach. *Fisheries Oceanography*. Soumis.

(2) C. Loots, S. Vaz, B. Planque, P. Koubbi
Spatial distribution of spawning population of
North Sea plaice and whiting from 1980 to 2007.
En préparation.

(3) C. Loots, S. Vaz, P. Koubbi, B. Planque, F. Coppin, Y. Verin
Inter-annual variability of North Sea plaice potential spawning habitat.
Journal of Sea Research. Soumis.

(4) C. Loots**,** S. Vaz, B. Planque, P. Koubbi
What control the spatial distribution of North Sea plaice spawning population.
Confronting ecological hypotheses through a model selection framework.
ICES Journal of Marine Science. Sous presse.

(5) C. Loots, B. Planque, S. Vaz, P. Koubbi
Identifying processes that control the spawning distribution of
North Sea whiting using multi model inference.
Fisheries Oceanography. Soumis.

Le chapitre 1 correspond à l'article (1) et se présente sous la forme d'une revue bibliographique. Il définit les concepts et modèles numériques relatifs à la distribution spatiale des populations de poissons. Tout d'abord, les grandes hypothèses de contrôle sont décrites.

Puis, les modèles numériques qui permettent de modéliser ces hypothèses sont présentés. Enfin, les différentes approches qui permettent de comparer les modèles construits et d'en déduire l'importance des différentes hypothèses de contrôle sont décrites.

Le chapitre 2 présente les données et les méthodes qui ont été utilisées au cours du travail de thèse. Il est basé principalement sur les articles (2), (3) et (4). Les deux premières parties sont consacrées à la présentation des zones d'étude, des campagnes d'observations scientifiques et des données qui en sont issues et qui ont été utilisées pour représenter la distribution spatiale de la reproduction pour les trois espèces cibles. La troisième partie présente les méthodes utilisées pour l'approche descriptive. La quatrième partie présente la méthodologie basée sur le chapitre 1 et qui a été utilisée et appliquée au cas de la distribution de la reproduction au cours de l'approche fonctionnelle.

Les chapitre 3, 4 et 5 sont l'application des approches descriptive et fonctionnelle aux trois cas d'études, respectivement la plie (chapitre 3), le merlan (chapitre 4) et l'anchois (chapitre 5). Ils rappellent succinctement les principaux traits de vie et les modalités de reproduction adoptées par chaque espèce, décrivent les résultats obtenus par chaque approche et discutent de ces résultats par rapport aux connaissances actuelles concernant chaque espèce. Le chapitre 3 est basé sur les articles (2), (3) et (4) et le chapitre 4 est basé sur les articles (2) et (5).

Le chapitre 6 correspond à l'approche intégrative et fait la synthèse des résultats obtenus pour les trois espèces. Il discute dans un premier temps des apports et des limites de l'approche de modélisation utilisée pour déterminer le rôle des différentes hypothèses de contrôle. Il replace ensuite les résultats écologiques obtenus pour les trois espèces dans le contexte plus général et théorique de l'habitat de reproduction. Enfin, les résultats obtenus permettent d'envisager et de discuter des approches qui pourraient être menées afin de venir compléter les connaissances acquises au cours de ce travail.

CHAPITRE 1

Control of the spatial distribution of fish population: theoretical concepts and numerical models

---

This chapter is based on a submitted article, from which only the concepts related to the present work are presented here.

B. Planque, C. Loots, P. Petitgas, U. Lindstrom, S. Vaz
Modelling the spatial distribution of fish populations using a statistical multi-model approach. *Fisheries Oceanography*. Submitted.

1.1. Introduction

Animal and plant populations are not distributed randomly in space but exhibit spatial patterns, and fish populations are no exception. Understanding what drives these patterns of spatial distribution is both a fundamental ecological question and a requirement for sound management when fish species are of commercial or conservation interest. The various controls on fish spatial distributions can be both external and internal to the populations. External controls, such as environmental forcing, are mainly known to affect the suitability of fish habitats. On the other hand, internal controls such as population size, age structure, fish condition, diversity and behaviour, can also modulate the spatial distribution of fish population, through mechanisms such as density dependence, age or stage dependent habitat preference, differential migration capacities, and so on.

In recent decades, a large body of literature has focused on species distribution models (SDMs, see Rushton *et al.*, 2004 for a review). SDMs are mostly concerned with the control of species spatial distribution by the environment. Such approach finds deep theoretical roots in the niche theory as defined by Hutchinson (1957). The theory, which is defined at the species level, is attractive as it can be easily translated into a wide range of statistical models that associate environmental conditions with species presence or abundance. Environment based models are often referred to as habitat models (Guisan and Zimmermann, 2000), species distribution models (Guisan and Thuillier, 2005) or habitat suitability models (Hirzel *et al.*, 2006) and their aim is to relate observations of animal presence or abundance to various attributes of the environment using a statistical method (Araujo and Guisan, 2006; Austin, 2002; Austin, 2007; Boisclair, 2001; Guisan *et al.*, 2006; MacNally, 2000; Randin *et al.*, 2006; Rushton *et al.*, 2004). In essence, SDMs are restricted to modelling species response to the environment but environmental forcing alone cannot be expected to fully explain the spatial distribution of fish populations, unless such forcing is so strong that it over-rides all other drivers. Other mechanisms such as site fidelity, learning behaviour, life history changes, competition (intra- and inter specific), predator avoidance, spawning behaviour, and exploitation are expected to affect the spatial distribution of fish populations, possibly to a large extent. Ignoring these mechanisms can lead to incomplete or erroneous representations of the processes controlling fish distribution. These processes are complex, interacting, and evolving as populations constantly adapt to new ecological conditions. For this reasons SDMs are often (but not always) used to model the ‘potential’ spatial distribution of species, that is the distribution resulting from environmental constraints only, which is the geographical realisation of the fundamental niche. Because they are based on niche theory, these models are primarily applicable to species rather than populations. They constitute very useful and powerful tools to model and predict the spatial distribution of species’ potential habitats, but are obviously limited in modelling and predicting actual spatial distribution of animals at the population level.

Modelling the spatial distribution of fish populations can be achieved by a group of models that can incorporate many hypotheses on the control of fish spatial distribution. We term these population distribution models (PDMs). Like SDMs, PDMs are constructed using statistical models, but in addition to SDMs, they can incorporate explicit modelling of non-environmental control processes such as site fidelity, spatial dependency, density dependence, species interactions or population memory. Such models may be advantageous when the information necessary to adequately model the movement of fish is poorly known (which is often the case, e.g. vertical and horizontal swimming behaviour), but when information on current and past spatial distribution of the population of interest, spatial distribution of other species and environmental conditions are known. Population distribution models are specifically designed at the population level, rather than species or individual level.

The intention here, is to revisit the ecological rational behind the conceptual models of spatial distribution of fish populations and to examine and discuss the corresponding numerical applications that have been developed for fish populations. The following conceptual models for explaining fish spatial distribution are explored: (i) control through geographical attachment, (ii) environmental conditions, (iii) density-dependent habitat selection, (iv) spatial dependency, (v) population demographic structure and (vi) population memory. Numerical applications exist, to variable degrees of development, for each of these hypotheses. We present some of the most common numerical applications, with a focus on models that are amenable to statistical inference and prediction. When developing such spatial distribution models, the issue of scale needs to be carefully considered (e.g. Levin, 1992). Here, we assume that investigation at the mesoscale is relevant and encompasses the extent of population distribution and that the resolution is sufficient to discriminate between areas where important differences in mean abundance or environmental conditions do occur. We argue that a common modelling framework in which numerical models associated with each hypothesis can be constructed will provide the basis to evaluate, compare and combine models in an objective manner. Under such conditions, it will be possible to test and challenge the ecological hypotheses and to progress in the understanding of what controls the spatial distribution of fish populations. This will also provide tools directly applicable for forecasting fish population distribution under a wide range of ecological scenarios.

1.2. The multiple controls of population spatial distribution

Below, we review seven hypotheses of control that are commonly considered for explaining fish population spatial distribution and which include both internal and external controls.

1.2.1. Absence of control

The absence of control, neither internal nor external is the simplest hypothesis concerning the spatial distribution of fish populations. In such case, the distribution is expected to be unstructured. At the mesoscale, this situation is never observed in the real world, but we explicitly state this hypothesis here because it is implicitly used in standard statistical tests. When population distribution models are tested using standard statistical inference, it is often this unrealistic situation that is taken as the null hypothesis against which other hypotheses may be tested.

1.2.2. Geographical attachment

In the hypothesis of a geographical control of fish distribution, it is assumed that the presence or abundance of fish is solely determined by the geographical location, i.e. site attachment. Contrary to the absence of control, the site attachment hypothesis can often constitute a valid null hypothesis against which other hypotheses can be tested. Physical or biological process which can lead to persistence or variations in population geographical distribution are discussed in the following sections.

1.2.3. Environmental conditions

The environment is usually thought as the major factor that controls the spatial distribution of fish populations. This constitutes an extension – in the geographical space - of the concept of ecological niche (Hutchinson, 1957). The fundamental niche is defined as a hypervolume of *n* environmental dimensions, each of them representing a environmental factor which may constrain species to survival, growth and reproduction. Identification of the environmental factors that should be retained is one of the many challenges for ecologists when studying the environment influence on the spatial distribution of fish populations. These may be classified according to their nature, as resources, direct and indirect gradients and also according to their function, as proximal and distal gradients (Austin, 2002; Austin, 2007). Once the environmental factors have been identified, the form of the relationship that links the environment to the species needs to be specified. In Hutchinson's niche theory, the species response (either biological, physiological and biochemical) along the environmental gradient is bell-shaped. This complies with the continuum theory in which the response shape of the species is assumed to have a unimodal and symmetric form. However, the response to environmental factors does not always match this theoretical pattern, and interactions between factors may seriously complicate investigations on their respective roles. This can lead to species response taking more complicated shape, e.g. skewed or non-unimodal form (Oksanen and Minchin, 2002). In any case, the notion of habitat may be directly derived from the ecological niche concept whereby the habitat is often considered as the geographical realisation of the niche. Indeed, the ecological niche may be represented by a map on which species-environment relationship has been applied to the environmental spatial pattern observed thus resulting in a habitat map.

1.2.4. Density dependent habitat selection

In their influential work, Fretwell and Lucas (1970) expressed that the suitability of a particular part of the environment will decrease with an increase in the density of individuals occurring locally. This has formed the basis for density dependent habitat selection (DDHS) and the original development of ideal free distribution (IFD) theory. In the ideal free distribution, individuals are "ideal" i.e. they possess a complete knowledge of the suitability of the environment and they seek access to habitat with the highest suitability. They are also "free" to access every habitat without any particular cost in time, energy or mortality risk. In the ideal free distribution world, individuals distribute themselves in habitats of varying quality in such way that the resulting suitability (often expressed as individual fitness) of every occupied habitat is even. One consequence of DDHS is that populations will tend to occupy wider geographical areas as their size increases while retract to refuge areas when their size decreases, leading to a population abundance-area relationship. However, observed abundance-area relationships may not always result from true DDHS and these should be analysed with caution (Shepherd and Litvak, 2004). DDHS models have found applications in fish ecology mainly through the development of the "basin model" (McCall, 1990) and its application to the Pacific "northern anchovy" (*Engraulis mordax*).

1.2.5. Spatial dependency

Fish spatial distribution in the ocean is directly affected by intraspecific interactions (e.g. schooling or shoaling) and interspecific interactions (e.g. prey searching and predator avoidance behaviours). These interactions result in non-random spatial distributions where groups of individuals tend to aggregate in specific areas whilst being nearly absent from others. Dense fish schools, as observed for many small pelagic fish, are extreme example of aggregated spatial distribution occurring at relatively small scale (few meters to few hundred meters). Shoaling fish will often generate less dense aggregations but will still result in aggregated spatial distributions at relatively small spatial scales. The spatial dependency (i.e. existence of areas of higher and lower densities of fish) can therefore be explained, at least partially, by interactions occurring within the population or between preys and predators. However, when environmental control is spatially structured, fish distribution will also be spatially structured, even if there is no true aggregative process at play. In such situation it is difficult to assign the origin of the spatial dependency observed to endogenous processes (population biological interactions) or to exogenous processes (spatially structured environmental controls). Endogenous spatial contagion may be important to consider and it may vary greatly with time. For example in fish populations with contracted spawning season, the spatial distribution may be highly aggregative during spawning and much less during other life-stages.

1.2.6. Demographic structure

Fish populations are demographically structured, i.e. they are composed of individuals of different development stages (eggs, larvae, juveniles, adults), different sexes, different ages and different sizes. Habitat occupation by individuals may vary according to these traits. The concept of size-dependent habitat selection was formalised by Werner and Gilliam (1984). This was not done in a spatial context, but rather in a theoretical context in which several habitats with distinct properties are available to a size-structured population. Particular habitats correspond to optimal sizes of individuals, so it is advantageous for individual fishes to move from one habitat to another, as they grow larger. The principles of the size-dependent habitat selection model outlined by Werner and Gilliam (1984) is that fitness (which they measured as the ratio between natural mortality and growth) is size dependent and that individuals will inhabit the most suitable habitats according to their size. Experimental work on coral-dwelling goby show that size dependent spatial organisation resulting from inter-cohort competition can be modified by removing old/large individuals (Hobbs and Munday, 2004). Similarly, habitat modification experiments coupled with variations in large fish densities have shown that juvenile flatfishes may change habitat preference and therefore spatial distribution as a result of increased large fish densities (Ryer *et al.*, 2007). Aside from theoretical and experimental work, empirical evidence for size-dependent spatial organisation is common for marine fish. Small individuals are often found close to the coast in shallow waters whilst larger individuals are found further offshore in deeper waters (see e.g. Gordoa and Duarte, 1991; Chen *et al.*, 1997; Koubbi *et al.*, 2006). Long range migrating populations such as the Norwegian spring spawning herring display differential distribution with larger/older individuals expanding their distribution much further offshore than do smaller/younger ones (Holst *et al.,* 2004).

There is evidence of age-specific spatial distribution in fish (Swain, 1993) but strict age-dependent habitat selection seems rare and there is, to our knowledge, little empirical support for it. Instead, age-dependent spatial distribution mostly results from size-dependent habitat selection (above), because older fish tend to be larger than younger ones. Age may also play an important role in the ability of fish individuals and fish populations to memorise and reproduce particular spatial distribution patterns. This is further discussed in the section on memory.

Sex-dependent spatial distribution may be expected when a species displays sex-specific life history tactics. There is evidence of sex dependent spatial distribution patterns for a number of fish species. For example, in the St Lawrence, Swain and Morin (1996) observed that depth distribution of American Plaice (*Hippoglossoides platessoides*) tended to be more spread out for females than for males. Sex-dependent spatial distribution may also be connected with size/age-dependency for hermaphrodite species in which individuals change sex at a given size or age in their life.

In summary, the demographic structure of a given fish population will influence its spatial distribution, primarily through size distribution. Observed differences in distribution at age are most likely to result from size effects rather than from strict age effects. Sex specific spatial distributions are observed for a number of populations, and can be expected when sex-specific life-history tactics are at play or when sex determination is environmentally influenced (Han and Tzeng, 2007).

1.2.7. Memory

The spatial distribution of fish populations tends to have a degree of persistence over years and recurrent spawning, nursery or feeding grounds are common for many fish stocks. The maintenance of spatial population patterns across years and across generations is likely driven by individual memory as well as social behaviour between individuals of the same population or contingent. Fish are able to memorise and learn from other fish (Helfman and Schultz, 1984; Brown and Laland, 2003). These two fundamental properties allow newcomers to be entrained into traditional migrations and habitats by older fish (McQuinn, 1997; Corten, 2002). This lead Petitgas *et al.* (2006) to develop the entrainment hypothesis, a mechanism by which fish population can maintain persistent migratory and spatial distribution patterns over generations while allowing for innovation in particular circumstances. The screening of 11 fish populations (including herring, sardine, whiting, anchovy, bluefin tuna, white perch and striped bass), undertaken during the ICES[1] workshop WKTEST[2] (ICES, 2007b) revealed that in all cases there was some support for the entrainment hypothesis. Entrainment was demonstrated for Canadian herring (McQuinn, 1997) by showing crossover of individuals between spring- and autumn-spawning contingents. Conditions for entrainment to happen are met when generations overlap in space, length and condition. Because entrainment results in conservatism in habitat use, the occupancy of particular habitats may continue over the life time of several fish generations, even though the suitability of these habitats has become sub-optimal. Numerical dominance of naïve fish relative to experienced fish may lead to the colonisation of novel habitats (e.g., variation in wintering habitats of Norwegian Spring Spawning herring as a result of incoming strong year classes, Dragesund *et al.*, 1997; Huse *et al.*, 2002). On the contrary, stock collapse is often associated with spatial memory collapse and contingent diversity collapse (e.g. McQuinn, 1997; Corten, 2002; ICES, 2007b). The above studies suggest that spatial distributions of fish populations in the past may partially control their spatial distributions in the future, as long as the conditions that allow for population memory to exist are met (individual memory and trans-generational information sharing).

---

[1] ICES: International Council for the Exploration of the Sea

[2] WKTEST: Workshop on testing the entrainment hypothesis

1.2.8. Multiple controls

Other controls such as species interactions may also be evoked to explain the spatial distribution of fish population but only the most studied and accepted ones relative to spawning distribution are developed here. As much as possible we have tried to isolate each individual hypothesis from the others, for the sake of clarity, but mainly because it is only when the hypotheses are clearly and strictly defined that they can be discussed and challenged in an objective manner. Fish spatial distribution in the wild is complex and the hypotheses presented above are not exclusive of each other. On the contrary, it is expected that fish distribution patterns observed in the wild result from the interplay between several of the hypothesised processes: environmental controls, density dependent habitat selection, spatial dependency, demographic structure or memory. As an example, experimental studies on juvenile flatfish have shown that habitat selection can depend upon the combined effects of ontogeny (development), temperature, sediment type and the density of individuals (Laurel *et al.*, 2007), thereby combining three hypotheses (demographic structure, environment and density dependence). Strict abiotic environmental control of species distribution is likely to depend upon the species abundance (at high abundance competition is more likely to have an effect, while at low abundance abiotic factors may dominate). Thus, abiotic control may be more applicable under some conditions (low abundance) than others (Mitchell, 2005).

1.3. Population distribution models

Many statistical models can ultimately be used to predict fish population spatial distribution, under observed, forecasted or hypothesized conditions. We have chosen to group model types according to the underlying control hypotheses listed above. This is helpful in understanding how each hypothesis is being addressed through numerical models and to highlight the respective strength and weaknesses in model developments. Below, we describe some of the numerical applications currently used for modelling fish population distributions. We have primarily focussed on PDMs that can be used to predict population distribution, rather than models restricted to strict inference. We then discuss how these models can be evaluated, compared and combined.

1.3.1. Geographical models

Geographical models can be constructed from past empirical evidence. For example, regular stock assessment surveys carried out over a number of years can be used to construct empirical average spatial distributions for particular species. This is commonly done for constructing traditional atlases (e.g. online maps of fish species spatial distribution in the North Sea provided by the ICES-Fishmap project, http://www.ices.dk/marineworld/fishmap/ices/). Geographical models are not totally hypothesis-free because they require that the spatial and temporal scales at which the data aggregation is made be defined and that a particular statistical distribution be selected (e.g. Normal, Poisson, etc.). However, because the geographical hypothesis is not process-related, geographical models remain fairly simple to construct and in many cases can serve as the benchmark against which more complex models can be evaluated. Geographical models can be formulated by continuous functions in which presence or abundance is a direct function of continuous geographical coordinates (longitude and latitude). This can be achieved using a suite of possible techniques, including ordinary least square (OLS) regression, generalised additive models (GAMs), regression trees or other. Alternatively, the geographical space may be divided into discrete spatial elements, either regular or not. Regular grids are commonly found in fisheries statistics (e.g. ICES statistical rectangles are 0.5° latitude and 1.0° longitude). Regions of variable sizes can also be defined, on the basis of known environmental characteristics. As a result of these constructions, either the rectangle or region identity may be used as predictors of abundance in a given location.

1.3.2. Environment based models

Environment based models aim to relate observations of animal presence or abundance to various attributes of the environment using a statistical method. These are SDMs in essence. Many statistical methods can be applied to relate animal presence or abundance to environmental factors. Choice among these methods depends on the nature of biological observations (or biological response). The biological response that is modelled can be presence-only data, presence/absence, count data or continuous data such as density or biomass. For presence only data, methods pertaining to the family of climatic envelopes are generally used (Elith *et al.*, 2006; Hirzel *et al.*, 2002; Hirzel *et al.*, 2006; Pearce and Boyce, 2006; Pearson and Dawson, 2003). Pearce and Boyce (2006) and Elith *et al.* (2006) provide a synthesis and critique of the methods currently available. For binary, counts and continuous data, classifications and regressions methods have been widely used and most methods are now easily available, either as individual software or often as a library developed in R (R Development Core Team, 2008, http://www.r-project.org/). Ordinary Least Square regression (OLS), Generalised Linear Models (GLMs, McCullagh and Nelder, 1989) and Generalised Additive Models (GAMs, Hastie and Tibshirani, 1990; Guisan *et al.*, 2002; Wood and Augustin, 2002) have been used to construct environmental models with various statistical distributions and degree of complexity in the response shape. Several methods based on GAMs have been implemented to allow for the automatic selection of environmental predictors (e.g. the R packages or functions *mgcv*, *GRASP and BRUTO*) and the automatic inclusion of interactions between predictors (the Hyperniche software, McCune, 2006). Alternative approaches include Multi Adaptive Regression Splines (MARS, Friedman, 1991), Artificial Neural Networks (ANN, Maravelias *et al.*, 2003; Joy and Death, 2004; Wieland and Jarre-Teichmann, 1997), regression trees (De'ath and Fabricius, 2000; De'ath, 2002; Stratoudakis *et al.*, 1998; Sutton *et al.*, 2005; Usio, 2007) and their boosted form (Friedman *et al.*, 2000; Friedman, 2001; Friedman and Meulman, 2003; Leathwick *et al.*, 2006b).

Whilst all the methods described above are designed to model the average biological response to the environment, alternative techniques have been developed to model specific parts (or quantiles) of the response distribution. Quantile regressions (Koenker and Basset Jr., 1978; Koenker, 2005) are particularly well suited for modelling environment-population relationships in the case of limiting environmental conditions (in the sense of the law of limiting factors, van der Ploeg *et al.*, 1999). Quantile regressions have been applied for various ecological problems (Cade *et al.*, 1999; Cade and Noon, 2003; Hiddink, 2005; Planque and Buffaz, 2008). When applied to the upper distribution quantiles, the method is well suited to model potential habitat in which environment variables can be used to predict potentially suitable areas but are not used to predict the average population response (Eastwood *et al.*, 2001; Eastwood *et al.*, 2003; Vaz *et al.*, 2008).

Geographical Weighted Regressions (GWR, Fotheringham *et al.*, 2002) address the problem of space-dependent relationships. They are based on the idea that the link between the species and the environment can vary locally, reflected by the parameters that can vary among the space using a kernel smoothing function. A very rich and recent literature on comparing all these techniques is available (Maggini *et al.*, 2006; Moisen and Frescino, 2002; Moisen *et al.*, 2006; Segurado and Araujo, 2004; Thuiller, 2003).

1.3.3. Density dependent habitat selection and Ideal Free Distribution models

The IFD model (Fretwell and Lucas, 1970; Fretwell, 1972) is expressed by a suitability index which depends on the basic local suitability of the habitat and a function of the local population density. At equilibrium, the population is distributed in such way that the suitability is equal in all occupied habitats (and equal or lower in unoccupied ones). Basic suitability is often not directly accessible but it can be derived from empirical relationships such as those obtained in environment-based models. In the specific situation of animal spatial distribution being controlled by the availability of preys and density-dependence, the equation can be written as in Sutherland (1983), where local density depends on the proportion of individuals in a given habitat and the proportion of available resource in this habitat. The equation can be applied for any type of resource (Tyler and Hargrove, 1997). DDHS models that are developed to predict habitat suitability can be written to predict species presence of abundance at given spatial locations. Particular adaptations of the original IFD model have been developed to account for predator-preys dynamics, long-term memory, individual difference in competitive ability, energetic cost of travelling between habitats or anti-predator tactics. Accounts of these developments can be found in Tregenza (1995) and Schilling (2005).

1.3.4. Spatial dependency in models

Spatial dependency in fish distribution occurs when locations close to each other exhibit more similar values than those further apart. This results in spatial autocorrelation, a statistical issue that has long been recognized (Student, 1914) but which has been integrated in ecological studies only more recently (Legendre, 1993). There has been an extensive development of the tools available to measure, test and account for autocorrelation in species distribution models during the recent decade (see review in Dormann *et al.*, 2007a).

It is not the intention here to review the list of possible models available to deal with autocorrelation in habitat models. Such review can be found in Dormann *et al.*, 2007a; Dormann, 2007b; Keitt *et al.*, 2002; Lichstein *et al.*, 2002; Segurado *et al.*, 2006; Wintle and Bartos, 2006). The following section is restricted to the presentation of models that appear to be particularly suited to deal with the modelling of fish population spatial distribution.

Habitat models which include spatial autocorrelation are usually extensions of models from two opposite origins. They can originate from geostatistical models to which an environmental component is added or from environment-based models to which a spatial component is added (Nishida and Chen, 2004). The result is a combination of environmental control and spatial dependence which can be estimated sequentially or simultaneously. Geostatistical models are specifically developed to model spatially autocorrelated processes and have been applied to the study and modelling of fish spatial distribution (Petitgas, 2001). In geostatistical models, the spatial prediction is analogous to spatial interpolation and is referred to as “kriging”. In co-kriging spatial dependence on variables and co-variables are considered simultaneously when modelling spatial distribution. Regression Kriging (RK) is an example of sequential method that combines predictions from a regression model with kriging of the model residuals. RK methods allow for flexible regression models to be used (e.g. GAMs, GLMs), and a wide variety of predictor variables to be included (Miller et al., 2007; Maxwell *et al.,* 2009).

Autoregressive models (AR) allow for the inclusion of a term accounting for spatial autocorrelation in addition to environmental factors. One of the first AR method to applied was the autologistic regression (Augustin *et al.*, 1996). The environmental effects can be modelled with multi-regression implemented with GLM or GAM, and can be applied to normal (auto-gaussian) and Poisson (auto-Poisson) data. However, autologistic regression have a poor capacity for parameter estimation and limited predictive power in comparison with other methods (Dormann, 2007c; Carl and Kuhn, 2007). Conditional autoregressive models (CAR) and simultaneous autoregressive models (SAR) have been derived from conventional AR models to better estimate the regression and autoregression parameters. In these models, endogenous and exogenous autocorrelation can be included using a term accounting for each of them (Kissling and Carl, 2008). However, as they are based on OLS, only Gaussian distribution can be modelled. For binomial distribution, Generalised Estimating Equation (GEE) seems to be a promising method (Carl and Kuhn, 2007).

1.3.5. Models that include population structure

To our knowledge, there is little or no application of statistical fish distribution models which include population demographic structure. When demographic structure is believed to influence the geographical distribution of a fish population, a valid strategy is to separate the population in entities that are then modelled separately (e.g. different age/size or sex groups).

Spatial distribution models can be constructed, for each sub-group of the population, using the modelling approaches outlined in this review. The spatial distribution of the whole fish population can then be reconstructed by assembling sub-group distributions afterwards.

1.3.6. Memory

Memory can be modelled as a factor controlling the spatial distribution of fish populations in a variety of ways. We shall here consider only how it can be implemented in regression models. In a regression modelling approach in which the occurrence of fish is fitted as a response to explicative covariates, the concept of memory can be directly translated by a variable accounting for population memory into covariate. Such variable can be similar to that constructed and used in the Eulerian displacement modelling approach. Accordingly, the conditions for entrainment to happen (sensu Petitgas *et al.*, 2006) may also be formulated as covariates. For instance the occurrence of older fish in previous years can be used as covariates to predict the spatial distribution of recruits in the current year (Rindorf and Lewy, 2006). The effects of past spatial distributions of the population can be included one or several time steps back in time, in a manner similar to what is commonly done for temporal autoregressive models.

1.4. Evaluation and selection of fish spatial distribution models

1.4.1. Evaluating single models

The conventional approach for testing the statistical significance of a model is to compare a selected metrics (for example, the parameters of a regression between a given environmental variable and local fish abundance) associated to the model with that same metrics under a null hypothesis (e.g. no relationship between the selected environmental variable and abundance). While this approach is valid if model uncertainty is small (i.e. there is only one or few models to choose from) and the null hypothesis is clearly defined and makes biological sense, it may not be appropriate if one or both of these assumptions are violated (Johnson and Omland, 2004). This is unfortunately a common situation when modelling spatial distribution of fish populations. The null hypotheses are rarely explicitly stated in the literature and they often implicitly refer to the total absence of spatial structure and control of fish presence or abundance (the hypothesis of ‘absence of control’ abovementioned). Models performing better than the “no control” hypothesis are often taken as being statistically significant. However, since most models, even when wrong, can do better that the “no control” one, most studies which only test a single hypothesis will turn out to find their model statistically significant. This is true even when the model displays very poor fitting or predictive performances.

An additional difficulty when testing for individual models is related to spatial and temporal autocorrelation in the data. Fish data is very often autocorrelated in space, time or both, resulting in non-independence between observations. In most statistical tests, individual observations are assumed to be independent from one another, and failure to meet this assumption results in biased statistical tests (Legendre, 1993; Dale and Fortin, 2002). This effect is particularly worrying because it often remains unrecognised and lead to high rejection rates of the null hypothesis, a situation which encourages publication of biased results and use of wrongly specified numerical models.

The performance of spatial distribution models relates to their prediction capabilities. Whilst model-fitting performance can be constantly improved by adding complexity to the model formulation, model-predicting performance will decline after a certain degree of complexity has been reached because added complexity is modelling noise (or error terms) rather than signal in the observation data. This is commonly known as the bias/variance trade-off (Hastie *et al.*, 2001). Model evaluation should therefore be performed on a testing data set independent from the training data set. Truly independent data sets may yet be difficult to obtain because of autocorrelation in the data collected. This point is discussed further in the following section.

1.4.2. Selecting among single hypotheses based models

Even when a single ecological hypothesis can be selected as the principal object of investigation, this ecological hypothesis may possibly be expressed in many different numerical formulations. In such case null hypothesis statistical testing (NHST) become ill-adapted because the question is no longer to test a unique numerical model against a null model (for the null hypothesis), but rather to compare different models of the same family. As an example, models developed along the lines of environmental control of population spatial distribution may be formulated in very different manners (e.g. GLMs, GAMs, regression trees and so on.) and may include different environmental variables or combination thereof. NHST done on each individual model can result in several competing models being significantly better than the null hypothesis. However, comparison of the models cannot be achieved through NHST and requires instead multimodel inference.

This can be accomplished by an information-theoretic approach as prescribed by Burnham and Anderson (2002), Johnson and Omland (2004), Stephens *et al.* (2007) and Diniz-Filho *et al.* (2008). In multimodel inference, several models are compared on the basis of a distance metrics between the model predictions and the observations. One commonly used metric is the Akaike Information Criterion (AIC, Akaike, 1974), a measure of the likelihood between the fitted values of the model and the observed values, penalised by the number of parameters in the model. Model selection based on AIC is done as a compromise between model complexity and model ability to reproduce the observations. It is possible to compare models by calculating their difference of AIC (Δ_i) with the best model an the probability (w_i) of each model to be the best model. The selection procedure can result in several models being reasonably good candidates for ‘best model’, so the outcome of multimodel inference may not always be the identification of a unique solution but rather an ensemble of likely solutions.

In multimodel selection, the null hypothesis is no longer needed and mis-specification of H_0 is therefore no longer an issue. However, model selection relies on several important assumptions that need to be carefully checked. First, the set of available observations has to reflect the underlying ‘truth’ (e.g. the true spatial distribution of animals). In the case of fish populations, this is not always simple because of limitations in the observational methods. Second, the best model needs to be present in the set of candidate models. This implies that the set of candidate models has been accurately defined and that all the reasonable hypotheses are reflected in one or several numerical models. Third, individual observations should be independent. Non-independence due to autocorrelation in the data will tend to generate longer (with more explanatory variables) models, so that the problem is analogous to the inflated Type I errors in the NHST approach, which tend to give more statistical importance to some explanatory variables than they really have (Diniz-Filho *et al.*, 2008). Quality of observation data, construction of the set of candidate models and data autocorrelation should be carefully checked before multimodel inference is performed.

Following Hastie *et al.* (2001), we have argued that model evaluation should be performed on data sets independent from those that have been used for model calibration (i.e. parameter estimation). By doing so, model prediction error rather than model fitting error is used as a measure of model performance. This should allow for a better evaluation of model performance when used in a predictive mode. Cross-validation methods are doing this by splitting the observation data set into a “training” and a “test” data set and evaluating model performance on the latter. The operation is usually repeated several times, so that each observation is used alternatively for training and testing. A fundamental assumption of cross-validation is that the training and testing data sets are independent. This is often violated because field surveys observations are spatially and temporally autocorrelated. This results in apparent predictive performance being closely related to model fitting, and will usually lead to selection of complex models with low bias and high variance, i.e. with low predictive power (Telford and Birks, 2005).

1.4.3. Towards a modelling framework for combining multiple hypotheses

The causes behind the true spatial distribution of fish populations are numerous and our perception of such distribution is only accessible through limited observations. The processes controlling fish spatial distribution at the population level are many, complex and are likely to vary with specific populations and geographical regions. Any attempt to describe and model fish spatial distribution with only one of these hypotheses is likely to generate poor results or even fail. More promising strategies include the construction of models which include several hypotheses. This can be achieved by combining the individual models presented above into more complex models that can be tested through multimodel evaluation and selection procedures. We recommend that models of different levels of complexity be build from the existing knowledge and hypotheses that are specific to the population studied. Constructing such candidate models is not simple, but following recommendations which apply for model selection is useful. These include 1) the selection of meaningful variables (which can be associated with causal factors), 2) the (theoretical) justification of functions that defines the relationships between controls and responses, and 3) the error structure of the model. Ultimately, the number of candidate models should be small enough to avoid generating so many models that spurious findings become likely (Johnson and Omland, 2004). A strategy consisting in the construction of all possible combination of models should be proscribed (Burnham and Anderson, 2002). The principles of model selection and the associated theoretical approach are now well developed and technically available to the community of marine and fish ecology researchers (see e.g. Hilborn and Mangel, 1997; Burnham and Anderson, 2002). So, if models can be constructed to match working hypotheses, these will be amenable to model selection and the hypotheses - of combination thereof - may be compared and selected. What remains largely unexplored today is the combination of hypotheses through integrated spatial distribution models. Evidence exists that multiple controls of spatial distribution or habitat selection is at play (see e.g. Laurel *et al.*, 2007) but population distribution models which integrate several controlling hypotheses are still the exception. This may result from differences in the mathematical formulation of the various types of models which makes them difficult to assemble. For example, whilst all models can be applied to predict abundance (or presence) at a given geographical location (i.e. they have the same output variable), they may request input information of different type: geographical models rely on geographical coordinates only, environmental models necessitate additional observed or modelled environmental conditions, DDHS and spatial dependency models require local or neighbouring densities to be known (a difficult situation since this is what the model is trying to predict), and memory models require knowledge about past population states. How these models can be assembled remains a challenge to marine and fisheries ecologists. Selection of models from different ‘families’ or hybrid models is also challenging, mostly because the amount of independent information used in these models is often poorly estimated as a result of strong spatial and temporal autocorrelation in the data. When this is the case, model selection techniques will tend to favour complex models which have lower bias but also less predictive power. Finally, if and when appropriate models can be developed to predict fish spatial distribution, model selection procedure may lead to the selection of more than one ‘best’ model. This happens when the data available is not sufficient to resolve model uncertainty. Researchers are then left with a subset of possible models, structurally different from one another but with similar performances. In such situation, multiple models can be used within an appropriate ensemble forecasting framework. The technique, if used appropriately, can lead to robust predictions without the need to identify a unique best model (Araujo and New, 2007).

CHAPITRE 2

Material and methods

The present chapter aims to describe the data of the three case studies, North Sea plaice and whiting and Bay of Biscay anchovy, and the various analyses that were performed to study their spawning distribution. To summarise this section, the data from plaice, whiting and anchovy were processed by geostatistical and multivariate analyses and spatial modelling. For each data set, a set of six hypotheses were investigated. Models were constructed, which can allow for complex combinations of the core hypotheses. The models were then evaluated based on there fitting and predictive performance and the hypotheses of control mechanisms were ranked using variance partitioning.

2.1. North Sea plaice and whiting

2.1.1. The North Sea

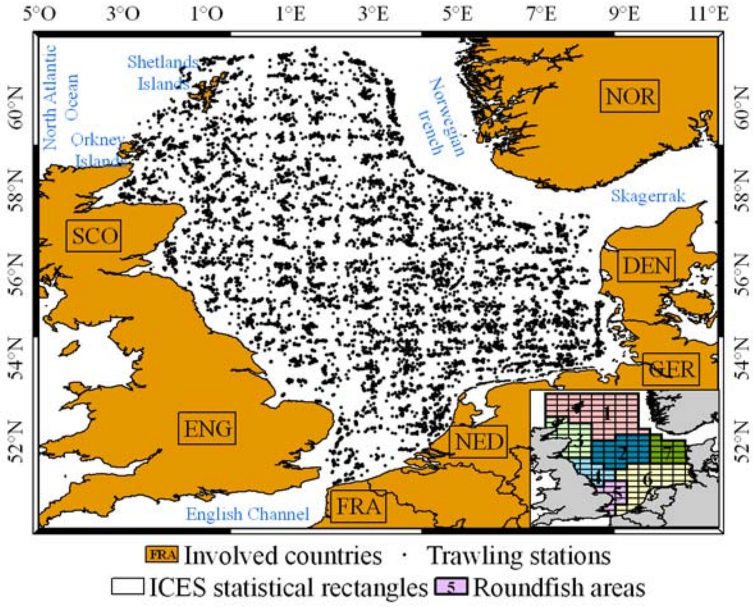


Figure 2.1. The North Sea international bottom trawl survey (IBTS). Main : countries that are involved in the first quarter (January-March) of the survey are: DEN : Denmark, ENG : England, FRA : France, GER : Germany, NED : Netherlands, NOR : Norway, SCO : Scotland. Black dots indicate the location of bottom trawls hauls since 1980. Box inlet: location of the ICES statistical rectangles and roundfish areas.

The North Sea (fig. 2.1) is an epicontinental sea with depth ranging between 15 and 200 m and characterised by a macrotidal regime. This basin of 575,300 km^2 surface area with a volume of 42,300 km^3 is limited to the northwest by the continental slope near the Shetlands and Orkney Islands, and by the Nowergian trench (700 m in depth) to the northeast. During the winter, North Sea waters are cold and desalted in the east because of cold water inflow coming from the Baltic Sea and large river inflows from the French to Danish coasts. Warmer and saltier waters enter the North Sea from the North Atlantic Ocean through the English Channel to the southwest and from the North Atlantic current to the northwest. During the winter, waters are vertically mixed with homogeneous temperature and salinity across the water column whereas during summer, a tidal front is found in the central North Sea with mixed water in the southern part and stratified water in the northern part (Hill *et al.*, 1993).

2.1.2. The IBTS survey

The first quarter International Bottom Trawl Survey (IBTS) has been carried out each year (since 1980) in the North Sea, from January to March (table 2.1), to collect data necessary for the stock assessment of several important demersal fishes (ICES, 2007a; ICES, 2008a). The sampling network is designed according to statistical rectangles of 1° of longitude by 0.5° of latitude (fig. 2.1) defined by the International Council for the Exploration of the Sea (ICES). Each rectangle is visited by two different countries which each perform one standardised 30 min trawl using a 36/47 GOV (Grande Ouverture Verticale) bottom trawl. Trawling locations are randomly chosen by the country among a pre-defined set of three or four trawling locations inside the rectangle. This results in a minimum of two trawl hauls per rectangle per survey quarter. Trawling depth (m), temperature (°C) and salinity are recorded at each trawling location. All specimens in each trawl sample are sorted by species and counted. Length measurements, otolith sampling and sexual maturity staging are performed for several key species, on a representative sample of individuals within seven standard “roundfish” areas (ICES, 2004a, fig. 2.1). Individuals are classified into four stages of maturity, with stage 3 corresponding to spawning individuals with fluent gonads (ICES, 2004a). From 1980 to 2007, 11,343 bottom trawls were performed, which represents an average of 405 bottom trawl hauls per year. These data are available through the DATRAS database (DAtabase of TRAwl Surveys, http://datras.ices.dk/Home/Default.aspx) coordinated by ICES.

Table 2.1. Time periods of the first quarter of the IBTS surveys since 1980 and number of trawling stations available for analyses performed on plaice and whiting.

Year	**Time of the IBTS**		**Number of stations**		**Year**	**Time of the IBTS**		**Number of stations**	
	Start	End	Plaice	Whiting		Start	End	Plaice	Whiting
1980	22/1	6/3	270	353	**1994**	6/1	21/3	248	322
1981	21/1	5/3	180	199	**1995**	7/1	28/2	224	291
1982	28/1	13/3	234	275	**1996**	16/1	27/2	206	280
1983	1/2	10/3	265	320	**1997**	17/1	27/2	242	316
1984	26/1	10/3	306	391	**1998**	6/1	25/2	281	364
1985	28/1	5/3	334	435	**1999**	10/1	25/2	242	313
1986	11/1	17/3	340	446	**2000**	13/1	4/3	262	337
1987	24/1	22/3	356	457	**2001**	14/1	22/3	303	380
1988	24/1	2/3	288	381	**2002**	18/1	28/2	300	379
1989	24/1	2/3	285	360	**2003**	8/1	27/2	294	374
1990	21/1	24/2	248	329	**2004**	21/1	26/2	254	330
1991	7/1	28/2	224	324	**2005**	21/1	7/3	228	305
1992	13/1	26/2	256	330	**2006**	12/1	23/2	246	326
1993	17/1	1/3	222	301	**2007**	16/1	26/2	179	217

2.1.3. Abundance of spawning adults

Abundance data of spawning adults from the first quarter of the IBTS (January-March) were computed from the available proportion of spawning adults within any given length class. For whiting, these proportions were calculated for each of the seven standard roundfish areas, by pooling together maturity data available from 1990 to present. For plaice, the determination of sexual maturity has only been carried out since 2001, so there was not enough data to calculate the proportion of stage 3 adults within each length class and for each year and roundfish area. Consequently, data on sexual maturity from 2001 to present were merged to calculate these proportions for northwest areas (areas 2, 3 and 4 pooled together, fig. 2.1) and southeast areas (areas 5, 6 and 7 pooled together, fig. 2.1). This choice was based on the time of spawning and spawning activity of plaice in each area (Daan *et al.*, 1990). No data on the sexual maturity of plaice in area 1 was available. For both species, data on males and females were merged as there was not enough data to calculate proportions of spawning adults for each sex. This reduced the total number of trawling stations, available for analyses, from 11,343 to 7,317 and 9,435, respectively for plaice and whiting (table. 2.1). For each of these stations, total abundance of spawning adults (in individuals.km^{-2}) were calculated from the summed product of the total abundance within each length class and the corresponding proportions of spawning adults inside that length class.

2.2. Bay of Biscay anchovy

2.2.1. The Bay of Biscay

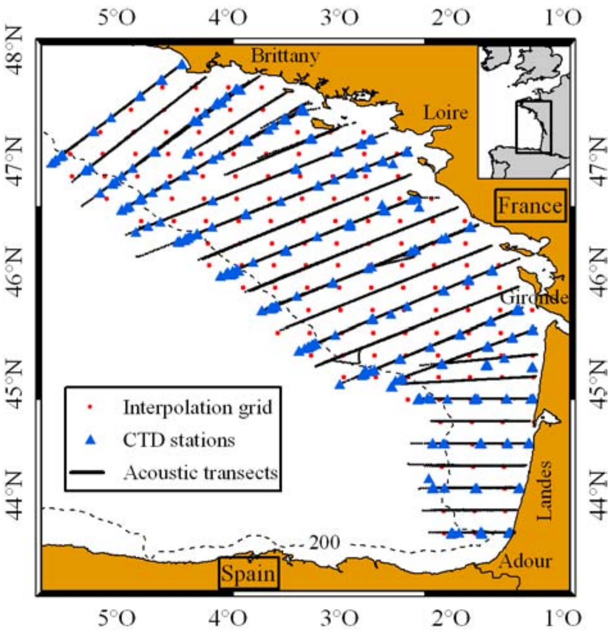


Figure 2.2. The Bay of Biscay PELGAS surveys: acoustic transects (lines) and hydrological stations (triangles) performed between 2000 and 2007. Red dots show the interpolation grid on which abundances of spawning adults of anchovy and environmental variables were estimated. Locations of the Loire, Gironde and Adour estuaries, Britanny and Landes coasts, and the 200 m isobath (dotted line) are also indicated. Box inlet: location of the French continental shelf of the Bay of Biscay along the French and Spanish coasts.

The Bay of Biscay belongs to the North-Atlantic Ocean and is located along the west coast of France from Britanny to Pyrénées-Atlantiques regions, and along the north coast of Spain, from the Basque country to Galicia. It covers an area of 225,000 km² and has a maximal depth of 4,735 m. The French continental shelf of the Bay of Biscay (43°N30’-48°N30’, fig. 2.2) is narrower in its southern part (wide of only a few km) than in its northern part (150 km in width) that opens to the Celtic Sea. It is limited by the 200 m isobath on its western part. From an hydrological point of view, it can be characterised by eight regions based on temperature and salinity regimes (Planque *et al.*, 2004). These typologies are relatively stable in time and offers different environmental conditions. The peak of planktonic production is maximal in spring, with a weak coastal upwelling between the Adour and Gironde estuaries where primary production is high due to high nutrient input from river discharge. Although the influence of the general circulation regime is weak on the shelf, meso-scale hydrological structures such as fronts, eddies and river plumes are present (Koutsikopoulos and Le Cann, 1996).

2.2.2. The PELGAS survey

The PELGAS survey (PELagique GAScogne) has taken place each year since 2000 in the spring (table 2.2) on the French continental shelf of the Bay of Biscay. It is mainly designed for stock assessment of small pelagic fish (anchovy and sardine) using hydroacoustic methods (ICES, 2006; ICES, 2008b). During the day, acoustic prospecting is performed along transects that are separated by 12 nautical miles (fig. 2.2). These transects are oriented perpendicularly to the coast, from 25 m bottom depth to the continental slope (250 m), and cover the whole the continental shelf from the Spanish coast in the south to the coast of Brittany in the north. Along these transects, acoustic energy is recorded at each nautical mile using an echo-sounder (SIMRAD EK500 38 kHz) and located as a point called an EDSU (Elementary Distance Sampling Unit). Fish species composition is determined from pelagic trawl hauls. All individuals captured (or a representative sample) are identified, counted and weighted. Otolith sampling, sexing and maturity staging are performed on a representative sample of individuals. Records of each EDSU are then combined following the method outlined in Petitgas *et al.* (2003) to allocate fractions of the total energy recorded (s_A: nautical area scattering coefficient) in term of biomass to the various species captured in the trawl. For each EDSU, this biomass is reallocated in term of number of individuals per size and age according to the mean weight observed in the reference trawl. During the night, vertical profiles using a CTD (Conductivity Temperature Depth) probe (Seabird 19) are performed along the same transects to measure temperature, salinity and density along the water column (fig. 2.2).

Table 2.2. Time period of the PELGAS survey for each year between 2000 and 2007. Number of CTD stations and EDSU available for kriging interpolation are shown for each year, along with the final number of stations of the interpolation grid.

Year	**Time period of the PELGAS survey**		**Number of stations**		
	Beginning	End	CTD station	EDSU	Interpolation grid
2000	18/04	14/05	93	1500	149
2001	3/05	3/06	60	1636	149
2002	10/05	6/06	80	1593	149
2003	30/05	24/06	75	1625	149
2004	29/04	23/05	79	1736	149
2005	5/05	24/05	66	1381	124
2006	2/05	30/05	62	1355	113
2007	27/04	19/05	66	1447	124

2.2.3. Abundance of spawning adults

During the PELGAS survey (April-June), all adult anchovies were sexually mature and had started spawning (Motos *et al.*, 1996). Total anchovy abundance calculated at each EDSU were used to show the spatial distribution of spawning adults between 2000 and 2007.

2.3. Spatial distribution of spawning adults

2.3.1. Geostatistics

Spatial structures can be identified and described quantitatively using geostatistics (Matheron, 1962; Petitgas, 1993; Petitgas, 2001; Mello and Rose, 2005; Woillez *et al.*, 2007). Geostatistical methods embody a suite of methods for analysing spatial data using a correlogram or a variogram and allow the estimation of the values of a variable of interest at unsampled locations using interpolation by ‘kriging’ (Webster and Oliver, 2001). The variogram (fig. 2.3) is a function that measures the relation between pairs of observations a certain distance apart. It summarises the way in which the variance of a variable changes as the distance and direction separating any two points vary. Typically, for spatially structured data, the variance is small at short lags and increases with larger separating distance. Kriging interpolation produces optimal unbiased estimates that can be used for mapping by taking into account the way in which a variable varies in space to predict the values at unsampled locations.

Geostatistics (Webster and Oliver, 2001) were used to map the spatial distribution of spawning adults of plaice and whiting and also to interpolate on the same grid the environmental variables and the abundance of spawning adults for anchovy. Abundance of spawning adults were log-transformed ($\log(x+1)$, where x = abundance) to be closer to a normal distribution. For each year and each variable, an experimental variogram was computed according to formula (1).

$$(1)\ \gamma(h) = \frac{1}{2 \times n(h)} \times \sum \left[z(x+h) - z(x) \right]^2$$

where $\gamma(h)$ is the experimental variogram, $n(h)$ is the number of pairs of observations for the distance h, $z(x)$ is the observed abundance at location x and h is the distance between two locations (fig. 2.3). The distance h was calculated using the latitude and corrected longitude of the trawling location for plaice and whiting, and that of the EDSU for anchovy. The longitude correction ($longitude \times \cos((latitude \times \pi)/180)$) transforms decimal degrees of longitude into decimal degrees of latitude which are of constant distance in a Mercator-like projection formula. In case of non stationary of $z(x)$, it can be split into two components following (2).

$$(2) \quad z(x) = m(x) + R(x)$$

where $m(x)$ is the spatial trend and $R(x)$ are the residuals.

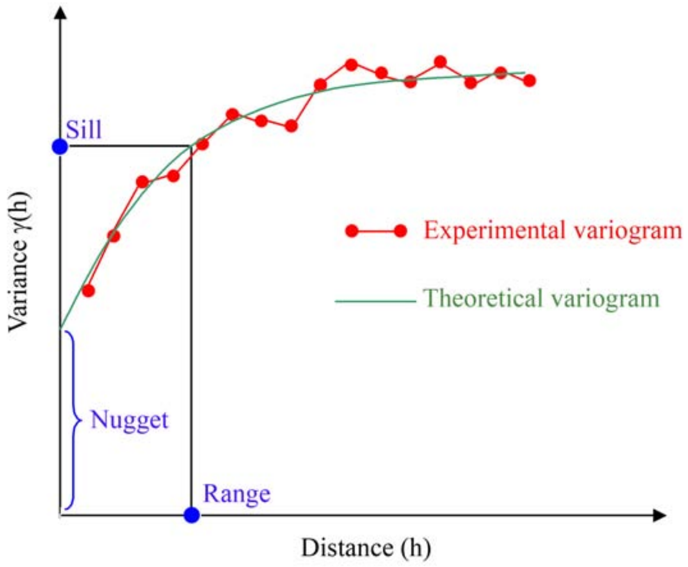


Figure 2.3. Example of experimental and theoretical variograms. The nugget is the local variance. The range indicates the distance after which observations become independent. At this distance, variance reaches a maximum which corresponds to the sill. Here, a combined model with a nugget effect and an exponential model is used as the theoretical variogram.

When accounting for more than 20% of the variation of the data, the spatial trend was modelled by fitting a low-order polynomial (linear or quadratic regression) to the spatial coordinates using the least-square regression method. The experimental variogram was then calculated on the residuals. A theoretical model, chosen among exponential, circular, spherical and penta-spherical authorised functions was adjusted to this experimental variogram to determine the nugget, sill and range (fig. 2.3). These four models were adjusted using the least-square regression method (Webster and Oliver, 2001) and the one with the best visual and statistical fit to the experimental variogram was retained as the chosen theoretical variogram model. This theoretical variogram was then used to estimate $z(x)$ on the mesh of a regular grid by using the ordinary kriging interpolation method, or the universal kriging in the presence of a spatial drift. A mesh size of 0.2 decimal degrees was chosen for the interpolation grid, as the mean survey resolutions were 0.16 decimal degrees for IBTS and 0.18 decimal degrees (CTD stations) for PELGAS. Geostatistics were implemented using Genstat (GenStat Release 7.1., 2004).

Using previously interpolated abundance, the ratio of spatial variance over temporal variance was calculated as a criteria to examine how the distribution varies in both space and time (Planque *et al.*, 2006). The spatial variance was calculated as the mean variance between all grid cells over all years and the temporal variance as the mean variance between years over all grid cells. A ratio exceeding 1 means that the population distribution is more variable in space than in time, which indicates a spatial structure that tends to persist in time. A ratio less than 1 indicates that there is more variation in time than in space, which means that temporal variance greatly exceeds geographical difference (Planque *et al.*, 2006).

2.3.2. Distribution mapping

Interpolated abundance of spawning adults were mapped in a geographical information system using ArcMap 9.1 (ESRI, 2005). Extrapolated areas (located outside of the surveyed area) were removed from the maps. The average map summarises the average spawning area and was computed as the mean of the maps between 1980 and 2007 for plaice and whiting, and between 2000 an 2007 for anchovy. The variability map represents the inter-annual variability in spawning and was produced by summing the variance of maps between 1980 and 2007 for plaice and whiting, and between 2000 and 2007 for anchovy. Preferential, occasional and unfavourable spawning sites were defined following Bellier *et al.* (2007a) using the average and variability maps. Preferential spawning sites can be characterised by a high mean and a low variance, occasional spawning sites by a high mean and a high variance and unfavourable spawning sites by a low mean and a low variance. The thresholds used to split the maps into high and low values of mean and variance were chosen visually according to their distribution histograms. The raster calculator of the Spatial Analyst extension of ArcMap 9.1 was used to combine the average and variability maps into the final map of preferential, occasional and unfavourable spawning sites.

2.3.3. Inter-annual variability

The inter-annual variability of spawning adult distribution for plaice and whiting was studied using a correspondence analysis performed on kriged abundance from 1980 to 2007. A contingency table of 28 years (columns) and 692 stations (lines) for plaice and 807 stations for whiting was extracted from the interpolated grid. A cluster analysis based on a Chi-square distance and using a flexible link (β=-0.25) was performed to group similar stations together. The correspondence analysis was performed using CANOCO 4.5 (ter Braak and Smilauer, 2002) and the cluster analysis using PC-ORD 4.25 (McCune and Mefford, 1999).

2.4. The multi-model approach

2.4.1. Modelling the hypotheses of control

The spatial distribution of spawning adults can result from a variety of controls, as discussed in chapter 1. To investigate the role of these controls and how they can possibly interact, six families of models were constructed: geographical position, environment, spatial dependency, age structure, population size and population memory. These models were then assembled with different level of complexity, from single hypotheses (e.g. pure environmental control) to complex interactions (e.g. density-dependence combined with population memory and environmental controls). The core models are detailed below.

Environment

For plaice and whiting, five environmental variables were used: depth, bottom temperature and salinity, sea-bed stress ($N.m^{-2}$) and sediment type. Depth, sea-bed stress and seabed sediment type were used to depict persistent environment, i.e. environmental conditions that are spatially structured but have not changed during the study period. Bottom temperature and salinity were used to depict non-persistent environment, i.e. environmental conditions that are also spatially structured but may have changed during the study period. Salinity was transformed using the equation: $\exp(Salinity)/10^{15}$, to be closer to normality. Sea-bed stress is a measure of the shear friction of water on the seabed due to the tidal currents. It was estimated using a 2D hydrodynamic model (Aldridge and Davies, 1993) from the Proudman Oceanographic Laboratory (POL), and running on a regular grid of 1/8° of longitude by 1/12° of latitude (WGS 1984 datum). The grid of points was then interpolated using ArcMap’s Spatial Analyst extension (ESRI, 2005) to create a continuous raster layer of 1 km^2 resolution. The bedstress was log-transformed to be closer to normality. Sediment type originated from the seafloor sediment of the North Sea built during the MARGIS project (Schlüter and Jerosch, 2008). Sediment types were classed into five categories : fine sand, coarse sand, mud, pebbles and gravels. A category of sea-bed stress and sediment type was allocated to each trawl location by resampling the corresponding map at that location using ArcMap’s Hawth’s Analysis Tools extension (Beyer, 2004).

For anchovy, 14 environmental variables were used. Seven were directly extracted from CTD profiles: bottom depth, bottom and surface temperature and salinity, and temperature and salinity differences between the surface and the bottom. Two were used to characterise the vertical stratification of the water column, mixed layer depth and potential energy deficit. Potential energy deficit is the amount of energy needed to vertically mix the water column, whilst mixed layer depth is the depth where the gradient of density is the highest. These were calculated using the densities extracted from CTD profiles using the formulas in Planque *et al.* (2006). Because of temporal differences in the time periods of the survey (table 2.2), anomalies were calculated and used for surface temperature, temperature difference and potential energy deficit. These nine variables were interpolated using kriging on the same grid as the one used to interpolate spawning adults abundance (see section 2.3.1). Five hydrological variables: depth of the pycnocline (m), primary production index ($gC.m^{-2}$), upwelling index ($m.d^{-1}$), frontal index ($kg.m^{-2}.s^{-2}$) and eddies index (s^{-2}) are outputs of the Eco-Mars 3D model that were extracted for the period 2000-2007 (Huret *et al.*, 2009). The Eco-Mars 3D model (Lazure and Dumas, 2008) combines an hydrodynamic model with a biological model and can be used to simulate the main hydrological features for the Bay of Biscay and western part of the English Channel. These indices are available on a 3-days period with a resolution of 4 km. They were averaged on the time period of the PELGAS survey for each year. Upwelling index, frontal index and eddies index were recoded in term of occurrences (0/1 for absence/presence) following thresholds of, respectively, 0.5 $m.d^{-1}$, 0.002 $kg.m^{-2}.s^{-2}$ and -10 s^{-2} (Huret M., pers. comm.). For each year, the closest value of these indexes was attributed to each cell of the interpolation grid of spawning adults’ abundance.

A principal component (PCA) analysis was performed on these 14 environmental variables (columns) available for the 1,106 stations (lines) of the eight years, to study and group the variables according to their collinearity (fig. 2.4). The different groups of variables were used as the various environmental hypotheses of control. Five groups of environmental variables were defined from the three first axes of the PCA. The first group (Envt 1) gathers frontal index, upwelling index and salinity difference. The second group (Envt 2) consists of pycnocline depth, eddies index and primary production. The third group (Envt 3) represents potential energy deficit, surface temperature and temperature difference. The fourth group (Envt 4) is made of mixed layer depth, surface salinity, bottom salinity and depth. Bottom temperature (BT) did not belong to any of these groups.

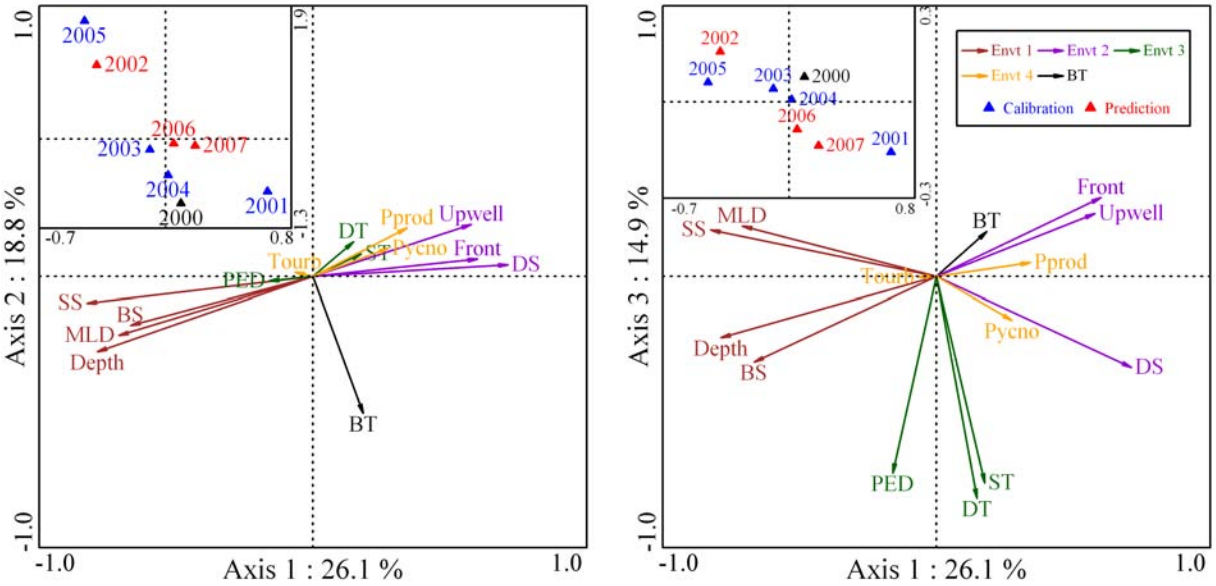


Figure 2.4. Principal component analysis performed on the 14 environmental variables of the Bay of Biscay used in the multi-model approach. Left : plot of the first and second axes. Right : plot of the first and third axis. Five groups of variables were determined (Envt 1, Envt 2, Envt 3, Envt 4, BT). Box inlets: years are represented at their respective stations' centroid. Years 2001, 2003, 2004 and 2005 were used for the calibration dataset and years 2002, 2006 and 2007 were used for the prediction dataset. BS : bottom salinity, BT : bottom temperature, DS : difference of salinity, DT : difference of temperature, Front : frontal index, MLD : mixed layer depth, PED : potential energy deficit, Pprod : primary production, Pycno : pycnocline depth, SS : surface salinity, ssb : spawning stock biomass, ST : surface temperature, Upwell : Upwelling index.

Spatial dependency

Spatial dependency (also termed spatial autocorrelation) describes the spatial structure present in the spatial distribution of a species. This structure may potentially be shared by other controls such as the environment. Spatial dependency in the spatial distribution of spawning adults was modelled using principal coordinates of neighbour matrices (PCNM). PCNM are particularly well designed to describe the spatial structure present in the distribution of an organism over a wide range of scales (Borcard and Legendre, 2002; Dray *et al.*, 2006; Bellier *et al.*, 2007b). PCNM variables result from the diagonalisation of a spatial weighting matrix and represent a spectral decomposition of the spatial relationships among the study sites. They produce orthogonal maps that maximise spatial auto-correlation and that can be directly linked to the spatial patterns of species distribution. Extraction of these PCNM was performed following Borcard and Legendre (2002). An Euclidean distance matrix was built using latitude and corrected longitude (see section 2.3.1) of trawling stations for plaice and whiting, and grid cells for anchovy. The resulting distance matrix was then truncated by recoding all the distances above a particular threshold to four times this threshold. The threshold was taken as the maximum distance between two neighbouring stations in order to keep all the stations connected together. A principal coordinate analysis was then performed on this truncated distance matrix leading to the extraction of all the potential PCNM from which only positive PCNM (eigenvectors with positive eigenvalues) were kept. Generally, the eigenvectors associated with the higher eigenvalues represent broader-scale variations, while eigenvectors derived from the smaller eigenvalues represent finer-scale variations.

PCNM that describe the spatial structure present in plaice, whiting and anchovy distribution were then selected. Following Borcard and Legendre (2002), log-transformed abundance were detrended prior to PCNM analysis; if not, half of the available PCNM variables would be forced to recover the trends, and so their role in modelling finer structures might go unnoticed (Borcard *et al.*, 2004). This was done using a quadratic trend based on latitudinal and corrected longitudinal coordinates of the trawling stations. PCNM were then confronted to these detrended log-transformed abundances following Blanchet *et al.* (2008) multiple regression procedure. In the first step, a global test including all the positive PCNM was performed to model the biological response using a 999 Monte Carlo permutations test (ter Braak and Smilauer, 2002). In the Monte-Carlo permutation test, the reference distribution is simulated by repeatedly permuting the samples. A statistical test (F-ratio) is computed for the original data and compared with those of permuted data. The value of the significance test is the probability that the response is independent from the tested explanatory variable (each PCNM). As this test was significant, the analysis was persued to the next step which consists in the forward selection. In the forward selection, the significance of each PCNM was assessed using a 999 Monte Carlo permutations test resulting to a p-value and the adjusted-R^2 (explained variance) of the model calculated with all the PCNM. PCNM were successively tested and the selection procedure stopped either if the newly added PCNM was not significant at a chosen p-value threshold, or if it resulted in the adjusted-R^2 of the whole model exceeding the one calculated for the model with all the PCNM. In order to avoid building spatial models with too many variables (and hence risk biasing the weight of the different hypotheses), only the highly significant PCNM (i.e. that explained most of the variability in the biological response) were kept. That explains why a low value for the alpha criterion (p-value) was chosen. Indeed, once this value was chosen, the adjusted-R^2 did not play any significant role into reducing the number of selected PCNM.

Spatial scale of these PCNM was determined following the method outlined in Bellier *et al.* (2007b). A variogram with a Gaussian model was adjusted, using least square regression (Webster and Oliver, 2001), to determine the range of the variogram which represents the spatial scale described by each PCNM. The eigenvalue of each PCNM were plotted versus their range, so as to group PCNM describing equivalent spatial scale into three sub-models: broad, medium and fine scale sub-models (Bellier *et al.*, 2007b). These sub-models were used as three distinct hypotheses of control. The R open-source software (R Development Core Team, 2008) combined to the PrCoord Tool 1.0 of CANOCO 4.5 software (ter Braak and Smilauer, 2002) were used for PCNM extraction. The ‘forward.sel’ function of the ‘packfor’ package (Dray *et al.*, 2007) was used for the forward selection of the PCNM. Experimental and theoretical variograms were calculated using the geoR package (Ribeiro Jr and Diggle, 2001).

Age structure

Age structure was defined by annual and spatial age structure. Annual age structure (referred to as annual demography) represents the proportion of each cohort, in each year, for the whole population, whereas spatial age structure (referred to as spatial demography) is the spatial distribution of each cohort at each station for each year. For plaice and whiting, annual abundance of each cohort are available for each year since 1957 through the work of an ICES Working Group (ICES, 2008a) that is in charge of stock assessment using the Virtual Population Analysis (VPA) method. The age-length key was applied to the total number of individuals in each trawl so as to calculate the proportion of each cohort at each station. Only the proportions of age 1 to 5 were considered, as they were the most abundant during the study period. For anchovy, annual abundance of age classes from 1 to 3 years-old are provided by the ICES working group on anchovy (WGANC, ICES, 2008b). Abundance of each age class at each EDSU (based on the length distribution observed in the respective reference trawls) were interpolated using kriging (see section 3.2.1) on the same interpolation grid as that used for spawning adults abundance and environmental variables.

Population size

Population size is the size of the population estimated each year. For plaice and whiting, it was estimated using the spawning stock biomass (in tonnes) of the year, as well as the total size (including juveniles) of the population (in number of individuals). These indices are also provided by the IBTS Working Group (ICES, 2008a). For anchovy, only the spawning stock biomass (in tonnes) was used as it is the same as the total stock; it is estimated, each year, by the ICES working group on anchovy (WGANC, ICES, 2008b).

Geographical location

Geographical location can be used to depict the spatial persistence of a population distribution over years, and corresponds to geographical coordinates. For plaice and whiting, the identity of the ICES statistical rectangles (fig. 2.1) were used as a categorical variable to represent the geographical location, respectively 114 and 157 rectangles for plaice and whiting. For anchovy, geographical position of each cell of the interpolation grid was used (fig. 2.2, see table 2.2 for the number of grid cells for each year). Geographical location models are time-invariant and assume that spatial distribution can be modelled without reference to any specific year.

Population memory

Population memory models depict the degree of persistence of population distributions from one year to the next, and indicate that the current spatial distribution depends on past distributions. Contrary to geographical models (see above), these models are not time-invariant, but instead, they assume some degree of persistence in the spatial distribution over years. Population memory was taken as the observed spawning adult abundance of the preceding year. Each station (trawling stations or grid cell) was allocated the observed abundance of the previous year using the spatial joining tools of ArcMap 9.1. For positions for which there was no observation during the preceding years (e.g. due to poor sea conditions), the observed abundance of the closest location was used. This reduced the total available number of stations from 7,317 and 9,435 to 6,313 and 8,098 for plaice and whiting, respectively, from 1981 to 2007, and from 1,106 to 946 for anchovy, from 2001 to 2007.

2.4.2. Model structure

Generalised additive models (GAM, Hastie and Tibshirani, 1990; Guisan *et al.*, 2002; Wood and Augustin, 2002) have been used for building the models. They were first introduced in terrestrial ecology (Yee and Mitchell, 1991) and later in marine ecology (Lehmann, 1998; Fox *et al.*, 2000). They are often used because they are data-driven rather than model-driven, and can relate in a non-linear way a biological response to several explanatory variables using smoothing functions (3).

$$(3)\ g(E[Y]) = \beta_0 + f_1 X_1 + f_2 X_2, ..., + f_n X_n = \beta_0 + \sum_{i=1}^{n} f_i X_i$$

where $g()$ is the link function, Y the response, β_0 the intercept, X_{1-n} the explanatory variables and f_{1-n} the smoothing functions.

GAM have been proved as robust as other regression technique such as Generalised Linear Models (GLM, McCullagh and Nelder, 1989) and are not more computationally intensive nor time-consuming compared with other modelling methods (Leathwick *et al.*, 2006a). As it is often the case with ecological data, the abundance of spawning adults displayed zero-inflated distributions. As a result, the modelling procedure was split into two steps. A binomial model with a logit link and a Gaussian model with an identity link were built using presence-absence data and non-null log-transformed abundance, respectively. For binomial models, several hypotheses of control were confronted: geography, environment, the three spatial sub-models, annual demography, population size and memory. For Gaussian models, spatial demography was also used. Probabilities of presence predicted by binomial models and log-transformed abundance predicted by Gaussian models were combined in a delta approach in which one was multiplied by the other (Stefánsson, 1996; Le Pape *et al.*, 2003) so as to predict the spatial distribution.

2.4.3. Model fitting and prediction

Models were fitted (i.e. adjusted) on one portion of the dataset (called calibration dataset) and applied for prediction on the remaining portion of the dataset (called prediction dataset). Degrees of smoothing (using a smoothing spline) were determined simultaneously for all the variable of the same hypothesis. Four degrees of smoothing (1 to 4) were tested for each explanatory variable. Finally, for each class of model (either binomial or Gaussian), all the combinations of the hypotheses were calculated resulting in 2^n-1 candidate models (with n being the number of hypotheses). Models were implemented using the R open source. The ‘gam’ package (Hastie, 2006) was used to construct binomial and Gaussian models. The ‘step.gam’ function of the MASS library (Venables and Ripley, 2002) was used to adjust the degrees of smoothing of the explanatory variables.

For plaice, models were confronted according to two strategies, so as to see the effect of the dataset on models and hypotheses selection. For strategy (a) called ‘Predict 2000s’, models were fitted on data from the 1980s and 1990s and applied on data from the 2000s. In total, 4,410 and 3,190 trawl hauls were used for fitting binomial and Gaussian models, respectively. For the prediction stage, 1,088 trawl hauls (binomial models) and 843 (Gaussian models) were used. For strategy (b) called ‘Predict 1980s’, models were fitted on data from the 1990s and 2000s and applied on data from the 1980s. In total, 4,268 and 3,229 trawls were used to fit binomial and Gaussian models. For the prediction stage, 962 stations were used for binomial models and 647 stations for Gaussian models.

For building the calibration and prediction datasets of whiting, data were selected so as to obtain a larger prediction dataset, but also, so as keep model prediction within the environmental and biological ranges covered by model adjustment. Nine years (1986-1989 and 2003-2007) were used for the calibration dataset, which corresponds to 2,943 and 2,753 stations used for fitting binomial and Gaussian models. For the prediction stage, the remaining 18 years (1981-1985 and 1990-2002) were used for the prediction dataset, which represents 4,150 stations used for binomial models and 3,894 for Gaussian models.

For anchovy, the data used for the calibration and prediction datasets were selected using the PCA outputs of the analysis described earlier in section 2.4.1 (fig. 2.4). Each year of data is represented by the centroid of its stations' scores, so as to illustrate its relative position in the ordinated multidimentional space. The years' centroid distribution highlighted the groupings of given years with similar environmental conditions. Four years (2001, 2003, 2004 and 2005) were used to build the calibration dataset, representing 571 and 295 stations used to fit binomial and Gaussian models, respectively. Three years (2002, 2006 and 2007) were used for the prediction dataset, which represents 375 and 195 stations, respectively, used for binomial and Gaussian models.

2.4.4. Model ranking

Models were confronted and selected by their ability to both fit and predict the spatial distribution of spawning adults. For model fitting, models were compared based on their goodness of fit using the Akaike Information Criteria (AIC, Akaike, 1974) and AIC differences (Δ_i, Burnham and Anderson, 2002; Johnson and Omland, 2004) which were calculated according to equations (4) and (5).

$$(4)\ AIC_i = -2 \times \log L_i + 2 \times p_i$$

where AIC_i is the AIC of the ith model, $\log L_i$ the log-likelihood of the ith model and p_i the number of parameter (explanatory variables) of the ith model.

$$(5)\ \Delta_i = AIC_i - AIC_{\min}$$

where Δ_i is the AIC difference for the ith model and $AIC_{\min}$ the lower AIC over all candidate models (the AIC of the "best" model). According to equations (4) and (5), the best among all candidate models is the one with the lowest AIC and with a Δ_i equal to 0; it is the model that displays the best adjustment to observed data and parsimony in the number of explanatory variables. Models were ranked in increasing order of Δ_i. In the binomial model, the log-likelihood of equation (4) was calculated between predicted probabilities of presence (from 0 to 1) and binary observed presence-absence (0 or 1) using equation (6) (McCullagh and Nelder, 1989).

$$(6)\ -2 \times \log L_i = -2 \times (\sum_{j=1}^{n} Y_j \times \log \hat{Y}_{i,j} + (1 - Y_j) \times \log(1 - \hat{Y}_{i,j}))$$

where Y is the vector of observed presence-absence and $\hat{Y}_i$ the vector of predicted probabilities of presence of the ith model for j observations and predictions. In the Gaussian model, the log-likelihood was calculated between predicted and observed log-transformed abundance using formula (7) (McCullagh and Nelder, 1989).

$$(7)\ -2\times\log L_i = -2\times(\sum_{j=1}^{n}(Y_j - \hat{Y}_{i,j})^2/\sigma^2 - 2\times j\times\log(\sigma)+C)$$

with Y being the vector of observed non-null abundance, $\hat{Y}_i$ the vector of predicted abundance of the ith model, σ^2 the residuals sum of square between predicted and observed abundance, j the number of observations and C a constant.

Models were also compared in their ability to predict the observed distribution (when applied on the prediction dataset) using the log-likelihood between observations and prediction according equations (6) and (7). The log-likelihood value being negative, the best model was the one with the highest log-likelihood. Other models were compared to this "best model" using a likelihood ratio test (8) (LRT, Burnham and Anderson, 2002).

$$(8)\ LRT_i = -2\times(\log L_i - \log L_{\max})$$

with LRT_i being the likelihood ratio test of the ith model and $\log L_{\max}$ the log-likelihood of the best model (with a LRT equal to 0). Models were ranked according to their LRT from smallest to largest.

AIC values were calculated with the AIC function of the ‘stats’ package in R (R Development Core Team, 2008). The log-likelihood was calculated using the ‘dbinom’ and ‘dnorm’ functions respectively in R (Bolker, B., pers. com., 2008).

2.4.5. Model evaluation

Binomial and Gaussian models selected based on the likelihood ratio test were re-applied to the corresponding prediction dataset, within the same data range as that used to fit Gaussian models (which are more restrictive as they are based on fewer observations). Distributions predicted by delta models, originating from the combination of selected binomial and Gaussian models, were graphically compared to observed distributions by the means of a Taylor diagram (fig. 2.5, Taylor, 2001) and the use of three similarity indices. The latter were the standard deviation, the root mean squared error (RMSE) and the Spearman correlation coefficient. A plot of the standard deviation indicates whether the model is able to reproduce the same variability in its prediction as those variations in the observations. The RMSE was computed as the root of the mean of the squared differences between each prediction and each observation. It incorporates both the variance of the model and its bias. The Spearman correlation coefficient and the RMSE quantify the correspondence between the observed and predicted patterns. Observations were plotted as a point called the reference point and were normalised so that their standard deviation was equal to 1, their correlation equal to 1 (the correlation between the observations and themselves) and their RMSE equal to 0 (the difference between the observations and themselves). Values of the indices for the predictions were plotted with the standard deviation on the x-axis, the correlation on the radials of the circle and the RMSE on concentric circles around the reference point. The Taylor diagram was computed using the 'plotrix' package in R (Lemon *et al.*, 2008).

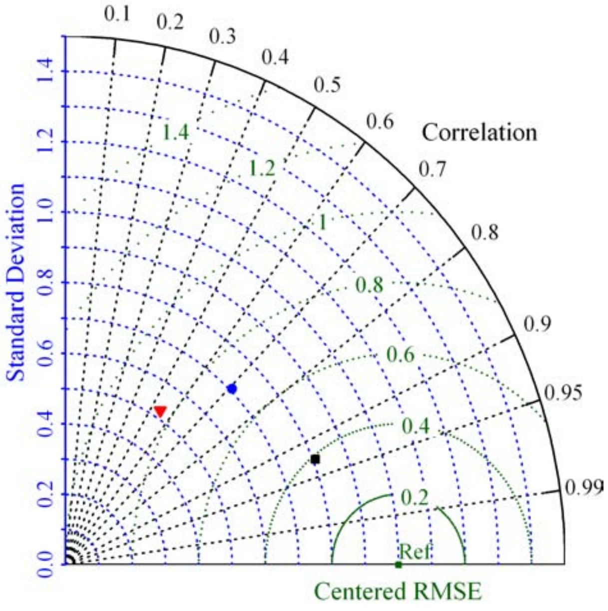


Figure 2.5. Example of a semi Taylor diagram. Predictions of three models are evaluated against the observations using the RMSE, the standard deviation and the Spearman correlation. These models are plotted according to their values for each of the three indices. Here, the 'black square' model is the best model as it is the closest of the 'Ref' point and has the lower RMSE and the highest standard deviation and correlation values.

2.4.6. Variance partitioning

Variance partitioning was used to explore the percentage of variation explained by each of the hypotheses within models selected from prediction. For each species, redundancy analysis (RDA, Legendre and Legendre, 1998) was used to relate the retained hypotheses to presence-absence and positive abundance of the prediction dataset. Although RDA is the constrained version of PCA (in that the ordination axes are constrained to become linear combinations of the variables), it as been used here in a single response context, making it equivalent to the multiple regression model described in the PCNM forward selection procedure (section 2.4.1). Each hypothesis was tested successively and those that were determined as significant ($p<0.05$) were kept for partial analysis. Partial analysis is a way of estimating how much of the variation of the response variable (either presence-absence or positive abundance) can be attributed exclusively to one hypothesis (variable), once the effect of the other hypotheses (covariables) has been removed (Borcard *et al.*, 1992). For a set of tested hypotheses, each variable was firstly regressed linearly by the covariables to partial out the shared part of variation. Once the shared component had been extracted, the partialled out variable was linearly regressed on the response to measure the amount of variation it purely explained on its own. Hence the amount of variation attributable or shared among the various hypotheses tested could be measured. The ‘rda’ and ‘varpart’ functions of the ‘vegan’ package (Oksanen *et al.*, 2009) of R were used to perform the Monte Carlo permutation test and partial analysis.

CHAPITRE 3

Spawning distribution of a flatfish: North Sea plaice

3.1. Introduction

The common Plaice (*P. platessa)* is one of the most studied flatfish species in the North Sea because of its economical importance for several countries (Rijnsdorp and Millner, 1996). Its life cycle has been widely described (Cushing, 1990; Wegner *et al.*, 2003). The spawning period is known for the English Channel and North Sea from late December to April with a peak of spawning in January-February. Locations of spawning grounds have been studied using eggs surveys (Harding *et al.*, 1978; ICES, 2005). The adults spawn near the bottom mainly during night (Simpson, 1971; Nichols, 1989). Because of their positive buoyancy, eggs are pelagic and can be found mainly in the upper layer of the water column (Coombs *et al.*, 1990). Eggs and larvae derived to the coastal nursery grounds were juveniles will develop during one year (Talbot, 1977). They will progressively leave the coastal zone to go offshore and be recruited to the adult population where they will become first time spawners at the age 4-5 years for females and 2-3 years for males (Rijnsdorp, 1989). As the capacity of the plaice to maintain and replace its population mainly depends on its ability to end up its life cycle, pelagic larvae and juveniles have been considered as the main critical stages (van der Veer *et al.*, 1990; van der Veer *et al.*, 1998; Pastoors *et al.*, 2000). However, little attention has been paid to adult population except with some recent studies related to population structure (Hunter *et al.*, 2004; Metcalfe, 2006) and migrations (Hunter *et al.*, 2003; Hunter *et al.*, 2004).

The aim of this chapter is to study the spatial distribution of spawning plaice in the North Sea. In section 3.2, distribution of spawning adults has been described and mapped in order to study its temporal variability during the last thirty years. In section 3.3., the spatial distribution of spawning adults has been related to environmental variables through habitat modelling to study how the spawning habitat of North Sea plaice has varied among years. In section 3.4, results are provided on the multi-model approach to identify the hypotheses that influence the spatial distribution of spawning plaice. In section 3.5, results of the three previous parts are discussed.

This chapter is based on three articles:

(2) C. Loots, S. Vaz, B. Planque, P. Koubbi
Spatial distribution of spawning population of
North Sea plaice and whiting from 1980 to 2007
Article in preparation

(3) C. Loots, S. Vaz, P. Koubbi, B. Planque, F. Coppin, Y. Verin
Inter-annual variability of North Sea plaice potential spawning habitat.
Journal of Sea Research. Submitted.

(4) C. Loots**,** S. Vaz, B. Planque, P. Koubbi
What control the spatial distribution of North Sea plaice spawning population.
Confronting ecological hypotheses through a model selection framework.
ICES Journal of Marine Science. In press.

3.2. Spatial distribution of spawning adults

The spawning distribution of plaice has been studied through the spatial distribution of spawning adults between 1980 and 2007. Spatial structures were described using geostatistical analyses (see section 2.3.1 for the methodology) and PCNM (see section 2.4.1 for the methodology). Abundances of spawning adults were interpolated for each year using kriging and were mapped into a geographical information system (see section 2.3.2 for the methodology). Interpolated abundances were then used to explore the inter-annual variability of spawning plaice distribution using a correspondence analysis during the last thirty years (see section 2.3.3 for the methodology). This section is derived from the article in preparation (2).

3.2.1. Proportions of spawning adults

Proportions of spawning adults were calculated per size class for areas 2, 3 and 4 and for areas 5, 6 and 7 (fig. 3.1). Proportions of spawning adults were higher among larger fish. They were also higher in the southern part (areas 5, 6 and 7) than in the northern part of the North Sea (areas 2, 3 and 4). In the northern part, proportions have reached a sill of 20%, whereas in the southern part it seems that the proportions still increased with the size of fish.

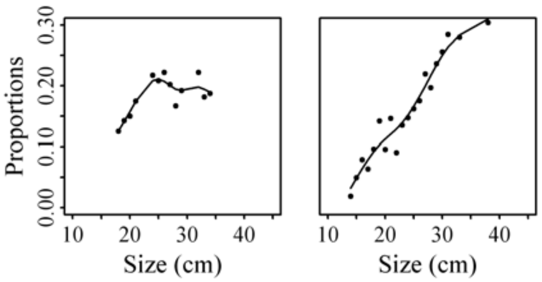


Figure 3.1. Proportions of adults of stage 3 (spawning adults) of plaice per size class. Left : areas 2, 3 and 4. Right : areas 5, 6 and 7. Symbols are the observed proportions and lines are the adjusted proportions using a smoothing spline for missing size class. See fig. 2.1 for location of the areas.

3.2.2. Geostatistics

Twenty eight variograms (one per year) were computed and adjusted on the abundances of spawning adults between 1980 and 2007. Values of the spatial trend, the sill and the nugget extracted from the variogram of each year are given in fig. 3.2. A significant quadratic trend was detected for years until 1994 (except 1982) and years 1995 and 1999 as it explained more than 20% of the variation in the data. For other years, percentage explained by the spatial trend was below 20 % at the end of the nineties and decreased until 10% in the

2000’s. Percentage explained by the spatial trend had the same evolution across years than that of the spawning stock biomass (SSB). There was an increase of the SSB close to 450,000 t during the eighties followed by a decrease at the beginning of the nineties. Since 1994, the SSB was less variable and was comprised between 150 and 250,000 t. For years 1980 and 1981, no spatial structure could be detected in the residuals and a pure nugget effect model was used. For other years, a combined model with a nugget component and a circular component were used to adjust experimental variograms. Nugget had values between 0.3 and 0.6 for the first four years and lower ones, between 0.1 and 0.3, for the following years. Sill was low for years 1982-1985 and 1990-1993 and higher for years 1986-1989 and 1994-2007. Ranges are not shown here as they depend if the spatial trend was removed or not. Therefore, ranges were not directly comparable between years (see appendix A.2.1).

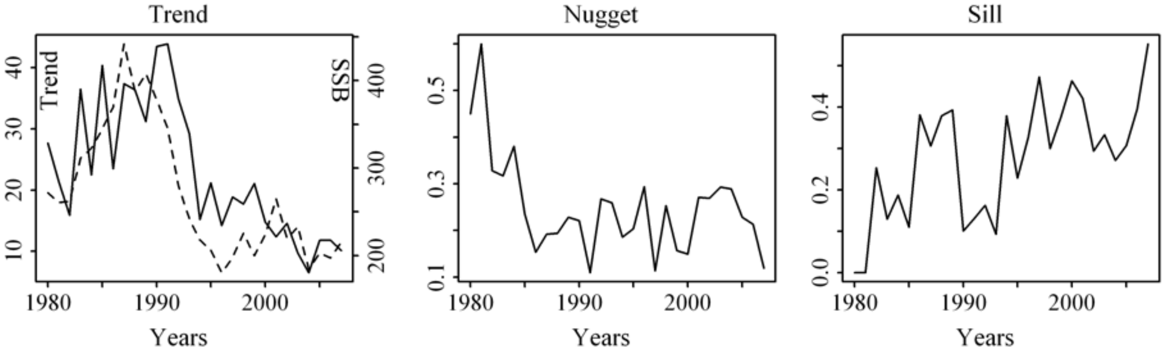


Figure 3.2. Parameters of the variogram. The spatial trend (in percentage explained), the nugget and the sill extracted from the theoretical variogram are plotted for each year between 1980 and 2007. ICES estimation of spawning stock biomass (SSB, 10^3 tons, dotted line) are also indicated.

3.2.3. Distribution of spawning adults

The theoretical variogram adjusted for each year was used to map the spatial distribution of spawning adults using kriging interpolation (fig. 3.3). Maps indicate areas of high abundances of adults were located in the southern part of the North Sea and along the east coast of United Kingdom. Very low abundances were found in the central part of the North Sea. The ratio of the spatial variance over the temporal variance was 2.64, indicating that the pattern of distribution of spawning plaice was quite constant in time. However, there was also a temporal variation in both the maximal abundance observed and in the spatial extent of high abundances. Abundances were higher and occupied a wider area at the end of the eighties than in the nineties and 2000’s.

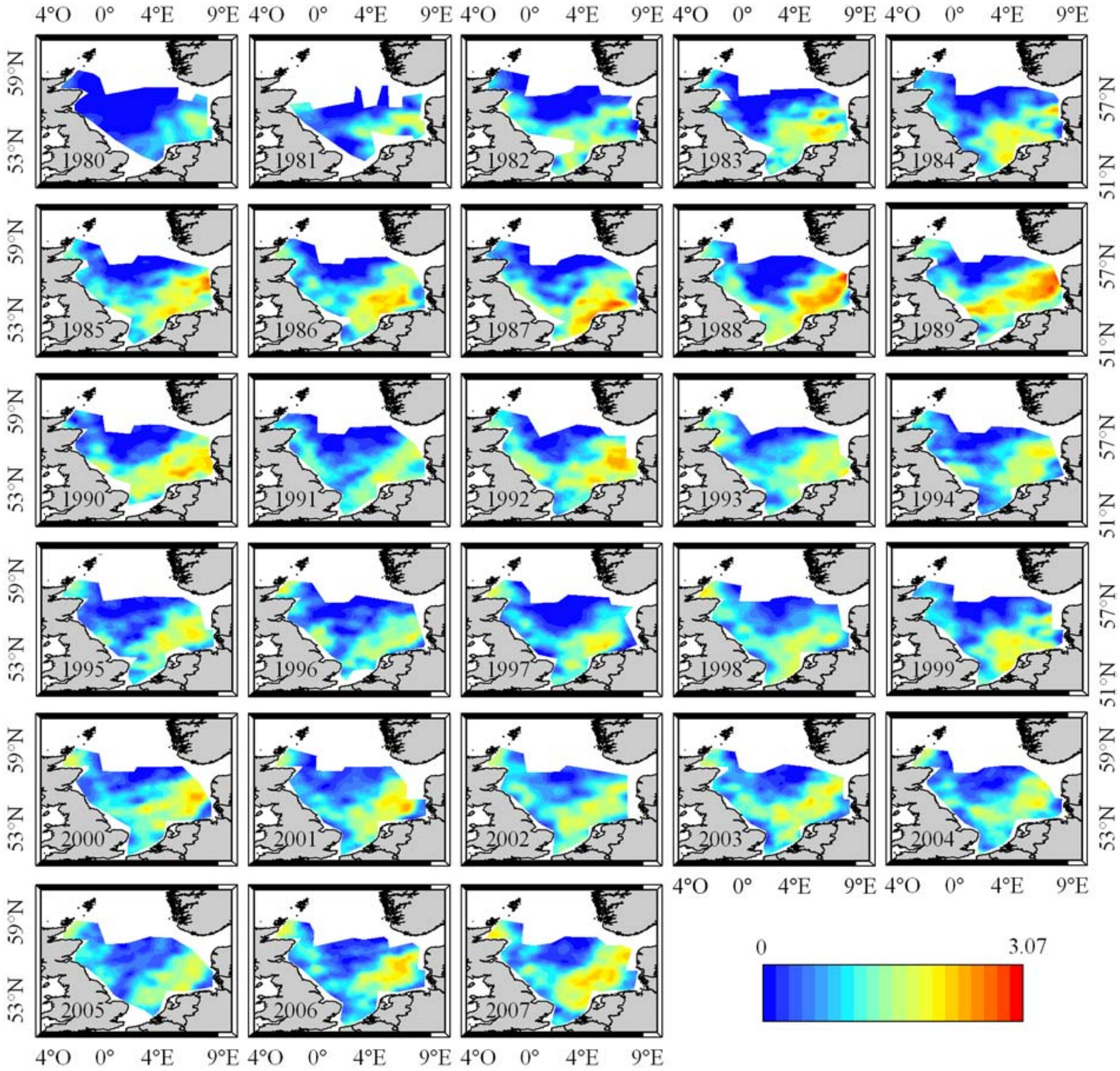


Figure 3.3. Spatial distribution of spawning adults of North Sea plaice. Maps were computed from the abundances of spawning adults for each year between 1980 and 2007. Abundances (log-transformed) are classified into twenty classes of equal interval between 0 and 3.07.

Distribution maps of each year were used to calculate average and variability maps. Thresholds of 1.05 for the mean and 0.1 for the standard deviation were used to combine these maps to define preferential, occasional and unfavourable spawning sites for plaice (fig. 3.4). Preferential spawning sites were located in a narrow area just off the northern coast of Netherlands. Occasional spawning sites were surrounding preferential spawning sites and were located in the southern and western part of the North Sea. In particular, they were located in the Southern and German Bights areas, in the Transition area in the south of the

Dogger Bank and in the Flamborough Head and Moray Firth. Unfavourable spawning sites were always located in the central part of the North Sea.

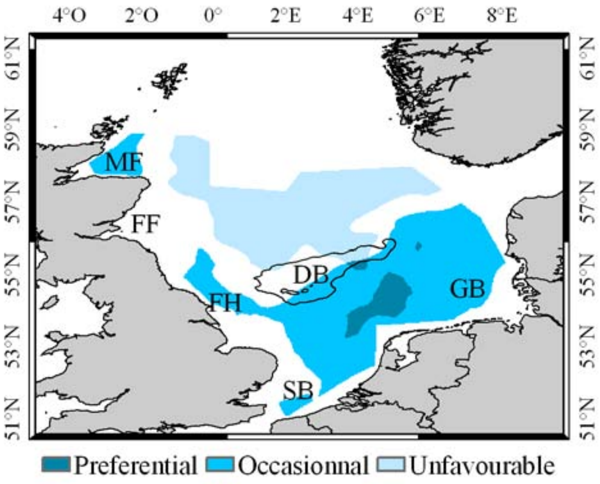


Figure 3.4. Location of preferential, occasional and unfavourable sites for spawning plaice. Map was computed from average and variability maps calculated from maps of spatial distribution of each year between 1980 and 2007 (fig. 3.5). SB = Southern Bight, GB = German Bight, FH = Flamborough Head, DB = Dogger Bank, FF = Firth of Forth, MF = Moray Firth.

PCNM were used to describe the structure in the spatial distribution of spawning plaice (fig. 3.5).

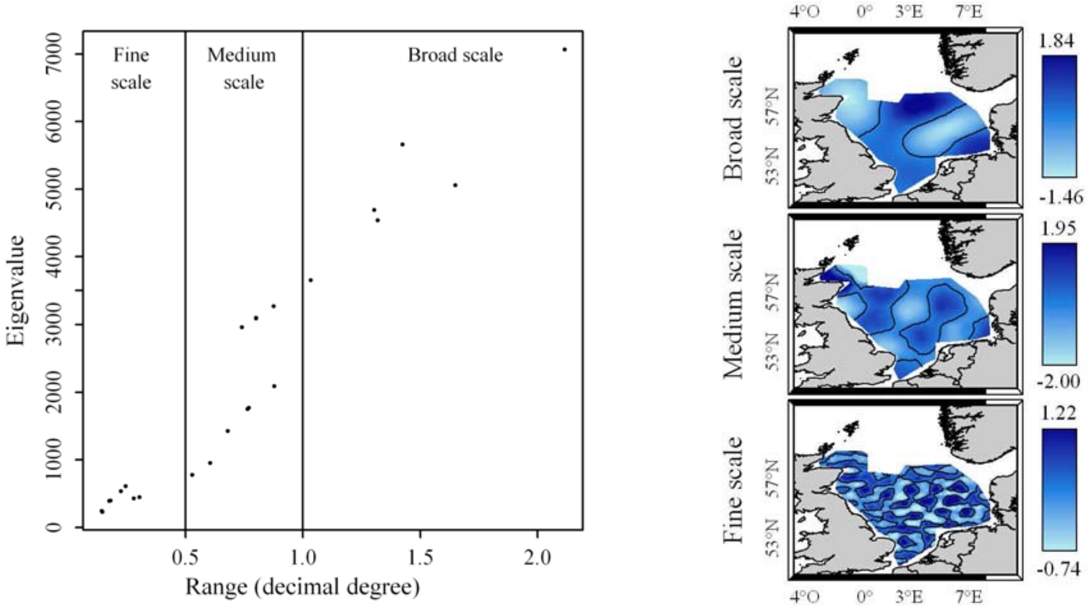


Figure 3.5. PCNM and spatial sub-models. Left : plot of the range values of the 23 selected PCNM against their eigenvalue. Range values were obtained by fitting a Gaussian model to the variogram computed on the eigenvectors of each PCNM. Three scales were identified : broad scale : eigenvectors with a range between 111 and 233 km (1 to 2.1 decimal degrees of latitude); medium scale : eigenvectors with a range between 55 km and 111 km (0.5 to 1 decimal degree of latitude); fine scale : eigenvectors with a range between 0 and 55 km (less than 0.5 decimal degrees of latitude). Right : maps of the three spatial scales. PCNM 1, 18 and 49 were used to depict respectively broad, medium and fine scales. Black line indicates the zero value.

A threshold of 1.13 (decimal degrees of latitude) was used to truncate the distance matrix between stations, corresponding to four times the maximum distance between two neighbouring stations. Overall, 2506 positive PCNM were extracted from the principal correspondence analysis and 23 were determined as significant by the forward selection (p-value < 0.001 with an adjusted-R^2 of 0.26). From these 23 PCNM (see appendix B.1), three spatial scales were identified : broad scale (111-233 km), medium scale (55-111 km) and fine scale (< 55 km). Six PCNM accounted for the broad scale, nine for the medium scale and eight for the fine scale. The broad scale allowed separating the main areas of plaice distribution like the southern and central parts and Moray Firth. The medium scale described smaller areas like the Southern Bight, Firth of Forth and Flamborough Head regions. The fine scale reflected the sampling strategy and described the variability of plaice distribution within the statistical rectangles.

3.2.4. Inter-annual variability of spawning adults distribution

A correspondence analysis was performed on interpolated abundances of plaice to describe how its spawning distribution has varied during the last thirty years. The first axis explained 18.9 % and the second axis explained 9.4 % of the variability (fig. 3.6). Years were mainly distributed along the first axis with the eighties on the left part, the nineties in the middle part and the 2000’s on the right part. Scores of the years on axis 1 showed an increase of their value with time, with negative scores for years 1980 to 1992 and positive scores for the following ones. Distribution and scores of the years along the second axis were less clear.

Three groups of stations were identified using the cluster analysis. The first group gathered 240 stations. These stations are located on the south-eastern part of the North Sea and along the east coast of Scotland (fig. 3.6). This group of stations was located on the negative part of the first axis close to the eighties. The second group of stations gathered 349 stations. These stations surrounded the first group of stations and were more close to the nineties and 2000’s. The third group contained 103 stations that were located in the central pat of the North Sea. This group of stations was located on the right part of the first axis.

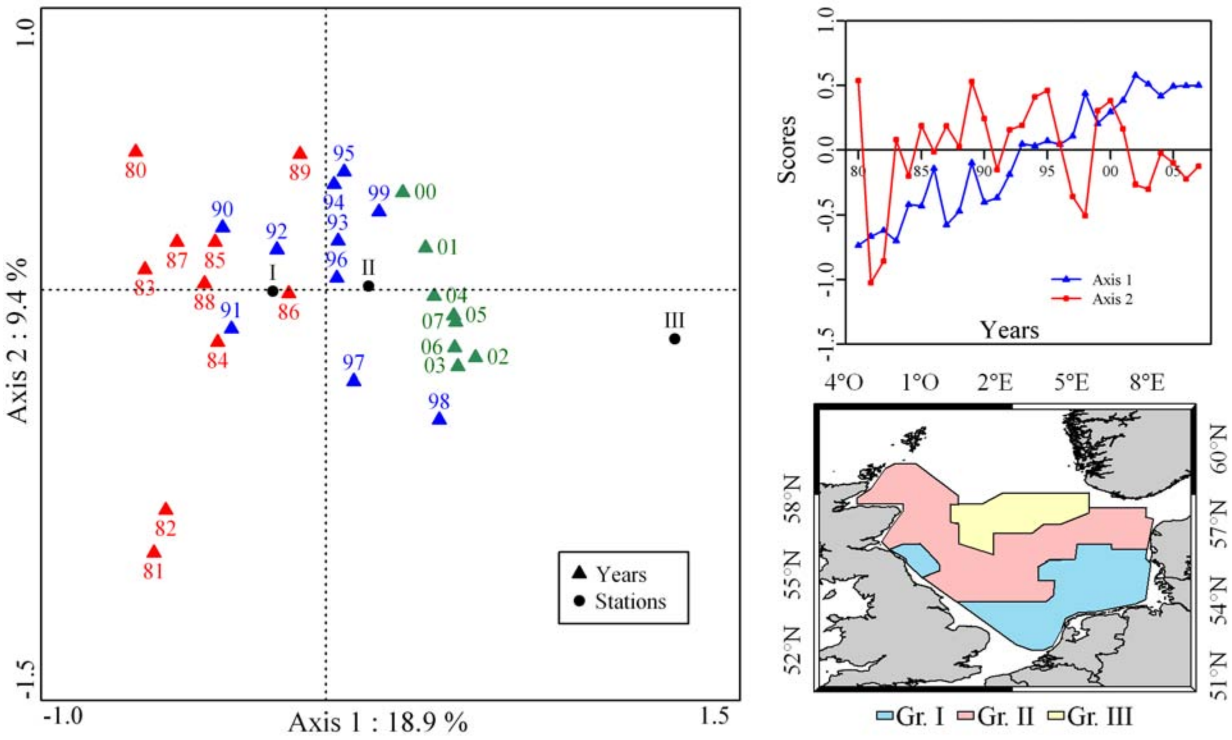


Figure 3.6. Correspondence analysis performed on abundances of plaice. Left : years and groups of stations are plotted on the factorial plan of axis 1 and axis 2. the three groups of stations are indicated by their centroids in the factorial plan. Top-right : scores of the years on axis 1 and 2. Bottom-right : location of the groups of stations.

3.3. Habitat modelling

An habitat model for North Sea plaice during its reproduction period time has been built and mapped for each year in order to see how it varies in time and to determine how this habitat model may reflect the observed distribution. The habitat is defined by environmental relationship between the species and the environment (Coyne and Christensen, 1997). This section was presented at the 7th international flatfish symposium (Sesimbra, Portugal) in 2008 (http://www.flatfish2008.fc.ul.pt/). This part is derived from article (3) submitted to Journal of Sea Research.

3.3.1. Methodology

Generalised additive models (see section 2.4.2) were used to relate the spatial distribution of North Sea plaice spawning adults to the five environmental variables : depth, bottom temperature and salinity, seabed stress and sediment types. Binomial and Gaussian models were adjusted using data from 1980 to 2007. Significance of each predictor was assessed in a stepwise selection procedure using the AIC (see section 2.4.4). In the stepwise selection, variables are successively tested by backward and forward elimination. The procedure stops when the AIC increases after removing an additional. Final binomial and Gaussian models were reapplied to the environmental values of each year from 1980 to 2007 and predicted distribution was obtained using a delta model. The delta model (see section 2.4.2) was also applied on the environmental values for the seventies that were extracted from the DATRAS database. High predicted values indicate high suitable areas whereas low values depict low suitable areas.

Evaluation of the habitat model has been made using the Taylor diagram (see section 2.4.5) where prediction of the habitat model were compared to observation for years 1980 to 2007. Predictions of the model were evaluated for all the years together and for each year separately in order to see how the model accuracy varies from year to year.

Predicted values of each year were interpolated and mapped on a regular grid using geostatistics. A contingency table of 38 years and 678 points was extracted from the interpolation grid of predicted values of the habitat model. A cluster analysis based on a Khi² distance and using a flexible link (β=-0.25) combined to a correspondence analysis was carried out on years 1980-2007 to study the temporal variations in the spatial distribution of spawning habitat of plaice. Because there was no observed distribution for the seventies, predictions of the habitat model for these years were used as supplementary variables in the correspondence analysis. In this way, they do not interfere in the building of the axes but can be projected on the factorial plane.

3.3.2. The habitat model

The five environmental variables were determined as significant in improving the adjustment of the model based on the AIC. They were kept at the end of the stepwise selection procedure for both the final binomial (1) and Gaussian (2) models.

$$(1)\ 0/1 = s(\text{Depth}, 4) + s(\text{Temperature}, 4) + s(\text{Salinity}, 4) + s(\text{Bedstress}, 4) + \text{as.factor}(\text{Sediments})$$

$$(2)\ \log(\text{Ab}>0) = s(\text{Depth}, 4) + s(\text{Temperature}, 3) + s(\text{Salinity}, 4) + s(\text{Bedstress}, 4) + \text{as.factor}(\text{Sediments})$$

The binomial model explained 17 % of the total deviance of the presence/absence data and the Gaussian model 21 % of the deviance of the positive abundance data. Plot of the fitted values for each smoothing terms (fig. 4) clearly showed a non-linear relationship with a degree 4 of smoothing except for bottom temperature for which a concave relationship was modelled using a degree 3 of smoothing. For each environmental variable, the response shape look like the same either for the binomial or the Gaussian model. High values of probability of presence or abundances were found for shallow depth (0-50 m), mean salinity (around 34.3), low bedstress (close to 0) and relatively high bottom temperature (4-8°C).

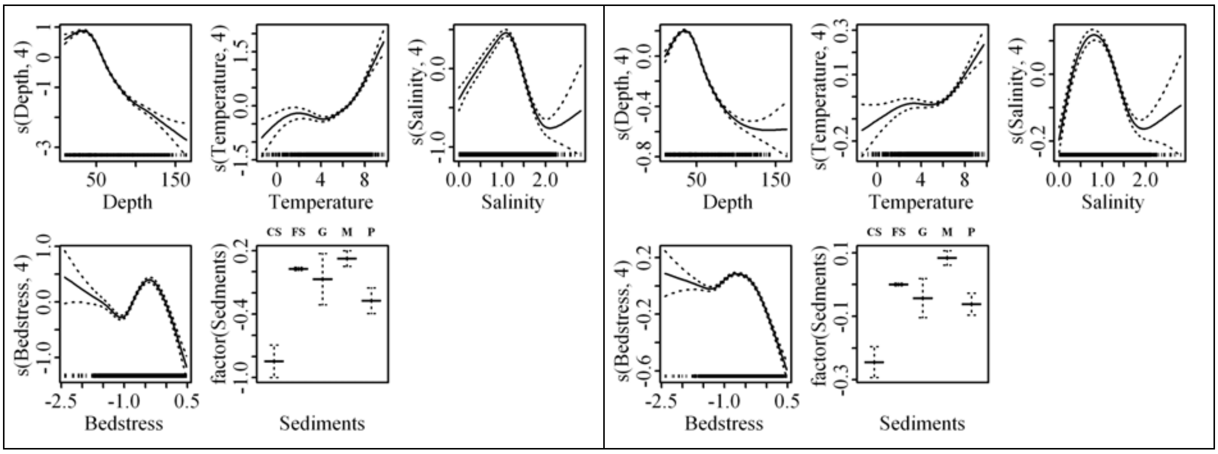


Figure 3.7. Response curves of the habitat model of plaice for the five environmental variables. Left : binomial model. Right : Gaussian model. Smoothing degree is indicated in bracket. Solid line denotes fitted values and dotted lines denotes the 95% confidence interval. CS=coarse sand, FS=fine sand, G=gravels, M=mud, P=pebbles.

3.3.3. Evaluation of the habitat model

Evaluation of the habitat model by the Taylor diagram is given in fig. 3.8. Evaluation of the predictions for each year surrounded the mean evaluation of the model, indicating a difference in accuracy of the model depending on the year that is considered. The correlation between predictions and observations was lower for the nineties than for the eighties and lower for the 2000’s than for the nineties. Predictions of the 2000’s had a higher RMSE value than the nineties and the eighties. This means that there was a general decrease in model accuracy along the three decades indicating that the model was less able to reproduce observed pattern of the 2000’s than for the nineties or the eighties.

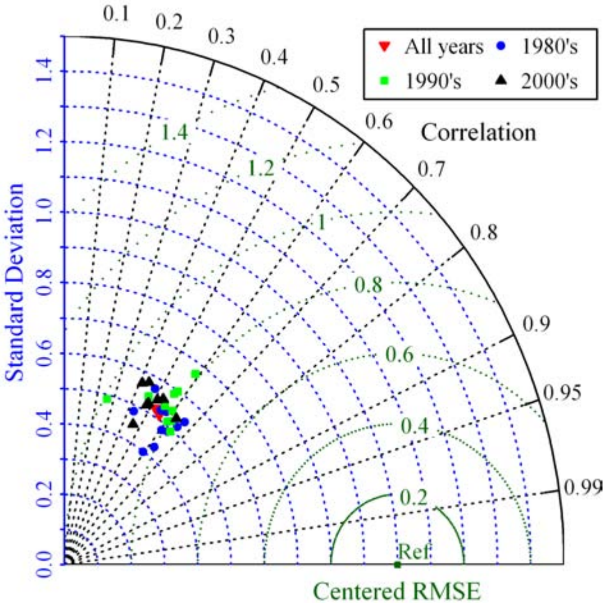


Figure 3.8. Evaluation of the habitat model by the semi Taylor diagram (only the positive correlation part). Observations are symbolised as a point (Ref) and are normalised so that their standard deviation is equal to 1, their correlation equal to 1 (the correlation between the observations and themselves) and their RMSE equal to 0 (the difference between the observations and themselves). Values of the indexes for the predictions are plotted with the standard deviation on the x-axis, the correlation on the radials of the circle and the RMSE on concentric circles around the reference point. The three indexes were calculated on predictions for all the years together and for each year separately.

3.3.4. Temporal variations of plaice spawning habitat

The classification allowed to separate four groups of samples at the threshold of 20%. These groups were plotted on the first two axes of the correspondence analysis (fig. 3.9). The first axis explained 23 % of the variation and the second axis explained 11.7 %. Groups of stations were mainly distributed along the first axis. The first group is plotted on the left part of the first axis. It gathered samples which were located on the south-eastern part of the North Sea, in the Flamborough Head along the English coast and in the Firth of Forth and Moray Firth along Scotland (fig. 3.9). These samples corresponded to high predicted abundance values indicating highly suitable areas. The second group was close to the first group on the first axis, it surrounded areas of the first group and had lower values of predicted abundance (i.e. less suitable habitat). The third and fourth group were on the right part of the first axis and gathered areas that were located respectively in the middle and north western parts of the North Sea. These areas showed very low or almost null predicted abundance values, indicating non-suitable habitats. Therefore, these four groups along axis 1 on the first axis were explained by a latitudinal gradient with samples of high values close to the coast and samples of lower values in the central part of the North Sea.

Three groups of year were identified using a threshold of 20% and were plotted on the same factorial plan as for the groups of samples (fig. 3.9). These groups of years were also distributed along the first axis. The first group was located on the left part of the first axis and was close to the first group of samples. It gathered mainly the eighties and the year 1991. The second group mainly gathered the nineties and the beginning of the 2000's and was located in the middle of the factorial plane close to the second group of samples. The third group gathered particular years and was located on the right part of the first axis closed to the third and fourth groups of samples. The seventies, used as supplementary variables, were distributed on the left part of the first axis closed to the eighties except for years 1973, 1978 and 1979 which were more closed to the second group of years.

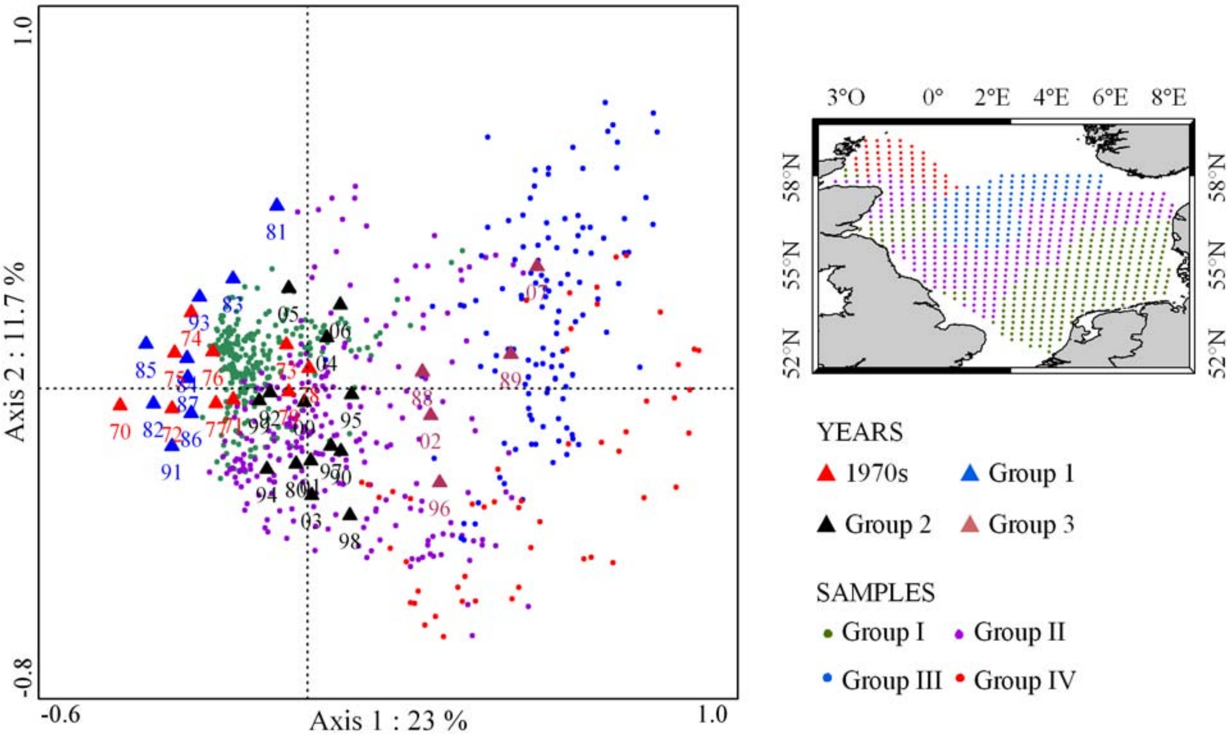


Figure 3.9. Correspondence analysis performed on predictions of the habitat model. Left : years and samples are plotted on the factorial plane of axes 1 and 2. Right : location of the groups of samples on the map of the North Sea.

Habitat maps for the three years (1987, 1997, 2007) belonging to each group and the supplementary variables (1970) are presented in fig. 3.10. It illustrates suitable spawning areas for each group of years that were identified in the correspondence analysis. These maps show that (1) there was a spatial stability of the predicted suitable areas during the three decades in the southeastern part and along the English and Scottish coasts and (2) there was also a spatial expansion of suitable areas between 1970 and 2007. This expansion began with the Moray Firth and Firth of Forth that became more suitable in the eighties compared to the seventies. This was more pronounced in the nineties with a wider suitable area along the English coast that also began to expand in the central part. In the 2000's, the central part was a more suitable spawning areas than in previous years. Maps of observed distribution for the same years (fig. 3.3 and 3.10) indicate that higher values of the observed abundances were still located in the areas of higher suitability during the three decades but did not extend more Northerly as supposed by the prediction of the habitat model.

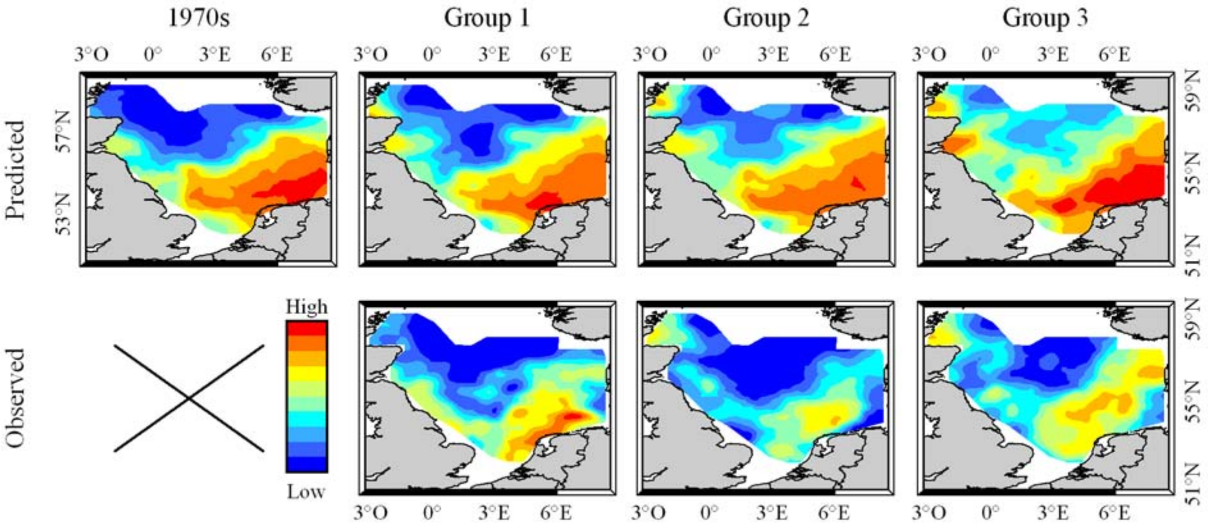


Figure 3.10. Maps of predicted and observed distribution of spawning plaice by the habitat model for the three groups of years determined by the correspondence analysis. 1987, 1997 and 2007 were used as illustrative years for each of the three groups. Year 1970 was used to illustrate the potential habitat distribution for 1970s used as supplementary variables in the correspondence analysis. All maps of predictions are on the same scale as well as all maps of observed distribution.

3.4. Multi-model approach

Multi-model approach has been used to identify among the ten hypotheses (geographical location, persistence and non persistence of the environment, the three spatial scales, populations size, annual and spatial demography and population memory) those that play a role in the spatial distribution of spawning plaice. Nine hypotheses were confronted using binomial models and ten using Gaussian models (see formulas in table 3.1). Combination of these hypotheses has resulted respectively in 511 and 1023 possible models. In strategy (a) ‘Predict 2000s’, models were fitted on the eighties and nineties and applied on 2000’s. In strategy (b) ‘Predict 1980s’, models were fitted on the nineties and 2000’s and applied on the eighties. Models were compared and ranked according to their AIC in the case of fitting and according to their LRT in the case of prediction. Only models with an AIC or LRT value between 0 and 10 were selected. This section is derived from the article (4) that is in press in ICES Journal of Marine Science.

Table 3.1. Formulas of the various hypotheses used in binomial and Gaussian models for strategy (a) 'Predict 2000s' and (b) 'Predict 1980s'.

Strategy (a) Predict 2000s

Hypotheses	**Binomial models**	**Gaussian models**
G	as.factor(StatisticalRectangle)	as.factor(StatisticalRectangle)
Ep	s(Depth, 4) + s(Bedstress, 4) + as.factor(Sediments)	s(Depth, 4) + s(Bedstress, 4) + as.factor(Sediments)
Enp	s(Temperature, 4) + s(Salinity, 4)	s(Temperature, 4) + s(Salinity, 4)
Bs	s(Ax1, 4) + s(Ax3, 4) + s(Ax4, 4) + s(Ax5, 4) + Ax6 + s(Ax8, 4)	Ax1 + s(Ax3, 4) + s(Ax4, 4) + s(Ax5, 4) + s(Ax6, 4) + s(Ax8, 4)
Ms	Ax9 + s(Ax10, 4) + s(Ax11, 4) + s(Ax14, 4) + s(Ax15, 4) + s(Ax16, 4) + s(Ax18, 1) + s(Ax21, 4) + s(Ax23, 4)	s(Ax9, 4) + s(Ax10, 4) + Ax11 + s(Ax14, 4) + s(Ax15, 4) + s(Ax16, 4) + s(Ax18, 4) + s(Ax21, 4) + s(Ax23, 2)
Fs	s(Ax26, 4) + s(Ax32, 4) + s(Ax41, 4) + s(Ax43, 4) + s(Ax47, 3) + s(Ax49, 4) + s(Ax84, 4) + s(Ax88, 3)	s(Ax26, 4) + s(Ax32, 4) + s(Ax41, 4) + s(Ax43, 4) + s(Ax47, 4) + s(Ax49, 4) + s(Ax84, 4) + Ax88
Ps	s(PopulationSize, 4) + SpawningStockBiomass	s(PopulationSize, 4) + s(SpawningStockBiomass, 4)
Da	s(Age1a, 2) + s(Age2a, 4) + s(Age3a, 4) + Age4a + Age5a	Age1a + s(Age2a, 4) + s(Age3a, 4) + Age4a + s(Age5a, 4)
Ds	-	s(Age1s, 4) + s(Age2s, 4) + s(Age3s, 4) + s(Age4s, 4) + s(Age5s, 4)
M	s(PreviousYearAbundance, 4)	s(PreviousYearAbundance, 4)

Strategy (b) Predict 1980s

Hypotheses	**Binomial models**	**Gaussian models**
G	as.factor(StatisticalRectangle)	as.factor(StatisticalRectangle)
Ep	s(Depth, 4) + s(Bedstress, 4) + as.factor(Sediments)	s(Depth, 4) + s(Bedstress, 4) + as.factor(Sediments)
Enp	s(Temperature, 4) + s(Salinity, 3)	s(Temperature, 4) + s(Salinity, 4)
Bs	s(Ax1, 4) + Ax3 + s(Ax4, 4) + s(Ax5, 4) + s(Ax6, 4) + s(Ax8, 4)	s(Ax1, 4) + s(Ax3, 4) + s(Ax4, 4) + s(Ax5, 4) + s(Ax6,) + s(Ax8, 4)
Ms	Ax9 + s(Ax10, 4) + s(Ax11, 4) + s(Ax14, 4) + s(Ax15, 4) + s(Ax16, 4) + s(Ax18, 1) + s(Ax21, 4) + Ax23	s(Ax9, 3) + s(Ax10, 4) + s(Ax11, 4) + s(Ax14, 4) + s(Ax15, 4) + s(Ax16, 4) + s(Ax18, 4) + s(Ax21, 4) + Ax23
Fs	s(Ax26, 4) + s(Ax32, 4) + s(Ax41, 4) + s(Ax43, 4) + s(Ax47, 4) + s(Ax49, 4) + s(Ax84, 4) + s(Ax88, 4)	s(Ax26, 4) + s(Ax32, 4) + s(Ax41, 4) + s(Ax43, 4) + s(Ax47, 4) + s(Ax49, 3) + s(Ax84, 4) + s(Ax88, 3)
Ps	PopulationSize + s(SpawningStockBiomass, 4)	PopulationSize + s(SpawningStockBiomass, 4)
Da	Age1a + s(Age2a, 4) + s(Age3a, 3) + Age4a + Age5a	Age1a + Age2a + s(Age3a, 4) + Age4a + s(Age5a, 4)
Ds	-	s(Age1s, 4) + s(Age2s, 4) + s(Age3s, 4) + s(Age4s, 4) + s(Age5s, 4)
M	PreviousYearAbundance	s(PreviousYearAbundance, 4)

G = geography, Ep = persistent environment, Enp = non-persistent environment, Bs = broad scale sub-model, Ms = medium scale sub-model, Fs = fine scale sub-model, Ax = number of the PCNM, Da = annual demography, Age1a = annual proportion of age 1, Age1s = spatial proportion of age 1, Ds = spatial demography, M = population memory, "s" means that a smoothing function has been used and the corresponding degree of smoothing is indicated in brackets.

3.4.1. Characteristics of selected models

A total number of 41 models were selected (table 3.2) using a selection threshold of 10 for the Δ_i and LRT. Selected models were not the same depending on the selection criteria which means that best fitted models to the data were not those with the best predictive power. Models selected based on their predictive power were more numerous. Selected models were also not the same according to the strategy and type of model.

Table 3.2. Formulas of selected models. Models were selected according to their capability to fit (Δ_i) and predict (LRT) the observed distribution. Models selected from fitting are those with a Δ_i lower than 10 and models selected from prediction are those with a LRT lower than 10. In the case of selection based on fitting, the LRT is also given as an informative value as well as the Δ_i in the case of selection based on prediction. Formulas are given for selected binomial and Gaussian models for strategy (a) ‘Predict 2000s’ and strategy (b) ‘Predict 1980s’. See table 3.1 for formulas of the hypotheses.

Strategy (a) Predict 2000s

Selection	Model	Δ_i	LRT
Fitting (n=6)	***Binomial (n=4)***		
	Ep + Enp + Bs + Ms + Ps + Da + M	0	40.9
	Ep + Enp + Bs + Ms + Da + M	0.7	71.6
	Ep + Bs + Ms + Ps + Da + M	3.9	39.7
	Ep + Bs + Ms + Da + M	8.9	103.5
	Gaussian (n=2)		
	Ep + Enp + Bs + Ms + Ps + Da + Ds + M	0	157.3
	Enp + Bs + Ms + Ps + Da + Ds + M	7.5	153.6
Prediction (n=13)	***Binomial (n=9)***		
	Ep + Enp + Bs + Ps + M	130.3	0
	Ep + Enp + Bs + M	149.4	1.2
	Ep + Enp + Bs + Fs + Ps + M	128.3	3.6
	Ep + Enp + Bs + Fs + M	148.2	4.6
	Enp + Bs + Ms + Ps + M	74.3	8.2
	Enp + Bs + Ms + M	97.5	8.6
	Ep + Enp + Bs + Ms + Ps + M	56.3	9.1
	Ep + Enp + Bs + Ms + M	81.2	9.5
	Enp + Bs + Ms + Fs + Ps + M	87.1	9.6
	Gaussian (n=4)		
	Bs + Ms + Fs + Ps + Ds + M	55.2	0
	Ep + Bs + Ms + Fs + Ps + Ds + M	50.6	1.5
	Bs + Ms + Ps + Ds + M	45.5	4.9
	Ep + Bs + Ms + Ps + Ds + M	38.1	6.0

Strategy (b) Predict 1980s

Selection	Model	Δ_i	LRT
Fitting (n=9)	***Binomial (n=5)***		
	Ep + Enp + Bs + Ms + Ps + M	0	61.4
	Ep + Bs + Ms + Ps + M	2.3	79
	Ep + Enp + Bs + Ms + Fs + Ps + M	6.6	66.3
	Ep + Bs + Ms + Fs + Ps + M	7.8	83.2
	Ep + Bs + Ms + Ps + Da + M	8.9	40.5
	Gaussian (n=4)		
	Ep + Enp + Bs + Ms + Fs + Ps + Da + Ds + M	0	28
	Ep + Bs + Ms + Fs + Ps + Da + Ds + M	3	13.3
	Ep + Enp + Bs + Ms + Fs + Ps + Ds + M	3.8	31.4
	Ep + Bs + Ms + Fs + Ps + Ds + M	6.8	18.5
Prediction (n=13)	***Binomial (n=1)***		
	Ep + Ms + Fs + Ps + Da + M	149.7	0
	Gaussian (n=12)		
	G + Ep + Ms + Ps + Da + Ds + M	105.5	0
	G + Ms + Ps + Da + Ds + M	109.7	2
	Ep + Bs + Ms + Ps + Da + Ds + M	23.4	3.3
	G + Ep + Ps + Da + Ds + M	82.6	4.7
	G + Ep + Bs + Ms + Ps + Da + Ds + M	118.3	4.7
	G + Bs + Ms + Ps + Da + Ds + M	120.1	6.8
	G + Ep + Bs + Ps + Da + Ds + M	83.8	7
	G + Ep + Ms + Ps + Ds + M	106.8	7.3
	G + Ep + Ms + Ps + Da + Ds	157.9	7.8
	G + Ps + Da + Ds + M	88.1	8.2
	Bs + Ms + Ps + Da + Ds + M	45.1	8.9

Selected models also differed in their complexity (fig. 3.11). Models selected based on their predictive power were less complex in term of number of hypotheses and explanatory variables than those selected based on their adjustment to the data. Selected binomial models were also less complex compared to selected Gaussian models.

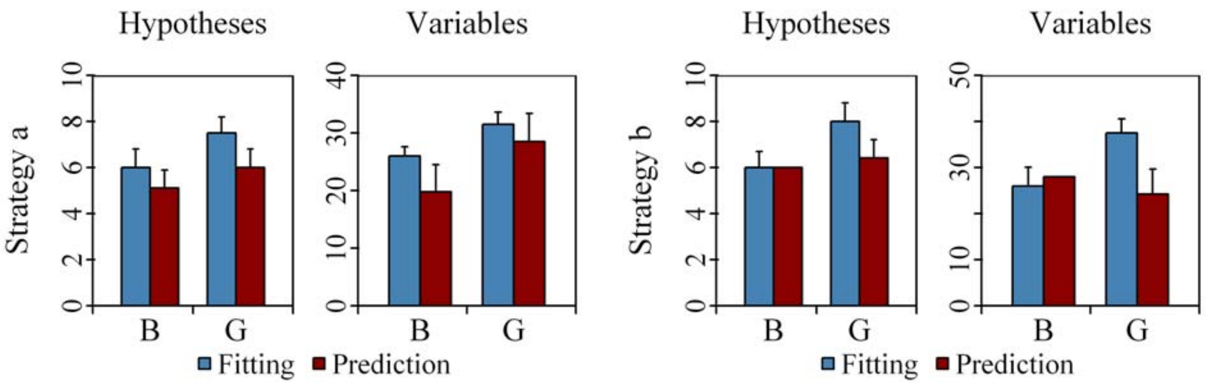


Figure 3.11. Complexity of selected models. Complexity is represented by the mean number of hypotheses and explanatory variables within binomial (B) and Gaussian (G) models. Complexity is given for models selected either on fitting or prediction and for strategy (a) 'Predict 2000s' and strategy (b) 'Predict 1980s'.

Frequencies of each hypothesis within the set of selected models have been calculated (table 3.3). Three hypotheses dominated: population size, population memory and spatial demography. Persistent environment and broad and medium scale sub-models were also important but less frequent. Geography, non-persistent environment, fine scale sub-model and annual demography showed strong differences in selected models according to the type of model (binomial of Gaussian) and selection procedure (fitting or prediction). Geography was more frequent in models selected from prediction and especially in Gaussian models. Non-persistent environment and fine scale sub-model were more frequent in Gaussian models selected from fitting and more frequent in binomial models selected from prediction. Annual demography was more frequent in selected Gaussian models than in selected binomial models.

Table 3.3. Frequencies of the various hypotheses within selected models. They give the percentage of model in which each hypothesis is present either for fitting based selection of prediction based selection. Frequencies are indicated for all models selected, for each model class (binomial or Gaussian) and each strategy (a 'Predict 2000s' and b 'Predict 1980s').

	Fitting based selection					**Prediction based selection**				
Hyp.	**All models**	**Strategy (a) Predict 2000s**	**Strategy (b) Predict 1980s**	**Binomial models**	**Gaussian models**	**All models**	**Strategy (a) Predict 2000s**	**Strategy (b) Predict 1980s**	**Binomial models**	**Gaussian models**
G	0	0	0	0	0	++	0	+++	0	+++
Ep	++++	++++	++++	++++	++++	+++	+++	+++	+++	+++
Enp	+++	+++	++	++	+++	++	+++	0	++++	0
Bs	++++	++++	++++	++++	++++	+++	++++	++	++++	+++
Ms	++++	++++	++++	++++	++++	+++	+++	+++	+++	++++
Fs	++	0	+++	+	+++	+	++	0	++	+
Ps	++++	+++	++++	++++	++++	++++	+++	++++	+++	++++
Da	+++	++++	++	+++	+++	++	0	++++	+	+++
Ds	++++	++++	++++	-	++++	++++	++++	++++	-	++++
M	++++	++++	++++	++++	++++	++++	++++	++++	++++	++++

0 (absence), + (<25%), ++ (25-50%), +++ (50-75%), ++++ (75-100%). See table 3.1 for formulas of the various hypotheses.

3.4.2. Evaluation of selected models

For strategy (a) ‘Predict 2000s’, combination of the nine binomial and four Gaussian models selected from prediction, were reapplied on the 2000-2007 period which resulted in 32 predicted distributions. For strategy (b) ‘Predict 1980s’, combination of the unique binomial model with the 12 Gaussian models selected from prediction, were reapplied on the 1981-1989 period which resulted in 12 predicted distributions. These predictions were compared using the Taylor diagram to the observations and to the prediction of a pure environmental model (that corresponds to the habitat model of section 3.3) and a control model (classified as the latest among all models). For strategy (a) ‘Predict 2000s’, the control model is the combination of the binomial model containing non-persistent environment and annual demography and the Gaussian model containing annual demography. For strategy (b) ‘Predict 1980s’, it’s the combination of geography, non-persistent environment, medium and fine scale sub-models and annual demography with the annual demography.

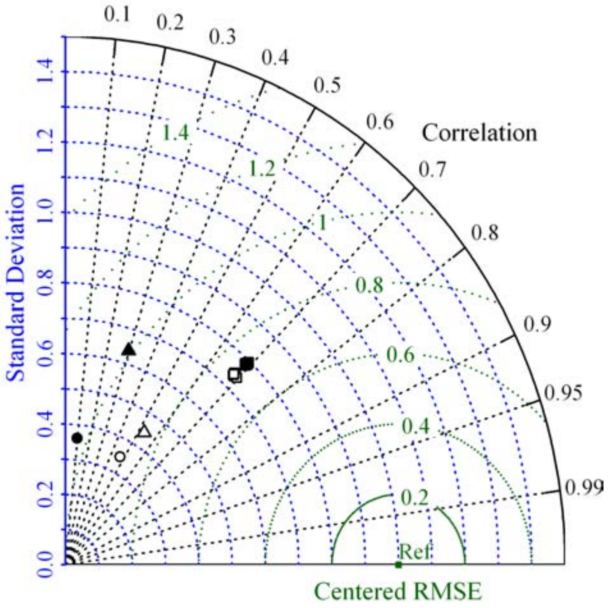


Figure 3.12. Evaluation of selected models by the Taylor diagram. The standard deviation, the RMSE and the Spearman correlation were calculated between observations and predictions of selected (square), habitat (triangle) and control (circle) models. Models are also evaluated for strategy (a) ‘Predict 2000s’ (black symbols) and strategy (b) ‘Predict 1980s’ (open symbols). Observations (Ref) are normalised so that their standard deviation is equal to 1, their correlation equal to 1 (the correlation between the observations and themselves) and their RMSE equal to 0 (the difference between the observations and themselves).

The Taylor diagram comparing these predicted distributions to observed distribution is presented in figure 3.12. Selected models were clearly apart from environmental and control models. Predictions from selected models have higher correlation, higher standard deviation and lower RMSE with observations than the environmental and control model, confirming that selected models have a better predictive power. Selected models have very close values of correlation, standard deviation and RMSE which make them hard to distinguish on the Taylor diagram. Selected models of strategy (a) ‘Predict 2000s’ and (b) ‘Predict 1980s’can be distinguished on the diagram however they have similar values of correlation (r² close to 0.7), RMSE (around 0.7) and standard deviation (between 0.75 and 0.80).

Maps of predicted distribution for the three types of model are indicated in fig. 3.13. Predicted distribution pattern from the best among selected models look more similar to the observed distribution compared to the others. The best model was able to depict more variability in the spatial distribution and to locate more accurately areas of high abundances that the habitat and control models.

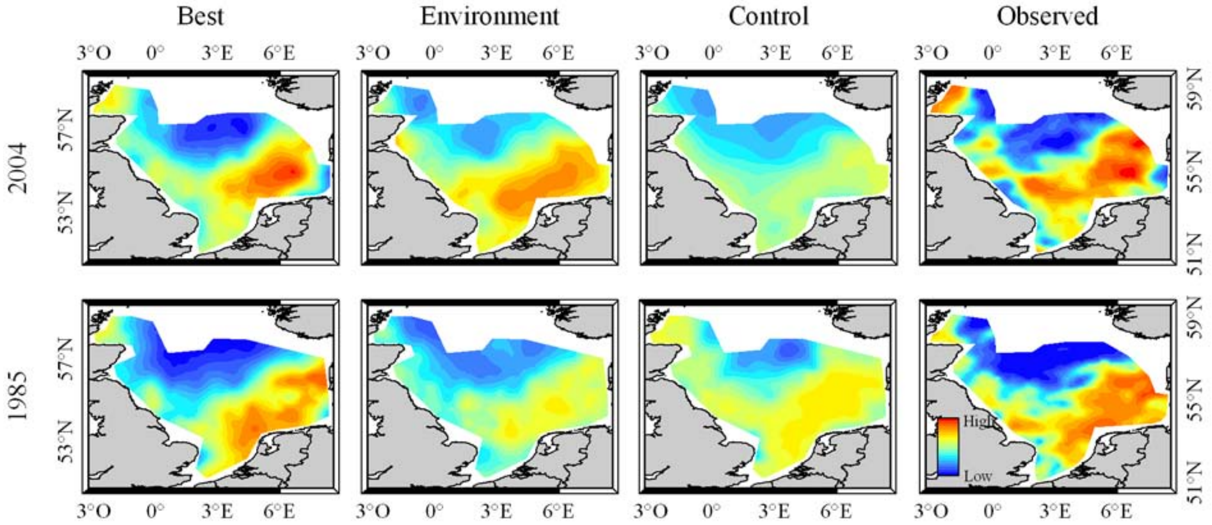


Figure 3.13. Maps of predicted distribution by the best among selected models, the habitat model and the control model, for the year 2004 (strategy (a) 'Predict 2000s') and year 1985 (strategy (b) 'Predict 1980s'). Maps of observed distribution for these two years are also represented. All maps of predicted distribution of the same year are on the same scale.

3.4.3. Variance partitioning

Percentage of variation explained by each hypothesis in the prediction dataset was measured using variance partitioning (table 3.4). Geography was not tested for presence-absence as it was absent in selected binomial models. Among the others hypotheses, non persistent environment and fine scale sub-model were not significant. Persistent environment, space and population memory accounted for most of the variation explained in the presence-absence and geography, space, spatial demography and population memory accounted for most of the explained variation in the positive abundances. This confirmed the interpretation of the frequencies of the hypotheses within selected models. Relevance of population size varied according to the type of response and the period considered. It explained a significant part of the variation in presence-absence of the eighties but not in the 2000’s and in positive abundances in the 2000’s but not in the eighties. Persistent environment explained a part of variation in the abundances in the eighties but not in the 2000’s. The analysis also revealed that the significant hypotheses shared an important part of the explained variation around 15 % for the presence-absence and 35 % for abundances. Once the effect of other hypotheses was partialled out, pure space and geography effect explained the most important part of variance respectively in the presence-absence and abundances.

Table 3.4. Summary table of the part of variation explained by each hypothesis in presence-absence and abundances of the prediction dataset for both strategies.

	Presence-absence		**Abundances**	
Component studied	**Strategy (a) Predict 2000s**	**Strategy (b) Predict 1980s**	**Strategy (a) Predict 2000s**	**Strategy (b) Predict 1980s**
All variation	100	100	100	100
All explained variation	24.2	34.9	43.3	38.4
Geography	**	**	37.8	38
Persistent environment	7.7	13.9	*	17.7
Non-persistent environment	*	*	*	*
Space (Bs + Ms + Fs*)	20.6	23.8	32.8	28.1
Population size	*	3.7	0.1	*
Population demography	*	*	18.8	18.1
Population memory	15.2	14.2	24.3	11.4
Purely by geography	-	-	3.9	3.6
Purely by persistent environment	1.9	4	-	0.04
Purely by space	6.9	12.8	2.5	0.04
Purely by population size	-	3.4	0.01	-
Purely by demography	-	-	2.6	0.03
Purely by population memory	1.1	0.4	0.2	0.1
Shared component	14.3	15.9	34.1	34
Unexplained variation	75.8	65.1	56.7	61.6

* not significant at p=0.05

** not tested

3.5. Discussion

3.5.1. Spatial distribution of spawning adults

Maps of the distribution of spawning adults showed a clear separation between the central part and the southern and western parts of the North Sea. The areas of high abundances in the southern and western parts were characterized as preferential and occasional spawning sites and corresponded to the main spawning grounds that were described from location of the eggs (Harding *et al.*, 1978), in the German Bight, in the Flamborough Head region and Moray Firth. Also, the proportions of spawning adults were higher in the southern part of the North Sea which confirmed a more important spawning activity in the German and Southern Bights which contribute the most and the Flamborough Head being the less productive region (Daan *et al.*, 1990). This has remained quite constant during the 20th Century except that the German Bight has become more productive than the Southern Bight and has expanded since the seventies (Harding *et al.*, 1978). In contrast, the observation revealed that the Southern Bight was not always occupied. This was probably due to the timing of the survey which began at the end of January whereas the peak of spawning in the Southern Bight occurred earlier in January (Harding *et al.*, 1978). Moreover, the amount of observations was limited in this area because of unfavourable substrate for bottom trawling (Verin Y., pers. comm.). Adults were absent each year in the central part of the North Sea which appeared not to be used by plaice for spawning.

The ratio of the spatial variance over the temporal variance exceeded 1 which indicates that the distribution of spawning plaice was more variable in space than in time. It seems that location of the spawning grounds in geographically well defined areas promoted this highly structured spatial distribution. Interpolated maps of each year revealed that the whole spatial structure of spawning plaice distribution has remained stable at the end of the last century. This was also confirmed by location of the spawning grounds from eggs surveys (ICES, 2005). However, the width of the occasional sites of spawning indicated a high variance between years inside the areas of high abundances, reflecting temporal variations in the observed abundance levels and in the extent of spawning areas. The similar behaviour between the spatial trend and the spawning stock biomass along years suggests an abundance-area relationship (Swain and Morin, 1996; Fisher and Frank, 2004; Blanchard *et al.*, 2005) where abundances and extent of spawning areas are higher for years of high spawning biomass which could suggest a basin model (McCall, 1990).

The correspondence analysis performed on observed distribution of spawning adults revealed a linear trend in the scores of the years which is associated with a spatial shift in the stations. Years of the period 1980-1992 were more associated with southernmost areas of the North Sea whereas years of the period 1993-2007 were more associated with central areas. We believe that this shift in the centre of gravity was the result of the contraction of plaice distribution due to the spawning stock biomass that started to decrease in 1992. For herring, it has been shown that such a decrease in the size of the population could lead to a spatial shift or a loss of past spawning grounds (Corten, 2002). For plaice, it seems that such an event has a more limited effect and may be counterbalanced by the strong attachment of this benthic fish to its spawning site. Moreover there was no collapse in the plaice population during the study period; the spawning stock biomass has returned to its original level which might have prevented dramatic change in the distribution of spawning adults.

3.5.2. The habitat model

Projection of the environmental relationship in the geographical space allows to locate suitable spawning areas. Highly suitable areas reflected quite well locations of spawning grounds described in the previous section. All the environmental predictors have been retained in the final adjusted binomial and Gaussian models. This indicated that depth, bottom temperature and salinity, seabed stress and sediment type play an important role in determining the spawning habitat of North Sea Plaice. Adjustment of response through the use of smoothing function of generalised additive models enabled to identify complex response curves that were clearly not linear. Moreover, response shapes were quite similar for each response towards environmental factor whether it was with the binomial or Gaussian model. For depth, bottom salinity and seabed stress, adjusted responses had a bell-shaped form which indicated an environmental window of preference of North Sea Plaice for these three factors during its reproduction. Spawning adults tended to concentrate on shallow areas of less than 50 metres depth which was in good agreement with what has been described for areas of high concentrations of eggs (Harding *et al.*, 1978). Little has been described about the relationship between salinity and spawning except for the Southern Bight where high salinity seemed to offer optimal conditions for developing eggs (Cushing, 1990). High salinity in the Southern Bight was due to inflow of central saline water from the English Channel trough the Dover Strait. In this study, we found an optimal response for salinity around 34.3; however, it was not the maximal salinity observed in the whole North Sea area. Relationship with temperature was more linear with a higher preference for temperatures between 4 to 8°C which was in the tolerance range of 2-8°C for plaice eggs survival (Wegner *et al.*, 2003).

Very few models used seabed stress as an environmental predictor. Spawning adults seemed to prefer areas with values of seabed stress close to zero, which corresponded to areas with low tidal impacts. These areas can be considered as having minimum turbulence which was associated with a minimum of turbidity that constituted optimal conditions for eggs (Harding *et al.*, 1978). The input of sediments in the model was less clear. There was a high standard deviation for each type of sediment in the adjusted response and there was a preference for muddy seabed. Moreover, sediments distribution in the North Sea was not very contrasted because of the lack of fine resolution of the map and was mainly dominated by fine sand. The sediment map may need to be updated with a finer spatial resolution to be more useful for habitat modelling.

3.5.3. Expansion of the habitat versus observed distribution

Predictions of the habitat model for each years allow seeing how the spawning habitat varied. Years of similar spawning habitat distribution can be grouped together with the cluster analysis combined to the correspondence analysis. Years of the eighties and seventies were separated from the nineties and the 2000’s which depicted a temporal trend of spatial expansion of suitable areas into the central and north-western part (south of Orkney Islands) of the North Sea. It can be assessed that the study period can be split at least into two periods with 1970-1995 at the beginning and the end of nineties and 2000’s after. If environmental conditions were less suitable in the seventies, it became more suitable in the 2000’s due to environmental change. In the habitat model, bottom temperature and salinity were the only predictors that varied from year to year. This meant that temporal change in suitability may be due to one of this parameter. However, the maps show that location of the main suitable spawning areas remained quite constant in time. This supported the fact that spawning grounds have remained spatially stable (ICES, 2005). We argue that the spatial stability of spawning grounds location was the consequence of environmental conditions remaining within the environmental range tolerated by the species. As a result, the habitat distribution was less variable in time than in space (Planque *et al.*, 2007). This low variability of suitable areas in the North Sea during the winter period may explain why this species use this period to reproduce.

Unlike what was shown by the habitat model, the observed distribution did not seem to have expanded during the last three decades. Plaice was always concentrated in areas of higher suitability in the southern part and along the English and Scottish coasts whereas it remained absent in the central part of the North Sea. This was illustrated through the Taylor diagram showing better predictions for the beginning of the period than for the end. These results implied that 1) the potential habitat was not always fully occupied, 2) change in habitat suitability did not automatically lead into a change in observed distribution of the species and 3) habitat modelling by environmental predictors was necessary but insufficient to understand what control the spatial distribution of a species.

The first point implies that other factors, related to the population itself, should affect the way in which a species is distributed and occupy its potential habitat. The observed distribution was usually less spread than expected, especially during the reproduction period when individuals tended to aggregate to favour the encounter of gametes of both sexes. The bell-shaped form of response curves found for depth, salinity and bedstress was due to the aggregative behaviour (Sutherland, 1983) of the individuals which tended to concentrate at the same place and was within the real environmental preferendum. The other concept which might control the occupation of the potential habitat was related to the density dependent effect (Fretwell and Lucas, 1970; Shepherd and Litvak, 2004) which states that the population will contract in smaller area at low level size and expand its distribution at higher level size.

The second point refers to the persistent character of a population that will play a key role in its spatial distribution. Difference in predicted and observed distribution tended to prove that spawning plaice persisted at the same place from year to year whatever environmental change that can occur at a larger temporal scale. This fidelity to spawning areas has also been observed in other studies (Hunter *et al.*, 2003) and other areas (Solmundsson *et al.*, 2005). This could suggest that more than the environmental conditions, it was the geographical position that was important for plaice reproduction. However, homing mechanisms remained unclear (Harding *et al.*, 1978; Gibson, 1997; Hunter *et al.*, 2003). Adults were not always distributed in the best environmental conditions for themselves but more in order to allow eggs and larvae to encounter favourable conditions thereafter. This should lead to discrepancies between predicted and observed spawning suitable areas. Therefore, in the light of these results, climate change impact is not expected to affect very much the spawning spatial distribution but should have consequences on reproduction success and survival of eggs and larvae.

Fidelity of the adult population of plaice to its spawning grounds may have adverse effect on population survival under climate change as adults will continue to spawn at the same place but in environmental conditions that will no more be suitable for eggs and larvae. Therefore, understanding the spatial distribution of a species by using environmental explanatory variables was not sufficient. This raised the problem about the prediction evaluation of habitat models when applied to climatic scenarios. In this study, we calibrated a model on thirty years and we were not able to depict the observed distribution of the last years showing the maximum environmental changes. So, we should wonder how reliable will be abundance predictions of the model when extrapolated in the future because of climate change or when the model is used into another area. Predictions under climate change can be highly misleading for spawning habitats and should be taken carefully as this species seemed to show a strong attachment to its reproduction sites. Finally, we recommend rather than adding more environmental variables in order to improve the habitat model, to consider and test other factors related to the state of the population such as the aggregation behaviour, density dependent effect and site attachment.

3.5.4. Control of the spatial distribution of spawning plaice

The analysis of the hypotheses retained in selected models from prediction reflects differences in the relevance of some hypotheses according to the type of model (either binomial or Gaussian). This suggests that the presence-absence of spawning plaice at a certain location was not controlled in exactly the same way as its level of abundance. This was also reflected in the variance partitioning analysis which allowed to quantify the percentage of variation explained by the most frequent hypotheses. The percentage of explained variation was not the same according to the type of response (presence/absence or abundances). Also, this analysis has shown that a large part of explained variation was shared by several hypotheses like geography, spatial dependency, persistent environment and memory. This reflects that these hypotheses were not mutually exclusive which may explain why a great number of models containing hypotheses of different nature had the same predictive power.

The importance of persistent environment for determining the spatial distribution of both presence/absence and abundances of spawning plaice means that its spatial distribution was primarily controlled by geographical variation in depth, bedstress and sediment types. Whereas non-persistent environment was present in most of selected models, we think that it was not a so relevant hypothesis to explain the spatial distribution of spawning plaice. We argue that the difference of frequency of both geographical location and non-permanent environment according to the time period considered reflects an overlapping effect in the 1980s between them on the spatial distribution of plaice. In the 1980s, the effect of geography was embedded in non-persistent environment and, although the non-persistent environment changed in the later period, the spatial distribution of plaice did not really change, revealing the importance of geography. This means that it is more the spatial structure of the non-persistent environment (highlighted by the importance of the geographical location) that was important rather than the pure environmental effect of non-persistent environment. The lesser importance of non-persistent environment is also supported by the fact that it was determined as not significant by the partial analysis. Therefore, the present study does not allow to prove that temporal changes in temperature and salinity could explain temporal variations in plaice spatial distribution during reproduction time. This lesser importance of non-permanent environment differs with results of most of the studies based on habitat modelling (Koubbi *et al.*, 2003; Koubbi *et al.*, 2006; Planque *et al.*, 2007; Loots *et al.*, 2007; Vaz *et al.*, 2008; present study). In these models, environment was generally the only hypothesis tested. As a result, one or several environmental variables were often kept in the final habitat model as they explained some variation in the distribution. The present study shows that when control by fluctuating environment is confronted to other hypotheses, it may not be as important as often believed. This means that non-persistent environment may not directly control the spatial distribution of a population, especially during the reproduction time, but can be spatially correlated with other more direct factors that were not (or not adequately) taken into account by habitat modelling studies. This is a critical issue in the context of hypotheses inference because it reveals that classical habitat modelling may attribute too much importance to environmental factors only because they only consider the environmental hypothesis. The low importance of non-permanent environment compared to other factors may reflect that this study focuses on the spatial distribution of a population and not on that of the species. This implies that the spatial scale involved was not the same and suggests that the temporal variations in temperature and salinity experienced by the North Sea population were within the tolerance range of plaice. Also, as the present study was restricted to the particular phase of spawning in the plaice life cycle, this may also suggest that the whole North Sea provided suitable hydrological conditions for plaice reproduction and that temperature and salinity were not as important for spawning adults distribution as they may later be for the successful survival of eggs and larvae.

Combined to persistent environment, control through the size of the population suggests a density dependent effect of the spatial distribution of plaice population during its reproduction time. For small population size, individuals contracted their distribution to preferred areas which can be characterised mostly by their persistent environmental characteristics. At high population size, fishes tended to expand their distribution. Population size was slightly more important in determining the areas of high abundance of plaice population than its areas of presence. This could reflect that variations in population size was not sufficient to fully detect its effect on the occupation of the spawning areas. Therefore, these results revealed that temporal variations in the size of the population were an important feature to explain temporal variations in the spatial distribution of spawning plaice.

We argue that the low frequency of the annual demography hypothesis during the 2000s only reflects the low diversity in the age structure at this time. This was supported by the fact that when the age structure was more diverse (as during the 1980s because the population size was larger), the role of the annual demography became evident through its frequency in selected models. Moreover, population demography seems to be more relevant in controlling level of abundances of spawning plaice than it is in controlling its presence/absence at a certain location. This implies that with the actual age structure of the population, spawning areas are occupied whatever the proportion of old and young adults that are observed each year. The use of the “old” and “young individuals” terminology is relative for plaice as only a small part of the natural age-pyramid is represented. The current heavily exploited population is dominated by younger ages (1 to 8 years, ICES, 2008a) whereas plaice may live up to 20-30 years old (Froese and Pauly, 2009). This may lead to underestimate the full role of the population age structure in determining the presence of spawning plaice. Despite the current truncated age structure, both annual and spatial population demography seem to be important to explain temporal variations in the spatial distribution of abundances. Young and old individuals of plaice are not occurring in the same areas. The older ones are usually distributed more offshore compared to youngest ones more near to the coast (Cushing, 1990). There is also a difference in time spent on the spawning grounds according to the age of the individuals. Old individuals arrive first on spawning grounds and stay longer compared to young individuals (Rijnsdorp, 1989). This implies that, depending on the annual age structure of the population observed in different years, several areas will be more or less occupied, which will lead to substantial variations in the observed distribution pattern from one year to another year. That is why, we argue that both the annual and spatial demography were important hypotheses to explain temporal variations in the spawning plaice distribution.

On the three spatial scales that have been detected in the spatial distribution of plaice spawning population, when confronted to other hypotheses, two are important : the broad and medium scales. They explain an important part of the variation in both the presence-absence and in the abundances of spawning plaice. Broad and medium scales describe the size of the spatial pattern of distribution. Broad scale is more important to characterise the areas of presence whereas medium scale is more important to characterise the areas of high abundances. This suggests that areas of presence are wider than areas of high abundances and thus not fully occupied. Moreover, the lower ability of the broad scale to explain the distribution at the beginning of the time period when adjusted to the end suggests a difference in the size of the pattern of distribution between the two periods. Patterns of presence were much larger at the beginning of the time period when population size was also much larger than at the end. This also confirms what was previously said about the control by the size of the population through the density dependent effect. Finally, the resolution of 55 km being similar to that of the survey, the absence of control of the fine scale sub-model could simply results from the sampling strategy. The actual form of the survey (both sampling strategy and gear) may not be suitable to explore processes at scales smaller than 55km.

3.5.5. The role of population memory: an evidence of conservatism for plaice?

Both presence-absence and abundance of spawning plaice also depends on the memory hypothesis. Geographical attachment is also relevant and explains the most important part of the variation in the spatial distribution of abundances. High concentrations of spawning adults are mainly located in the south-east of the North Sea and along the east coast of United Kingdom. Spawning adults are not abundant in the central and Northern part of the North Sea. These areas of high abundances coincides with the described spawning grounds in the literature that have been located from eggs surveys (Harding *et al.*, 1978). These spawning grounds are occupied by distinct sub-population that have been identified using tag experiment (Hunter *et al.*, 2004). Three main sub-populations have been identified, one in the western part, one in the south-eastern part and one more in the northern part of the North Sea. The hypothesis of the existence of a fourth sub-population in the Moray Firth region has also been advanced. The three sub-populations of the southern part of the North Sea are located on distinct feeding grounds during summer and mix on southern spawning grounds during winter (Hunter *et al.*, 2004). Northern and eastern sub-populations will predominantly migrate to the German Bight and Transition areas whereas the western sub-population will migrate to the Flamborough Head region, Southern Bight area and the eastern English Channel. This migration pattern seems to be repeated each year (Hunter *et al.*, 2003) which explains why several studies have shown from egg surveys that location of the spawning grounds did not vary a lot during the 20th century (ICES, 2005). Similar fidelity for their spawning areas was also shown for the Icelandic plaice (Solmundsson *et al.*, 2005).

Conservatism has been proposed to explain the ability of a population to maintain its spawning location in time (Corten, 2002). It implies that fish have learned a certain behaviour that they have turned into a habit and that this habit has been transmitted across generations. Whereas it has been suggested for several pelagic species (McQuinn, 1997; Corten, 2002; Petitgas *et al.*, 2006), one can ask if this concept can be valid for a benthic fish. In the case of plaice, the habit corresponds to the annual migration from northern summer feeding grounds to southern winter spawning grounds. Whereas the habit formation part of conservatism could be favoured by using external clues such as physical characteristics of migrations routes and tidal current (Hunter *et al.*, 2003), orientation mechanisms implied in plaice migration remain unclear. Corten (2002) argues that to prove the existence of habit formation, it must be demonstrated that (1) the same fish visit the same location in successive years, (2) that this behaviour is not genetically determined and (3) that the return of the fish is not simply due to environmental constraints. Point (1) is supported by the fact that tag experiments have shown that the different identified sub-populations of North Sea plaice tends to return to their specific spawning grounds from one year to another (Hunter *et al.*, 2004). The present study allows validating point (3) as it has proved that space, persistent environment and population memory are the key factors and that non-persistent environment is not so important since the observed distribution does not appear to follow its changes. According to us, because the permanent environment is spatially structured but does not vary in time, it should act more like a spatial constraint rather like a pure environmental constraint (i.e. in the sense of direct physiological effect that temperature and salinity may have on spatial distribution). That is why we argue that the importance of permanent environment does not call into question the role that conservatism may play in the spatial stability of North Sea plaice spawning areas. The mechanism of transmission of this habit between generations still remains unclear for plaice but the encounter between first time spawners and spawning adults on spawning or nursery grounds (Cushing, 1990) could be a preliminary supportive evidence for the existence of tradition (sensu Corten, 2002).

Three main events may change the tradition (Corten, 2002). First, if environmental conditions become really bad (Rindorf and Lewy, 2006). For the considered study period, it does not seem that the main pattern of spawning distribution has changed as spawning plaice is still located in the western and southern part of the North Sea. This could suggest that until now plaice did not experience enough strong environmental changes to modify its spawning behaviour. This reinforces the thought that the North Sea does not provide really unsuitable environmental conditions for plaice to spawn. However, this can be biased by the fact that conservatism creates a certain inertia in the spawning behaviour which can create a time lag between the environmental change and the change in fish migration (Corten, 2002). It is highly probable that current distribution of North Sea plaice reflects the environmental conditions in the past rather than those at present. Then, tradition can also change following a collapse of the stock. North Sea plaice has been evaluated since 1957 (ICES, 2008a) and no stock collapse has been noticed which also could explain why the distribution pattern could have been maintained successfully across generations. Finally, the last explanation of change in tradition is an abnormal distribution of the recruiting year-class or a scarcity of older individuals. In both case, older individuals are not enough to entrain all the pool of young individuals, which can lead to the loss of a part of the population recruit that can potentially colonise past or new spawning areas. As Rijnsdorp (1989) noticed, individuals of plaice now become mature at younger age and smaller size compared to the beginning of the century. The fact that, all individuals do not become mature at the same age, greatly limits the risk of non-entrainment and lost of a part of the first time spawners. This is true only under the condition that older individuals are still numerous enough to entrain young individuals even in the case of a small pool of recruits. In the case of plaice, strong one-year class has been noticed for plaice since 1957 (ICES, 2008a) which was followed by important recruitment to spawning population. If conservatism in plaice is true, the fact that spawning tradition has been maintained across generations could suggest that the pool of older individuals should have always been sufficient to entrain all the young individuals. However, as it has been previously said, the term “old” is relative for plaice and one have to admit that the reproduction in plaice is nowadays predominantly ensured by younger individuals. In the context of global warming, this can have strong impact on the future distribution of North Sea spawning plaice population if, as Corten (2002) suggested, young individuals are expected to respond more directly to environmental change than older ones.

CHAPITRE 4

Spawning distribution of a demersal fish: North Sea whiting

4.1. Introduction

North Sea whiting (*M. merlangus*) is one of the key species of the North Sea demersal fish assemblages (Greenstreet and Hall, 1996) as well as the third most exploited gadoid fish in Scotland (Stratoudakis *et al.*, 1999). Not as much studied as cod and haddock, there are few researches on its population structure (Hislop and MacKenzie, 1976; Rico *et al.*, 1997; Charrier *et al.*, 2007), stock assessment (Pope and Macer, 1996), diet (Hislop *et al.*, 1991; Pedersen, 1999) and spatial distribution (Zheng *et al.*, 2001; Zheng *et al.*, 2002; Atkinson *et al.*, 2004). Whiting has one of the longest spawning period among North Sea species, going from February to June with a peak in April (Gibb *et al.*, 2004). Spawning occurs later in the season as the latitude increases. Whiting begins to reproduce at 2 years old. Adults prefer shallow waters with temperatures between 6 and 9°C and have a protracted spawning between 6 weeks and 2 months. Eggs are pelagic and released in the water in the form of many batches (ICES, 2005). As for many fish in the North Sea, whiting stock abundance has strongly declined since the seventies (Hislop, 1996; Serchuk *et al.*, 1996) because of an increasing fishing mortality since the forties (Greenstreet and Hall, 1996), a high discarding rate (Stratoudakis *et al.*, 1999) and a poor recruitment in the eighties and nineties (ICES, 2008a). Gibb *et al.* (2004) indicate that this decline may be associated with a contraction in the spatial distribution and a loss of distinct reproductive populations. However, Daan *et al.* (1990) highlight that little information on spawning distribution and factors that determine the choice of spawning grounds is available (Coull *et al.*, 1998; Gibb *et al.*, 2004; ICES, 2005).

The aim of this chapter is to describe and map the spatial distribution of spawning whiting in the North Sea in the last thirty years and to identify factors that control the spatial pattern in the distribution of spawning adults. In section 4.2, the spatial distribution of spawning whiting is described and mapped through the spatial distribution of spawning adults. Results on the inter-annual variability of spawning whiting distribution are also given.

In section 4.3, results of the application of the multi-model approach used to identify factors that control its spawning distribution are presented. In section 4.4, results of sections 4.2 and 4.3 are discussed.

This chapter is based on the two following articles:

(2) C. Loots, S. Vaz, B. Planque, P. Koubbi
Spatial distribution of spawning population of
North Sea plaice and whiting from 1980 to 2007
Article in preparation

(5) C. Loots, B. Planque, S. Vaz, P. Koubbi
Identifying processes that control the spawning distribution of
North Sea whiting using multi model inference.
Fisheries Oceanography. Submitted

4.2. Distribution of spawning adults

The spatial distribution of spawning whiting at the beginning of its reproduction period has been studied through the distribution of spawning adults during the first quarter of the IBTS surveys between 1980 and 2007. Geostatistical analyses (see section 2.3.1 for the methodology) and PCNM (see section 2.4.1 for the methodology) were used to describe the spatial structure in the spawning distribution of whiting. For each year between 1980 and 2007, observed abundances were interpolated using kriging and mapped into a geographical information system. Maps of each year were used to calculate the average and variability maps from which preferential, optional and unfavourable spawning site were located (see section 2.3.2 for the methodology). Interpolated abundances for each year were also used in a correspondence analysis to explore how the distribution of spawning whiting has varied spatially among years (see section 2.3.3 for the methodology). This section is the second part of the article (2) in preparation.

4.2.1. Proportions of spawning adults

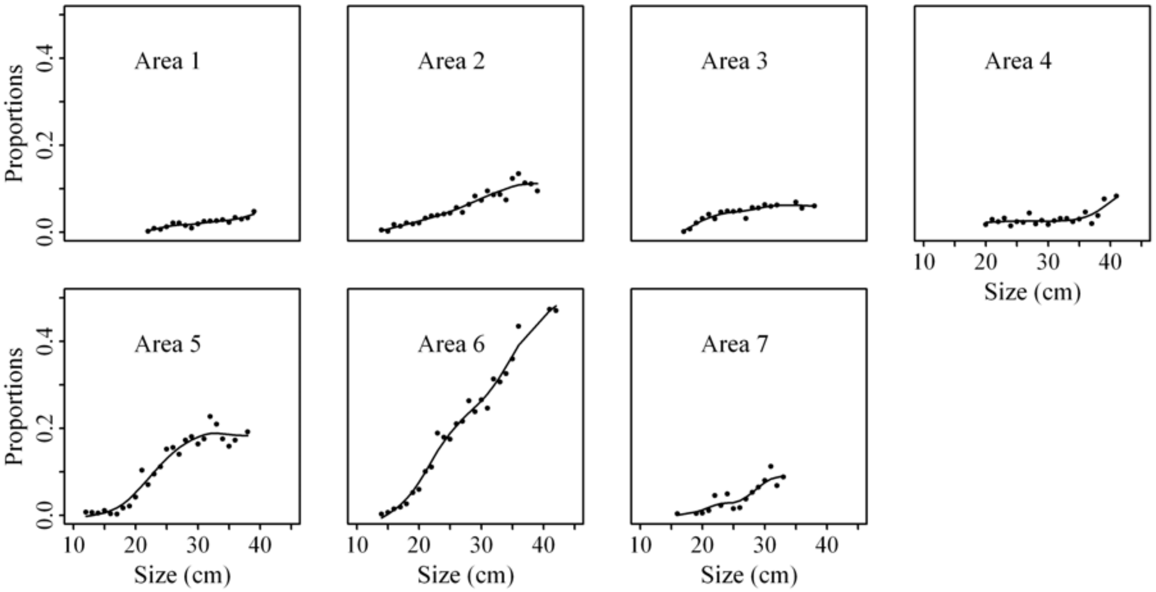


Figure 4.1. Proportions of stage 3 adults (spawning adults) of whiting per size class for each area. Symbols are the observed proportions and lines are the adjusted proportions using a smoothing spline for missing size class. See fig. 2.1 for the location of the areas.

Proportions of spawning whiting were calculated for each roundfish areas (fig. 4.1). Proportions strongly vary according to the area. They are high in the southern part of the North Sea (5, 6 and 7), intermediate in the central part (2 and 4) and low in the northern part

(1 and 3). Proportions of spawning adults are higher for larger size class of fish. They reach their maximum only in area 5 with 20% of spawning adults for size greater than 30 cm.

4.2.2. Geostatistics

Twenty eight variograms were computed and adjusted on abundances of spawning adults between 1980 and 2007. Values of the spatial trend, nugget and sill are given for each year in figure 4.2. The percentage explained by the spatial trend decreases between 1980 and 1991 from 35 % to 12%, then increases at about 30 % until 2000 and remains stable between 20 and 25 % in 2001-2007. For years 1980-1991, the evolution of the estimation of the SSB of whiting follows the evolution of the percentage explained by the spatial trend. However, this is not the case for following years where the SSB continues to decrease.

No spatial structure was detected in the residuals for years 1981 and 1983. A pure nugget model was used for these years. For other years, a combined model with a nugget component and one of the four tested model among exponential, circular, spherical and penta-spherical was used as the theoretical variogram. The circular model was used in more than 50 % of the theoretical variogram models (see appendix A.2.3). Nugget values varied a lot between years but showed a general decrease. Values of sill along years remained constant between 0.2 and 0.5 except for a peak in year 1987 and lower values below 0.2 at the end of the 2000’s.

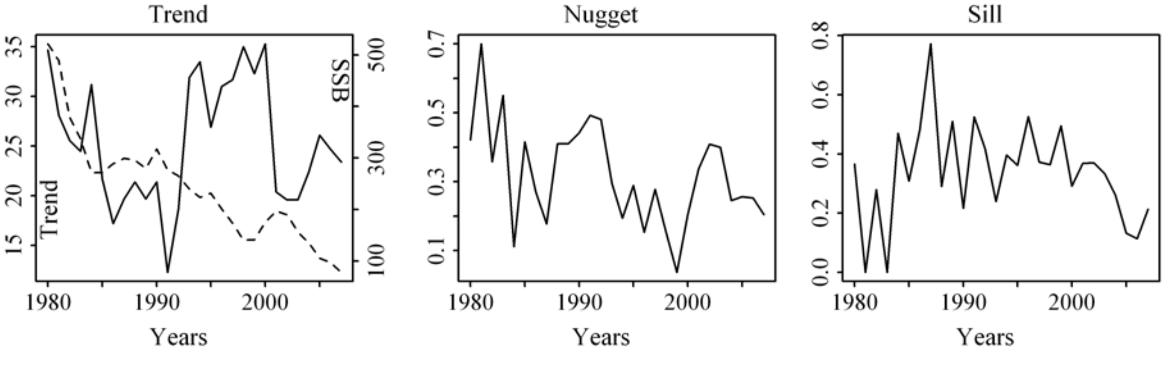


Figure 4.2. Parameters of the variogram. The spatial trend (in percentage explained), the nugget and the sill determined from geostatistical analyses are plotted for each year between 1980 and 2007. ICES estimation of spawning stock biomass (SSB in 10^3 tons, dotted line) is also indicated.

4.2.3. Distribution of spawning adults

Based on the previous adjusted variograms, abundances were mapped using kriging interpolation (fig. 4.3). Maps of distribution indicated that there were two main areas of high abundance of spawning adults. One was located north of the Dogger Bank along the east coast of Scotland, from the south of Shetlands Islands to the Flamborough Head region. The other was located south of the Dogger Bank in the Southern Bight and in front of the Flamborough Head region. Very low abundances were located in the central and eastern parts of the North Sea as well as in the Firth of Forth and Moray Firth regions. The distribution of spawning adults was more variables in space than in time with a ratio of 1.78. This indicated that the distribution pattern was constant in time; however there were temporal variations in the maximal abundance observed as well as in the extent of spawning areas. At the end of the 2000's, abundances were much lower and spawning adults were distributed in narrower areas than at the beginning of the nineties.

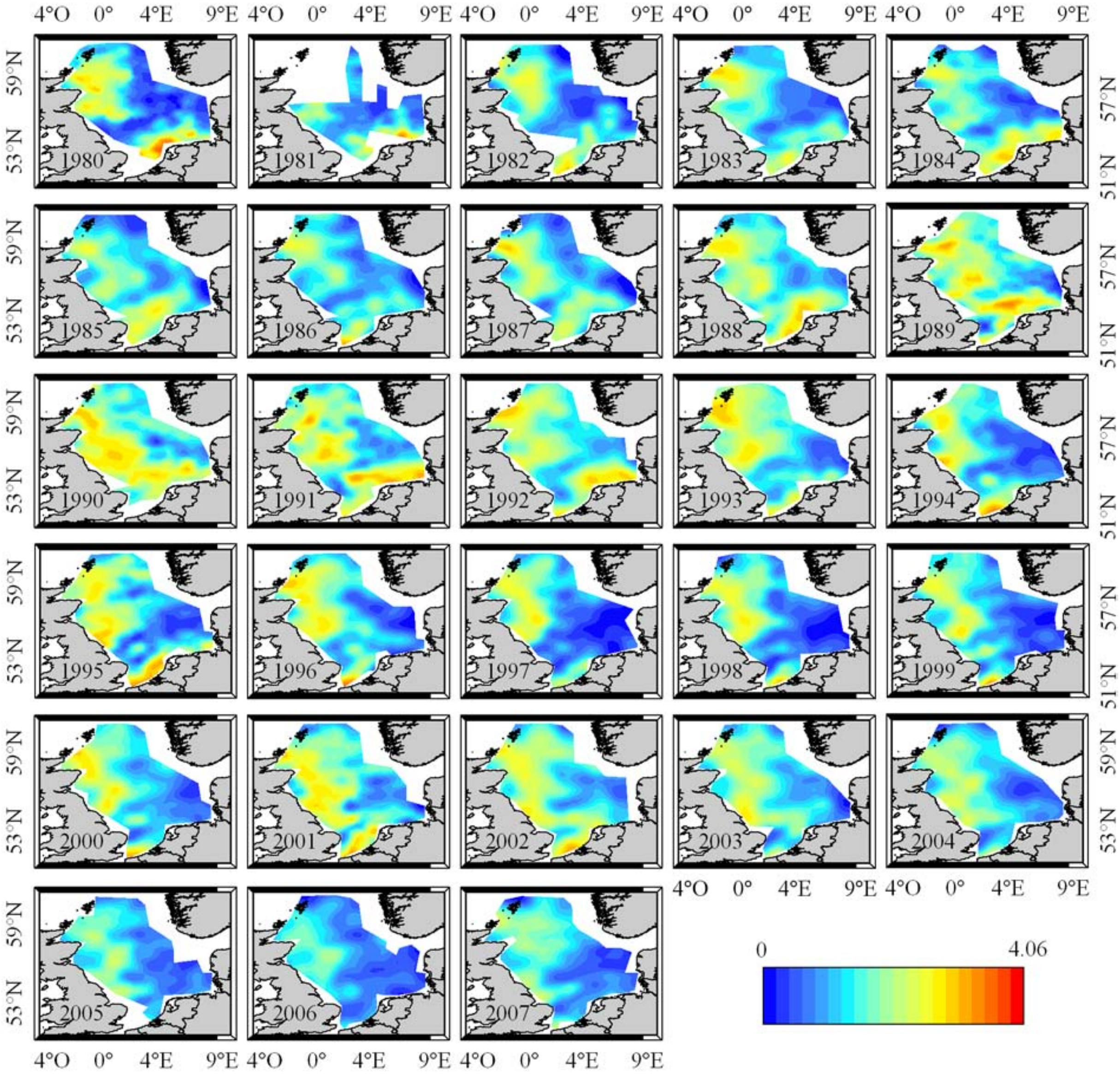


Figure 4.3. Maps of the spatial distribution of spawning adults of North Sea whiting from 1980 to 2007.

Mean and variance of the distribution maps of each year between 1980 and 2007 were calculated. A threshold of 1.60 for the mean and 0.5 for the variance were used to combine these two maps in order to locate preferential, occasional and unfavourable spawning sites (fig. 4.4). Preferential spawning sites were located in the northern and southern part of the North Sea. Occasional sites correspond to a small area in the Southern part whereas unfavourable areas were located in the central and eastern part of the North Sea.

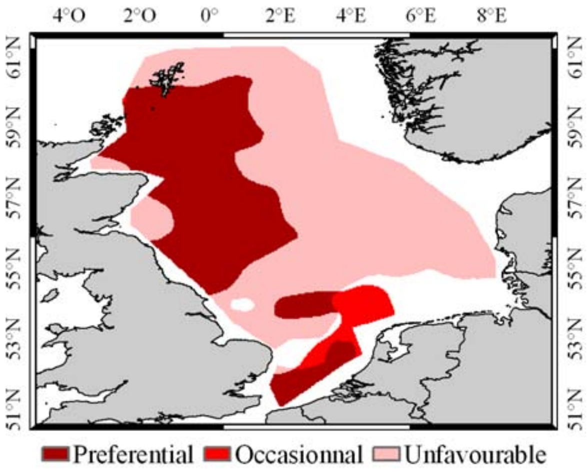


Figure 4.4. Location of preferential, occasional and unfavourable spawning sites for North Sea whiting. Map was computed by combining average and variability maps that were calculated as the mean and standard deviation of all maps of distribution between 1980 and 2007 (see fig. 4.3).

PCNM were used to describe spatial scales relevant to the spatial distribution of spawning whiting. A threshold of 1.13 (decimal degrees of latitude), corresponding to four times de maximum distance between two neighbouring observations, was used to truncate the distance matrix between stations. In total, 3832 positive PCNM were extracted from the principal coordinate analysis and 22 were determined as significantly correlated to whiting abundances by the forward selection (p-value $\leq$ 0.001 with an adjusted R^2 of 0.156). From these 22 PCNM (see appendix B.2), three spatial scales were identified (fig. 4.5): broad scale (166-355 km), medium scale (55-166 km) and fine scale (< 55 km). Three PCNM accounted for the broad scale, ten for the medium scale and nine for the fine scale. The broad scale represented the main areas of high and low concentrations of spawning adults. The medium scale was more difficult to interpret but could be related to main hydrodynamical and physical features of the North Sea. The fine scale seemed to describe the variability in the spatial distribution of spawning adults detected at the scale of statistical rectangles.

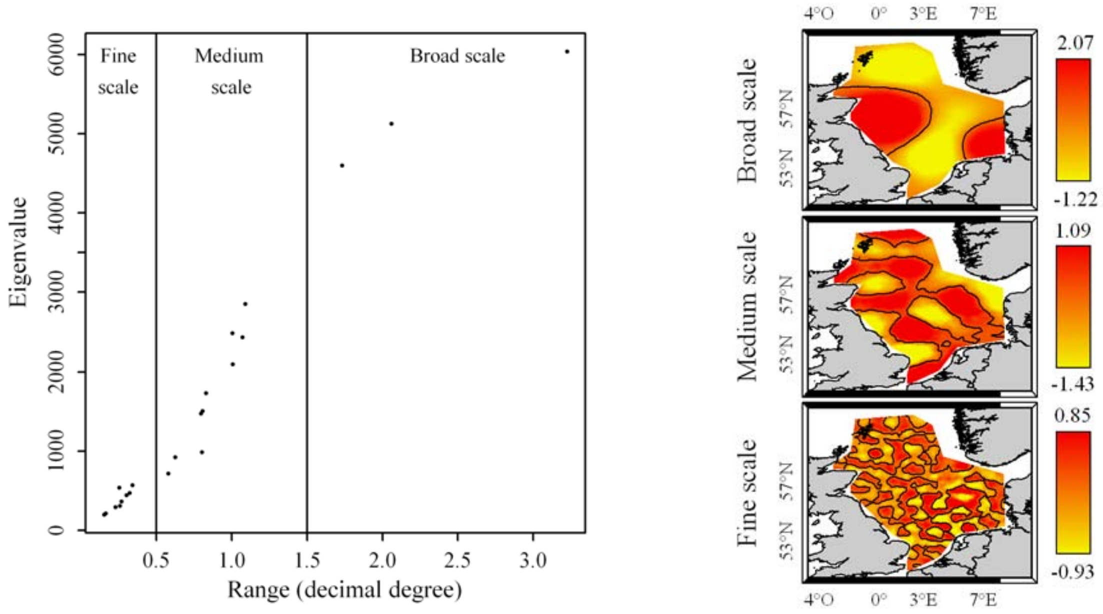


Figure 4.5. PCNM and spatial sub-models. Left : plot of the range values of the 22 selected PCNM against their eigenvalues. Ranges were obtained by fitting a Gaussian model as a theoretical variogram on each spatial predictor. PCNM describing similar scale were grouped together. Three scales were determined; broad scale : eigenvectors with a range between 166 and 233 km (1.5 to 3.2 decimal degrees of latitude); medium scale : eigenvectors with a range between 55 km and 166 km (0.5 to 1.5 decimal degree of latitude); fine scale : eigenvectors with a range between 0 and 55 km (less than 0.5 decimal degrees of latitude). Right : maps of the three scale of variation. PCNM 4, 25 and 73 were respectively used to depict broad, medium and fine scale. Black line indicates the zero value.

4.2.4. Inter-annual variability of spawning distribution

A correspondence analysis was performed on previously interpolated abundances to study the inter-annual variability of whiting distribution (see section 2.3.3 for the methodology). The first axis of the CA explained 26.4 % and the second axis explained 13.8 % (fig. 4.6). Distribution of the years in the factorial plan show a Guttman effect indicating a strong gradient along axis 1 and 2. Along axis 1, years 1980 to 1992 are located on the right part and years 1993 to 2007 on the left part. Scores of the years for the first axis show a decrease in their value from 1980 to 2007 (fig. 4.6). Scores of the years for the second axis alternate between negative values (years 1980-1987 and 1993-1999) and positive values (1988-1992 and 2000-2007). Four groups of stations were identified through the cluster analysis with respectively 104, 232, 270 and 201 stations for group I, II, III and IV. These groups of stations are distributed along a Southeast-Northwest direction (fig. 4.6) with group I and II in the south part, group IV in the central part and group III in the north part. These groups of stations are also distinct in the factorial plan with group 1 and 2 on the right part of the first axis, and group 3 and 4 on the left part. Group 1 is closer to years 1980 to 1984, group 2 is close to years 1985-1992, group 3 is close to years 1993-1999 and group 4 is close to years 2000-2007.

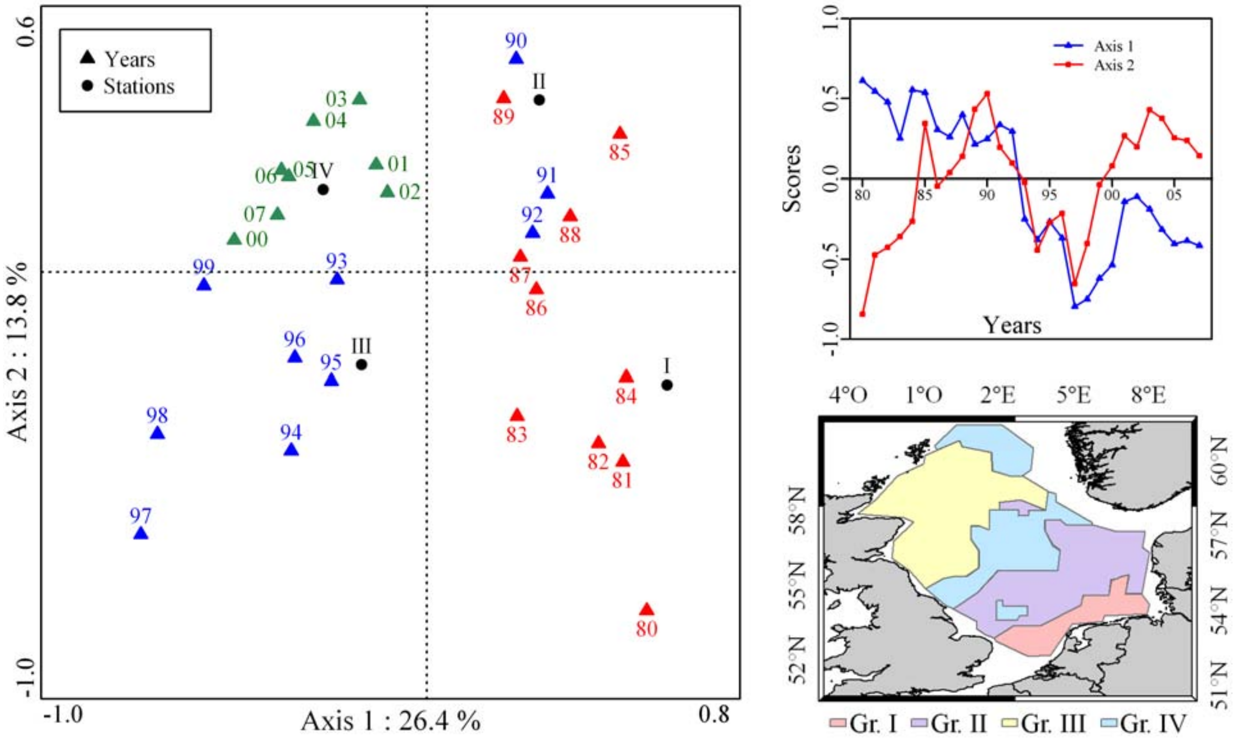


Figure 4.6. Correspondence analysis performed on abundances of whiting. Left : years and groups of stations are plotted on the factorial plane of axis 1 and axis 2. Groups of stations are represented by their centroids on the factorial plan. Top-right : scores of the years on axes 1 and 2. Bottom right : location of the four groups of stations on the map of the North Sea.

4.3. Multi-model approach

Factors that determine the spatial distribution of spawning whiting have been identified using the multi-model approach (see section 2.4 for the methodology). Nine hypotheses were confronted in binomial models to discriminate among factors that control the presence-absence : geographical position, persistent and non persistent environment, the three sub-models of spatial dependency, population size, yearly age structure and population memory. In Gaussian models, spatial distribution of age proportions was also included to study factors that determine the distribution of positive abundances (table 4.1). Respectively, 511 and 1023 possible binomial and Gaussian models were built, compared and ranked. A prediction dataset larger than the calibration dataset has been used to study its impact on the complexity of selected models from prediction. Models have been selected using a maximum threshold value of 60 for Δ_i and LRT in order to see the effect of different values of selection threshold on model complexity and hypothesis frequency. This section is base on article (5).

Table 4.1. Formulas of the various hypotheses used in binomial and Gaussian models in the multi-mode approach applied on whiting.

Hypotheses	**Binomial models**	**Gaussian models**
G	as.factor(StatisticalRectangle)	as.factor(StatisticalRectangle)
Ep	s(Depth, 4) + s(Bedstress, 4) + as.factor(Sediments)	s(Depth, 4) + s(Bedstress, 4) + as.factor(Sediments)
Enp	Temperature + s(Salinity, 3)	s(Temperature, 4) + s(Salinity, 4)
Bs	s(Ax1, 3) + s(Ax2, 4) + s(Ax4, 4)	s(Ax1, 4) + s(Ax2, 4) + s(Ax4, 4)
Ms	Ax11 + s(Ax13, 2) + Ax14 + s(Ax16, 4) + s(Ax18, 3) + Ax20 + Ax21 + Ax25 + s(Ax26, 3) + s(Ax28, 1)	s(Ax11, 4) + s(Ax13, 4) + s(Ax14, 4) + s(Ax16, 4) + s(Ax18, 4) + Ax20 + s(Ax21, 4) + s(Ax25, 4) + Ax26 + s(Ax28, 4)
Fs	s(Ax31, 4) + Ax33 + s(Ax36, 4) + s(Ax39, 1) + s(Ax54, 2) + Ax70 + s(Ax73, 1) + Ax104 + s(Ax110, 2)	s(Ax31, 4) + Ax33 + s(Ax36, 4) + s(Ax39, 1) + s(Ax54, 2) + Ax70 + s(Ax73, 1) + Ax104 + s(Ax110, 2)
Ps	PopulationSize + s(SpawningStockBiomass,1)	PopulationSize + s(SpawningStockBiomass, 4)
Da	s(Age1a, 4) + Age2a + Age3a + Age4a + Age5a	Age1a + Age2a + s(Age3a, 3) + s(Age4a,4) + Age5a
Ds	-	s(Age1s, 2) + s(Age2s, 4) + s(Age3s, 1) + s(Age4s, 3) + s(Age5s, 4)
M	s(PreviousYearAbundance, 2)	s(PreviousYearAbundance, 4)

G = geography, Ep = persistent environment, Enp = non-persistent environment, Bs = broad scale sub-model, Ms = medium scale sub-model, Fs = fine scale sub-model, Ax = number of the PCNM, Da = annual demography, Age1a = annual proportion of age 1, Age1s = spatial proportion of age 1, Ds = spatial demography, M = population memory, “s” means that a smoothing function has been used and the corresponding degree of smoothing is indicated in brackets.

4.3.1. Impact of the selection threshold on selected models

Using a limit threshold of selection of 60 for Δ_i and LRT, a total number of 320 models were finally selected. Selected models were not the same according to the selection procedure (fitting or prediction) and the type of model (binomial or Gaussian). Complexity of selected models was studied using the mean number of hypotheses and variables they contain (Fig. 4.7). Models selected on data fitting performances were in average more complex, in term of number of hypotheses and explanatory variables, than models selected on predictive capabilities. Selected Gaussian models were also more complex than selected binomial models.

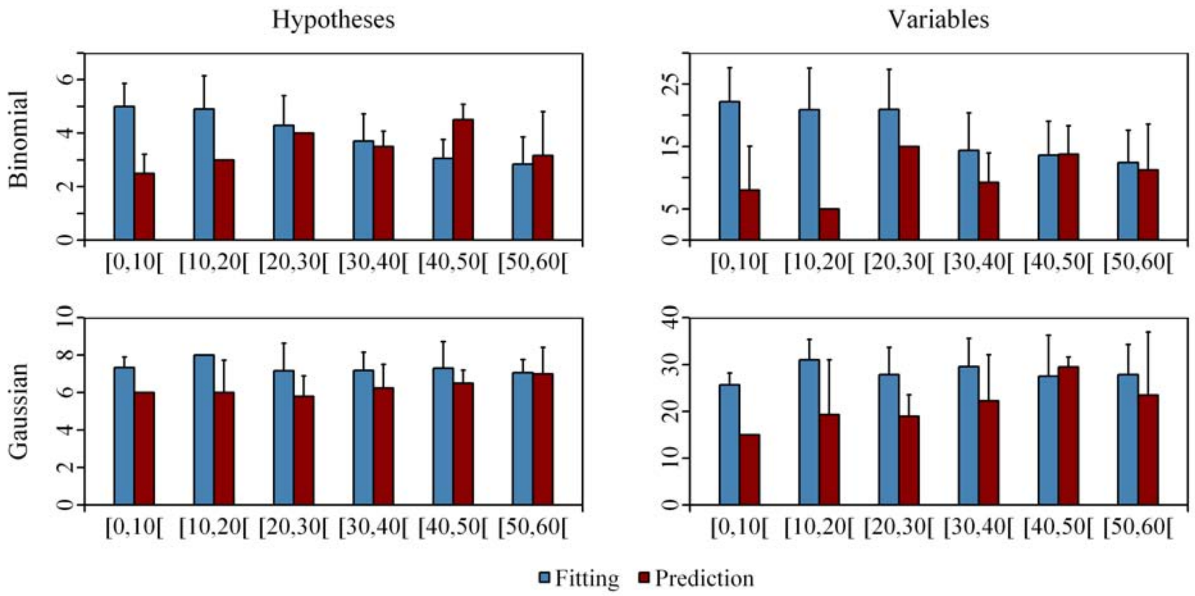


Figure 4.7. Complexity of selected models. Complexity is represented by the mean number of hypotheses and explanatory variables in binomial and Gaussian models. Complexity of models selected either on fitting or prediction is given for different threshold of selection between 0 and 60.

Complexity of selected models was not the same according to the threshold of selection (fig. 4.7). In binomial models selected from fitting, the number of hypotheses and explanatory variables decreased as the selection threshold increases. This was not the case for binomial models selected on prediction. Selected models with lower value of LRT (0-20) contained fewer hypotheses and explanatory variables than models with higher LRT values (20-60). Gaussian models selected on fitting contained an equivalent number of hypotheses and explanatory variables whatever the threshold considered. This was not the case for models selected on prediction. Models with low LRT values (0-30) contained less hypotheses and explanatory variables than models with higher LRT values (30-60).

4.3.2. Impact of the selection threshold on hypotheses frequency

Frequencies of each hypothesis of control were calculated from the 24 binomial and 17 Gaussian models selected from prediction (table 4.2).

Table 4.2. Formulas of models selected from prediction. Binomial and Gaussian models were selected with an LRT value between 0 and 60. The Δ_i was also given as an informative value to show that these models were not determined as the best models by fitting based selection. The number in bracket indicated models from which predictions have been mapped : (1) are simple models and (2) are complex ones. See table 4.1 for formulas of the hypotheses.

Binomial models	**LRT**	**Δ_i**	**Gaussian models**	**LRT**	**Δ_i**
Enp + M *(1)*	0.0	54.3	G + Ep + Enp + Ps + Ds + M	0.0	42.7
Enp + Ms + M	9.2	10.1	G + Ps + Ds + M *(1)*	11.2	160.7
Enp + Ps + M	17.5	47.0	Ep + Enp + Ms + Fs + Ps + Ds + M	16.5	244.8
Enp + Ms + Ps + M	27.5	5.7	G + Ep + Enp + Bs + Ps + Ds + M	16.9	56.0
Enp + Bs + M	31.3	12.4	G + Fs + Ps + Ds + M	21.7	106.1
Ep + Enp + Bs + M	33.8	17.5	G + Ep + Enp + Ms + Ps + Ds + M	21.7	50.9
Ep + Enp + M	38.5	36.8	G + Bs + Ps + Ds + M	23.5	166.8
Ep + Enp + Ms + M	38.9	10.0	G + Ms + Ps + Ds + M	23.7	162.0
Enp + Bs + Ps + M	42.6	9.9	G + Ep + Enp + Bs + Ms + Ps + M	26.9	61.7
Ep + Enp + Bs + Ms + M	44.4	11.4	G + Bs + Ms + Ps + Ds + M	30.2	166.7
Ep + Enp + Bs + Ps + M	45.6	15.1	G + Bs + Fs + Ps + Ds + M	34.0	116.0
Enp + Bs + Ms + M	45.7	1.0	Ep + Enp + Bs + Ms + Fs + Ps + Ds + M *(2)*	37.3	245.3
M	50.4	73.4	G + Enp + Ps + Ds + M	39.1	90.3
Enp + Ms	50.5	21.2	G + Ms + Fs + Ps + Ds + M	40.4	118.8
Ep + Enp + Ps + M	51.4	32.8	G + Bs + Ms + Fs + Ps + Ds + M	47.1	129.3
Ep + Enp + Ms + Ps + M	52.4	7.6	G + Ep + Enp + Ms + Fs + Ps + Ds + M	53.4	0.1
Enp + Fs + M	53.5	38.2	G + Enp + Bs + Ps + Ds + M	58.7	97.8
Ps + M	53.8	61.1			
Ep + Enp + Bs + Ms + Ps + M *(2)*	53.8	10.5			
Enp + Bs + Ms + Ps + M	54.6	0.0			
Ms + M	55.5	28.0			
Ep + Bs + M	57.9	36.5			
Enp + Ms + Fs + M	58.9	8.7			
Enp	59.6	84.0			

Frequencies of the various hypotheses differed between binomial and Gaussian models and varied according to the value of the threshold of selection (table 4.3). For binomial models, three hypotheses dominated regardless of the considered threshold: non-persistent environment, medium spatial scale and population memory. Control by population size was found in models with LRT values between 10 and 20. The number of hypotheses raised from 4 to 6 in models with a LRT value between 20 and 40 with persistent environment and broad scale sub-model as supplementary hypotheses. Control by fine spatial scale was only present and not frequent (less than 50%) in models with a LRT value between 50 and 60. Geography and annual demography were not present in selected binomial models.

Table 4.3. Frequencies of the various hypotheses among binomial and Gaussian models selected from prediction (see table 4.2). They illustrated the percentage of model in which each hypothesis is present. They are given for LRT values between 0 and 60.

	Binomial models						**Gaussian models**					
	[0-10[	**[0-20[**	**[0-30[**	**[0-40[**	**[0-50[**	**[0-60[**	**[0-10[**	**[0-20[**	**[0-30[**	**[0-40[**	**[0-50[**	**[0-60[**
G	0	0	0	0	0	0	++++	++++	++++	++++	++++	++++
Enp	++++	++++	++++	++++	++++	++++	++++	++++	+++	+++	++	+++
Ep	0	0	0	++	++	++	++++	++++	+++	++	++	++
Bs	0	0	0	+	+++	++	0	+	++	++	++	++
Ms	+++	++	+++	++	++	++	0	+	++	++	++	++
Fs	0	0	0	0	0	+	0	+	+	++	++	++
Ps	0	++	+++	+	++	++	++++	++++	++++	++++	++++	++++
Da	0	0	0	0	0	0	0	0	0	0	0	0
Ds	-	-	-	-	-	-	++++	++++	++++	++++	++++	++++
M	++++	++++	++++	++++	++++	++++	++++	++++	++++	++++	++++	++++

0 (absence), + (<25%), ++ (25-50%), +++ (50-75%), ++++ (75-100%). See table 4.1 for formulas of the various hypotheses.

For Gaussian models (table 4.3), six hypotheses dominated in models regardless of the selection threshold value: geography, persistent and non-persistent environment, population size, spatial demography and population memory. Persistent and non-persistent environment were less frequent in models with LRT values between 20 and 60. The three scales' sub-models were not present in models with LRT values between 0-10 whereas they were increasingly frequent in models with higher LRT values. Annual demography was not present in selected models.

4.3.3. Model evaluation

The 17 binomial and 24 Gaussian models selected from prediction were reapplied to the prediction dataset. They were combined into delta-models where probability of presence from each selected binomial model were multiplied by abundances of each selected Gaussian model, which has resulted in 408 predictions. These predictions were compared to observations using the Taylor diagram (fig. 4.8). The Taylor diagram indicated that all the models had comparable values of Spearman correlation (close to 0.8), Root Mean Square Error (close to 0.6) and standard deviation (between 0.7 and 0.8).

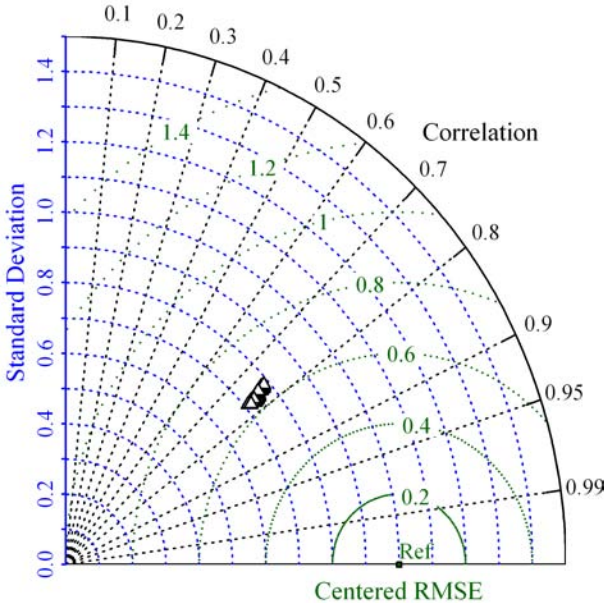


Figure 4.8. Evaluation of selected models by the Taylor diagram (only the positive correlation part is represented). Observations and predicted values of delta models resulting from the combination of binomial and Gaussian models selected from prediction are compared using the standard deviation, the RMSE and the Spearman correlation. Observations (Ref) are normalised so that their standard deviation is equal to 1, their correlation equal to 1 (the correlation between the observations and themselves) and their RMSE equal to 0 (the difference between the observations and themselves). See table 4.2 for formulas of selected models.

Predictions of two delta models of contrasting complexity were mapped for two years of the prediction dataset in order to check if there was a real difference of predictive capacity between them (fig. 4.9).

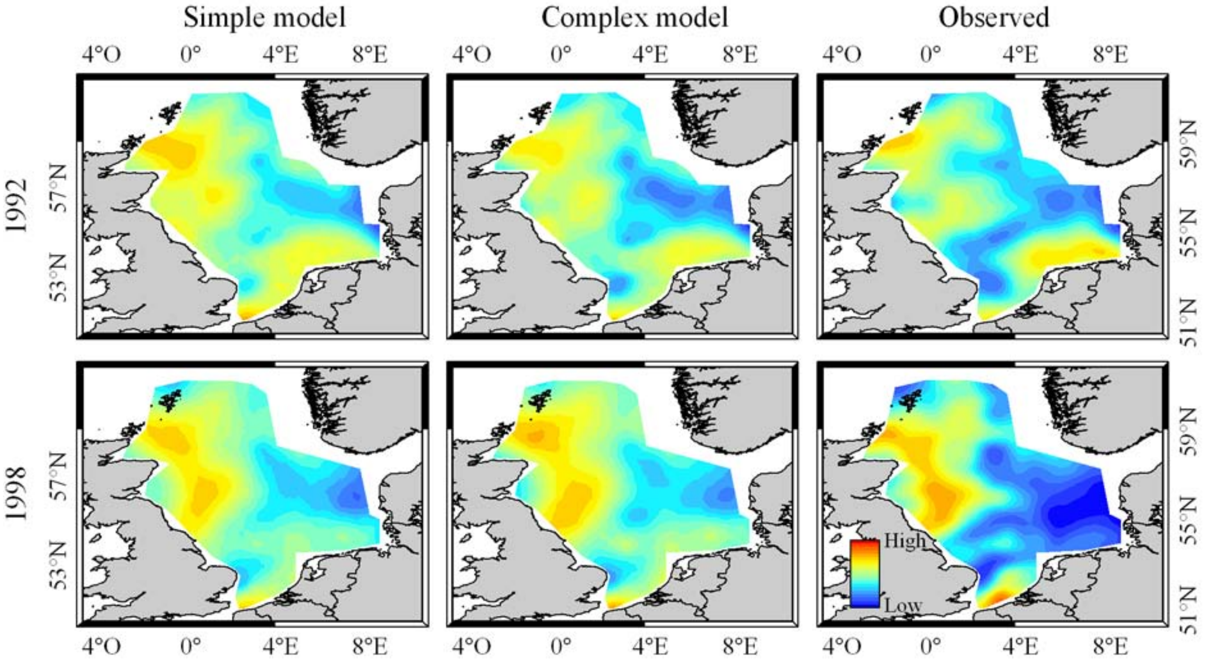


Figure 4.9. Maps of predicted distribution of spawning whiting by two delta models of different complexity. The complex model is the combination of the most complex selected binomial and Gaussian models. The simple model is the combination of the less complex selected binomial and Gaussian models. Observations and predictions are mapped for two contrasting years of the prediction dataset : 1992 and 1998. For a same year, maps of predicted distribution are on the same scale. See table 4.2 for formulas of binomial and Gaussian models used.

Maps of observed distribution for years 1992 and 1998 showed similar pattern of distribution with higher abundances in the northern and southern part of the North Sea. However, areas of distribution in the southern part were not as extensive in 1998 (fig. 4.9) as in 1992. Although the two models were of different complexity levels, the most simple model was as efficient as the most complex one to correctly depict the observed distribution for each year. This suggested that predictions were equivalent whether medium scale in the binomial model and persistent and non-persistent environment in the Gaussian model have been taken into account or not. This proved that only the hypotheses contained in the most simple model played a significant role in the spatial distribution of spawning whiting, i.e. non-persistent environment and population memory for the presence-absence and geography, population size, spatial demography and population memory for the abundances.

4.3.4. Variance partitioning

The percentage explained by the most important hypotheses in the spatial distribution of presence-absence and abundances of whiting were measured using a partial analysis (table 4.4). The amount of presence-absence data variance explained by non-persistent environment and population memory and of positive abundance data variance explained by geography, population size, spatial demography and population memory were computed. All the hypotheses were determined to explain a significant part of the total variation by the RDA analysis (p=0.05). For presence-absence data, however, the total variation explained was very low due to the low proportion of absence in the data. Non-persistent environment and population memory explained an equivalent part of the variation of about 2%. For abundances, 60 % of the observed variation was explained by the four hypotheses with geography, spatial demography and population memory that explained the larger part. These three hypotheses shared a large proportion of the explained variation and once the shared component was removed, only geography and spatial demography still explained more than 10% of the variation. Population size and population memory explained less than 2% of the variation.

Table 4.4. Percentage of variation explained by the most relevant hypotheses in the presence-absence and abundances of whiting in the prediction dataset.

Component studied	Variation (%)	Component studied	Variation (%)
Presence-absence		**Abundances**	
All variation	100	All variation	100
All explained variation	3.4	All explained variation	59.6
Non persistent environment	2.1	Geography	41.8
Population memory	1.9	Population size	2.1
Purely by non persistent environment	1.5	Spatial demography	33.8
Purely by population memory	1.3	Population memory	36.5
Shared component	0.6	Purely by geography	10.5
Unexplained variation	96.6	Purely by population size	1.1
		Purely by spatial demography	10.7
		Purely by population memory	1.8
		Shared component	36.3
		Unexplained variation	40.4

4.4. Discussion

4.4.1. Spatial distribution of spawning adults

Higher proportions of whiting spawning adults are found in the south than in the north of the North Sea which confirms that the spawning varies with the latitude, occurring in February-March in the English Channel and southern North Sea and in April-June in the northern North Sea (Daan *et al.*, 1990; Zheng *et al.*, 2001). In principle, the IBTS should offer an incomplete view of the spawning activity of whiting in the northern area and may only be used to depict its distribution at the beginning of its spawning period. However, at the time of the IBTS, high abundances of spawning adults were already observed both in the north and in the south. These areas were identified as recurrent spawning areas and correspond to historical spawning grounds that have been described (Coull *et al.*, 1998; Gibb *et al.*, 2004). Charrier *et al.* (2007) have shown that these areas are occupied by two distinct sub-populations that can be genetically identified. They have concluded that the Dogger Bank might be a natural barrier which prevents mixing of the two populations and maintains genetic differentiation between them. Spawning of whiting remains clearly absent in the central part and eastern part of the North Sea that may be characterised as unsuitable spawning areas.

4.4.2. Inter-annual variability of whiting spawning distribution

The ratio of the spatial variance over the temporal variance exceeded 1 which indicated that distribution of spawning adults was more variable in space than in time. However, interpolated maps of each year showed temporal variations in observed abundances in the southern and northern areas, in particular a recent depletion in the southern area and a strong contraction of adults distribution in the northern area. The similar behaviour between the spawning stock biomass and the spatial trend during the eighties that became irrelevant during the nineties and 2000's could suggest that (1) there was no density dependent control of whiting distribution, (2) there was a density dependent effect that was only partially shown by the present study, (3) the SSB failed to reflect the current state of the population. The second hypothesis suggested that the SSB may have reached a low threshold after the eighties where density dependence had become a less important process than others factors. Variations in sea surface temperature through the North Atlantic current (Zheng *et al.*, 2002) and the spatial segregation between young individuals confined to shallow areas and old individuals in deep areas (Zheng *et al.*, 2001) have also been proposed to explain these variations in the spatial distribution of whiting. The third hypothesis is supported by the low precision of the estimates of spawning stock biomass which do not correlate well with survey data (Serchuk *et al.*, 1996). Also, considering the whiting population as one stock whereas two distinct populations have been genetically identified (Charrier *et al.*, 2007) may lead to misinterpretation of the respective evolution of their spawning biomass in time.

In the correspondence analysis, the scores of the years on the first axis allowed to distinguish between two distinct periods: 1980-1992 and 1993-2007. The second period being more correlated with stations located in the northern part of the North Sea, suggesting that the linear trend on the first axis highlighted the general decrease of the abundances of spawning adults in the Southern part between 1980 and 2007. On the second axis, the two periods could be further split in two resembling a cyclic pattern. This illustrated that the centroid of the years depicted the centre of gravity of the distribution an alternately moved between the southern and northern parts of the North Sea. This may reveal a temporal shift in the centre of the distribution of spawning adults of whiting during the study period towards the north. This shift was due to alternating period of extension and contraction of adults distribution in the Southern part and to the contraction of the distribution in the Northern part at the end of the study period.

4.4.3. Control of the spatial distribution of spawning whiting

The low number of null abundances (about 6%) revealed that whiting had a wide area of presence and very localised areas of absence. This may explain why only two hypotheses were determined as important to explain the variation in the pattern of presence-absence. The low percentage of explained variation revealed by the variance partitioning highlights that there was a high degree of uncertainty on (1) the relevance of these two hypotheses and (2) the validity of the absence recorded. We argue that the low percentage of explained variation corresponded mostly to noise. Usually, a high number of absence is a statistical problem as it makes the distribution being non normal. However, the present study suggested that a low number of absence may be also a problem as there is not enough variation to extract a significant pattern of distribution.

Geography, population size, age structure and population memory were determined as relevant hypotheses in determining the spatial distribution of whiting abundance. The variance partitioning indicated that geography and spatial demography explained the larger part of variation in the pattern of abundance distribution. The importance of geography combined to population memory highlighted a strong attachment and a high fidelity of whiting to its spawning sites and showed that these areas tended to be maintained in time as abundance of whiting for one year was highly related to that of the previous year. This strategy creates a temporal persistence in the spatial distribution of spawning whiting which could result in a lag period between environmental variations and changes in the spatial distribution, allowing it to face exceptional events (Corten, 2002). In the short run, this can result in a poor annual recruitment because eggs and larvae may not be released each year in optimal conditions for their successful drift; but such strategy may ensure good average recruitment over a long period.

Surprisingly, the present study revealed that non persistent environment expressed through bottom temperature and salinity, was not as so relevant as it was suggested by Zheng *et al.* (2001, 2002). In their study, they used GAM to relate the spatial distribution of whiting to several environmental variables and concluded that the spatial pattern of sea surface temperature had an important influence on the spatial distribution of whiting, in particular during winter and spring. They also found that whiting abundances were higher when the waters were warmer and suggested that this reflected an indirect influence of North Atlantic waters entering the North Sea trough the North Atlantic current. Because the relationship with temperature was no longer significant in summer (Zheng *et al.*, 2002), they concluded that whiting select particular areas which offer suitable conditions of temperature for spawning. In the present study, the non-persistent environment was frequent in selected models, however, it was not present in the most simple one. We argue that non-persistent environment mainly acts as a supplementary hypothesis that do not explain more variation than that has already been explained by the other hypotheses like geography. In this context, we suggest that it was more the spatial effect of the non-persistent environment rather that its pure environmental effect that was determined as relevant by Zheng *et al.* (2002).

Our results also indicated that the demographic status of the population expressed through its size and age structure were relevant to explain the spatial distribution of whiting abundance. Control through population size suggested a density dependent effect that was also partially shown by geostatistical analyses, where areas of high concentrations of spawning adults may vary in their spatial extent according to the annual population abundance. Moreover, there is a spatial segregation between young individuals, that are mainly distributed in shallow areas, and older individuals distributed in deeper areas (Zheng *et al.*, 2001). This age-dependent segregation pattern may be important to explain and predict the spatial distribution of abundance of spawning whiting.

In summary, the spawning distribution of whiting tends to be persistent in time, however, it depicts strong variations in term of abundance distribution, especially in the southern part of the North Sea. Although population size and spatial age segregation have been determined as important, they could not explain on their own the alternating periods of low and high abundances in this area. Furthermore, the present study does not allow to confirm the role of non-persistent environment on whiting distribution. Therefore, the mechanisms that control its spawning distribution and should explain the observed variations in the distribution pattern remain unclear. Several aspects related to the observation time period can make difficult the identification of these mechanisms of control. The start of the spawning period is subjected to the environmental conditions encountered during the previous year and may vary from several days to several weeks. Therefore, studying the spawning distribution of fish population at the beginning of their spawning season might fail to clearly identify mechanisms that determine their distribution. Although whiting has a long spawning period, the present study partially described it by focussing only on its earlier part. All spawning adults may not have completed their spawning migrations. In this context, they may not be installed yet on their spawning grounds which might explain why they are almost present everywhere in the North Sea at the time of the IBTS survey. Because of that, there may be no need to distinguish between presence-absence and positive abundance when studying whiting distribution at the beginning of its spawning period. Moreover, we suggest that studying the distribution of whiting during the peak of spawning in March-April may be more adapted to precisely determine what controls its spawning distribution.

CHAPITRE 5

Spawning distribution of a small pelagic fish: Bay of Biscay anchovy

5.1. Introduction

The European anchovy (*E. encrasicolus*) is a short lived species which mostly grows and become mature during its first year. Its life traits in the Bay of Biscay have been reviewed and it is described as a fast growing, early reproductive and highly fecund species that may live up to 4 or 5 years (ICES, 2004b). Spawning occurs in spring/summer, mainly from April to July with a maximal intensity in May. One year old individuals spawn mainly in June whereas the 2-year old and older individuals spawn earlier in late April-May (Motos *et al.*, 1996).

Spawning stock biomass of anchovy over the continental shelf of Bay of Biscay has strongly decreased between 2000 and 2007 (fig. 5.1).

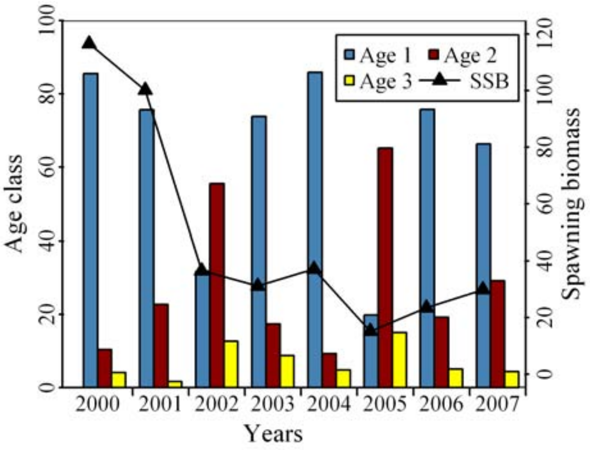


Figure 5.1. Time series of spawning stock biomass (thousand of tonnes) and proportions of one to three years age classes (%) of anchovy in the Bay of Biscay between 2000 and 2007. These estimations were provided by the ICES working group on anchovy (WG ANC, ICES, 2008b).

It was estimated to more 100,000 tons in 2000 and 2001 and fell down to less than 40 000 tons in 2002 and less than 20,000 tons in 2005. Since 2005, the population has still not recovered to an acceptable exploitation level and therefore fisheries have remained closed. Over all the period, spawning population was composed of over 70% of 1-year old individuals (fig. 5.1), except in 2002 and 2005 where 2-years old individuals dominated (between 55 and 65 %) because of poor recruitment in these years. Three-years old individuals were not abundant during this period and contributed to less than 15% to the total number of individuals in the population.

Because anchovy is a short lived and highly exploited species, recruitment plays a major role in setting year-to-year changes in the stock level. Several environmental factors like North-easterly winds and upwelling intensity have been identified to influence recruitment variability (Borja *et al.*, 1998; Borja *et al.*, 2008). Bellier *et al.* (2007a) suggest that the configuration of the spawning habitat also affects recruitment. Therefore, identifying factors that control the spawning habitat of anchovy could bring relevant information on what influences recruitment variability through the spawning process. Spawning habitat of anchovy has been mainly described (Motos *et al.*, 1996; Bellier *et al.*, 2007a) and modelled (Planque *et al.*, 2007) through the spatial distribution of eggs. Recently, Petitgas (2008) confronted several factors related to both environmental and population size control to model the mean spatial pattern of distribution of spawning biomass of adults and its residual variability.

The present chapter will describe, map and model the spatial structure of the Bay of Biscay anchovy in order to identify factors that affect its spawning distribution. In section 5.2, the spatial distribution of adults of anchovy is described and mapped in order to study how its spawning distribution has varied among years between 2000 and 2007. Results on the modelling of the spawning distribution of anchovy using the multi-model approach are given in section 5.3 and discussed in section 5.4.

5.2. Spatial distribution of spawning adults

Total abundances of adults were used to depict the spawning distribution of anchovy in the Bay of Biscay for the period 2000-2007. Its spatial distribution has been described using geostatistical analyses (see section 2.3 for the methodology) and mapped into GIS. Moreover, spatial structures arising from the survey design and relevant to the population distribution have been quantified using PCNM (see section 2.4.1 for the methodology). These PCNM have been used as spatial sub-model for the multi-model approach.

5.2.1. Geostatistics

Eight variograms have been computed and adjusted on the spatial distribution of anchovy between 2000 and 2007 (fig. 5.2). A quadratic trend has been used for most of the years except for 2003. The importance of this spatial trend decreased between 2000 and 2001, followed by a highly variable period between 2001 and 2005 and an increase since 2005. It can be observed that the variations of the spawning stock biomass seem to be correlated to those of the spatial trend, except in 2002 where the value of the trend decreased a year before the spawning stock biomass. With the exception of 2005 where a pure nugget model had to be used, for the remaining years, a combined model with a nugget component and one of the four models of variograms were used to adjust the experimental variogram (see appendix A.2.4). From 2000 to 2003, an exponential model component was used whereas a spherical and pentaspherical models were used for years 2004 and 2006 and a circular model for 2007. There was no trend in the values of nugget according to the years whereas values of the sill showed a clear decrease between 2003 and 2005 and a beginning of increase after 2005.

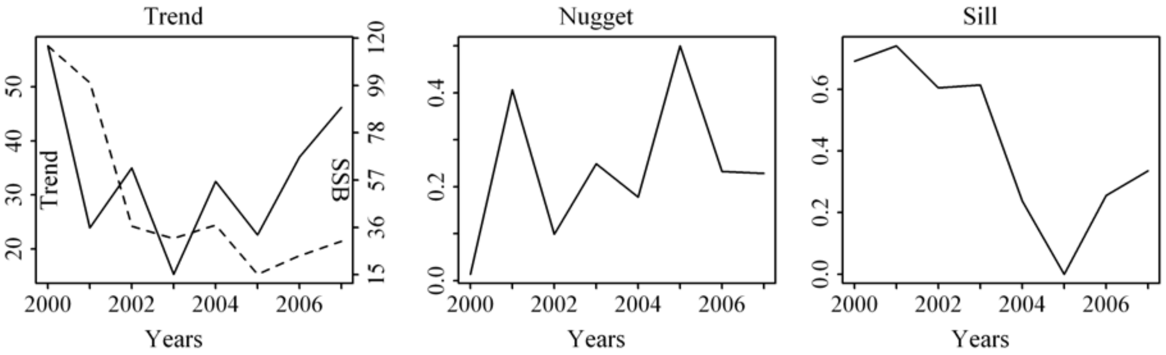


Figure 5.2. Parameter of the variogram. Plots of the percentage of variation explained by the spatial trend, nugget and sill values are plotted for each year between 2000 an 2007. Estimation of the spawning stock biomass (10^3 tonnes, dotted line) for each year is also indicated.

5.2.2. Distribution of spawning adults

Maps of the spatial distribution of spawning adults of anchovy from 2000 to 2007 are shown in fig. 5.3. Overall the study period, adults were mainly concentrated south to 46°N of latitude, whilst being mostly absent of the Northern part, except in 2001 where high concentrations were located along the coast of Brittany. In the southern part, spawning adults were abundant along the coast of les Landes. There was a high degree of inter-annual variability in both the level of abundances and the extent of the spawning areas. The highest abundances were observed in 2000 and 2001 and the lowest abundances in 2005. The spatial distribution in 2003 was patchy and spread over most of the shelf whereas it was much narrower and concentrated to the coast between 2004 and 2006.

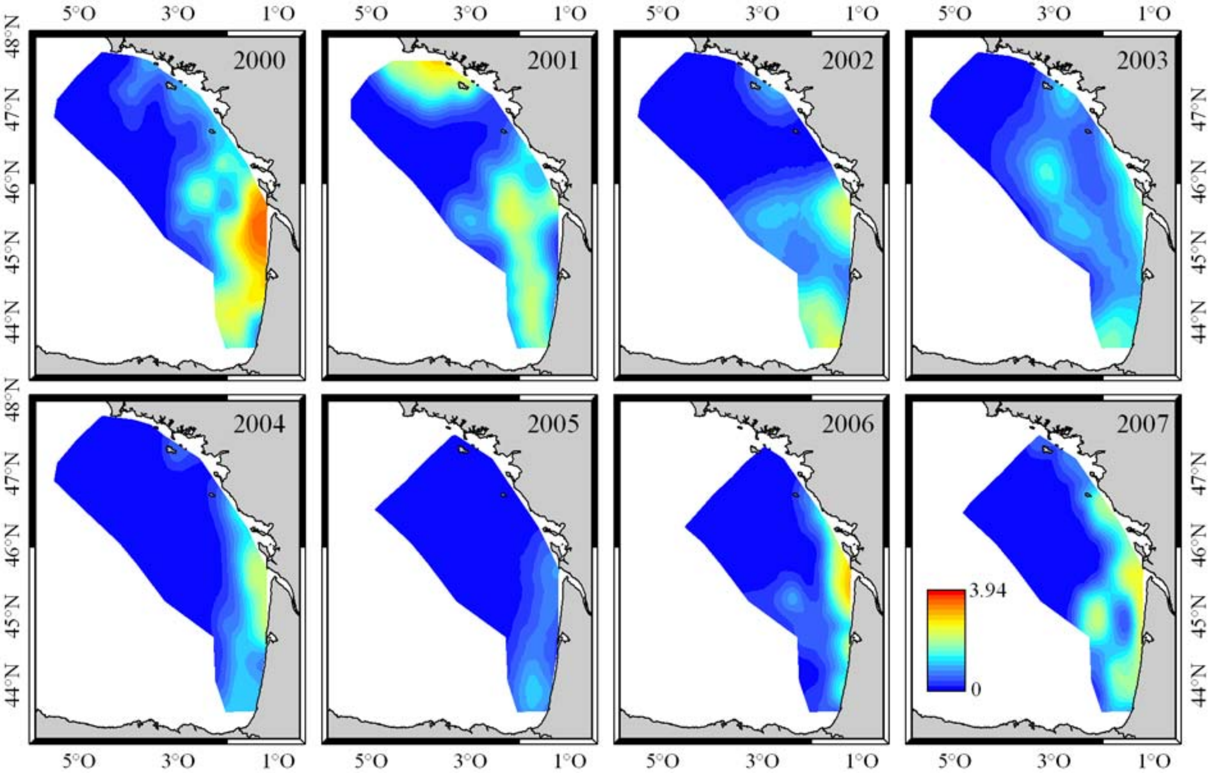


Figure 5.3. Maps of the spatial distribution of spawning anchovy in the Bay of Biscay between 2000 and 2007. A classification of twenty classes of equal interval has been used.

Although the pattern of distribution denotes temporal variations, distribution of spawning adults was more variable in space than in time, with a ratio of the spatial variance over the temporal variance of 1.28. Maps of each year between 2000 and 2007 were used to calculate the average and variance maps. These maps were combined using a threshold of 1 for the mean and 0.8 for the variance to identify preferential, occasional and unfavourable sites for anchovy to spawn (fig. 5.4). Three well-defined preferential spawning sites could be identified. Two were coastal and located in front of the Adour and Gironde estuaries. The third one was more off the coast. Occasional spawning sites consisted in a continuous area that surrounded the preferential sites from the French coast of les Landes to the Gironde estuary. Because observed abundances were low in the Northern part of the Bay of Biscay during most of the study period, this area was characterised as an unfavourable site for spawning.

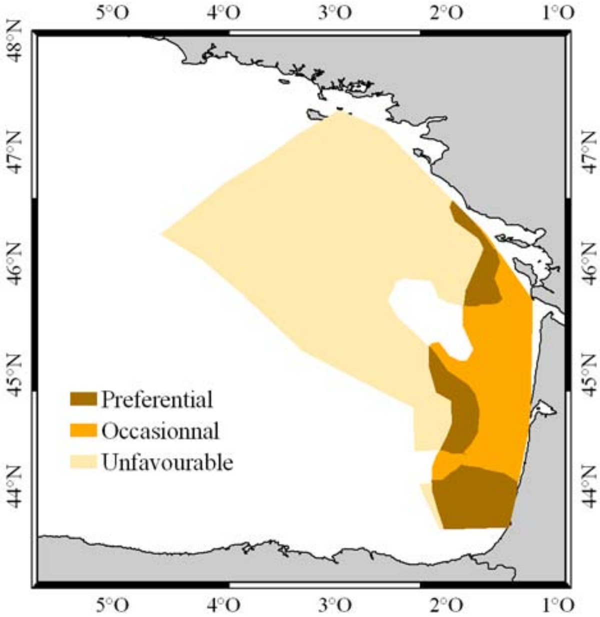


Figure 5.4. Location of preferential, occasional and unfavourable spawning sites for anchovy in the Bay of Biscay. Map was computed from the average and variability maps that were calculated from all distribution maps between 2000 and 2007 (see fig. 5.3).

PCNM were used to describe the scale of spatial variation relevant to the distribution of spawning anchovy. PCNM were computed and extracted using a threshold of 0.2 (decimal degrees of latitude) to truncate the distance matrix between stations. Overall, 105 positive PCNM were extracted from the principal coordinate analysis and 15 were kept at the end of the forward selection (p-value < 0.05 with an adjusted-R^2 of 0.091). From these 15 PCNM (see appendix B.3), 3 spatial scales were identified in the spatial distribution of spawning adults of anchovy (fig. 5.5). Two PCNM were used to build the broad scale sub-model (66-114 km), five for the medium scale sub-model (33-66 km) and eight for the fine scale sub-model (< 33 km). In particular, the first PCNM in the broad scale sub-model depict a clear distinction between the northern and southern part of the Bay of Biscay (fig. 5.5). The medium scale describes smaller structures like the river plume of la Gironde and Adour estuaries. Fine scale is more difficult to interpret but may be related to minor hydrodynamic structures such as upwelling areas along les Landes coast, eddies along the shelf break and the ‘cold pool’, a cold bottom body of water in the northern part of the shelf.

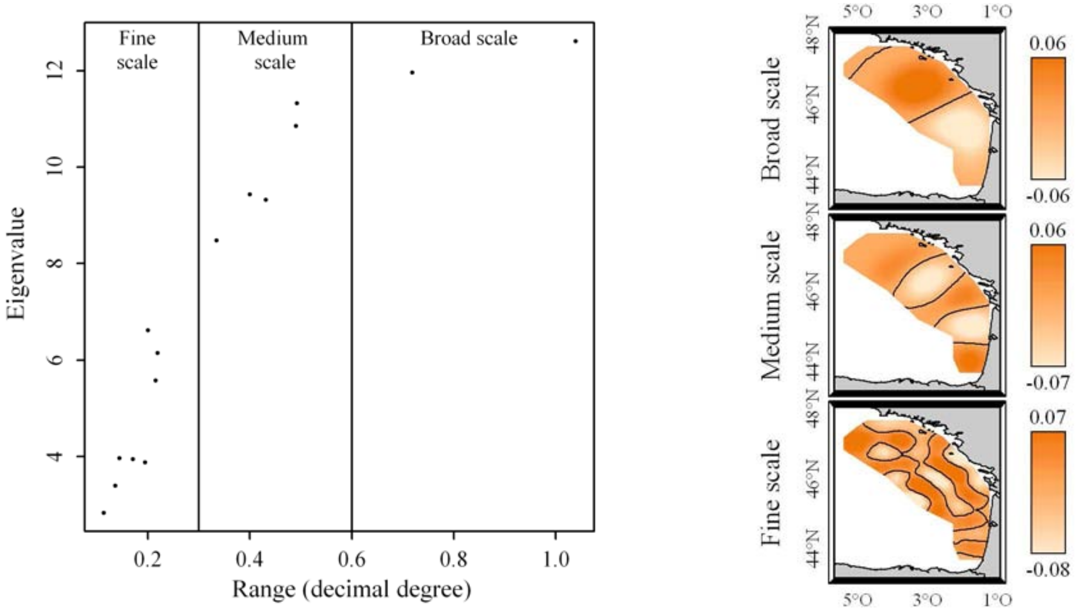


Figure 5.5. Spatial scale sub-models of spawning anchovy distribution. Left : plot of the range values of the 15 selected PCNM against their eigenvalue. Ranges were obtained by fitting a Gaussian model as a theoretical variogram on each spatial predictor. Three scales were determined. Broad scale : eigenvectors with a range between 66 and 114 km (0.6 to 1.1 decimal degrees of latitude); medium scale : eigenvectors with a range between 33 km and 66 km (0.3 to 0.6 decimal degree of latitude); fine scale : eigenvectors with a range between 0 and 33 km (less than 0.3 decimal degrees of latitude). Right : maps of the three scale sub-models. PCNM 1, 6 and 34 were used to depict each scale. Black line indicates the zero value.

5.3. Modelling the spatial distribution of spawning adults

The multi model approach (see section 2.4 for methodology) has been applied to identify factors that are relevant in the control of the spatial distribution of spawning adults of anchovy in the Bay of Biscay. Twelve hypotheses were used to model the distribution of presence-absence of anchovy and thirteen were used to model the distribution of its abundances (see formulas of the various hypotheses in table 5.1). The five environmental hypotheses were determined from the PCA performed on the environmental variables between 2000 and 2007 (see section 2.4.1). Years of the calibration and prediction datasets were also chosen from the PCA according to the variability in term of environmental conditions and population structure (size and age) they encompass. Years 2001, 2003, 2004 and 2005 were used for the calibration dataset and years 2002, 2006 and 2007 were used for the prediction dataset. Overall, 4095 and 8191 possible binomial and Gaussian models were built. Models were selected either on their fitting or predictive ability using a threshold of selection of 60 for the Δ_i and LRT.

Table 5.1. Summary table of the hypotheses used in binomial and Gaussian models for the multi-model approach applied on the spatial distribution of spawning anchovy between 2000 and 2007.

Hypotheses	Binomial models	Gaussian models
G	as.factor(GKn)	as.factor(GKn)
Envt1	FI + UpI + s(SD, 4)	s(FI, 4) + s(UpI,2) + s(SD, 4)
Envt2	s(PD, 4) + s(Ed, 3) + s(PP, 4)	s(PD, 4) + Ed + s(PP, 4)
Envt3	s(PED, 4) + s(ST, 4) + s(TD, 4)	s(PED, 4) + s(ST, 3) + TD
Envt4	s(MLD, 3) + SS + s(BS, 2) + Depth	s(MLD, 4) + s(SS,4) + BS + s(Depth,4)
BT	s(BT, 3)	s(BT, 4)
Bs	s(Ax1, 1) + s(Ax3, 1)	s(Ax1, 2) + s(Ax3, 3)
Ms	s(Ax6, 4) + Ax7 + s(Ax13, 2) + s(Ax14, 4) + s(Ax18, 2)	s(Ax6, 4) + Ax7 + s(Ax13, 4) + Ax14 + s(Ax18, 1)
Fs	s(Ax30, 1) + s(Ax34, 4) + s(Ax39, 4) + s(Ax53, 3) + Ax54 + Ax56 + Ax62 + s(Ax69, 1)	s(Ax30, 2) + s(Ax34, 3) + s(Ax39, 1) + Ax53 + s(Ax54,2) + Ax56 + Ax62 + Ax69
Ps	s(SSB, 3)	s(SSB, 2)
Da	Age1a + Age2a + Age3a	Age1a + Age2a + Age3a
Ds	-	s(Age1s, 4) + Age2s + s(Age3s, 4)
M	s(PYAb, 2)	s(PYAb, 3)

G = geography, GKn = grid knot, FI = frontal index, Upl = upwelling index, SD = salinity difference, PD = pycnocline depth, Ed = eddies, PP = primary production, PED = potential energy deficit, ST = surface temperature, TD = temperature difference, MLD = mixed layer depth, SS = surface salinity, BS = bottom salinity, BT = bottom temperature, Bs = broad scale sub-model, Ms = medium scale sub-model, Fs = fine scale sub-model, Ax = number of the PCNM, Ps = population size, SSB = spawning stock biomass, Da = annual demography, Ds = spatial demography, M = population memory, PYAb = previous year abundance. "s" means that a smoothing function has been used and the corresponding degree of smoothing is indicated in brackets.

5.3.1. Complexity of selected models

In total, 1109 binomial models and 2032 Gaussian models with a Δ_i lower than 60 were selected from fitting. But, 92 binomial models and 2855 Gaussian models with a LRT lower than 60 were selected from prediction. For the two selection procedures, selected binomial models were less complex than selected Gaussian models in term of number of hypotheses and explanatory variables (fig. 5.6). Binomial models selected from fitting were largely more complex than those selected from prediction in term of number of hypotheses and explanatory variables implied. When selected from fitting, their complexity decreased with an increasing Δ_i value, whereas when selected from prediction, models with a LRT value lower than 20 are less complex than those with a LRT value higher than 20. It has to be noticed that the most simple binomial model selected from prediction was the broad scale sub-model with a LRT value of 17.8. All other binomial models selected from prediction within a range of LRT between 0 and 40 contained at least the broad scale sub-model with 1 to 5 additional hypotheses (corresponding to 1 to 18 supplementary variables). Only 12 selected models did not contain the broad scale sub-model and those had a LRT value higher than 40. For selected Gaussian models, there was not so much difference in term of complexity between the two selection procedures (fig. 5.6). Complexity of Gaussian models was higher for models with low value of the selection threshold (between 0 and 10) and slightly decreased with increasing value of the selection threshold.

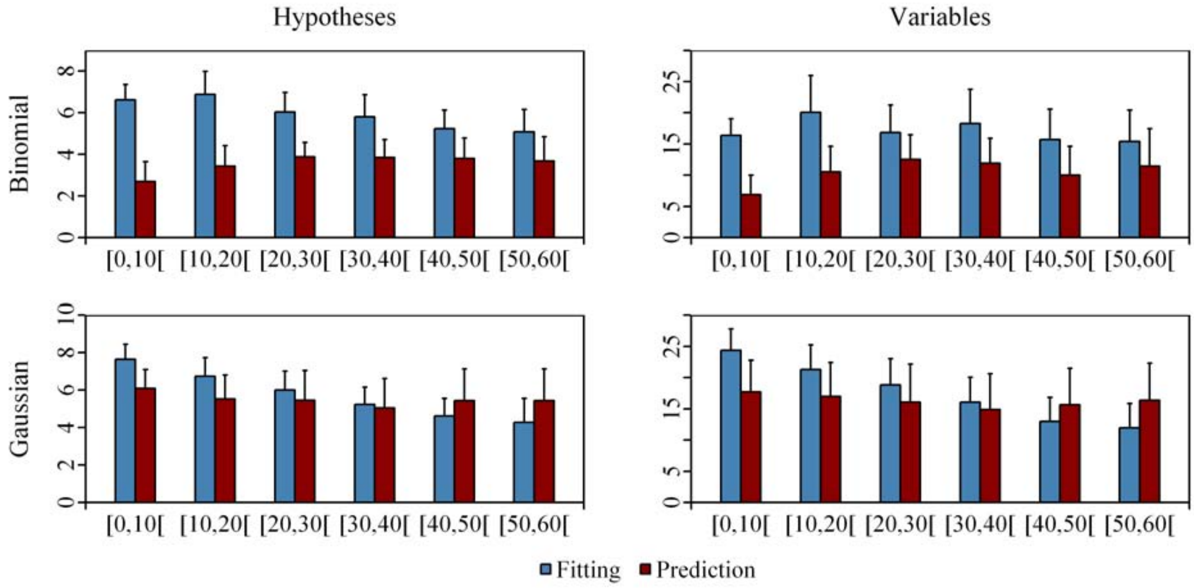


Figure 5.6. Complexity of selected binomial (top) and Gaussian models (bottom) in term of number of hypotheses (left) and explanatory variables (right). Mean and standard deviation are given for models selected from the two selection procedures either based on fitting or prediction and for different range of selection threshold between 0 end 60.

5.3.2. Hypotheses frequencies

Frequencies of the hypotheses within models selected from prediction were given in table 5.2. Five hypotheses were the most frequent in binomial models. The broad scale sub-model was present in 100 % of models selected with a LRT value between 0 and 40 and in more than 75 % of selected models with a LRT value higher than 40.

The hypothesis envt 1 was the second most frequent hypothesis which was present in more than 75% of selected models that also had a LRT value lower than 10. Its frequency decreased as the threshold of selection increased. Medium scale sub-model and population memory were found in less than 75% of selected models with a LRT value lower than 10 whereas their frequency fell under 50% for LRT values higher than 10. Bottom temperature was found in less than 50% of selected model with a LRT value between 0 and 40 and in more than 50% of selected models with a LRT value between 40 and 60.

Seven hypotheses (geography, envt 2, 3 and 4, fine scale, population size and annual demography) were not found in first selected model with a LRT value between 0 and 10. Five of them, the hypotheses envt 2 (depth of pycnocline, index of eddies, index of primary production) and envt 3 (surface temperature, difference of temperature, potential energy deficit), the fine scale sub-model, population size and annual demography, became more frequent in models with higher LRT value. However, they did not reach a frequency higher than 50%. Two of them, geography and envt 4 (mixed layer depth, bottom and surface salinity and depth) were not present in the whole set of selected binomial models with a LRT value between 0 and 60.

Table 5.2. Summary table of the frequencies of the various hypotheses in binomial and Gaussian models selected from prediction for Bay of Biscay anchovy. They give the percentage in which each hypothesis is present in selected models within a LRT range between 0 and 60.

	Binomial models						**Gaussian models**					
	[0-10[	**[0-20[**	**[0-30[**	**[0-40[**	**[0-50[**	**[0-60[**	**[0-10[**	**[0-20[**	**[0-30[**	**[0-40[**	**[0-50[**	**[0-60[**
G	0	0	0	0	0	0	0	0	+	++	++	++
Envt 1	++++	+++	++	++	++	++	++++	+++	+++	++	+++	+++
Envt 2	0	0	0	+	+	+	+++	+++	+++	++	++	++
Envt 3	0	0	+	+	+	+	+	++	++	++	++	++
Envt 4	0	0	0	0	0	0	0	0	+	+	+	+
BT	++	++	++	++	+++	+++	+++	++	+++	+++	++	+++
Bs	++++	++++	++++	++++	++++	++++	++++	+++	+++	+++	+++	+++
Ms	+++	++	++	++	++	++	++++	++++	++++	+++	+++	+++
Fs	0	0	++	++	++	++	++	++	+++	+++	+++	+++
Ps	0	0	0	0	+	+	+++	+++	+++	+++	+++	+++
Da	0	+	+	+	++	++	++++	++++	+++	++	++	++
Ds	-	-	-	-	-	-	++++	++++	++++	++++	++++	++++
M	+++	++	++	++	+++	+++	++	++	+++	++	++	++

0 (absence), + (<25%), ++ (25-50%), +++ (50-75%), ++++ (75-100%). See table 5.1 for formulas of the various hypotheses.

For Gaussian models (table 5.2), 6 hypotheses were present in more than 50 % of selected models with a LRT values between 0 and 60, the hypotheses envt 1, broad and medium scale sub-models, population size and population demography (annual and spatial). Spatial demography was present in 100 % of selected models with an LRT value between 0 and 30 whereas annual demography was present in 100 % of selected models with an LRT value between 0 and 20. Five hypotheses had their frequency varying between 25 and 75% depending on the value of the selection threshold: these were envt 2 and 3, bottom temperature fine scale-sub model and population memory. Geography and envt 4 were not present in first selected models (with a LRT between 0 and 10) and became more frequent in following models. However, their frequency did not reached 50%.

5.3.3. Variance partitioning

Percentage of variation explained by the hypotheses in the presence-absence and abundances of spawning anchovy, was measured using the RDA and partial analysis (see section 2.4.6 for the methodology). Only the most frequent hypotheses within models selected from prediction with a LRT value between 0 and 10 (see table 5.2) were tested for their significance of explained variation. For presence-absence, the hypotheses envt 1, bottom temperature, broad and medium scale sub-models and population memory were used. For abundances, hypotheses envt 1 and 2, bottom temperature, broad, medium and fine scale sub-models, population size, age structure and population memory were tested. Results of the partial analysis are presented in table 5.3.

For presence-absence, only broad scale and bottom temperature were determined as significant by the RDA analysis ($p<0.05$) whereas the others were not found significant ($p>0.05$). These two hypotheses explained 30% of the variance in the presence-absence data of the prediction dataset. The pure effect of the broad scale accounted for 15% of the variance and the pure effect of bottom temperature accounted for 7% of the variance. Within the broad scale, the first PCNM (Ax1) accounted for the major part of the explained variation. About 8% of the variance explained was shared by the broad scale and bottom temperature.

For abundances, spatial demography, the hypothesis envt 1 (frontal and upwelling indexes and difference of salinity) and the broad scale hypothesis were highly significant ($p < 0.01$). The hypothesis envt 2 (pycnocline depth, eddies and primary production), bottom temperature and population size were significant ($p < 0.05$) whereas the others were not ($p > 0.05$). The six hypotheses explained 52.5 % of the variation and shared 19% of the explained variation. Pure environmental effect and pure demographical effect accounted for most of the explained variation, followed by space and population size. Within the environment compartment, frontal index and primary production explained most of the variation.

Table 5.3. Summary table of the percentage of variation explained by the hypotheses in the spatial distribution of presence-absence and abundances of spawning anchovy.

Component studied	Variation (%)	Component studied	Variation (%)
Presence-absence		**Abundances**	
All variation	100	All variation	100
All explained variation	30.4	All explained variation	52.5
Broad scale	23.4	Purely by environment	18.9
Bottom temperature	15.2	*Frontal Index*	7.2
Purely by broad scale	15.1	*Difference of salinity*	1.6
Ax1	15	*Bottom temperature*	1.4
Ax3	0.1	*Primary Production*	5.9
Purely by bottom temperature	6.9	*Shared component*	3
Shared component	8.3	Purely by space (Ax1)	5.1
Unexplained variation	69.6	Purely by spatial demography	19.2
		Purely by population size	3.8
		Shared component	19
		Unexplained variation	47.5

5.3.4. Evaluation of delta-models

Delta-models, resulting from the combination of binomial and Gaussian models selected from prediction, were evaluated for each range of the LRT between 0 and 60 using a Taylor diagram (fig. 5.7). Diagrams showed high variability in term of predictive performance between models of the same range of LRT. The cloud of points was even more spread out as the range of LRT was high. This primarily reflected that models within higher LRT range were more numerous but also that the predictive performance was more variable between these models than between models selected with low LRT values. Although there was a progressive shift towards lower correlation for models with higher LRT range, many models belonging to different LRT range had the same evaluation which reflected that they had an equivalent predictive power. This highlighted that several models with high LRT values performed as well as models with low LRT values.

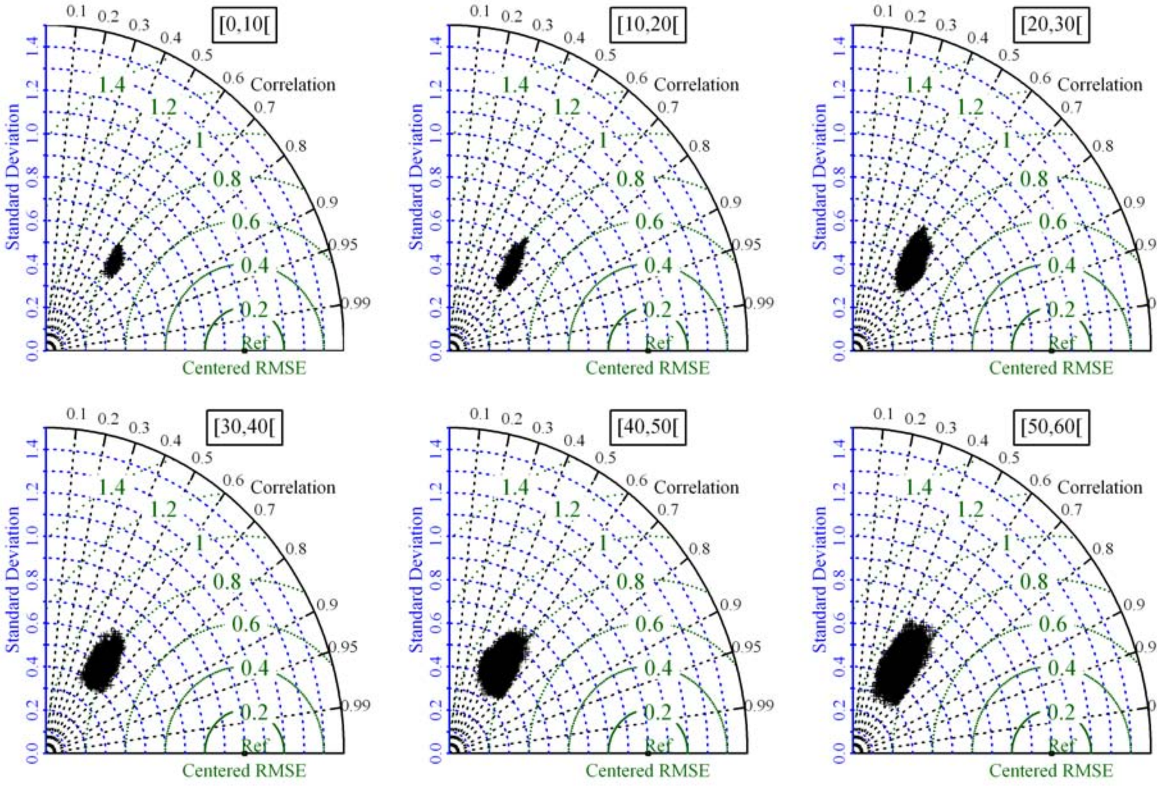


Figure 5.7. Evaluation of delta models resulting from the combination of binomial and Gaussian models selected from prediction for different ranges of LRT value. Observations and predictions of the delta-models are compared using the standard deviation, the root mean square error (RMSE) and the Spearman correlation that are plotted on a Taylor diagram (only the positive correlation part is represented). Observations are symbolised as a point called the reference point and are normalised so that the standard deviation is equal to 1, the correlation equal to 1 (the correlation between the observations and themselves) and the RMSE equal to 0 (the difference

between the observations and themselves). Values of the indexes for the predictions are plotted with the standard deviation on the x-axis, the correlation on the radials of the circle and the RMSE on concentric circles around the reference point.

Evaluation of three delta models of decreasing complexity was also performed (table 5.4). In model (1), all hypotheses that were the most frequent in binomial and Gaussian models selected from prediction were kept. In model (2), only the hypotheses of model (1) that were determined as significant by the RDA analysis were kept. In model (3), only the factors contained in the hypotheses of model (2) that explain on their own a significant part of the variation in observed presence-absence and abundances were kept.

Table 5.4. Summary table giving the formulas of binomial and Gaussian models that were evaluated in fig. 5.8.

Selection based on	Models
(1) Model selection	
Binomial	s(BT,3) + FI +UpI + s(SD, 4) + s(Ax1, 1) + s(Ax3, 1) + s(Ax6, 4) + Ax7 + s(Ax13, 2) + s(Ax14, 4) + s(Ax18, 2) + s(PYB, 2)
Gaussian	s(FI, 4) + s(UpI,2) + s(SD, 4) + s(PD, 4) + Ed + s(PP, 4) + s(BT, 4) + s(Ax1, 2) + s(Ax3, 3) + s(Ax6, 4) + Ax7 + s(Ax13, 4) + Ax14 + s(Ax18, 1) + s(SSB, 2) + Age1a + Age2a + Age3a + s(Age1s, 4) + Age2s + s(Age3s, 4) + s(PYAb, 3)
(2) Model selection + RDA	
Binomial	s(BT,3) + s(Ax1, 1) + s(Ax3, 1)
Gaussian	s(FI, 4) + s(UpI,2) + s(SD, 4) + s(PD, 4) + Ed + s(PP, 4) + s(BT, 4) + s(Ax1, 2) + s(Ax3, 3) + s(SSB,2) + s(Age1s, 4) + Age2s + s(Age3s, 4)
(3) Model selection + RDA + variance partitioning	
Binomial	s(BT, 3) + s(Ax1, 1)
Gaussian	s(FI, 4) + s(SD, 4) + s(PP, 4) + s(BT, 4) + s(Ax1, 2) + s(SSB, 2) + s(Age1s, 4) + Age2s + s(Age3s, 4)

FI = frontal index, Upl = upwelling index, SD = salinity difference, PD = pycnocline depth, Ed = eddies, PP = primary production, PED = potential energy deficit, ST = surface temperature, TD = temperature difference, MLD = mixed layer depth, SS = surface salinity, BS = bottom salinity, BT = bottom temperature, Bs = broad scale sub-model, Ms = medium scale sub-model, Fs = fine scale sub-model, Ax = number of the PCNM, Ps = population size, SSB = spawning stock biomass, Da = annual demography, Ds = spatial demography, M = population memory, PYAb = previous year abundance. “s” means that a smoothing function has been used and the corresponding degree of smoothing is indicated in brackets.

For each of the three models, predictions of the binomial and Gaussian models, applied on the prediction dataset, were multiplied and compared to the observations using a Taylor diagram (fig. 5.8). The Taylor diagram showed a higher standard deviation for model (1) than for model (2) and (3). Also, predictions of model (1) had slightly better correlation and lower RMSE with observations than those of model (2) and (3). Model (2) and (3) had the same correlation however model (2) had a higher standard deviation and a slightly lower RMSE than model (3).

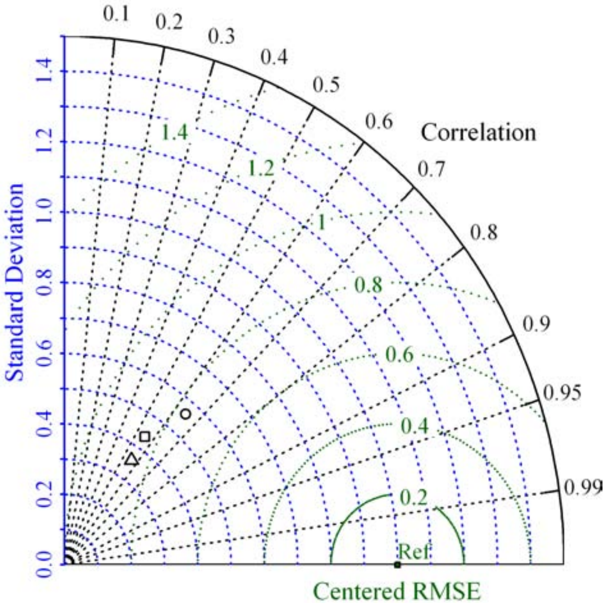


Figure 5.8. Evaluation using the Taylor diagram of the three delta-models of decreasing complexity, model 1 (circle), model 2 (square) and model 3 (triangle) resulting from the combination of binomial and Gaussian models (see table 5.4 for formulas of the models). Observations (Ref) are normalised so that the standard deviation is equal to 1, the correlation equal to 1 (the correlation between the observations and themselves) and the RMSE equal to 0 (the difference between the observations and themselves). Values of the indexes for the predictions are plotted with the standard deviation on the x-axis, the correlation on the radials of the circle and the RMSE on concentric circles around the reference point.

Predictions of the three delta models were mapped for each year of the prediction dataset (fig. 5.9). For each year, predictions in areas outside of the environmental and biological ranges encompassed by the calibration dataset were withdrawn from the maps. Each model was able to locate the two coastal recurrent spawning areas, however they were less efficient to locate the more offshore one. The three models predicted a wider area of distribution than the map of observed data. In 2002, they all succeeded in reflecting the small area of abundances in the south of Brittany and in 2007, model (2) and model (3) predicted the abundances in the North of the Gironde estuary. The three models mostly differed in the variability they were able to reflect which confirms that the more complex the model is, the higher is the standard deviation of the predictions. This showed that withdrawing the non-significant hypotheses mostly impact the variability rather than the predicted global pattern of distribution.

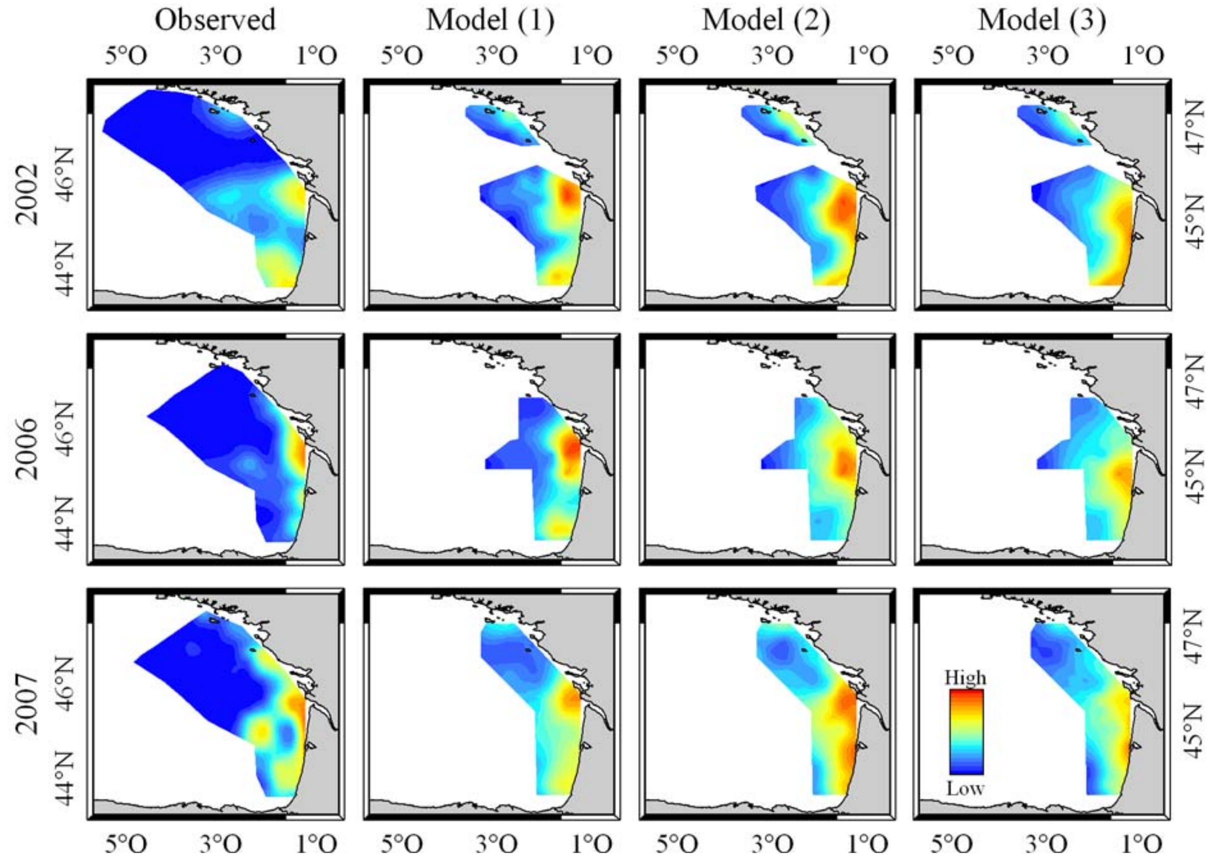


Figure 5.9. Maps of anchovy distribution predicted by the three delta-models of decreasing complexity (model 1, 2 and 3) for the three years of the prediction dataset (see table 5.4 for formulas of the models). Maps of predicted distribution are on the same scale for each year. Maps of observed distribution for the three years are also given.

5.4. Discussion

The spawning distribution of anchovy has shown spatio-temporal variations between 2000 and 2007. However, a mean pattern of distribution with areas of high abundances has been extracted from maps of the different years. Three preferential spawning areas may be identified and corresponded to the distribution of eggs (Motos *et al.*, 1996; Bellier *et al.*, 2007a). At the time of the survey, in late spring and early summer, the spawning distribution of anchovy was preferentially located in the southern part of the Bay of Biscay and high abundances were found in front l'Adour and la Gironde estuaries and along the shelf break.

The spatial pattern of presence-absence of spawning anchovy was determined by two main factors: spatial autocorrelation at broad scale and bottom temperature. Relevance of the broad scale was shown through its high occurrence frequency in selected binomial models as well as its presence in the first selected binomial models. The larger part of the variability it explained in comparison to bottom temperature suggested that the spatial pattern of occupation of anchovy in the Bay of Biscay was better described by large scale autocorrelation patterns than by the environment. The large part of the variation explained by the first PCNM highlighted that it was the spatial segregation between the northern and the southern part of the shelf of the Bay of Biscay that explained most the variability in the pattern of presence-absence. In the present study, the northern part was characterised as an unfavourable site for spawning which is in agreement with what was found in other studies on eggs distribution (Bellier *et al.*, 2007a). Although, another hypothesis may also explain the anchovy observed pattern of occupation. Bottom temperature shared a common effect with spatial segregation at broad scale and was also shown to explain a significant part of variation in the anchovy presence distribution. Whereas it may be surprising that an environmental measure close to the sea-bed may explain a part of the variability in the spatial distribution of a pelagic fish, this result was also found in other studies. Petitgas (2008) has also demonstrated that bottom temperature is an important factor to explain the mean spatial distribution of anchovy biomass in Biscay. Planque *et al.* (2007) modelled the potential spawning habitat of anchovy using the distribution of eggs and found that bottom temperature was the best predictor. They argue that bottom temperature appears as a good proxy for the conditions between the bottom and the thermocline but that it is unlikely that anchovy should be directly dependent on temperature close to the bottom. During the day anchovy forms small soft schools that are distributed between 10 and 25m above the bottom (ICES, 2004b) whereas during the night, it forms small dense schools at the surface. Moreover, at the time of the survey, bottom temperature was warmer south of 46°N and colder on the northern part. The role of bottom temperature and the progressive warming from south to north of the Bay of Biscay might explain the northward extent of the potential spawning habitat noticed by Planque *et al.* (2007). However, the percentage of variation explained by space and bottom temperature showed that the latter had not the same capacity of prediction than space may have had. This may be due to the coverage of the survey which was not sufficient to sample all the spatial variation of bottom temperature and detect its relationship with the spawning distribution of anchovy. Because this may lead to underestimate the full role of bottom temperature, we argue that the study area should be extended to the Spanish coast, to take into account the full gradient of the environmental variables.

Inside the area occupied by spawning anchovy, its abundance spatial distribution seemed mainly explained by environmental conditions. On the thirteen environmental variables used in the present study and grouped into five hypotheses, only a few explained a relevant part of the observed variation in abundances distribution. The present study also offered the opportunity to use environmental factors that were hydrodynamic model outputs. This revealed itself useful as it was shown that some factors related to hydrographic structures may be more relevant to understand the spatial distribution of spawning anchovy compared to classical factors like temperature and salinity. The relevance of frontal index and primary production confirmed that food production and concentration are important features to locate high spawning concentrations. Fronts belong to the 'Ocean triads' (Agostini and Bakun, 2002) of which nutrient enrichment through upwelling system and local retention by eddies are the two other components. In the Bay of Biscay, primary production is mainly enhanced by river outflows that provide nutrients directly from terrigenous sources and through the coastal upwelling between estuaries of Adour and Gironde (Motos *et al.*, 1996). The upwelling system is due to easterly winds that induce horizontal displacement of water masses along the shelf, resulting in a vertical entrainment of deeper ocean water into the offshore flowing plume. However, the present study revealed that upwelling index was a poor predictor which highlighted that high productivity resulting from nutrient enrichment, even if necessary, was not sufficient to provide suitable spawning conditions for anchovy. Spawning distribution of anchovy shows a pattern of concentration/expansion (Motos *et al.*, 1996) where reproductive activity begins into the two coastal preferential spawning grounds identified in the present study and disperse afterwards to occupy the whole Bay of Biscay area. Preferential areas were located in the river plumes of Adour and Gironde estuaries and constituted the two main cores of the spawning distribution of anchovy. More than simply highly productive areas, these river plumes also provide an important stability of the water column which favours nutrient enrichment and concentration of food particles (Motos *et al.*, 1996). Therefore, associated to highly productive areas, thermohaline fronts support high zooplankton productivity and concentration. Eddies are not as important oceanographic features for determining the areas of high concentrations of anchovy eggs as they may be for larvae (Bakun, 2006). This is in good agreement with the observed spatial distribution of eggs, larvae and juveniles described by Irigoien *et al.* (2008) who showed a segregation pattern in the distribution of the larvae according to their size, i.e. larger larvae are more distant to the coast than smaller ones. This supports the idea that it is likely that anchovy favours more the areas that will provide a sufficient food to ensure feeding of adults during the spawning period rather than areas that will provide suitable conditions for released eggs and larvae and their maintaining on the French shelf (Irigoien *et al.*, 2007). This strategy seems to be better adapted than the one where spawning location would vary each year according to the environmental conditions. In a highly variable environment like the Bay of

Biscay, the spawning of anchovy relies on the occurrence each year of some perennial pattern of distribution related to food distribution associated with fronts. Because it primarily ensures the survival of the adults rather than that of their eggs, this strategy may not always result in successful recruitment on unfavourable years but may insure the reproductive success of adults over longer time periods.

Spatial demography related to how the individuals of the various age classes are distributed in space appeared to be a key factor. Motos *et al.* (1996) have described the segregation pattern of distribution between young individuals of 1-year old mostly distributed to the coast in front of the Gironde Estuary and old individuals (2-year old and more) that are distributed further offshore, in front of the Adour estuary and along the shelf break. This spatial segregation between young and old individuals is more related to the size of the individuals than strictly to their age, i.e. young but large individuals may be found offshore mixed with older ones. In this context, the demographic structure of anchovy population act through a size dependant habitat selection more than an age dependent habitat selection. Since large individuals benefit from enhanced energy storing capacity linked to their larger body size, old individuals are able to travel larger distance compared to young individuals. This ability offers an advantage to old individuals over young ones, however it does not explain why they systematically spawn in more offshore areas each year. If competition interaction between young and old individuals is probably one of the causal factors, the presence of old individuals on the shelf break may be explained in term of food resources adapted to the planktinovorous feeding regime of this species. Irigoien *et al.* (2009) have shown a clear spatial pattern between 1998 and 2006 in zooplankton distribution in the Bay of Biscay. The highest biomass of the smaller width class of mesozooplankton were observed closer to the coast whereas the highest biomass of the largest width classes were found near the shelf break. Since larger anchovies need bigger preys, the persistent spatial difference in the size structure of anchovy preys may favour and constrain the segregation pattern between old and young individuals. Potentially, the spawning on the shelf break may lead to an absence of retention on the continental shelf. However, Irigoien *et al.* (2007) argue that this absence of retention does not automatically imply bad trophic conditions and they also suggest that the absence of retention could provide a spatial loophole of lower predation for eggs and larvae. This lower predation risk may lead to low mortality rates experienced by larvae outside of the continental shelf compared to those that are present in the areas of high food concentration on the shelf. Irigoien *et al.* (2007) argue that in a variable environment like the Bay of Biscay, this may be a way to ensure the survival of an offspring and can be an adaptive strategy to prevent from total extinction of the species by providing some individuals for the recruitment even in bad annual environmental conditions. This may also increase the likeliness that released eggs drifting in a large variety of water masses favours the encounter of the hatching larvae with more diverse oceanographic conditions, some of which being more suitable for their survival and growth. This supports the idea that the spatial configuration of the spawning distribution plays a role on the success of the recruitment of the year whilst in turn the recruitment will influence the spawning distribution through the pool of new recruits that will enter in the population (Bellier *et al.*, 2007a).

Planque *et al.* (2007) pointed out that the relationship between the spawning biomass of anchovy and the probability of presence of eggs is almost null. The present study partly confirmed this result by showing that population size was only relevant to explain the spatial distribution of abundance levels but not that of the population presence-absence. This means that a higher number of individuals in the population will affect the level of abundance of spawning anchovy in a given location but will not affect its probability of presence in that place. Bellier *et al.* (2007a) found that the spatial distribution of anchovy has expanded northward between two distinct periods 1967-1972 and 2000-2004, which may suggest that on the long term, year to year variation in the spawning biomass may affect the extent of the occupied spawning area. This extent should not only result from variations in spawning biomass but also from changes in environmental conditions. The fact that population size was not relevant to predict the presence-absence of anchovy between 2000 and 2007 may reflect that anchovy did not experience enough environmental variations and/or was not present in sufficient biomass to cause such important and detectable changes in the extent of the occupied spawning area.

CHAPITRE 6

Discussion générale

Le but de l’étude était de tester les hypothèses de contrôle de la structuration spatiale de la distribution des adultes en ponte. Une approche multi-modèles a été utilisée afin de confronter et classer ces différentes hypothèses. La particularité de l’étude réside dans le fait que quatre grands types de contrôle ont été considérés alors qu’habituellement n’est testé que le contrôle par l’environnement; ces contrôles sont l’attachement géographique ou environnemental des individus, la dépendance spatiale et l’état démographique de la population. Chaque grand type de contrôle a ensuite été décliné en plusieurs hypothèses.

Pour la plie et le merlan, dix hypothèses ont été utilisées: la position géographique, l’environnement persistant et non persistant, trois échelles de dépendance spatiale, la taille de la population, sa structure en âge dans l’espace et dans le temps et l’effet mémoire de la population. Pour l’anchois, le contrôle par l’environnement a été représenté par cinq hypothèses ce qui a aboutit à un nombre total de treize hypothèses utilisées.

Chacune des hypothèses a été modélisée selon plusieurs variables grâce aux modèles additifs généralisés. Ces derniers ont été choisis car ils sont une des techniques les plus couramment utilisées du fait de leur fonction de lissage qui leur donne l’avantage de pouvoir utiliser des courbes de réponses plus souples et mieux adaptées aux données. Les hypothèses modélisées ont ensuite été combinées en modèles plus complexes qui ont été comparés entre eux, classifiés et sélectionnés d’après leur capacité à reproduire les distributions observées. Les hypothèses importantes sont alors extraites à partir d’un ensemble de modèles déterminés comme les meilleurs modèles.

La construction de deux types de modèles, d’une part les modèles binomiaux et d’autre part les modèles Gaussiens a permis de distinguer les hypothèses qui sont importantes pour la présence-absence de celles qui sont importantes pour l’abondance des populations en train de se reproduire. Enfin, ces deux types de modèles ont été évalués par un diagramme de Taylor en les combinant en modèle delta où les probabilités de présence prédites par chaque modèle binomial sélectionné sont multipliées par les abondances prédites par chaque modèle Gaussien sélectionné.

6.1. Apports de l’approche multi-modèles

Le principe de l’approche multi-modèles est de comparer un ensemble de modèles indépendamment les uns des autres en utilisant une mesure de distance, entre les données observées et les données prédites par chaque modèle. La capacité des modèles à refléter les distributions observées a été étudiée selon leur ajustement et leur capacité de prédiction. L’ajustement des modèles correspond à leur capacité à reproduire les données avec lesquelles ils ont été construits. Il a été mesuré pour chaque modèle grâce au critère d’AIC qui permet de pénaliser l’ajustement des modèles selon leur complexité. Le degré de prédiction des modèles correspond à leur capacité à prédire la distribution observée d’un autre jeu de donnée que celui qui a été utilisé pour les construire. Le critère de log de vraisemblance a été utilisé pour mesurer cette capacité prédictive de chaque modèle. Une fois la valeur du critère calculé pour chaque modèle, les modèles sont alors comparés entre eux en mesurant la distance qui les sépare du modèle déterminé comme le meilleur, qui est celui qui possède la plus petite distance. Les modèles sont classés selon l’augmentation de distance qu’ils engendrent vis à vis de ce modèle. Pour l’ajustement, l’augmentation de distance entre les modèles a été mesurée en utilisant le Δ_i tandis que pour la prédiction c’est le rapport des log de vraisemblance qui a été utilisé.

6.1.1. La sélection basée sur l’ajustement

Pour les trois espèces étudiées, il a été montré que les modèles déterminés comme les meilleurs d’après leur ajustement aux données existantes n’étaient pas ceux déterminés comme les meilleurs d’après leur capacité à prédire. De plus, bien que le critère d’AIC soit spécialement conçu pour pénaliser les modèles en fonction de leur degré de complexité, les modèles sélectionnés par l’ajustement sont au moins aussi, voire plus complexes que les modèles sélectionnés d’après la prédiction (Loots *et al.*, sous presse; sections 4.3.1 et 5.3.1). Ceci suggère que malgré son aptitude à favoriser les modèles les plus parcimonieux, le critère d’AIC ne permet pas totalement d’éviter le sur-ajustement aux données due aux caractéristiques intrinsèques du processus de sélection basée sur l’ajustement. Notamment, l’autocorrélation temporelle et spatiale dans les données d’ajustement fait que le nombre de degrés de liberté (nombre d’observations indépendantes) est surestimé (Legendre, 1993). Ce phénomène va alors surévaluer le premier terme de l’AIC (l’ajustement du modèle) par rapport au deuxième terme (la parcimonie du modèle) et ainsi promouvoir les modèles bien ajustés aux données sans une pénalisation suffisante de leur complexité (parcimonie). Dans ce contexte, les modèles les plus complexes étant ceux qui s'ajustent le mieux aux données, ils seront alors classés et sélectionnés en priorité.

Burnham et Anderson (2002) n'excluent pas la possibilité que le critère d'AIC puisse mener à la sélection de modèles trop complexes. Pour limiter ce problème, ils conseillent de ne pas utiliser plus de paramètres qu'il n'en est réellement nécessaire pour la construction des modèles. Dans cette étude, pour éviter un nombre trop important de modèles (Diniz-Filho *et al.*, 2008), les modèles ont été construits sur la base d'hypothèses plutôt qu'à partir de variables explicatives prises individuellement. Si l'hypothèse est retenue, toutes les variables qui y sont incluses sont prises en compte. Deuxièmement, dans les fonctions de lissage utilisées dans les GAM, tous les degrés de lissage n'ont pas été considérés lors de la construction des modèles mais ont été déterminés avant la combinaison des différentes hypothèses, ceci permet également de limiter le nombre final de modèles.

La méthode de sélection basée sur l'ajustement aux données ne permet pas d'identifier correctement les hypothèses importantes. En effet, elle peut amener à surestimer l'importance d'hypothèses plus complexes comme la fine échelle spatiale (issue des PCNM), et à l'inverse sous-estimer l'importance d'hypothèses très simples comme la position géographique (Loots *et al.*, sous presse). En effet, la sélection basée sur l'ajustement a tendance à promouvoir les modèles les plus complexes, soit ceux qui contiennent les hypothèses contenant un grand nombre de variables, sans pour autant signifier que ces hypothèses expliquent une part de variation pertinente dans la distribution observée. En revanche, pour des variables catégoriques comme le carré statistique, utilisé pour représenter la position géographique dans le cas de la plie et du merlan, chaque catégorie est comptabilisée comme un paramètre du modèle, ce qui tend fortement à pénaliser cette hypothèse car les modèles qui la contiennent voient leur AIC s'accroître rapidement.

6.1.2. La sélection basée sur la prédiction

Il a été montré que les modèles sélectionnés d'après la prédiction étaient plus simples (Loots *et al.*, sous presse; section 4.3.1) ou au moins aussi simples (section 5.3.1) que ceux sélectionnés d'après l'ajustement. On pouvait alors supposer qu'il s'agissait là d'un biais dû à la taille du jeu de données de calibration qui était plus grand que celui de prédiction. Cependant, dans le cas d'application sur le merlan, le jeu de donnée de prédiction était beaucoup plus grand et pourtant les modèles sélectionnés d'après la prédiction étaient également en moyenne plus simples que ceux sélectionnés d'après la calibration.

La sélection des modèles par leur application à un jeu de données autre que celui utilisé pour les calibrer permet donc d'éliminer en partie le problème de l'autocorrélation spatiale des données qui tendait à faire que les modèles les plus complexes étaient sélectionnés en priorité. Ce problème ne se pose pas lorsque l'on teste la performance de la prédiction en utilisant le rapport des log de vraisemblance car ce critère de sélection ne dépend pas du nombre réel d'observations indépendantes dans le jeu de données de calibration. De plus, lorsque la procédure de sélection des modèles est basée sur la prédiction, le nombre de variables n'est pas pris en compte dans le critère de sélection et donc des hypothèses comme la position géographique ne sont pas plus pénalisées que les autres hypothèses.

L'utilisation du log de vraisemblance comme critère de sélection des modèles d'après la prédiction ne permet pas pour autant d'identifier directement les hypothèses les plus importantes. Pour les trois espèces étudiées, il a été montré que les hypothèses qui avaient été déterminées comme les plus fréquentes à partir des modèles sélectionnés d'après la prédiction n'ont pas toujours été déterminées, par la partition de variance et par l'analyse de redondance, comme expliquant une part significative de la variance de la réponse de l'espèce. Il faut noter que l'analyse de redondance est plus restrictive que les modèles additifs généralisées dans le sens où elle basée sur une combinaison linéaire des variables. Elle ne permet donc pas de prendre en compte des types de réponse plus complexes alors que cela s'avère souvent nécessaire comme l'ont montré les degré de lissage utilisés au cours de cette étude pour la modélisation des hypothèses. En effet, certaines hypothèses générant des réponses non linéaires de la distribution de l'espèce, elles risquent d'être considérées comme non significatives dans une analyse de redondance. Cependant l'objectif n'étant pas de comparer les deux méthodes, ces résultats suggèrent que certaines hypothèses, bien que fréquentes dans les modèles sélectionnés d'après la prédiction, expliquent en fait une part très faible et peu-être non significative de la variabilité dans la distribution observée.

La sélection basée sur la prédiction ne permettrait donc pas d'isoler efficacement les hypothèses qui déterminent la distribution spatiale pour les trois espèces étudiées. En effet, dans le cas du merlan et de l'anchois, les modèles delta, issus de la combinaison de modèles ayant des vraisemblances très différentes, ont une évaluation similaire (diagramme de Taylor) et ont produit des distributions prédites de qualités équivalentes (section 4.3.3 et 5.3.4). De plus, les modèles binomiaux et Gaussiens sont développés indépendamment les uns des autres pour prédire la présence-absence et l'abondance et le fait qu'ils contiennent des hypothèses qui n'expliquent que peu de variation n'a pas d'influence sur leur classification car le nombre d'hypothèses ne pénalise pas le rapport des log de vraisemblance. Quand ces modèles sont finalement re-combinés en modèle delta, les prédictions produites par des modèles contenant ou non ces hypothèses sont de qualité équivalentes, ces dernières n'apportant pas d'information supplémentaire.

6.1.3. Influence du jeu de données

Nous avons montré que les modèles sélectionnés pouvaient être différents selon la période de temps sur laquelle ils étaient calibrés ou appliqués. Ceci a permis de montrer que certaines hypothèses n'avaient pas la même importance en fonction de la période de temps considérée (Loots *et al.*, sous presse). Par exemple, les conditions de l'environnement non-persistant pour le début de la période étudiée pour la plie (1980-1990) sont capables d'expliquer la distribution observée de la plie dans les années 2000. Cependant, lorsque les modèles sont calibrés sur la fin de la période et appliqués sur la période des années 1980, l'environnement persistant n'a plus aucune capacité explicative. A l'inverse, la position géographique et la démographie annuelle ont été déterminées comme plus importantes pour expliquer la distribution spatiale des années 1980 que celle des années 1990 et 2000.

Il a également été montré qu'un jeu de donnée de prédiction trop petit comme celui utilisé pour l'anchois pouvait conduire à sélectionner des modèles Gaussiens très complexes alors que des modèles plus simples avaient une capacité prédictive équivalente (section 5.3.4). Le faible degré de variabilité contenu dans un jeu de donnée trop petit, rend inefficace la sélection basée sur la prédiction pour séparer et classer les modèles correctement; ceci aboutit au final à la sélection d'un très grand nombre de modèles. La taille des jeux de données de calibration et de prédiction doit donc être suffisante de façon à capturer une variabilité suffisamment grande dans la réponse modélisée et dans les variables explicatives utilisées pour permettre de séparer les hypothèses et sélectionner un ensemble correct de modèles.

6.1.4. Seuil de sélection

Après avoir été classés selon leur performance d'ajustement ou de prédiction, les modèles sont ensuite sélectionnés selon un seuil dans la distance (Δ_i ou LRT) qui les sépare. En dessous de ce seuil, les modèles sont considérés comme équivalents et le meilleur modèle peut potentiellement se trouver parmi eux. Burnham et Anderson (2002) affirment qu'une augmentation de distance entre les modèles supérieure à 10 en terme de Δ_i traduit une différence significative dans la capacité des modèles à refléter correctement les données observées. Les résultats obtenus ont montré que les modèles sélectionnés selon un seuil de sélection inférieur à 10 avaient une capacité de prédiction équivalente (Loots *et al.*, sous presse). De plus, il a également été montré que des modèles ayant un seuil de sélection supérieur à 10 pouvaient non seulement être plus simples mais également avoir une capacité de prédiction équivalente avec des modèles ayant un seuil inférieur à 10 (voir section 4.3.3 et 5.3.4). En effet, pour le merlan et l'anchois, nous avons choisi une gamme de seuil de sélection de 0 à 60 afin de comprendre comment cela pouvait influencer la sélection des modèles. Les modèles issus de cette sélection ont alors été combinés en modèles delta en multipliant les probabilités de présence prédites par les modèles binomiaux par les abondances prédites par les modèles Gaussiens. L'évaluation de ces modèles delta par le diagramme de Taylor a montré que beaucoup de prédictions avaient des corrélations semblables avec les observations. Ainsi, une différence entre les modèles dans leur valeur de critère ne traduit pas automatiquement des différences significatives dans leur capacité à refléter les distributions observées.

On peut alors s'interroger sur la valeur de seuil à utiliser pour sélectionner les modèles. Dans le cas du merlan, nous avons montré que le seuil de sélection de 10 permettait de choisir un ensemble de modèles binomiaux contenant le modèle le plus simple. Dans ce cas précis, la restriction à 10 du seuil de sélection permet donc de retenir un modèle contenant un petit nombre d'hypothèse et qui a un pouvoir prédictif équivalent avec d'autre modèles plus complexes. Néanmoins, pour les modèles Gaussiens, le modèle le plus simple n'était pas parmi les modèles ayant un seuil de sélection inférieur à 10. Dans ce cas, choisir un seuil trop restrictif conduit à exclure un modèle simple qui a une capacité prédictive similaire à d'autres modèles plus compliqués. Si le seuil de sélection peut influencer la complexité des modèles sélectionnés, en revanche, il a été montré qu'il n'avait que peu d'impact sur l'importance des hypothèses (section 4.3.2 et 5.3.2). En effet, les hypothèses les plus fréquentes, parmi les modèles sélectionnés pour des valeurs de seuil faibles, restent les plus fréquentes dans les modèles avec des valeurs supérieures de seuil. Ainsi, le fait d'augmenter le seuil de sélection ne conduira pas à mésestimer les hypothèses les plus importantes mais pourra en revanche conduire à attribuer plus d'importance à certaines d'entre elles qu'elles n'en ont en réalité.

6.1.5. Amélioration de l’approche multi-modèles

De manière à sélectionner en priorité les modèles qui prédisent le mieux la distribution observée avec un minimum d’hypothèses, il serait intéressant d’utiliser un autre critère que le log de vraisemblance si on veut pouvoir déterminer directement l’importance des hypothèses à partir des modèles sélectionnés avec un seuil raisonnable de sélection entre 0 et 10. Le critère d’AIC semble plus adapté. Il permettrait de pénaliser les modèles par leur complexité et il ne conduirait pas à la sélection de modèles sur-ajustés puisqu’il a été montré que la sélection basée sur la prédiction était un bon moyen d’atténuer l’effet de l’autocorrélation dans les données. On peut alors imaginer une approche où les modèles sont ajustés sur un jeu de données de calibration mais où la sélection des hypothèses se fait sur un jeu de données de prédiction basée sur le critère d’AIC.

De plus, l’approche multi-modèles utilisée au cours de cette étude s’est révélée fastidieuse car elle implique de classer et de sélectionner les modèles binomiaux d’une part et les modèles Gaussiens d’autre part, du fait de l’absence de normalité dans les données. Cependant, il peut arriver que l’on soit intéressé uniquement par les facteurs qui contrôlent la distribution globale des individus sans forcément distinguer ceux qui contrôlent leur présence-absence de ceux qui contrôlent leur abondance. Une manière de s’affranchir de cette sélection en deux temps serait d’utiliser directement l’évaluation des modèles delta plutôt que de l’utiliser à posteriori comme il a été fait au cours de l’étude. Ainsi, les modèles delta pourraient alors être comparés suivant les trois critères du diagramme de Taylor et l’identification des hypothèses importantes se ferait alors d’après l’ensemble des modèles delta déterminés comme les meilleurs d’après ces trois critères. Ceci permettrait également de sélectionner les meilleurs modèles binomiaux et Gaussiens non plus de manière indépendante mais simultanée.

6.2. Habitat de reproduction

Le terme générique d’habitat désigne entre autres la projection géographique sur une carte des conditions environnementales favorables à la réalisation d’une fonction biologique (Coyne et Christensen, 1997). La sélection d’un habitat de reproduction repose sur l’aptitude des adultes à sélectionner les milieux afin de maximiser leur aptitude à se reproduire (Orians et Wittenberger, 1991). Cette sélection n’implique pas de mécanisme conscient de la part des individus mais signifie que la distribution observée des individus n’est pas le résultat du hasard (Boulinier *et al.*, 2005). Le but de l'étude ici n'est pas de démontrer qu'il y a effectivement une sélection consciente de la part des individus mais de montrer que l'habitat de reproduction est structuré dans l'espace et que cette structuration est influencée par plusieurs types de contrôle. Même si cette structuration importante de l'habitat de reproduction dans l'espace peut varier pendant la saison pour des espèces comme l'anchois qui ont une période de ponte étalée, elle tend néanmoins à persister à l'échelle de plusieurs années. Si l'on veut correctement appréhender les changements qui peuvent se produire dans la distribution des adultes reproducteurs, il est nécessaire d'identifier les facteurs qui créent la structure spatiale de l'habitat de reproduction et ceux qui tendent à favoriser sa persistance au fil des années.

6.2.1. Distribution des oeufs et distribution des adultes en reproduction

Plusieurs termes sont habituellement utilisés pour décrire la distribution spatiale de la reproduction chez les poissons. Ainsi, on parle généralement de frayères, d'habitat de ponte et d'habitat de reproduction. Ces trois notions assument que les moyens utilisés pour identifier et localiser géographiquement les habitats de ponte ne sont pas les mêmes. Comme nous l'avons vu pour les trois espèces étudiées, la localisation spatiale de la reproduction se fait généralement grâce à la distribution des œufs. Le terme de frayères désigne les zones géographiques où a lieu le frai, c'est à dire la rencontre des mâles et des femelles en un même endroit pour se reproduire. Dans le cas où ils sont pélagiques, les œufs dérivent selon les conditions hydrodynamiques locales de la zone dans laquelle ils ont été pondus. A moins d'être certain que les œufs collectés aient été récemment pondus et aient peu dérivé, il est donc risqué d'utiliser leur distribution pour en déduire la localisation des zones de frayères. L'identification des conditions environnementales affectant la survie des œufs et larves est donc à relier à l'habitat de succès de ponte tandis que la distribution observée des adultes reproducteurs permettra de caractériser les habitats potentiel et réalisé de ponte. Les termes habitat de ponte et habitat de reproduction peuvent être équivalents dans la mesure où les mâles et les femelles doivent être au même endroit (à l'exception peut être des sélaciens) pour favoriser la rencontre entre les gamètes mâles et femelles et permettre la fécondation des œufs quand ils sont émis dans l'eau.

6.2.2. Le rôle des différentes hypothèses

Parmi les différentes hypothèses qui ont été modélisées, toutes ne sont pas importantes dans le contrôle de la distribution des adultes reproducteurs (table 6.1). L’importance des différentes hypothèses varie selon l’espèce et selon le type de réponse étudiée. C’est majoritairement l’environnement, qu’il soit stable ou variable dans le temps, qui détermine la présence des espèces. Pour la plie et l’anchois, la grande échelle spatiale est également importante tandis que pour la plie, l’effet mémoire joue également un rôle. Le niveau d’abondance avec lequel les zones sont occupées est, quant à lui, déterminé par des facteurs internes à la population tels que la taille et la structure en âge. Pour la plie et le merlan, l’effet mémoire et la position géographique sont également très importants, ce qui montre leur attachement très fort à leurs sites de reproduction. Dans le cas de l’anchois, l’environnement non persistant joue, de nouveau, un rôle important et notamment la quantité de nourriture disponible pour soutenir les pontes successives pendant toute la période de reproduction. Pour l’anchois, ni l’effet mémoire ni la position géographique n’ont été déterminés comme importants, ce qui dénote le caractère moins conservateur de l’anchois pour la sélection des sites de reproduction.

Table 6.1. Tableau récapitulatif de l’importance relative des différentes hypothèses de contrôle pour la présence-absence (*p-a*) et l’abondances positive (*log(Ab+)*) pour les trois espèces étudiées.

	Plie		**Merlan**		**Anchois**	
	p-a	**log(Ab+)**	**p-a***	**log(Ab+)**	**p-a**	**log(Ab+)**
Position géographique	0	+++	0	+++	0	0
Environnement						
persistant	+++	0	0	0	0**	0**
non persistant	0	0	+	0	+++	+++
Echelle spatiale						
Grande	+++	+++	0	0	+++	++
Moyenne	+++	+++	0	0	0	0
Petite	0	0	0	0	0	0
Taille de la population	+	+++	0	++	0	++
Structure en âge						
annuelle	0	++	0	0	0	0
spatiale	-	+++	-	+++	-	+++
Mémoire de la population	+++	+++	+	+++	0	0

* faible proportion d’absences
** uniquement la profondeur testée

6.2.3. Un habitat structuré dans l'espace…

La cartographie de la distribution spatiale des adultes reproducteurs a montré que l'habitat de reproduction était très structuré dans l'espace avec des zones où les individus sont très abondants (habitats préférentiels), des zones où les individus sont moins abondants (habitats occasionnels) et des zones où les individus sont peu abondants voire absents (habitats non favorables).

Il est généralement admis que c'est l'hétérogénéité spatiale du milieu dans lequel vivent les individus qui structure les habitats et en particulier l'habitat de reproduction (Orians et Wittenberger, 1991). En effet, si le milieu était homogène, il n'y aurait pas lieu de choisir une zone plutôt qu'une autre car partout la capacité des individus à maximiser leur aptitude à se reproduire serait la même. L'hétérogénéité du milieu dans lequel vivent les individus peut être exprimée soit à travers des conditions qui varient seulement dans l'espace soit par des conditions qui varient à la fois dans l'espace et dans le temps. L'habitat de reproduction, pour les populations benthiques et démersales comme la plie et le merlan, est majoritairement déterminé par des contraintes environnementales qui ne varient que dans l'espace. Des variables telles que la température ou la salinité, qui varient à la fois dans l'espace et dans le temps, semblent ne pas être déterminantes pour les individus, tout du moins au moment de se reproduire. En revanche, l'habitat de reproduction d'une espèce pélagique, comme l'anchois, est structuré par des caractéristiques de l'environnement qui varient à la fois dans l'espace et dans le temps (exemple de la température de fond). Même dans ce cas, il a été montré que la température de fond dans le Golfe de Gascogne était l'un des paramètres dont la structure dans l'espace variait le moins dans le temps (Planque *et al.*, 2006; Planque *et al.*, 2007), ce qui relativise son caractère non persistant.

Les fortes abondances d'individus qui caractérisent les habitats préférentiels sont dues à la concentration des individus au niveau des frayères. Le caractère agrégatif des individus à micro-échelle est donc également un facteur qui participe à la structuration spatiale à méso-échelle de l'habitat de reproduction. Cette structure à méso-échelle a été démontrée notamment à travers l'importance de la grande et la moyenne échelle décrite par les PCNM. Les adultes reproducteurs de la plie et du merlan s'agrègent dans des zones de faibles profondeurs, dans le sud de la mer du Nord et le long des côtes du Royaume-Uni. Dans ces zones, les conditions de courant permettent la dérive des œufs et des jeunes larves jusqu'aux nourriceries côtières. Par contre, le comportement de vie en banc de l'anchois crée une structure de l'habitat de reproduction qui est beaucoup plus fine où les adultes sont concentrés au niveau des zones de forte production planctonique, ce qui permet d’assurer la survie des adultes et l’émission d’une quantité importante de gamètes.

Le terme d’habitat « non favorable », utilisé par Bellier *et al.* (2007a), serait mal adapté dans le cas de la reproduction, car il sous-entend que l’absence ou la faible abondances d’individus dans certaines zones traduit le fait que ces zones n’offrent pas les conditions nécessaires pour la reproduction. Ainsi, Certain *et al.* (2007) préfère utiliser le terme d’ « habitat évité » plutôt que celui de non favorable. Le terme d’ « habitat évité » traduit une absence d’individus à un endroit donné sans pour autant vouloir dire que cet endroit n’est pas favorable. En effet, l’absence d’individus enregistrée à certains endroits est toujours difficile à interpréter car il existe des situations où des zones favorables à la reproduction peuvent ne pas être occupées ou faiblement occupées. Pour que les zones favorables soient occupées, il faut tout d’abord que les individus aient connaissance de l’existence de ces zones, ce qui nécessite que la capacité de prospection des individus soit au moins égale à l’échelle de l’hétérogénéité spatiale de l’environnement. De cette manière, les individus rencontrent des zones de qualité différente au sein du milieu et sélectionnent alors celles qui leur sont le plus favorables (Boulinier et Lemel, 1996). Si des zones favorables se trouvent en dehors de la zone que les individus sont capables de couvrir par leur migration, elles ne seront pas occupées. De même, l’existence de barrières naturelles ainsi que la présence de compétiteurs ou de prédateurs peut limiter non seulement la connaissance des zones favorables mais également l’accès à ces zones. Enfin, le but de la reproduction étant de rencontrer son partenaire sexuel et non pas de coloniser un milieu, il est probable que toutes les zones où les conditions environnementales sont potentiellement favorables ne soient pas occupées ou faiblement occupées. Pour la plie, cela s’est traduit par une différence importante entre la distribution prédite par le modèle d’habitat et la distribution observée; cette dernière étant moins étendue que celle des conditions favorables. Cela est également suggéré dans le cas de l’anchois où la grande échelle spatiale explique plus de variabilité que la température de fond. Ainsi, il apparaît que l’anchois n’occupe pas toutes les zones où la température de fond est favorable à sa reproduction car si cela était le cas, la température de fond expliquerait plus de variabilité que la grande échelle. Enfin, des zones favorables peuvent ne pas être occupées à cause d’un nombre insuffisant d’individus dans la population. Cet aspect est traité dans le point suivant.

6.2.4. … suivant l’état de la population

Pour des populations en équilibre ou pseudo-équilibre avec leur environnement, il est probable qu’à court terme cela soit majoritairement l’état de la population qui fasse varier la distribution des adultes reproducteurs. Il a été montré pour les trois espèces étudiées que la taille de la population ainsi que la ségrégation spatiale des différentes classes d’âge contribuaient à la structuration spatiale de l’habitat de reproduction. Ainsi, la structuration spatiale de l’habitat de reproduction montré à travers l’existence de zones préférentielles, occasionnelles et non favorables est largement influencée par la taille de la population a travers des phénomènes de densité dépendance qui suivent le modèle constant ou de basin.

La structure en âge de la population joue également un rôle dans la structuration spatiale de l’habitat de reproduction car généralement les individus les plus jeunes ne sont pas distribués dans les mêmes zones que les individus les plus âgés. On peut donc espérer que des variations inter-annuelles importantes des proportions des différentes classes d’âges conduiront à une occupation plus ou moins importante de certaines zones préférentiellement occupées par ces classes d’âges. De plus, dans le cadre de l’hypothèse de l’entraînement (Petitgas *et al.*, 2006), la connaissance des zones de reproduction est assurée par les vieux individus qui la transmette alors aux jeunes individus en les entraînant avec eux jusqu’aux frayères. Lorsque certains jeunes individus ne sont pas entraînés, ils peuvent alors être perdus pour la population, ou peuvent trouver de nouvelles zones reproduction, ou redécouvrir des zones qui étaient anciennement occupées.

6.2.5. et prévisible dans le temps

La persistance temporelle de l’habitat de reproduction provient du fait que l’hétérogénéité spatiale de l’environnement qui crée sa structuration dans l’espace est prévisible dans le temps (Orians et Wittenberger, 1991). Avec un environnement dont l’hétérogénéité est prévisible dans le temps, on peut espérer que les critères sur lesquels ont reposé le choix du lieu de reproduction des populations seront toujours valables d’une année sur l’autre. Pour des espèces comme la plie et le merlan qui ne se reproduisent qu’une fois dans l’année, ceci est un avantage car sinon cela impliquerait de réévaluer chaque année la qualité des zones de reproduction et d’attendre l’année suivante pour constater les conséquences de la sélection de telle ou telle zone. En revanche, pour une espèce comme l’anchois qui plusieurs fraies successives pendant sa période de reproduction, il peut y avoir une réévaluation régulière du milieu qui va lui permettre de se distribuer constamment dans les zones lui assurant une quantité de nourriture suffisante. Ceci permettrait d'expliquer pourquoi la ponte a lieu en priorité dans le sud du Golfe de Gascogne au début de la saison de reproduction et couvre ensuite une zone beaucoup plus large du plateau continental (Motos *et al.*, 1996; Planque *et al.*, 2007).

La plupart des populations de poissons tempérés étant eurythermes et euryhalines, elles sont capables de supporter des variations environnementales relativement importantes. Ceci est d'autant plus valable pour les trois populations étudiées ici qui ne sont pas en limite de distribution dans leur milieu, car les variations interannuelles subies par les individus de la population lors de la reproduction ont généralement lieu à l'intérieur du gradient environnemental toléré à l'échelle de l'espèce. Il est donc probable que les variations interannuelles des conditions hydrologiques n'aient que peu d'influence à court terme sur la distribution des adultes reproducteurs. Néanmoins, dans le cas de la plie et du merlan, les conditions environnementales de température et de salinité agissent en amont de la reproduction au moment de la maturation des gonades et du déclenchement de la ponte et en aval de la reproduction sur la survie des œufs et des larves. Pour l'anchois, la quantité de nourriture disponible est également un facteur important car elle permet d'assurer les maturations répétées de ses gonades nécessaires aux pontes successives au cours de sa période de reproduction (ICES, 2004b). Des variations interannuelles de ces conditions peuvent alors entraîner des décalages temporels dans la période de reproduction ainsi que des taux de mortalité plus au moins importants des œufs et des larves du fait d'une asynchronie temporelle (mismatch) avec le pic de production des proies planctoniques.

6.2.6. Le retour aux zones de reproduction

La prévisibilité de l'hétérogénéité de l'environnement dans le temps favorise l'attachement des populations à leur site de reproduction, ce qui tend à faire persister dans le temps la distribution spatiale de l'habitat de reproduction des espèces. Cette capacité des adultes reproducteurs à revenir sur les mêmes zones d'une année sur l'autre est également un moyen de limiter l'effet des variations temporelles des conditions environnementales de manière à ce que les individus ne modifient pas systématiquement leur distribution au moindre changement des conditions rencontrées. Ainsi, Corten (2002) a montré que l'attachement du hareng pour ses zones de reproduction, qui est favorisé par le phénomène de conservation et d'entraînement, a tendance à limiter l'effet d'évènements exceptionnels. Ceci crée alors un décalage de temps entre le moment où les conditions environnementales changent et le moment où les répercussions sur la distribution sont visibles.

La stabilité de la distribution spatiale de l’habitat de reproduction dans le temps suggère un mécanisme de « homing », permettant aux individus de revenir au niveau des zones favorables pour la reproduction d’une année sur l’autre. Le terme de homing (Papi, 1992) désigne le retour des individus au niveau de zones où ils ont déjà exercé une activité. Les différentes sortes de homing peuvent être définies et classées en fonction du type d’information qui est utilisé pour retourner sur la zone souhaitée. Brochier (2009) teste deux types de homing: le homing géographique et le homing environnemental. Dans le homing géographique, les individus retournent se reproduire aux mêmes endroits où ils sont nés d’une année sur l’autre; on parle également de « natal homing » (Papi, 1992). Le homing environnemental (Cury, 1994; Baras, 1996) est une généralisation du natal homing dans le sens où les individus retournent dans les conditions environnementales où ils sont nés et ont l’habitude de se reproduire. Ces deux types de homing impliquent une phase « d’imprinting » qui a lieu pendant les premiers stades du cycle de vie et pendant laquelle les individus apprennent à reconnaître les conditions du milieu (Dodson, 1988). Un des intérêts de ce genre d’étude est de chercher à déterminer vers quel type de homing s’orientent les espèces pour revenir vers leurs zones de reproduction.

L’importance de la position géographique pour les espèces bentho-démersales étudiées comme la plie et le merlan suggère un homing de type géographique. Il est probable qu’en mer du Nord, le retour aux zones de reproduction soit déterminé par l’orientation du courant résiduel de marée qui est prévisible dans le temps et qui permet aux individus des populations benthiques de retrouver les mêmes conditions hydrodynamiques permettant la dérive des œufs et les larves vers les nourriceries côtières. D’ailleurs, des études de marquage ont montré que la plie parcourait des distances importantes pour revenir se reproduire au niveaux des frayères situées dans le sud de la mer du Nord alors que les zones de nutrition pour les adultes sont situées plus au nord (Hunter *et al.*, 2004). Si les mécanismes à l’origine du départ des adultes pour les frayères restent pour l’instant obscures, Hunter *et al.* (2003) ont montré que les adultes de plie se servaient du cycle de marée pour s’orienter et faciliter leur migration sur de si longues distances. Pour des populations de petits pélagiques comme l’anchois, Brochier (2009) fait remarquer que le homing géographique n’est pas réaliste dans le sens où il ne permet pas de reproduire la variabilité inter-annuelle de la ponte de l’anchois observée dans les systèmes d’upwelling. Il a ainsi montré que c’est le homing environnemental qui semble à la fois la stratégie la plus plausible en termes de processus, et réaliste en terme de distribution de ponte observée (Brochier *et al.*, sous-presse). Cependant, il est nécessaire de rappeler que l’espèce d’anchois étudiée ici (*E. encrasicolus*) est différente de celle étudiée par Brochier (2009) pour le Courant de Humbolt (*E. ringens*) et que les environnements étudiés sont très différents. Les régions du courant des Canaries, du Courant de Humbolt et du Benguela sont dominées par des systèmes d’upwelling qui sont la source principale de nourriture pour les individus. Dans le Golfe de Gascogne, l’upwelling est d’une intensité beaucoup plus faible et l’apport de matières nutritives se fait principalement au niveau des panaches d’estuaires et du talus continental. Enfin, les facteurs qui déterminent la présence-absence des individus d’anchois dans le Golfe de Gascogne ne sont pas les mêmes que ceux qui déterminent leur niveau d’abondance. Ainsi, nous suggérons que pour des espèces pélagiques, le phénomène de homing peut permettre d’expliquer la variabilité dans l’occupation de certaines zones mais en aucun cas la variabilité dans le degré d’abondance avec lequel les individus occupent ces zones. En effet, nous avons vu que ce niveau d’abondance était déterminé par d’autres facteurs qui sont plus liés à l’état de la population.

Doligez *et al.* (2003) ont montré que le succès de reproduction selon différentes stratégies variait en fonction de la prévisibilité de l’environnement dans le temps. Ainsi, le succès de reproduction d’une stratégie basée sur le homing géographique sera d’autant plus fort que l’environnement sera très corrélé dans le temps. A l’extrême inverse, dans un milieu où les conditions seraient aléatoires dans le temps, une stratégie basée sur une reproduction au hasard donnerait de meilleurs résultats. Il est possible qu’une stratégie basée sur le homing environnemental soit plus adaptée dans des milieux où les conditions environnementales rencontrées y varient fortement d’une année sur l’autre (comme les systèmes d’upwellings). En revanche, dans le Golfe de Gascogne, il est possible que le homing environnemental ait été utilisé au début de la sélection de l’habitat de reproduction et que la stabilité dans le temps des conditions environnementales et trophiques ait permis ensuite aux individus de s’orienter vers un homing géographique, leur permettant ainsi de trouver les conditions environnementales favorables à la reproduction. Ainsi, l’importance conjointe de la grande échelle spatiale et de la température de fond dans la distribution des zones de présence-absence de l’anchois dans le Golfe de Gascogne peut traduire le fait que c’est bien la température de fond qui contraint la distribution des adultes mais que la stratégie adoptée vise à occuper en priorité la partie sud du Golfe de Gascogne là où l’anchois est certain de retrouver les conditions favorables de température de fond. Ceci permettrait d’expliquer pourquoi l’ensemble des zones où la température de fond est favorable n’est pas occupé et aussi pourquoi la température de fond a un pouvoir prédictif moins important que l’espace.

6.3. Critiques et perspectives

La théorie de l'information sur laquelle repose l'approche multi-modèles utilisée au cours de ce travail, permet d'extraire des propriétés sur les mécanismes qui contrôlent la distribution spatiale des individus à partir de la modélisation de données d'observations (Hilborn et Mangel, 1997; Burnham et Anderson, 2002). Ceci suppose d'une part que la distribution observée des individus soit capable de refléter correctement leur distribution réelle dans le milieu et d'autre part que la structure spatiale de certains facteurs qui est partagée par la distribution des individus, traduise une relation de cause à effet entre les deux. Or, aucune de ces deux conditions n'est jamais totalement remplie. En effet, les distributions observées obtenues grâce au procédé d'échantillonnage n'offrent qu'une vision filtrée et donc partielle de la distribution spatiale des individus car elles sont une image statique prise à un moment donné d'une distribution réelle qui elle est dynamique. Ce phénomène limite d'une part l'information qui peut être retirée de ces données observées et d'autre part les prédictions qui sont faites à partir des modèles basés sur ces données.

6.3.1. Les modèles environnementaux

Même si l'hétérogénéité spatiale de l'environnement du milieu dans lequel vivent les individus possède un certain degré de prévisibilité dans le temps, il n'en demeure pas moins que des variations temporelles des conditions environnementales, notamment dans le cadre des changements à long terme, peuvent affecter la configuration spatiale de la qualité de l'habitat de reproduction. Les modèles développés au cours de ce travail pourraient permettre d'évaluer l'impact des changements environnementaux sur la distribution des adultes reproducteurs. Dans le cadre du projet RECLAIM[3], un modèle d'habitat pour la plie a été construit en utilisant les variables environnementales du modèle ECOSMO (ECOSystem MOdel, Schrum *et al.*, 2006). Les forçages climatiques suivant ont été appliqués au modèle ECOSMO pour 2002, 2003 et 2004: augmentation de 3°C de la température atmosphérique, augmentation de 30% du vent et augmentation de 20% de l'irradiance solaire. Les variables environnementales ont alors été recalculées selon ces contraintes et les cartes de distribution de la plie ont été produites en appliquant le modèle d'habitat sur les nouvelles valeurs des variables environnementales. Les cartes de distribution pour les trois années de forçage

---

[3] RECLAIM : REsolving CLimAtic IMpacts on fish stocks. STREP-FP6, contract n°44133. http://www.climateandfish.eu/

climatique (fig. 6.1) ont montré d'une part que le sud de la mer du Nord restait la partie majoritairement favorable à la reproduction des adultes de plie mais que d'autre part les zones favorables sont beaucoup plus réduites que celles observées actuellement. Les forçages climatiques ont été appliqués à trois années pour lesquelles les distributions observées étaient disponibles afin de les comparer aux données prédites selon le scénario. Néanmoins, l'intérêt de ce genre de modèle réside surtout dans le fait de pouvoir les appliquer à plusieurs scénarios possibles pour en comparer les prédictions et également à la prévision d'années pour lesquelles on ne dispose pas de distributions observées mais dont on connaît la distribution des variables pertinentes de l'environnement.

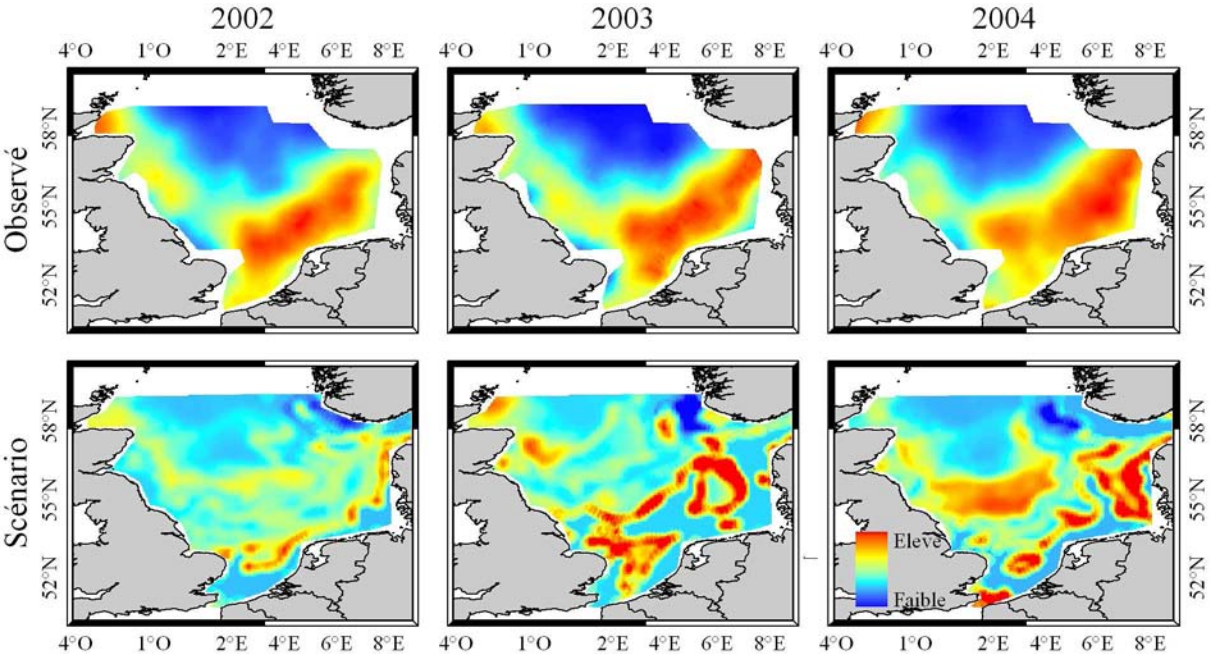


Figure 6.1. Cartes de distribution observée et prédite des adultes de plie en période de reproduction pour trois années de référence. La distribution sous scénario correspond à la distribution prédite par le modèle d'habitat d'après les conditions environnementales du scénario issues du modèle ECOSMO.

Le fait que seuls les facteurs environnementaux soient pris en compte, comme c'est majoritairement le cas, dans les modèles d'habitats classiques (Guisan et Zimmermann, 2000; Austin, 2002; Austin, 2007) se base sur l'hypothèse généralement erronée que c'est uniquement l'environnement qui façonne la distribution des adultes reproducteurs. Certaines courbes de réponse environnementales comme celles construites pour la plie peuvent suggérer des préférences qu'ont les individus pour certaines valeurs de conditions environnementales. Or, l'importance du phénomène d'agrégation des individus qui caractérise la reproduction peut aboutir à ce que seule une partie du gradient environnemental couvert par la zone soit occupé et donc conduire à la modélisation de courbes de réponse sans qu'il n'y ait de relation réelle avec l'environnement. La validité de l'application des modèles d'habitat reposant en premier lieu sur l'hypothèse que la distribution observée reflète bien la dépendance des individus vis à vis des conditions environnementales actuelles, l'utilisation de ces modèles en mode prédictif est limitée dès lors que les modèles déterminés comme les meilleurs ne contiennent pas d'hypothèses liées à l'environnement non-persistant comme ce fut le cas pour les populations benthiques de la plie et du merlan.

6.3.2. Les modèles structurés en âge

Si l'on veut évaluer correctement l'effet des changements de taille et de structure de population sur la distribution spatiale, il faut envisager d'avoir recours à des scénarios biologiques dans lesquels il est possible de faire varier la taille totale de la population, le nombre de géniteurs et les proportions des différentes classes d'âge. On peut envisager d'appliquer directement les modèles sélectionnés pour les trois espèces à de nouvelles proportions de classe d'âge à chaque station afin de prédire la distribution résultante. Il serait également intéressant de pouvoir injecter directement dans les modèles des valeurs annuelles de taille et de structure d'âge.

Or, même si les modèles développés dans le cadre de ce travail intègrent des variables de taille et de proportions annuelles des classes d'âges, ils ont surtout été construits pour être utilisés dans un cadre exploratoire afin de mettre en relation de manière simple des variables explicatives avec une distribution observée en utilisant des multi-régressions. De ce fait, ce sont des modèles statiques où la manière dont une variation temporelle de l'état de la population va influencer la distribution dans l'espace des adultes n'est pas exprimée de manière dynamique. En effet, il n'y a pas de relation directe entre la quantité d'individus observée à une année donnée et l'abondance qui est observée à chaque station (McCall, 1990). De même, la manière dont la proportion d'individus d'un certain âge se distribue dans l'espace et occupe certaines zones préférentielles n'est pas représentée de façon explicite dans les modèles.

Une manière de remédier à ce problème serait de construire des modèles pour chaque classe de taille ou d'âge. La réponse modélisée ne serait alors plus la présence ou l'abondance totale observée à un endroit mais celle de chaque classe d'âge expliquée par les différentes hypothèses utilisées au cours de cette étude. Les prédictions de ces modèles pourraient alors être combinées à la fin pour retrouver l'abondance totale des individus au niveau des différentes zones. Ceci pourrait notamment permettre de constater l'influence des facteurs environnementaux sur la répartition des individus en fonction de leur âge ou de leur taille (Koubbi *et al.*, 2006).

Une structure démographique tronquée et une taille de population qui présentent peu de variabilité au cours du temps limitent également l'évaluation par les modèles de l'impact réel que peuvent avoir des variations importantes dans ces deux paramètres. En effet, le fait que ces deux hypothèses n'aient pas été retenues dans certains modèles sélectionnés d'après la prédiction au cours de cette étude n'implique pas nécessairement qu'ils n'ont pas d'influence sur la distribution spatiale de la reproduction. Cela peut simplement traduire le fait que les variations rencontrées dans les données observées qui ont été utilisées pour construire les modèles n'étaient pas suffisantes pour détecter un effet significatif de ces facteurs.

Les mesures de gestion des populations de poissons basées sur les quotas et les mesures techniques comme la taille des mailles des chaluts visent principalement à limiter la capture des jeunes individus pour leur laisser le temps de se reproduire tandis que les individus les plus vieux subissent une exploitation par la pêche toujours aussi forte. Les populations exploitées ont donc souvent une structure démographique peu diversifiée, largement dominée par les plus jeunes individus et peu variable dans le temps. Les caractéristiques biologiques des jeunes individus peuvent avoir des conséquences dommageables sur le succès de la reproduction (Solemdal, 1997).

Pour la plie et le merlan, il a été montré au cours de cette étude que seule une faible proportion d'individus pour les classes de taille rencontrées actuellement étaient capables de se reproduire. L'exploitation par la pêche a tendance à favoriser les individus qui sont capables de se reproduire très rapidement (Rijnsdorp, 1989; Rijnsdorp *et al.*, 2005; Jorgensen *et al.*, 2007). Cependant, il n'en demeure pas moins que les jeunes individus se reproduisent généralement un peu plus tard dans la saison et moins longtemps que les individus plus âgés (Rijnsdorp, 1989), leur plus petite taille fait qu'ils ont une capacité énergétique plus faible à allouer à la reproduction, ils pondent une quantité d'œufs moins importante et de moins bonne qualité. Les jeunes individus sont également plus sensibles à la mortalité par la prédation et aux variations climatiques que les individus les plus âgés (Bundy et Fanning, 2005), ce qui entraîne des variations de leur abondance totale d'une année sur l'autre qui sont alors relativement importantes. Cet aspect est d'autant plus essentiel pour des populations de petits pélagiques comme l'anchois où les individus d'âge 1 sont déjà capables de se reproduire et assurent la majeure partie de la reproduction (ICES, 2008b). Une structure démographique diversifiée permet alors de couvrir une plus grande diversité des conditions environnementales pouvant être rencontrées pendant la période de reproduction et d'absorber plus facilement les variations environnementales notamment pour des populations benthiques comme la plie ou le merlan qui ont une longue durée de vie (Planque *et al.*, sous presse).

Enfin, la manière dont les individus se répartissent au sein de l'habitat de reproduction en fonction de leur âge ainsi que la taille de la population peuvent également avoir des conséquences sur le succès de reproduction. En effet, il a été montré au cours de l'étude que les zones occupées à des niveaux très bas de taille de population sont moins étendues et les zones où se distribuent les jeunes individus sont généralement plus côtières; ceci ne permet pas une occupation spatiale de la reproduction très étendue. Potentiellement, ce phénomène peut alors limiter les chances pour les œufs et larves de rencontrer des conditions environnementales diversifiées qui pourraient augmenter leur chance de survie. Enfin, sous l'hypothèse de l'entraînement, une structure démographique peu diversifiée peut entraîner une perte de la connaissance des zones de reproduction et un changement important dans la distribution spatiale des zones occupées (Corten, 2002).

6.3.3. Approches complémentaires

L'utilisation de données d'observation doit nécessairement être combinée à d'autres approches. La prise en compte du comportement des individus à l'échelle de la population semble essentielle. Ainsi, les individus, au lieu d'évaluer séparément la qualité d'un habitat de reproduction (information privée), peuvent se servir de l'évaluation des conditions que les autres individus ont fait et des connaissances qu'ils ont acquises sur leur milieu. On parle alors d'information publique (Valone et Templeton, 2002), qui est de l'information obtenue à partir de l'échantillonnage vicariant (par les autres). Plus que l'échantillonnage en lui seul, ce sont les résultats obtenus par les individus à partir de cet échantillonnage, qui est utilisé comme source d'information par d'autres individus. Il semble qu'une stratégie basée sur l'information publique plutôt que sur l'information privée soit plus pertinente dans le cas de la sélection d'un habitat de reproduction car elle évite à chaque individu d'avoir à prospecter indépendamment les zones à occuper (Doligez *et al.*, 2003). Ceci représente une économie d'énergie non négligeable et aboutit généralement à une agrégation très importante des individus qui, puisqu'ils utilisent la même source d'information, ont tendance à occuper les mêmes zones. Par exemple, il a été montré que le succès de reproduction des congénères obtenus durant l'année en cours pouvait être utilisé par les individus en échec de reproduction pour décider de rester ou de quitter les zones qu'ils occupaient actuellement (Doligez *et al.*, 2003). De plus, il a été montré qu'un bon succès de reproduction obtenu dans le passé augmentait la fidélité des individus à leur site de reproduction (Switzer, 1997). De ce fait, Boulinier et Danchin (1997) suggèrent d'étudier le succès de reproduction des diverses zones de reproduction plutôt que la distribution des abondances. Ceci peut se faire à travers l'étude de l'habitat efficace de reproduction. Planque *et al.* (2007) suggèrent que l'habitat efficace pour des populations de poissons ne peut être déterminé au moment de la reproduction et doit être déduit d'informations sur les adultes et les jeunes individus une fois que ces derniers ont été recrutés dans la population. Les modèles développés dans l'approche multi-modèles doivent donc nécessairement être combinés à des approches qui permettent de prendre en compte le comportement des individus notamment en ce qui concerne l'information utilisée pour décider d'occuper un lieu pour se reproduire.

Une approche par l'expérimentation pourrait également permettre de déterminer la partie de la niche fondamentale relative à la reproduction, c'est à dire l'ensemble des facteurs environnementaux limitant la distribution des adultes reproducteurs. On peut alors imaginer étudier une réponse physiologique en rapport avec la reproduction, telle que le taux de fécondité, en fonction de plusieurs variables environnementales et ensuite appliquer ces relations aux distributions observées de ces variables afin de cartographier l'habitat potentiel de reproduction. Il serait également intéressant d'étudier quelle part de la ressource énergétique emmagasinée par les individus est dédiée à la reproduction, par rapport à la survie et à la croissance, suivant les conditions environnementales rencontrées au cours de la maturation des gonades. Ceci est d'autant plus essentiel dans le cas des jeunes individus qui sont déjà capables de se reproduire mais pour qui de l'énergie doit également être allouée à la croissance. Ce type d'information est également nécessaire pour permettre le développement de modèles individus centrés (modèles IBM).

Planque *et al.* (2007) suggèrent que les IBM peuvent être utilisés pour suivre les cohortes ou les individus et les conditions environnementales dans lesquelles ils vivent au début de leur cycle de vie (Allain *et al.*, 2003). On pourrait alors cartographier l'habitat efficace pour plusieurs années et ainsi déterminer si les individus tendent à revenir d'une année sur l'autre sur les mêmes zones qui leur ont offert un succès de reproduction important. Des modèles théoriques et conceptuels peuvent également être développés pour représenter les différentes stratégies utilisées par les individus pour sélectionner un habitat de reproduction. A partir de ces modèles, il est possible de prédire les distributions qui seraient obtenues en fonction des divers choix qui s’offrent aux individus. Dans ce contexte, les données d’observation, plutôt que d’être utilisées pour calibrer les modèles, sont alors utilisées à posteriori pour tester des prédictions qui sont faites par ces modèles conceptuels. Si des différences sont constatées, on peut alors essayer d’expliquer quelles en sont les raisons. C’est notamment ce qui a été réalisé dans la thèse d’Irisson (2008) et de Brochier (2009) sur l’ichthyoplancton où un IBM a été utilisé pour intégrer le comportement des individus. Brochier *et al.* (2008) et Brochier *et al.* (sous-presse) ont utilisé un IBM évolutif pour permettre de comparer aux données de distribution observées, les distributions obtenues à partir de trois stratégies de ponte différente: une stratégie basée sur le homing environnemental, une stratégie basée sur le homing géographique et une stratégie de ponte aléatoire. Ces approches à la fois différentes et complémentaires de celle utilisée au cours de l’étude permettront de faire avancer la compréhension des mécanismes qui déterminent la distribution spatiale des populations de poissons.

Conclusion

Le but de la thèse était de caractériser et d'identifier les facteurs qui influencent la distribution spatiale de la phase de reproduction chez les populations de poissons. Du fait des multiples hypothèses qui peuvent contraindre cette distribution, l'utilisation et le développement de modèles basés sur des hypothèses prises individuellement se révèlent être insuffisants et peuvent conduire à une interprétation erronée. En effet, pour la plie il a été montré que le modèle d'habitat était incapable de refléter correctement la distribution observée lorsque la taille de la population variait. L'objectif de cette étude n'était pas de déterminer si chacune des hypothèses prises individuellement était valide mais d'évaluer la contribution relative des différents processus et d'évaluer la fiabilité et l'utilité des hypothèses lorsque celles ci étaient utilisées pour représenter les distributions observées. Une des approches possibles consiste à transcrire ces différentes hypothèses sous la forme de modèles numériques qui peuvent être ajustés à des données de terrain et qui peuvent également être évalués vis à vis de données qui sont indépendantes de celles utilisées pour leur ajustement. Une approche multi-modèles basée sur la théorie de l'information a donc été utilisée pour modéliser plusieurs hypothèses de contrôle en utilisant des variables explicatives. Ces modèles numériques ont ensuite été comparés et classés afin de sélectionner ceux qui sont capables de refléter au mieux les distributions observées.

Des différences ont été mises en évidence concernant la complexité des modèles en fonction de leur capacité à s'ajuster ou à prédire des données d'observation. L'autocorrélation spatiale présente dans les données agit comme une nuisance statistique sur la sélection basée sur l'ajustement et conduit à sélectionner des modèles trop complexes et à donner trop d'importances aux hypothèses de contrôle qui contiennent le plus de variables explicatives. Le sur-ajustement des modèles sélectionnés aux données utilisées pour les calibrer crée du bruit dans les prédictions lorsqu'ils sont appliqués à un autre jeu de données, ce qui aboutit alors à un taux d'erreur important entre les données prédites et les données observées. Les modèles déterminés comme les mieux ajustés aux données ne sont pas ceux capables de prédire au mieux les distributions observées lorsqu'ils sont appliqués sur un autre jeu de données. En revanche, l'effet de l'autocorrélation sur la complexité des modèles est moindre lorsque ceux ci sont sélectionnés d'après la prédiction. Pour cette raison, il est probable que la sélection basée sur la prédiction soit plus adaptée que celle basée sur l’ajustement et qu’elle devrait être utilisée pour classifier et sélectionner correctement les modèles.

Un des enjeux majeur de l’approche multi-modèles est de pouvoir tester la valeur des différents modèles d’après un jeu de données indépendant. Au cours de cette étude, les données d’observations ont été divisées en deux et un de jeu données différent de celui utilisé pour calibrer les modèles a été utilisé dans le cadre de la sélection basée sur la prédiction. Ces deux jeux de données n’en demeurent pas moins non-indépendants car les distributions observées sont autocorrélées à la fois dans le temps et dans l’espace. Il apparaît alors que l’obtention d’un jeu de donnée complètement indépendant pour évaluer les modèles est difficile car elle impliquerait de se débarrasser de l’autocorrélation qui est une propriété intrinsèque des données biologiques. De ce fait, les modèles sélectionnées par la prédiction grâce au log de vraisemblance contiennent des hypothèses qui ont été déterminées comme non significatives. Ceci prouve que cette l’approche n’est pas totalement capable d’éliminer certaines hypothèses expliquant uniquement du bruit dans les observations. Ceci n’aurait que peu d’influence dans une approche purement prédictive où le but serait de sélectionner les modèles qui prédisent le mieux la distribution observée sans se soucier de leur niveau de complexité. Néanmoins, lorsque le but recherché est de déterminer quelles sont les hypothèses qui sont réellement importantes, cette méthode de sélection basée sur le log de vraisemblance ne permet pas d’avoir accès directement aux modèles qui prédisent le mieux avec le moins d’hypothèses possibles. Nous préconisons, dans le cadre de la sélection basée sur la prédiction, l’utilisation d’un critère de type AIC qui pourrait permettre de résoudre ce problème.

La sélection des modèles les plus pertinents repose sur l’utilisation d’une valeur seuil de distance entre deux mesures de l’ajustement des modèles comparés. La valeur 10 est classiquement utilisée comme valeur limite pour distinguer les meilleurs modèles sur la base de l’AIC. Cependant, il a été montré au cours de cette étude que d’une part les modèles ayant une valeur inférieure à ce seuil avaient une capacité de prédiction équivalente et que d’autre part des modèles avec une valeur largement supérieure à 10 avaient également une capacité prédictive équivalente aux premiers. Ainsi, une différence de seuil de sélection, même importante, entre les modèles n’implique pas obligatoirement une différence significative dans la capacité à prédire des modèles. Il est donc difficile de conclure quant au choix d’une valeur de seuil à utiliser mais cette valeur ne devrait pas être trop restrictive et surtout le seuil de sélection ne devrait pas être utilisé comme unique critère pour sélectionner les modèles. Encore une fois, l’utilisation d’un critère comme l’AIC permettrait lors de la sélection basée sur la prédiction de classer en priorité et dans un seuil restreint, les meilleurs modèles contenant le moins d'hypothèses possibles.

L'habitat de reproduction possède une distribution qui est hautement structurée dans l'espace. Cette structuration est influencée d'une part par la manière dont les individus occupent le milieu et d'autre part par le degré d'abondances avec lequel les zones sont occupées. Les adultes reproducteurs se concentrent en priorité dans certaines zones géographiques bien localisées tandis qu'ils sont absents des zones adjacentes. Il a été montré au cours de cette étude que les facteurs influençant la présence-absence des individus en train de se reproduire étaient différents et moins nombreux que ceux influençant la distribution de leurs abondances. L'occupation des zones est déterminé principalement par des facteurs environnementaux qui créent l'hétérogénéité du milieu. Cette hétérogénéité aboutit à l'existence de zones favorables et non favorables dont le degré d'occupation des individus va être modulé par des caractéristiques liées à la population elle même.

L'importance des facteurs environnementaux dans la structuration spatiale de l'habitat de reproduction dépend des traits de vie et des modalités de reproduction des espèces. Pour des espèces itéropares annuelles comme la plie, les variations temporelles des conditions environnementales vont agir en amont de la reproduction, tandis qu'au moment de se reproduire, les adultes vont se répartir selon les caractéristiques de l'environnement qui varient uniquement dans l'espace. En revanche, pour des espèces comme l'anchois qui se reproduit à plusieurs reprises au cours de la période de reproduction, les individus vont se distribuer selon les facteurs environnementaux qui varient à la fois dans le temps et dans l'espace comme la température et la quantité de nourriture disponible.

La taille de la population va influencer la manière dont les zones favorables sont occupées en jouant sur la contraction et l'extension de la distribution de l'habitat de reproduction. Cet effet est détectable uniquement si les populations ont subi pendant la période d'observation de fortes variations de leur abondance qui ont amené à une réorganisation ou à une modification de la distribution. Ainsi, il est possible de conclure sur un effet de densité dépendance dans le cas de la plie, alors que dans le cas du merlan et de l'anchois, l'absence de relation nette entre la distribution et la taille de la population n'implique pas forcément l'absence de ce phénomène de densité dépendance. En effet, les biomasses de ces deux espèces peuvent avoir atteint un niveau où l'effet de densité dépendance est devenu moins important que l'effet du à d'autres facteurs. L'appréciation de la qualité des zones selon l'âge des individus va également influencer la distribution spatiale des abondances observées. Certaines zones seront occupées de manière plus ou moins abondante selon l’abondance des différentes classes d’âges qui constituent la population. La détection et la quantification de l’importance de cet effet de l’âge sur la sélection de l’habitat peuvent être sous-estimées par l’état actuel des populations dont la structure démographique tronquée par l’activité de pêche n’offrent qu’une vision partielle de la distribution réelle de la reproduction.

La structure spatiale de l’habitat de reproduction tend à se répéter dans le temps. Ce phénomène est favorisé par un attachement très fort des individus pour leurs sites de reproduction. Il a été montré que cet attachement était d’autant plus fort pour des populations benthiques comme la plie et le merlan que pour des populations pélagiques comme l’anchois. Cependant, même dans le cas d’un petit pélagique comme l’anchois, on constate un degré de persistance de l’organisation spatiale de l’habitat de reproduction qui est plus important dans le Golfe de Gascogne qu’il ne l’est dans d’autres zones où les conditions environnementales varient fortement d’une année sur l’autre. Cet attachement des populations pour leur site de reproduction est à la fois une adaptation vis à vis des conditions environnementales qui sont prévisibles dans le temps et une manière de faire face aux changements environnementaux éphémères en s’imposant une certaine inertie dans leur distribution spatiale.

La nature des facteurs influençant la distribution des individus au moment de se reproduire confirme que les stratégies de reproduction adoptées différent selon l’espèce considérée. La reproduction hivernale de la plie est un moyen d’éviter la compétition avec les autres espèces. Le fait que l’espace, reflété à travers la position géographique, la grande et moyenne échelle spatiale ainsi que l’environnement persistant, ait été déterminés comme importants montre que cette espèce reproduit chaque année le même schéma de distribution et mise avant tout sur une dérive des œufs et larves jusqu’au nourriceries côtières sans pour autant assurer à ces derniers des conditions de production maximale en nourriture. Contrairement à la plie, le merlan, qui émet ses œufs à la fin de l’hiver, compte ainsi sur l’utilisation de la nourriture apportée par le pic printanier pour assurer la croissance des larves. Il montre également un attachement géographique très fort et un schéma de distribution qui se reproduit d’une année sur l’autre. Un tel schéma peut être dangereux puisque sa période de reproduction a lieu au moment où les risques de décalage temporel dans le pic de production sont les plus élevés. Ces risques sont compensés par une période de reproduction qui est l’une des plus étalée parmi les espèces de Mer du Nord. L’anchois, lui, se reproduit clairement pendant le pic de production de nourriture. La distribution adoptée possède un caractère conservateur qui vise à assurer un apport de nourriture suffisant pour la nutrition des adultes plutôt que pour celle des larves. Elle possède donc une plus grande flexibilité spatiale qui la rend plus difficile à prédire et qui résulte en plus de variabilité dans le temps que pour les deux espèces benthiques. Pour les trois espèces, la distribution de l'habitat de reproduction ne vise donc pas forcément à déposer les œufs et larves dans des conditions environnementales leur assurant une survie maximale. Elle vise en premier lieu à assurer aux adultes une capacité à se reproduire et à retrouver les zones favorables pour l'émission des œufs. Les facteurs qui vont affecter la survie des œufs et des larves seront donc différents de ceux qui contraignent la distribution des adultes reproducteurs.

L'importance des différents facteurs a été quantifiée à travers leur capacité à expliquer une part de la variabilité dans les données observées de distribution. Cette manière de faire ne permet pas de déterminer quels sont les processus mis en jeu pouvant aboutir à de telles distributions. Dans le cas du merlan et de la plie, il a été mis en évidence que la position géographique était importante, cependant il est difficile de savoir quelle source d'information est utilisée par les individus pour revenir aux mêmes endroits d'une année sur l'autre. De manière générale, la distribution spatiale des individus ne repose pas forcément sur les critères du milieu auxquels les abondances d'individus sont corrélées. Par exemple, le fait que la production primaire ait été déterminée comme importante reflète que la nourriture est abondante dans les zones favorables pour l'anchois sans pour autant que cela soit ce critère qui soit utilisé pour reconnaître ces zones. Lorsque deux facteurs sont confondus dans l'espace (comme c'est souvent le cas pour des co-réalisations distinctes issues d'un même processus sous-jacent), il devient difficile de déterminer si c'est l'un ou l'autre qui contraint la distribution des individus. De plus, les mêmes structures de distribution peuvent être obtenues par des processus de choix différents. Par exemple, nous avons montré que la localisation spatiale et l'environnement étaient tous les deux importants pour expliquer la présence-absence de l'anchois. Cependant, il reste difficile de savoir si ces deux paramètres agissent de manière complémentaire ou bien si un seul des deux est utilisé par les individus pour revenir sur leurs zones de reproduction. L'étude des distributions observées peut donc donner quelques indications quant aux sources d'informations utilisées ou sur les facteurs contraignant la distribution, mais elle n'est pas suffisante pour en déduire les mécanismes de prises d'information. Des approches différentes de celle qui a été utilisée au cours de cette étude sont donc nécessaires pour venir compléter les résultats obtenus et ainsi mieux appréhender le rôle des différents facteurs dans le contrôle de la distribution spatiale de la reproduction, de chaque phase en particulier et du cycle de vie en général.

Bibliographie

Agostini, V.N. and Bakun, A. 2002. Ocean triads in the Mediterranean Sea: physical mechanisms potentially structuring reproductive habitat suitability (with example application to European anchovy, *Engraulis encrasicolus*). Fisheries Oceanography, 11: 129-142.

Akaike, H. 1974. A new look a the statistical model identification. IEEE Transactions on Automatic Control, 19: 716-723.

Aldridge, J.N. and Davies, A.M. 1993. A high resolution three-dimensional hydrodynamic tidal model of the Eastern Irish Sea. Journal of Physical Oceanography, 23: 207-224.

Allain, G., Petitgas, P., Grellier, P. and Lazure, P. 2003. The selection process from larval to juvenile stages of anchovy (*Engraulis encrasicolus*) in the Bay of Biscay investigated by Lagrangian simulations and comparative otolith growth. Fisheries Oceanography, 12: 407-418.

Araujo, M.B. and Guisan, A. 2006. Five (or so) challenges for species distribution modelling. Journal of Biogeography, 33: 1677-1688.

Araujo, M.B. and New, M. 2007. Ensemble forecasting of species distributions. Trends in Ecology and Evolution, 22: 42-47.

Atkinson, C.J.L., Bergmann, M. and Kaiser, M.J. 2004. Habitat selection in whiting. Journal of Fish Biology, 64: 788-793.

Augustin, N.H., Mugglestone, M.A. and Buckland, S.T. 1996. An autologistic model for the spatial distribution of wildlife. Journal of Applied Ecology, 33: 339-347.

Austin, M.P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling, 157: 101-118.

Austin, M.P. 2007. Species distribution models and ecological theory: a critical assesment and some possible new approaches. Ecological Modelling, 200: 1-19.

Bakun, A. 2006. Fronts and eddies as key structures in the habitat of marine fish larvae: opportunity, adaptive response and competitive advantage. Scientia marina, 70S2: 105-122.

Baras 1996. Commentaire à l’hypothèse de l’éternel retour de Cury (1994) : proposition d’un mécanisme fonctionnel dynamique. Canadian Journal of Fisheries and Aquatic Sciences, 53: 681-684.

Bellier, E., Planque, B. and Petitgas, P. 2007a. Historical fluctuations in spawning location of anchovy (*Engraulis encrasicolus*) and sardine (*Sardina pilchardus*) in the Bay of Biscay during 1967-73 and 2000-2004. Fisheries Oceanography, 16: 1-15.

Bellier, E., Monestiez, P., Durbec, J. and Candau, J. 2007b. Identifying spatial relationships at multiple scales: principal coordinates of neighbour matrices (PCNM) and geostatistical approaches. Ecography, 30: 385-399.

Benaka, L. 1999. Fish habitat: essential fish habitat and rehabilitation. American Fisheries Society. 400 pp.

Beyer, H.L. 2004. Hawth's Analysis Tools for ArcGIS. Available at http://www.spatialecology.com/htools.

Blanchard, J.L., Mills, C., Jennings, S., Fox, C.J., Rackham, B.D., Eastwood, P.D. and O'Brien, C.M. 2005. Distribution-abundance relationships for North Atlantic cod (*Gadus morhua*): observation versus theory. Canadian Journal of Fisheries and Aquatic Sciences, 62: 2001-2009.

Blanchet, F.G., Legendre, P. and Borcard, D. 2008. Forward selection of explanatory variables. Ecology, 89: 2623-2632.

Boisclair, D. 2001. Fish habitat modeling: from conceptual framework to functional tools. Canadian Journal of Fisheries and Aquatic Sciences, 58: 1-9.

Borcard, D., Legendre, P. and Drapeau, P. 1992. Partialling out the spatial component of ecological variation. Ecology, 73: 1045-1055.

Borcard, D. and Legendre, P. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling, 153: 51-68.

Borcard, D., Legendre, P., Avois-Jacquet, C. and Tuomisto, H. 2004. Dissecting the spatial structure of ecological data at multiple scales. Ecology, 85: 1826-1832.

Borja, A., Uriarte, A., Egana, J., Motos, L. and Valencia, V. 1998. Relationships between anchovy (*Engraulis encrasicolus*) recruitment and environment in the Bay of Biscay (1967-1996). Fisheries Oceanography, 7: 375-380.

Borja, A., Fontan, A., Saenz, J. and Valencia, V. 2008. Climate, oceanography, and recruitment: the case of the Bay of Biscay anchovy (*Engraulis encrasicolus*). Fisheries Oceanography, 17: 477-493.

Boulinier, T. and Lemel, J. 1996. Spatial and temporal variations of factors affecting breeding habitat quality in colonial birds: some consequences for dispersal and habitat selection. Acta Oecologica, 17: 531-522.

Boulinier, T. and Danchin, E. 1997. The use of conspesific reproductive success for breeding patch selection in terrestrial ecology. Evolutionnary Ecology, 11: 505-517.

Boulinier, T., Mariette, M. and Danchin, E. 2005. La sélection d'un lieu de reproduction. In Ecologie comportementale: Une perspective évolutive, pp. 171-198. Ed. by E. Danchin, L.A. Giraldeau, and F. Cézilly. Dunod, Paris.

Brochier, T., Lett, C., Tam, J., Fréon, P., Colas, F. and Ayon, P. 2008. Anchovy early life in the northern Humboldt Current system : an individual-based model study. Progress in Oceanography, 79: 313-325.

Brochier, T. 2009. Stratégie de reproduction des petits poissons pélagiques dans les zones d'upwelling: une approche par modélisation individu-centrée appliquée aux systèmes de courant de Humbolt et des Canaries. Thèse d'université. Paris VI. 170 pp.

Brochier, T., Colas, F., Lett, C., Echevin, V., Cubillos, L.A., Tam, J., Chlaida, M. et al. sous-presse. Small pelagic fish reproductive strategies in upwelling systems : a natal homing evolutionary model to study environmental constraints. Progress in Oceanography.

Brown, C. and Laland, K.N. 2003. Social learning in fishes : a review. Fish and Fisheries, 4: 280-288.

Bundy, A. and Fanning, P.L. 2005. Can Atlantic cod (*Gadus morhua*) recover? Exploring trophic explanations for the non-recovery of the cod stock on the eastern Scotian Shelf, Canada. Canadian Journal of Fisheries and Aquatic Sciences, 62: 1474-1489.

Burnham, K.P. and Anderson, D.R. 2002. Model selection and multimodel inference. A practical information-theoritic approach. Springer-Verlag, USA. 488 pp.

Cade, B.S., Terrell, J.W. and Schroeder, R.L. 1999. Estimating effects of limiting factors with regression quantiles. Ecology, 80: 311-323.

Cade, B.S. and Noon, B.R. 2003. A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and Environment, 1: 412-420.

Carl, G. and Kuhn, I. 2007. Analyzing spatial autocorrelation in species distributions using Gaussian and logit models. Ecological Modelling, 207: 159-170.

Certain, G., Bellier, E., Planque, B. and Bretagnolle, V. 2007. Characterising the temporal variability of the spatial distribution of animals : an application to seabirds at sea. Ecography, 30: 695-708.

Charrier, G., Coombs, S.H., McQuinn I H and Laroche, J. 2007. Genetic structure of whiting *Merlangius merlangus* in the northeast Atlantic and adjacent waters. Marine Ecology Progress Series, 330: 201-211.

Chen, Y., Liggins, G.W., Graham, K.J. and Kennelly, S.J. 1997. Modelling the length-dependent offshore distribution of redfish, *Centroberyx affinis*. Fisheries Research, 29: 39-54.

Coombs, S.H., Nichols, J.H. and Fosh, C.A. 1990. Plaice eggs (*Pleuronectes platessa* L.) in the southern North Sea: abundance, spawning area, vertical distribution, and buoyancy. ICES Journal of Marine Science, 47: 133-139.

Corten, A. 2002. The role of "conservatism" in herring migrations. Reviews in Fish Biology and Fisheries, 11: 339-361.

Coull, K.A., Johnstone, R. and Rogers, S.I. 1998. Fisheries sensitivity maps in British waters.. 63 pp.

Coyne, M.S. and Christensen, J.D. 1997. Biogeography program : habitat suitability index modeling : species habitat suitability index values, technical Guidelines. NOAA, : 1-19.

Cury, P. 1994. Obstinate nature: an ecology of individuals. Thoughts on reproductive behaviour and biodiversity. Canadian Journal of Fisheries and Aquatic Sciences, 51: 1664-1673.

Cushing, D.H. 1990. Hydrographic containment of a spawning group of plaice in the Southern Bight of the North Sea. Marine Ecology Progress Series, 58: 287-297.

Daan, N., Bromley, P.J., Hislop, J.R.G. and Nielsen, N.A. 1990. Ecology of North Sea Fish. Netherlands Journal of Sea Research, 26: 343-386.

Dale, M.R.T. and Fortin, M. 2002. Spatial autocorrelation and statistical tests in ecology. Ecoscience, 9: 162-167.

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

De'ath, G. 2002. Multivariate regression trees: a new technique for modeling species environment relationship. Ecology, 83: 1105-1117.

Diniz-Filho, J.A.F., Rangel, T.F.L.V.B. and Bini, L.M. 2008. Model selection and information theory in geographical ecology. Global Ecology and Biogeography, 17: 479-488.

Dodson, J.J. 1988. The nature and role of learning in the orientation and migratory behavior of fishes. Environmental Biology of Fishes, 23: 161-182.

Doligez, B., Cadet, C., Danchin, E. and Boulinier, T. 2003. When to use public information for breeding habitat selection? The role of environmental predictability and density dependence. Animal Behaviour, 66: 973-988.

Dormann, C.F., McPherson, J.M., Araujo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G. et al. 2007a. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography, 30: 609-628.

Dormann, C.F. 2007b. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. Global Ecology and Biogeography, 16: 129-138.

Dormann, C.F. 2007c. Assessing the validity of autologistic regression. Ecological modelling, 207: 234-242.

Dragesund, O., Johannessen, A. and Ulltang, O. 1997. Variation in migration and abundance of Norwegian spring spawning herring (*Clupea harengus* L.). Sarsia, 82: 97-105.

Dray, S., Legendre, P. and Peres-Neto, P.R. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling, 196: 483-493.

Dray, S., Legendre, P. and Blanchet, G. 2007. packfor: Forward Selection with permutation (Canoco p.46). R package version 0.0-7.

ESRI 2005. Environmental Systems Research Institute. ArcGis. Version 9.1. Redlands, CA : Environmental Research Institute Inc.

Eastwood, P.D., Meaden, G.J. and Grioche, A. 2001. Modelling spatial variations in spawning habitat suitability for the sole *Sole solea* using regression quantiles and GIS procedures. Marine Ecology Progress Series, 224: 251-266.

Eastwood, P.D., Meaden, G.J., Carpentier, A. and Rogers, S.I. 2003. Estimating limits to the spatial extent and suitability of sole (*Solea solea*) nursery grounds in the Dover Strait. Journal of Sea Research, 50: 151-165.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J. et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography, 29: 129-151.

Fisher, J.A.D. and Frank, K.T. 2004. Abundance–distribution relationships and conservation of exploited marine fishes. Marine Ecology Progress Series, 279: 201-213.

Fotheringham, A.S., Brunsdon, C. and Charlton, M. 2002. Geographically weighted regression: the analysis of spatially varying relationships. Chichester: Wiley.

Fox, C.J., O'Brien, C.M., Dickey-Collas, M. and Nash, R.D.M. 2000. Patterns in the spawning of cod (*Gadus morhua* L.), sole (*Solea solea* L.) and plaice (*Pleuronectes platessa* L.) in the Irish Sea as determined by generalized additive modelling. Fisheries Oceanography, 9: 33-49.

Fretwell, S.D. and Lucas, H.L. 1970. On territorial behavior and other factors influencing habitat distribution in birds. Acta Biotheoritica, 19: 16-36.

Fretwell, S. 1972. Populations in a seasonal environment. Princeton University Press, New Jersey. 217 pp.

Friedman, J.H. 1991. Multivariate adaptive regression splines. The Annals of Statistics, 19: 1-67.

Friedman, J.H., Hastie, T. and Tibshirani, R. 2000. Additive logistic regression: a statistical view of boosting. Annals of Statistics, 28: 337-407.

Friedman, J.H. 2001. Greedy function approximation: the gradient boosting machine. Annals of Statistics, 29: 1189-1232.

Friedman, J.H. and Meulman, J.J. 2003. Multiple additive regression trees with application in epidemiology. Statistics in Medicine, 22: 1365-1381.

Froese, R. and Pauly, D. 2009. FishBase. World Wide Web electronic publication. www.fishbase.org. Version (01/2009).

GenStat Release 7.1. 2004. Lawes Agricultural Trust (Rothamsted Experimental Station). R.W Payne, S.A Harding, D.A. Murray, D.M. Soutar, D.B. Baird, S.J. Welham, A.F. Kane, A.R. Gilmour, R. Thompson, R. Webster, G. Tunnicliffe Wilson. VSN International, Wilkinson House, Jordan Hill Road, Oxford, UK.

Gibb, F.M., Wright, P.J., Gibb, I.M. and O'Sullivan, M. 2004. Haddock and whiting spawning areas in the North Sea and Scottish west coast. 26 pp.

Gibson, R.N. 1997. Behaviour and the distribution of flatfishes. Journal of Sea Research, 37: 241-256.

Gordoa, A. and Duarte, C.M. 1991. Size-dependent spatial distribution of hake (*Merluccius capensis* and *Merluccius paradoxus*) in Namibian waters. Canadian Journal of Fisheries and Aquatic Sciences, 48: 2095-2099.

Greenstreet, S.P.R. and Hall, S.J. 1996. Fishing and the ground-fish assemblage structure in the north-western North Sea: an analysis of long-term and spatial trends. Journal of Animal Ecology, 65: 577-598.

Guisan, A. and Zimmermann, N.E. 2000. Predictive habitat distribution models in ecology. Ecological Modelling, 135: 147-186.

Guisan, A., Edwards, T.C. and Hastie, T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling, 157: 89-100.

Guisan, A. and Thuillier, W. 2005. Predicting species distribution: offering more than simple habitats models. Ecology Letters, 8: 993-1009.

Guisan, A., Lehmann, A., Ferrier, S., Austin, M.P., Overton, J.M.C., Aspinall, R. and Hastie, T. 2006. Making better biogeographical predictions of species' distributions. Journal of Applied Ecology, 43: 386-392.

Han, Y.S. and Tzeng, W.N. 2007. Sex-dependent habitat use by the Japanese eel *Anguilla japonica* in Taiwan. Marine Ecology Progress Series, 338: 193-198.

Harden Jones, F.R. 1968. Fish migration. Edward Arnold, London. 325 pp.

Harding, D., Nichols, J.H. and Tungate, D.S. 1978. The spawning of plaice (*Pleuronectes platessa*) in the southern North Sea and English Channel. Rapports et Procès Verbaux des Réunions du Conseil International pour l'Exploration de la Mer, 172: 102-113.

Hastie, T.J. and Tibshirani, R.J. 1990. Generalized additive models. Chapman and Hall, London. 335 pp.

Hastie, T., Tibshirani, R.J. and Friedman, J.H. 2001. The elements of statistical learning. Springer-Verlag. 533 pp.

Hastie, T. 2006. gam: generalized additive models. R package version 0.98.

Helfman, G.S. and Schultz, E.T. 1984. Social transmission of behavioural traditions in a coral reef fish. Animal Behaviour, 32: 379-384.

Hiddink, J.G. 2005. Implications of Liebig’s law of the minimum for the use of ecological indicators based on abundance. Ecography, 28: 264-271.

Hilborn, R. and Mangel, M. 1997. The ecological detective. Confronting models with data. Monographs in population biology, 28. Princeton University Press, Princeton. 315 pp.

Hill, A.E., James, I.D., Linden, P.F., Matthews J P, Prandle D, Simpson, J.H., Gmitrowicz, E.M. et al. 1993. Dynamics of tidal mixing fronts in the North Sea [and discussion]. Philosophical Transactions: Physical Sciences and Engineering, 343: 431-446.

Hirzel, A.H., Hausser, J., Chessel, D. and Perrin, N. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? Ecology, 83: 2027-2036.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C. and Guisan, A. 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling, 199: 142-152.

Hislop, J.R.G. and MacKenzie, K. 1976. Population studies of the whiting *Merlangius merlangus* (L.) of the northern North Sea. ICES Journal of Marine Science, 37: 98-110.

Hislop, J.R.G., Robb, A.P., Bell, M.A. and Armstrong, D.W. 1991. The diet and food consumption of whiting (*Merlangius merlangus*) in the North Sea. ICES Journal of Marine Science, 48: 139-156.

Hislop, J.R.G. 1996. Changes in North Sea gadoid stocks. ICES Journal of Marine Science, 53: 1146-1156.

Hobbs, J.P. and Munday, P.L. 2004. Intraspecific competition controls spatial distribution and social organisation of the coral dwelling goby, *Gobiodon histrio*. Marine Ecology Progress Series, 278: 253-259.

Holst, J.C., Rottingen, I. and Melle, W. 2004. The herring. In The Norwegian Sea ecosystem, pp. 203-226. Ed. by H.R. Skjoldal, R. Saetre, A. Faerno, O.A. Misund, and I. Rottingen. Trondheim, Norway: Tapir Academic Press.

Hunter, E., Metcalfe, J.D. and Reynolds, J.D. 2003. Migration route and spawning area fidelity by North Sea plaice. Proceedings of the Royal Society of London, Series B: Biological Sciences, 270: 2097-2103.

Hunter, E., Metcalfe, J.D., Arnold, G.P. and Reynolds, J.D. 2004. Impacts of migratory behaviour on population structure in North Sea plaice. Journal of Animal Ecology, 73: 377-385.

Huret, M., Struski, C., Léger, F., Petitgas, P., Lazure, P. and Sourisseau, M. 2009. Modélisation couplée physique-biogéochimie du Golfe de Gascogne sur la période 1971-2007. 26 pp.

Huse, G., Railsback, S. and Feronö, A. 2002. Modelling changes in migration pattern of herring: collective behaviour and numerical domination. Journal of Fish Biology, 60: 571-582.

Hutchinson, G.E. 1957. Concluding remarks. Cold Spring Harbor Symposium on Quantitative Biology, 22: 415-427.

ICES 2004a. Manual for the international bottom trawl surveys. Revision VII. The International Bottom Trawl Survey Working Group. Resource Management Committee. ICES CM 2004. 51 pp.

ICES 2004b. Report of the study group on regional scale ecology of small pelagic fish (SGRESP), 23-26 February 2004, Nantes, France. ICES CM 2004/G:06 Ref. ACFM, ACE. 55 pp.

ICES 2005. Report of the Planning Group on North Sea Cod and Plaice Egg Surveys in the North Sea (PGEGGS), 10–12 May 2005, Lowestoft, UK. ICES Document CM 2005/G:11. 85 pp.

ICES 2006. Report of the working group on acoustic and egg surveys for sardine and anchovy in ICES areas VIII and IX, 24-28 October 2005, Vigo, Spain. ICES CM 2006/LRC:01 Ref. ACFM. 126 pp.

ICES 2007a. Report of the International Bottom Trawl Survey Working Group (IBTSWG), 27-30 March 2007, Sète, France. ICES CM 2007/RMC:05. 200 pp.

ICES 2007b. Report of the workshop on testing the entrainment hypothesis (WKTEST). ICES CM, 2007/LRC:10. 124 pp.

ICES 2008a. Report of the working group on the assessment of demersal stocks in the North Sea and Skagerrak - combined spring and autumn (WGNSSK), 7-13 May 2008, Copenhagen, Denmark. ICES CM 2008/ACOM:09. 921 pp.

ICES 2008b. Report of the working group on anchovy, 13-16 June 2008, Copenhagen, Denmark. ICES CM 2008/ACOM:04. 230 pp.

Irigoien, X., Fiksen, O., Cotano, U., Uriarte, A., Alvarez, P., Arrizabalaga, H., Boyra, G. et al. 2007. Could Biscay Bay Anchovy recruit through a spatial loophole? Progress in Oceanography, 74: 132-148.

Irigoien, X., Cotano, U., Boyra, G., Santos, M., Alvarez, P., Otheguy, P., Etxebeste, E. et al. 2008. From egg to juvenile in the Bay of Biscay: spatial patterns of anchovy (*Engraulis encrasicolus*) recruitment in a non-upwelling region. Fisheries Oceanography, 17: 446-462.

Irigoien, X., Fernandes, J.A., Grosjean, P., Denis, K., Albaina, A. and Santos, M. 2009. Spring zooplankton distribution in the Bay of Biscay from 1998 to 2006 in relation with anchovy recruitment. Journal of Plankton Research, 1: 1-17.

Irisson, J. 2008. Approche comportementale de la dispersion larvaire en milieu marin. Thèse de l'Ecole Pratique des Hautes Etudes. Perpignan. 214 pp.

Johnson, J.B. and Omland, K.S. 2004. Model selection in ecology and evolution. Trends in Ecology and Evolution, 19: 101-108.

Jorgensen, C., Enberg, K., Dunlop, E.S., Arlinghaus, R., Boukal, D.S., Brander, K., Ernande, B. et al. 2007. Ecology: managing evolving fish stocks. Science, 318: 1247-1248.

Joy, M.K. and Death, R.G. 2004. Predictive modelling and spatial mapping of freshwater fish and decapod assemblages using GIS and neural networks. Freshwater Biology, 49: 1036-1052.

Keitt, T.H., Bjornstad, O.N., Dixon, P.M. and Citron-Pousty, S. 2002. Accounting for spatial pattern when modeling organism-environment interactions. Ecography, 25: 616-625.

Kissling, W.D. and Carl, G. 2008. Spatial autocorrelation and the selection of simultaneous autoregressive models. Global Ecology and Biogeography, 17: 59-71.

Koenker, R. and Basset Jr., G. 1978. Regression quantiles. Econometrica, 46: 33-50.

Koenker, R. 2005. Quantile regression. Cambridge University Press, New York, USA. 349 pp.

Koubbi, P., Duhamel, G., Harlay, X., Durand, I. and Park, Y.H. 2003. Distribution of larval *Krefftichtys anderssoni* (Myctophidae, Pisces) at the Kerguelen Archipelago (Southern Indian Ocean) modelled using GIS ans Habitat suitability. In Antarctic Biology in a Global Context, pp. 215-223. Ed. by A.H.L. Huiskes, W.W.C. Giskes, J. Rozema, R.M.L. Schomo, S.M. van der vies, and W.J. Wolf. Backhuys, Leiden.

Koubbi, P., Loots, C., Cotonnec, G., Harlay, W., Grioche, A., Vaz, S., Martin, C. et al. 2006. Spatial patterns and GIS habitat modelling of *Solea solea*, *Pleuronectes flesus* and *Limanda Limanda* fish larvae in the eastern English Channel during the spring. Scientia Marina, 70S2: 147-157.

Koutsikopoulos, C. and Le Cann, B. 1996. Physical processes and hydrological structures related to the Bay of Biscay anchovy. Scientia Marina, 60: 9-19.

Laurel, B.J., Stoner, A.W. and Hurst, T.P. 2007. Density-dependent habitat selection in marine flatfish: the dynamic role of ontogeny and temperature. Marine Ecology Progress Series, 338: 183-192.

Lazure, P. and Dumas, F. 2008. An external-internal mode coupling for a 3D hydrodynamical model for applications at regional scale (MARS). Advances in Water Resources, 31: 233-250.

Le Pape, O., Chauvet, F., Mahévas, S., Lazure, P., Guérault, D. and Désaunay, Y. 2003. Quantitative description of habitat suitability for the juvenile common sole (*Solea solea*, L.) in the Bay of Biscay (France) and the contribution of different habitats to the adult population. Journal of Sea Research, 50: 139-149.

Leathwick, J.R., Elith, J. and Hastie, T. 2006a. Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. Ecological Modelling, 199: 188-196.

Leathwick, J.R., Elith, J., Francis, M.P., Hastie, T. and Taylor, P. 2006b. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. Marine Ecology Progress Series, 321: 267-281.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology, 74: 1659-1673.

Legendre, P. and Legendre, L. 1998. Numerical Ecology. Elsevier, Amsterdam. 853 pp.

Lehmann, A. 1998. GIS modeling of submerged macrophyte distribution using generalized additive models. Plant Ecology, 139: 113-124.

Lemon, J., Bolker, B., Oom, S., Klein, E., Rowlingson, B., Wickham, H., Tyagi, A. et al. 2008. plotrix: Various plotting functions. R package version 2.5.

Levin, S.A. 1992. The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. Ecology, 73: 1943-1967.

Lichstein, J.W., Simons, T.R., Shriner, S.A. and Franzreb, K.E. 2002. Spatial autocorrelation and autoregressive models in ecology. Ecological Monographs, 72: 445-463.

Loots, C., Koubbi, P. and Duhamel, G. 2007. Habitat modelling of *Electrona antarctica* (Myctophidae, Pisces) in Kerguelen by generalized additive models and geographic information systems. Polar Biology, 30: 951-959.

Loots, C., Vaz, S., Planque, B. and Koubbi, P. sous presse. What controls the spatial distribution of North Sea plaice spawning population? Confronting ecological hypotheses through a model selection framework. ICES Journal of Marine Science.

MacNally, R. 2000. Regression and model-building in conservation biology, biogeography and ecology: The distinction between – and reconciliation of – ‘predictive’ and ‘explanatory’ models. Biodiversity and Conservation, 9: 655-671.

Maggini, R., Lehmann, A., Zimmermann, N.E. and Guisan, A. 2006. Improving generalized regression analysis for the spatial prediction of forest communities. Journal of Biogeography, 33: 1729-1749.

Maravelias, C.D., Haralabous, J. and Papaconstantinou, C. 2003. Predicting demersal fish species distributions in the Mediterranean Sea using artificial neural networks. Marine Ecology Progress Series, 255: 249-258.

Martin, C.S., Lelièvre, S. and Vaz, S. 2007. Identification des frayères du Détroit du Pas-de-Calais et des zones maritimes adjacentes, Rapport final du projet ISADO, Programme Interreg IIIa. Canterbury (RU): Canterbury Christ Church University. 114 pp.

Matheron, G. 1962. Traité de géostatistique appliquée. *In* Mémoires du Bureau de Recherches Géologiques et Minières. Ed. by Editions Technip. Paris, 504 pp.

Maxwell, D.L., Stelzenmuller, V., Eastwood, P.D. and Rogers, S.I. 2009. Modelling the spatial distribution of plaice (*Pleuronectes platessa*), sole (*Solea solea*) and thornback ray (*Raja clavata*) in UK waters for marine management and planning. Journal of Sea Research, 61: 258-267.

McCall, A.D. 1990. Dynamic geography of marine fish populations. University of Washington Press, United States. 153 pp.

McCullagh, P. and Nelder, J.A. 1989. Generalized linear models, Monographs on statistics and applied probability 37. Chapman and Hall, London. 511 pp.

McCune, B. and Mefford, M.J. 1999. Multivariate Analysis of Ecological Data. Version 4.25. MjM Software, Gleneden Beach, Oregon, U.S.A. 200 pp.

McCune, B. 2006. Non-parametric habitat models with automatic interactions. Journal of Vegetation Science, 17: 819-830.

McQuinn, I.H. 1997. Metapopulations and the Atlantic herring. Reviews in Fish Biology and Fisheries, 7: 297-329.

Mello, L.G.S. and Rose, G.A. 2005. Using geostatistics to quantify seasonal distribution and aggregation patterns of fishes: an example of Atlantic Cod (*Gadus morhua*). Canadian Journal of Fisheries and Aquatic Sciences, 62: 659-670.

Metcalfe, J.D. 2006. Fish population structuring in the North Sea: understanding processes and mechanisms from studies of the movements of adults. Journal of Fish Biology, 69: 48-65.

Mitchell, S.C. 2005. How useful is the concept of habitat? - a critique. Oikos, 110: 634-638.

Moisen, G.G. and Frescino, T.S. 2002. Comparing five modelling techniques for predicting forest characteristics. Ecological Modelling, 157: 209-225.

Moisen, G.G., Freeman, E.A., Blackard, J.A., Frescino, T.S., Zimmermann, N.E. and Edwards Jr, T.C. 2006. Predicting tree species presence and basal area in Utah: A comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. Ecological Modelling, 199: 176-187.

Motos, L., Uriarte, A. and Valencia, V. 1996. The spawning environment of the Bay of Biscay anchovy (*Engraulis encrasicolus* L.). Scientia Marina, 60: 117-140.

Nichols, J.H. 1989. The diurnal rhythm in spawning of plaice (*Pleuronectes platessa* L.) in the southern North Sea. Journal Conseil International pour l'Exploration de la Mer, 45: 277-283.

Nishida, T. and Chen, D. 2004. Incorporating spatial autocorrelation into the general linear model with an application to the yellowfin tuna (*Thunnus albacares*) longline CPUE data. Fisheries Research, 70: 265-274.

Oksanen, J. and Minchin, P.R. 2002. Continuum theory revisited: what shape are species responses along ecological gradients? Ecological Modelling, 157: 119-129.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G.L., Solymos, P., Stevens, M.H.H. et al. 2009. vegan: Community Ecology Package. R package version 1.15-3.

Orians, G.H. and Wittenberger, J.F. 1991. Spatial and temporal scales in habitat selection. The American Naturalist, 137: S29-S49.

Papi, F. 1992. Animal homing. Chapman & Hall, London. 390 pp.

Pastoors, M.A., Rijnsdorp, A.D. and Van Beek, F.A. 2000. Effects of a partially closed area in the North Sea ("plaice box") on stock develpment of Plaice. ICES Journal of Marine Science, 57: 1014-1022.

Pearce, J.L. and Boyce, M. 2006. Modelling distribution and abundance with presence-only data. Journal of Applied Ecology, 43: 405-412.

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

Pedersen, J. 1999. Diet comparison between pelagic and demersal whiting in the North Sea. Journal of Fish Biology, 55: 1096-1113.

Petitgas, P. 1993. Geostatistics for fish stock assessment : a review and acoustic application. ICES Journal of Marine Science, 50: 285-298.

Petitgas, P. 2001. Geostatistics in fisheries survey design and stock assessment: models, variances and applications. Fish and Fisheries, 2: 231-249.

Petitgas, P., Massé, J., Beillois, P., Lebarbier, E. and Le Cann, A. 2003. Sampling variance of species identification in fisheries-acoustic surveys based on automated procedures associating acoustic images and trawl hauls. ICES Journal of Marine Science, 60: 437-445.

Petitgas, P., Reid, D., Planque, B., Nogueira, E., O'Hea, B. and Cotano, U. 2006. The entrainment hypothesis: an explanation for the persistence and innovation in spawning migration and life cycle patterns. ICES CM 2006/B:07. 9 pp.

Petitgas, P. 2008. Fish habitat mapping with Empirical Orthogonal Functions. ICES CM 2008/M:07.

Planque, B., Lazure, P. and Jégou, A. 2004. Detecting hydrological landscapes over the Bay of Biscay continental shelf in spring. Climate Research, 28: 41-52.

Planque, B., Lazure, P. and Jegou, A. 2006. Typology of hydrological structures modelled and observed over the Bay of Biscay shelf. Scientia Marina, 70S1: 43-50.

Planque, B., Bellier, E. and Lazure, P. 2007. Modelling potential spawning habitat of sardine (*Sardina pilchardus*) and anchovy (*Engraulis encrasicolus*) in the Bay of Biscay. Fisheries Oceanography, 16: 16-30.

Planque, B. and Buffaz, L. 2008. Quantile regression models for fish recruitment-environment relationship - four case studies. Marine Ecology Progress Series, 357: 213-223.

Planque, B., Fromentin, J., Cury, P., Drinkwater, K.F., Jennings, S., Perry, R.I. and Kifani, S. Sous presse. How does fishing alter marine populations and ecosystems sensitivity to climate? Journal of Marine Systems.

Pope, J.G. and Macer, C.T. 1996. An evaluation of the stock structure of North Sea cod, haddock, and whiting since 1920, together with a consideration of the impacts of fisheries and predation effects on their biomass and recruitment. ICES Journal of Marine Science, 53: 1157-1169.

R Development Core Team 2008. R : a language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria.

Randin, C.F., Dirnbock, T., Dullinger, S., Zimmermann, N.E., Zappa, M. and Guisan, A. 2006. Are niche-based species distribution models transferable in space? Journal of Biogeography, 33: 1689-1703.

Ribeiro Jr, P.J. and Diggle, P.J. 2001. geoR: a package for geostatistical analysis. 15-18 pp.

Rico, C., Ibrahim, K.M., Rico, I. and Hewitt, G.M. 1997. Stock composition in North Atlantic populations of whiting using microsatellite markers. Journal of Fish Biology, 51: 462-475.

Rijnsdorp, A.D. 1989. Maturation of male and female North Sea plaice (*Pleuronectes platessa* L.). Journal Conseil International pour l'Exploration de la Mer, 46: 35-51.

Rijnsdorp, A.D. and Millner, R.S. 1996. Trends in population dynamics and exploitation of North Sea plaice (*Pleuronectes platessa* L.) since the late 1800s. ICES Journal of Marine Science, 53: 1170-1184.

Rijnsdorp, A.D., Grift, R.E. and Kraak, S.B.M. 2005. Fisheries-induced adaptive change in reproductive investment in North Sea plaice (*Pleuronectes platessa*)? Canadian Journal of Fisheries and Aquatic Sciences, 62: 833-843.

Rindorf, A. and Lewy, P. 2006. Warm, windy winters drive cod north and homing of spawners keeps them there. Journal of Applied Ecology, 43: 445–453.

Rushton, S.P., Ormerod, S.J. and Kerby, G. 2004. New paradigms for modelling species distribution ? Journal of Applied Ecology, 41: 193-200.

Ryer, C.H., Stoner, A.W., Spencer, M.L. and Abookire, A.A. 2007. Presence of larger flatfish modifies habitat preference by Age-0 northern rock sole *Lepidopsetta polyxystra*. Marine Ecology Progress Series, 342: 227-238.

Schilling, N.S. 2005. Survival of the fittest: fish in patchy environments show ideal free distribution (IFD). Eukaryon, 1: 11-16.

Schlüter, M. and Jerosch, K. 2008. Digital atlas of the North Sea (DANS). Geo-information regarding geology, geochemistry, oceanography and biology. Available at http://www.awi.de/en/research/research_divisions/geosciences/marine_geochemistry/marine_gis/digital_atlas_of_the_north_sea/.

Schrum, C., Alekseeva, I. and St. John, M. 2006. Development of coupled physical–biological ecosystem model ECOSMO Part I: Model description and validation for the North Sea. Journal of Marine Systems, 61: 79-99.

Segurado, P. and Araujo, M.B. 2004. An evaluation of methods for modelling species distributions. Journal of Biogeography, 31: 1555-1568.

Segurado, P., Araujo, M.B. and Kunin, W.E. 2006. Consequences of spatial autocorrelation for niche-based models. Journal of Applied Ecology, 43: 433-444.

Serchuk, F.M., Kirkegaard, E. and Daan, N. 1996. Status and trends of the major roundfish, flatfish, and pelagic fish stocks in the North Sea: thirty-year overview. ICES Journal of Marine Science, 53: 1130-1145.

Shepherd, T.D. and Litvak, M. 2004. Density-dependent habitat selection and the ideal free distribution in marine fish spatial dynamics: considerations and cautions. Fish and Fisheries, 5: 141-152.

Simpson, A.C. 1971. Diel spawning behaviour in populations of Plaice, Dab, Sprat and Pilchard. Journal Conseil International pour l'Exploration de la Mer, 34: 58-64.

Solemdal, P. 1997. Maternal effects -- a link between the past and the future. Journal of Sea Research, 37: 213-227.

Solmundsson, J., Palsson, J. and Karlsson, H. 2005. Fidelity of mature Icelandic plaice (*Pleuronectes platessa*) to spawning and feeding grounds. ICES Journal of Marine Science, 62: 189-200.

Stefánsson, G. 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES Journal of Marine Science, 53: 577-588.

Stephens, P.A., Buskirk, S.W. and del Rio, C.M. 2007. Inference in ecology and evolution. Trends in Ecology & Evolution, 22: 192-197.

Stratoudakis, Y., Gallego, A. and Morrison, J.A. 1998. Spatial distribution of developmental egg ages within a herring *Clupea harengus* spawning ground. Marine Ecology Progress Series, 174: 27-32.

Stratoudakis, Y., Fryer, R.J., Cook, R.M. and Pierce, G.J. 1999. Fish discarded from Scottish demersal vessels: estimators of total discards and annual estimates for targeted gadoids. ICES Journal of Marine Science, 56: 592-605.

Student 1914. The elimination of spurious correlation due to position in time or space. Biometrika, 10: 179-181.

Sutherland, W.J. 1983. Aggregation and the ideal free distribution. Journal of Animal Ecology, 52: 821-828.

Sutton, C.D., Rao, C.R., Wegman, E.J. and Sodka, J.L. 2005. Classification and regression trees, bagging, and boosting. *In* Handbook of Statististics, pp. 303-329. Ed. by Elsevier.

Swain, D.P. 1993. Age- and density-dependent bathymetric pattern of Atlantic Cod (*Gadus morhua*) in the southern Gulf of St. Lawrence. Canadian Journal of Fisheries and Aquatic Sciences, 50: 1255-1264.

Swain, D.P. and Morin, R. 1996. Relationships between geographic distribution and abundance of American plaice (*Hippoglossoides platessoides*) in the southern Gulf of St. Lawrence. Canadian Journal of Fisheries and Aquatic Sciences, 53: 106-119.

Switzer, P.V. 1997. Past reproductive success affects future habitat selection. Behavioral ecology and sociobiology, 40: 307-312.

Talbot, J.W. 1977. The dispersal of plaice eggs and larvae in the Southern Bight of the North Sea. Journal Conseil International pour l'Exploration de la Mer, 37: 221-248.

Taylor, K.E. 2001. Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research, 106: 7183-7192.

Telford, R.J. and Birks, H.J.B. 2005. The secret assumption of transfer functions: problems with spatial autocorrelation in evaluating model performance. Quaternary Science Reviews, 24: 2173-2179.

Thuiller, W. 2003. BIOMOD – optimizing predictions of species distributions and projecting potential future shifts under global change. Global Change Biology, 9: 1353-1362.

Tregenza, T. 1995. Building on the ideal free distribution. Advances in Ecological Research, 26: 253-302.

Tyler, J.A. and Hargrove, W.H. 1997. Predicting spatial distribution of foragers over large resource landscapes: a modeling analysis of the ideal free distribution. Oikos, 79: 376-386.

Usio, N. 2007. Endangered crayfish in northern Japan: Distribution, abundance and microhabitat specificity in relation to stream and riparian environment. Biological Conservation, 134: 517-526.

Valone, T.J. and Templeton, J.J. 2002. Public information for the assesment of quality: a widespread social phenomenon. Philosophical transaction of the royal society of London B, 357: 1549-1557.

Vaz, S., Martin, C.S., Eastwood, P.D., Ernande, B., Carpentier, A., Meaden, G.J. and Coppin, F. 2008. Modelling species distributions using regression quantiles. Journal of Applied Ecology, 45: 204-217.

van der Veer, H.K., Pihl, L. and Bergman, M.J.N. 1990. Recruitment mechanisms in North Sea plaice *Pleuronectes platessa*. Marine Ecology Progress Series, 64: 1-12.

van der Veer, H.W., Ruardij, P., van den Berg, A.J. and Ridderinkhof, H. 1998. Impact of interannual variability in hydrodynamic circulation on egg and larval transport of plaice *Pleuronectes platessa* L. in the southern North Sea. Journal of Sea Research, 39: 29-40.

Venables, W.N. and Ripley, D.B. 2002. Modern applied statistics with S. Springer, New York. 495 pp.

Webster, R. and Oliver, M.A. 2001. Geostatistics for environmental scientists. Statistics in practice. Wiley, Chichester, England. 271 pp.

Wegner, G., Damm, U. and Purps, M. 2003. Physical influences on the stock dynamics of plaice and sole in the North Sea. Scientia Marina, 67: 219-234.

Werner, E.E. and Gilliam, J.F. 1984. The ontogenic niche and species interactions in size-structured populations. Annual Review of Ecology and Systematics, 15: 393-425.

Wieland, K. and Jarre-Teichmann, A. 1997. Prediction of vertical distribution and ambient development temperature of Baltic cod, *Gadus morhua* L., eggs. Fisheries Oceanography, 6: 172-187.

Wintle, B.A. and Bartos, D.C. 2006. Modeling Species-habitat relationships with spatially autocorrelated observation data. Ecological Applications, 16: 1945-1958.

Woillez, M., Poulard, J., Rivoirard, J., Petitgas, P. and Bez, N. 2007. Indices for capturing spatial patterns and their evolution in time, with application to European hake (*Merluccius merluccius*) in the Bay of Biscay. ICES Journal of Marine Science, 64: 537-550.

Wood, S.N. and Augustin, N.H. 2002. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. Ecological Modelling, 157: 157-177.

Yee, T.W. and Mitchell, N.D. 1991. Generalized additive models in plant ecology. Journal of Vegetation Science, 2: 587-602.

Zheng, X., Pierce, G.J. and Reid, D.G. 2001. Spatial patterns of whiting abundance in Scottish waters and relationships with environmental variables. Fisheries Research, 50: 259-270.

Zheng, X., Pierce, G.J., Reif, D.G. and Jolliffe, I.T. 2002. Does the North Atlantic current affect spatial distribution of whiting? Testing environmental hypotheses using statistical and GIS techniques. ICES Journal of Marine Science, 59: 239-253.

ter Braak, C.J.F. and Smilauer, P. 2002. CANOCO reference manual and CanoDraw for windows User's guide: software for canonical community ordination (version 4.5). Microcomputer Power, Ithaca, New York. 500 pp.

van der Molen, J., Rogers, S.I., Ellis, J.R., Fox, C.J. and McCloghrie, P. 2007. Dispersal patterns of the eggs and larvae of spring-spawning fish in the Irish Sea, UK. Journal of Sea Research, 58: 313-330.

van der Ploeg, R.R., Bohm, W. and Kirkham, M.B. 1999. On the origin of the theory of mineral nutrition of plants and the law of the minimum. Soil Science Society of America Journal, 63: 1055-1062.

Annexe A

Analyses géostatistiques

A.1 Rappel des paramètres

Données

Variable: *type de variable analysé*

Année: *année d'analyse*

Trans: *transformation appliquée à la variable*

Tendance spatiale

Tendance: pourcentage de variation expliqué par la tendance spatiale

Tendance quadratique : $variable = a + bX + cY + dX^2 + eY^2 + f(X*Y)$

Tendance linéaire: $variable = a + bX + cY$

X = *latitude*

Y= *longitude corrigée*

Variogramme

X_{max}: *distance maximale étudiée*

Pas: *intervalle des classes de distance*

Modèle: *modèle utilisé comme variogramme théorique*

ModVar: *mesure d'ajustement du variogramme théorique (%)*

Pépite: *valeur de la pépite (nugget)*

Pallier: *valeur du pallier (sill)*

Portée: *valeur de la portée (range)*

Paramètres de krigeage

Maille: *maille de la grille d'interpolation (mesh)*

Rayon: *rayon de recherche des voisins*

Min/Max: *nombre minimum/maximum de voisins utilisés pour le krigeage*

A.2 Tableaux récapitulatifs des valeurs obtenues pour les trois espèces

A.2.1 Abondances observées pour la plie de 1980 à 2007

Données			**Tendance spatiale**							**Variogramme**							**Krigeage**			
Variable	Année	Trans	Tendance	a	b	c	d	e	f	Xmax	Pas	Modèle	ModVar	Pépite	Pallier	Portée	Maille	Rayon	Min	Max
Abondances	1980	log(x+1)	27,7	0,23	-0,09	0,04	0,05	0,00	-0,031	-	-	Pure pépite	-	0,45	0,00	2,00	0,2	0,4	1	5
Abondances	1981	log(x+1)	21,4	-0,06	-0,18	0,60	0,07	-0,08	-0,066	-	-	Pure pépite	-	0,60	0,00	2,00	0,2	0,4	1	4
Abondances	1982	log(x+1)	15,9	-	-	-	-	-	-	2,5	0,350	Circulaire	99,9	0,33	0,25	1,51	0,2	0,4	1	4
Abondances	1983	log(x+1)	36,5	-0,84	0,43	0,54	0,02	-0,05	-0,105	1,3	0,240	Circulaire	64,7	0,32	0,13	0,67	0,2	0,4	1	4
Abondances	1984	log(x+1)	22,5	-0,45	0,59	0,37	-0,01	-0,03	-0,097	1,5	0,300	Circulaire	96,2	0,38	0,19	0,88	0,2	0,4	1	6
Abondances	1985	log(x+1)	40,4	1,02	-0,30	0,33	0,09	-0,05	-0,048	0,6	0,115	Circulaire	82,3	0,23	0,11	0,26	0,2	0,4	1	6
Abondances	1986	log(x+1)	23,5	-1,14	0,64	0,73	-0,02	-0,07	-0,095	2,0	0,350	Circulaire	99,4	0,15	0,38	1,35	0,2	0,4	1	7
Abondances	1987	log(x+1)	37,4	0,40	0,42	0,24	0,00	-0,03	-0,086	1,2	0,245	Circulaire	99,8	0,19	0,31	0,86	0,2	0,4	1	7
Abondances	1988	log(x+1)	36,4	2,94	-0,76	-0,22	0,13	-0,01	0,002	1,6	0,350	Circulaire	99,4	0,19	0,38	1,26	0,2	0,4	1	6
Abondances	1989	log(x+1)	31,2	1,34	-0,48	0,33	0,10	-0,05	-0,009	3,0	0,500	Circulaire	99,7	0,23	0,39	2,53	0,2	0,4	1	6
Abondances	1990	log(x+1)	43,5	2,30	-0,63	-0,08	0,11	-0,02	-0,002	0,6	0,140	Circulaire	93,8	0,22	0,10	0,30	0,2	0,4	1	6
Abondances	1991	log(x+1)	43,9	1,72	-0,46	0,02	0,10	-0,02	-0,029	2,2	0,460	Circulaire	99,9	0,11	0,13	1,49	0,2	0,4	1	5
Abondances	1992	log(x+1)	34,9	1,15	-0,44	0,41	0,09	-0,06	-0,024	3,2	0,670	Circulaire	99,4	0,27	0,16	2,06	0,2	0,4	1	6
Abondances	1993	log(x+1)	29,2	0,17	-0,04	0,45	0,08	-0,04	-0,085	4,0	0,500	Circulaire	98,5	0,26	0,09	2,96	0,2	0,4	1	5
Abondances	1994	log(x+1)	15,2	-	-	-	-	-	-	-	-	Pentasphérique	99,9	0,19	0,38	4,64	0,2	0,4	1	5
Abondances	1995	log(x+1)	21,5	1,26	-0,27	0,00	0,07	-0,01	-0,013	4,0	0,680	Circulaire	99,3	0,20	0,23	2,44	0,2	0,4	1	5
Abondances	1996	log(x+1)	14,2	-	-	-	-	-	-	-	-	Circulaire	98,8	0,29	0,33	3,85	0,2	0,4	1	4
Abondances	1997	log(x+1)	18,9	-	-	-	-	-	-	-	-	Circulaire	99,6	0,11	0,47	2,86	0,2	0,4	1	5
Abondances	1998	log(x+1)	17,7	-	-	-	-	-	-	-	-	Circulaire	99,0	0,25	0,30	3,24	0,2	0,4	1	6
Abondances	1999	log(x+1)	21,1	0,41	-0,06	0,16	0,07	-0,01	-0,055	3,2	0,500	Circulaire	99,6	0,16	0,38	2,28	0,2	0,4	1	6
Abondances	2000	log(x+1)	14,8	-	-	-	-	-	-	-	-	Circulaire	99,6	0,15	0,46	3,08	0,2	0,4	1	6
Abondances	2001	log(x+1)	12,4	-	-	-	-	-	-	-	-	Circulaire	99,0	0,27	0,42	3,43	0,2	0,4	1	6
Abondances	2002	log(x+1)	14,6	-	-	-	-	-	-	-	-	Circulaire	99,0	0,27	0,29	2,92	0,2	0,4	1	6
Abondances	2003	log(x+1)	9,7	-	-	-	-	-	-	-	-	Circulaire	98,7	0,29	0,33	2,68	0,2	0,4	1	6
Abondances	2004	log(x+1)	6,5	-	-	-	-	-	-	-	-	Circulaire	99,7	0,29	0,27	3,22	0,2	0,4	1	5
Abondances	2005	log(x+1)	11,8	-	-	-	-	-	-	-	-	Circulaire	99,5	0,23	0,31	3,15	0,2	0,4	1	5
Abondances	2006	log(x+1)	11,8	-	-	-	-	-	-	-	-	Circulaire	99,6	0,21	0,40	2,57	0,2	0,4	1	5
Abondances	2007	log(x+1)	10,1	-	-	-	-	-	-	-	-	Circulaire	99,3	0,12	0,55	2,55	0,2	0,4	1	5

A.2.2 Abondances prédites par le modèle d'habitat construit pour la plie de 1970 à 2007

Données			Tendance spatiale							Variogramme							Krigeage			
Variable	Année	Trans	Tendance	a	b	c	d	e	f	Xmax	Pas	Modèle	ModVar	Pépite	Pallier	Portée	Maille	Rayon	Min	Max
Abondances	1970	log(x+1)	73,8	1,369	-0,15	-0,111	0,05	-0,021	0,003	2,5	0,37	Circulaire	99,8	0,012	0,05	1,92	0,2	0,4	1	5
Abondances	1971	log(x+1)	71,1	1,606	-0,34	-0,027	0,071	-0,012	-0,026	2,6	0,37	Circulaire	99,7	0,005	0,07	2,11	0,2	0,4	1	6
Abondances	1972	log(x+1)	81,2	1,665	-0,38	-0,136	0,09	-0,006	-0,005	2,9	0,38	Circulaire	98,2	0,008	0,05	2,25	0,2	0,4	1	5
Abondances	1973	log(x+1)	76,1	1,973	-0,54	-0,115	0,09	-0,007	-0,005	3	0,38	Circulaire	97,5	0,014	0,03	2,24	0,2	0,4	1	6
Abondances	1974	log(x+1)	82,1	1,897	-0,49	-0,224	0,09	0,011	-0,021	2,4	0,2	Circulaire	84,9	0,017	0,02	1,58	0,2	0,4	1	4
Abondances	1975	log(x+1)	83,4	2,345	-0,50	-0,353	0,08	0,013	0,014	1,8	0,26	Circulaire	100	0,001	0,04	1,08	0,2	0,4	1	6
Abondances	1976	log(x+1)	76,3	1,970	-0,33	-0,251	0,06	0,004	0,005	1,85	0,31	Circulaire	96,5	0,021	0,04	1,26	0,2	0,4	1	5
Abondances	1977	log(x+1)	72,9	0,843	-0,11	0,096	0,05	-0,020	-0,037	2,5	0,39	Circulaire	99,9	0,005	0,06	1,69	0,2	0,4	1	4
Abondances	1978	log(x+1)	70,4	2,469	-0,59	-0,272	0,08	0,003	0,018	1,7	0,31	Circulaire	99,8	0,000	0,06	0,98	0,2	0,4	1	4
Abondances	1979	log(x+1)	66,0	1,706	-0,39	-0,070	0,09	-0,006	-0,036	2	0,4	Circulaire	98,9	0,032	0,05	1,23	0,2	0,4	1	4
Abondances	1970	log(x+1)	73,8	1,369	-0,15	-0,111	0,05	-0,021	0,003	2,5	0,37	Circulaire	99,8	0,012	0,05	1,92	0,2	0,4	1	5
Abondances	1971	log(x+1)	71,1	1,606	-0,34	-0,027	0,071	-0,012	-0,026	2,6	0,37	Circulaire	99,7	0,005	0,07	2,11	0,2	0,4	1	6
Abondances	1980	log(x+1)	63,7	1,539	0,32	0,004	0,06	0,022	0,0023	3,5	0,69	Circulaire	99,8	0,014	0,06	2,51	0,2	0,4	6	11
Abondances	1981	log(x+1)	64,2	1,822	-0,32	-0,206	0,05	-0,002	0,0105	3,5	0,67	Circulaire	100,0	0,006	0,06	2,07	0,2	0,4	3	7
Abondances	1982	log(x+1)	70,3	0,676	-0,14	0,197	0,71	-0,025	-0,0455	2,4	0,5	Circulaire	99,0	0,003	0,07	1,89	0,2	0,4	1	5
Abondances	1983	log(x+1)	66,5	1,357	-0,28	0,067	0,06	-0,024	-0,0075	2,8	0,55	Circulaire	100,0	0,009	0,06	1,93	0,2	0,4	3	10
Abondances	1984	log(x+1)	68,6	0,0649	-0,06	0,225	0,04	-0,034	-0,0300	2,7	0,55	Circulaire	99,7	0,012	0,06	1,98	0,2	0,4	2	8
Abondances	1985	log(x+1)	68,8	0,705	-0,15	0,260	0,06	-0,037	-0,0399	2,5	0,55	Circulaire	99,9	0,019	0,05	1,87	0,2	0,4	2	6
Abondances	1986	log(x+1)	68,8	0,642	-0,16	0,298	0,08	-0,038	-0,0514	2,5	0,52	Circulaire	99,8	0,015	0,05	1,83	0,2	0,4	5	11
Abondances	1987	log(x+1)	71,4	0,959	-0,28	0,217	0,09	-0,033	-0,0468	2,5	0,48	Circulaire	100,0	0,013	0,06	1,73	0,2	0,4	7	12
Abondances	1988	log(x+1)	63,4	0,983	-0,15	0,185	0,05	-0,029	-0,0252	2,5	0,5	Circulaire	98,5	0,032	0,03	1,76	0,2	0,4	3	11
Abondances	1989	log(x+1)	53,4	0,883	-0,07	0,197	0,04	-0,028	-0,0345	2,4	0,51	Sphérique	100,0	0,010	0,06	1,84	0,2	0,4	1	7
Abondances	1990	log(x+1)	60,0	1,493	-0,32	0,081	0,06	-0,029	-0,0073	2,7	0,53	Circulaire	99,6	0,010	0,08	1,91	0,2	0,4	5	9
Abondances	1991	log(x+1)	66,2	0,8	-0,16	0,229	0,06	-0,038	-0,0357	2,8	0,55	Circulaire	99,0	0,018	0,05	1,91	0,2	0,4	3	9
Abondances	1992	log(x+1)	57,8	0,969	-0,22	0,247	0,06	-0,037	-0,0258	3,3	0,65	Circulaire	100,0	0,015	0,07	2,32	0,2	0,4	2	9
Abondances	1993	log(x+1)	62,7	1,144	-0,25	0,141	0,07	-0,025	-0,0308	2,6	0,5	Circulaire	99,8	0,005	0,07	1,56	0,2	0,4	1	7
Abondances	1994	log(x+1)	57,0	1,004	-0,18	0,886	0,07	-0,017	-0,0305	2,5	0,45	Circulaire	99,3	0,014	0,08	1,91	0,2	0,4	6	11
Abondances	1995	log(x+1)	64,6	1,464	-0,26	0,038	0,06	-0,019	-0,0119	2,4	0,5	Circulaire	100,0	0,013	0,05	1,62	0,2	0,4	1	9
Abondances	1996	log(x+1)	41,4	0,658	-0,17	0,275	0,06	-0,033	-0,0361	2,6	0,48	Circulaire	99,7	0,013	0,07	1,60	0,2	0,4	5	8
Abondances	1997	log(x+1)	62,1	0,267	-0,06	0,379	0,06	-0,040	-0,0563	2,3	0,5	Circulaire	99,0	0,018	0,04	1,59	0,2	0,4	5	10
Abondances	1998	log(x+1)	54,0	0,935	-0,12	0,178	0,05	-0,027	-0,0239	2,8	0,51	Circulaire	99,7	0,018	0,07	1,98	0,2	0,4	1	9
Abondances	1999	log(x+1)	62,8	0,952	-0,18	0,222	0,06	-0,033	-0,0362	2,4	0,5	Circulaire	99,9	0,019	0,05	1,84	0,2	0,4	6	10
Abondances	2000	log(x+1)	58,4	1,146	-0,26	0,156	0,07	-0,026	-0,0257	2,3	0,5	Circulaire	99,7	0,014	0,06	1,60	0,2	0,4	6	11
Abondances	2001	log(x+1)	51,8	0,439	0,07	0,234	0,04	-0,025	-0,0494	2,7	0,5	Circulaire	99,8	0,009	0,09	1,98	0,2	0,4	1	9
Abondances	2002	log(x+1)	58,5	0,87	-0,07	0,212	0,05	-0,029	-0,0421	2,7	0,57	Circulaire	99,8	0,023	0,05	1,89	0,2	0,4	7	11
Abondances	2003	log(x+1)	47,4	1,081	-0,24	0,194	0,06	-0,050	-0,0214	3	0,58	Circulaire	99,9	0,013	0,09	1,93	0,2	0,4	7	11
Abondances	2004	log(x+1)	63,0	1,01	-0,14	0,215	0,05	-0,034	-0,0312	3	0,57	Exponentiel	100,0	0,005	0,07	0,79	0,2	0,4	6	10
Abondances	2005	log(x+1)	65,8	1,047	-0,23	0,140	0,07	-0,022	-0,0334	3,2	0,55	Circulaire	99,4	0,025	0,04	2,75	0,2	0,4	7	10
Abondances	2006	log(x+1)	64,2	0,892	-0,24	0,288	0,07	-0,040	-0,0303	3	0,57	Circulaire	99,3	0,018	0,05	2,09	0,2	0,4	6	11
Abondances	2007	log(x+1)	60,5	1,27	-0,28	0,135	0,07	-0,022	-0,0965	2,5	0,5	Circulaire	98,9	0,024	0,02	1,16	0,2	0,4	2	9

A.2.3 Abondances observées pour le merlan de 1980 à 2007

Données			**Tendance spatiale**							**Variogramme**							**Krigeage**			
Variable	Année	Trans	Tendance	a	b	c	d	e	f	Xmax	Pas	Modèle	ModVar	Pépite	Pallier	Portée	Maille	Rayon	Min	Max
Abondances	1980	log(x+1)	34,7	-0,52	1,02	0,47	-0,0580	-0,002	-0,213	2,9	0,6	Circulaire	98,1	0,42	0,37	2,24	0,2	0,4	1	5
Abondances	1981	log(x+1)	28,1	0,13	0,47	0,40	0,0253	0,024	-0,218	-	-	Pure pépite		0,70	0,00	2,00	0,2	0,4	1	4
Abondances	1982	log(x+1)	25,6	0,30	0,77	0,17	-0,0604	0,005	-0,120	2	0,4	Circulaire	100,0	0,36	0,28	0,91	0,2	0,4	1	4
Abondances	1983	log(x+1)	24,5	0,44	0,50	0,26	-0,0116	0,002	-0,130	-	-	Pure pépite		0,55	0,00	2,00	0,2	0,4	1	4
Abondances	1984	log(x+1)	31,2	0,48	0,68	0,18	-0,0079	0,006	-0,147	2,2	0,45	Exponentiel	99,4	0,11	0,47	0,43	0,2	0,4	1	5
Abondances	1985	log(x+1)	21,7	-0,35	1,12	0,44	-0,1196	-0,028	-0,114	2,1	0,43	Circulaire	100,0	0,42	0,31	1,00	0,2	0,4	1	5
Abondances	1986	log(x+1)	17,2	-	-	-	-	-	-	1	0,2	Circulaire	98,7	0,27	0,48	0,63	0,2	0,4	1	5
Abondances	1987	log(x+1)	19,7	-	-	-	-	-	-	2,7	0,44	Exponentiel	99,0	0,18	0,77	0,62	0,2	0,4	1	5
Abondances	1988	log(x+1)	21,4	-0,97	1,83	0,61	-0,0891	-0,025	-0,157	2,8	0,46	Circulaire	99,4	0,41	0,29	2,03	0,2	0,4	1	5
Abondances	1989	log(x+1)	19,7	-	-	-	-	-	-	2,8	0,44	Circulaire	99,3	0,41	0,51	2,17	0,2	0,4	1	5
Abondances	1990	log(x+1)	21,4	0,58	0,66	0,51	-0,0544	-0,031	-0,122	1,7	0,35	Circulaire	99,7	0,44	0,22	1,21	0,2	0,4	1	5
Abondances	1991	log(x+1)	12,3	-	-	-	-	-	-	3,3	0,39	Circulaire	98,3	0,49	0,53	2,68	0,2	0,4	1	5
Abondances	1992	log(x+1)	18,7	-	-	-	-	-	-	2	0,37	Circulaire	91,8	0,48	0,42	1,12	0,2	0,4	1	5
Abondances	1993	log(x+1)	31,9	0,06	0,76	0,38	-0,0740	-0,007	-0,111	1,5	0,13	Circulaire	99,4	0,30	0,24	0,95	0,2	0,4	1	5
Abondances	1994	log(x+1)	33,5	2,84	0,15	-0,24	-0,0286	0,030	-0,057	1,3	0,27	Circulaire	99,7	0,19	0,40	1,03	0,2	0,4	1	5
Abondances	1995	log(x+1)	26,9	2,30	0,04	-0,08	0,0084	0,019	-0,078	1,5	0,29	Circulaire	99,7	0,29	0,36	1,17	0,2	0,4	1	5
Abondances	1996	log(x+1)	31,0	1,30	0,23	0,26	-0,0497	-0,013	-0,059	1,7	0,28	Circulaire	98,8	0,31	0,32	1,28	0,2	0,4	1	5
Abondances	1997	log(x+1)	31,7	-2,12	-0,17	-0,05	-0,0003	0,012	-0,037	1,4	0,27	Circulaire	96,6	0,29	0,32	1,07	0,2	0,4	1	5
Abondances	1998	log(x+1)	35,0	2,74	-0,38	-0,07	-	-	-	2,5	0,41	Pentasphérique	99,7	0,15	0,36	1,25	0,2	0,4	1	5
Abondances	1999	log(x+1)	32,3	1,90	0,04	-0,01	-0,0311	0,004	-0,035	1,6	0,28	Pentasphérique	96,2	0,13	0,42	1,38	0,2	0,4	1	5
Abondances	2000	log(x+1)	35,3	3,06	-0,39	-0,08	-	-	-	2,4	0,4	Circulaire	81,8	0,20	0,29	1,08	0,2	0,4	1	5
Abondances	2001	log(x+1)	20,4	2,11	0,20	0,06	-0,0547	-0,008	-0,028	1,5	0,25	Circulaire	98,6	0,41	0,30	1,26	0,2	0,4	1	5
Abondances	2002	log(x+1)	19,6	-	-	-	-	-	-	2,5	0,39	Circulaire	99,1	0,41	0,37	1,52	0,2	0,4	1	5
Abondances	2003	log(x+1)	19,6	-	-	-	-	-	-	2,5	0,43	Circulaire	99,8	0,40	0,33	1,89	0,2	0,4	1	5
Abondances	2004	log(x+1)	22,6	1,04	0,27	0,26	-0,0602	-0,023	-0,029	2,2	0,33	Sphérique	98,6	0,24	0,26	1,63	0,2	0,4	1	5
Abondances	2005	log(x+1)	26,1	2,70	-0,22	-0,19	-0,0155	0,007	0,010	2	0,33	Circulaire	99,8	0,26	0,13	1,70	0,2	0,4	1	5
Abondances	2006	log(x+1)	24,7	1,39	-0,13	0,16	-0,0144	-0,019	-0,004	1,7	0,29	Circulaire	94,7	0,25	0,12	1,37	0,2	0,4	1	5
Abondances	2007	log(x+1)	23,4	1,96	-0,07	-0,07	-0,0279	0,002	0,002	2,2	0,45	Circulaire	99,4	0,22	0,19	1,37	0,2	0,4	1	4

A.2.4 Abondances totales et les trois classes d’âge de l’anchois de 2000 à 2007

Données			Tendance spatiale							Variogramme							Krigeage			
Variable	Année	Trans	Tendance	a	b	c	d	e	f	Xmax	Pas	Modèle	ModVar	Pépite	Pallier	Portée	Maille	Rayon	Min	Max
Abondance	2000	log(x+1)	57,6	-3,316	0,97	1,533	0,45	-0,144	-0,47	0,70	0,110	Exponentiel	99,7	0,014	0,69	0,12	0,2	0,2	3	7
Abondance	2001	log(x+1)	23,9	-6,440	3,35	0,666	-0,27	0,183	-0,42	1,00	0,190	Exponentiel	99,2	0,407	0,74	0,22	0,2	0,2	3	7
Abondance	2002	log(x+1)	35,0	3,730	-1,38	-1,743	0,32	0,212	0,21	0,80	0,160	Exponentiel	99,5	0,099	0,60	0,15	0,2	0,2	3	7
Abondance	2003	log(x+1)	15,3	-	-	-	-	-	-	0,80	0,150	Exponentiel	99,8	0,249	0,61	0,16	0,2	0,2	2	8
Abondance	2004	log(x+1)	32,5	2,261	-2,00	-0,771	0,53	0,101	0,16	0,50	0,093	Sphérique	97,8	0,178	0,24	0,15	0,2	0,2	3	9
Abondance	2005	log(x+1)	22,6	1,322	-0,68	-0,825	0,27	0,146	0,06			Pure pépite		0,500	0,00	0,08	0,2	0,2	3	9
Abondance	2006	log(x+1)	37,0	1,401	-2,88	-0,082	1,19	-0,036	0,24	1,00	0,159	Pentasphérique	99,6	0,232	0,26	0,45	0,2	0,2	3	9
Abondance	2007	log(x+1)	46,2	2,246	-2,39	-1,057	0,97	0,144	0,31	0,50	0,105	Circulaire	99,8	0,229	0,34	0,35	0,2	0,2	2	8
Age 1	2000	log(x+1)	50,7	-5,470	2,29	2,783	0,34	-0,321	-0,80	0,90	0,160	Exponentiel	99,7	0,167	0,72	0,19	0,2	0,2	3	7
Age 2	2000	log(x+1)	53,5	-0,216	-0,21	0,335	0,45	-0,060	-0,13	0,70	0,140	Exponentiel	99,2	0,000	0,42	0,09	0,2	0,2	3	7
Age 3	2000	log(x+1)	46,7	1,227	-0,84	-0,378	0,39	0,022	0,07	0,70	0,145	Circulaire	96,7	0,108	0,19	0,25	0,2	0,2	3	7
Age 1	2001	log(x+1)	8,7	-	-	-	-	-	-	1,00	0,200	Pentasphérique	99,9	0,640	0,89	0,75	0,2	0,2	3	7
Age 2	2001	log(x+1)	12,9	-	-	-	-	-	-	1,00	0,200	Pentasphérique	99,2	0,442	0,53	0,64	0,2	0,2	3	7
Age 3	2001	log(x+1)	6,7	-	-	-	-	-	-	1,00	0,200	Pentasphérique	98,8	0,154	0,15	0,57	0,2	0,2	3	7
Age 1	2002	log(x+1)	26,5	2,520	-1,80	-1,539	0,81	0,265	0,25	0,90	0,142	Pentasphérique	99,8	0,279	0,55	0,68	0,2	0,2	3	7
Age 2	2002	log(x+1)	17,1	-	-	-	-	-	-	0,90	0,174	Exponentiel	99,5	0,192	0,22	1,20	0,2	0,2	3	7
Age 3	2002	log(x+1)	14,5	-	-	-	-	-	-	0,96	0,177	Exponentiel	97,6	0,105	0,22	0,75	0,2	0,2	3	7
Age 1	2003	log(x+1)	3,6	-	-	-	-	-	-	0,70	0,145	Exponentiel	99,1	0,357	0,23	0,72	0,2	0,2	2	8
Age 2	2003	log(x+1)	5,8	-	-	-	-	-	-	0,80	0,150	Exponentiel	99,0	0,206	0,18	0,29	0,2	0,2	2	8
Age 3	2003	log(x+1)	7,1	-	-	-	-	-	-	0,60	0,120	Exponentiel	98,9	0,105	0,10	0,17	0,2	0,2	2	8
Age 1	2004	log(x+1)	30,9	4,482	-6,33	-2,214	2,83	0,317	1,10	0,60	0,110	Circulaire	91,6	0,522	0,26	0,34	0,2	0,2	3	9
Age 2	2004	log(x+1)	17,6	-	-	-	-	-	-	0,50	0,100	Circulaire	91,2	0,263	0,22	0,14	0,2	0,2	3	9
Age 3	2004	log(x+1)	14,9	-	-	-	-	-	-	0,50	0,099	Circulaire	71,2	0,185	0,22	0,09	0,2	0,2	3	9
Age 1	2005	log(x+1)	11,1	-	-	-	-	-	-	0,80	0,159	Exponentiel	95,5	0,105	0,13	0,24	0,2	0,2	3	9
Age 2	2005	log(x+1)	13,7	-	-	-	-	-	-	0,65	0,125	Circulaire	96,4	0,226	0,17	0,34	0,2	0,2	3	9
Age 3	2005	log(x+1)	14,0	-	-	-	-	-	-	0,65	0,130	Pentasphérique	76,5	0,001	0,16	0,29	0,2	0,2	3	9
Age 1	2006	log(x+1)	45,7	2,913	-6,39	-1,366	3,71	0,142	1,34	0,62	0,110	Pentasphérique	98,8	0,220	0,45	0,38	0,2	0,2	3	9
Age 2	2006	log(x+1)	29,2	1,696	-4,13	-0,367	2,51	-0,046	0,74	0,80	0,170	Exponentiel	99,9	0,191	0,13	0,38	0,2	0,2	3	9
Age 3	2006	log(x+1)	21,2	0,794	-2,35	0,026	1,58	-0,066	0,30	0,85	0,160	Exponentiel	99,4	0,156	0,15	0,20	0,2	0,2	3	9
Age 1	2007	log(x+1)	35,1	1,750	-2,40	-0,648	1,20	0,073	0,32	0,50	0,100	Circulaire	99,9	0,276	0,34	0,50	0,2	0,2	2	8
Age 2	2007	log(x+1)	30,5	1,440	-1,39	-0,344	0,72	0,001	0,17	0,50	0,101	Circulaire	99,4	0,253	0,34	0,39	0,2	0,2	2	8
Age 3	2007	log(x+1)	27,3	1,552	-1,91	-0,395	0,76	0,019	0,23	0,50	0,100	Circulaire	98,6	0,100	0,32	0,15	0,2	0,2	2	8

A.2.5 Facteurs environnementaux de surface et de fond du Golfe de Gascogne entre 2000 et 2007

Données			Tendance spatiale							Variogramme							Krigeage			
Variable	Année	Trans	Tendance	a	b	c	d	e	f	Xmax	Pas	Modèle	ModVar	Pépite	Pallier	Portée	Maille	Rayon	Min	Max
Temp Surf	2000	-	32,5	8,230	2,80	2,351	-0,37	-0,218	-0,86	0,90	0,211	Exponentiel	99,6	0,099	0,26	0,14	0,2	0,4	2	8
Temp Surf	2001	-	56,7	29,190	-5,09	-5,500	0,35	0,184	1,50	1,20	0,280	Circulaire	99,5	0,843	0,98	0,73	0,2	0,4	1	4
Temp Surf	2002	-	24,6	10,740	2,25	1,300	-0,31	-0,128	-0,42	1,40	0,273	Circulaire	95,4	0,053	0,09	0,78	0,2	0,4	2	6
Temp Surf	2003	-	14,6	-	-	-	-	-	-	2,00	0,380	Sphérique	95,2	0,668	0,25	1,37	0,2	0,4	2	6
Temp Surf	2004	-	52,1	7,360	3,69	1,710	-0,44	0,099	-0,92	0,60	0,110	Circulaire	97,5	0,000	0,30	0,30	0,2	0,4	2	6
Temp Surf	2005	-	54,2	11,770	3,46	0,444	-0,80	0,082	-0,71	0,60	0,153	Sphérique	88,1	0,000	0,17	0,14	0,2	0,4	2	6
Temp Surf	2006	-	24,8	10,380	4,37	2,160	-1,41	-0,345	-0,56	2,00	0,441	Exponentiel	98,6	0,000	0,41	0,23	0,2	0,4	2	6
Temp Surf	2007	-	63,2	11,100	5,19	1,940	-1,42	-0,433	-0,65	1,20	0,254	Exponentiel	87,5	0,166	0,35	0,28	0,2	0,4	2	6
Temp Fond	2000	-	68,3	12,454	-0,10	-0,013	0,01	-0,069	-0,01			Pure pépite		0,053	0,00	4,00	0,2	0,4	2	8
Temp Fond	2001	-	71,9	7,340	1,87	1,689	-0,13	-0,169	-0,35	1,50	0,280	Circulaire	96,9	0,027	0,02	0,76	0,2	0,4	1	4
Temp Fond	2002	-	38,8	14,700	-1,84	-2,040	0,48	0,320	0,32	2,35	0,423	Exponentiel	69,6	0,187	0,23	0,47	0,2	0,4	2	6
Temp Fond	2003	-	43,5	12,150	-0,45	-0,157	0,37	0,056	-0,13	0,65	0,149	Exponentiel	80,8	0,000	0,16	0,05	0,2	0,4	2	6
Temp Fond	2004	-	37,6	15,190	-2,19	-1,173	0,42	0,056	0,36			Pure pépite		0,150	0,00	4,00	0,2	0,4	2	6
Temp Fond	2005	-	58,5	13,470	-0,65	-0,665	0,04	0,025	-0,09	0,75	0,139	Sphérique	97,3	0,000	0,26	0,53	0,2	0,4	2	6
Temp Fond	2006	-	47,5	13,665	1,35	1,125	0,23	0,091	0,34	0,70	0,144	Circulaire	94,9	0,000	0,55	0,25	0,2	0,4	2	6
Temp Fond	2007	-	33,8	7,070	3,43	1,886	-0,51	-0,172	-0,63	0,90	0,150	Circulaire	94,6	0,114	0,14	0,70	0,2	0,4	2	6
Sal Surf	2000	-	73,1	34,040	2,11	1,737	-0,90	-0,436	-0,45	3,30	0,700	Circulaire	84,7	0,159	0,13	2,21	0,2	0,4	2	8
Sal Surf	2001	-	49,6	36,440	0,04	-0,220	-0,14	0,038	-0,20	0,90	0,178	Sphérique	97,3	0,000	0,53	0,44	0,2	0,4	1	4
Sal Surf	2002	-	80,3	35,632	0,49	0,189	-0,26	-0,058	-0,20	1,10	0,220	Sphérique	98,1	0,000	0,07	0,77	0,2	0,4	2	6
Sal Surf	2003	-	69,0	36,950	-0,49	-0,415	-0,18	0,005	0,02	1,10	0,210	Circulaire	98,6	0,034	0,11	0,61	0,2	0,4	2	6
Sal Surf	2004	-	-	-	-	-	-	-	-	2,20	0,450	Circulaire	98,0	0,000	1,01	1,62	0,2	0,4	2	6
Sal Surf	2005	-	64,9	33,000	2,03	2,020	-0,68	-0,386	-0,59	1,10	0,204	Circulaire	98,7	0,149	0,18	0,81	0,2	0,4	2	6
Sal Surf	2006	-	-	-	-	-	-	-	-	2,10	0,400	Sphérique	99,8	0,000	1,06	1,62	0,2	0,4	2	6
Sal Surf	2007	-	33,5	36,070	-0,86	-0,154	-	-	-	0,60	0,115	Circulaire	67,2	0,000	0,36	0,21	0,2	0,4	2	6
Sal Fond	2000	-	84,3	32,143	3,21	1,904	-0,79	-0,280	-0,74	2,00	0,439	Exponentiel	74,8	0,000	0,04	0,15	0,2	0,4	2	8
Sal Fond	2001	-	70,2	34,195	1,05	0,548	-0,20	-0,065	-0,17	5,70	0,131	Circulaire	98,0	0,000	0,02	0,45	0,2	0,4	1	4
Sal Fond	2002	-	80,6	32,872	1,87	1,402	-0,34	-0,182	-0,46	1,70	0,318	Exponentiel	96,2	0,010	0,00	0,42	0,2	0,4	2	6
Sal Fond	2003	-	86,3	34,819	0,82	0,452	-0,23	-0,065	-0,21	0,70	0,134	Circulaire	96,3	0,000	0,01	0,59	0,2	0,4	2	6
Sal Fond	2004	-	75,1	34,593	1,44	0,396	-0,44	-0,047	-0,21	0,70	0,132	Circulaire	95,1	0,000	0,04	0,29	0,2	0,4	2	6
Sal Fond	2005	-	90,2	34,289	1,43	0,678	-0,40	-0,087	-0,34	0,75	0,139	Sphérique	93,9	0,000	0,01	0,58	0,2	0,4	2	6
Sal Fond	2006	-	62,0	33,570	1,77	1,215	-0,38	-0,170	-0,51	0,70	0,149	Pentasphérique	74,6	0,000	0,01	0,31	0,2	0,4	2	6
Sal Fond	2007	-	63,6	32,577	2,43	1,657	-0,49	-0,215	-0,33	0,60	0,120	Pentasphérique	92,2	0,000	0,02	0,50	0,2	0,4	2	6

Temp Surf : température de surface, Temp Fond : température de fond, Sal Surf : salinité de surface, Sal Fond : salinité de fond.

A.2.6 Autres facteurs environnementaux du Golfe de Gascogne entre 2000 et 2007

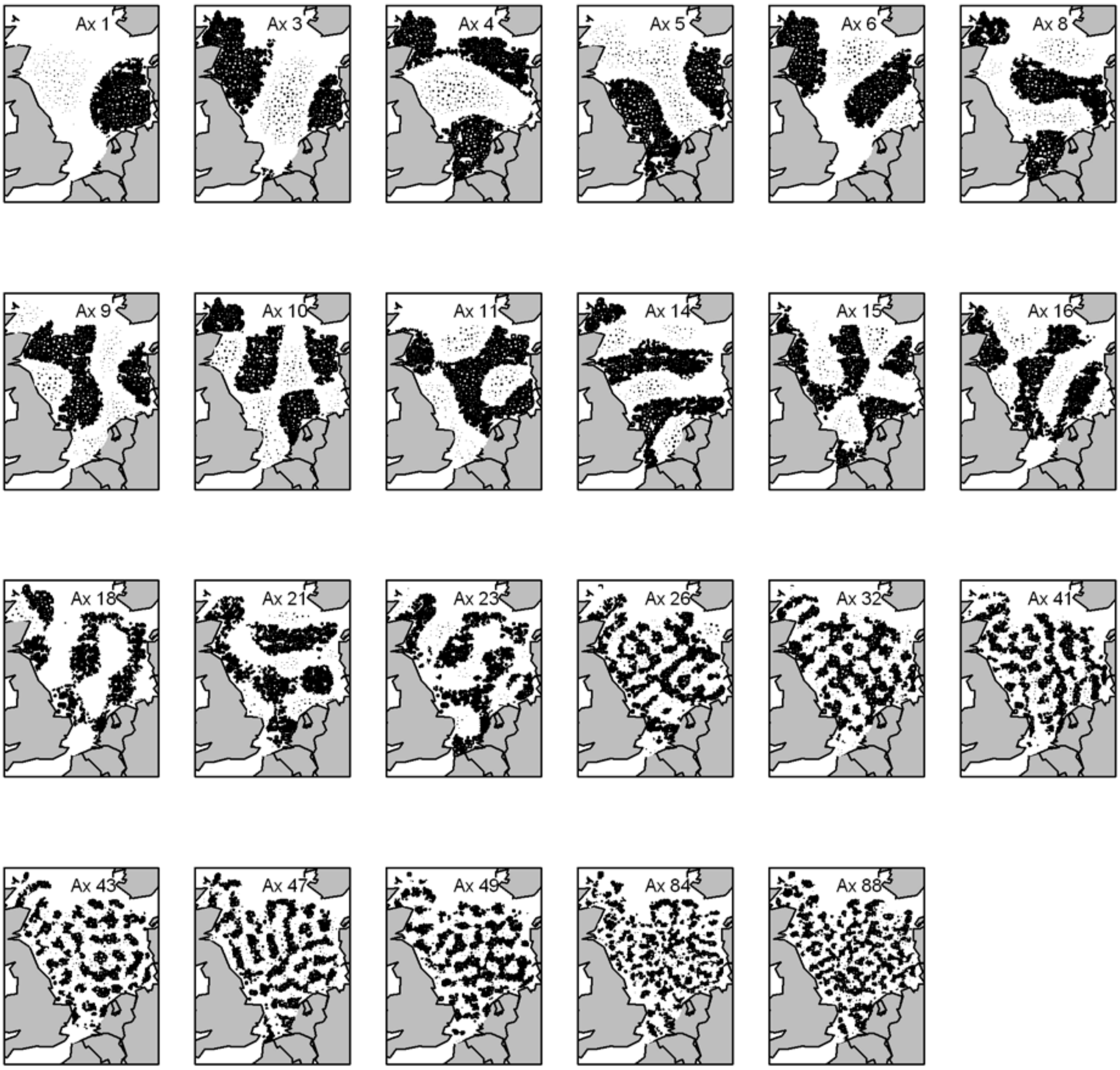
Données			**Tendance spatiale**							**Variogramme**							**Krigeage**			
Variable	Année	Trans	Tendance	a	b	c	d	e	f	Xmax	Pas	Modèle	ModVar	Pépite	Pallier	Portée	Maille	Rayon	Min	Max
MLD	2000	-	48,2	27,40	31,50	7,00	-13,390	-2,92	4,14	0,9	0,168	Pentasphérique	98,3	0,000	51,44	0,14	0,2	0,4	1	8
MLD	2001	-	63,6	126,30	-66,90	-27,90	10,350	2,32	6,22	1,0	0,192	Pentasphérique	72,5	0,000	71,85	0,24	0,2	0,4	1	4
MLD	2002	-	70,7	145,50	-5,75	-28,10	2,920	-0,19	9,63	1,2	0,246	Circulaire	98,2	0,000	55,82	0,82	0,2	0,4	1	4
MLD	2003	-	71,9	111,40	-45,90	-27,73	3,270	1,09	8,43	2,0	0,371	Circulaire	99,1	15,171	16,08	1,08	0,2	0,4	1	5
MLD	2004	-	58,8	75,10	-1,20	-8,70	-8,700	-2,52	3,26	2,0	0,410	Circulaire	97,4	4,310	88,73	1,42	0,2	0,4	1	5
MLD	2005	-	89,5	42,00	2,40	13,50	-6,670	-4,17	-2,58	2,0	0,380	Exponentiel	98,1	0,000	15,08	0,18	0,2	0,4	1	5
MLD	2006	-	73,6	51,85	-17,67	-6,42	-	-	-	0,9	0,919	Pentasphérique	82,6	0,000	17,93	0,38	0,2	0,4	1	5
MLD	2007	-	67,0	75,60	-56,60	-10,50	11,090	0,23	5,53	1,2	0,235	Circulaire	99,2	0,000	38,80	0,93	0,2	0,4	1	5
DiffTemp	2000	-	-	-	-	-	-	-	-	0,6	0,113	Circulaire	97,0	0,128	0,22	0,36	0,2	0,4	2	8
DiffTemp	2001	-	46,7	23,13	-7,37	-7,68	0,491	0,39	1,95	2,1	0,397	Circulaire	99,2	0,643	1,57	1,25	0,2	0,4	1	4
DiffTemp	2002	-	21,3	-40,80	4,14	3,39	-0,791	-0,45	-0,75	2,1	0,422	Circulaire	99,9	0,290	0,27	1,15	0,2	0,4	2	6
DiffTemp	2003	-	28,9	-3,89	7,26	4,86	-1,648	-0,66	-1,20	2,2	0,428	Circulaire	98,4	0,557	0,48	1,64	0,2	0,4	2	6
DiffTemp	2004	-	70,2	-6,33	4,99	2,16	-0,740	0,13	-1,06	0,9	0,151	Pentasphérique	96,9	0,012	0,36	0,80	0,2	0,4	2	6
DiffTemp	2005	-	26,7	-1,70	4,11	1,11	-0,837	0,06	-0,62	0,5	0,101	Circulaire	87,9	0,000	0,42	0,20	0,2	0,4	2	6
DiffTemp	2006	-	35,1	-3,28	5,72	3,28	-1,641	-0,44	-0,90	2,2	0,474	Pentasphérique	93,4	0,000	0,54	0,85	0,2	0,4	2	6
DiffTemp	2007	-	61,4	3,31	2,18	0,61	-0,974	-0,35	-0,13	1,3	0,250	Circulaire	91,7	0,285	0,31	1,04	0,2	0,4	2	6
DiffSal	2000	-	-	-	-	-	-	-	-	2,2	0,367	Pentasphérique	83,7	0,000	0,47	1,88	0,2	0,4	2	8
DiffSal	2001	-	34,7	-1,89	0,88	0,63	-0,050	-0,09	0,06	1,8	0,346	Circulaire	99,3	0,085	0,58	0,98	0,2	0,4	1	4
DiffSal	2002	-	49,1	-2,74	1,37	1,21	-0,083	-0,13	-0,26	1,0	0,161	Circulaire	81,8	0,024	0,07	0,80	0,2	0,4	2	6
DiffSal	2003	-	49,1	-1,82	1,07	0,77	0,006	-0,06	-0,19	1,1	0,216	Circulaire	96,0	0,390	0,10	0,55	0,2	0,4	2	6
DiffSal	2004	-	-	-	-	-	-	-	-	2,4	0,443	Circulaire	94,5	0,084	0,42	0,16	0,2	0,4	2	6
DiffSal	2005	-	45,0	1,29	-0,61	1,33	0,287	0,30	0,25	1,1	0,205	Circulaire	96,6	0,131	0,20	0,81	0,2	0,4	2	6
DiffSal	2006	-	-	-	-	-	-	-	-	1,2	0,285	Circulaire	92,9	0,000	0,87	1,18	0,2	0,4	2	6
DiffSal	2007	-	24,0	-4,87	4,42	2,51	-0,721	-0,30	-0,90	0,8	0,142	Circulaire	97,1	0,143	0,32	0,67	0,2	0,4	2	6
NRJpot	2000	-	45,0	331,60	247,00	146,60	-37,470	-10,40	-59,70	1,0	0,150	Sphérique	96,3	115,000	330,60	0,79	0,2	0,4	1	7
NRJpot	2001	-	32,4	-484,00	445,00	237,80	-89,000	-27,40	-69,10	1,2	0,220	Circulaire	99,6	0,000	2196,00	0,65	0,2	0,4	1	4
NRJpot	2002	-	61,8	-182,60	-199,10	129,00	-45,130	-19,15	-33,62	1,3	0,246	Circulaire	93,4	136,700	112,20	0,73	0,2	0,4	1	6
NRJpot	2003	-	53,8	69,00	125,20	71,20	-47,700	17,77	-20,80	2,0	0,350	Exponentiel	98,1	183,000	955,00	0,37	0,2	0,4	1	6
NRJpot	2004	-	50,4	-191,90	157,60	90,80	-24,600	-30,80	-38,70	1,3	0,267	Sphérique	97,8	225,000	161,00	0,75	0,2	0,4	1	6
NRJpot	2005	-	-	-	-	-	-	-	-	0,9	0,133	Exponentiel	73,4	0,000	919,10	0,21	0,2	0,4	1	6
NRJpot	2006	-	28,7	-152,10	239,20	117,20	-70,700	-15,55	-45,90	1,0	0,216	Circulaire	99,0	286,900	393,40	0,84	0,2	0,4	1	5
NRJpot	2007	-	40,4	-242,00	284,80	181,60	-73,700	-28,63	-54,90	0,9	0,191	Circulaire	95,2	50,900	653,60	0,64	0,2	0,4	1	5
Profondeur	-	-	89,7	424,40	-101,80	-52,20	-10,290	-6,22	15,02	1,5	0,329	Exponentiel	99,8	0,000	208,87	0,24	0,2	0,4	3	7

MLD : profondeur de la couche de mélange, DiffTemp : différence de température entre la surface et le fond, DiffSal : différence de salinité entre la surface et le fond, NRJ pot : déficit d’énergie potentielle.

Annexe B

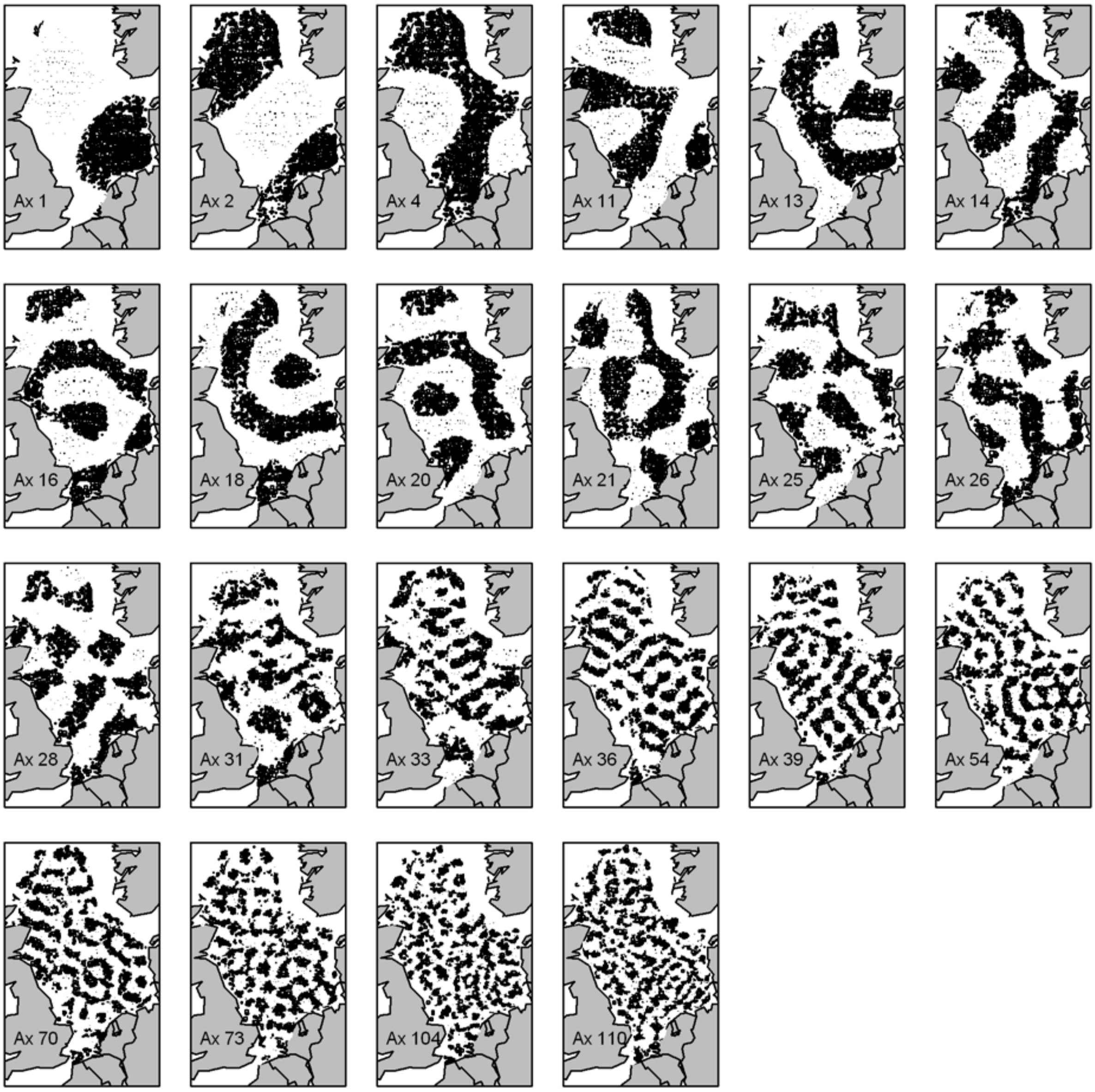
Cartographie des coordonnées principales de matrices de voisinage (PCNM)

B.1 La plie de Mer du Nord



Grande échelle : Ax 1 à 8; Moyenne échelle : Ax 9 à 23; Petite échelle : Ax 26 à 88.

B.2 Le merlan de Mer du Nord



Grande échelle : Ax 1 à 4; Moyenne échelle : Ax 11 à 28; Petite échelle : Ax 31 à 110.

B.3 L’anchois du Golfe de Gascogne

Grande échelle : Ax 1 et 3; Moyenne échelle : Ax 6 à 18; Petite échelle : Ax 30 à 69.

Annexe C

Liste des communications

C.1 Articles

Loots C., Vaz S., Planque B., Koubbi P. What control the spatial distribution of North Sea plaice spawning population. Confronting ecological hypotheses through a model selection framework. *ICES Journal of Marine Science*. Sous presse.

Loots C., Planque B., Vaz S., Koubbi P. Identifying processes that control the spawning distribution of North Sea whiting using multi model inference. *Fisheries Oceanography*. Soumis.

Loots C., Vaz S., Koubbi P., Planque B., Coppin, F., Verin, Y. Inter-annual variability of North Sea plaice potential spawning habitat. *Journal of Sea Research*. Soumis.

Planque B., **Loots C.**, Petitgas P., Lindstrom, U., Vaz, S. Modelling the spatial distribution of fish populations using a statistical multi-model approach.. *Fisheries Oceanography*. Soumis.

C.2 Conférences

Loots, C. Vaz, S., Koubbi, P., Planque, B., Coppin, F., Verin, Y., 2008 Inter-annual variability of North Sea plaice potential spawning habitat. In: 7th International Flatfish Symposium, Sesimbra, Portugal. (Communication orale).

Loots C., Vaz, S., Koubbi, P., Planque, B., Coppin, F., Verin, Y., 2007. Spawning habitat of North Sea Plaice. From non-spatial to spatial modelling. ICES Annual Science Conference, Helsinki, Finland. (Poster).

C.3 Deliverables pour le projet Reclaim

D31 on cross-mapping & habitat statistical models

D43 on the effects of climate change on the distribution, migration, growth and reproduction on fish and shellfish stocks