
Mémoire de fin d'études
Pour l'obtention du « Diplôme d'Agronomie Approfondie » (DAA) Spécialisation
Sciences Halieutiques et Aquacoles

**Modélisation statistique bayésienne d'un modèle
de production de biomasse. Application à la
pêche de poulpe (*Octopus vulgaris*) de
Mauritanie**

Présenté par : ONO Kotaro

Soutenu le : 14/09/07





Mémoire de fin d'études
Pour l'obtention du « Diplôme d'Agronomie Approfondie » (DAA) Spécialisation
Sciences Halieutiques et Aquacoles

**Modélisation statistique bayésienne d'un modèle
de production de biomasse. Application à la
pêcherie de poulpe (*Octopus vulgaris*) de
Mauritanie**

Présenté par : ONO Kotaro

Soutenu le : 14/09/07

Devant le Jury

M. François HUSSON, AGROCAMPUS RENNES.
M. Richard SABATIE, AGROCAMPUS RENNES
M. Didier GASCUEL, AGROCAMPUS RENNES
M. Etienne RIVOT, AGROCAMPUS RENNES

Remerciements

Je tiens à adresser en premier mes plus sincères remerciements à mon maître de stage M. Etienne RIVOT qui, je tiens à le préciser, a été un maître de stage formidable. Non seulement je voudrai le remercier pour sa grande disponibilité et ses nombreux aides et conseils toujours très clairs qu'il a pu me donner au cours de ce stage mais aussi pour sa grande gentillesse. Ca a été un vrai moment de plaisir de pouvoir travailler avec lui pendant ces 6 mois. De plus, ce stage m'a permis d'acquérir de nouvelles compétences statistiques (surtout le Bayésien) et Etienne m'a aussi appris à être plus rigoureux dans mon travail. Encore merci Etienne !

Je tiens aussi à remercier Sylvain Bonhommeau qui m'a apporté de très précieuse aide pour les calculs d'indices d'upwelling. Je le remercie aussi pour toutes les fois où il a pu m'emmener faire du squash entre midi et deux et accepter de jouer contre moi malgré mon faible niveau de jeu... cela m'a bien permis de changer les idées au cours des dures journées de travail.

Merci aussi à M. Bambaye Hamady de l'IMROP pour m'avoir fourni les données concernant les indices d'upwelling de la Mauritanie et M. Abdelmalek Faraj de l'INRH pour ceux du Maroc.

Un grand merci aussi à toute l'équipe du Laboratoire d'Ecologie Halieutique d'Agrocampus Rennes sans qui je n'aurai pas pu passer un aussi agréable stage. Les grands épisodes sportifs du jeudi midi, les apéros du vendredi soir, ou les moments de discussions cinématographiques pendant le repas au Sésame m'ont grandement aidé à profiter de tout mon temps pendant mon stage au laboratoire : merci Hervé, Olivier, Richard, Cathy, Catherine, Marie L., Marie H., et Guy!!! Merci aussi à Béatrice pour ses aides bibliographiques.

Enfin, je remercie aussi très vivement mes deux amies d'école : Anne et Enora qui ont été toujours là pour venir à mon aide, me décontracter pendant des moments de stress, et me faire retrouver mon sourire !!! Un BIG thank you en plus à Anne qui a été aussi ma voisine de bureau pendant tout le stage et qui a réussi à supporter mes « tremblements de jambes » sans craquer et pour les moments de grands délires qu'on a pu avoir devant « Zoo Keeper », et les « Têtes à claques » !!!

Je n'oublierai jamais ces 6 mois que j'ai passé dans ce labo où j'ai vraiment appris énormément de choses. Et j'espère que je vous reverrai d'une manière ou d'une autre, pendant ou après mes études aux USA !

Sommaire

INTRODUCTION.....	4
1.1 CONTEXTE GENERAL - PROJET ISTAM	4
1.2 LES PECHERIES DE POULPE DANS LA ZONE COPACE – CAS DE LA MAURITANIE.....	4
1.3 METHODES CLASSIQUES D’EVALUATION DES STOCKS DE POULPE.....	5
1.4 PROBLEMES METHODOLOGIQUES SOULEVES.....	6
1.5 LES OBJECTIFS DU TRAVAIL	7
1.6 PLAN DU MEMOIRE	8
2 MATERIEL ET METHODES.....	8
2.1 DONNEES.....	8
2.2 MODELE DE BIOMASSE – <i>STATE-SPACE MODEL (SSM)</i>	10
2.3 HYPOTHESE D’EQUILIBRE	12
2.4 AJUSTEMENT DANS LE CADRE BAYESIEN.....	12
2.5 EVALUATION DE LA METHODE PAR UNE APPROCHE DE SIMULATION.....	14
2.6 EVALUATION DU STOCK DE POULPE MAURITANIE - MODELE DE BIOMASSE AVEC EFFET DE L’ENVIRONNEMENT.....	19
3 RESULTATS	22
3.1 EVALUATION DE LA METHODE PAR UNE APPROCHE DE SIMULATION.....	22
3.2 EVALUATION DU STOCK DE POULPE MAURITANIE	31
4 DISCUSSION ET CONCLUSION	39
4.1 INTERETS ET LIMITES DU CADRE DE MODELISATION <i>SSM</i> BAYESIEN	39
4.2 APPLICATION AU CAS DE LA MAURITANIE	39
4.3 PERSPECTIVES	40
5 BIBLIOGRAPHIE	42
6 ANNEXES.....	45
ANNEXE 1 - CAPTURES ET D’INDICE D’ABONDANCE (IA) DE LA PECHERIE DE POULPES MAURITANIENNES.	46
ANNEXE 2 - DONNEES D’INDICE D’UPWELLING COTIER CALCULE A PARTIR DE L’EQUATION DE TRANSPORT D’EKMAN.....	47
ANNEXE 5 - DIC ET FACTEURS DE BAYES	52
ANNEXE 6 - RESUME DES AJUSTEMENTS DYNAMIQUES ET A L’EQUILIBRE DES 13 SCENARII, TRAVAIL AVEC LES IC BAYESIENS	54
ANNEXE 7 - CODE WINBUGS	55

Introduction

1.1 Contexte général - Projet ISTAM

Les côtes Nord-Ouest africaines (Fig. 1) sont d'une grande richesse halieutique et suscitent un intérêt croissant des pêcheurs à la fois locaux mais aussi étrangers (FAO 2006 ; Kane 2008).

En Mauritanie par exemple, dès le début du vingtième siècle, des chalutiers de Boulogne et d'Arcachon venaient y exploiter les ressources (Beauté et le Gall 1936 dans Inejih 2002). Ces exploitations n'ont ensuite cessé de s'intensifier après la seconde guerre mondiale. Aujourd'hui, les ressources côtières de bons nombres de ces pays sont en état de pleine ou de sur-exploitation (COPACE 2002), et il est donc important d'améliorer leur gestion.

Ce travail s'inscrit dans le cadre de l'action concertée ISTAM (Improve Scientific and Technical Advice for fisheries Management – <http://projet-istam.org/>). L'action ISTAM a pour but de coordonner les activités scientifiques et techniques afin de renforcer les systèmes d'information sur les pêches, et d'améliorer les méthodes d'évaluation pour viser à une gestion plus adaptée de la ressource au sein de la zone COPACE. Ce travail fait partie du WP3 (*Working Package 3*) qui porte sur le développement de méthodes d'évaluation de stock. Plus particulièrement, cette étude s'inscrit dans les activités « Modèles d'évaluation dans un environnement *data-poor* » et « Modèles avec effet de l'environnement » qui ont notamment les pêcheries de poulpe (*Octopus vulgaris*) comme cas d'étude.

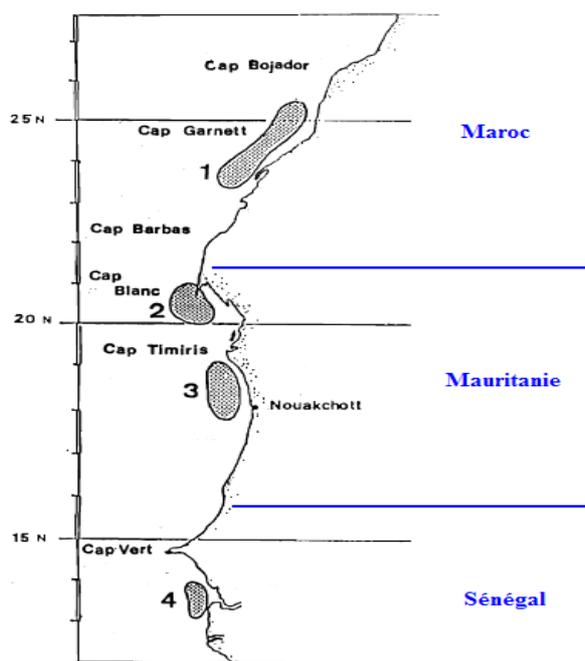


Figure 1 : Les principaux stocks de poulpes *Octopus vulgaris* au large du Maroc / Mauritanie / Sénégal. (source : FAO 1979b)

1.2 Les pêcheries de poulpe dans la zone COPACE – cas de la Mauritanie

Le poulpe commun (*Octopus vulgaris*) est une espèce benthique à cycle de vie court (environ un an et demi) (Inejih 2000). Il fréquente des habitats très variables (roches, récifs coralliens, ou herbiers) et colonisent des milieux allant de 0 à 200 m (FAO 1984). Comme la plupart des espèces à vie courte, le poulpe est également sensible aux fluctuations environnementales, notamment au niveau de la phase de pré-recrutement (Cury et Roy 1987 ; Faure *et al.* 2000).

Le poulpe constitue une espèce de grande importance dans la zone COPACE et notamment en Mauritanie (Balguerias *et al.* 2000) (Fig. 1). C'est quasiment la seule espèce cible des pêcheries nationales Mauritaniennes, industrielles et artisanales (Inejih 2000) et elle constitue une ressource économique importante pour le pays. Aussi, de nombreux bateaux de pêche étrangers viennent exploiter cette ressource (Espagne, Japon, Chine). Les céphalopodes en général génèrent près de 80% du chiffre d'affaire des exportations totales de produits halieutiques congelés Mauritanies (CMAP 2005) dont plus de 70% par le poulpe seul. Les principaux engins de pêche sont le chalut industriel et la pêche au pot artisanale.

Les captures actuelles environnent les 30.000 tonnes/an (Fig 2). Comme la majorité des espèces démersales des côtes Nord-Ouest Africaines, le stock de poulpe Mauritanie montre aujourd'hui des signes de surexploitation. Alors que l'effort de pêche a été multiplié par 3 entre 1991 et 2002 (CMAP 2002), les captures montrent de grandes fluctuations et une tendance générale à la baisse depuis 1970, passant d'environ 28000 tonnes à 23000 tonnes sur la période (Fig. 2).

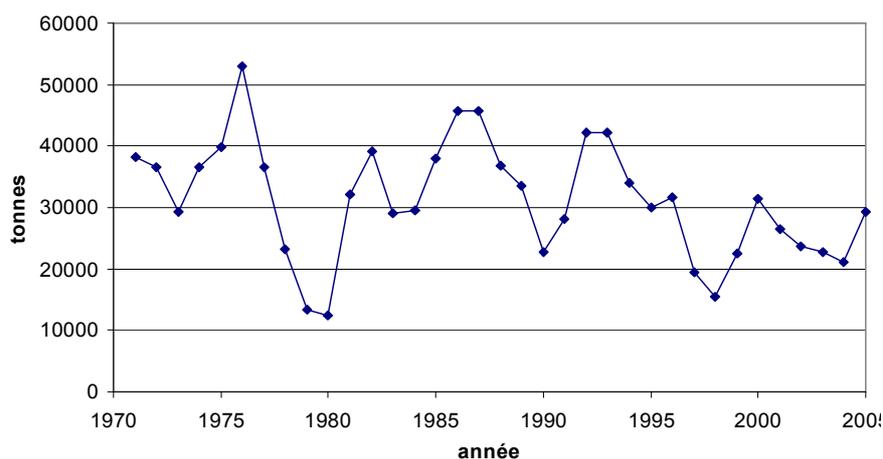


Figure 2 : Capture total du poulpe *Octopus vulgaris* en Mauritanie (source : IMROP 2006)

1.3 Méthodes classiques d'évaluation des stocks de poulpe

Plusieurs méthodes ont été mises en œuvre pour évaluer les stocks de céphalopodes (Pierce et Guerra 1994).

- Modèles de déplétion sur pas de temps court (mensuel ou hebdomadaire).

Ces modèles sont utilisés pour représenter, sur un pas de temps court (*e.g.* mois ou semaine), la décroissance de l'effectif d'un stock au cours d'une saison de pêche en fonction des captures et de la mortalité naturelle. Ils sont très utilisés pour les espèces à vie courte comme les céphalopodes (McAllister *et al.* 2004 ; Young *et al.* 2004), et sont adaptés pour une gestion de la pêcherie en cours de saison de pêche. Mais ils soulèvent notamment deux problèmes : *i*) Ils nécessitent des données fines (captures et indices d'abondance exprimés en effectifs et sur un pas de temps court) qui ne sont pas disponibles pour cette étude ; *ii*) Ils ne permettent pas d'estimer des points de référence pour les pêcheries : F_{msy} , C_{msy} ou B_{msy} .

- Modèles structurés (analyse de cohortes), type VPA sur pas de temps mensuel ou hebdomadaire.

L'intérêt de ces modèles est qu'ils représentent précisément les processus démographiques au sein de la population. Mais tout comme les modèles de déplétion, ils nécessitent des données fines non disponibles ici.

- Modèles de production de biomasse (pas de temps annuel)

C'est l'approche suivie par les groupes d'évaluation de l'IMROP (IMROP 2006 ; Chassot et al. 2006 ; Gascuel et Inejih 2006) et du COPACE (FAO 2006). Le modèle de production de biomasse (MPB dans la suite) est l'un des modèles les plus utilisés pour les évaluations de stock lorsque la quantité de données est très limitée (données de capture en poids et indice d'abondance, IA dans la suite) (Hilborn et Walters 1992). S'il est moins détaillé que les modèles de déplétion ou de VPA sur pas de temps court, le MPB offre des avantages : *i*) Il est parcimonieux (peu de paramètres) et peu gourmand en données (données de capture et d'IA) ; *ii*) Il permet de modéliser la dynamique sur le moyen-long-terme (série > 30 ans) ; *iii*) Il permet d'estimer les points de référence pour la gestion basés sur des points d'équilibre à long terme comme le C_{msy} ou le B_{msy} .

Dans ce travail, l'approche par modèle de production de biomasse sur un pas de temps annuel a été choisie.

1.4 Problèmes méthodologiques soulevés

Dans ce travail, nous prenons appui sur le cadre statistique Bayésien (Gelman *et al.* 1995) pour aborder plusieurs problèmes méthodologiques dans l'analyse du MPB.

- Modèle dynamique ou hypothèse d'équilibre ?

Classiquement, les groupes de travail pour l'évaluation des stocks en Mauritanie utilisent le MPB sous l'hypothèse d'équilibre, selon laquelle les captures (en poids) au cours de l'année t sont égales à la production de Biomasse au cours de cette même année. L'hypothèse d'équilibre montre certains avantages (notamment l'estimation des paramètres est plus aisée) mais reste conceptuellement discutable car elle considère que la biomasse de l'année $t+1$ est indépendante des biomasses des années précédentes. C'est pourquoi de nombreux auteurs anglo-saxons préconisent plutôt l'utilisation de ces modèles dans un cadre dynamique (Hilborn et Walters 1992). Dans ce travail, nous proposons d'adopter une approche dynamique de la production de biomasse.

- Prise en compte des incertitudes dans les diagnostics et les décisions (cadre statistique Bayésien)

Cette étude est menée dans un cadre statistique Bayésien (Gelman *et al.* 1995), particulièrement adapté à la prise en compte des incertitudes dans les diagnostics et les décisions. En effet, le cadre statistique Bayésien est entièrement basé sur la manipulation de distributions de probabilité. Il est en ce sens mathématiquement rigoureux, conceptuellement élégant et dans la pratique adapté pour la prise en compte des incertitudes et des risques dans l'évaluation et la gestion des ressources halieutiques (Punt et Hilborn 1997 ; McAllister et Kirkwood 1998 ; Harwood et Stockes 2003).

- Prise en compte des erreurs d'observation et de la stochasticité dans les modèles

De nombreuses études montrent l'importance de la prise en compte de multiples sources d'incertitude dans les analyses quantitatives, et notamment des erreurs dans les données (erreur d'observation) et de la variabilité dans les modèles (erreur de processus) (Harwood et Stockes 2003). Dans des travaux de simulation déjà anciens réalisés sur le MPB,

Punt (Punt 1989a dans Punt 2003) et Polacheck *et al.* (Polacheck *et al.* 1993) pointent les différences dans les qualités d'ajustement (biais et précision des estimations) selon que seule l'erreur de processus est considéré (« *process error model* ») ou que seule l'erreur d'observation est considérée (« *observation error model* »). Plus récemment, de nombreux auteurs ont montré que dans le cas de modèles dynamiques, la prise en compte simultanée des deux sources d'incertitude dans le cadre théorique des modèles stochastiques à états latents (*state-space models* ou *SSM*) permettait en moyenne d'obtenir les estimations les moins biaisées (Punt 2003 ; Meyer et Millar 1999 ; Millar et Meyer 2000).

Dans ce travail, nous proposons d'intégrer le modèle de production de biomasse dans un cadre de *SSM* afin d'intégrer différentes sources d'incertitude : l'erreur de processus et l'erreur d'observation. Ce travail s'appuiera sur les premières analyses menées dans ce sens par Chassot *et al.* (2006). Meyer et Millar (1999) et Millar et Meyer (2000) ont montré comment les méthodes numériques de simulation de Monte Carlo pour l'inférence Bayésienne offraient une grande souplesse dans le traitement statistique des *SSM* par rapport aux méthodes classiques d'ajustement. Notamment, ces méthodes permettent de traiter des modèles de grande dimension, et de s'affranchir des hypothèses de linéarité des équations de processus et d'observation et offrent la possibilité d'utiliser des termes aléatoires non Gaussiens.

- Effet de l'environnement

Enfin les MPB ne prennent généralement pas en compte de façon explicite l'effet des fluctuations environnementales dans la dynamique de la biomasse. Or, la sensibilité de la dynamique du poulpe (recrutement ou phase larvaire) aux fluctuations environnementales est bien connue (Caddy *et al.* 1983 ; Boyle et Boletzky 1996 dans Faure *et al.* 2000). Ceci est exacerbé par le fait qu'il s'agit d'une espèce à vie courte.

Gascuel et Inejih (2002) avait proposé une méthode d'intégration de l'effet de l'environnement dans un MPB sous l'hypothèse d'équilibre. Dans ce travail, nous proposons une approche similaire mais pour un MPB dynamique dans le cadre Bayésien.

1.5 Les objectifs du travail

Ce travail poursuit plusieurs objectifs :

1 – Evaluer la faisabilité d'une approche statistique Bayésienne et évaluer la qualité des estimations fournies.

Les méthodes d'estimation de modèles complexes (notamment les *SSM*) par des méthodes Bayésiennes sont aujourd'hui conceptuellement bien établies. Leur mise en oeuvre pratique est largement facilitée par la mise à disposition de logiciels dédiés. Néanmoins, en pratique, cela demande un investissement méthodologique et des simulations qui peuvent être assez lourds. Il est donc nécessaire de tenter d'évaluer la fiabilité des résultats obtenus.

On traitera ce problème par une approche simulation-ajustement adaptée au cas d'étude de la Mauritanie. La démarche générale consiste à simuler des trajectoires de biomasse et d'IA selon les équations du MPB à paramètres connus (en introduisant de l'aléa dans le processus et dans les observations), puis à comparer les valeurs (connues) des paramètres ayant servi lors de la simulation des données avec les estimations de ces paramètres obtenues par une méthode statistique Bayésienne. On s'intéressera notamment aux problèmes suivants :

- quelle est la faisabilité technique de l'approche (programmation, temps de calcul, choix des distributions a priori) ;

- quelle est la qualité (biais et dispersion) des estimations des paramètres obtenues et l'influence de la quantité d'information disponible sur la qualité de ces estimations ;
- quel est le biais éventuel dans les estimations introduit par l'hypothèse d'équilibre (production = captures) ;
- quelle est la qualité des estimations obtenues par un ajustement dans un cadre de *SSM* Bayésien par rapport à un ajustement dans un cadre d'*observation error model* dans le cadre statistique Bayésien et fréquentiste (ajustement par maximum de vraisemblance).

2 – Application aux pêcheries de poulpe Mauritaniennes

Dans un deuxième temps, il s'agira d'appliquer la méthodologie développée précédemment pour réaliser l'évaluation du stock de la pêche Mauritanienne à partir des séries de captures et d'indice d'abondance de pêche Mauritanienne.

3 – Prise en compte de covariables environnementales dans les modèles

Dans cette troisième partie, on s'intéressera à la prise en compte de covariables environnementales pour tenter d'expliquer une partie de la variabilité environnementale dans le MPB. On s'intéressera plus particulièrement à l'influence de l'indice d'upwelling dont l'impact sur la dynamique des stocks de poulpe a déjà été étudié (Caddy *et al.* 1983 ; Inejih 2000 ; Faure *et al.* 2000 ; Gascuel et Inejih 2002).

1.6 Plan du mémoire

La seconde partie de ce rapport est consacrée aux Matériels et Méthodes. Après une brève présentation des données de la pêche Mauritanienne, et des covariables environnementales, l'approche de simulation-ajustement sera détaillée. La méthodologie suivie pour l'application au cas Mauritanie sera ensuite présentée. L'essentiel des résultats concerne l'approche méthodologique de simulation-ajustement. Les résultats concernant l'application au cas Mauritanie seront ensuite présentés. Le rapport se termine par une discussion qui évoquera les limites et les perspectives ouvertes par ce travail.

2 Matériel et Méthodes

2.1 Données

2.1.1 Pêche Mauritanienne

Les données de capture et d'indice d'abondance de la pêche Mauritanienne sont issues des travaux de l'IMROP (IMROP 2006) et s'étendent sur 35 ans : de 1971 à 2005. Les deux grandes zones de pêche Nord et Sud (correspondant sans doute à deux sous-stocks) ne sont pas distinguées dans ce travail.

- Captures

Les captures totales (en poids) correspondent à une somme des captures industrielles (nationales et étrangères) et artisanales des pêcheries Nord et Sud de la Mauritanie (cf. Annexe 1).

- Données d'indice d'abondance

La série d'IA utilisée provient du dernier groupe de travail de l'IMROP (IMROP 2006) (cf. Annexe 1). Elle est construite à partir d'une moyenne géométrique de trois séries d'IA issues elles-mêmes des campagnes scientifiques et des données de CPUE de la pêche industrielle et artisanale. Cette série est donc sans unité.

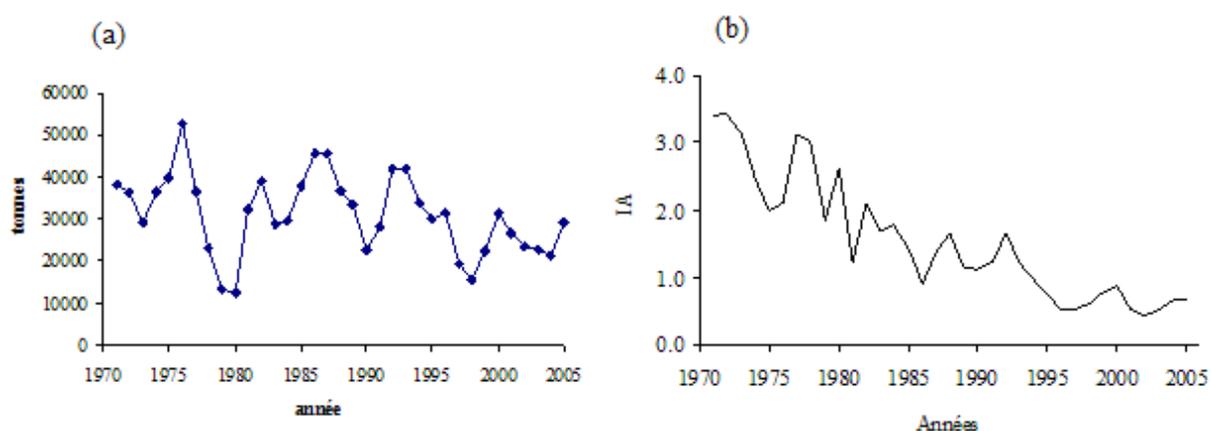


Figure 3 : (a) : Captures totales (en tonnes) de poulpes en Mauritanie. (b) : Indice d'abondance de poulpes en Mauritanie. (Source : IMROP 2006)

2.1.2 Données environnementales

On s'est intéressé en premier lieu à l'intégration de l'indice d'upwelling côtier (IUC) dans le MPB. L'upwelling côtier est le principal élément qui conditionne l'hydro-climat de ces régions Nord-Ouest Africaines et a un fort impact sur la production biologique. En effet, cette remontée d'eau froide des profondeurs apporte avec elle des sels minéraux nécessaires à la production planctonique qui forme la base de la chaîne alimentaire.

Deux séries différentes d'indice d'upwelling ont été considérées. La première a été calculée à partir de l'équation de transport d'Ekman à partir des données de vitesse des vents. Cette série est issue des groupes de travail de l'IMROP et elle s'étend de janvier 1970 à avril 2004 avec un pas de temps mensuel (Bambaye, communication personnelle) (cf. Annexe 2). La deuxième série d'IUC est calculée à partir des données de température de surface de la mer (SST) du satellite AVHRR (Advanced Very High Resolution Radiometer). Les données s'étendent de janvier 1985 à décembre 2005 avec un pas de temps mensuel. On s'appuie sur les travaux de Demarcq et Faure (2000) et Faure *et al.* (2000) pour le calcul de l'indice (cf. Annexe 3).

Les séries d'IUC ont été calculées en se concentrant sur les périodes clés du cycle de vie du poulpe. Deux pics de ponte ont été identifiés pour le poulpe de Mauritanie (Dia 1988 dans Inejih 2000 ; Inejih 1996 ; Hatanaka 1979). La période varie légèrement selon les auteurs mais cela reste concentré sur la période suivante :

- Pic principal : Avril – Juin
- Pic secondaire : Septembre-octobre

Les œufs mettent quelques mois pour éclore avant de donner naissance à des larves pélagiques. Selon des études menées par différents auteurs, cette durée d'incubation varie en fonction de la SST mais pour des températures d'eau comprises entre 15 et 25°C, les larves apparaissent environ 1 mois après la ponte (Guerra 1975 ; Itami 1963 dans Guerra 1975 ; Mangold & Boletzky 1973 dans Faure *et al.* 2000). Puisque la SST de la zone Mauritanienne se situe dans cette tranche de température tout au long de l'année, on peut estimer la durée d'incubation à environ 1 mois. Ce qui nous permet de déterminer deux périodes d'émergence des larves pélagiques (correspondant aux deux pics de ponte) :

- Pic principal : Mai – Juillet
- Pic secondaire : Octobre – Novembre

Dans ce travail, le choix a été fait d'étudier l'influence de l'IUC au cours de la période Mai - Juillet correspondant à la phase d'incubation pour le pic principal de ponte. Pour cela, une moyenne géométrique de l'IUC pendant ces 3 mois a été calculée pour chaque année et pour les deux types d'indice (créé à partir des données de vent, et créé à partir des SST). On exprimera les séries de la façon suivante : IUC₁ la série créée à partir des données de vent et IUC₂ celle calculée à partir des SST.

Des relations avec les séries de température de surface de la mer SST ont aussi été recherchées. En effet, la température serait susceptible d'influencer la production de poulpe de part son rôle dans la reproduction et la croissance (Guerra 1975 ; Wood *et al.* 2000 ; Semmens *et al.* 2004).

2.2 Modèle de Biomasse – *State-Space Model (SSM)*

2.2.1 Principe général d'un SSM

Les *state-space models* (Hilborn and Peterman 1996 dans Lucato 2006 ; Buckland *et al.* 2004 ; Meyer et Millar 1999 ; Punt 2003) constituent un cadre général de modélisation qui permet de prendre en compte deux sources d'incertitude dans des modèles dynamiques :

- *l'erreur de processus* qui correspond à l'incertitude liée à la dynamique du système modélisé (e.g. la dynamique des populations) (voir après)

- *l'erreur d'observation*, associée à des imprécisions de mesure dans les données.

Le SSM est structuré par deux processus qui sont liés l'un à l'autre :

- le processus d'état qui intègre l'erreur de processus dans la dynamique du système.

- le processus d'observation qui intègre l'erreur d'observation i.e le fait que les variables d'état du système soient indirectement observés et avec une certaine incertitude.

2.2.2 Processus d'état

C'est l'équation qui décrit l'évolution de la biomasse B_t d'une année sur l'autre. Le MPB que l'on a utilisé considère la biomasse d'un stock comme une unité simple qui n'est ni structurée en âge ni en taille. Dans un cadre déterministe, ce modèle décrit la biomasse en début d'année t+1 (B_{t+1}) comme une fonction de la biomasse en début d'année t (B_t), de la production de l'année t ($g(B_t)$), et de la capture de l'année t (notée C_t^{obs} car on la considère ici comme une série de données observées sans erreur) :

$$(1) \quad B_{t+1} = B_t + g(B_t) - C_t^{obs}$$

La fonction de production $g(B_t)$ traduit le changement global de la biomasse au cours de l'année, dû aux phénomènes de recrutement, croissance et de mortalité naturelle des individus (Ricker 1975). Dans le cadre de notre étude, on a choisi d'utiliser la fonction de production de Fox (1970) utilisée dans les groupes de travail de l'IMROP.

$$(2) \quad g(B_t) = r \cdot B_t \cdot \left(1 - \frac{\ln(B_t)}{\ln(K)}\right)$$

où r est le taux intrinsèque de croissance de la population et K la capacité biotique du milieu. Cette fonction montre une relation de densité-dépendance donc de compétition intra-spécifique (qui augmente avec la biomasse) (terme $(1 - \ln(B_t)/\ln(K))$).

Cette équation dynamique de production de biomasse est écrite dans un cadre stochastique en y adjoignant un terme aléatoire log-Normal (qui décrit l'erreur de processus). Ce terme d'erreur correspond, entre autres, à des fluctuations de biomasse d'un stock en fonction des conditions environnementales.

$$(3) \quad B_{t+1} = \left(B_t + r \cdot B_t \cdot \left(1 - \frac{\ln(B_t)}{\ln(K)}\right) \right) \cdot e^{\varepsilon 1(t)} - C_t^{obs}$$

avec $e^{\varepsilon 1(t)}$ l'erreur de processus où $\varepsilon 1(t) \stackrel{iid}{\sim} N(0, \sigma_p^2)$.

2.2.3 Processus d'observation

La biomasse B_t est une variable non observée. On dit que c'est une variable d'état du modèle. Le processus d'observation permet de faire le lien entre la série de données d'indice d'abondance et cette variable d'état. On fait l'hypothèse classique que l'indice d'abondance est proportionnel à la biomasse B_t , et que le coefficient de proportionnalité q est constant dans le temps. L'erreur d'observation est supposé log.Normale :

$$(4) \quad I_t^{obs} = (q \cdot B_t) \cdot e^{\varepsilon 2(t)}$$

avec I_t^{obs} l'indice d'abondance, q la « capturabilité » et $e^{\varepsilon 2(t)}$: l'erreur d'observation où $\varepsilon 2(t) \stackrel{iid}{\sim} N(0, \sigma_{IA}^2)$.

Les séries d'IA utilisées sont sans unités. Par conséquent, le paramètre q du modèle ne correspond pas à l'idée de capturabilité (au sens traditionnel) *i.e* la probabilité, pour un individu pris au hasard dans un ensemble, d'être capturé par une unité de pêche (Fréon 1991). Ce sera plutôt un paramètre de forme sans unité.

2.2.4 SSM, process error model, observation error model

Le SSM conjugue donc les aléas dans le processus et les erreurs de mesures :

$$(5) \quad \begin{cases} B_{t+1} = (B_t + g(B_t)) \cdot e^{\varepsilon 1(t)} - C_t^{obs} \\ I_t^{obs} = (q \cdot B_t) \cdot e^{\varepsilon 2(t)} \end{cases}$$

On désigne par λ , le rapport des variances de processus et d'observation.

$$(6) \quad \lambda = \sigma_p^2 / \sigma_{IA}^2$$

Le « *process error model* » est un modèle qui n'intègre que l'erreur de processus ($\sigma_{IA}^2 = 0$ et $\sigma_p^2 > 0$).

L'« *observation error model* » est un modèle qui suppose que l'équation de la dynamique de population est déterministe mais que l'observation est réalisée avec erreur ($\sigma_p^2 = 0$ et $\sigma_{IA}^2 > 0$).

2.2.5 Hypothèses supplémentaires nécessaires pour assurer l'identification des paramètres

Dans le cadre du *SSM*, deux hypothèses sont nécessaires pour assurer l'identification des paramètres :

- La biomasse de la première année est supposée proportionnelle à la biomasse à l'état vierge ; Dans cette étude, l'hypothèse par défaut a été choisie, soit $B_1=K$ (Punt 2003 ; Millar *et al.* 2000).

- Le rapport des variances λ est supposé connu. En effet, à moins de disposer d'une idée a priori de la valeur d'au moins une des deux variances (processus ou observation), λ ne peut pas être estimé et doit donc être fixé a priori. L'hypothèse $\lambda=1$ est fixée par défaut, ce qui correspond à une répartition de l'incertitude à part égale entre le processus et l'observation. Mais différentes valeurs de λ seront testées au cours de l'étude. $\lambda \gg 1$ correspond à un modèle proche d'un *process-error-model*. Au contraire, $\lambda \ll 1$ permet de se rapprocher d'une configuration de type *observation-error-model*.

2.3 Hypothèse d'équilibre

A l'équilibre, on suppose que la production générée par le stock l'année t est équivalente à la capture réalisée la même année.

$$(7) \quad g(B_t) \cdot e^{\varepsilon I(t)} = C_t^{obs} \quad \Leftrightarrow \quad r \cdot B_t \cdot \left(1 - \frac{\ln(B_t)}{\ln(K)}\right) \cdot e^{\varepsilon I(t)} = C_t^{obs}$$

L'hypothèse $B_1 = K$, nécessaire dans le modèle dynamique, est relaxée dans le cas d'un ajustement sous l'hypothèse d'équilibre.

Sous l'hypothèse d'équilibre, il n'a pas été possible de modéliser simultanément les erreurs de processus et d'observation. Seule l'erreur de processus sera conservée, ce qui correspond à des hypothèses équivalentes à celle réalisée dans le groupe de travail IMROP (IMROP 2006).

2.4 Ajustement dans le cadre Bayésien

2.4.1 Principe

Le cadre statistique Bayésien est fondé sur une représentation de l'incertitude autour de l'estimation des inconnues d'un modèle par des distributions de probabilité (Gelman *et al.* 1995). L'analyse demande de fixer une distribution a priori (ou prior) qui représente la connaissance que l'on a autour de chaque inconnue avant l'analyse des données en cours. L'analyse Bayésienne permet de passer du prior au posterior grâce à une mise à jour apportée par les données via le modèle.

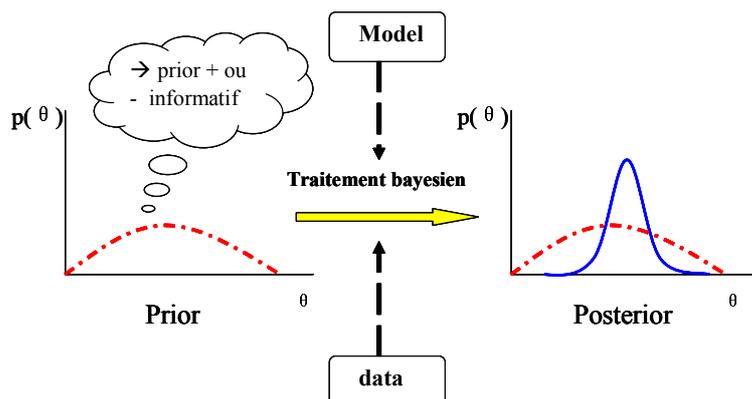


Figure 4 : Représentation schématique de la chaîne de traitement Bayésien

En notant θ , le vecteur des inconnues du modèle, la distribution a posteriori de θ , notée $p(\theta|Data)$, s'obtient en combinant la distribution a priori $p(\theta)$ et la distribution conditionnelle des données $p(Data|\theta)$ (= vraisemblance) par la formule de Bayes :

$$(8) \quad p(\theta/Data) = \frac{p(Data/\theta) \cdot p(\theta)}{\int_{\theta} p(Data/\theta) \cdot p(\theta) d\theta}$$

2.4.2 Choix des priors

La démarche générale a consisté à travailler autant que faire se peut avec des priors non informatifs sur les paramètres. Néanmoins, lorsque des priors non informatifs sont sélectionnés pour tous les paramètres du modèle, on aboutit à une indétermination du fait du manque d'information dans les données. C'est pourquoi nous avons choisi d'affecter un prior informatif pour le paramètre r . En effet, fixer un prior informatif sur K semblait plus délicat du fait du caractère très stock-dépendant de ce paramètre : le transfert d'information à partir d'autres stocks pour définir un prior sur K est moins évident.

Choix des priors informatifs sur r

Le choix d'un prior informatif de forme gamma a été fait pour le paramètre r . La nature du prior choisi dépend des cas d'étude. Pour plus de précision, se référer aux paragraphes 2.5.3. et 2.6.

Prior « non informatifs » sur les autres paramètres ($r, K, q, \sigma_p^2, \sigma_{IA}^2$)

Différentes formes de priors non informatifs sont proposées dans la littérature (Myers *et al.* 2002 ; Punt 2003 ; Millar et Meyer 2000 ; McAllister et Kirkwood 1998) pour les paramètres du MPB. Les choix suivants ont été faits :

- $K \sim \text{unif}(50000 ; 10000000)$ Non informatif au vue de l'intervalle de variation des captures : $\max(C_t^{obs}) \approx 50000$
- $\log(q) \sim \text{unif}(-15 ; 3)$ Proportionnel à $1/q$ entre $3 \cdot 10^{-7}$ et 20
- $\log(\sigma_p^2) \sim \text{unif}(-9 ; 3)$ Proportionnel à $1/\sigma_p^2$ entre $1 \cdot 10^{-4}$ et 20
- $\sigma_{IA}^2 = 1/\lambda \cdot \sigma_p^2$

2.4.3 Principe de l'estimation des distributions par MCMC

Les distributions de probabilité ont été estimées à partir des méthodes d'échantillonnage par simulations MCMC (Monte Carlo Markov Chain) (Gelman *et al.* 1995) par l'intermédiaire du logiciel BUGS (Bayesian inference Using Gibbs sampling) (téléchargeable gratuitement au <http://mathstat.helsinki.fi/openbugs/>). L'automatisation des chaînes de traitement est réalisée à l'aide de l'interface R – WinBUGS fournie par la librairie BRUGS (OpenBUGS 2.2.0 beta).

Afin de ne pas alourdir le texte de ce rapport, des détails techniques de l'estimation par MCMC sont donnés en Annexe 4 et l'Annexe 7 fournit les codes WinBUGS pour les principaux modèles.

2.5 Evaluation de la méthode par une approche de simulation

Une grande partie du travail de ce stage a été consacré au développement d'une approche de simulation-ajustement afin d'évaluer la faisabilité de la méthode et la fiabilité des inférences statistiques. Ce travail se décompose en 3 étapes successives (Fig. 5).

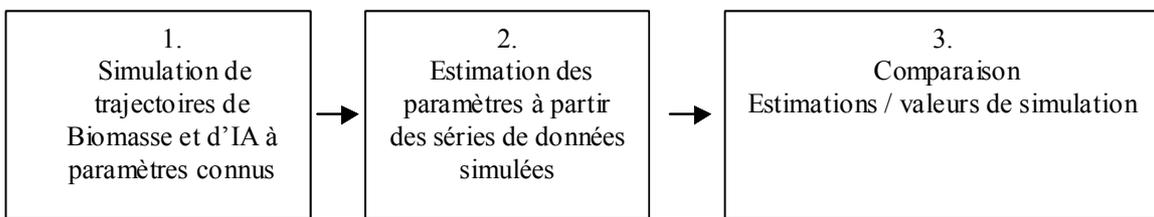


Figure 5 : Grandes étapes de l'approche simulation-ajustement

2.5.1 Simulation à paramètres connus

Dans cette première étape, des trajectoires de Biomasse et d'IA sont simulés à partir des équations du modèle (eq. 3 et 4) en fixant les valeurs des paramètres r , q , σ_p et λ . En s'inspirant de Punt (2003), il a été choisi de fixer le taux de réduction de biomasse $B_{n=35}/K$ plutôt que de fixer la capacité d'accueil K . Mais $B_{n=35}/K$ et tous les autres paramètres ainsi que la série de captures étant fixés, la valeur de K se déduit directement.

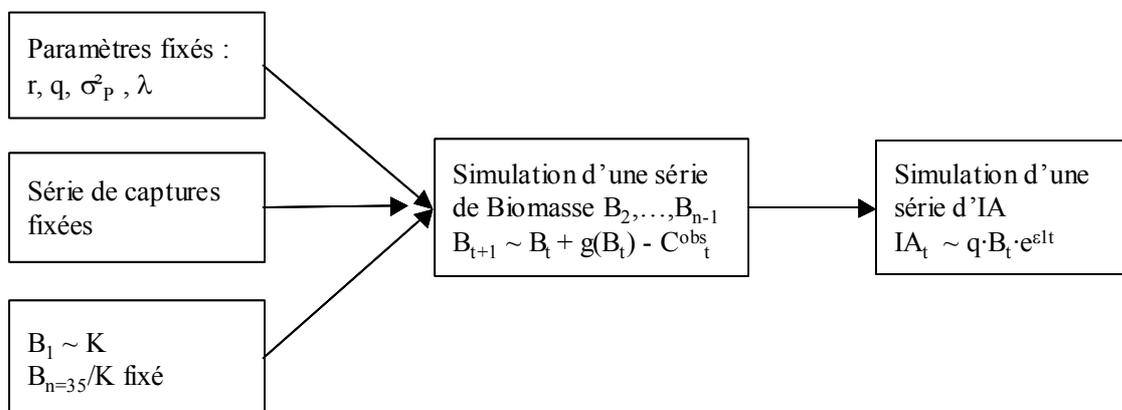


Figure 6 : Procédure de simulation des séries d'IA

Chaque combinaison de paramètres constitue « un scénario » et permet de créer une trajectoire d'indice d'abondance qui lui est spécifique. 13 scénarii différents ont été simulés, chacun à partir de combinaisons différentes de paramètres et représentant des situations variées de l'évolution du stock (Tableau 1).

- Chaque scénario emprunte la même série de captures totales des pêcheries Mauritanienes, considérée comme connue.
- Afin de se rapprocher du cas d'étude Mauritanie, on génère ici des séries d'IA du type « one-way trip » en choisissant les valeurs de $B_{n=35}/K$ suivantes : 0.8, 0.5 et 0.2 (Fig. 7) :

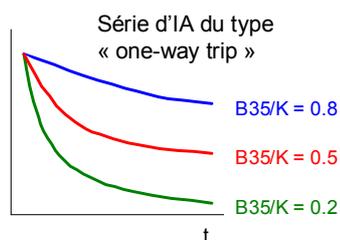


Figure 7 : Les séries d'IA du type « one-way » trip

- 3 valeurs du paramètre r ont été utilisées : $r=0.6$ (faible), $r=1.2$ (moyenne) (Balguerias dans IMROP 2006) et $r=3$ (forte) ;
- 3 valeurs de λ ont été testées. $\lambda = 0.1$, $\lambda = 1$ et $\lambda = 10$;
- Le paramètre de capturabilité a été fixé à $q = 10^{-5}$ pour tous les cas. (Balguerias dans IMROP 2006).

Numéro de scénario	B_{35}/K	r	K	q	λ	σ_p^2	
1	0,2		624959	10^{-5}	1	0.0049	
2	0,5	1,20	816503	10^{-5}	1	0.0049	
3	0,8		1775133	10^{-5}	1	0.0049	
4	0,2		870489	10^{-5}	1	0.0049	
5	0,5	0,60	1252907	10^{-5}	1	0.0049	
6	0,8		2936590	10^{-5}	1	0.0049	
7	0,2		323623	10^{-5}	1	0.0049	
8	0,5	3,00	351177	10^{-5}	1	0.0049	
9	0,8		639299	10^{-5}	1	0.0049	
Numéro de scénario	B_{35}/K	r	K	q	λ	σ_p^2	σ_{IA}^2
10					1	0.0004	0.0004
11					1	0.04	0.04
12	0,2	1,20	624959	10^{-5}	0.1	0.00049	0.0049
13					10	0.0049	0.00049

Tableau 1 : Tableau récapitulatif des valeurs de simulations

Pour chaque scénario, 10 trajectoires de Biomasse et d'IA (=10 simulations) sont générées aléatoirement selon les équations du modèle (eq 3 et 4, série de captures considérée comme connue) (La valeur de 10 simulations peut paraître faible pour ce genre d'étude mais pour des raisons techniques de temps de calcul, il n'a pas été possible de réaliser plus de simulations).

Par exemple, la Figure 8 présente les trajectoires de Biomasse et d'IA obtenues dans le cas du scénario 1.

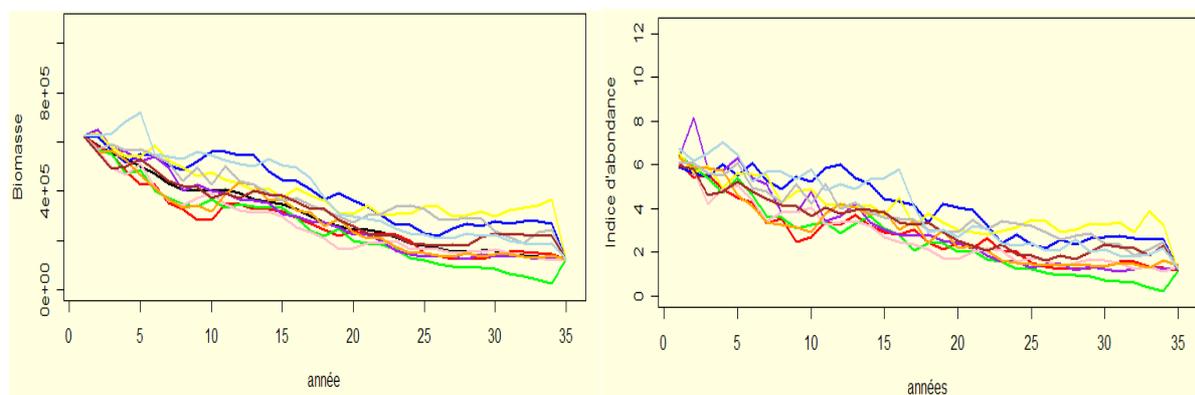


Figure 8 : Tracés des trajectoires de Biomasse et d'IA générés aléatoirement

2.5.2 Estimation des paramètres à partir des données simulées

Pour chaque scénario, la méthode Bayésienne a été mise en œuvre pour estimer les paramètres à partir des 10 exemples de trajectoires d'IA générées aléatoirement. Pour chaque scénario, on dispose donc de 10 estimations différentes (sous la forme de distribution de probabilité) qui sont alors comparées aux « vraies valeurs des paramètres ».

Les inférences Bayésiennes ont été réalisées dans les conditions suivantes :

- Dans tous les cas, sauf lorsque c'est explicitement mentionné (pour les cas équilibre vs dynamique et comparaison avec une méthode classique de maximum vraisemblance - voir ci-après), l'estimation Bayésienne est réalisée sous les mêmes hypothèses de modèle (équation, forme des aléas, $B_1=K$, et λ fixé) que celles ayant servies aux simulations.

- La trajectoire de captures, déjà considérée comme connue dans les simulations, est aussi considérée comme connue dans l'étape d'ajustement.

- Pour chaque scénario, l'ajustement est réalisé en n'utilisant que les données simulées des 34 premières années (IA entre 1971 et 2004). Le fait de fixer le taux de réduction de biomasse $B_{n=35}/K$ lors de la simulation des trajectoires contraint les séries de biomasse à prendre des valeurs spécifiques à la 35^{ème} année. Cela entraîne alors des sauts de valeurs entre B_{34} et B_{35} (donc entre IA_{34} et IA_{35}) dont la crédibilité est très faible conditionnellement aux valeurs de σ_p^2 choisies dans les scénarii (cf. l'exemple de la Fig. 8). De ce fait, les valeurs de σ_p^2 (et donc de σ_{IA}^2) sont systématiquement surestimées dans les ajustements et entraînent par conséquent des biais dans les inférences des autres paramètres. Ainsi, il a fallu exclure les données de la dernière année au moment des ajustements.

- L'estimation Bayésienne nécessite de définir des priors sur les paramètres. Des priors non informatifs sont définis sur les paramètres K , q , σ_p^2 (voir 2.4.2.). Un prior informatif de type « gamma » est défini sur le paramètre r . Le prior sur r est adapté à chaque scénario. Il est défini de telle sorte que la variance soit toujours égale à 1 mais que le mode de la distribution a priori soit égal à la valeur de r ayant servie à simuler les données dans le scénario considéré (Fig. 9) :

- mode = 0,6 : $r \sim \text{gamma}(1.81 ; 1.34)$
- mode = 1,2 : $r \sim \text{gamma}(3.12 ; 1.77)$.
- mode = 3 : $r \sim \text{gamma}(10.91 ; 3.30)$

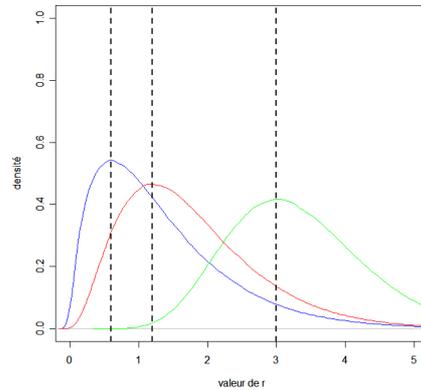


Figure 9 : Forme des 3 priors mis en place sur le paramètre r

2.5.3 Objectifs de l'étude de comparaison simulation/estimation

2.5.3.1 Sensibilité aux priors

Comme un prior informatif a été mis en place sur r , une étude de sensibilité aux priors a été réalisée pour analyser l'impact de ces priors informatifs sur l'ajustement. 2 priors différents sur r ont été testés pour les deux scénarii 4 et 7 :

- Un prior effectivement centré sur la valeur de r utilisée pour les simulations : $r = 0.6$, $r = 3$ (appelé par la suite : « prior adapté »)
- Un prior centré sur 1.2 donc a priori décalé de la vraie valeur de r utilisée dans les scénarii 4 et 7.

2.5.3.2 Sensibilité à la quantité d'information contenue dans les données

Les données disponibles pour les évaluations de stock ne sont pas toujours très informatives. L'approche simulation-ajustement permet de quantifier l'influence de l'information contenue dans les données sur la qualité des estimations. On comparera ici les estimations obtenues *i*) pour différentes valeurs de l'indice de réduction de stock $B_{n=35}/K$. Les résultats seront illustrés en s'appuyant sur les scénarii 1, 2 et 3 ; *ii*) pour différentes valeurs de variance (σ_p^2 et σ_{IA}^2) qui viennent « bruyter » les trajectoires de biomasse et d'indice d'abondance. Les résultats seront illustrés avec les scénarii 1, 10 et 11.

2.5.3.3 Influence de l'hypothèse d'équilibre

Les résultats d'ajustement des paramètres r , K , C_{msy} , q , σ_p^2 , B_2 , B_{34} et B_{34}/K , de tous les scénarii (1 à 13) ont été comparés entre l'ajustement à partir des équations dynamiques et l'ajustement sous hypothèse d'équilibre. On rappelle que dans ce dernier, le modèle utilisé dans l'ajustement Bayésien (Hyp. d'équilibre) est différent de celui qui a servi à générer les données (dynamique). Le but de cette étude est alors de montrer les éventuels biais et incertitudes générés par cette erreur de modèle.

2.5.3.4 Comparaison SSM Bayésien, *observation error model* dynamique Bayésien et *observation error model* dynamique en max. vrais.

La méthode d'ajustement en *observation error model* est aujourd'hui la méthode la plus utilisée en pratique pour les modèles dynamiques. L'une des raisons majeures est la facilité de la mise en œuvre du modèle dans un cadre classique fréquentiste du maximum de vraisemblance. Polacheck avait également souligné ses atouts vis-à-vis d'un *process error model* (Polacheck 1993).

Les résultats obtenus par l'ajustement Bayésien sous l'hypothèse SSM ont alors été comparés à ceux obtenus par un ajustement sous l'hypothèse d'*observation error model* dans le cadre Bayésien et dans le cadre fréquentiste (méthode de maximum de vraisemblance). Pour ce dernier cas, la fonction NLS de R (moindres carrés non linéaires) a été utilisée. Le but de ce travail est alors d'analyser le niveau d'erreur que l'on commet en réalisant un ajustement à partir d'un modèle différent de celui qui a servi à générer les données. Punt (2003) avait signalé que cela pouvait entraîner une augmentation des biais dans les estimations (Punt, 2003).

Ce travail de comparaison sera fondé sur l'étude du cas du scénario 12 dont les données ont été simulées dans un cas proche d'un *observation error model*.

2.5.4 Méthodes de comparaison simulations / estimations

Les comparaisons entre estimations et valeur simulées ont été menées afin de quantifier le biais et les incertitudes dans les estimations. 3 méthodes complémentaires ont été combinées pour analyser les résultats :

- Analyse « visuelle » des profils de densité a posteriori
- Etude des écarts quadratiques moyens (\overline{EQM}) et des écarts relatifs moyens (\overline{ERM}) (Casella et Berger 1990 dans Hammond *et al.* 2005).
- Etude des intervalles de crédibilité (IC) Bayésiens.

2.5.4.1 Observation des profils de distributions a posteriori

On compare ici la forme des distributions a posteriori des différents paramètres et leurs « décalages » par rapport aux valeurs « vraies » utilisées dans la simulation des données.

2.5.4.2 Les écarts quadratiques moyens et les écarts relatifs moyens

L' \overline{EQM} est une valeur synthétique et spécifique à chaque paramètre et donne des informations sur les biais et/ou les incertitudes d'un ajustement Bayésien pour ce paramètre. On ne peut pas comparer les valeurs d' \overline{EQM} entre les différents paramètres. En règle générale, pour un paramètre donné, une faible valeur de l' \overline{EQM} montre une distribution a posteriori peu biaisée et/ou peu dispersée. L' \overline{EQM} a été calculé pour chaque scénario et pour les paramètres r , K , q , C_{msy} , σ^2_p , B_2 , B_{34} , et B_{34}/K par les formules suivantes :

Pour chaque simulation s ($s = 1, \dots, 10$), on calcule d'abord pour chaque paramètre l' EQM_s :

$$(9) \quad EQM_s = \frac{1}{G} \cdot \sum_{i=1}^G (\theta_s^i - \theta_{true})^2$$

où G est la taille de l'échantillon MCMC. L'écart quadratique moyen correspond à la moyenne sur les 10 simulations EQM_s :

$$(10) \quad \overline{EQM} = \frac{1}{10} \cdot \sum_{s=1}^{10} EQM_s$$

L' \overline{ERM} est un indice complémentaire de \overline{EQM} qui renseigne sur la tendance des biais des estimations de chaque paramètre. Un paramètre a tendance à être surestimé si $\overline{ERM} > 0$ ou sous-estimé si $\overline{ERM} < 0$. Pour chaque paramètre d'une simulation s ($s = 1, \dots, 10$), on calcule :

$$(11) \quad ERM_s = \frac{1}{G} \cdot \sum_{i=1}^G \left(\frac{\theta_s^i - \theta_{true}}{\theta_{true}} \right) \quad \text{et} \quad (12) \quad \overline{ERM} = \frac{1}{10} \cdot \sum_{s=1}^{10} ERM_s$$

Dans le cas où les estimations des paramètres sont réalisés dans le cadre du maximum de vraisemblance, le calcul des quantités \overline{EQM} et \overline{ERM} est simplifié car on compare alors la « vraie valeur » du paramètre avec son estimation ponctuelle. On calculera alors l'écart quadratique moyen ponctuel (\overline{EQMP}) et l'écart relatif moyen ponctuel (\overline{ERMP}) de la façon suivante :

$$(13) \quad \overline{EQMP} = \frac{1}{10} \cdot \sum_{s=1}^{10} (\bar{\theta}_s - \theta_{true})^2 \quad \text{et} \quad (14) \quad \overline{ERMP} = \frac{1}{10} \cdot \sum_{s=1}^{10} \frac{(\bar{\theta}_s - \theta_{true})}{\theta_{true}}$$

avec $\bar{\theta}_s$ l'estimation du maximum de vraisemblance dans le cas fréquentiste ou le mode de la distribution a posteriori dans le cas Bayésien.

Cependant, l'étude de ces 2 indices (\overline{EQM} et \overline{ERM}) ne permet pas à elle seule de déterminer si le « décalage » entre les estimations et les valeurs simulées est dû principalement à un biais dans les estimations ou à une incertitude dans les estimations ou aux deux. Par exemple, la valeur d' \overline{EQM} pour un paramètre peut être élevée parce que les distributions a posteriori de ce paramètre est très biaisé mais avec une faible incertitude mais aussi parce que les posteriors sont peu biaisés mais avec une grande incertitude.

2.5.4.3 Les IC Bayésiens

L'IC Bayésien de niveau $1-\alpha$ renseigne sur les biais des estimations. Pour chaque scénario, on dénombre le nombre de simulations dans lesquelles l'IC de niveau $1-\alpha$ comprend la vraie valeur de simulation. Par exemple, avec un IC à 75%, dans le cas d'un ajustement non biaisé, on s'attend à ce que 7 ou 8 simulations donnent des IC qui contiennent la « vraie » valeur du paramètre. Si ce nombre est moindre, alors les estimations sont biaisées. Par contre, cet indice ne renseigne pas sur le niveau d'incertitude des estimations.

2.6 Evaluation du stock de poulpe Mauritanie - Modèle de biomasse avec effet de l'environnement

Le cadre méthodologique développé dans la partie précédente a été adapté pour construire un modèle dynamique de production de biomasse et l'ajuster sur les données de la pêcherie Mauritanienne.

Dans un premier temps, le modèle de production de biomasse a été ajusté sur les données sans prendre en compte les covariables environnementales. Dans un deuxième temps, nous avons cherché à introduire explicitement l'effet des indices d'upwelling dans le modèle.

2.6.1 Modèle de production de Biomasse dynamique sans effet de l'environnement

Cette application reprend les mêmes hypothèses de modèle que les travaux de simulation (équation dynamique, modèle *SSM*, forme des aléas log-Normale, $B_1=K$, $\lambda=1$, priors non informatifs) à l'exception du prior choisis sur le paramètre r . Pour ce paramètre, un prior de type gamma centré sur 1.2 (Balguerias dans IMROP 2006) avec une variance de 10 a été choisi : $r \sim \text{gamma}(1.46 ; 0.38)$. Ce dernier est donc moins informatif que le prior utilisé dans les simulations.

2.6.2 Effet de l'environnement dans le modèle de production de biomasse

2.6.2.1 Corrélation entre les résidus du modèle et les indices environnementaux

La démarche mise en œuvre dans cette étude a consisté dans un premier temps à rechercher des corrélations entre les indices environnementaux et les résidus de l'estimation obtenus dans l'ajustement du MPB sans effet de l'environnement.

Selon Fréon (1991), l'environnement peut agir de deux façon dans un modèle de production halieutique : soit il agit directement sur la production du stock, soit il agit sur la capturabilité des individus. Dans notre cas, nous avons privilégié la première hypothèse, car le terme de capturabilité évoqué dans notre modèle est sans unité, ni sens technique (voir 2.2.3.).

Dans cette hypothèse, dans le modèle de production de biomasse utilisé jusque là, l'effet de l'environnement, s'il existe, est intégré implicitement dans le terme d'erreur de processus (cf. eq. 3 et 4). De ce fait, les effets potentiels de l'environnement peuvent être renvoyés dans les résidus à l'issue de l'ajustement.

D'après les connaissances sur le cycle de vie des poulpes Mauritanien, les juvéniles qui naissent entre mai-juillet de l'année t constitueront le futur stock d'adultes de l'année $t+1$. L'hypothèse formulée est que des conditions environnementales favorables au cours de l'année t , et particulièrement au cours de la période clé de pré-recrutement de mai-juillet de l'année t , seraient susceptibles de favoriser une forte production de poulpe au cours de l'année t .

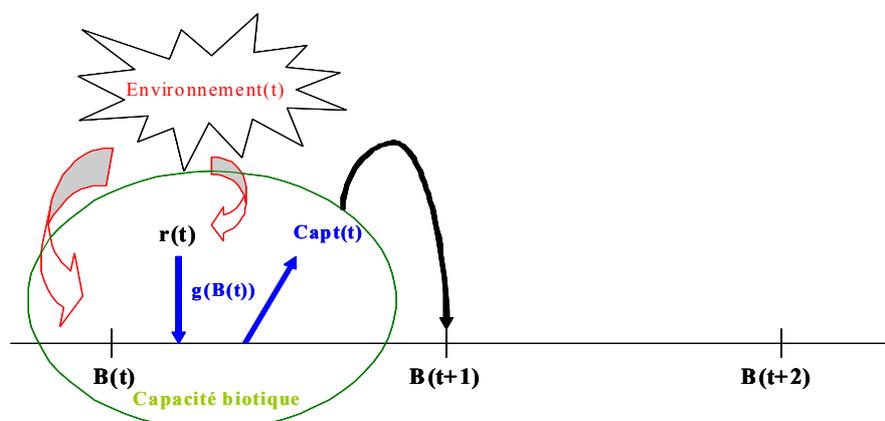


Figure 10 : Influence de l'environnement sur la production de biomasse

Dans ce contexte, les conditions environnementales de l'année t seraient susceptibles d'influencer la biomasse de l'année t+1 et pourraient donc être corrélées avec les résidus de Biomasse de l'année t+1, Res_B_{t+1} , qui s'expriment comme la différence entre la biomasse de l'année t+1 et la biomasse « prédite » de l'année t+1 en fonction de l'équation déterministe.

$$(15) \quad Res_B_{t+1} = (B_t + r \cdot B_t \cdot (1 - \frac{\ln(B_t)}{\ln(K)})) \cdot e^{\varepsilon_{I(t)}} - (B_t + r \cdot B_t \cdot (1 - \frac{\ln(B_t)}{\ln(K)}))$$

La première étape de la démarche de construction d'un modèle avec effet de l'environnement a donc été de rechercher des corrélations entre les séries de covariables environnementales et la série de résidus de biomasse calculée comme dans l'eq (9). Cette recherche a été réalisée pour les deux séries d'indice d'upwelling IUC_1 et IUC_2 ainsi que la série de SST moyen (moyenne géométrique entre mai-juillet) SST.

2.6.2.2 Intégration des indices d'environnement dans le modèle sous forme paramétrique

Dans l'hypothèse où l'environnement agit sur la fonction de production, son action peut être modélisée à plusieurs niveaux : soit sur le paramètre K (la capacité biotique) uniquement, soit sur r (le taux de croissance intrinsèque), soit sur les deux.

La démarche adoptée a consisté à introduire l'effet de l'environnement sur r. En effet, l'environnement semble influencer les premières phases de vie du poulpe en agissant plutôt sur la survie des larves (Faure *et al.* 2000). De plus, l'analyse statistique doit être conduite dans un souci de parcimonie qui conduit à chercher des modèles simples.

En s'inspirant des travaux de plusieurs auteurs, plusieurs types de relations paramétriques ont été testées pour intégrer l'effet de l'environnement dans le paramètre r, sous la forme d'un effet linéaire, soit sous la forme d'un effet puissance (Cury *et al.* 1987 ; Fréon 1991 ; Jacobson *et al.* 2005 ; Gascuel et Inejih dans IMROP 2002).

En profitant de la souplesse du cadre de modélisation Bayésien, quatre modèles différents ont pu être testés et comparés avec le modèle sans environnement (M1):

- $r = \text{constante}$ modèle M1
- $r^2_t = b \cdot IUC_t$ modèle M2
- $r^2_t = IUC_t^b$ modèle M3
- $r^2_t = a + b \cdot IUC_t$ modèle M4
- $r^2_t = a \cdot IUC_t^b$ modèle M5

Afin d'intégrer ces différentes formes d'influence de l'environnement dans le modèle Bayésien, un prior de type gamma a été choisi sur r (comme précédemment) :

$$(16) \quad r_t \sim \text{gamma}(\text{mode}_t, \text{variance})$$

où l'effet de l'environnement est modélisé au niveau du mode de la distribution gamma :

$$(17) \quad \begin{aligned} M2: \quad \text{mod } e_t &= b \cdot IUC_t \\ M3: \quad \text{mod } e_t &= IUC_t^b \\ M4: \quad \text{mod } e_t &= a + b \cdot IUC_t \\ M5: \quad \text{mod } e_t &= a \cdot IUC_t^b \end{aligned}$$

La variance a été fixée à 10 pour ces 4 modèles M2, M3, M4, M5.

Des lois uniformes assez restreintes ont été choisies sur les hyper-paramètres a et b , en choisissant les bornes de telle sorte que le prior généré sur les différentes valeurs de r soient réalistes en fonction des valeurs des indices environnementaux.

$$(18) \quad \begin{aligned} M2 : & b_2 \sim \text{unif}(0,3) \\ M3 : & b_3 \sim \text{unif}(-1,2) \\ M4 : & a_4 \sim \text{unif}(0,2) \text{ et } b_4 \sim \text{unif}(0,1.5) \\ M5 : & a_5 \sim \text{unif}(0,3) \text{ et } b_5 \sim \text{unif}(-1,3) \end{aligned}$$

Une comparaison de ces 4 modèles a ensuite été réalisée en prenant comme critères de sélection : le DIC (Deviance Inference Criterion) (Spiegelhalter *et al.* 2002) et le facteur de Bayes (Kass et Raftery 1995) (cf. ANNEXE 5).

3 Résultats

3.1 Evaluation de la méthode par une approche de simulation

3.1.1 Simulation MCMC, convergence, corrélation entre les paramètres et sensibilité aux priors

3.1.1.1 Simulation MCMC, convergence

Le modèle est complexe et l'information apportée par les données (IA et C^{obs}_t) reste relativement pauvre. Ainsi, la convergence des chaînes MCMC vers la loi a posteriori est longue à obtenir (cf. ANNEXE 4 pour les détails relatifs aux tests de convergence)

Ce travail de simulation a nécessité un temps de calcul considérable. Les inférences sont réalisées sur 10 trajectoires différentes par scénario et chaque simulation a nécessité un temps de calcul d'environ 1h30 (Pentium IV, 3.00Ghz) et occupe 40Mo d'espace disque.

3.1.1.2 Corrélation entre les paramètres du modèle

Il existe une très forte corrélation entre les paramètres du modèle, illustrée ci-dessous par la corrélation a posteriori entre (q, K) et (r, K) (Fig 11).

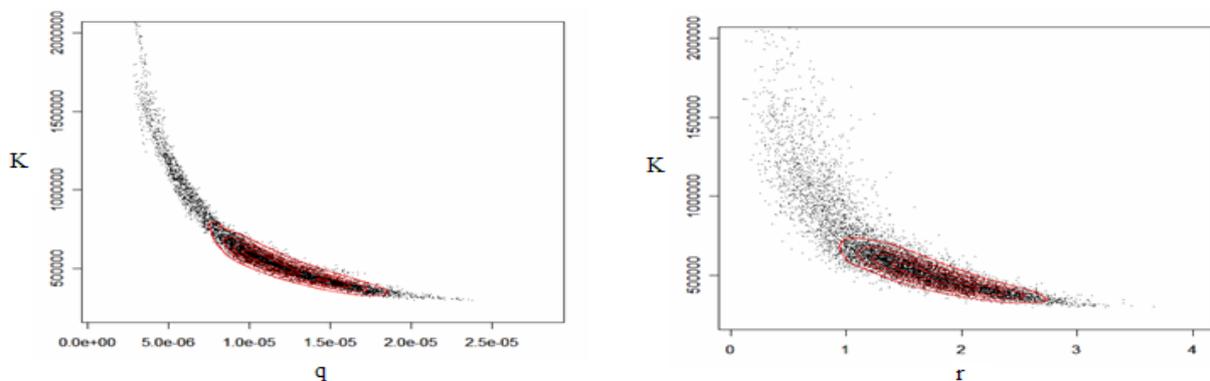
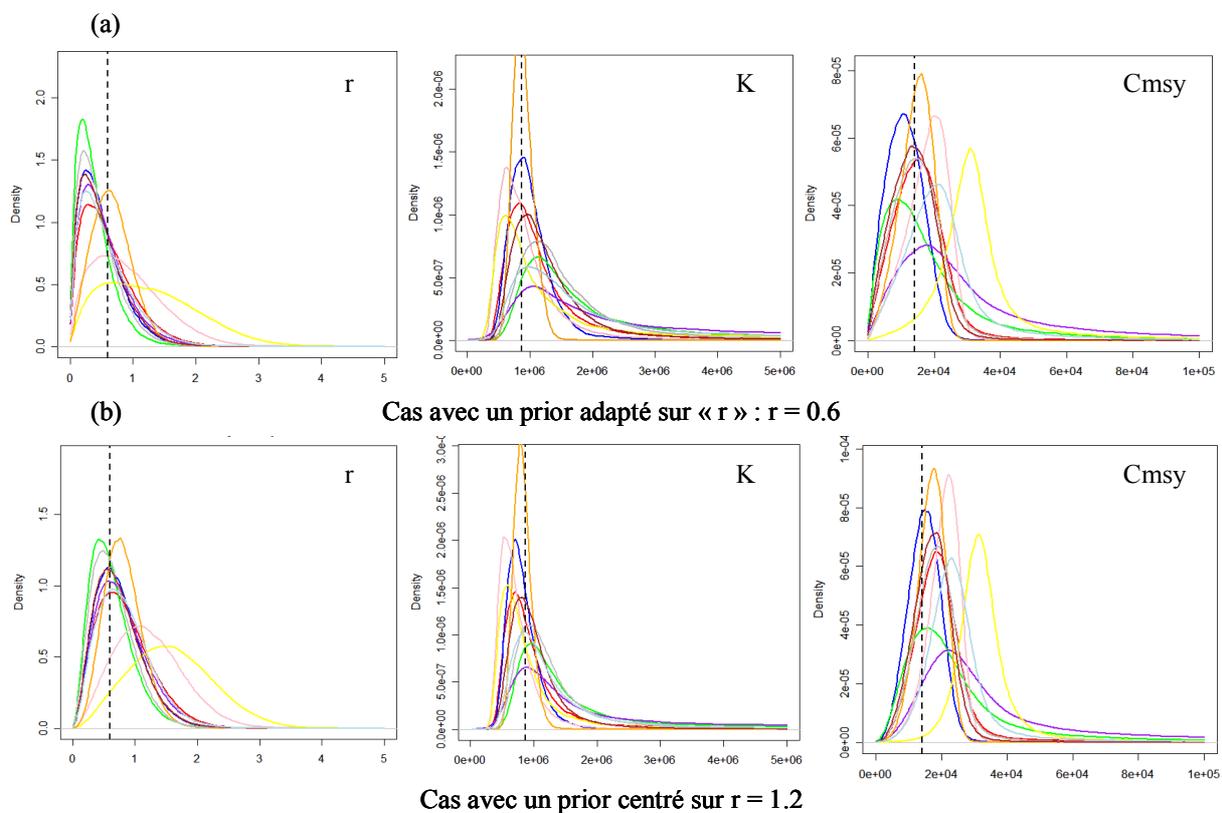


Figure 11 : Tirages MCMC joints pour les paramètres (q, K) et (r, K)

Cela constitue un signe de manque d'information dans les données. En effet, les séries d'indice d'abondances du type « *one way trip* » sont peu informatives. Malheureusement, cette situation de « *one way trip* » est très fréquente lorsque l'intensité de l'exploitation de la ressource est maintenue à un haut niveau, ce qui est le cas pour le stock de poulpe Mauritanie.

3.1.1.3 Sensibilité aux priors

Cette analyse montre que le modèle est effectivement sensible au prior sur r (Fig. 12). Les distributions a posteriori de r sont biaisées entre le cas d'ajustement avec un prior adapté (mode du prior correspondant à la valeur « vraie » du paramètre) et le cas d'ajustement avec un prior centré sur 1.2 (décalage des modes de distribution). En revanche, la sensibilité des inférences sur les autres paramètres comme K , σ_p^2 , les biomasses ou C_{msy} semble être moins marquée et se manifeste surtout par une augmentation des incertitudes.



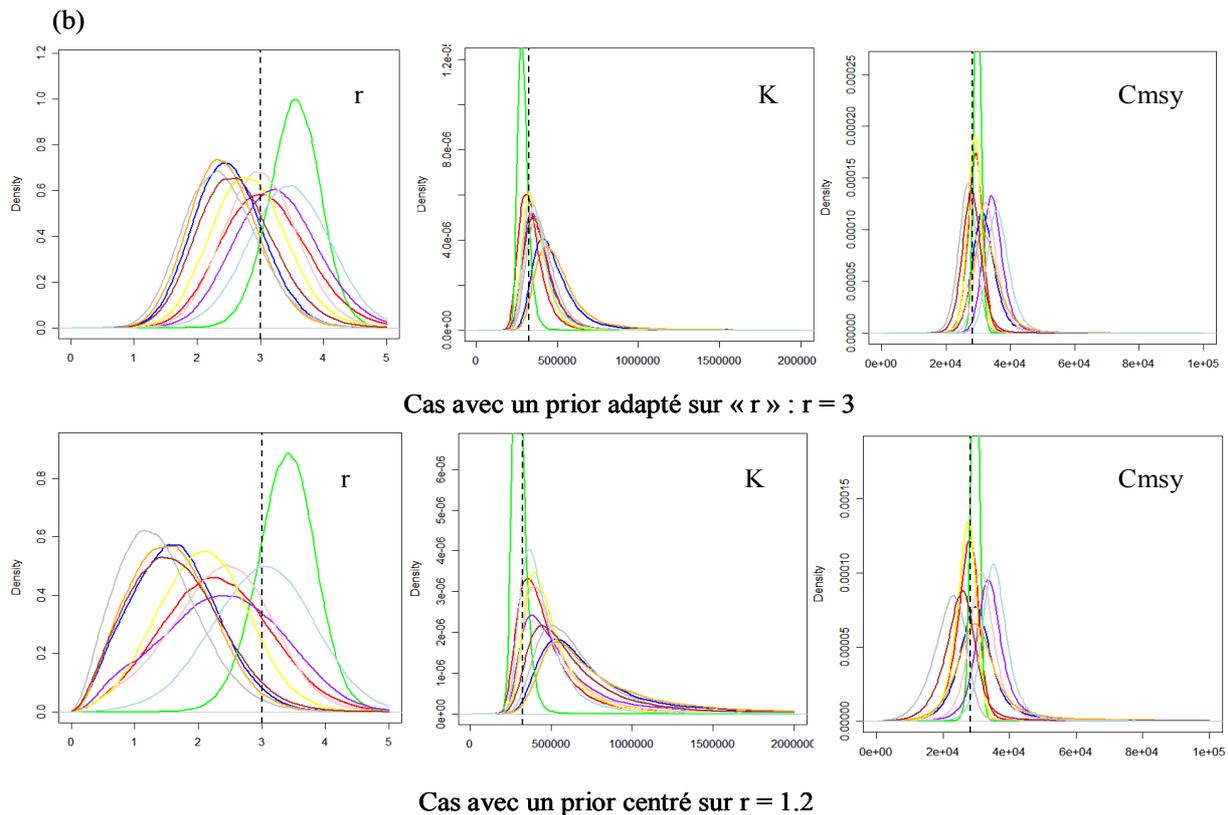


Figure 12 : Test de sensibilité aux priors. Comparaison des distributions a posteriori des paramètres d'intérêt r , K , et C_{msy} pour un ajustement avec un prior adapté et un prior centré sur 1.2. (a) : Cas du scénario 4 ; (b) : Cas du scénario 7

Bilan : simulations MCMC, convergence et sensibilité aux priors

Les résultats montrent que la mise en œuvre pratique de l'ajustement des modèles dynamique de production de biomasse dans un cadre de *SSM* Bayésien peut être délicate.

Les simulations MCMC demandent un temps de calcul important et la convergence des chaînes est longue à obtenir.

Les inférences nécessitent de définir des lois a priori informatives au moins sur un des deux paramètres r , K (r a été choisi ici). Les inférences sont sensibles au choix des distributions a priori.

3.1.2 Sensibilité des ajustements à l'information contenues dans les données

3.1.2.1 Niveau de réduction du stock $B_{n=35}/K$

La figure 13 illustre l'influence du niveau de réduction du stock sur la qualité des inférences sur les paramètres r , K et C_{msy} dans le cas des scénarii 1 et 3.

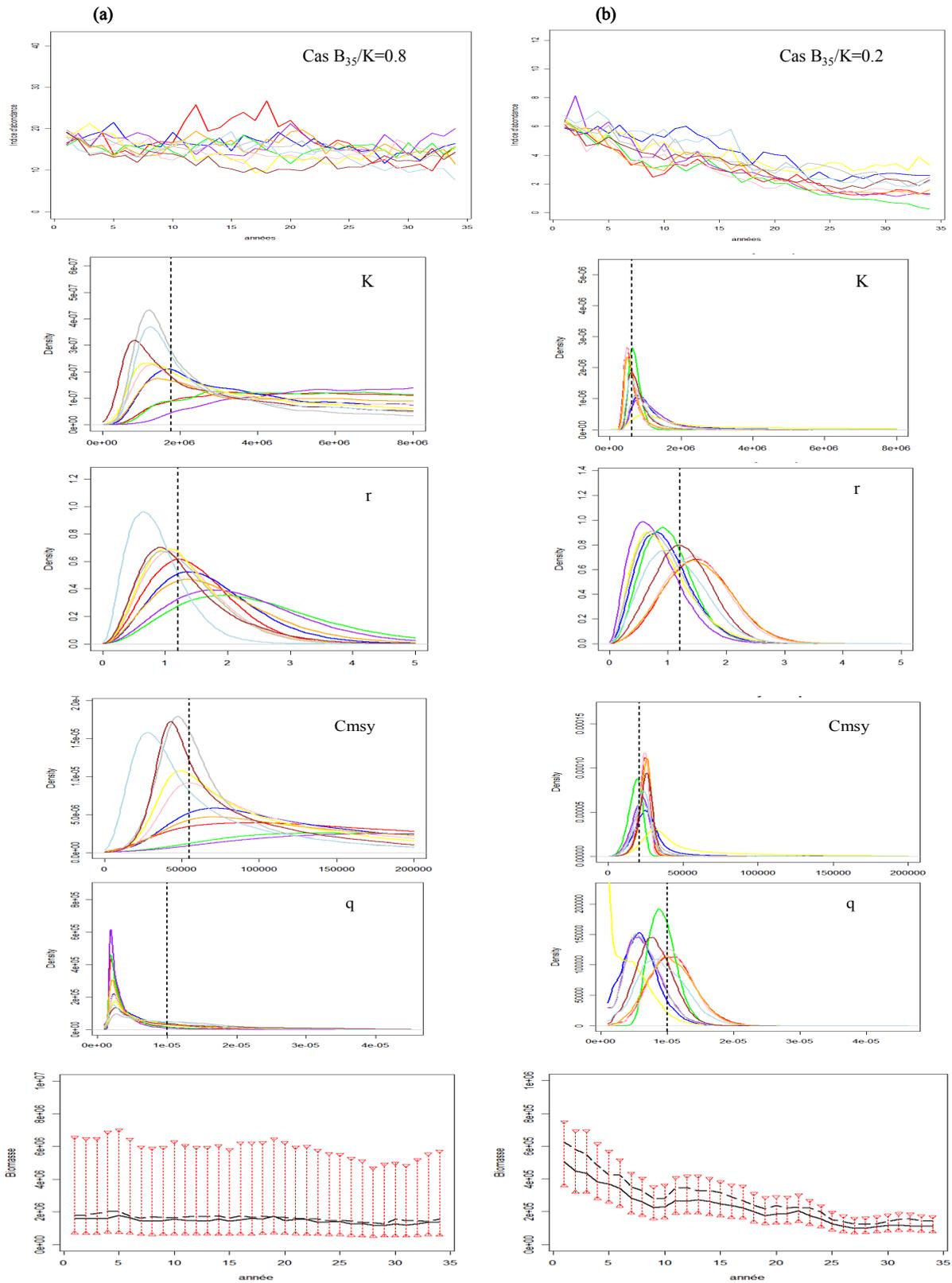


Figure 13 : Influence du niveau de déplétion des séries d'indice d'abondance sur les inférences a posteriori des paramètres d'intérêt r , K , C_{msy} , q , et les biomasses (faute de place, seule une trajectoire de biomasse parmi les 10 simulations est représentée). Le mode (trait plein), la vraie valeur (tiret long), et l'intervalle de crédibilité Bayésien à 75% (pointillé) (a) : Scénario 3 ; (b) : Scénario 1.

La figure 13, complétée par la quantification des écarts quadratiques moyens et écarts relatifs moyens donnés à la figure 15, montre que plus le niveau de réduction de stock est élevé, plus les données sont « informatives », moins l'incertitude et le biais sur les estimations des paramètres r , K , C_{msy} et q sont importants et moins les inférences sont impactées par la forme du prior (notamment inférences sur q). Aussi, l'incertitude autour de la trajectoire des biomasses reconstituées est d'autant plus grande que les données sont peu informatives (Fig 13).

Ces résultats sont généralisables à l'ensemble des scénarii et des paramètres testés (cf. figure 15 ci-après).

3.1.2.2 Niveaux de variance

La figure 14 montre l'influence du niveau de bruit dans le processus et dans les données (valeurs de σ^2_p et σ^2_{IA}) sur la qualité des ajustements dans le cas de la comparaison des scénarii 10 (variance faible), 1 (variance intermédiaire), et 11 (variance forte). Plus le niveau de bruit dans le processus et dans les données est grand, plus les distributions a posteriori des paramètres r , K , et C_{msy} et q sont dispersées et biaisées et plus l'incertitude autour de l'estimation des biomasses est grande. Ces illustrations « visuelles » sont confirmées par la quantification des écarts quadratiques moyens et écarts relatifs moyens donnés à la figure 15.

La comparaison du scénario 1 ($\lambda=1$) avec les scénarii 12 ($\lambda=0.1$, faible variance de processus) et 13 ($\lambda=10$, faible variance d'observation) montre que la réduction d'une des deux sources de variance (processus ou observation) par rapport au cas $\lambda=1$ induit une augmentation de la précision et une diminution du biais dans les estimations (cf. figure 15 ci-après).

Ces résultats sont également généralisables à tous les scénarii et pour tous les paramètres.

Bilan : Sensibilité des ajustements à la quantité d'information contenue dans les données

Ces différentes analyses de sensibilité montrent que le niveau de biais et d'incertitudes (surtout) des ajustements obtenus dans la configuration du MPB dynamique traité comme un *SSM* augmentent fortement lorsque la quantité d'information contenue dans les données diminue (mesurée ici par le niveau de bruit dans le processus ou dans les données et par le niveau de réduction du stock). Moins les données sont informatives, plus les distributions a posteriori sont proches des lois a priori.

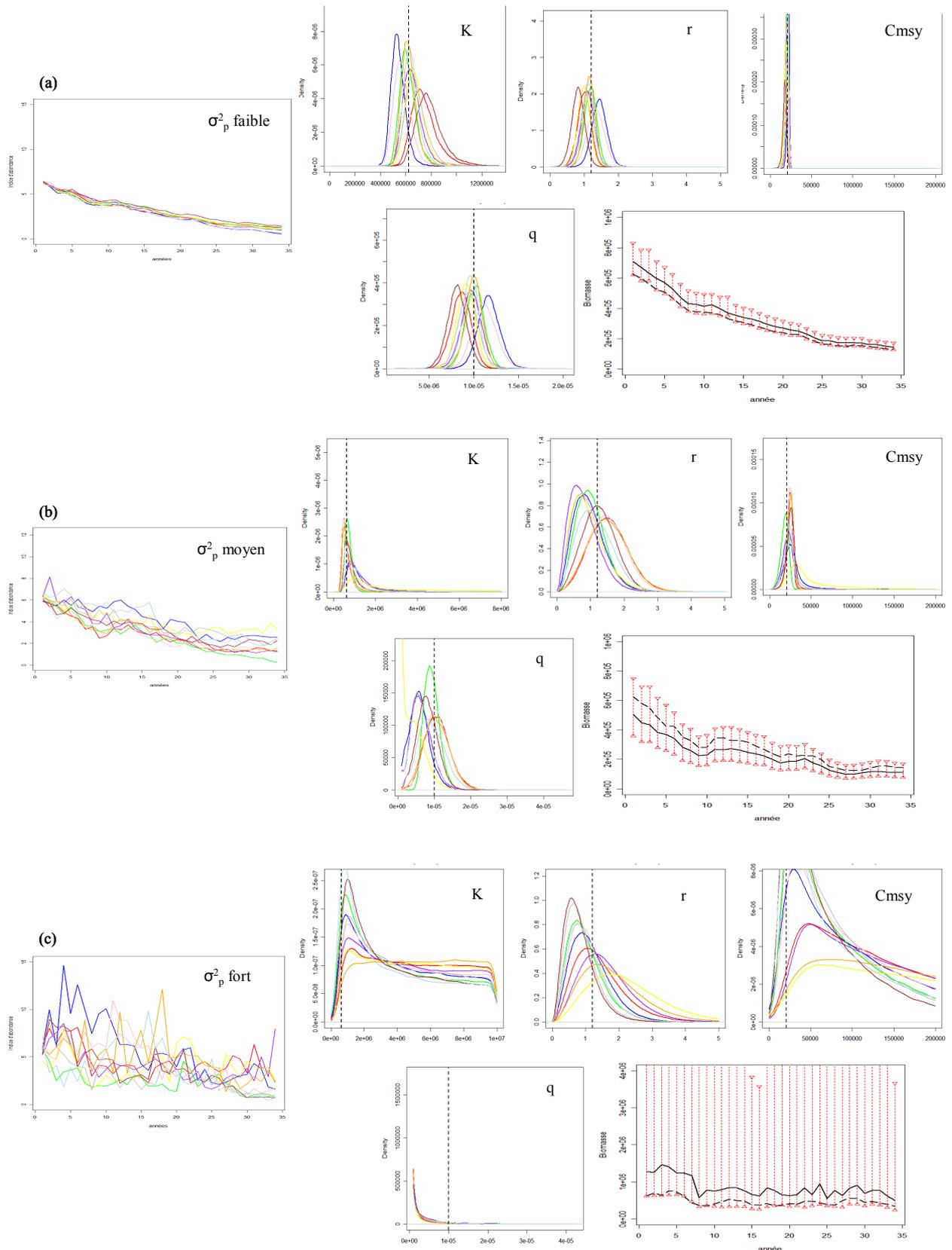


Figure 14 : Influence du niveau de variance sur les distributions a posteriori des paramètres K , r , et C_{msy} , q et sur les biomasses (faute de place, seule une trajectoire de biomasse parmi les 10 simulations est représentée) Le mode (trait plein), la vraie valeur (tiret long), et l'intervalle de crédibilité Bayésien à 75% (pointillé) (a) : Scénario 10 ; (b) : Scénario 1 ; (c) : Scénario 11.

3.1.3 Influence de l'hypothèse d'équilibre

La qualité (biais et incertitudes) des inférences obtenues dans les deux configurations *i)* SSM dynamique et *ii)* modèle de production ajusté sous l'hypothèse d'équilibre ont été étudiés au travers de trois critères : les IC Bayésiens (Annexe 6) et les valeurs des écarts quadratiques moyens (\overline{EQM}) et relatifs (\overline{ERM}) (Fig. 15).

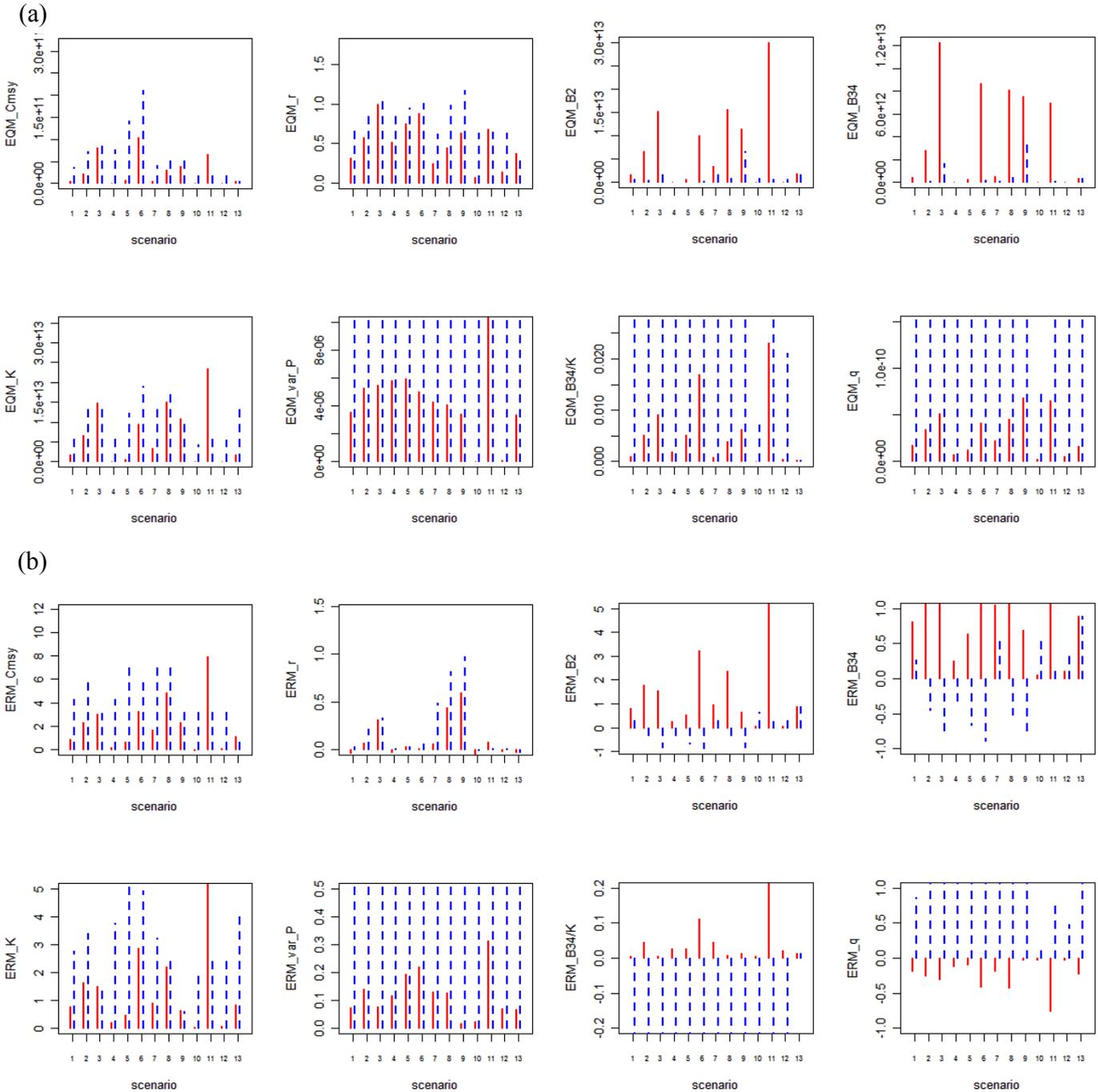


Figure 15 : Comparaison des \overline{EQM} (Pannel a) et \overline{ERM} (Pannel b) des 8 paramètres d'intérêt, pour les 13 scénarii, entre l'ajustement dynamique (barre pleine) et l'ajustement sous l'hypothèse d'équilibre (barre pointillée).

La figure 15 montre que pour la majorité des paramètres (sauf pour les biomasses) et pour la majorité des scénarii (sauf le scénario 11) l'ajustement en *SSM* dynamique donne des estimations de meilleure qualité que celles obtenues sous l'hypothèse d'équilibre. En effet, les valeurs des écarts quadratiques moyens \overline{EQM} sont dans la majorité des cas bien plus faibles dans la configuration *SSM* dynamique, et les biais sont plus réduits comme le montre la valeurs des écarts relatifs moyens \overline{ERM} .

Ces diagnostics sont confirmés par l'examen des IC Bayesiens (Annexe 6). Pour les paramètres r , K , C_{msy} , σ_p^2 , les biomasses, et la valeur de la réduction de biomasse (B_{34}/K) et dans tous les scénarii, la configuration *SSM* dynamique donne en général des IC Bayesiens qui contiennent souvent la « vraie » valeur des paramètres.

En revanche, même dans la configuration dynamique, le paramètre q est souvent très mal estimé. Les distributions a posteriori sont souvent non informatives (posterior semblable au prior). Les inférences sont souvent très dispersées et biaisées négativement (sous-estimation des modes). Cela se répercute sur les estimations de K et des biomasses qui s'en retrouvent logiquement surestimées. Ce résultat est à nuancer par l'examen des résultats des scénarii 1, 4, 10 et 12 qui donnent des inférences de bonne qualité pour le paramètre q . Cela est à relier au fait que ces scénarii correspondent à des données informatives (forte réduction de la biomasse et variances modérées).

L'ajustement à l'équilibre quant à lui, donne également des résultats fiables en termes d'IC Bayésien pour les paramètres « K », et « r ». Par contre, une étude combinée de l'IC Bayésien avec l' \overline{ERM} montre que les estimations du paramètre q sont souvent biaisées positivement, et les biomasses s'en retrouvent biaisées négativement dans tous les scénarii (Pannel (b) de la fig 15). Plus la valeur de K utilisée pour simuler les données est grande, plus le biais sur les biomasses est grand. La valeur de la réduction de biomasse B_{34}/K est elle, systématiquement sous-estimée.

L'hypothèse d'équilibre induit aussi une surestimation très importante du paramètre de variance du processus σ_p^2 .

Enfin, la surestimation du paramètre de gestion C_{msy} est très marquée sous l'hypothèse d'équilibre. Cela peut avoir des conséquences importantes en termes de gestion.

Bilan de l'influence de l'hypothèse d'équilibre

Cette étude comparative permet de montrer que l'ajustement d'un modèle de production sous l'hypothèse d'équilibre à partir de données générées par un modèle dynamique donne des inférences qui, en moyenne, sont de moins bonne qualité (en termes de biais et d'incertitudes) que celles obtenues lorsque les relations dynamiques entre les biomasses sont maintenues dans le modèle statistique.

Les biais induits par l'hypothèse d'équilibre sont particulièrement marqués pour le C_{msy} , la variance de processus et les biomasses.

Ces résultats présentent ainsi les différents risques occasionnés par l'utilisation de l'hypothèse d'équilibre dans le cas où les données analysées sont effectivement générées par un modèle dynamique. Cela souligne le fait que la bonne identification du fonctionnement du système (dynamique ou à l'équilibre) est cruciale pour une bonne évaluation de l'état du stock.

Il faut rappeler cependant que l'hypothèse d'un fonctionnement à l'équilibre permet de s'affranchir de l'hypothèse $B_1=K$, ce qui présente un intérêt majeur lorsque cette hypothèse est très incertaine.

3.1.4 Comparaison des ajustements entre *SSM*, et *observation error model* ajusté en Bayésien et en Maximum de Vraisemblance

		EQMP	ERMP	EQM	ERM	IC bayésien à 75%
C_{msy}	SSM			1.03E+08	7.05E-02	8
	obs error	2.14E+07	1.38E-01	2.06E+08	2.13E-01	4
	max vrais	2.70E+07	1.40E-01			
r	SSM			1.43E-01	-2.14E-02	8
	obs error	3.40E-01	2.21E-01	4.41E-01	1.97E-01	5
	max vrais	6.45E-01	3.79E-01			
K	SSM			2.62E+10	0.07e-02	8
	obs error	3.70E+10	-2.72E-02	5.85e+10	5.04e-02	4
	max vrais	5.42E+10	-4.73E-02			
q	SSM			5.65E-12	-2.73E-02	8
	obs error	9.51E-12	7.24E-02	1.55E-11	8.22E-02	7
	max vrais	2.11E-11	2.33E-01			
B₂	SSM			2.59E+10	7.44E-02	9
	obs error	4.16E+10	-8.46E-02	5.94E+10	4.86E-02	4
	max vrais	5.37E+10	-5.86E-02			
B₃₄	SSM			2.42E+09	1.04E-01	8
	obs error	1.45E+10	-7.11E-02	1.07E+10	4.86E-02	5
	max vrais	5.32E+09	-1.12E-01			
B₃₄/K	SSM			3.67E-04	2.05E-02	8
	obs error	1.69E-02	-1.08E-01	4.02E-03	5.70E-02	5
	max vrais	1.90E-03	-5.72E-02			
σ_{IA}	SSM			5.22E-08	6.91E-02	6
	obs error	5.89E-05	1.44E+00	1.33E-03	2.53	0
	max vrais					

Tableau 2 : Critères synthétiques de qualité d'ajustement des paramètres obtenus à partir de 3 méthodes statistiques différentes dans le cas du scénario 12.

Le tableau 2 récapitule les résultats des critères synthétiques de qualité d'ajustement obtenus à partir de trois méthodes statistiques différentes dans le cas du scénario 12 (données simulées dans le cas d'un *SSM* mais avec une variance de processus faible donc se rapprochant d'un *observation error model*).

L'ajustement en *SSM* donne des résultats les plus fiables et les moins biaisés pour les paramètres C_{msy} , r , K , q , et σ_p^2 (Tableau 2).

L'*observation error model* ajusté dans le cadre statistique classique du Maximum de Vraisemblance donne des résultats les plus biaisés pour les paramètres C_{msy} , K , r au vue de la valeur d' $EQMP$.

Bilan

Les résultats confirment l'idée que lorsque les données sont générées avec du bruit dans le processus et dans les observations, l'ajustement dans le cadre des modèles stochastique d'état *SSM* donne en moyenne de meilleurs résultats que l'ajustement sous l'hypothèse d'un modèle contenant uniquement des erreurs d'observation.

Par défaut, le traitement statistique dans le cadre des *SSM* est donc recommandé (Punt 2003 ; Meyer et Millar 1999).

3.2 Evaluation du stock de poulpe Mauritanie

Dans cette partie, on s'appuie sur le cadre méthodologique du traitement Bayésien du *SSM* développé dans la partie précédente pour analyser les données de la pêche Mauritanienne, avec ou sans effet de l'environnement. Les données utilisées sont les séries d'indice d'abondance et de captures totales disponibles sur la période 1971-2005, ainsi que les séries d'indice d'upwelling de 1971-2003.

3.2.1 Modèle de production de biomasse sans effet de l'environnement

3.2.1.1 Cas en dynamique

Les figures 16 (a) et 17 (a) donnent les distributions a posteriori des principaux paramètres du modèle obtenus dans le cadre d'un *SSM* dynamique.

L'estimation du paramètre r est à interpréter avec précaution car les inférences sur ce paramètre sont sensibles au prior informatif choisis pour r (cf. 3.1.1.3). En revanche, les paramètres K et C_{msy} peuvent être interprétés avec plus de sérénité. En effet, les résultats de l'étude de simulation/ajustement montrent que les inférences sur ces paramètres sont généralement assez peu biaisées et moins sensibles au choix du prior sur r .

Le mode de la distribution a posteriori de r (l'estimation la plus probable) se situe autour de 0.4. La distribution a posteriori du paramètre K reste assez lourde et l'influence du prior uniforme se fait ressentir. Le mode de la distribution de K est néanmoins bien marqué et il se situe autour de 700000 tonnes. La variance du processus, σ_p^2 est estimée à près de 0.02 (variance assez forte par rapport aux scénarii testés dans la première partie).

La loi a posteriori du paramètre de gestion C_{msy} est riche d'enseignement. Il apparaît que le mode de la distribution se situe autour de 30000 tonnes. Etant donnée une valeur seuil fixée arbitrairement, notée x , la distribution a posteriori permet de calculer directement la probabilité a posteriori $p(C_{msy} < x \mid \text{Données})$. Dans le cas où la valeur x est un quotas décidé par le gestionnaire, $p(C_{msy} < x \mid \text{Données})$ s'interprète comme un risque.

On peut calculer ce risque pour deux valeurs seuil $x = 20000$ tonnes et $x = 30000$ tonnes. Pour un quota fixé à 30000 tonnes, il y aurait 44.7% de risque que la « vraie » valeur de C_{msy} soit en dessous de ce quota. Si le quota est à 20000 tonnes, ce risque diminue jusqu'à 22.3%.

La figure 17 (a) retrace la trajectoire de la biomasse estimée entre 1971 et 2004. L'incertitude autour de l'estimation de la trajectoire de biomasse est grande mais logiquement, cette trajectoire retrace la forme de la trajectoire de l'indice d'abondance. L'estimation du paramètre B_{34}/K montre que le niveau de biomasse a diminué de près de 4/5 en 34 ans (Fig. 16 (a)).

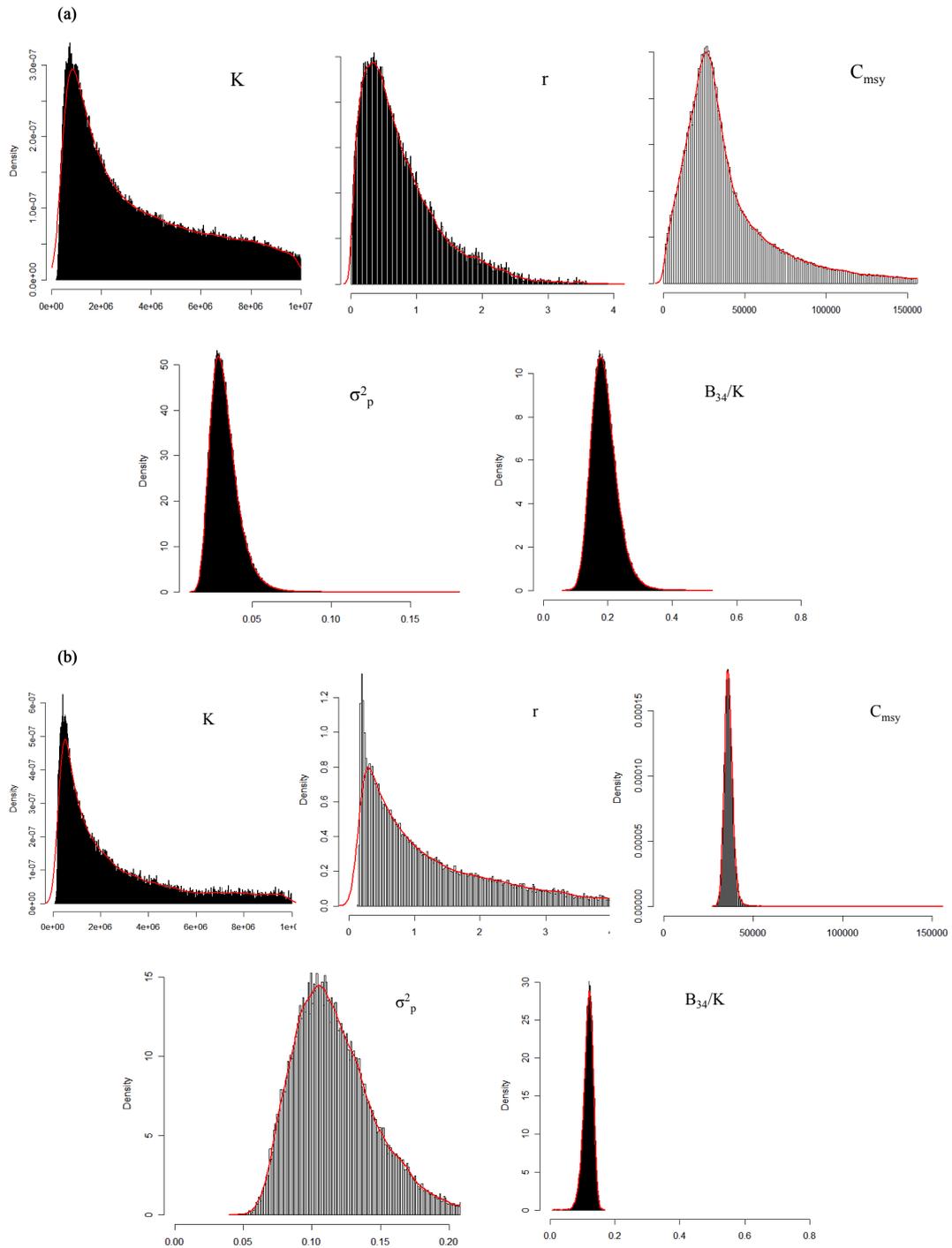


Figure 16 : Distribution a posteriori des paramètres K , r , C_{msy} , σ_p^2 , et B_{34}/K du modèle. (a) : ajustement dynamique ; (b) : ajustement sous l'hypothèse d'équilibre

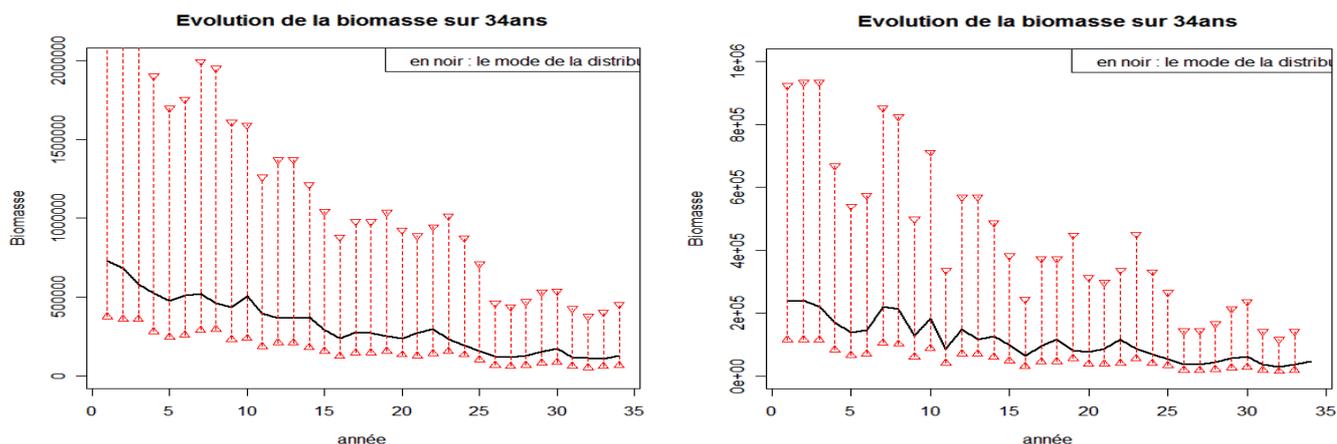


Figure 17 : Reconstruction de la trajectoire de biomasse entre 1971 et 2004.
 Trait plein : modes de la distribution a posteriori des biomasses ;
 Pointillé : intervalles de crédibilité à 50% Bayésien. (a) : ajustement dynamique ; (b) : ajustement sous l'hypothèse d'équilibre

3.2.1.2 A l'équilibre

Ce modèle a été également ajusté sous l'hypothèse d'équilibre. Ce dernier correspond alors au modèle mis en place par les groupes de travail de Mauritanie pour l'évaluation du stock de poulpe. Les résultats d'ajustement sont résumés dans les figures 16 (b) et 17 (b).

Ici aussi, la distribution a posteriori de K reste assez dispersée mais avec un mode bien marqué à 500000 tonnes cette fois (contre 70000 tonnes en dynamique). Le mode de la distribution de r quant à lui se situe autour de 0.3 (contre 0.4 en dynamique). Ce qui conduit à une estimation du mode de C_{msy} à 35000 tonnes. Cette valeur du mode de C_{msy} évaluée par la méthode Bayésienne est très proche de la valeur de C_{msy} déterminée par les groupes de travail de Mauritanie (IMROP 2006) qui estime par maximum de vraisemblance le paramètre à 34862 tonnes. Mais cette valeur est supérieure à celle estimée précédemment avec le *SSM* en dynamique (autour de 30000 tonnes) et surtout, l'incertitude autour de son estimation est beaucoup plus faible. Cette dernière constatation rejoint l'analyse qui a été faite dans le paragraphe 3.1.3.2.

Les inférences sur B_{34}/K suggèrent que la biomasse aurait diminué de plus de 4/5 en 34 ans (Fig. 17 (b)).

Bilan :

Ces résultats suggèrent l'idée que le stock de poulpe est aujourd'hui largement surexploité en Mauritanie. Le niveau de biomasse aurait baissé de près de 3/4 entre 1971-2004 et le C_{msy} est estimé à environ 30000 tonnes pour le modèle dynamique et à 35000 tonnes pour le modèle à l'équilibre. Ces résultats sont cohérents avec les estimations de Chassot et al. (2006) et Gascuel et Inejih (2006).

Les ajustements Bayésiens de ce modèle de production de biomasse dans un cadre dynamique pour l'un et sous l'hypothèse d'équilibre pour l'autre, conduisent à des inférences sur les paramètres r , K , C_{msy} , σ_p^2 , et B_{34}/K sensiblement différentes. Cela confirme l'idée que le choix du modèle est crucial pour une bonne évaluation des ressources (3.1.3.)

3.2.2 Essai de construction d'un modèle avec effet de l'environnement

3.2.2.1 Séries d'indices environnementaux : IUC₁ et IUC₂ et SST

Les deux méthodes de calcul de l'IUC pour le poulpe de Mauritanie (voir Matériels et Méthodes) donnent des résultats sensiblement différents (Fig. 18). Les deux séries évoluent de façon différente et ne sont pas corrélées.

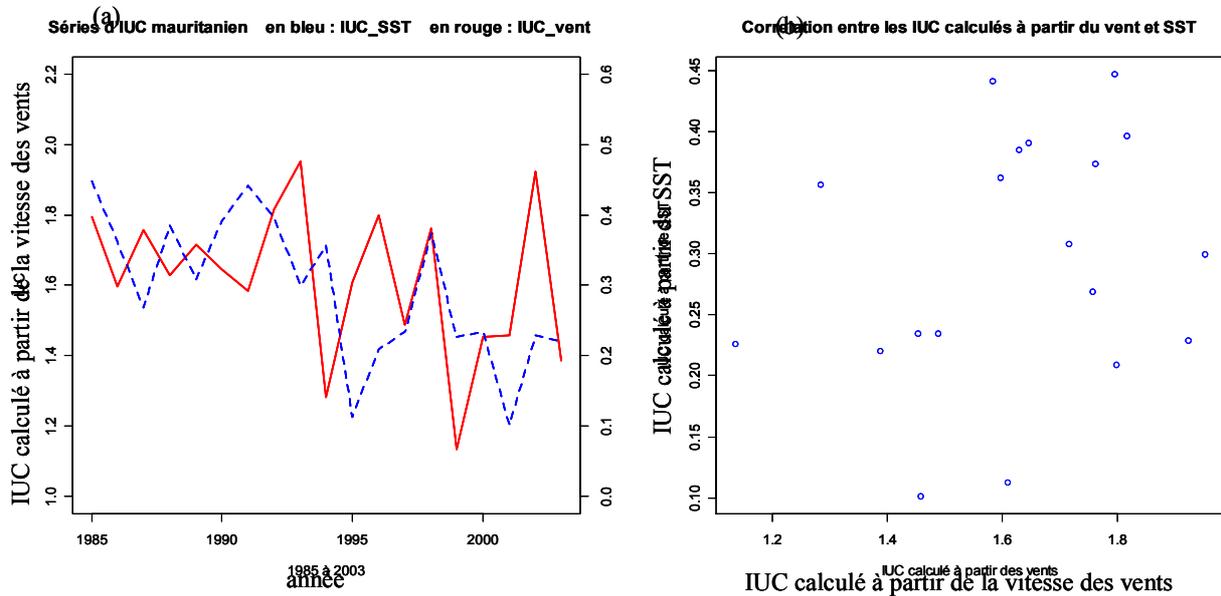


Figure 18 : Comparaison des deux séries d'IUC. (a) : Tracé des séries d'IUC obtenues par les données de vents (trait plein) et par les données de SST (trait pointillé) ; (b) : Corrélation entre ces deux séries d'IUC.

Par ailleurs, les séries d'IUC₁ et de SST sont en anti-phase (Fig. 19). Cela est cohérent avec le fait que l'upwelling correspond à une remontée d'eau froide en surface.

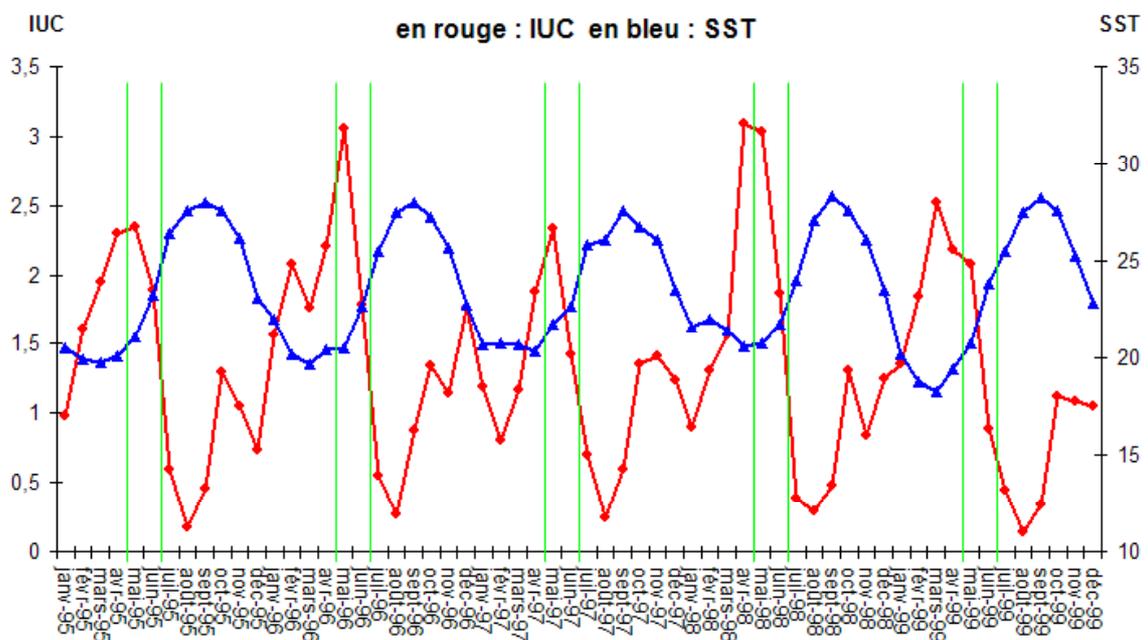


Figure 19 : Variation saisonnière de l'IUC₁ et de la SST entre 1995 et 1999

3.2.2.2 Corrélation entre séries environnementales et résidus du modèle

Les résultats montrent que la série d'IUC calculée à partir des données de vitesse du vent est significativement et positivement corrélée avec les modes des distributions a posteriori des résidus du modèle de production de biomasse dynamique (Fig. 20).

En revanche, aucune corrélation n'est détectée avec la série d'IUC calculée à partir des SST tout comme avec la série des SST (mai-juillet). Ce résultat, bien qu'inexpliqué, est cohérent avec le fait qu'il n'existe pas de corrélation nette entre les deux séries d'IUC ni entre la série d'IUC calculée à partir des vents et celle de SST.

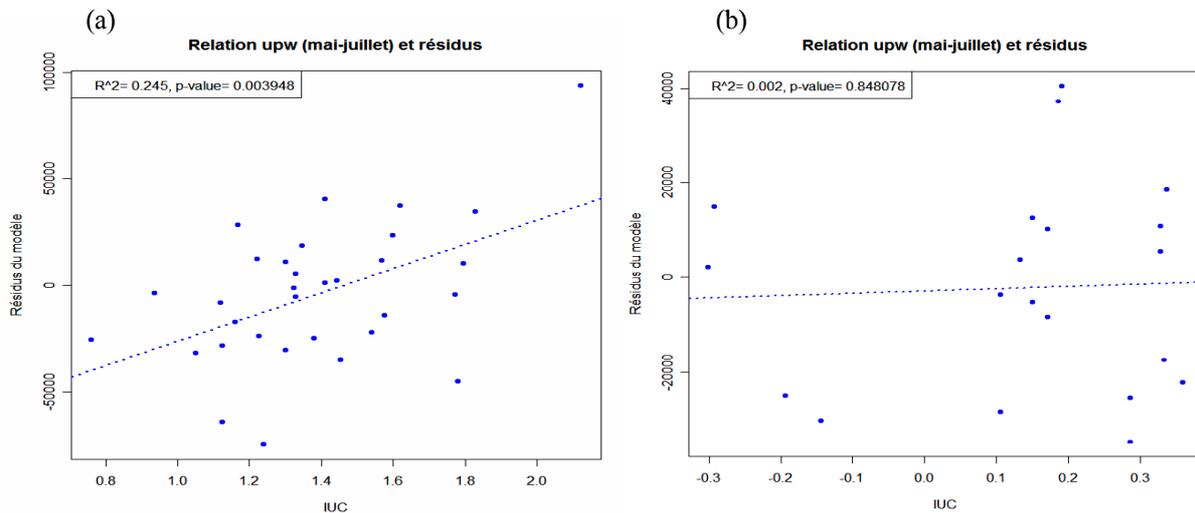


Figure 20 : Corrélation entre les modes des distributions a posteriori des résidus du modèle dynamique de production de biomasse et les séries d'IUC₁ (panel a) (calculée à partir des vents) et IUC₂ (panel b) (calculée à partir de la SST).

Bilan

L'analyse montre que la série d'IUC calculé à partir des vents sur la période mai-juillet, correspondant à la période principal de pré-recrutement, est significativement et positivement corrélée aux résidus du modèle.

En revanche, aucune corrélation significative n'est détectée avec les séries de SST (mai-juillet) et d'IUC calculée à partir des SST.

Cette analyse préliminaire suggère que l'IUC entre mai et juillet, joue effectivement un rôle dans la détermination inter-annuelle de la biomasse du stock de poulpes : la production de biomasse serait d'autant plus forte que l'IUC est important. Mais cette relation semble néanmoins assez tenue.

3.2.2.3 Intégration de l'IUC dans les paramètres du modèle

En se basant sur les résultats de l'analyse préliminaire, on ne considèrera dans la suite que les données d'IUC calculées à partir du vent (IUC₁) pour intégrer l'effet de l'environnement dans le modèle.

3.2.2.3.1 Sélection des différents modèles M2,...,M5

Les valeurs du critère DIC obtenues pour les différents modèles M1 (sans effet de l'environnement) et M2 à M5 (avec effet de l'environnement) montrent que le modèle de

référence sans le paramètre environnemental (M1) semble « meilleur » (au sens de ce critère qui s'interprète comme le critère d'Akaike, *i.e.* comme un compromis entre qualité d'ajustement et parcimonie) que les autres modèles avec effet de l'environnement (Tableau 4).

	DIC	Bayes Factor
M1	-173.1	
M2	-61.32	1.669
M3	-90.7	3.916
M4	-17.28	1.018
M5	-120.0	2.165

Tableau 4 : Tableau récapitulatif des résultats de comparaison de modèles avec DIC

En revanche, une analyse complémentaire à partir du facteur de Bayes (qui s'interprète comme la crédibilité relative des différents modèle conditionnellement aux données) montre qu'aucun des 5 modèles ne semble vraiment « meilleur » qu'un autre. En effet, selon les critères de choix du modèle donnés par Jeffrey (Jeffrey 1961), un Facteur de Bayes n'indique une différence significative que s'il est >5 .

Finalement, les résultats de DIC et de Facteur de Bayes montrent au final que l'intégration de l'effet de l'environnement dans le modèle ne permet pas, d'améliorer le modèle de production de biomasse au sens de ces critères statistiques formels.

Pour autant, la construction d'un modèle avec effet de l'environnement reste conceptuellement intéressante et porteuse de sens biologique. Cette voie ne doit donc pas être écartée uniquement sur la base des critères statistiques formels.

3.2.2.3.2 Estimations des paramètres dans le cadre du modèle avec effet de l'environnement M5

Le modèle M5 est repris ici afin d'estimer les valeurs des paramètres d'intérêt du modèle. Selon les résultats de la sélection de modèle (notamment critère DIC), le modèle M5 apparaît comme le meilleur des modèles avec effet de l'environnement.

Les distributions a posteriori des paramètres K , σ_p^2 , q , sont représentées dans la figure 21. Ces paramètres sont indépendants de l'indice d'upwelling et sont donc constants tout au long de la série.

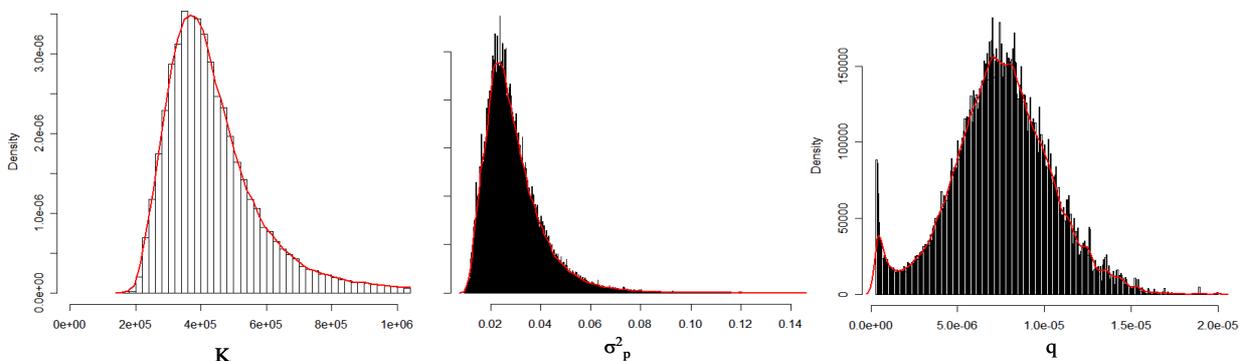


Figure 21 : Distributions a posteriori des paramètres indépendants de l'IUC, K , σ_p^2 et q .

Comparées au cas d'estimation avec le modèle sans effet de l'environnement, les distributions a posteriori sont moins dispersées pour tous les paramètres (sauf r et C_{msy}). Par contre, les estimations du mode de certains paramètres ont changé. C'est le cas de la distribution de K qui montre maintenant un mode autour de 400000 tonnes.

Pour ce qui est de la trajectoire de la biomasse entre 1971 et 2003, elle est estimée avec beaucoup moins d'incertitudes que dans le modèle sans effet de l'environnement (Fig. 22).

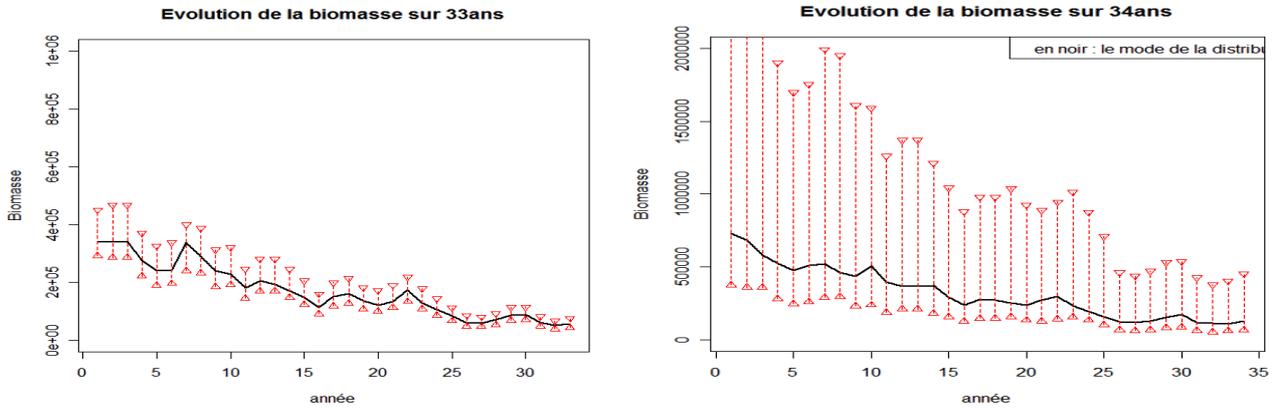


Figure 22 : Reconstruction de la trajectoire de biomasse entre 1971 et 2003. En trait plein : le mode de la distribution a posteriori des biomasses ; en pointillé : les intervalles de crédibilité à 50% Bayésien. (a) : modèle avec environnement ; (b) : modèle sans environnement

Le principal changement apparaît au niveau des estimations des paramètres r et C_{msy} (surtout). En effet, ces paramètres dépendent de l'indice d'upwelling et sont donc évalués chaque année (Fig. 23).

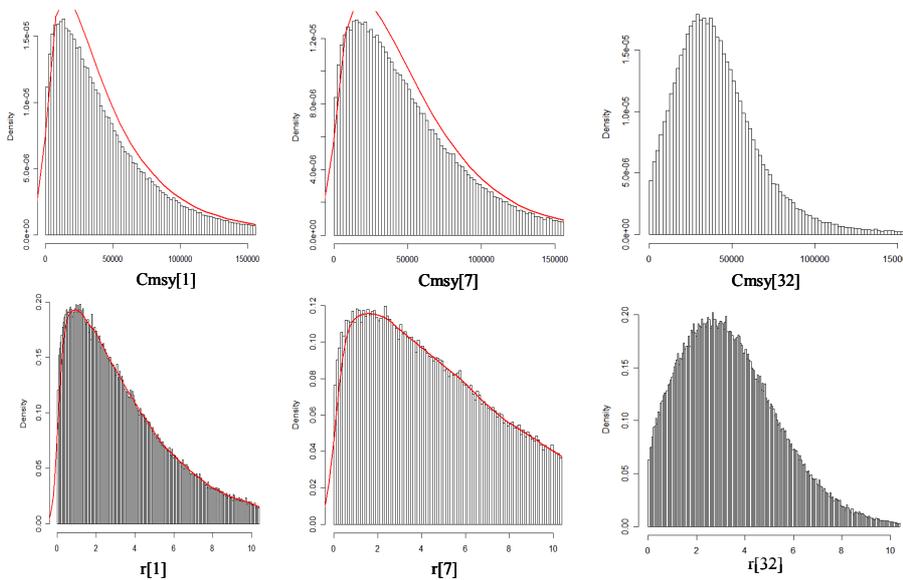


Figure 23 : Distributions a posteriori des paramètres r et C_{msy} pour différentes années 1, 7 et 32.

De manière générale, les résultats de l'inférence montrent que l'on retrouve bien une relation positive entre l'IUC et l'estimation des paramètres r et C_{msy} . Cela conforte l'analyse réalisée sur les résidus du modèle sans effet de l'environnement (cf. 3.2.2.1) (Fig. 24).

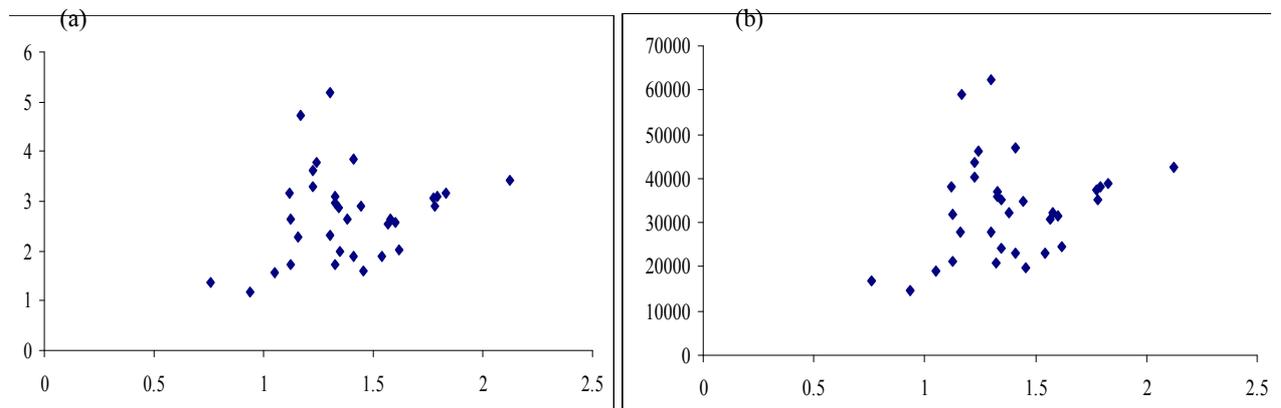


Figure 24 : Relation entre la valeur de l'IUC₁ et les estimations ponctuelles (médiane de la distribution a posteriori) des paramètres r (panel a) et C_{msy} (panel b).

Bilan de l'intégration de l'IUC dans les paramètres du modèle

La tentative de sélection des modèles à l'aide de critères statistiques formels (DIC et Facteur de Bayes) a montré que l'intégration de l'effet de l'environnement sous forme paramétrique dans le paramètre r ne permettait pas d'améliorer significativement le modèle.

Cependant, le modèle avec effet de l'environnement est conceptuellement intéressant. La tentative de paramétrisation réalisée dans le modèle M5 ($r = a \cdot IUC^b$) permet bien de retrouver une relation positive entre l'IUC et les paramètres r et C_{msy} .

4 Discussion et conclusion

4.1 Intérêts et limites du cadre de modélisation *SSM* bayésien

Les Modèles Stochastiques d'Etat (*SSM*) traité dans un cadre statistique Bayésien constituent un cadre de modélisation flexible d'un grand intérêt pour la modélisation quantitative dans les sciences halieutiques. Ce cadre permet notamment de construire et d'ajuster les paramètres de modèles dynamiques en tenant compte de deux sources d'incertitude majeures : les erreurs de processus et les erreurs d'observation (Meyer et Millar 1999 ; Millar et Meyer 2000 ; Harwood et Stockes 2003 ; Buckland et al. 2004).

La partie méthodologique de ce travail a permis d'évaluer la faisabilité de la mise en œuvre de ce cadre de modélisation dans le contexte d'un modèle dynamique de production de biomasse, et d'évaluer la qualité des inférences produites. L'approche simulation-ajustement mise en place dans cette étude représente un travail original. Ce genre d'approche a été déjà largement mis en œuvre dans un cadre de Maximum de Vraisemblance (voir par exemple Punt 2003) mais il a été très rarement réalisé dans un cadre Bayésien (voir un exemple dans Hammond *et al.* 2005).

Les résultats ont permis de montrer que la mise en œuvre pratique de l'ajustement des modèles dynamiques de production de biomasse dans un cadre de *SSM* Bayésien peut être délicate. D'un point de vue technique, les simulations MCMC demandent un temps de calcul important et la convergence des chaînes est longue à obtenir. De ce fait, seules 10 simulations n'ont pu être réalisées par scénario. Par ailleurs, les inférences nécessitent de définir des lois a priori informatives au moins sur un des deux paramètres clé (r , K) et elles se révèlent sensibles au choix de ces distributions a priori. Le paramètre q du modèle d'observation est difficile à estimer et il est quasiment impossible de définir une loi a priori informative pour ce paramètre. Ces difficultés se répercutent au niveau des estimations du paramètre K et des biomasses. En outre, il est nécessaire de poser des hypothèses contraignantes (dans cette étude, $B_1 = K$, $\lambda = \text{fixé}$) pour assurer l'identifiabilité des paramètres. Il convient donc d'être très prudent dans la chaîne de traitement et dans l'interprétation des résultats obtenus.

Au delà de ces limites, les résultats obtenus sont très encourageants et ont permis de montrer que dans l'hypothèse où les données sont générées à partir d'un modèle dynamique comptant à la fois une stochasticité dans le processus et des erreurs de mesure, le respect de ces hypothèses dans le modèle statistique utilisé pour l'ajustement permettait de donner, en moyenne, des inférences de meilleure qualité que d'autres méthodes plus classiques comme un modèle avec erreur d'observation uniquement ou un modèle à l'équilibre.

4.2 Application au cas de la Mauritanie

Le cadre méthodologique développé ci-dessus a été appliqué pour l'évaluation du stock de poulpe mauritanien.

Dans un premier temps, un modèle sans effet de l'environnement a été développé. Ce modèle a permis d'estimer les paramètres clé comme le C_{msy} ou le taux de réduction du stock ainsi que leur incertitude. Le taux de réduction entre le début des années 1970 et 2005 est estimé à plus de 80% et le C_{msy} à environ 30000 tonnes. Ce résultat diffère quelque peu de l'estimation de C_{msy} obtenue dans le cadre d'un ajustement à l'équilibre : 35000 tonnes. On

retrouve ici un résultat général de la partie simulation-ajustement qui montre que les estimations de C_{msy} étaient systématiquement positivement biaisées (surestimées) lorsqu'on ajustait un modèle à l'équilibre. Cette valeur du C_{msy} obtenue sous l'hypothèse d'équilibre est très proche de l'estimation de C_{msy} obtenue à partir d'un modèle à l'équilibre par la méthode du maximum de vraisemblance par le groupe de travail IMROP 2006 (Chassot et al. 2006, Gascuel et Inejih 2006).

Dans un deuxième temps, grâce à la flexibilité du cadre de modélisation statistique Bayésien, une tentative de paramétrisation d'un modèle dynamique avec effet de l'environnement a pu être réalisée. Un travail préliminaire d'étude de la corrélation entre les résidus du modèle sans effet de l'environnement et les indices d'upwelling a montré une corrélation significative et positive entre les indices d'upwelling côtier de Mai-Juillet (période de pré-recrutement de poulpe) et les résidus du modèle. Cette relation n'a pu être mise en évidence que pour l'indice d'upwelling calculé à partir des vents IUC₂ (Bambaye communication personnelle). L'absence de corrélation avec la deuxième série d'indice d'upwelling calculée à partir des SST pourrait être due à des différences dans la spatialisation des données utilisées pour calculer les séries d'IUC. Enfin, les premiers essais de paramétrisation de l'effet de l'environnement sur le paramètre r ne sont pas très concluants. Les critères statistiques formels DIC et Facteur de Bayes ne permettent pas de montrer que les différents modèles avec effet de l'environnement sont meilleurs (d'un point de vue statistique) que le modèle de référence sans environnement.

Dans l'état actuel d'avancement de ce travail et en l'absence de résultats plus probants, il paraît donc plus raisonnable de continuer à baser les évaluations de stock sur le modèle sans effet de l'environnement. Le modèle avec effet de l'environnement reste cependant conceptuellement très intéressant et d'un plus grand réalisme écologique. Il doit faire l'objet d'études complémentaires.

4.3 Perspectives

Ce travail est très riche d'enseignements et a ouvert de nombreuses perspectives.

4.3.1 Redéfinir les zones d'étude et les stocks ?

En Mauritanie, il est probable qu'il existe deux sous-stocks de poulpes assez séparés l'un de l'autre (FAO 1979b) : un situé au Nord vers Cap Blanc et l'autre plus au Sud vers Nouakchott. Une étude séparée de chacun de ces stocks pourrait être entreprise à l'image des dernières études réalisées par le groupe d'évaluation IMROP 2006. Mais cela soulève plusieurs problèmes, notamment : *i*) le stock Nord est partagé entre le Maroc et la Mauritanie ; *ii*) les données permettant de distinguer les deux zones Nord et Sud ne sont disponibles que depuis le début des années 1990.

Les études pourraient également être étendues à l'évaluation des stocks de poulpes du Maroc et du Sénégal. Une première ébauche de ce travail a été réalisée sur le cas du Maroc, à partir des données de la pêcherie et de séries d'indices d'upwelling fournies par l'INRH (Abdel Malek Faraj, communication personnelle). Ce travail préliminaire mérite d'être poursuivi.

4.3.2 Poursuivre les études à partir du modèle de production de biomasse

Le travail de simulation a permis de montrer la nécessité d'intégrer de l'information dans le modèle sous forme de prior informatif. En contre partie, les inférences se révèlent

sensibles aux priors choisis. Le choix du prior informatif est donc une étape cruciale dans la mise en œuvre de ce type de modèle, et doit donc être réalisé avec précaution (McAllister et Kirkwood 1998). Par la suite, on pourrait par exemple envisager de construire des priors informatifs sur les paramètres (par exemple r) en s'inspirant par exemple des méthodes développées par (McAllister *et al.* 2001) qui permettent de valoriser les connaissances sur la démographie de l'espèce.

Dans cette étude, la tentative de paramétrisation des effets de l'environnement a porté sur le paramètre r . D'autres auteurs ont porté leur choix sur le paramètre K (Cury *et al.* 1987 ; Jacobson *et al.* 2005). La flexibilité du cadre de modélisation Bayésien permettrait sans problèmes techniques apparent de conduire d'autres essais dans ce sens.

L'influence dynamique de l'environnement pourrait aussi être modélisée sous la forme de sauts (*shifts*) entre différents paliers. C'est l'approche qui est poursuivie par d'autres partenaires du projet ISTAM (A. Solari 2007, communication personnelle). D'après les analyses de données des captures et d'indice d'abondance mauritaniennes, Solari montre que le système de production de poulpe mauritanien pourrait osciller autour de trois états de stabilité correspondant chacun à une valeur spécifique de la capacité d'accueil du milieu. Ainsi, au lieu d'exprimer l'influence de l'environnement de manière continue sur un paramètre, d'autres modèles basés sur plusieurs niveaux discrets des paramètres (par exemple, trois niveaux de r (r_1, r_2, r_3) ou de K (K_1, K_2, K_3)) pourraient être développés. On pourrait par exemple s'appuyer sur la rupture constatée dans la série d'upwelling entre 1994 et 1995 (cf. Fig. 18)

4.3.3 Envisager d'autres modèles ?

D'autres modèles que le modèle de production de biomasse sur un pas de temps annuel devraient être envisagés. En effet, ce modèle ne paraît pas le plus adéquat pour étudier les variations d'abondance des espèces à vie courte comme le poulpe.

Les travaux d'intégration de l'effet de l'environnement dans le modèle de production à pas de temps annuel n'a pas permis d'obtenir de résultats très probants quant à l'influence de l'environnement sur la production de biomasse. Une réflexion autour d'un modèle de production de biomasse sur un pas de temps de 6 mois mériterait d'être menée. En effet, cette espèce présente deux pics de ponte annuels (avril - juin, septembre - octobre). Un modèle sur un pas de temps de 6 mois permettrait de mieux prendre en compte ces deux pics de recrutement ainsi que les effets de l'environnement.

Enfin, un modèle de déplétion, sur un pas de temps mensuel, intégrant un recrutement variable (McAllister *et al.* 2004 ; Young *et al.* 2004) permettrait encore mieux de représenter la dynamique du stock au cours des saisons de pêche. Mais un tel modèle nécessite de disposer de données beaucoup plus fines (pas de temps court, données de captures en effectif, indices de recrutement). Un travail avec les gestionnaires des bases de données des pêcheries concernées serait nécessaire pour établir si ces données sont disponibles.

5 Bibliographie

- Best, N., Cowles, M. K., and Vines K. 1996. Convergence Diagnosis and Output Analysis Software for Gibbs sampling output Version 0.30. MRC biostatistics Unit, 1-41.
- Boyle, P.R and Boletzky, S.V. 1996. Cephalopod populations: definition and dynamics. *In* Faure V. Inejih C.A., Demarcq H., and Cury P. 2000. The importance of retention processes in upwelling areas for recruitment of *Octopus vulgaris*: the example of the Arguin Bank (Mauritania). *Fisheries Oceanography* 9:4, 343-355.
- Brooks, S. P., and Roberts, G. O. 1998. Assessing Convergence of Markov Chain Monte Carlo Algorithms. *Statistics and Computing* 8: 319-335.
- Buckland, S. T., Newman, K. B., Thomas, L., and Koesters, N. B. 2004. State-space models for the dynamics of wild animal populations. *Ecological modelling* 171: 157-175.
- Caddy, J.F. and Rodhouse, P.G. 1998. Cephalopod and groundfish landings: evidence for ecological change in global fisheries? *Fish Biology and Fisheries* 8: 431-444.
- Caverivičre, A. and Demarcq, H. (2002) Indices d'abondance du poulpe commune et intensité de l'upwelling côtier au Sénégal. in *Le poulpe Octopu vulgaris. Sénégal et côtes nord-ouest africaines. Colloques et séminaires. IRD Editions*, 143-156.
- Chassot E., Balguerías E., Guitton J., Jouffre D., Tfeil B. and Gascuel D. 2006. Diagnostic de l'état du stock de poulpe (*Octopus vulgaris*) mauritanien. Synthèse et nouvelles évaluations par approche globale. Rapport de la sous commission Poulpe, groupe d'évaluation des stocks démersaux, IMROP, Nouadhibou 11-16 janvier 2006.
- Cury, P. and Faure, V. 1998. Environmental constraints and pelagic fisheries in upwelling areas: the Peruvian puzzle. *South African Journal of Marine Science* 19: 159-167
- Cury, P. and Roy, C. (1987) Upwelling et pêche des espèces pélagiques côtières de Côte-d'Ivoire : une approche globale. *Oceanologica acta*, 3, 347-357.
- Demarcq, H. and Faure, V. 2000. Coastal upwelling and associated retention indices derived from satellite SST. Application to *Octopus vulgaris* reruitment. *Oceanologica acta* 23(4), 391-408.
- Dia, M. A. 1988. Biologie et exploitation du poulpe (*Octopus vulgaris*, Cuvier, 1797) des côtes Mauritanienes. *In* Inejih, C.A (thèse). 2000. Dynamique spatio-temporelle et biologique du poulpe (*Octopus vulgaris*) dans les eaux Mauritanienes : modélisation de l'abondance et aménagement des pêcheries.
- Balguerías, E., Quintero, M. E. and Hernandez-Gonzalez, C. L.. 2000. The origin of the Saharan Bank cephalopod fishery. *ICES Journal of Marine Science*, 57: 15–23.
- Failler, P., Diop, M., Dia, M.A., Inejih, C.A., Tous, P. (éds). Evaluation des stocks et aménagement des pêcheries de la ZEE Mauritanienne. Rapport du cinquième Groupe de travail IMROP. Nouadhibou, Mauritanie, 9-17 décembre 2002. COPACE/PACE Séries. No. 06/66. Rome, FAO. 2006. 197p.
- FAO. 1984. FAO Species catalogue Vol. 3. Cephalopods of The World. An Annotated and Illustrated Catalogue of Species of Interest to Fisheries. FAO Fisheries Synopsis, 211-212.
- FAO. 2006. Rapport du groupe de travail FAO/COPACE sur l'évaluation des ressources démersales - Sous-groupe Nord. Saly, Sénégal, 14-23 septembre 2004. FAO, COPACE/PACE Séries 06/68, p372.

- Faure, V., Inejih, C.A., Demarcq, H., and Cury, P. 2000. The importance of retention processes in upwelling areas for recruitment of *Octopus vulgaris*: the example of the Arguin Bank (Mauritania). *Fisheries Oceanography* 9:4, 343-355.
- Fréon, P. 1991. L'introduction d'une variable climatique dans les modèles globaux de production. *ORSTOM*, 395-424.
- Gascuel, D., Inejih, C.A. 2002. Evaluation du stock de poulpe par le modèle global et analyse des effets de l'environnement Rapport du groupe d'évaluation IMROP 2002.
- Gascuel, D., Inejih, C.A. 2006. Evaluation du stock de poulpe par le modèle global et analyse des effets de l'environnement. Rapport du groupe d'évaluation des stocks démersaux, IMROP 2002.
- Gelman A., Carlin, J.B., Stern, H.S., and Rubin, D.B. 1995. Bayesian data analysis. Chapman & Hall p526.
- Guerra, A. 1979. Edad y crecimiento de *Octopus vulgaris* del Atlantico Centro-oriental (26°10'N-23°30'N). *CECAF/ECAF* 78/11, 113-126.
- Hammond, T.R. and Trenkel, V.M. 2005. Censored catch data in fisheries stock assessment. *ICES Journal of Marine Science*, 1118-1130.
- Hatanaka, H. 1979. Spawning season of common octopus, *Octopus vulgaris* CUVIER, off the northwestern coast of Africa. *FAO, COPACE/PACE* 78/11, 135-146.
- Hilborn, R. and Walters, C. J. 1992. Quantitative Fisheries Stock Assessment. Choice, Dynamics and Uncertainty. Chapman and Hall, 297-329.
- IMROP. 2006. Unpublished. Rapport du groupe d'évaluation des ressources démersales en Mauritanie, IMROP, Nouadhibou 11-16 décembre 2006.
- Inejih, C.A. and Dedah, S. 1996. Reproduction et recrutement du poulpe (*Octopus vulgaris*, Cuvier 1797) dans la région du Cap Blanc, Mauritanie. Communication à l'atelier INOC/CNROP sur la circulation des masses d'eau, pollution et leur impact sur les ressources.
- Inejih, C.A. (thèse). 2000. Dynamique spatio-temporelle et biologique du poulpe (*Octopus vulgaris*) dans les eaux Mauritanienues : modélisation de l'abondance et aménagement des pêcheries.
- Jacobson, L.D., Bograd, S.J., Parrish, R.H., Mendelsohn, R., and Schwing, F.B. (2005) An ecosystem-based hypothesis for climatic effects on surplus production in California sardine (*Sardinops sagax*) and environmentally dependent surplus production models. *Can. J. Fish. Aquat. Sci.*, 1782-1796.
- Jeffreys, H., 1961. *Theory of Probability* (3e). Oxford Eds, p432
- Kane, E.A.I. (thèse). 2008. Comportements Opportunistes et Négociation de Accords de Partenariat Pêche entre l'Union Européenne et les Organisations Régionales de Pêches d'Afrique, des Caraïbes et du Pacifique : Enjeux et défis. Cas de la Commission Sous Régionales des Pêches (CSR) de l'Atlantique Centre-Est.
- Kass, R.E., and Raftery, A.E. 1995. Bayes factors. *J. Am. Stat. Assoc.* 90(430): 773-795.
- Hilborn, R. and Peterman, R.M. 1996. The development of scientific advice with incomplete information in the context of the precautionary approach. *In* Lucato, S.H.B. (thèse). 2006. An Improved Mixed-Error Non-Equilibrium Stock-Production Model and its Application to some Brazilian Fish Stocks. p162.
- McAllister, M.K, Hill, S.L., Agnew, D.J., Kirkwood, G.P., and Beddington, J.R.. 2004. A Bayesian hierarchical formulation of the De Lury stock assessment model for abundance

- estimation of Falkland Islands' squid (*Loligo gahi*). *Can. J. Fish. Aquat. Sci.* 61: 1048–1059.
- McAllister, M. K. and Kirkwood, G. P. 1998. Bayesian stock assessment: a review and example application using the logistic model. *ICES Journal of Marine Science* 55, 1031-1060.
- McAllister, M.K., Pikitch, E.K., Babcock, E.A., 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. *Can. J. Fish. Aquat. Sci.* 58, 1871-1890
- McAllister, M.K., Pikitch, E.K., Punt, A.E., and Hilborn, R. 1994. A Bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. *Can. J. Fish. Aquat. Sci.* 51, 2673-87.
- Meyer, R. and Millar, R.B. 1999. Bayesian stock assessment using a state-space implementation of the delay difference model. *Can. J. Fish. Aquat. Sci.*, 56:37-52.
- Meyer, R. and Millar, R.B. 1999. BUGS in bayesian stock assessment. *Can. J. Fish. Aquat. Sci.*, 56: 1078-1086.
- Millar, R. M. and Meyer, R. . 2000. Non linear state-space modelling of fisheries biomass dynamics by using Metropolis-Hastings within-Gibbs sampling. *Appl. Statist.* 49(Part 3), 327-342.
- Myers, R.A. and Mertz, G. 1997. Maximum population growth rates and recovery times for Atlantic cod, *Gadus morhua*. *Fishery Bulletin* 95, 762-772.
- Ould Mahmoud, E., Ould Mohamed Fall, K., Gascuel, D., Corten, A., Groupe de travail IMROP-RIVO, Nouadhibou 19-20 janvier 2006, *In IMROP-RIVO*. 2006.
- Pierce, G.J., Guerra, A., 1994. Stock assessment methods used for cephalopod fisheries. *Fish. Res.* 21, 255–285.
- Pollacheck, T., Hilborn, R., and Punt, A.E. 1993. Fitting Surplus Production Models: Comparing Methods and Measuring Uncertainty. *Can. J. Fish. Aquat. Sci.* 50, 2597-2607.
- Punt, A.E. 2003. Extending production models to include process error in the population dynamics. *Can. J. Fish. Aquat. Sci.* 60: 1217-1228.
- Punt, A.E. and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. *Rev. Fish. Biol. Fish.* (Reviews in Fish Biology and Fisheries) 7: 35-63.
- Semmens, J.M. , Pecl, G.T. , Villanueva, R. , Jouffre, D. , Sobrino, I. , Wood, J.B. , and Rigby, P.R. 2004 Understanding octopus growth: patterns, variability and physiology. *Marine and Freshwater Research*, 367-377.
- Spiegelhalter, D.J., BestBradley, N.G., Carlin, P., and Van der Lite, A. 2002 Bayesian measures of model complexiy and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64(4), 583-639.
- Young, I.A.G., Pierce, G.J.,, Daly, H.I., Santos, M.B., Key , Bailey ,N., Robin, J.-P., Bishop ,A.J., Stowasser, G., Nyegaard, M., Cho, S.K., Rasero M., Pereira, J.M.F. 2004. Application of depletion methods to estimate stock size in the squid *Loligo forbesi* in Scottish waters (UK). *Fisheries Research* 69 : 211–227

6 Annexes

ANNEXE 1 - Captures et d'indice d'abondance (IA) de la pêcherie de poulpes Mauritaniennes

IA-WG98 : issus du groupe de travail IMROP 1998 ; IA-PI : CPUE des flottes industrielles (CPUE) ; IA-Scient : issus des campagnes scientifique ; IA final : moyenne géométrique des 3 séries. (source : IMROP 2006).

Année	Captures	indice WG 98	IA - PI	IA - Scient	IA final
1971	38200	3.40			3.40
1972	36600	3.44			3.44
1973	29200	3.13			3.13
1974	36600	2.46			2.46
1975	39900	1.99			1.99
1976	52900	2.11			2.11
1977	36600	3.14			3.14
1978	23300	3.04			3.04
1979	13400	1.84			1.84
1980	12500	2.62			2.62
1981	32200	1.24			1.24
1982	39100	1.46		3.02	2.10
1983	29000	0.89		3.16	1.67
1984	29500	0.94		3.39	1.79
1985	38000	0.93		2.14	1.41
1986	45600	0.90		0.89	0.90
1987	45700	1.25		1.49	1.37
1988	36700	1.21		2.26	1.65
1989	33500	1.00		1.32	1.15
1990	22840	0.77		1.58	1.10
1991	28237	1.09	1.11	1.56	1.24
1992	42208	1.88	2.03	1.18	1.66
1993	42182	1.42	1.29	0.98	1.22
1994	33883	0.90	0.89	1.18	0.98
1995	29992	0.67	0.77	0.80	0.75
1996	31631	0.48	0.52	0.59	0.53
1997	19448	0.55	0.39	0.70	0.53
1998	15456		0.48	0.76	0.61
1999	22453		0.63	0.99	0.79
2000	31389		0.75	1.02	0.87
2001	26531		0.52	0.51	0.52
2002	23556		0.49	0.37	0.43
2003	22700		0.60	0.46	0.52
2004	21181		0.61	0.70	0.66
2005	29267		0.69	0.64	0.67

ANNEXE 2 - Données d'indice d'upwelling côtier calculé à partir de l'équation de transport d'Ekman

(Source : Bambaye, communication personnelle, IMROP)

Les données utilisées pour le calcul de cette série d'IUC proviennent de deux sources :

- Une série de 45 ans de données mensuelles du vent dans les deux stations météorologiques de Nouakchott et de Nouadhibou sur la période de 1960 à 2005.
- Une série de 45 ans de données mensuelles des composantes du vent sur la ZEE Mauritanienne pendant la même période (1960 – 2005), ces données sont tirées de la base NOAA NCEP - NCAR CDAS – 1 avec une résolution de deux degrés ($2^{\circ} \times 2^{\circ}$).

L'indice d'upwelling (CUI) est ainsi calculé par la formule suivante :

$$CUI = (r \cdot Cd \cdot V^2) / (2 \cdot \omega \cdot \sin y) \text{ [T/sec/100m littoral]}$$

r – densité de l'air ($1,22 \cdot 10^3 \text{ d/cm}$)

Cd - coefficient de viscosité à l'interface air-mer ($1.25 \cdot 10^3$).

V – composante du vent parallèle a la côte au carré (m/sec) sur une surface de 1000 mb.

ω – vitesse angulaire de la rotation de la terre ($0.729 \cdot 10^{-2} / \text{sec}$)

Y – latitude du point.

La composante parallèle du vent a été calculée selon les formules suivantes :

$$V_y = -V \cdot \cos \Phi \quad 0 \geq \Phi \leq 90$$

$$V_y = V \cdot \cos (180 - \Phi) \quad 90 \geq \Phi \leq 180$$

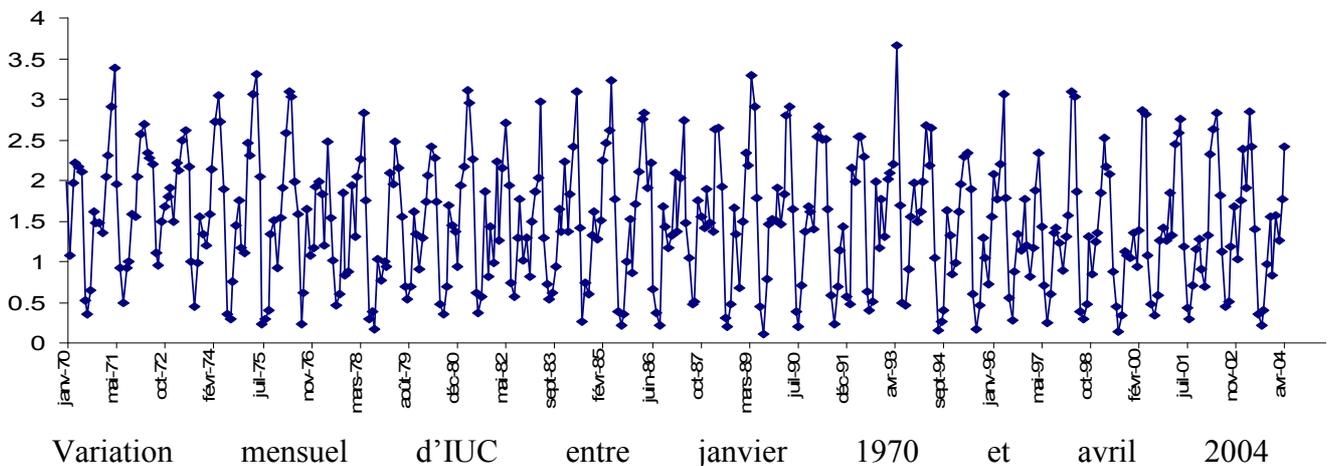
$$V_y = V \cdot \cos (\Phi - 180) \quad 180 \geq \Phi \leq 270$$

$$V_y = -V \cdot \cos (360 - \Phi) \quad 270 \geq \Phi \leq 360$$

Avec : V_y – composante parallèle du vent

V = Vitesse du vent

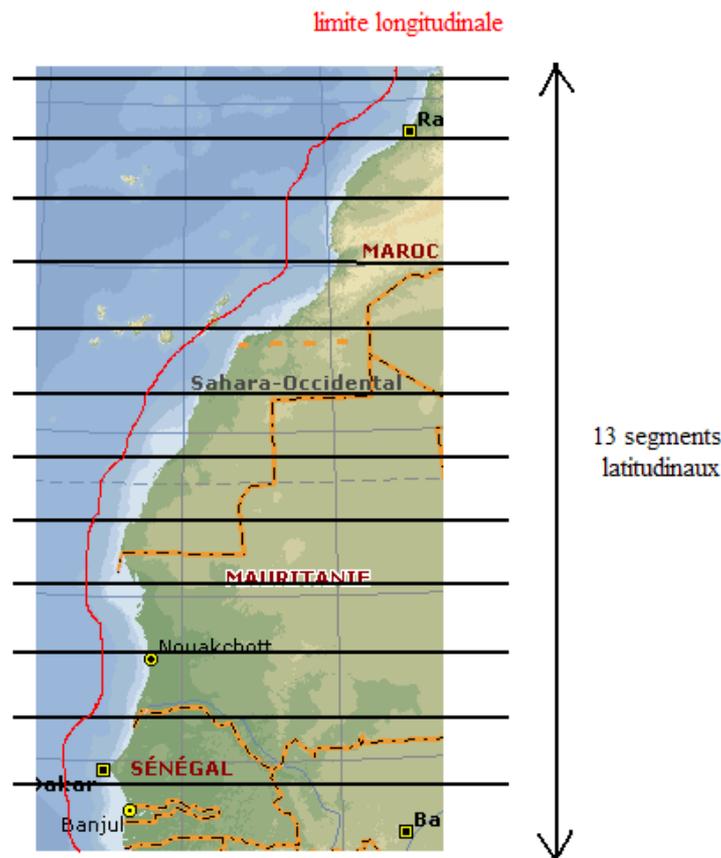
Φ = Direction du vent



ANNEXE 3 - Calcul des IUC à partir des données de SST

Les données proviennent du satellite AVHRR (résolution spatiale de 0.7miles x 0.7miles) et s'étendent de janvier 1985 à décembre 2005 avec un pas de temps mensuel. Les données de SST s'étendent sur une grande échelle : depuis le Maroc jusqu'au Sénégal, entre 36°N et 12°N et entre 20°W et 6°W.

Un découpage latitudinal de 2° par zone a été réalisé. De cette manière, 13 segments ont été créés entre le Nord et le Sud de la zone d'étude. Au niveau longitudinal, seules les valeurs de SST comprise dans une zone allant de la côte à la limite bathymétrique des 200m (ce qui constitue la limite biologique de vie du poulpe commun) ont été conservées.



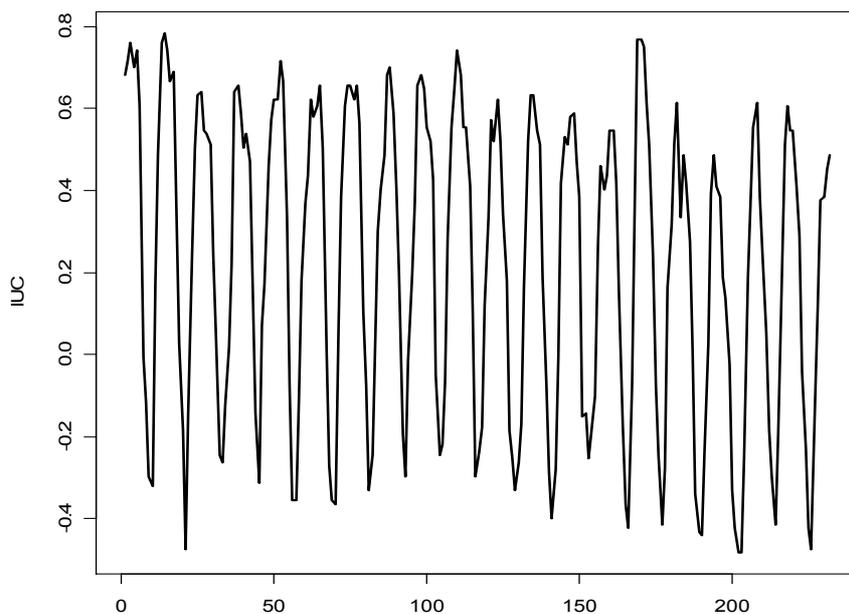
L'IUC a été ensuite calculé selon l'équation suivante :

$$IUC = \frac{SST_{mean} - SST_{min}}{SST_{mean} - SST_{upw}}$$

- SST_{mean} correspond à une moyenne de SST sur la totalité des 13 segments. De cette manière, on évite les potentielles influences des anomalies climatiques associées à des changements environnementaux de grande échelle (Demarcq *et al.* 2000).
- SST_{min} correspond à la température minimale dans les segments 3 et 4 confondues (segments en partant du bas)
- SST_{upw} est la température des eaux profondes des upwellings. Elle a été choisi égale à 14°C, température du South Atlantic Central Water (SACW) qui est la composante principale des eaux d'upwelling en Mauritanie (Faure *et al.* 2000 ; Demarcq *et al.* 2000)

Lorsque l'upwelling est fort, la SST_{min} est alors très proche du SST_{upw} et on a donc des valeurs d'IUC qui se rapprochent de 1. A contrario, lorsque l'upwelling est faible, cette valeur est également faible.

Remarque : Dans la série d'IUC ainsi obtenues, on trouve parfois des valeurs d'IUC négatives. La raison est que la SST_{mean} a été calculée en moyennant les SST sur les 13 zones. De ce fait, il arrive que la SST_{min} de certaines zones (surtout du Sud) puisse avoir des valeurs supérieures à la SST_{mean} .



Variation mensuel de l'IUC entre janvier 1985 et décembre 2005

ANNEXE 4 - Détails techniques de l'estimation Bayésienne

Reparamétrisation du modèle suggérée par Millar *et Meyer* (2000)

$$P_t = B_t / K$$

Cette reparamétrisation réduit considérablement l'intervalle des valeurs prises par B_t et le temps de calcul.

Convergence

Dans le cadre Bayésien, on doit faire face au problème de la vitesse de convergence des chaînes MCMC et de la taille du "burn-in period" : longueur des premiers tirages MCMC à rejeter (Gelman *et al.* 1995 ; R-News, 2006 ; Cowles *et al.* 1996, Brooks *et al.* 1998) car ils correspondent à une « période de chauffe » de l'algorithme. Si la convergence n'est pas atteinte, la distribution estimée par l'échantillon obtenu n'est pas représentative de la distribution a posteriori.

La théorie markovienne montre que la distribution de ces tirages MCMC converge vers la distribution *a posteriori* « vraie » de ses paramètres lorsque le nombre d'itérations tend vers l'infini. (Brooks *et al.*, 1998 ; R-News, 2006). Mais en pratique, on ne peut pas répéter les itérations un nombre infini de fois ; il faut donc utiliser des techniques statistiques pour tenter d'évaluer ce point de convergence : c'est ce que l'on appelle des « diagnostics de convergence » (Brooks, 1998 ; Brooks *et al.* 1998, Cowles *et al.* 1996).

Plusieurs auteurs comme Brooks, Carlin ou Asmussen affirment qu'il n'existe pas une méthode unique capable de donner de bons résultats dans tous les cas de figure, et qu'il est donc préférable de multiplier les méthodes de diagnostics pour étudier la convergence et (Brooks *et al.* 1998 ; Asmussen *et al.* 1992 dans Brooks, 1998 ; Kass *et al.* 1998).

Valeur de « thinning »

Les tirages MCMC sont par définition auto-corrélés (car ils proviennent d'une chaîne de Markov). Parallèlement, il existe également des limites techniques telles que la capacité de stockage ou le temps de calcul. Afin de limiter l'autocorrélation dans les tirages et de limiter la taille des chaînes MCMC stockées (pb. de mémoire), il est souvent nécessaire d'adopter un « thinning », c'est-à-dire de ne conserver que 1 tirage tout les x tirages.

Choix méthodologiques adoptés dans ce travail

Au final, dans notre travail, pour chaque simulation, 3 chaînes de 100.000 itérations (en moyenne) ont été lancées avec un thinning de 5 après un burn-in period de 4000.

Dans notre cas d'étude, on a mis en place 3 diagnostics de convergence pour estimer le point de convergence :

- le diagnostic de Gelman&Rubin (1992)
- le diagnostic de Raftery&Lewis (1992)
- le diagnostic de Geweke (1992)

Ces diagnostics de convergence ont été conduits grâce au package CODA sous R.

Diagnostic de Gelman&Rubin (source : Gelman et Rubin, 1992)

C'est le seul diagnostic multi-chaîne parmi ces 3 méthodes et il s'intéresse au rapport de variance intra-chaîne/inter-chaîne pour un variable donné. Lorsque cette valeur converge vers 1, alors les chaînes MCMC ont convergé.

Ce diagnostic semble être assez fiable et assez facile à interpréter à condition que la distribution a posteriori de la variable soit unimodale et suive une loi normale. Chose qui est rarement vérifiée exactement dans la pratique.

Diagnostic de Raftery&Lewis (source : Raftery et Lewis,1992)

C'est un diagnostic mono-chaîne qui s'intéresse à la stabilisation de la valeur d'un quantile d'une chaîne. Une fois la convergence atteinte, la distribution a posteriori de la variable ne change plus même si le nombre d'itérations augmente. De ce fait, la valeur d'un quantile doit aussi se stabiliser une fois la convergence atteinte. Ce diagnostic fait donc un test de stationnarité avec un niveau de confiance α .

Récemment, avec la nouvelle version d'OpenBUGS 3.01, on peut visualiser directement ce diagnostic.

Diagnostic de Geweke (source : Geweke, 1992)

C'est un diagnostic mono-chaîne qui fait une comparaison de la moyenne entre 2 segments de la chaîne (chaîne de taille N) pris au début et à la fin de la chaîne MCMC. Le premier bout de chaîne est de taille $N_1 = \alpha\%$ de N et on fixe généralement la taille du deuxième segment à $N_2 = 50\%$ de N (seconde moitié de la longueur totale de la chaîne). Lorsque les moyennes de ces deux chaînes sont très proches, cela suggère que la convergence est atteinte. Il faut faire attention à ce que les deux segments de chaînes ne se croisent pas.

Il est à noter que ce diagnostic est surtout efficace pour montrer que la convergence n'a pas été atteinte mais pas pour affirmer que la convergence a été atteinte.

Dans toute notre étude, la convergence des chaînes a systématiquement été vérifiée à l'aide d'une combinaison de trois diagnostics. Une étude menée sur un cas simple a permis d'établir que la chaîne de traitement suivante donnait de très bons résultats :

- 1 - réaliser en premier le diagnostic de Gelman&Rubin
→ cela permet d'avoir une idée générale du moment de convergence.
- 2 - faire suivre cette étude du diagnostic de Raftery&Lewis
→ cela permet de donner des indications plus pointilleuses sur le point de convergence mais qui est à prendre avec précaution car selon la précision et le quantile sur lequel on base le traitement, les résultats peuvent varier.

Dans le cas Mauritanie, les réponses plus adaptées et concomitantes avec les résultats de Gelman ont été obtenues avec les quantiles 0.1 ou 0.05.

Brooks, Hilborn, et d'autres ont bien précisé que ces diagnostics sont à analyser avec précaution et qu'ils étaient « case-specific ».

- 3 - Une étape finale de vérification basée sur le diagnostic de Geweke qui permet surtout de confirmer que la taille de chaîne donnée par les diagnostics précédents donne un diagnostic de Geweke positif.

ANNEXE 5 - DIC et Facteurs de Bayes

DIC

Le calcul de DIC fait intervenir un certain nombre de paramètres :

$$\text{DIC} = \text{Dbar} + pD = \text{Dhat} + 2 \cdot pD$$

Avec

Dbar : il correspond à la moyenne de la deviance de tous les paramètres. $\overline{D(\theta)}$

Dhat : il correspond à la deviance ponctuelle calculée à la moyenne de tous les paramètres. $D(\bar{\theta})$.

Rq : Le Dhat (ou DIC) suppose que les moyennes constituent une bonne estimation ponctuelle des paramètres étudiés. C'est le cas lorsque le posterior est symétrique (loi normal) mais dès que cela n'est pas vérifié, comme notre cas, alors le DIC est à interpréter avec précaution. Mais cela permet de donner une idée générale.

pD : C'est une variable qui permet de décrire la complexité du modèle. Plus ce facteur est grand, plus le modèle est complexe et contient beaucoup de paramètres.

Le DIC permet de faire un compromis entre la qualité d'ajustement (exprimé par le terme de la vraisemblance) et la complexité du modèle.

→ Le modèle sélectionné est celui qui a le plus petit DIC

On choisit généralement, par défaut, une différence de DIC de plus de 5 comme critère de choix du « meilleur » modèle.

Rappel

La deviance correspond à : $deviance = -2 \cdot \log(p(y / \theta))$

Avec : θ le vecteur des paramètres du modèle.

$p(y / \theta)$ la vraisemblance du modèle en θ .

Bayes Factor

Le Bayes Factor s'exprime de la manière suivante :

$$(1) \quad \frac{p(M_i / data)}{p(M_j / data)} = \frac{p(M_i)}{p(M_j)} \cdot \frac{p(data / M_i)}{p(data / M_j)}$$

i.e : « Posterior odd » = « prior odd » · « Bayes Factor »

$p(data / M_i)$ est la vraisemblance marginale dans le modèle M_i qui peut s'estimer par Monte Carlo de la manière suivante :

$$(2) \quad p(\text{data} / M_i) \approx 1/G \left(\sum_{i=1}^G p(\text{data} / \theta_i^{\text{posterior}})^{-1} \right)^{-1}$$

avec G = taille de la chaîne MCMC (nombre d'itérations)

Grâce à WinBUGS, on dispose de tirages MCMC de valeurs de la deviance : Or :

$$(3) \quad \text{deviance} = -2 \cdot \log(p(y / \theta_i^{\text{posterior}}))$$

$$(4) \quad \rightarrow p(\text{data} / \theta_i^{\text{posterior}}) = e^{-1/2 * (\text{deviance})}$$

ANNEXE 6 - Résumé des ajustements dynamiques et à l'équilibre des 13 scénarii, travail avec les IC Bayesiens

Le tableau ci-dessous donne, pour chaque scénario, le nombre de fois (sur les 10 trajectoires simulées par scénario) que l'IC Bayésien à 75% contient effectivement la vraie valeur du paramètre

Paramètres de simulation	Type d'ajustement	r	q	σ^2_p	K	C_{msy}	B_2	B_{34}	B_{34}/K
scénario N°1	dynamique	9	6	8	10	9	10	10	10
	équilibre	10	10	0	10	0	9	8	0
scénario N°2	dynamique	8	6	8	10	8	10	10	7
	équilibre	10	10	0	10	2	2	2	0
scénario N°3	dynamique	10	5	8	7	9	7	7	7
	équilibre	10	3	0	8	8	0	0	0
scénario N°4	dynamique	8	7	6	8	6	9	9	9
	équilibre	10	8	0	9	3	3	3	0
scénario N°5	dynamique	8	8	8	8	5	8	8	6
	équilibre	10	1	0	10	8	1	0	0
scénario N°6	dynamique	9	1	9	10	10	10	10	4
	équilibre	10	0	0	10	10	0	0	0
scénario N°7	dynamique	9	6	9	10	9	10	10	8
	équilibre	10	9	0	9	0	7	8	0
scénario N°8	dynamique	10	3	9	8	10	8	7	8
	équilibre	10	10	0	8	0	2	2	0
scénario N°9	dynamique	10	7	7	8	9	8	9	8
	équilibre	10	6	0	8	9	0	0	0
scénario N°10	dynamique	9	7	7	8	8	8	8	9
	équilibre	10	10	0	10	0	10	10	0
scénario N°11	dynamique	9	0	7	9	10	9	9	8
	équilibre	10	10	0	10	0	7	8	0
scénario N°12	dynamique	8	8	6	8	8	9	8	8
	équilibre	10	10	0	10	0	9	9	0
scénario N°13	dynamique	8	7	10	10	7	10	10	9
	équilibre	10	10	0	7	0	5	5	0

ANNEXE 7 - Code WinBUGS

1. SSM

```
model ;
{

##### Les priors #####

## prior sur "q"
log_q ~ dunif(-15 , 3)
q <- exp(log_q)

## prior sur "r" :
mode <- 1.2
var <- 1
a1 <- pow(mode, 2)
a2 <- a1 + 4*var
a3 <- pow(a2, 0.5)
a4 <- mode + a3
a5 <- pow(a4, 2)
a <- a5 / (4*var)
b <- a4 / (2*var)
r ~ dgamma(a, b)

## prior sur "K"
K ~ dunif(50000, 10000000)

## prior sur la variance de l'erreur de processus
log_var_P ~ dunif(-9, 3)
var_P <- exp(log_var_P)
ivar_P <- 1 / var_P

## prior sur la variance de l'erreur d'observation
var_IA <- var_P / lambda
ivar_IA <- 1 / var_IA

### process model :  $P[t+1] = (P[t] - r / \ln(K) * P[t] * \ln P[t]) * e^{e1 - \text{Capt}/K}$ 

## initialisation
P[1] <- alpha
log_P[1] <- log(P[1])
B[1] <- K
```

```

## chaîne MCMC
for (t in 1: n-1)
{
Umean[t] <- max( P[t] - r/log(K) *P[t]*log_P[t] , 0.000001 )
log_Umean[t] <- log(Umean[t])
log_Utrue[t] ~ dnorm(log_Umean[t], ivar_P)

Utrue[t] <- exp(log_Utrue[t])
P[t+1] <- max(Utrue[t] -Capt[t]/K, 0.000001)
log_P[t+1] <- log(P[t+1])
}

Umean[34] <- max(P[34] - r/log(K)*P[34]*log_P[34], 0.000001 )
log_Umean[34] <- log(Umean[34])
log_Utrue[34] ~ dnorm(log_Umean[34], ivar_P)
Utrue[34] <- exp(log_Utrue[34])

## boucle pour Biomasse
for (t in 2: n)
{
B[t] <- P[t]*K
}

### observation model :     $IA[t] = (q * P[t] * K) * e^{-e2}$ 

for (t in 1:n)
{
log_Iobs[t] <- log(Iobs[t])
log_Iobs[t] ~ dnorm(log_Ipred[t], ivar_IA)
log_Ipred[t] <- log_P[t] + log(K) +log(q)
}

### Résidus du modèle
for (t in 1:n-1)
{
Res_B[t+1] <- Utrue[t]*K - Umean[t]*K
}

### Autres paramètres intéressants :
## Valeur de "Cmsy"
CMSY <- r * K / (exp(1) * log(K))

}

##### DATAS #####
list(
n= 34,

```

```

Capt = c(38200, 36600, 29200, 36600, 39900, 52900, 36600, 23300, 13400, 12500,
          32200, 39100, 29000, 29500, 38000, 45600, 45700, 36700, 33500, 22840, 28237,
          42208, 42182, 33883, 29992, 31631, 19448, 15456, 22453, 31389, 26531, 23556,
          22700, 21181),
alpha = 1,
lambda = 1
)

```

Remarque : La série d'Iobs sera prise des données de chaque simulation

```
##### Initialisation des paramètres : 1 exemple #####
```

```

list(
log_q = -11 ,
r = 4,
K = 400000,
log_var_P = -3 ,
log_Utrue = c(-0.116533816, -0.127833372, -0.139262067, -0.15082289, -0.162518929, -
0.174353387, -0.186329578, -0.198450939, -0.210721031, -0.223143551, -0.235722334, -
0.248461359, -0.261364764, -0.274436846, -0.287682072, -0.301105093, -0.314710745, -
0.328504067, -0.342490309, -0.356674944, -0.371063681, -0.385662481, -0.400477567, -
0.415515444, -0.430782916, -0.446287103, -0.46203546, -0.478035801, -0.494296322, -
0.510825624, -0.527632742, -0.544727175, -0.562118918,
-0.579818495)
)

```

2. Process error model sous l'hypothèse d'équilibre

```

model ;
{

##### Les priors #####

## prior sur "q"
log_q ~ dunif(-15 , 3)
q <- exp(log_q)

## prior sur "r"
mode <- 1.2
var <- 1
a1 <- pow(mode, 2)
a2 <- a1 + 4*var
a3 <- pow(a2, 0.5)
a4 <- mode + a3

```

```

a5 <- pow(a4, 2)
a <- a5 / (4*var)
b <- a4 / (2*var)
r ~ dgamma(a, b)

## prior sur "K"
K ~ dunif(50000, 10000000)

## prior sur la variance de l'erreur de processus
log_var_P ~ dunif(-9, 3)
var_P <- exp(log_var_P)
ivar_P <- 1 / var_P

## prior sur la variance de l'erreur d'observation
var_IA <- var_P / lambda
ivar_IA <- 1 / var_IA

# Processus à l'équilibre
for (t in 1:n)
{
Cpred[t] <- max( (- r/log(K) *P[t]*log(P[t])) * K , 0.000001)
log_Capt[t] ~ dnorm(log_Cpred[t], ivar_P)
log_Capt[t] <- log(Capt[t])
log_Cpred[t] <- log(Cpred[t])
}

## boucle pour Biomasse
for (t in 2: n)
{
B[t] <- P[t]*K
}

#### observation model

for (t in 1 : n)
{
log_Iobs[t] <- log(Iobs[t])
log(P[t]) <- log_Iobs[t] - log(K) - log(q)
}

```

3. Modèle avec environnement, exemple du M5

```
model ;
{

##### Les priors #####

## prior sur "q"
log_q ~ dunif(-15 , 3)
q <- exp(log_q)

## prior sur "r"
b ~ dunif(-1, 3)
abis ~ dunif(0, 3)
var <- 10
for (t in 1:n)
{
mode[t] <- abis * pow(upw[t], b)
a1[t] <- pow(mode[t], 2)
a2[t] <- a1[t] + 4*var
a3[t] <- pow(a2[t], 0.5)
a4[t] <- mode[t] + a3[t]
a5[t] <- pow(a4[t], 2)
a[t] <- a5[t] / (4*var)
b1[t] <- a4[t] / (2*var)
r[t]~ dgamma(a[t], b1[t])

## création du paramètre "Cmsy"
CMSY[t] <- r[t]*K/(exp(1)*log(K))
}

## "prior sur "K"
K ~ dunif(50000, 10000000)

## erreur de processus
log_var_P ~ dunif(-9, 3)
var_P <- exp(log_var_P)
ivar_P <- 1 / var_P

## erreur d'observation
var_IA <- var_P / lambda
ivar_IA <- 1 / var_IA

### process model :  $P[t+1] = (P[t] - r / \ln(K) * P[t] * \ln P[t]) * e^{-1} - \text{Capt}/K$ 

## initialisation
P[1] <- alpha
log_P[1] <- log(P[1])
```

```

B[1] <- K

for (t in 1: n-1)
{
Umean[t] <- max( P[t] - r/log(K) *P[t]*log_P[t] , 0.000001 )
log_Umean[t] <- log(Umean[t])
log_Utrue[t] ~ dnorm(log_Umean[t], ivar_P)

Utrue[t] <- exp(log_Utrue[t])
P[t+1] <- max(Utrue[t] -Capt[t]/K, 0.000001)
log_P[t+1] <- log(P[t+1])
}

Umean[34] <- max(P[34] - r/log(K)*P[34]*log_P[34], 0.000001 )
log_Umean[34] <- log(Umean[34])
log_Utrue[34] ~ dnorm(log_Umean[34], ivar_P)
Utrue[34] <- exp(log_Utrue[34])

## boucle pour Biomasse
for (t in 2: n)
{
B[t] <- P[t]*K
}

#### observation model :   IA[t] = ( q * P[t] * K ) * e ^e2

for (t in 1:n)
{
log_Iobs[t] <- log(Iobs[t])
log_Iobs[t] ~ dnorm(log_Ipred[t], ivar_IA)
log_Ipred[t] <- log_P[t] + log(K) +log(q)
}

##### DATAS #####

list(
n = 33,
Iobs = c(3.40, 3.44, 3.13, 2.46, 1.99, 2.11, 3.14, 3.04, 1.84, 2.62, 1.24, 2.10,
1.67, 1.79, 1.41, 0.90, 1.37, 1.65, 1.15, 1.10, 1.24, 1.66, 1.22, 0.98,
0.75, 0.53, 0.53, 0.61, 0.79, 0.87, 0.52, 0.43, 0.52),
Capt = c(38200, 36600, 29200, 36600, 39900, 52900, 36600, 23300, 13400, 12500,
32200, 39100, 29000, 29500, 38000, 45600, 45700, 36700, 33500, 22840, 28237,
42208, 42182, 33883, 29992, 31631, 19448, 15456, 22453, 31389, 26531, 23556,
22700) ,
upw = c(1.3445219, 1.8287681, 1.7727206, 1.7796560, 1.2259617, 1.1671215, 2.1223473,
1.5685448, 1.1246575, 1.3230609, 1.2408238, 1.5991602, 1.5765132, 1.4091757, 1.0501178,
1.3017985, 1.4112664, 1.6200699, 1.1597566, 1.3280729, 1.2228357, 1.3470333, 1.5403521,

```

```
1.4539891, 0.7595166, 1.3799876, 1.4433665, 1.3278796, 1.3012041, 0.9364284, 1.1244505,  
1.1201228, 1.7947848),  
alpha = 1 ,  
lambda = 1  
)
```

```
##### Initialisation des paramètres : 1 exemple #####
```

```
list(  
log_q = -9 ,  
abis = 1.5 ,  
b = 1 ,  
K = 4000000 ,  
log_var_P = 3 ,  
log_Utrue = c(-0.116533816, -0.127833372, -0.139262067, -0.15082289, -0.162518929, -  
0.174353387, -0.186329578, -0.198450939, -0.210721031, -0.223143551, -0.235722334, -  
0.248461359, -0.261364764, -0.274436846, -0.287682072, -0.301105093, -0.314710745, -  
0.328504067, -0.342490309, -0.356674944, -0.371063681, -0.385662481, -0.400477567, -  
0.415515444, -0.430782916, -0.446287103, -0.46203546, -0.478035801, -0.494296322, -  
0.510825624, -0.52763, -0.562118918,  
-0.5798184952742),  
r= c(1.3445219, 1.8287681, 1.7727206, 1.7796560, 1.2259617, 1.1671215, 2.1223473,  
1.5685448, 1.1246575, 1.3230609, 1.2408238, 1.5991602, 1.5765132, 1.4091757, 1.0501178,  
1.3017985, 1.4112664, 1.6200699, 1.1597566, 1.3280729, 1.2228357, 1.3470333, 1.5403521,  
1.4539891, 0.7595166, 1.3799876, 1.4433665, 1.3278796, 1.3012041, 0.9364284, 1.1244505,  
1.1201228, 1.7947848)  
)
```



Pôle : Halieutique

Spécialisation : Sciences Halieutiques et
Aquacoles

Enseignant Responsable du stage:

M. Sabatié RICHARD



Cadre réservé à la Bibliothèque
Centrale

Auteur : Kotaro ONO

Organisme d'accueil : Laboratoire d'Ecologie
Halieutique - Agrocampus Rennes

65 rue de Saint-Brieuc, CS84215, 35042, RENNES
Cedex

Année de soutenance : 2006 - 2007

Tél : 02.23.48.55.36

Titre du mémoire :

Fax : 02.23.48.55.35

« Modélisation statistique bayésienne d'un
modèle de production de biomasse. Application
à la pêche de poulpe (*Octopus vulgaris*) de
Mauritanie »

Courriel : halieut@agrocampus-rennes.fr

Maître de stage : M. Etienne RIVOT

Fonction : enseignant chercheur

Résumé :

Ce travail s'inscrit dans le *Working Package 3* de l'action concertée ISTAM (Improve Statistical and Technical Advice for fishery Management). Le WP3 vise à contribuer à l'amélioration des méthodes d'évaluation des stocks dans la zone COPACE. L'attention a été portée à la pêche de poulpe Mauritanien. Ce travail développe une approche Bayésienne de traitement du modèle dynamique de production de Biomasse appliquée à cette pêche de poulpe. Dans une première partie, la faisabilité et la qualité des inférences obtenues sont évaluées par une approche de simulation-ajustement. Cette étude a fait apparaître des difficultés et des limites d'une telle approche mais a permis de montrer que l'ajustement Bayésien donne en moyenne de bonnes inférences sur la plupart des paramètres d'intérêt. De plus, ces estimations sont souvent meilleures que celles des autres méthodes classiques d'ajustement (*observation error model* dynamique, ou modèle à l'équilibre). Dans une deuxième partie, le modèle a été appliqué au cas de la pêche de poulpe Mauritanien. L'estimation des paramètres clés comme le C_{MSY} ou le taux de réduction de biomasse a permis de montrer que ce stock est largement surexploité. Enfin, une tentative de paramétrisation de l'effet d'upwelling sur le même modèle a été effectuée et une corrélation positive a été trouvée entre la série d'indices d'upwelling et les résidus du modèle sans environnement. Cependant des critères statistiques ont montré que le modèle sans environnement reste meilleur.

Abstract :

This study takes part of the Working Package 3 of ISTAM (Improve Statistical and Technical Advice for fishery Management) which has the aim to improve the evaluation methods for COPACE region resources. The study is interested in the octopus fishery in Mauritania and a Bayesian nonlinear state-space framework using the Fox production model in the dynamic state process was used to explore the population dynamics of this resource. First, a simulation study was conducted to analyse the reliability of the estimations when data are generated in a SSM framework. Despite some limits, the analysis shows that the Bayesian SSM method gives reliable estimation on mostly all the model parameters. Moreover, SSM often performs best in parameter estimations compared to the dynamic observation error model and the equilibrium model. Then, the same model is used to estimate quantities of interest for the Mauritanian octopus stock (C_{MSY} and depletion index). The study shows that the stock is being largely over-exploited. Finally, we tried to integrate the effect of the upwelling index on a model parameter to better explain the population dynamics. A positive correlation was found between residuals of the model and this environmental index, but further statistical analysis suggested that this does not greatly improve the fit of the model.

Mots clés :

State-Space Model, Bayésien, Simulation, *Octopus vulgaris*, upwelling, ISTAM