



Institut de recherche
pour le développement

Mémoire de fin d'études

Pour l'obtention du « Diplôme d'Agronomie Approfondie » (DAA)

Spécialisation Halieutique

**Modélisation statistique Bayésienne de la
croissance du thon albacore (*Thunnus albacares*)
de l'océan Indien**

Présenté par : MASSIOT-GRANIER, Félix

Soutenu le : 16 septembre 2010



Institut de recherche
pour le développement

Mémoire de fin d'études

Pour l'obtention du « Diplôme d'Agronomie Approfondie » (DAA)

Spécialisation Halieutique

**Modélisation statistique Bayésienne de la
croissance du thon albacore (*Thunnus albacares*)
de l'océan Indien**

Présenté par : MASSIOT-GRANIER, Félix

Soutenu le : 16 septembre 2010

Devant le Jury

M. Etienne RIVOT, Agrocampus Ouest Pôle Halieutique.

M. Emmanuel CHASSOT, IRD Sète.

M. Didier GASCUEL, Agrocampus Ouest Pôle Halieutique.

M. Eric PARENT, AgroParisTech-ENGREF.

M. Alain FONTENEAU, IRD Sète.

Diffusion du mémoire

Aucune confidentialité ne sera prise en compte si la durée n'en est pas précisée.

Préciser les limites de la confidentialité ⁽¹⁾ :

Mémoire de fin d'études

Consultable sur place : oui non

Reproduction autorisée : oui non

Prêt autorisé : oui non

Confidentialité absolue : oui non

(ni consultation, ni prêt)

Diffusion de la version numérique : oui non

Durée de la confidentialité ⁽²⁾ :

Je soussigné Emmanuel Chassot, propriétaire des droits de reproduction de la dite version, autorise toutes les sources bibliographiques à le signaler et le publier.

Fiche de résumé du mémoire de fin d'études :

Résumé diffusable : oui non

Si oui, l'auteur complète l'autorisation suivante :

Je soussigné Emmanuel Chassot, propriétaire des droits de reproduction dudit résumé, autorise toutes les sources bibliographiques à le signaler et le publier.

Date :

Signature :

Rennes, le

Le Maître de stage⁽³⁾, Emmanuel Chassot (IRD Sète)

L'auteur, Félix Massiot-Granier

L'Enseignant responsable⁽³⁾, Didier Gascuel (Agrocampus Ouest)

(1) L'administration, les enseignants et les différents services de documentation du Pôle Agronomique de Rennes s'engagent à respecter cette confidentialité.

(2) La durée maximale de confidentialité est fixée à 10 ans.

(3) Signature et cachet de l'organisme.

Remerciements

Je tiens tout d'abord à remercier Etienne Rivot et Emmanuel Chassot. Il a été très agréable de travailler avec vous et j'ai beaucoup appris en votre présence.

Je tiens également à remercier Jean-Pierre Hallier pour le temps qu'il m'a consacré et les réponses qu'il a su m'apporter.

Je remercie également Eric Morize et toute l'équipe du LEMAR pour leur accueil et leurs explications précieuses concernant les lectures d'otolithes.

Un grand merci à toute l'équipe du labo halieut qui rend ce lieu agréable tout en restant propice au travail.

J'en profite aussi pour remercier ma famille pour son soutien et Juliette pour m'avoir obligé à profiter un peu du soleil Sétois.

Tables des matières

I	INTRODUCTION	1
II	MATERIELS ET METHODES	3
1.	MODELES DE CROISSANCE	3
1.1.	<i>Différents modèles proposés dans la littérature</i>	3
1.2.	<i>Modèles à stances concernant l'albacore</i>	3
2.	ESTIMATION DE LA COURBE DE CROISSANCE	5
2.1.	<i>Intégration de multiples sources d'information et des erreurs de mesures</i>	5
2.2.	<i>Le choix du cadre Bayésien</i>	5
2.3.	<i>Principes des modèles d'estimation de la courbe de croissance</i>	7
2.3.1.	<i>Ajustement à partir de couples âge-taille</i>	8
2.3.2.	<i>Ajustement à partir de couples de marquage-recapture</i>	8
2.3.3.	<i>Ajustement à partir de couples de marquage-recapture et des couples âge-taille</i>	10
3.	LES DONNEES DE MARQUAGE-RECAPTURE ET D'OTOLITHOMETRIE POUR L'ALBACORE DE L'OCEAN INDIEN.	10
3.1.	<i>Le jeu de données de marquage-recapture</i>	10
3.1.1.	<i>Présentation du jeu de données</i>	10
3.1.2.	<i>Sélection des données</i>	10
3.1.3.	<i>Sources d'erreurs de mesure de la taille</i>	12
3.2.	<i>Le jeu de données d'otolithométrie</i>	12
3.2.1.	<i>Présentation des données</i>	12
3.2.2.	<i>Lecture de l'âge</i>	12
3.2.3.	<i>Sources d'erreurs de mesure</i>	14
3.2.4.	<i>Modèle d'erreurs de mesure liées aux lectures de microstries d'otolithes</i>	14
4.	METHODES DE VALIDATION PAR SIMULATION-ESTIMATION	16
4.1.	<i>Démarche générale et modèle de simulation</i>	16
4.2.	<i>Scénario de référence</i>	16
4.3.	<i>Influence de la qualité et de la quantité des données simulées</i>	17
4.3.1.	<i>Influence de la qualité des données simulées</i>	17
4.3.2.	<i>Influence de la quantité de données</i>	18
4.4.	<i>Sensibilité au prior lors de la phase d'estimation Bayésienne</i>	18
4.4.1.	<i>Prior sur L_{inf}</i>	18
4.4.2.	<i>Prior sur les paramètres de transition α et β</i>	18
4.4.3.	<i>Prior sur les taux de croissance de la première stance (k_1) et de la deuxième stance (k_2)</i>	18
5.	AJUSTEMENT BAYESIEN DES MODELE DE CROISSANCE A PARTIR DES DONNEES DE L'ALBACORE DE L'OCEAN INDIEN	19
5.1.	<i>Le choix des priors</i>	19
5.2.	<i>Le modèle de croissance sur les données de sclérochronologie</i>	20

5.3.	<i>Le modèle de croissance intégrant les données de marquage-recapture et les données de sclérochronologie.....</i>	20
5.4.	<i>Le modèle de croissance adapté à des données de sclérochronologie couplé au modèle d'erreurs de mesures liées aux lectures de microstries d'otolithes.....</i>	20
III	RESULTATS	21
1.	EVALUATION DE LA PERFORMANCE DE LA METHODE PAR SIMULATION-ESTIMATION	21
1.1.	<i>Corrélation statistique des paramètres.....</i>	21
1.2.	<i>Influence du niveau d'information contenu dans les données.....</i>	22
1.2.1.	<i>Sensibilité à l'âge maximum des poissons.....</i>	22
1.2.2.	<i>Sensibilité au niveau de bruit dans l'âge des poissons.....</i>	22
1.3.	<i>Influence de la quantité de données.....</i>	22
1.4.	<i>Sensibilité au prior sur L_{inf}.....</i>	23
1.5.	<i>Sensibilité au prior sur les paramètres de transition α et β.....</i>	23
1.6.	<i>Sensibilité au prior sur k_1 et k_2.....</i>	23
1.7.	<i>Conclusions de l'évaluation de la performance de la méthode.....</i>	23
2.	APPLICATION A L'ALBACORE DE L'OCEAN INDIEN	24
2.1.	<i>Modèle basé sur les données de lectures de microstries d'otolithes.....</i>	24
2.1.1.	<i>Ajustement de référence à partir des couples âge-taille.....</i>	24
2.1.2.	<i>Sensibilité au prior sur L_{inf}.....</i>	25
2.1.3.	<i>Sensibilité aux données de poissons de petites tailles.....</i>	25
2.2.	<i>Ajustement intégrant les deux sources d'information: lectures de microstries d'otolithes et données de marquage-recapture.....</i>	27
2.3.	<i>Modèle d'erreurs de mesure des lectures microstries d'otolithes et influence sur le modèle de croissance.....</i>	29
2.3.1.	<i>Modèle d'erreur de mesure.....</i>	29
2.3.2.	<i>Couplage avec l'ajustement de la courbe de croissance.....</i>	30
IV	DISCUSSION ET CONCLUSION	31
1.	SYNTHESE	31
1.1.	<i>Avancées méthodologiques.....</i>	31
1.2.	<i>Difficultés techniques rencontrées.....</i>	32
1.3.	<i>Connaissance sur la croissance des thons albacores.....</i>	32
2.	LES LIMITES.....	33
2.1.	<i>Les données.....</i>	33
2.2.	<i>Le modèle de croissance.....</i>	33
3.	PERSPECTIVES	34
3.1.	<i>Régler les problèmes techniques.....</i>	34
3.2.	<i>Amélioration du jeu de données.....</i>	34
3.3.	<i>Choix de modélisation.....</i>	35
3.4.	<i>Application à la gestion.....</i>	35
V	BIBLIOGRAPHIE	36

Table des figures

Figure 1 : Evolution du taux de croissance k en fonction du temps et exemple de courbe de croissance simulée illustrant un changement du paramètre k au cours du temps	4
Figure 2 : L'approche Bayésienne	6
Figure 3 : DAG du modèle d'estimation de la croissance à partir de couples de données âge-taille . $A[i]$ = âge de l'individu i ; $l[i]$ = taille de l'individu i ; $\theta = \{L_{inf}, \alpha, \beta, k_1, k_2, t_0\}$; σ = variance de l'erreur d'observation	8
Figure 4 : DAG du modèle d'estimation de la croissance à partir de données de marquage-recapture ($A1[j]$, $A2[j]$, $l1[j]$, $l2[j]$ et $TaL[j]$) respectivement l'âge au marquage et à la recapture, la taille au marquage et à la recapture et le temps en liberté du poisson i ; $\theta = \{L_{inf}, \alpha, \beta, k_1, k_2\}$	9
Figure 5 : DAG du modèle d'estimation de la croissance à partir de couples âge-taille et de marquage-recapture (A = âge ; l = taille ; TaL = durée de liberté ; $\theta = \{L_{inf}, \alpha, \beta, k_1, k_2\}$)..	10
Figure 6 : Les différents comptages effectués lors des lectures d'otolithe de thon albacore (en rouge les zones sources d'incertitude).....	13
Figure 7 : Ajustement obtenu dans le cas du scénario de référence de la méthode de simulation-estimation. Distributions <i>a posteriori</i> jointes (nuages de tirages MCMC) et marginales (diagonale)	21
Figure 8 : Estimations des paramètres obtenues à partir du jeu de données de lecture de microstries d'otolithes et les priors de référence (Tableau 3). Distributions <i>a posteriori</i> jointes (nuages de tirages MCMC) et marginales (diagonale).....	24
Figure 9 : Sensibilité de l'ajustement de la courbe de croissance au choix du prior sur le paramètre L_{inf} . (noir : prior à 110 cm ; bleu : prior à 130 cm ; pointillé : prior à 144 ; vert : prior à 150 cm ; rouge : prior à 170 cm). Les courbes sont tracées à partir des estimateurs ponctuels (médianes <i>a posteriori</i>) des paramètres.....	25
Figure 10 : Sensibilité de l'estimation de la courbe de croissance aux différents jeux de données d'otolithométrie (rouge : marquage-recapture, vert : ajustement sur données WSTTP, noir : les deux jeux de données). Les courbes sont tracées à partir des estimateurs ponctuels (médianes <i>a posteriori</i>) des paramètres.	26
Figure 11 : Ajustement de la courbe de croissance obtenu à partir du mélange des deux sources de données : lectures de microstries et marquage-recapture. Le nuage de points correspond aux poissons du jeu de données de marquage-recapture dont la taille est connue mais dont l'âge à la capture et à la recapture a été re-estimé. Les courbes sont tracées à partir des estimateurs ponctuels (médianes <i>a posteriori</i>) des paramètres.	29
Figure 12 : Régression linéaire du nombre de stries comptées en fonction de du temps en liberté TaL . La régression est tracée en fonction des estimateurs ponctuels (médiane <i>a posteriori</i>) des paramètres B et C	30

Table des tableaux

Tableau 1 : Récapitulatif du nombre de poissons de la base de données et de la sélection	11
Tableau 2 : Distributions <i>a priori</i> de référence utilisées lors de l'étude de simulation-estimation	17
Tableau 3 : Priors sur les paramètres de croissance de l'ajustement de référence.....	19
Tableau 4 : Sensibilité de l'estimation de la courbe de croissance aux différents jeux de données (estimateurs ponctuels (médianes a posteriori) des paramètres).....	27
Tableau 5 : Sensibilité de l'estimation de la courbe de croissance aux différents jeux de données de 1000 marquage-recapture (estimateurs ponctuels (médianes a posteriori) des paramètres).....	28

Table des annexes

Annexe 1 : a) Histogrammes des distributions des tailles des thons albacores du jeu de données original au marquage et à la recapture b) après sélection	I
Annexe 2 : Graphique représentant la sélection sur le taux de croissance des thons albacores (en rouge : données exclues).	II
Annexe 3 : a) Histogrammes des distributions de tailles des thons albacores du jeu de données d'otolithométrie au marquage et à la recapture b) Histogramme de la distribution de tailles à la capture des thons albacores issus de la campagne complémentaire.....	III
Annexe 4 : EQM et ERM des ajustements des paramètres de la fonction de croissance en fonction de l'âge maximum de marquage.	IV
Annexe 5 : EQM et ERM des ajustements des paramètres de la fonction de croissance en fonction du coefficient de variation appliqué au bruit sur l'âge.	IV
Annexe 6 : Résidus obtenus à l'issue de l'ajustement du modèle à l'intégralité des données de marquage-recapture (L_{inf} fixé)	V
Annexe 7 : Distributions <i>a posteriori</i> du paramètre L_{inf} dans le cas d'un modèle intégrant 1000 données de marquage-recapture et intégrant seulement les données d'otolithes.....	V
Annexe 8: Code WinBUGS du modèle intégrant les données de marquage-recapture et les données de lectures de microstries d'otolithes.	VI
Annexe 9: Format des données de marquage-recapture et les données de lectures de microstries d'otolithes.....	IX

I Introduction

Les stocks d'albacore (*Thunnus albacares*), de listao (*Katsuwonus pelamis*) et de patudo (*Thunnus obesus*) de l'océan Indien sont intensément exploités depuis l'arrivée des premiers palangriers japonais au début des années 1950. Au cours des années 2000, les captures de ces trois espèces de thons tropicaux étaient de l'ordre de 1 million de tonnes par an, ce qui représentait environ 25% des captures mondiales de thons. La pêcherie thonière tropicale de surface de l'océan Indien représente des captures ayant varié entre 300 000 et 400 000 t au cours des dix dernières années. Elle est essentiellement représentée par les senneurs français et espagnols dont la flottille est constituée d'une cinquantaine de navires.

La gestion des stocks de grands pélagiques de l'océan indien est de la responsabilité de la Commission des Thons de l'Océan Indien (CTOI, <http://www.iotc.org/>). Le suivi de l'état des ressources de l'OI passe par des modèles qui intègrent les connaissances sur les processus biologiques, écologiques et démographiques qui gouvernent la productivité et la capacité de résilience des stocks à la pêche et aux changements globaux tels que le réchauffement climatique. La compréhension de la dynamique des populations de thons nécessite ainsi d'étudier des processus démographiques tels que la croissance, les mortalités naturelles et par pêche, les mouvements et migrations.

Dans ce travail nous nous focaliserons particulièrement sur la croissance. C'est un facteur majeur de la productivité des populations qui a été largement étudiée au cours des dernières décennies. La croissance des poissons peut être estimée à partir de différentes sources de données issues de la sclérochronologie (écailles, épines dorsales, otolithes, vertèbres ; Panfili et al., 2002), de l'analyse des structures de tailles des captures d'une pêcherie, e.g. le suivi de la progression du mode des histogrammes de taille des captures au cours du temps (Viera, 2005) et de données de marquage-recapture.

La démarche mise en œuvre dans cette étude est générale mais elle se focalise sur le cas du thon albacore. La croissance des thons albacores de l'océan Indien a été estimée à partir de données issues de lectures de stries vertébrales (Romanov et Korotkova, 1988), de lectures de microstries d'otolithes (Stequert et al., 1996) et de fréquences de taille obtenues à partir des différentes pêcheries de l'océan Indien (Marcille et Stéquert, 1976; Marsac et Lablache, 1985; Lumineau, 2002; Viera, 2005). Les ajustements de courbes de croissance du thon albacore de l'Océan Indien à partir de ces données ont mis en évidence des résultats contradictoires. Certaines estimations aboutissant à une courbe de croissance suivant une loi de Von Bertalanffy (Marcille et Stéquert, 1976; Stequert et al., 1996), alors que d'autres ont conclu à une croissance en deux stances comme pour le thon albacore de l'Atlantique et du Pacifique (Marsac et Lablache, 1985; Marsac, 1991; Lumineau, 2002). Ces résultats remettent en cause la qualité ou la représentativité des données utilisées et témoignent de la nécessité d'acquérir de nouveaux jeux de données afin d'aboutir à des conclusions plus fiables sur la forme et les paramètres de la courbe de croissance de l'albacore dans l'océan Indien.

La connaissance de la variabilité individuelle de croissance présente également un intérêt majeur pour convertir les matrices de captures disponibles en fréquences de taille à l'échelle des pêcheries, en matrices de fréquences d'âge qui sont utilisées comme données d'entrée dans les modèles d'évaluation utilisés au sein des groupes de travail de la CTOI. Un vaste projet régional de marquage de thons dans l'Océan Indien (Regional Tuna Tagging Project ; RTTP-IO; Hallier et Million, 2009) a été mis en place à partir de 2005, financé par l'Union Européenne à hauteur de 14 millions d'euros. L'objectif de ce projet était d'acquérir un jeu de

données indépendant de la pêche afin d'apporter de nouvelles informations sur la dynamique des populations de thons tropicaux. Ce projet mené par la Commission de l'Océan Indien (www.coi-ioc.org) et supervisé par la CTOI a permis de marquer entre mai 2005 et août 2007 168 158 thons, dont 32% d'albacore, 21% de patudo et 47% de listao. Au jour d'aujourd'hui, 27 384 thons marqués ont été recapturés dont 9 498 albacores, 5 319 patudos et 12 567 listaos. Au cours du projet, près de 4 400 albacores et patudos ainsi qu'environ 1 500 listaos ont reçu une injection d'oxytétracycline qui laisse une marque fluorescente dans les pièces calcifiées (i.e. os, otolithes, épines dorsales), afin de valider les lectures d'âge réalisées sur les otolithes par le laboratoire de l'IRD/LEMAR de Brest (Morize et al., 2008).

Ce travail vise à développer un modèle statistique pour l'estimation des courbes de croissance dans une approche permettant d'intégrer différentes sources de données dans l'analyse et de prendre en compte les différentes sources d'incertitude liées à aléa de la croissance et aux erreurs de mesure. En effet, une approche intégrée permet aux différentes sources de données de se compléter et de fournir des estimations plus précises, plus complètes et plus robustes des paramètres de croissance (Laslett et al., 2002; Eveson et al., 2004; Polacheck et al., 2004).

Cette étude est menée dans un cadre statistique Bayésien (Gelman, 2004), mathématiquement rigoureux et adapté à la prise en compte des incertitudes (Clark, 2003). La méthode Bayésienne développée permet de modéliser la croissance des poissons en intégrant des données provenant de sources multiples. La démarche s'appuie sur des données issues des programmes de marquage-recapture et sur des données de détermination directe de l'âge par l'analyse des structures anatomiques dures, i.e. les otolithes.

Les objectifs de l'analyse sont triples:

- (1) Il s'agit tout d'abord d'évaluer la performance de la méthode d'estimation par une démarche de simulation-estimation. Ce premier volet permettra d'évaluer la faisabilité, l'intérêt et la pertinence d'un modèle de croissance intégrant à la fois des données de sclérochronologie (lecture d'otolithes) et de marquage-recapture dans un cadre statistique Bayésien ;
- (2) Le deuxième objectif consiste à appréhender l'incertitude autour des estimations d'âges à partir des lectures de microstries d'otolithes. Les lectures des pièces calcifiées sont sujettes à caution car les sources d'incertitudes sont multiples. Un modèle d'erreur de mesure pour la lecture des otolithes prenant en compte différentes sources d'incertitude dans le processus de lecture des microstries sera construit ;
- (3) Enfin, il s'agit d'appliquer cette méthodologie au cas du thon albacore de l'océan Indien à partir des données fournies par le RTTP-IO. L'ajustement est effectué sur la courbe de croissance « VB log K » (Laslett et al., 2002) qui a été considérée par le groupe de travail sur les données de marquage (IOTC 2008) comme la plus adaptée pour décrire la croissance de l'albacore de l'océan Indien.

II Matériels et méthodes

1. Modèles de croissance

1.1. Différents modèles proposés dans la littérature

Le modèle de croissance décrit l'évolution de la taille en fonction de l'âge. Différentes formes fonctionnelles de courbes de croissance ont été proposées dans la littérature : le modèle de Von Bertalanffy (VB) ; (Von Bertalanffy, 1938), le modèle de Gompertz (Gompertz, 1825) ou encore le modèle de Richards (Richards, 1959). Schnute (1981) a proposé un modèle générique de croissance basé sur 4 paramètres ayant une interprétation biologique et qui permet de représenter une large gamme de modèles de croissance. Cependant, pour les thons tropicaux, il a été suggéré depuis le début des années 1980 qu'une courbe de type VB ne permettait pas de représenter la complexité des processus de croissance des individus au cours du temps, notamment en raison de la succession de phases d'accélération et de ralentissement (stances) de la croissance (Fonteneau, 1980; Bard, 1984).

1.2. Modèles à stances concernant l'albacore

Des données historiques de marquage et de fréquence de taille des captures dans l'océan Atlantique ont en effet mis en évidence des ralentissements de croissance pour des tailles comprises entre 40 et 65 cm environ, i.e. pour des individus d'âge compris entre 6 et 22 mois environ (Fonteneau, 1980; Bard, 1984; Gascuel et al., 1992). Des résultats similaires ont été mis en évidence dans le sud-ouest Pacifique avec des taux de croissance faibles de l'ordre de 1.3 cm par mois pour des tailles à la fourche comprises entre 30 et 50 cm (Brouard et al., 1984). De faibles taux de croissance pour les jeunes albacores compris entre 35 et 57 cm ont également été estimés à partir de données de fréquences de taille dans les captures des flottilles de senneurs européens de l'océan Indien (Marsac et Lablache, 1985; Marsac, 1991). Les premières analyses conduites sur les données d'otolithométrie récoltées au cours du RTTP-IO suggèrent également une croissance en deux stances caractérisée par un ralentissement pour les poissons juvéniles (Morize et al., 2008).

Pour tenir compte de tels changements de croissance, Laslett et al. (2002) ont proposé une extension du modèle de VB avec un taux de croissance variant de façon logistique au cours de la vie de l'individu en s'appuyant sur le modèle de Wang (1998). Ce modèle baptisé « VB log K » a été utilisé pour le thon rouge du sud (*Thunnus macoyii*) qui présente tout comme les thons tropicaux une transition de croissance au cours de son cycle de vie (Laslett et al., 2002; Eveson et al., 2004). Ce modèle a été choisi comme cadre de référence pour la modélisation de la croissance de l'albacore de l'océan Indien au sein de la CTOI (Eveson et Million, 2008; Hillary et Million, 2008). Il sera considéré comme le modèle de référence dans toute notre étude.

Les hypothèses du modèle « VB log K » sont les suivantes :

Soit un poisson de taille nulle à l'âge t_0 , le taux de croissance à l'âge t s'exprime ainsi :

$$(1) \quad l'(t) = (L_{\text{inf}} - l(t))k(t)$$

Avec $k(t)$, la fonction de contrôle du taux de croissance pouvant dépendre du temps. La solution générale de cette équation est de la forme :

$$(2) \quad l(t) = L_{\text{inf}} (1 - \exp(-K(t_0, t)))$$

$$\text{Avec } K(t_0, t) = \int_{t_0}^t k(u) du$$

On sait que la croissance de certains poissons, e.g les thons albacores et patudo, change de façon importante entre la phase juvénile et la phase adulte (Hearn et Polachek, 2002). Le modèle proposé par (Wang, 1998) modélise cette variation de la croissance par un changement brutal du taux de croissance dans le temps :

$$(3) \quad k(t) = \begin{cases} k_1 & \text{pour } t \leq t_0 + \alpha \\ k_2 & \text{pour } t > t_0 + \alpha \end{cases}$$

Dans ce modèle, l'âge α , l'âge auquel la transition a lieu, est interprété comme l'âge moyen auquel les juvéniles deviennent adultes. D'un point de vue biologique, cette fonction par palier (*step function*) est difficile à justifier, de plus elle implique des problèmes lors de l'analyse statistique (Hearn et Polachek, 2002). Ainsi Laslett et al. (2002) ont proposé une extension continue du modèle de Wang (1998). Dans ce modèle, $k(t)$ suit une courbe logistique qui permet une transition continue entre les deux niveaux de taux de croissance.

$$(4) \quad k(t) = k_1 + (k_2 - k_1) \frac{1}{1 + \exp(-\beta(t - t_0 - \alpha))}$$

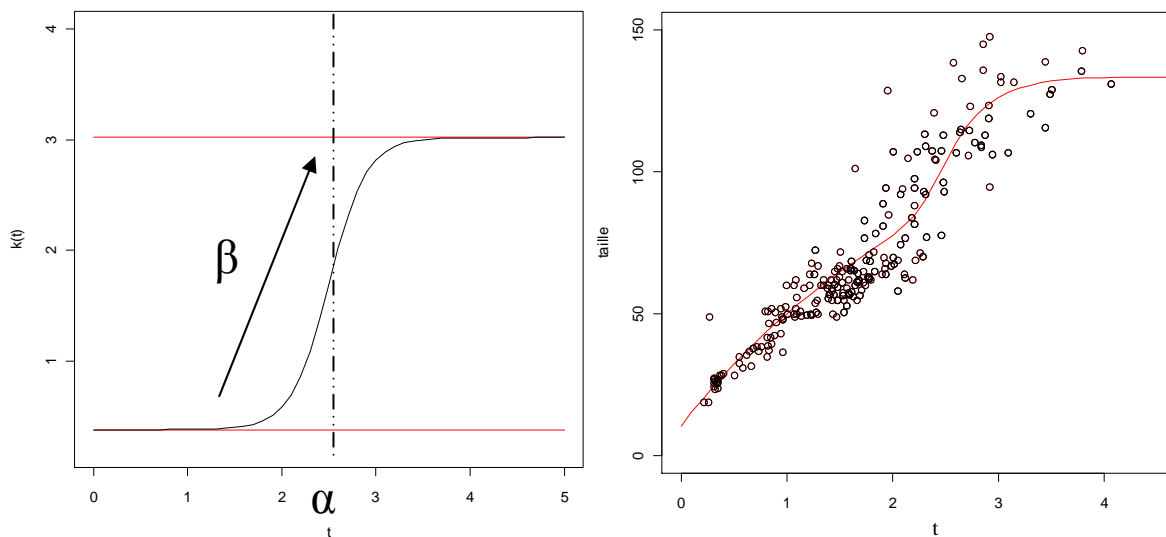


Figure 1 : Evolution du taux de croissance k en fonction du temps et exemple de courbe de croissance simulée illustrant un changement du paramètre k au cours du temps

La transition entre les paliers k_1 et k_2 est réglée par deux paramètres ; α fixe le point d'inflexion de la courbe de transition alors que la vitesse de transition est dirigée par β . Plus β est grand, plus la vitesse de transition est importante. Le cas limite $\beta \rightarrow \infty$ correspond à $k(t)$ suivant la fonction par palier (cf. précédent). Pour $t \ll t_0 + \alpha$, $k(t) = k_1$ et pour $t \gg t_0 + \alpha$,

$k(t) = k_2$. L'intérêt de la forme logistique (5) est que le calcul de l'intégrale sur le temps nécessaire à l'expression de la courbe de croissance (3) admet une solution analytique, dite courbe de croissance de « VB log K ».

$$(5) \quad l(t) = L_{\text{inf}} \left[1 - e^{-k_2(t-t_0)} \left\{ \frac{1 + \exp(-\beta(t-t_0 - \alpha))}{1 + e^{\beta\alpha}} \right\}^{-(k_2-k_1)/\beta} \right] \quad \text{pour } t > t_0$$

Le modèle apparaît comme flexible dans sa capacité à capturer l'existence de deux stances de croissance. Néanmoins, il comporte six paramètres, et on peut s'attendre à rencontrer des difficultés statistiques d'identifiabilité de ces paramètres.

2. Estimation de la courbe de croissance

2.1. Intégration de multiples sources d'information et des erreurs de mesures

D'une façon générale et en particulier pour l'albacore de l'océan Indien, des données de natures différentes sont disponibles pour estimer les courbes de croissance. Les données les plus riches sont les couples âge-taille obtenus par sclérochronologie (e.g. lecture d'écaille, comptage de microstries d'otolithes) grâce à la lecture des pièces calcifiées. Pour chaque poisson i on dispose d'une mesure de taille $l[i]$ et d'âge $A[i]$. Si le même poisson a été capturé et mesuré plusieurs fois (occasions de captures j) et qu'à chaque occasion son âge est estimé, (voir paragraphe 3.2.1) on dispose d'un ensemble de mesure $l[i,j]$ et d'âge $A[i,j]$. Néanmoins ces données sont très coûteuses et donc rares. De plus ces lectures sont largement sujettes aux erreurs de mesures, particulièrement dans le cas des poissons tropicaux où les variations saisonnières sont peu marquées. Les programmes de marquage-recapture fournissent des données plus nombreuses mais moins informatives sur l'âge. Pour chaque poisson i , on dispose de $l_1[i]$ et $l_2[i]$, respectivement la taille au marquage et la taille à la recapture, ainsi que de $\Delta t_{1 \rightarrow 2}[i]$, le temps écoulé entre le marquage et la recapture, aussi appelé « temps de liberté » et noté $Tal[i]$ (pour « *time at liberty* ») dans la suite. Dans ce cas, on dispose d'information sur le taux de croissance du poisson i mais à aucun moment son âge absolu n'est connu (Eveson et al., 2004). Il est donc intéressant de proposer des méthodes capables d'intégrer simultanément ces différentes sources d'information.

Par ailleurs ces différents types de mesures intègrent des erreurs d'observation (ou de mesure) qui leur sont propres. Ne pas les prendre en compte peut conduire à des estimations biaisées et/ou sur-confiantes des paramètres.

Une autre difficulté des modèles de croissance réside dans le peu de connaissance dont on dispose sur les individus les plus âgés. Bien souvent, les jeux de données contiennent peu d'information concernant ces individus, et ainsi les paramètres tels que la taille asymptotique à un temps infini (L_{inf}) sont souvent bien peu renseignés.

2.2. Le choix du cadre Bayésien

Les statistiques Bayésiennes, couplée aux méthodes d'estimation numériques par échantillonnage (e.g. méthodes de Monte Carlo sous Winbugs utilisant l'échantillonneur de Gibbs), offrent des solutions techniques et pragmatiques permettant d'intégrer les différentes sources d'information et les multiples sources d'incertitudes dans les modèles (Meyer et Millar, 1999; Rivot et al., 2004; Clark, 2005; Cressie et al., 2009). De plus, par l'intermédiaire des distributions a priori, l'outil Bayésien permet d'intégrer dans la l'analyse l'information

issue de l'expertise (Hilborn et Liermann, 1998; McCarthy et Masters, 2005). Cela peut permettre par exemple de compenser le manque d'information d'un jeu de données (Box 1-2).

Box 1 – Principe de l'analyse Bayésienne

Les statistiques Bayésiennes offrent un cadre de référence pour la modélisation stochastique et statistique, l'inférence et la prédiction dans de nombreux domaines : sciences sociales, économie, finance, biomédical, environnement, etc. (Efron, 1986; Gelman, 2004; Clark, 2005; Parent et Bernier, 2007). Dans le contexte de la dynamique des populations en halieutique, l'approche Bayésienne est très utilisée depuis les années 1990 (Punt et Hilborn, 1997; McAllister et Kirkwood, 1998).

Dans le cadre Bayésien, l'incertitude sur l'état des connaissances des variables inconnues d'un modèle, θ , s'exprime sous la forme d'une distribution de probabilité. La spécification du modèle nécessite de définir la fonction de vraisemblance, comme dans tout modèle statistique, mais également de définir une distribution de probabilité *a priori* (prior) pour les inconnues $p(\theta)$. Cette étape offre la possibilité d'introduire de l'information *a priori* dans le modèle à partir de connaissances d'experts ou de l'analyse de données externes (Punt et Hilborn, 1997). C'est un avantage essentiel dans le développement de nombreux modèles en environnement « data poor ». La formule de Bayes permet de mettre à jour la loi *a priori* par les données par l'intermédiaire de la fonction de vraisemblance pour obtenir une distribution *a posteriori* $p(\theta / y)$.

$$(6) \quad p(\theta / y) = \frac{p(\theta)p(y / \theta)}{\int p(\theta)p(y / \theta)d\theta}$$

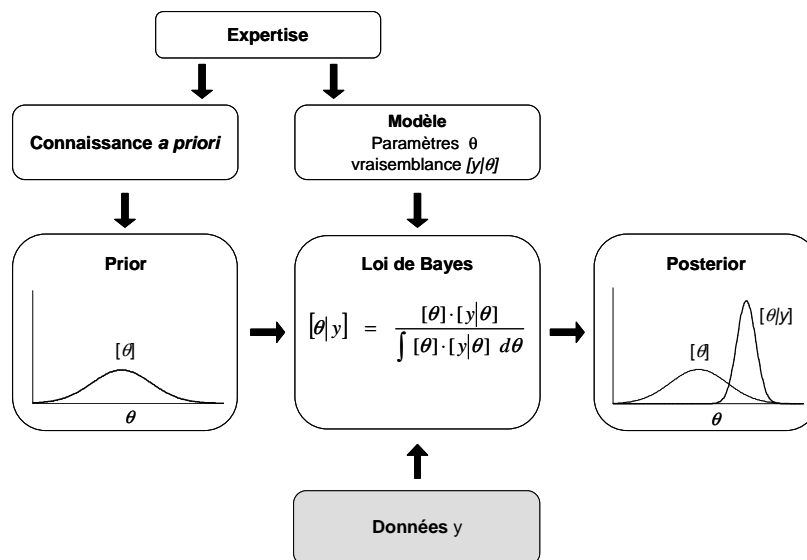


Figure 2 : L'approche Bayésienne

Box 2 - L'analyse Bayésienne en pratique

En pratique, l'estimation de la distribution jointe *a posteriori* des paramètres s'appuie sur de méthodes de simulation de Monte Carlo. L'échantillonnage s'effectue grâce à des algorithmes stochastiques. Deux grandes familles d'algorithmes existent (Gilks et al., 1996; Gelman, 2004) : MCMC (Monte Carlo par Chaîne de Markov) et S.I.R. (Sampling Importance Resampling). Le logiciel libre OpenBugs (Bayesian inference Using Gibbs Samples) (<http://www.openbugs.info/w/>) utilisé lors de cette étude fonctionne sur la base d'un algorithme de type MCMC. Il peut être interfacé avec le logiciel libre R via certains packages comme BRugs ou rbugs.

Dans la pratique, la construction des modèles Bayésiens s'appuie largement sur le raisonnement conditionnel et sur la modélisation graphique à l'aide des Graphes Acycliques Orientés (DAG). Dans un DAG, les ellipses et les ellipses grisées représentent respectivement les quantités aléatoires et les quantités aléatoires observables. Les flèches symbolisent les relations de conditionnement probabiliste. Les nœuds apparaissant au sommet d'un DAG représentent les paramètres sur lesquels des priors doivent être définis. Les constantes sont représentées par des rectangles grisés.

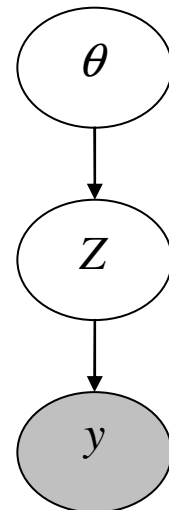
Ainsi un DAG de la forme suivante représente le modèle défini par :

$$\theta \sim p(\theta) \quad \text{prior}$$

$$z|\theta \sim \text{Distribution conditionnelle de la variable latente } Z \text{ sachant } \theta$$

$$y|z \sim \text{Distribution conditionnelle de l'observation } y \text{ sachant } Z$$

(fonction de vraisemblance)



L'inférence Bayésienne consiste à établir la distribution *a posteriori* jointe des paramètres θ et des variables latentes Z conditionnellement aux observations y : $p(\theta, z|y)$. Une fois le DAG construit, la programmation des modèles dans des logiciels type BUGS est simplifiée car le code prend une forme déclarative qui suit la structure de conditionnement du DAG.

2.3. Principes des modèles d'estimation de la courbe de croissance

Nota :

- (1) Dans ce travail, la variabilité inter-individuelle des paramètres de croissance n'est pas prise en compte. Tous les paramètres sont considérés comme fixes au sein de la population. Ce point sera soulevé en discussion.
- (2) Le modèle d'erreur de mesure sur la lecture d'âge sera présenté plus tard (paragraphe 3.2.4)

2.3.1. Ajustement à partir de couples âge-taille

Dans une première étape, un modèle est développé pour ajuster une courbe de croissance sur des couples de données âge-taille. Ce modèle ne prend pas en compte l'incertitude autour de l'estimation de l'âge par otolithométrie. L'équation d'observation (7) permet de relier les observations individuelles âge-taille à l'état moyen de la population en supposant que l'ensemble de la variabilité du modèle provient des erreurs d'observation. L'équation (7) permet d'exprimer la taille en fonction de l'âge (où i représente un poisson et j définit l'occasion de capture), en faisant l'hypothèse d'une relation de croissance de VB log K de paramètres $\theta = \{L_{\text{inf}}, \alpha, \beta, k_1, k_2, t_0\}$ et d'une variabilité log-normale autour du modèle moyen :

$$(7) \quad l[i, j] = L_{\text{inf}} \left[1 - e^{-k_2(A[i, j] - t_0)} \left\{ \frac{1 + \exp(-\beta(A[i, j] - t_0 - \alpha))}{1 + e^{\beta\alpha}} \right\}^{-(k_2 - k_1)/\beta} \right] \times e^{\varepsilon_i}$$

Avec $\varepsilon_{i, j} \sim N(0, \sigma^2)$

Pour chacun des paramètres de la courbes de croissance une distribution *a priori* est définie. L'ajustement Bayésien du modèle permet d'estimer le jeu de paramètres $\theta = \{L_{\text{inf}}, \alpha, \beta, k_1, k_2, t_0\}$ et de la variance σ^2 .

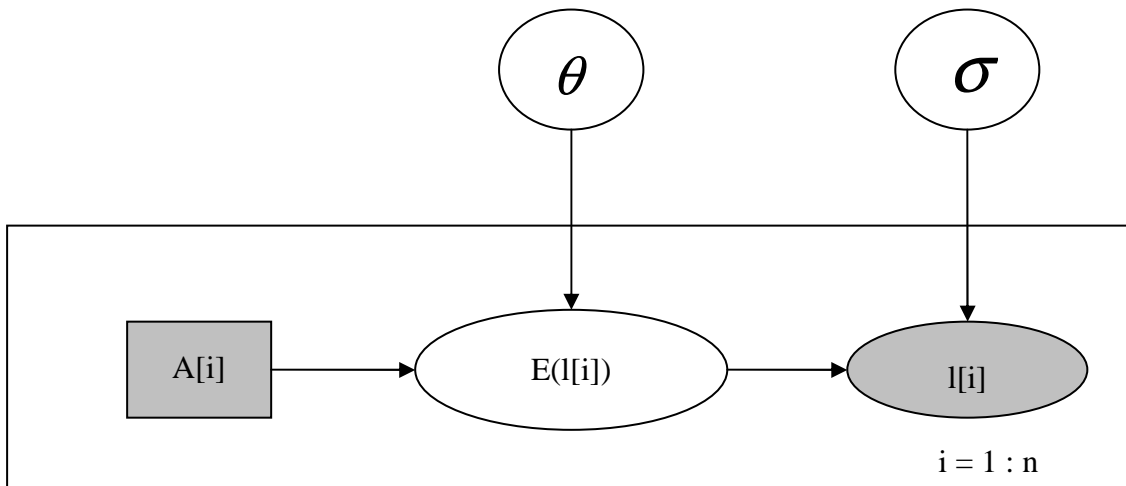


Figure 3 : DAG du modèle d'estimation de la croissance à partir de couples de données âge-taille . $A[i]$ = âge de l'individu i ; $l[i]$ = taille de l'individu i ; $\theta = \{L_{\text{inf}}, \alpha, \beta, k_1, k_2, t_0\}$; sigma = variance de l'erreur d'observation

2.3.2. Ajustement à partir de couples de marquage-recapture

Ce modèle permet d'ajuster une courbe de croissance sur des couples de données de marquage-recapture. Les données de marquage-recapture ne fournissent pas d'information directe sur l'âge des poissons. Le paramètre t_0 (qui fixe « l'origine des temps ») n'est pas estimable à partir de ces données. Cependant le changement de taille d'un poisson entre la date de marquage (l_1 au temps t_1) et de recapture (l_2 au temps t_2), i.e. au cours du temps de liberté (*TaL*), fournit une information importante pour la modélisation de la croissance.

L'ajustement du modèle permet d'estimer le jeu de paramètres $\theta_{mr} = \{L_{\text{inf}}, \alpha, \beta, k_1, k_2\}$. Le modèle construit est le suivant :

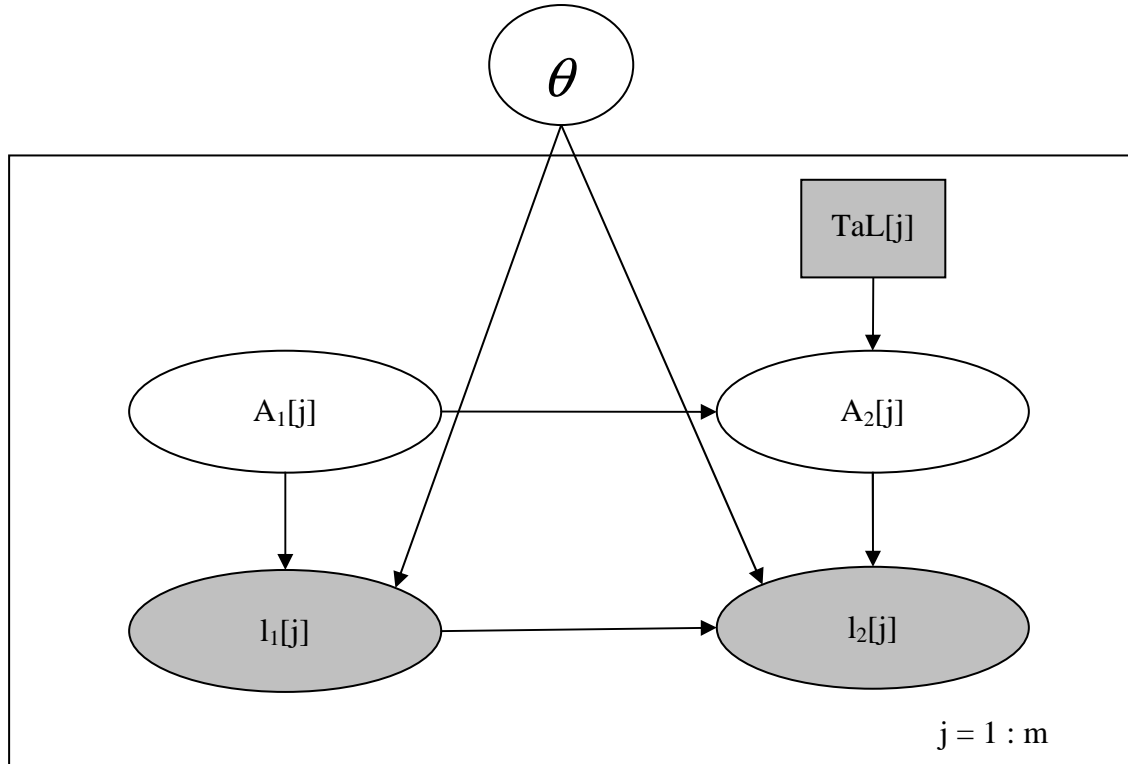


Figure 4 : DAG du modèle d'estimation de la croissance à partir de données de marquage-recapture ($A_1[j]$, $A_2[j]$, $l_1[j]$, $l_2[j]$ et $TaL[j]$) respectivement l'âge au marquage et à la recapture, la taille au marquage et à la recapture et le temps en liberté du poisson i ; $\theta = \{L_{\text{inf}}, \alpha, \beta, k_1, k_2\}$

$$(8) \quad \begin{cases} l_1[i] = L_{\text{inf}} \left[1 - e^{-k_2(A_1[i]-t_0)} \left\{ \frac{1 + \exp(-\beta(A_1[i]-t_0 - \alpha))}{1 + e^{\beta\alpha}} \right\}^{-(k_2-k_1)/\beta} \right] \times e^{\varepsilon_{1,i}} \\ l_2[i] = L_{\text{inf}} \left[1 - e^{-k_2(A_2[i]-t_0)} \left\{ \frac{1 + \exp(-\beta(A_2[i]-t_0 - \alpha))}{1 + e^{\beta\alpha}} \right\}^{-(k_2-k_1)/\beta} \right] \times e^{\varepsilon_{2,i}} \end{cases}$$

où l_1 et l_2 sont respectivement la taille du poisson au marquage et à la recapture et A_1 et A_2 respectivement l'âge estimé au marquage et à la recapture.

Les âges A_1 et A_2 sont des variables latentes aléatoires inconnues. L'âge A_1 est supposé distribué selon une loi Gamma peu informative et l'âge A_2 s'obtient en ajoutant le temps de liberté TaL supposé connu sans erreur :

$$(9) \quad \begin{aligned} A_1[i] &\sim \text{gamma}(\mu, CV) \\ A_2[i] &= A_1[i] + TaL[i] \end{aligned}$$

Considéré indépendamment, ce modèle ne permet pas l'estimation de l'âge des poissons. Il nécessite d'être couplé avec un modèle utilisant des données renseignant sur l'âge absolu ou

d'être complété par une hypothèse sur t_0 . Par ailleurs, il ne prend pas en compte l'incertitude autour de l'estimation de l'âge par otolithométrie.

2.3.3. Ajustement à partir de couples de marquage-recapture et des couples âge-taille

Ce modèle couple les deux modèles développés en 2.3.1 et en 2.3.2. Il permet d'intégrer simultanément les données des couples âge-taille et de marquage-recapture au sein d'un même modèle, afin d'estimer conjointement tous les paramètres de croissance.

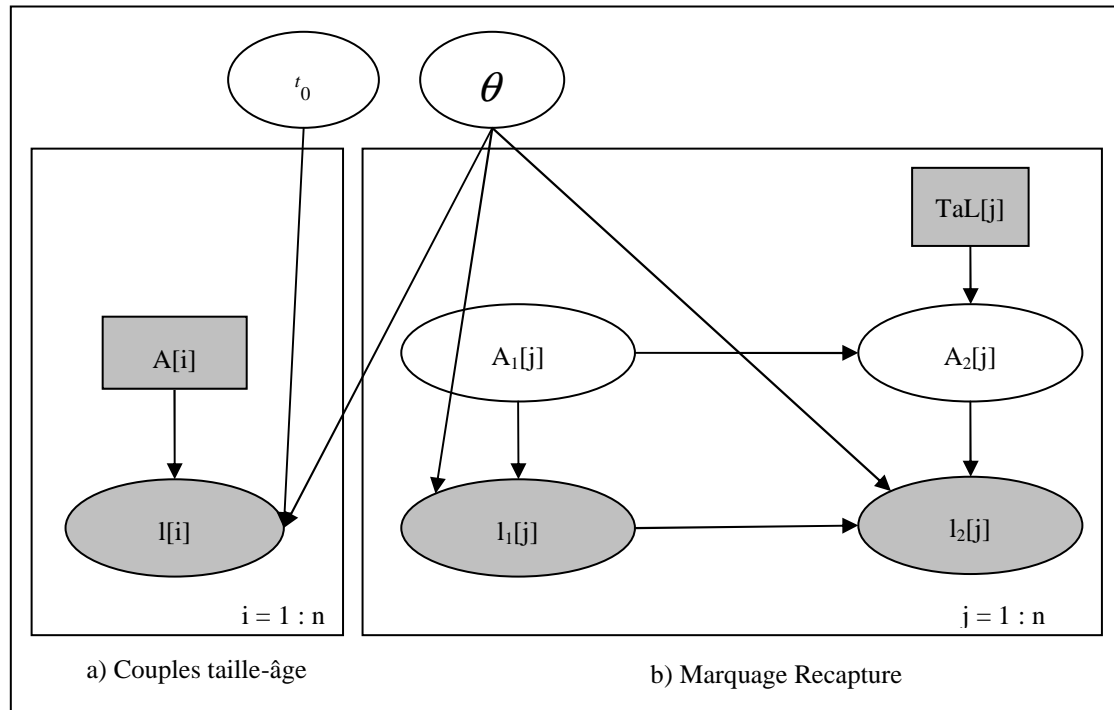


Figure 5 : DAG du modèle d'estimation de la croissance à partir de couples âge-taille et de marquage-recapture (A = âge ; l = taille ; TaL = durée de liberté ; $\theta = \{L_{inf}, \alpha, \beta, k_1, k_2\}$)

3. Les données de marquage-recapture et d'otolithométrie pour l'albacore de l'océan Indien.

3.1. Le jeu de données de marquage-recapture

3.1.1. Présentation du jeu de données

Le jeu de données issu de la campagne de marquage RTTP-IO (Tableau 1) comprenait au moment du stage un total de 27 384 thons recapturés dont 9 498 albacores (35 %), 5 319 patudos (19 %) et 12 567 listaos (46 %), soit plus de 16 % de recaptures recaptures essentiellement réalisées par les senneurs européens. Les données sont actuellement disponibles dans une base de données MS ACCESS gérée par J. P. Hallier (IRD) et J. Million (CTOI) et sont accessibles sur demande auprès de la CTOI.

3.1.2. Sélection des données

Le jeu de données de marquage-recapture est renseigné avec précision et permet donc une sélection franche. Plusieurs sources d'erreurs, potentielles ou évidentes, ont été identifiées et

supprimées du jeu de données original. Le but de cette sélection est d'obtenir un jeu de données homogène et fiable. Les critères de sélections sont les suivants :

- Les poissons dont l'une des mesures était considérée comme peu fiable dans la base de données selon les critères définis par l'équipe du RTTP ont été exclus de la sélection.
- Seuls les poissons mesurés de la tête à la fourche (FL) ont été sélectionnés. Les mesures de poissons réalisées en longueur pré-dorsale (FDL) ou longueur totale (CL) impliquant une conversion en longueur à la fourche ont été considérées peu précises et exclues de la sélection. Les poissons dont la mesure à la recapture n'a pas été effectuée avec un pied à coulisse ont également été exclus.
- La récupération des marques a essentiellement été réalisée par les senneurs (93% des recaptures) à bord desquels plusieurs coups de senne (ou calées), éventuellement à plusieurs jours d'intervalle, peuvent être placés dans une même cuve. Puisque l'information disponible est le numéro de la cuve, la récupération de poissons marqués lors du débarquement peut ainsi générer une incertitude sur la date de pêche en mer. Les poissons dont la date de recapture était non renseignée ou comportait une incertitude (date maximale – date minimale de recapture) supérieure à 10% du temps passé en mer ont été exclus de l'analyse.
- Les poissons dont l'espèce enregistrée au marquage différait de celle à la recapture ont été éliminés de l'analyse.
- Le dispositif de récupération des données a évolué au cours du RTTP-IO pour se stabiliser début 2007. Les informations concernant la date et la taille de recapture sont jugées plus fiables à partir de cette date (J.P. Hallier, pers. Com.). La sélection comprend donc exclusivement les poissons débarqués à partir de 2007.
- Certains couples de marquage-recapture ont un taux de croissance négatif, ces données ont été exclues du jeu de données utilisé lors de l'analyse. Pour limiter le biais engendré par cette suppression, il a été choisi de sélectionner les taux de croissance compris entre le quantile 2% et le quantile 98 % (Annexe 2).

La sélection obtenue à l'aide de ces critères aboutit à un jeu de données plus homogène et plus fiable tout en étant constitué d'un nombre d'individus importants (Tableau 1, Annexe 1).

Tableau 1 : Récapitulatif du nombre de poissons de la base de données et de la sélection

Espèce	Nombre total de recaptures	Recaptures sélectionnées
Albacore	9498	3781
Patudo	5319	2736
Listao	12567	4726
Total	27384	11243

Cette sélection préserve la structure des données (Annexe 1). Dans le cas de l'albacore, les tailles à la capture sont ainsi toujours distribuées autour d'un mode et 90% d'entre elles sont comprises entre 40 et 76 cm. Les tailles à la recapture ont quant à elles une distribution bimodale ayant pour modes 60 et 110 cm. Cette bimodalité serait due à une sélectivité de la senne. Cette structure est légèrement accentuée par la sélection.

3.1.3. Sources d'erreurs de mesure de la taille

La mesure des poissons lors du marquage est faite sur des poissons vivants à l'aide de table de mesure. Le cahier des charges de la campagne de marquage n'autorisait qu'une manipulation très rapide des poissons (< 15 secondes hors de l'eau). Les mesures de ces poissons sont donc relativement précises. En revanche, les mesures à la recapture sont souvent plus précises, l'animal est mort ce qui permet une manipulation plus aisée et l'utilisation d'outils plus précis (i.e. pied à coulisse). Certains effets de la congélation peuvent néanmoins affecter les mesures (courbure du poisson, etc.) mais ils ont généralement été pris en compte par l'équipe du RTTP lorsqu'indiqués lors des mesures.

3.2. Le jeu de données d'otolithométrie

3.2.1. Présentation des données

Des estimations d'âge absolu par lecture de microstries d'otolithes ont été réalisées au sein du LEMAR (Laboratoire des Sciences de l'Environnement Marin) de Brest par l'équipe d'Eric Morize (IRD) pour 207 poissons. Ces estimations viennent de deux jeux de données distincts :

(1) Les poissons issus du programme RTTP-IO sont des individus qui ont été marqués à l'oxytétracycline (OTC), injectée lors du marquage dans les parties musculaires à l'aide d'une seringue. L'injection est de 1.5 à 3 ml d'OTC en fonction de la taille du poisson et prend quelques secondes. L'OTC a la propriété de marquer les pièces calcifiées des poissons (os, écailles et otolithes). Les otolithes issus du jeu de données de marquage-recapture ont donc une marque visible résultant de l'injection de l'OTC lors du marquage. Parmi les 2018 thons albacores marqués, 230 ont été recapturés et 147 otolithes ont été lus. On dispose pour ces poissons d'information sur l'âge à la recapture estimé par otolithométrie, de la date de marquage, de la date de recapture ainsi que des tailles associées (Annexe 3a). En comparant le délai entre le marquage et la recapture et le nombre de stries comptées après la marque, il est possible de valider l'hypothèse qu'une strie soit déposée chaque jour (Wild et al., 1995; Stequert et al., 1996). On obtient l'âge au marquage de chacun des poissons. On dispose donc potentiellement de deux couples taille- âge par poisson.

(2) L'autre jeu de données est issu de la campagne West Sumatra Tuna Tagging Project (WSTTP) menée en 2006 et 2007. Ce jeu de données contient 60 individus dont des poissons de petites tailles (18 poissons entre 19 et 30 cm) et des individus de grandes tailles (10 individus compris entre 130 cm et 147.5 cm). Ne faisant pas partie d'un programme de marquage-recapture, ces poissons n'ont fait l'objet que d'une lecture de microstries lors de la capture. On ne dispose pour ces poissons que d'un seul couple âge-taille (Annexe 3b). Ces données apportent une information complémentaire aux données issues du RTTP-IO.

3.2.2. Lecture de l'âge

Les lectures de microstries d'otolithes ont été effectuées au sein du Laboratoire de Sclérochronologie des Animaux Aquatiques (LASAA) par l'équipe d'Eric Morize.

BOX 3 - Préparation des otolithes

Les otolithes (sagittae) sont préparés suivant les méthodes décrites par Secor et al. (1991) et al. (1991), Stéquert et al. (1996) et Panfili et al. (2002). Ils sont nettoyés à l'hypochlorite de sodium et rincés à l'eau distillé avant d'être inclus dans une résine et coupés transversalement de chaque côté du nucléus. Le morceau est alors fixé à une lame de verre pour être poncé jusqu'au nucléus. L'opération est effectuée de chaque côté jusqu'à obtenir une pièce de 100 microns d'épaisseur. La surface de la lame obtenue est décalcifiée avec de l'EDTA (acide diaminotétracarboxylique) pour améliorer le contraste. Les lames sont ensuite observées au grossissement 1000x afin de compter le nombre de stries de la sagitta.

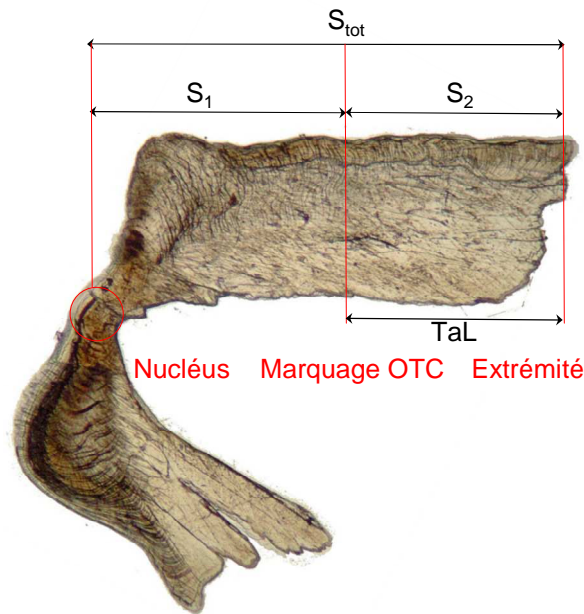


Figure 6 : Les différents comptages effectués lors des lectures d'otolithe de thon albacore (en rouge les zones sources d'incertitude)

Les données disponibles sont complexes (Figure 6). Les otolithes de l'échantillon ont été lus de diverses manières : (1) intégralement ; le nombre de stries est compté du nucléus à la partie terminale (S_{tot}). Les autres types de lectures, réservés aux otolithes marqués à l'OTC et donc issus du RTTP-IO, sont (2) le comptage du nombre de microstries entre le nucléus et la marque (S_1) et (3) le comptage du nombre de microstries entre la marque et l'extrémité de l'otolithe (S_2). Cette dernière lecture (S_2) peut être validée par comparaison au temps en liberté et permet d'appréhender les erreurs et les biais de lectures. Pour chaque otolithe, plusieurs lectures peuvent être faites (de 2 à 5 selon les otolithes), éventuellement par des lecteurs différents.

La méthode classique d'utilisation de ces données pour estimer l'âge est la suivante :

- Pour les poissons marqués à l'OTC, l'âge est calculé comme la somme du temps de liberté moyen (sur les différentes lectures) et du nombre moyen de stries comptées entre nucléus et le marquage à l'OTC (S_1).
- Pour les poissons qui n'ont pas été marqués à l'OTC ou pour lesquels le comptage S_1 n'a pas été effectué, l'âge est associé au nombre moyen de stries comptées du nucléus à la partie terminale de l'otolithe. Ce choix se justifie par le fait que la connaissance du temps de liberté est plus précise que l'information obtenue par lecture de l'otolithe.

Cette méthode fournit des estimateurs ponctuels de l'âge du poisson sans prendre en compte l'incertitude dans la lecture de l'âge.

3.2.3. Sources d'erreurs de mesure

Les sources d'incertitude autour des lectures d'otolithe sont multiples.

- La première source d'incertitude vient de la variabilité des lectures. Généralement, les otolithes sont lus plusieurs fois par différentes personnes ce qui permet d'appréhender la variabilité des lectures. Le comptage total a été répété de 2 à 5 fois. La répétition des lectures dépend de plusieurs facteurs allant de la qualité de la préparation de l'otolithe à la variabilité des lectures. Si les premières lectures ne concordent pas alors que l'otolithe est de bonne qualité, des lectures supplémentaires sont généralement effectuées. Le comptage du nombre de microstries entre le nucléus et la marque OTC et du nombre de microstries entre la marque et l'extrémité de l'otolithe sont répétés respectivement de 1 à 3 fois et de 1 à 4 fois.
- De plus, certaines zones de l'otolithe sont plus difficiles que d'autres à lire : la base de l'otolithe appelée nucléus est opaque et n'est pas aisément lisible, cette zone est estimée à 17 à plus ou moins 2 stries près (E.Morize, pers. Com.). De même, la partie terminale de l'otolithe est bien souvent détruite du fait de la préparation. Elle est rognée, la tendance est donc à la sous-estimation de cette partie.
- Enfin, il existe un décalage entre l'injection d'OTC et le marquage effectif de l'otolithe, la marque à l'OTC est donc un indicateur biaisé de la date de marquage. Cette marque s'étale sur plusieurs stries, la strie marquée est donc connue avec imprécision.

Du fait de la complexité des données et de la multiplicité des sources d'incertitudes, un modèle d'erreur de mesure sur les lectures d'otolithes a été développé.

3.2.4. Modèle d'erreurs de mesure liées aux lectures de microstries d'otolithes

La démarche a pour objectif de valoriser les lectures répétées des mêmes otolithes et l'expertise sur les autres sources d'incertitudes afin de quantifier les erreurs de lecture et l'incertitude sur l'estimation de l'âge qui en est issue. Ce modèle, développé dans le cadre Bayésien, peut alors fournir une structure de prior informatif pour les âges des poissons de la base de données de sclérochronologie, et peut être couplé au modèle développé au 2.3.1 Cela permet d'analyser l'influence de la prise en compte des erreurs de mesures sur l'âge sur l'estimation des paramètres de croissance. Le modèle d'erreur de mesure est construit dans une logique de Modèle Bayésien Hiérarchiques (Clark, 2005 ; Cressie et al., 2009) qui permet d'enchaîner des processus conditionnels aléatoires successifs.

La première composante de ce modèle valorise les données de marquage à l'OTC pour relier le temps de liberté TaL au nombre de stries lues entre le marquage OTC et l'extrémité de l'otolithe (S2). Ce modèle permettra de valider l'idée a priori qu'une strie est déposée chaque jour. On propose un modèle de régression linéaire entre le TaL et S2 :

$$(10.1) \quad TaL[i] = B \times S2_vrai[i] + C$$

où S2_vrai[i] n'est pas une mesure mais une variable latente intermédiaire qui sera reliée aux différentes mesures S2[i,j] dans une deuxième composante du modèle. Le paramètre C représente le décalage dû au temps d'action du marquage à l'OTC sur l'otolithe et B le rapport entre le temps en liberté et le nombre de stries une fois le décalage C pris en compte. Idéalement, l'estimation de B devrait être proche de 1 avec peu d'incertitude. Les priors utilisés sur les paramètres B et C suivent des lois Beta dilatées:

$$B' \sim dbeta(10,10) \quad \text{et} \quad B = 2 \times B'$$

$$C' \sim dbeta(1,1) \quad \text{et} \quad C = 5 \times C'$$

TaL[i] est une variable aléatoire distribuée selon une distribution uniforme dont les bornes sont déduites de l'expertise et de l'information dont on dispose sur la pêcherie. En effet, la date de recapture exacte du poisson (qui détermine le TaL) est souvent incertaine car la date fournie par les armateurs est associée à une cuve dans les cales du bateau qui correspond elle-même à plusieurs coups de seine s'étalant sur plusieurs jours.

$$(10.2) \quad TaL[i] \sim Unif(TaL[i]_{min}, TaL[i]_{max})$$

où TaL[i]_min, TaL[i]_max sont déduits de la base de données

l'équation (10.1) est couplée avec les données de lecture multiples des otolithes afin de faire le lien entre la variable latente S2_vrai[i] et les différentes lectures S2obs[i,j]. Ce lien comporte deux sources de variabilité bien distinctes : la variabilité entre les mesures S2obs[i,j] et l'incertitude au sein de chaque mesure S2[i,j].

L'incertitude dans chaque mesure est obtenue en combinant les sources d'incertitude dues au temps d'intégration de la marque OTC (ψ_2) et à l'extrémité rognée de l'otolithe (ψ_3) :

$$(10.3) \quad S2[i,j] = S2obs[i,j] + \psi_2 + \psi_3$$

$$\psi_2 \sim Normale(0, V = 4)$$

$$\psi_3 \sim Normale(0, V = 5)$$

La variabilité entre les mesures est représentée par une distribution Gamma :

$$(10.4) \quad S2[i,j] \sim Gamma(E=S2_vrai[i,j], CV)$$

Cette première composante du modèle a été traitée indépendamment à partir de données sélectionnées pour leur fiabilité, les individus concernés étant ceux pour lesquels le temps de liberté est connu assez précisément et pour lesquels au moins deux lectures de microstries ont été réalisées. Les sorties d'intérêt de cette première composante sont essentiellement l'estimation des paramètres B et C, permettant la validation de l'hypothèse que les stries se déposent bien de manière journalière.

La deuxième composante du modèle d'erreur de mesure vise à estimer l'incertitude autour de l'âge de chaque poisson. Elle combine les résultats de la première composante (les distributions a posteriori de B et C sont utilisées comme priors informatifs dans cette deuxième étape) avec les différentes lectures S1, S2 et Stot et/ou les données de TaL. Les détails du modèle ne sont pas décrits et seules les idées générales sont présentées. Pour tous

les poissons, l'âge est toujours estimé en s'appuyant sur la base d'un comptage de nombre de stries, qui seront converties en un âge en reprenant l'équation (10.1) et les distributions informatives des paramètres B et C obtenues dans la première étape. Pour certains poissons, cette conversion porte sur Stot, sur la somme S1 + S2, ou uniquement sur S1 lorsque le TaL est jugé plus précis que S2. Dans tous les cas l'incertitude liée aux flous dans les différentes zones de comptage (nucléus, marque, partie rognée) et la variabilité dans les différentes lectures sont prises en compte par une structure hiérarchique analogue à celle décrite dans les équations (10.3) et (10.4).

Les sorties d'intérêt de cette deuxième composante sont essentiellement les distributions a posteriori de l'âge de chaque poisson.

4. Méthodes de validation par simulation-estimation

4.1. Démarche générale et modèle de simulation

Les premiers essais réalisés sur les données ainsi que des éléments rapportés dans la bibliographie ont montré que l'ajustement du modèle de croissance VB log K pouvait poser des problèmes non triviaux comme la difficulté à identifier tous les paramètres se manifestant par exemple par une grande sensibilité à la taille du jeu de données ou aux priors. Ainsi une démarche de validation par simulation-estimation a été mise en œuvre afin de mieux comprendre l'origine des difficultés rencontrées et de valider la méthode de simulation.

La démarche consiste à simuler des données de croissance individuelle, soit de nature « couple-taille âge », soit de nature marquage-recapture, à partir d'une courbe de VB paramétrée de façon réaliste, puis à réaliser l'estimation à partir des modèles présentés en 2.3.12.3 (taille âge) ou en 2.3.3 (taille-age + marquage-recapture). Afin de tester la performance de la méthode d'estimation, les distributions *a posteriori* des paramètres sont comparées aux vraies valeurs des paramètres en croisant plusieurs critères. Le critère graphique des profils des densités *a posteriori* permet de comparer les distributions *a posteriori* entre elles et à la valeur simulée. L'étude des Ecarts Quadratiques Moyens (EQM) (Equation 11) et des Ecarts Relatifs Moyens (ERM) (Equation 12) permet une approche quantitative de ces différences d'estimation. L'EQM propose une mesure combinée du biais et de la précision. Un EQM faible signifie un ajustement peu biaisé ou peu dispersé. L'ERM quant à lui renseigne sur le biais de l'ajustement des paramètres, un ERM positif traduit une surestimation du paramètre et inversement.

Si la distribution *a posteriori* d'un paramètre est estimée par un échantillon MCMC $i = 1, \dots, G$ de taille G, l'EQM et l'ERM sont estimés par :

$$(11) \quad EQM = \frac{1}{G} \sum_{i=1}^G (\theta_i - \theta_{true})^2 \quad (12) \quad ERM = \frac{1}{G} \sum_{i=1}^G \left(\frac{\theta_i - \theta_{true}}{\theta_{true}} \right)$$

Les différents scénarios testés concernent soit les hypothèses servant à simuler les données soit les hypothèses utilisées pendant la phase d'ajustement Bayésien (i.e. les priors).

4.2. Scénario de référence

Le jeu de données simulé dans le scénario de référence est constitué de 100 couples de données de marquage-recapture et de 100 données de lectures de microstries d'otolithe ayant 3 ans d'âge maximum au marquage et jusqu'à 3 ans de plus à la recapture. Ces données ont

été simulées à partir d'une courbe de croissance suivant des paramètres réalistes de la croissance du thon (Tableau 2). L'aléa dans la croissance individuelle autour du modèle moyen est simulé en tirant L_{inf} dans une loi Gamma. Ce jeu de données propose un bon compromis entre le réalisme de la taille du jeu de données et un temps de calcul raisonnable pour l'ajustement (simulation MCMC). Dans la phase d'ajustement, des priors relativement non informatifs sont utilisés (Tableau 2).

Pour tous les autres scénarios ci-dessous, seules les caractéristiques détaillées ci-dessous spécifiques à chaque scénario sont modifiées par rapport au scénario de référence.

Tableau 2 : Distributions *a priori* de référence utilisées lors de l'étude de simulation-estimation

Paramètres	Valeurs simulées	Loi du prior	Paramètres du prior
L_{inf}	Mode = 141 CV = 0.1	Gamma	Mode = 141 CV = 0.4
α	2.4	Gamma	Mode = 2.5 CV = 0.5
β	8	Gamma	Mode = 8 CV = 0.4
k1	0.29	Gamma	Mode = 0.5 CV = 0.4
k2	1.28	Gamma	Mode = 1 CV = 0.4
t_0	-0.4	Uniforme (t_{min}, t_{max})	$t_{min} = -2$ $t_{max} = 0$

Chacun des scénarios ci-dessous est conduit dans le cadre de référence, i.e. les données ont été simulées avec les paramètres de référence et ajustées avec les priors de référence. Seules les caractéristiques détaillées dans chaque scénario sont modifiées.

4.3. Influence de la qualité et de la quantité des données simulées

4.3.1. Influence de la qualité des données simulées

i. Sensibilité à l'âge maximum atteint par les poissons

La qualité du jeu de données est hautement dépendante de l'âge maximum. En effet, un manque d'information sur les âges les plus grands peut entraîner une estimation biaisée de certains paramètres de la courbe de croissance et plus spécifiquement de paramètres tels que L_{inf} et k2. Pour répondre à cela, la démarche choisie est la suivante : 4 jeux de données de 100 marquage-recapture et 100 couples âge-taille sont simulés avec des âges maximaux à la capture de 7 ans, 5 ans, 3 ans et 1 an, sachant que les recaptures sont étalées sur 3 ans après le marquage. Ils correspondent respectivement à une information très importante sur les individus de grande taille, une information moyenne sur les individus de grande taille, peu d'information sur ces individus et aucune information sur ces individus.

ii. Sensibilité au niveau de bruit dans les données

- Bruit sur l'âge

Afin d'étudier l'impact d'une erreur de mesure sur l'âge des poissons, plusieurs jeux de données ont été simulés en bruitant l'âge par un bruit de distribution Normale centrée en 0 mais avec différentes valeurs de coefficient de variation (CV). Ce bruit retranscrit le bruit généré par la prise en compte des erreurs de comptage lors de l'analyse des microstries

d'otolithes, la paramétrisation en CV permet de retranscrire une variabilité qui augmente avec l'âge du poisson comme observée avec le jeu de données empirique. Les coefficients de variation testés sont les suivants : 0.05, 0.1, 0.2 et 0.3.

4.3.2. Influence de la quantité de données

Le but de cette démarche est de mettre en évidence l'intérêt de disposer de nombreux individus et de tester dans quelle mesure la méthode Bayésienne était capable de gérer le traitement conjoint d'un jeu de données de lectures d'âge absolu avec un jeu de données beaucoup plus conséquent de marquage-recapture. Plusieurs jeux de données ont été testés. Chacun de ces jeux de données comprenait 100 données de lectures de microstries d'otolithes simulées complété par un nombre variable de données de marquage-recapture : 100, 200, 500, 1000, 3000.

4.4. Sensibilité au prior lors de la phase d'estimation Bayésienne

4.4.1. Prior sur L_{inf}

Les inférences peuvent être sensibles au choix de la distribution de probabilité *a priori* sur L_{inf} en faisant varier le mode et le CV de la loi Gamma correspondante. L'effet de la variation du mode et du CV sont étudiés séparément, cela permet d'estimer respectivement, l'impact d'un prior faussement informatif et d'un prior peu voire non informatif. Les priors testés sont donc un mode de 110, 130, 150 et 170 et un coefficient de variation de 0.2, 0.4, 0.6 et 0.8.

4.4.2. Prior sur les paramètres de transition α et β

Le paramètre α se révèle très important car il représente le point d'inflexion de cette transition qui symbolise l'âge moyen du passage de stade juvénile au stade adulte (Laslett et al., 2002; Eveson et al., 2004). Une étude de sensibilité est aussi menée concernant ces paramètres. Sont testés les priors suivants : coefficient de variation de 0.2 à 0.8 par pas de 0.2 et des modes faussement informatifs.

4.4.3. Prior sur les taux de croissance de la première stance (k_1) et de la deuxième stance (k_2)

Une étude de sensibilité est aussi menée concernant ces paramètres. Les différents priors testés sont à la fois des priors faussement informatifs (mode : 0.5, 1, 1.5) et des priors plus ou moins informatifs (CV de 0.2, 0.4, 0.6 et 0.8).

5. Ajustement Bayésien des modèle de croissance à partir des données de l'albacore de l'océan Indien

5.1. Le choix des priors

L'analyse a été menée en utilisant des priors relativement peu informatifs. Les priors sont centrés sur des valeurs définies à partir d'une étude bibliographique (Tableau 3). La dispersion exprimée via le coefficient de variation (CV) des priors est relativement importante, i.e. $CV \geq 0.4$ pour la majorité des paramètres.

Tableau 3 : Priors sur les paramètres de croissance de l'ajustement de référence

Variable	Loi	Paramètres	
L_{inf}	Gamma	Mode = 144	CV = 0.2
α	Gamma	Mode = 2.5	CV = 0.5
β	Gamma	Mode = 8	CV = 0.4
k_1	Gamma	Mode = 0.5	CV = 0.4
k_2	Gamma	Mode = 1	CV = 0.4
t_0	Uniforme (t_{min}, t_{max})	$t_{min} = -2$	$t_{max} = 0$

Le prior le plus informatif concerne le paramètre L_{inf} . Une loi Gamma centrée sur 144 avec un CV de 0.2 a été définie. En effet, la configuration du jeu de données, et notamment le peu de poissons dans la partie asymptotique de la zone de croissance laisse présager la nécessité d'introduire de l'information a priori sur ce paramètre pour obtenir un ajustement cohérent avec les connaissances de la biologie de l'espèce. Ce prior a été établi à partir des données de débarquement des senneurs européens de 1991 à 2009, le mode de ce prior est le quantile 99% de ces fréquences de tailles au débarquement.

Des priors peu informatifs sont définis pour k_1 et k_2 (Tableau 3). Une information fiable pour renseigner les priors concernant ces paramètres est très difficile à trouver étant donné qu'ils sont spécifiques à la courbe de croissance de VB log K. Fonteneau et Gascuel (2008) ont montré qu'il existe une augmentation importante du taux de croissance au stade pré-adulte autour de 65-70 cm. Cette augmentation correspond à la présence de la deuxième stance de la croissance du thon albacore.

Le prior choisi pour le paramètre α est peu informatif et centré sur une valeur crédible. Le paramètre α qui représente l'âge auquel le taux de croissance atteint son maximum, i.e le point d'inflexion, est estimable avec un biais à partir de la littérature. En effet, seul l'âge absolu de transition peut être estimé à partir des études précédentes. Ici α est un âge relatif dépendant de t_0 . Le paramètre t_0 dépend largement des données dont on dispose sur les jeunes spécimens. On lui affecte un prior non informatif.

Le paramètre β est spécifique au modèle, aucune information *a priori* sur β n'est disponible dans la littérature. β est supposé a priori distribué selon une loi peu informative (gamma) centrée sur la valeur seuil 8.

5.2. Le modèle de croissance sur les données de sclérochronologie

Dans un premier temps nous avons réalisé l'ajustement du modèle de VB log K sur le jeu de données issu des lectures de microstries d'otolithes. Le jeu de données considéré ici consiste en 354 couples tailles-âge (estimateurs ponctuels sans erreur de mesure, cf. 3.2.2) issus de l'ensemble des données du RTTP-IO et de la campagne du WSTTP. Les priors sur les paramètres sont définis dans le Tableau 3.

Une étude de sensibilité au prior sur L_{inf} est menée où sont comparés les ajustements en fonction du mode du prior sur L_{inf} (110, 130, 144, 150, 170). La sensibilité de l'ajustement au jeu de données (mélange des données RTTP-IO et du WSTTP) sera aussi étudiée.

5.3. Le modèle de croissance intégrant les données de marquage-recapture et les données de sclérochronologie.

Les résultats de la démarche de simulation-estimation (on anticipe ici sur les résultats annoncés au paragraphe III1 des résultats) ont montré que l'intégration des deux sources de données dans l'ajustement (sur la base du modèle décrit dans le DAG représenté en Figure 1) pose des problèmes méthodologiques importants que nous n'avons pas réussis à résoudre entièrement.

En effet, la procédure d'estimation (implémentée dans le logiciel OpenBUGS) semble être limitée dans le nombre de données de marquage-recapture qu'il est possible d'intégrer dans le modèle. Cette limite s'établit aux alentours de 1000. Au-delà de cette limite, l'échantillonnage MCMC ne donne pas de résultats exploitables i.e. il n'y a pas de convergence des algorithmes.

Face à ce problème technique non résolu, la démarche suivante a été mise en œuvre :

1) d'une part, les ajustements ont été réalisés à partir de deux jeux de données de 1000 marquage-recaptures tirés aléatoirement dans le jeu de données. Les résultats obtenus ont été comparés.

2) par ailleurs, les hypothèses a priori sur les paramètres ont été modifiées. Le paramètre β , très difficile à estimer, a été fixé à 8 (Eveson, 2004). Pour pallier à la grande corrélation entre les paramètres k_1 et k_2 , une reparamétrisation des priors a été mise en œuvre. Les nouveaux priors ont été définis ainsi:

$$(13) \quad k_1 \sim \text{Gamma}(\text{mode} = 2.5, CV = 0.5)$$
$$k_2 = k_1 + \varepsilon \quad \text{avec} \quad \varepsilon \sim \text{Uniforme}(0,2)$$

5.4. Le modèle de croissance adapté à des données de sclérochronologie couplé au modèle d'erreurs de mesures liées aux lectures de microstries d'otolithes.

Le modèle utilisé est le même que celui décrit en 2.3.1. Il est couplé au modèle d'erreur de mesure permettant de prendre en compte les erreurs de mesure sur l'âge décrit au paragraphe 3.2.4.

III Résultats

1. Evaluation de la performance de la méthode par simulation-estimation

L'approche de simulation-estimation a permis d'appréhender la faisabilité des inférences statistiques et le degré de confiance attribué aux sorties du modèle. Les résultats de cette approche sont synthétisés ci-dessous.

1.1. Corrélation statistique des paramètres

La Figure 7 synthétise les corrélations statistiques entre les paramètres. On observe une forte corrélation positive entre k_1 et k_2 . Ces deux paramètres sont négativement corrélés à la longueur asymptotique (L_{inf}). Cela implique qu'une croissance rapide entraîne une taille asymptotique faible. L'existence d'une telle corrélation est classique dans des modèles tels que le modèle de croissance de Von Bertalanffy (Pilling et al., 2002), il n'est donc pas étonnant de retrouver ce type de corrélation entre paramètres. Les paramètres k_2 et α sont positivement corrélés. Une corrélation négative est également observée entre les paramètres α et β .

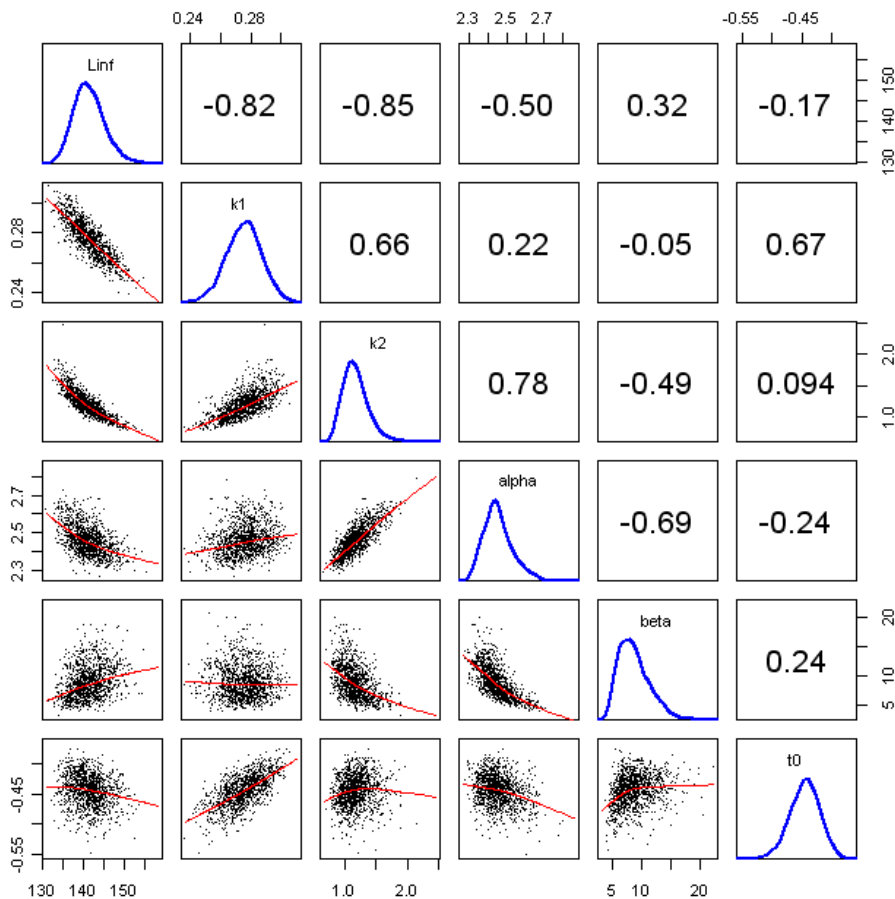


Figure 7 : Ajustement obtenu dans le cas du scénario de référence de la méthode de simulation-estimation. Distributions *a posteriori* jointes (nuages de tirages MCMC) et marginales (diagonale)

L'information contenue dans le jeu de données joue un rôle prépondérant sur les corrélations inter paramètres lors de l'ajustement. Toutes les corrélations, mise à part la corrélation entre

k_1 et t_0 , sont d'autant plus importantes que le jeu de données est « data poor ». Cela traduit des difficultés d'identifiabilité des paramètres.

1.2. Influence du niveau d'information contenu dans les données

1.2.1. Sensibilité à l'âge maximum des poissons

La qualité des ajustements dépend du niveau d'information disponible sur la taille asymptotique. Les différents jeux de données testés diffèrent par l'âge maximum à la capture des poissons (1 an, 3 ans, 5 ans et 7 ans au marquage et jusqu'à 3 ans de plus à la recapture). Les ajustements des paramètres L_{inf} , k_1 et k_2 y sont particulièrement sensibles.

On étudie dans un premier temps l'influence de ces différents jeux de données sur l'ajustement des paramètres caractérisant spécifiquement la seconde stance de croissance, L_{inf} et k_2 . L'écart quadratique moyen de L_{inf} est très important dans le cas où l'on ne dispose pas d'information sur les grands individus (Annexe 4), la valeur simulée n'est pas contenue par l'interquantile. Cet écart se réduit avec l'augmentation de l'âge maximum dans le jeu de données et stagne au-delà d'un âge maximum de marquage de 5 ans (Annexe 4). Le biais mesuré par l'ERM est largement positif pour le jeu de données ne disposant pas d'information sur les grands individus, le paramètre L_{inf} est surestimé de 18% en moyenne. Dans les autres cas, le biais est inférieur à 1%, l'estimation de L_{inf} est donc peu biaisée lorsque l'on dispose d'information sur les individus âgés. On remarque que le biais d'estimation est négatif pour les cas où l'on dispose de beaucoup d'information sur les grands individus (5 et 7 ans au marquage). Le modèle a donc tendance à surestimer le paramètre L_{inf} en cas de manque d'information sur les individus âgés et à le sous-estimer légèrement en cas de grande quantité d'information sur la taille asymptotique.

Le paramètre k_2 est le taux de croissance de la deuxième stance. Ce paramètre est estimé avec un biais négatif qui s'amenuise avec l'augmentation de la quantité d'information sur les grands individus. Le biais est largement négatif (ERM=28%) et la distribution de l'ajustement est plus dispersée dans le cas « data poor ». Le paramètre k_2 est estimé avec une précision comparable avec les jeux de données informatifs sur les individus âgés (5 et 7 ans).

On remarquera tout de même que le paramètre α est légèrement mais systématiquement sur estimé et que le biais de l'estimation de t_0 est invariablement négatif. Les autres paramètres ne montrent pas de tendances particulières dépendant de l'âge maximum de capture du jeu de données simulé.

1.2.2. Sensibilité au niveau de bruit dans l'âge des poissons

Les résultats mettent en évidence une importance prépondérante des erreurs de mesure sur l'âge sur l'estimation des paramètres. La dispersion et les biais des distributions *a posteriori* des paramètres augmentent avec l'amplitude du bruit des données (Annexe 5). Le paramètre L_{inf} est sous estimé, α et k_2 sont logiquement surestimés de part leur corrélation négative avec L_{inf} .

1.3. Influence de la quantité de données

La quantité de données influe fortement sur la performance de la méthode d'estimation. En effet, des difficultés techniques importantes ont été rencontrées pour mettre en œuvre l'échantillonnage MCMC (en utilisant OpenBugs) lorsque la taille du jeu de données de marquage-recapture augmente (>1000). Les ajustements conduits en intégrant simultanément 100 données de lectures d'otolithes (couples âge-taille) avec 100, 200 et 500 données de

marquage-recapture ne posent pas de problème et aboutissent à des ajustements quasi similaires. Le modèle avec 1000 marquage-recapture nécessite un nombre d'itérations plus beaucoup plus important pour converger. Lorsque le nombre de données de marquage-recapture excède 1000, les simulations MCMC convergent très difficilement et le temps de calcul devient trop important, sauf à utiliser des priors exagérément informatifs.

Notre expertise nous conduit à conclure que les difficultés d'ajustement sur l'ensemble du jeu de données sont de nature technique et inhérentes à l'utilisation d'OpenBUGS, et pas de nature théorique. Malheureusement, il n'a pas été possible de cerner l'origine exacte du problème pendant le stage.

1.4. Sensibilité au prior sur L_{inf}

Cette analyse de sensibilité menée à partir du scénario de référence met en évidence que le coefficient de variation de la distribution *a priori* de L_{inf} n'a pas d'impact sur l'ajustement des paramètres du modèle. Des priors relativement faussement informatifs ($CV > 0.2$) ne modifient pas ou peu l'ajustement des paramètres, mais influent sur la vitesse de convergence. Si le prior a un CV inférieur, l'estimation peut être largement modifiée. Le modèle est donc relativement robuste au prior sur L_{inf} .

1.5. Sensibilité au prior sur les paramètres de transition α et β

L'analyse de sensibilité a mis en évidence une faible influence de la précision du prior sur α . Une précision trop faible ou un prior trop éloigné de la valeur simulée empêchent la convergence des simulations MCMC en un temps raisonnable. Lorsque le paramètre β prend des valeurs importantes (> 8), son impact devient très faible sur la forme de la courbe. La conclusion de cette analyse est que le paramètre β est très difficile à estimer.

1.6. Sensibilité au prior sur k_1 et k_2

L'analyse de sensibilité révèle deux éléments : (1) Une faible sensibilité aux priors avec une estimation qui est peu modifiée par la précision ou la qualité faussement informative de l'information *a priori* malgré (2) l'existence de seuil qui entraîne une non-convergence des chaînes pour ($CV > 0.5$).

1.7. Conclusions de l'évaluation de la performance de la méthode

Cette étude de sensibilité a permis de mieux appréhender la performance de la méthode d'estimation. L'ajustement du modèle reste fragile malgré une relative faible sensibilité aux priors sur les paramètres. En effet, les biais des estimations sont en général relativement faibles, les seules estimations largement biaisées sont issues d'ajustements sur des jeux de données fortement bruités (cf. 1.2.2) ou très peu informatifs (cf. 1.2.1). Ainsi la prise en compte des erreurs de lecture du nombre de microstries d'un otolithe peut impacter grandement l'estimation des paramètres de croissance. Cependant les priors jouent un rôle prépondérant pour la convergence du modèle, des priors fortement faussement informatifs ou trop peu informatifs entraînent une non convergence du modèle, c'est particulièrement le cas des ajustements avec des jeux de données de taille très importante (cf. 1.3). Cette étude a aussi permis de mettre en évidence certaines sources de difficultés techniques. Le paramètre β , du fait de la structure de la courbe, se révèle difficile à estimer en raison de la faible influence de ce paramètre sur le profil de la courbe lorsqu'il est supérieur à 8 (cf. 1.5). De plus, on observe de fortes corrélations entre les paramètres, ces corrélations étant d'autant plus importantes que les données sont bruitées. C'est particulièrement le cas des paramètres L_{inf} , k_1 et k_2 qui sont fortement corrélés (cf. 1.1).

2. Application à l'albacore de l'Océan Indien

2.1. Modèle basé sur les données de lectures de microstries d'otolithes

2.1.1. Ajustement de référence à partir des couples âge-taille

L'ajustement met clairement en évidence une croissance en deux stances : une première stance jusqu'à l'âge de deux ans environ avec un taux de croissance relativement faible et quasi constant jusqu'à une taille de 70 cm, suivie d'une deuxième stance avec une forte accélération de croissance atteignant son maximum à un peu plus de deux ans (Figure 9). La croissance diminue alors jusqu'à atteindre la taille infinie ($L_{inf} = 131.2$ cm). Le modèle montre une croissance très rapide lors de la deuxième stance suivie d'une chute abrupte du taux de croissance une fois proche de la valeur du L_{inf} .

L'ajustement confirme les difficultés d'estimation liées aux corrélations fortes entre les estimations des paramètres du modèle qui sont bien du même ordre de grandeur de que celles observées lors de l'ajustement des valeurs simulées (Figure 8). Les données sont suffisamment informatives pour permettre une estimation des paramètres sans apport d'information par l'expertise. Cependant, le jeu de données contient peu d'individus âgés, il convient donc de compenser ce manque par de l'apport d'information issue de l'expertise.

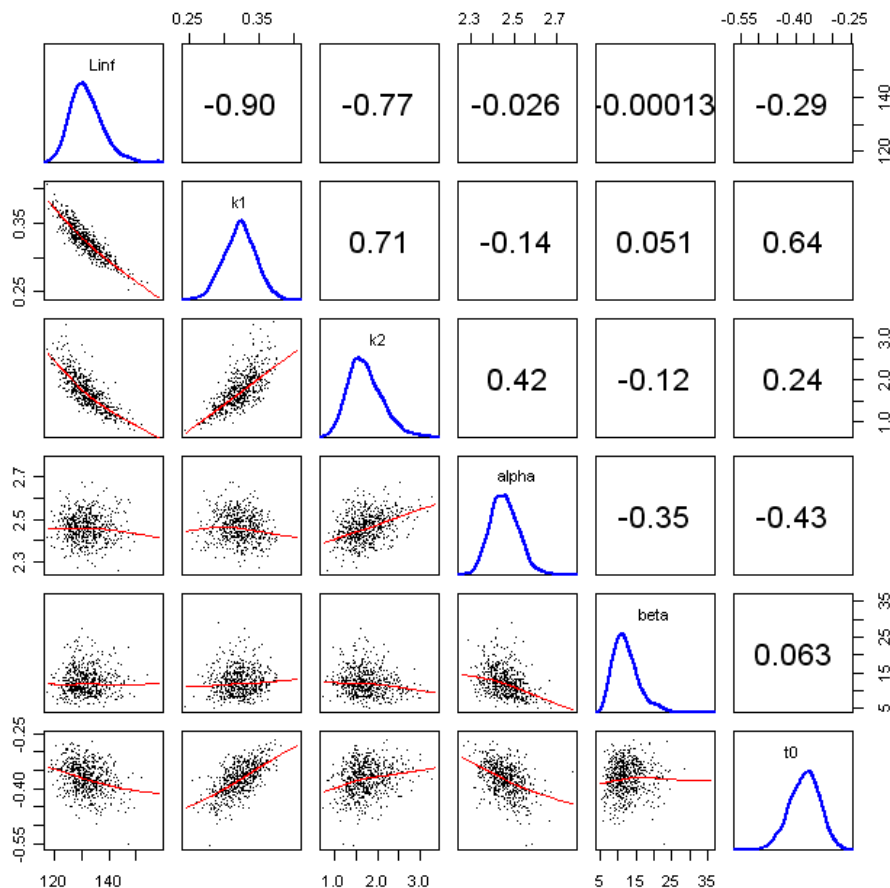


Figure 8 : Estimations des paramètres obtenues à partir du jeu de données de lecture de microstries d'otolithes et les priors de référence (Tableau 3). Distributions *a posteriori* jointes (nuages de tirages MCMC) et marginales (diagonale)

2.1.2. Sensibilité au prior sur L_{inf}

Les résultats montrent une sensibilité assez importante de l'ajustement au choix du prior sur L_{inf} . Lorsque le mode du prior est décalé vers les fortes valeurs (les valeurs 110, 130, 150, et 170 ont été testées), le plateau asymptotique se décale vers les fortes valeurs (Figure 9).

Du fait des corrélations statistiques fortes entre les estimations des paramètres, les estimations des autres paramètres sont aussi sensibles au choix du prior sur L_{inf} . L'estimation du paramètre L_{inf} est négativement corrélée à celles de k_1 et k_2 , un prior sur L_{inf} favorisant les fortes valeurs entraîne ainsi une estimation plus faible de k_1 et k_2 .

Mis à part la taille asymptotique, le reste de la courbe de croissance reste relativement inchangé lorsque le prior sur L_{inf} est modifié. Cependant les estimations de k_1 et k_2 sont bien modifiées, mais elles ont peu d'impact sur la forme de la courbe.

Cette analyse de sensibilité confirme donc que le jeu de données de lecture de microstries est relativement peu informatif sur la taille asymptotique des albacores de l'océan Indien. Un ajustement Bayésien d'une courbe de croissance sur ce jeu de données nécessite donc de porter la plus grande attention au choix du prior sur la taille asymptotique.

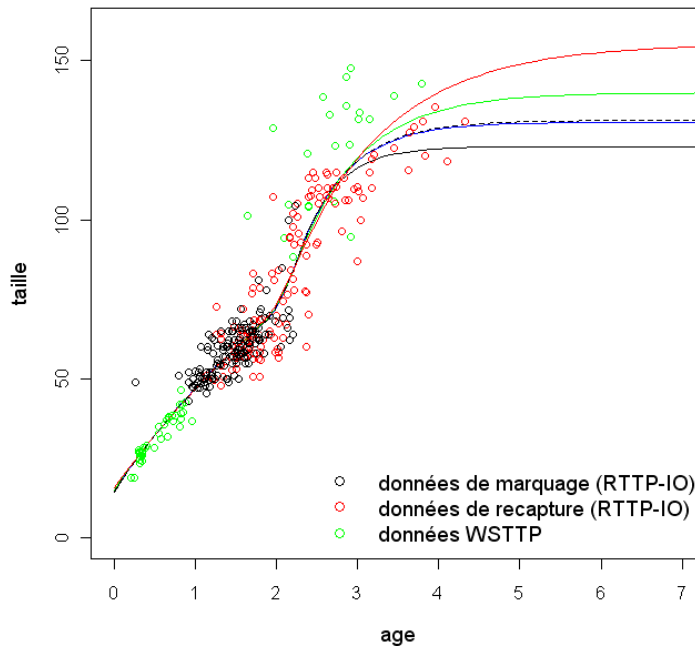


Figure 9 : Sensibilité de l'ajustement de la courbe de croissance au choix du prior sur le paramètre L_{inf} . (noir : prior à 110 cm ; bleu : prior à 130 cm ; pointillé : prior à 144 ; vert : prior à 150 cm ; rouge : prior à 170 cm). Les courbes sont tracées à partir des estimateurs ponctuels (médianes a posteriori) des paramètres.

2.1.3. Sensibilité aux données de poissons de petites tailles

Le jeu de données de lecture de microstries est issu de plusieurs programmes et l'information portée par les différents jeux de données est potentiellement différente. Afin d'appréhender la sensibilité de l'ajustement à l'origine du jeu de données, des ajustements de la courbe sur des

jeux de données séparés ont été réalisés et comparés aux résultats obtenus en mélangeant les jeux des deux programmes. (Tableau 4, Figure 10).

Les données issues de la campagne WSTTP (Figure 9, en vert) semblent différer des données du RTTP-IO. Les jeunes poissons semblent plus petits et les grands individus semblent plus jeunes relativement à leurs tailles. La croissance semble donc plus lente pour les jeunes stades et plus rapide pour les poissons adultes. Cependant, le jeu de données WSTTP est le seul à apporter de l'information sur les très jeunes poissons (< 1 an). L'influence de ce jeu de données est donc naturellement très forte sur l'ajustement de la première partie de la courbe.

L'ajustement obtenu avec les données de marquage-recapture seules (programme RTTP-IO ; Figure 10, rouge et noir) semble peu différent dans la deuxième partie de la courbe de celui obtenu précédemment avec l'ensemble des deux jeux de données. En revanche, comme mentionné précédemment, l'absence de mesure pour les jeunes poissons (les seuls points provenant du WSTTP) influe grandement sur la première partie de la courbe de croissance.

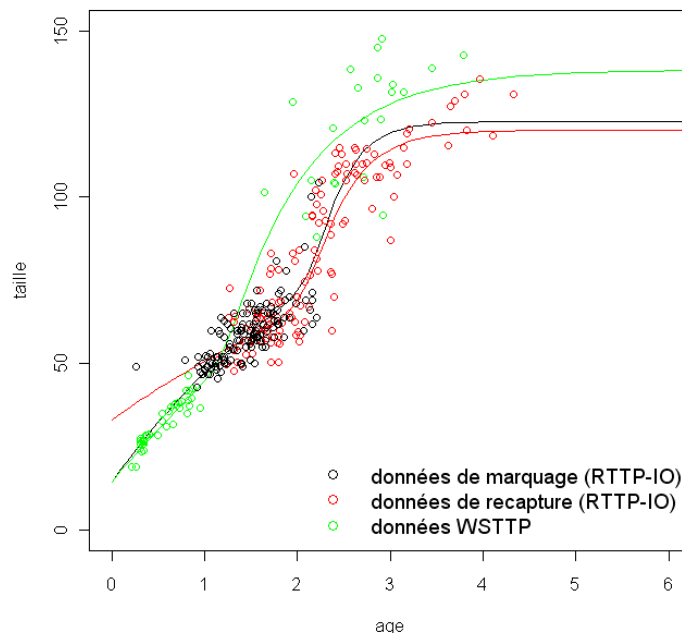


Figure 10 : Sensibilité de l'estimation de la courbe de croissance aux différents jeux de données d'otolithométrie (rouge : marquage-recapture, vert : ajustement sur données WSTTP, noir : les deux jeux de données). Les courbes sont tracées à partir des estimateurs ponctuels (médianes a posteriori) des paramètres.

Tableau 4 : Sensibilité de l'estimation de la courbe de croissance aux différents jeux de données (estimateurs ponctuels (médianes a posteriori) des paramètres)

Paramètres	Jeux de données ajustés		
	Marquage-recapture (en rouge)	Données complémentaires (en vert)	Les 2 jeux de données (en noir)
L_{inf}	120.1	138.2	122.7
α	3.56	1.6	2.55
β	8.47	11.04	9.98
k_1	0.23	0.27	0.36
k_2	2.58	1.20	3.33
t_0	-1.4	-0.4	-0.35

2.2. Ajustement intégrant les deux sources d'information: lectures de microstries d'otolithes et données de marquage-recapture

Comme montré lors de la démarche de simulation-estimation, le couplage de ces différentes sources d'information est ardu.

Face aux difficultés d'ajustement du paramètre β , il a été fixé à 8 (Eveson, 2004). De plus, les estimations des paramètres k_1 et k_2 , fortement corrélés, ont été reparamétrés afin de faciliter la convergence du modèle et l'estimation de ses paramètres. Les priors concernant les paramètres k_1 et k_2 sont les suivant :

$$k_1 \sim \text{gamma}(\text{mode} = 2.5, CV = 0.5)$$

$$k_2 = k_1 + \varepsilon \quad \text{avec} \quad \varepsilon \sim \text{uniforme}(0,2)$$

La démarche de simulation a montré qu'il existe une limite au nombre de marquage-recapture qu'OpenBUGS peut gérer et au delà de celle-ci le modèle ne converge pas. L'ajustement a donc été réalisé sur deux jeux de données différents obtenus par tirage aléatoire de 1000 données de marquage-recaptures dans le jeu de données total. Les résultats obtenus ont été comparés. Les estimations des paramètres obtenues sont quasi similaires. Ce résultat est intéressant car il semble indiquer que les estimations obtenues sont sans doute très proches de celles que l'on aurait obtenues en intégrant la totalité du jeu de données. Par ailleurs on n'observe pas de biais systématiquement positif ou négatif des résidus de l'ajustement. La structure générale correspond à une augmentation de la valeur absolue des résidus avec l'âge. Cependant leur structure met en évidence un goulot d'étranglement correspondant à une faible variabilité des estimations au niveau du pic de croissance de la deuxième stance (2,5 ans) (Annexe 6).

Les âges à la capture et à la recapture de chaque poisson du jeu de données de marquage-recapture sont estimés. Les estimations semblent cohérentes par rapport à la courbe de croissance obtenue par otolithométrie.

Les ajustements obtenus à partir des jeux de données de 1000 individus sont relativement proches de la courbe ajustée à partir des données de lectures de microstries d'otolithe. Les caractéristiques de la courbe de croissance obtenue avec 1000 données de marquage-recapture sont les suivantes : une taille à l'infini de 125.7 cm en moyenne. Les paramètres k_1 et k_2 valent en moyenne 0.28 et 2.18. Les paramètres α alpha et t_0 ont pour valeurs respectives 2.97 et - 0.78 (Tableau 5).

Tableau 5 : Sensibilité de l'estimation de la courbe de croissance aux différents jeux de données de 1000 marquage-recapture (estimateurs ponctuels (médianes a posteriori) des paramètres)

Paramètres	Jeu de données n°1	Jeu de données n°2
L_{inf}	125.2	126.3
α	2.95	2.98
B (fixé)	8	8
k_1	0.28	0.27
k_2	2.19	2.17
t_0	-0.76	-0.79

L'estimation du point d'inflexion de la deuxième stance est le même que celui estimé à partir des données de microstries d'otolithe (environ 2,2 ans). k_1 et k_2 sont respectivement surestimé et sous-estimé par rapport aux estimations précédentes. Le résultat marquant de cette comparaison est que le mélange des deux sources d'observation semble conduire à une sous-estimation de L_{inf} par rapport à celle obtenue en intégrant les données de microstries uniquement.

L'incertitude autour des estimations est plus faible avec les données de marquage-recapture ce qui montre l'intérêt d'intégrer les différentes sources de données au sein d'un même modèle (Annexe 7).

Ajustement à L_{inf} fixé

Afin d'intégrer l'intégralité des données de marquage-recapture et de fournir une courbe permettant la conversion des matrices de captures disponibles à l'échelle des pêcheries en fréquences de taille en matrices de fréquences d'âge qui sont utilisées comme données d'entrée dans les modèles d'évaluation utilisés au sein des groupes de travail de la CTOI, le modèle est ajusté avec le paramètre L_{inf} fixé à 144 cm, la valeur du mode du prior de référence. La courbe ajustée obtenue (Figure 11) met en évidence taux de croissance de la première stance plus faible que dans les analyses précédentes.

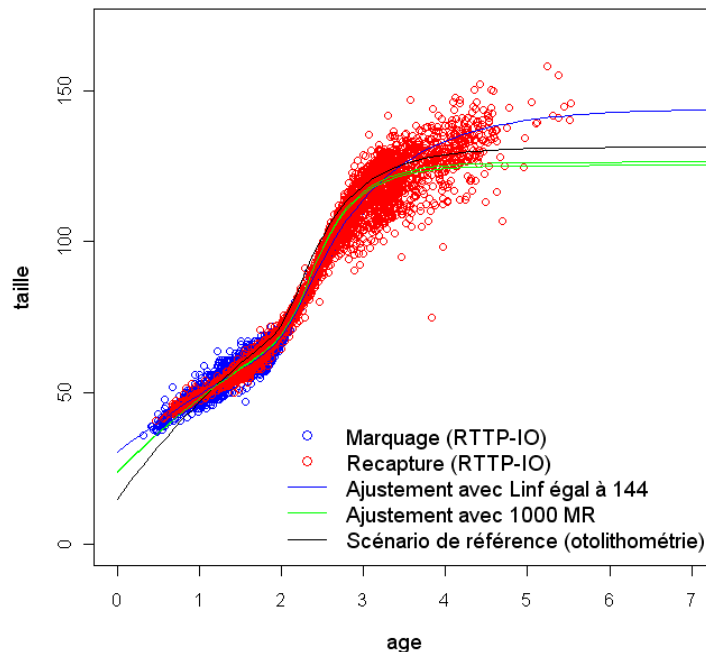


Figure 11 : Ajustement de la courbe de croissance obtenu à partir du mélange des deux sources de données : lectures de microstries et marquage-recapture. Le nuage de points correspond aux poissons du jeu de données de marquage-recapture dont la taille est connue mais dont l'âge à la capture et à la recapture a été re-estimé. Les courbes sont tracées à partir des estimateurs ponctuels (médianes a posteriori) des paramètres.

2.3. Modèle d'erreurs de mesure des lectures microstries d'otolithes et influence sur le modèle de croissance

2.3.1. Modèle d'erreur de mesure

Le modèle d'erreur de mesure en tant que tel est innovant et produit des résultats intéressants. L'ajustement de la relation linéaire (eq. 10.1) entre le nombre de stries comptées et le temps de liberté pour les poissons marqués à l'OTC (cf. matériels et méthodes II5.4) est un résultat fondamental (Figure 12) portant sur la validation de la lecture d'âge par otolithométrie. Le coefficient B est estimé à 1.08 avec coefficient de variation de 0.026 et le paramètre C est estimé à 0.614 avec un coefficient de variation de 0.99.

Ainsi le nombre de stries comptées semble légèrement inférieur au nombre de jours réels écoulés ($1/1.08 = 0.92$). L'estimation du paramètre C suggère une absence quasi-totale de délai (0.6 jours) entre l'injection d'OTC et le marquage effectif de l'otolithe.

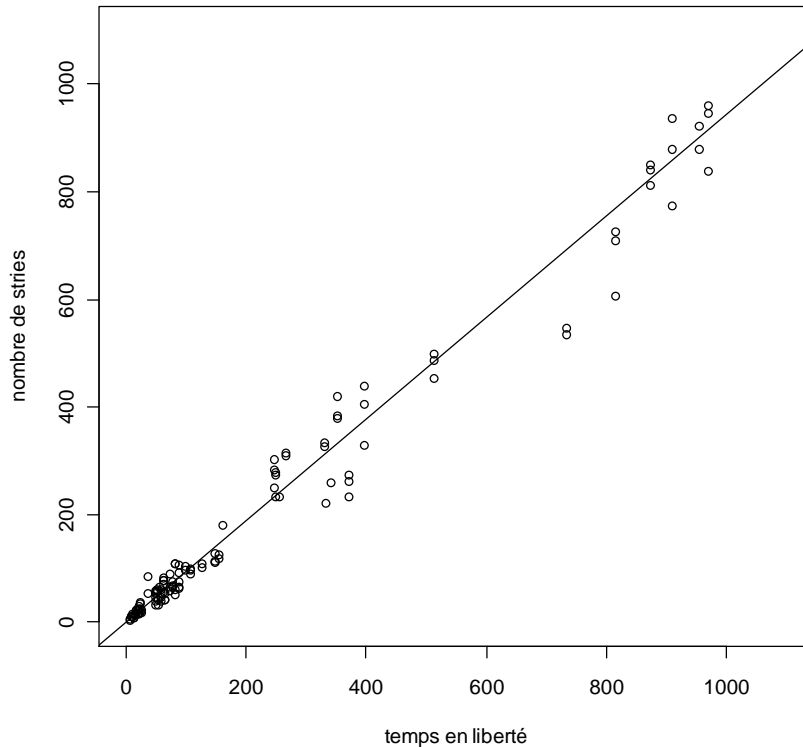


Figure 12 : Régression linéaire du nombre de stries comptées en fonction de du temps en liberté TaL. La régression est tracée en fonction des estimateurs ponctuels (médiane a posteriori) des paramètres B et C.

Le modèle estime en moyenne l'erreur de lecture à 8,7% du nombre de stries comptées. Du fait de la faible incertitude de l'estimation des paramètres B et C, cela se traduit par une incertitude légèrement supérieure autour de l'âge des poissons. Ce chiffre est relativement important et au vu des résultats de l'analyse de simulation-estimation, on peut s'attendre à un effet important de la prise en compte des erreurs de mesure sur l'ajustement de la courbe de croissance.

2.3.2. Couplage avec l'ajustement de la courbe de croissance

Au vu des difficultés rencontrés dans l'intégration des deux sources d'information (microstries et marquage-recapture), le couplage du modèle d'erreur de mesure avec le modèle de croissance a été réalisé sans intégrer les données de marquage-recapture (données de lecture de microstries uniquement).

L'ajustement obtenu à partir de ces données bruitées s'est révélé très fragile et non concluant en utilisant les priors peu informatifs du Tableau 3. En effet, si on observe la courbe de croissance ajustée, la croissance est très rapide entre 2 et 3 ans, le fait de bruiteur cette zone entraîne des estimations peu précises de l'intégralité des paramètres qui dépendent tous de cette partie de la courbe. Les résultats obtenus montrent que seuls les paramètres t_0 , k_1 et L_{inf} montrent des distributions a posteriori significativement actualisées par rapport aux priors. Les distributions *a priori* des autres paramètres ne sont presque pas mises à jour.

L'influence des erreurs de mesure sur la lecture d'âge des poissons est donc majeure sur notre capacité à estimer la courbe de croissance à partir des données de lecture de microstries. Les procédures classiques négligeant les erreurs de mesure sur la lecture d'âge surestiment grandement la qualité de l'information portée par les données et peuvent sans doute conduire

à des ajustements biaisés ou tout au moins sur-confiant en terme d'incertitude autour des estimations.

IV Discussion et conclusion

1. Synthèse

1.1. Avancées méthodologiques

A notre connaissance, ce travail constitue la première étude dédiée au développement d'une approche Bayésienne d'un modèle de croissance à partir de la courbe de Von Bertalanffy logK à partir de données de marquage-recapture et de lecture de microstries d'otolithes.

Le modèle mis au point est adapté à l'étude de la croissance du thon albacore de l'Océan Indien. Il est cependant applicable à tout jeu de données de marquage-recapture. En particulier, il est tout à fait envisageable de l'adapter aux autres espèces ciblées par le RTTP-IO, i.e. le patudo et le listao bien que les données de lectures d'otolithes pour le listao soient très difficiles ; ceci soulignant d'autant plus l'importance d'intégrer les erreurs de mesures liées à cette source d'information.

L'approche par simulation-estimation est une étape assez lourde mais primordiale de ce travail. Elle a permis de mieux appréhender les problèmes relatifs aux ajustements en termes de sensibilité aux hypothèses *a priori* et aux difficultés de convergence. Cette approche montre que dans le cas de données de marquage-recapture suivant une loi de croissance à deux stances, les paramètres de croissance sont identifiables malgré des limites de l'outil Openbugs.

Le principe de la méthode mise en œuvre réside dans la valorisation de la complémentarité des données de marquage-recapture et de lectures de microstries d'otolithes. En effet, la détermination de l'âge absolu n'est pas possible en utilisant des données de marquage-recapture seules, car cela nécessite d'avoir un référentiel sur l'axe des âges. Plusieurs méthodes peuvent être utilisées, comme par exemple une estimation de l'âge à la taille de recrutement. La seule source d'information concernant l'âge absolu utilisée dans notre étude est basée sur la sclérochronologie, e.g. la méthode des microstries journalières d'otolithes. Seules ces données sont utilisées pour estimer le t_0 l'âge auquel le poisson à une taille nulle et ainsi « ancrer » la courbe de croissance obtenue par les données de marquage-recapture en fonction des âges.

Le modèle d'erreur de mesure développé s'inscrit dans une démarche innovante permettant de prendre en compte les multiples sources d'incertitude dans la lecture des otolithes (i.e. la variabilité de la lecture et l'incertitude autour du nucléus, de la zone de marquage et de l'extrémité de l'otolithe). Ce modèle conduit à estimer des erreurs de mesure importantes, et leur influence sur notre capacité à estimer la courbe de croissance à partir des données de lecture de microstries s'est révélée majeure. Les procédures classiques négligeant les erreurs de mesure sur la lecture d'âge surestiment grandement la qualité de l'information portée par les données et peuvent sans doute conduire à des ajustements biaisés ou tout au moins sur-confiants en terme d'incertitude autour des estimations. Par ailleurs, le coefficient de 0.92 stries par jour semble indiquer qu'une standardisation serait nécessaire.

1.2. Difficultés techniques rencontrées

Les avancées méthodologiques réalisées ne doivent pas masquer les difficultés techniques rencontrées. Elles ont été nombreuses et ont conduit à mettre en œuvre une démarche d'essais-erreurs parfois fastidieuse.

L'analyse a révélé une limitation du nombre de données que l'on peut intégrer qui est liée au logiciel OpenBugs. En effet, la convergence est longue et très difficile dans des situations « data rich ». Il a été montré que des situations avec des jeux de données de plus de 1000 données entraînent ces difficultés. Cette limite peut être partiellement levée par des priors très informatifs limitant la phase exploratoire de l'ajustement.

Le modèle de croissance développé semble sous-estimer le L_{inf} par rapport aux L_{inf} estimés par les études précédentes. L'estimation du paramètre L_{inf} est un problème redondant (Bard, 1984; Eveson et Million, 2008) lors de l'ajustement de la courbe de croissance et encore plus spécifiquement dans le cas de données de marquage-recapture. La littérature suggère le rôle potentiel d'une forte variabilité inter annuelle aboutissant à des estimations de L_{inf} différents, dans notre cas il est plus vraisemblable d'imputer la faible valeur du L_{inf} estimée au peu de données dont on dispose sur les individus âgés et à la structure du modèle de croissance. La disponibilité des données de recapture pour les années 2009 et 2010 de gros albacores devrait favoriser considérablement l'ajustement du modèle et améliorer l'estimation de L_{inf} .

La structure des résidus des âges estimés par le modèle de croissance avec des données de marquage-recapture est problématique. La valeur absolue des résidus augmente avec l'âge des poissons ce que l'on peut expliquer par le fait que la variabilité individuelle de la croissance augmente avec l'âge. Cependant on observe un goulot d'étranglement des résidus autour de 0 au niveau de la deuxième stance. Or c'est précisément la période où le taux de croissance est le plus fort et où la variabilité inter individuelle devrait être la plus grande. La variabilité est sans doute transférée sur les autres parties de la courbe.

1.3. Connaissance sur la croissance des thons albacores

Malgré la fragilité des ajustements réalisés, l'étude apporte des éléments intéressants concernant la croissance des thons albacore de l'OI. La conclusion principale de cette étude est que le thon albacore de l'océan Indien a une croissance à deux stances.

Comme le suggèrent les données de lectures de microstries d'otolithes du programme WSTTP, les jeunes poissons (de 20 à 35 cm) ont un taux de croissance très important. Au-delà, le taux de croissance atteint un plateau et se stabilise voire diminue légèrement jusqu'à une taille de 65 cm en moyenne. Le taux de croissance augmente ensuite de façon importante jusqu'à un point d'inflexion à l'âge d'environ 2,2 ans pour une taille de 78 cm. Le taux de croissance diminue alors et tend vers 0 au bout d'un temps infini. La taille asymptotique estimée est de 125,7 cm.

2. Les limites

2.1. Les données

Les données de marquage du thon albacore issues de la campagne RTTP-IO concernent en grande majorité des individus de petites tailles comprises entre 40 et 70 cm. La taille de recrutement par la pêcherie étant d'environ 40 cm, ce jeu de données n'est donc pas informatif sur les individus avant leur entrée dans la pêcherie. Très peu des individus recapturés ont une taille supérieure à 140 cm (43 poissons), seulement 5 poissons dépassant 150 cm.

L'utilisation de données de marquage-recapture peut affecter l'estimation de la courbe de croissance par différents biais dus à la modification de la croissance et du comportement des poissons suite à la manipulation et au marquage. Dans le cas du thon albacore, largement pêché sous DCP, il n'est pas rare qu'un banc suive les bateaux, les poissons marqués ont donc potentiellement suivi le bateau scientifique lors de la campagne. La présence d'un bateau, e.g. le bateau scientifique, pourrait influencer sur les déplacements et le comportement alimentaire du banc donc sur son activité prédatrice et ainsi générer un biais sur la croissance individuelle des poissons. Cependant, compte tenu de l'expansion des zones de marquage et de recapture et du grand nombre de poissons recapturés tout au long de l'année, ce biais, s'il existe, est considéré négligeable.

Lors de la campagne, de nombreux poissons ont été marqués puis recapturés à quelques jours d'intervalle, à chaque fois les poissons ont été mesurés par l'équipe scientifique. Si on fait l'hypothèse que les poissons n'ont pas grandi entre chaque recapture, on peut estimer que les différences de mesure pour un même poisson sont dues à l'erreur de mesure et ainsi quantifier la variance de ces erreurs de mesure.

Il a été démontré qu'un biais systématique existe dans l'analyse de données de marquage-recapture due au rétrécissement des thons (shrinkage). En effet les thons sont toujours mesurés vivant lors du marquage alors qu'ils sont remesurés congelés lors de la recapture. Une diminution de 1.95% de la taille à la fourche des thons patudo a été observée (Schaefer et Fuller, 2006). Cette diminution est due au fait que les thons sont souvent compressés et tordus lors du stockage ce qui empêche une mesure aisée lors du débarquement. Ceci n'a pas été pris en compte lors de l'analyse mais il pourrait être intéressant de mener une analyse de sensibilité par rapport à ce critère.

Les données de lectures de microstries d'otolithes sont très bruitées. Le modèle d'erreurs de mesure des données d'otolithométrie mis au point montre une incertitude importante autour des estimations des âges et tend donc à montrer que les données de lectures d'otolithes sont des informations à utiliser avec précaution.

2.2. Le modèle de croissance

Le modèle de croissance VB log K a montré des limites liées au grand nombre de paramètres requis (6). La paramétrisation apporte une flexibilité à la forme de la courbe mais implique des difficultés d'estimation des paramètres. Ces difficultés d'estimation se manifestent par une forte corrélation statistique entre les paramètres (surtout entre L_{inf} , k_1 et k_2). La reparamétrisation du modèle peut être une solution mais s'avère très complexe. Des essais ont été conduits dans ce sens mais les méthodes de reparamétrisation développées par Schnute (1981) se sont révélées inapplicables à la courbe de croissance « VB log K ».

Les données de taille à la recapture suivent deux modes le premier à environ 55 cm et l'autre à environ 120 cm. C'est un cas particulier classique du thon albacore, capturé à la senne, que l'on retrouve aussi dans les autres océans (Marsac, 1991). L'explication avancée par la littérature est basée sur une vulnérabilité variable à la senne en fonction de la taille du poisson. Les poissons dont la taille est comprise entre 60 et 100 cm seraient moins vulnérables que les individus plus petits et plus grands. La courbe de croissance serait donc biaisée par un effet de sélectivité, les poissons intermédiaires relativement absents du jeu de données, la moindre vulnérabilité des poissons de taille de 60 cm et 110 cm pourrait aboutir à un ajustement d'une croissance en deux stances. Les deux plateaux étant des artéfacts de la variabilité interindividuelle de l'âge auquel les poissons atteignent ces deux tailles. Cette théorie considère donc la croissance individuelle du thon albacore comme une multitude de croissance individuelle suivant une loi de Von Bertalanffy. Cependant, cette théorie est invalidée par les données de lecture de microstries d'otolithes qui mettent bien en évidence une croissance suivant une courbe à deux stances marquées.

Ce faible nombre de poissons de taille intermédiaire dans le jeu de données peut s'expliquer partiellement par le fait que la taux de croissance est bien plus important entre 60 et 100 cm ainsi la population de thons est bien moindre dans cet intervalle de tailles que dans les intervalles inférieurs et supérieurs (mortalités non prises en compte). On a donc bien une croissance en deux stances avec une accélération de croissance entre 60 et 100 cm.

3. Perspectives

3.1. Régler les problèmes techniques

Les problèmes techniques sont de deux ordres. Ils dépendent en partie du logiciel utilisé, il serait intéressant d'adapter le modèle mis au point au logiciel JAGS afin de comparer ses performances d'estimation à celles du logiciel OpenBUGS. De plus les fortes corrélations confirment la nécessité de travailler sur la reparamétrisation du modèle afin d'améliorer l'identifiabilité des paramètres.

3.2. Amélioration du jeu de données

Les ajustements à partir des données de marquage-recapture ont révélé une relative non compatibilité des données de lecture d'otolithes du programme WSTTP. Ces données provenant de l'Est océan Indien, on peut se demander si la croissance des thons albacores diffère de celle de l'Ouest océan Indien. La croissance semble plus lente pour les premiers stades et plus rapides pour les poissons adultes, il serait bien entendu nécessaire de faire des recherches plus poussées sur le sujet pour répondre à cette question. Ce problème avait déjà été soulevé par Eveson et Million (2008). Il est donc important de renseigner ce jeu de données avec de l'information sur les jeunes et les vieux individus provenant de la zone du programme de marquage-recapture RTTP-IO. Les vieux individus seront sans doute disponibles à l'avenir suite aux recaptures des années 2009 et 2010 et des années futures. Une campagne visant les jeunes individus dans la zone ouest de l'océan Indien serait un plus indéniable pour l'étude des premières étapes de croissance. Cela permettrait de valider les données d'otolithométrie du WSTTP ainsi que de conclure sur une éventuelle croissance différente entre ces deux zones.

Une des perspectives de travail réside aussi dans l'intégration des fréquences de taille (Eveson et al., 2004) dans le modèle, ces données pouvant apporter de l'information sur la courbe de croissance. Les fréquences de taille des captures d'albacores ont en effet été l'une des

méthodes majeures d'estimation de la croissance de l'albacore dans le passé (Le Guen et Sakacawaj, 1973; Marcille et Stéquert, 1976).

Le sexage des grands thons albacores du RTTP-IO est en cours mais les données ne sont pas encore en nombre suffisant pour déterminer si oui ou non les mâles et les femelles suivent une courbe de croissance différente. Un dimorphisme sexuel dans la croissance des albacores a été suggéré par Wild (1986) avec des femelles qui grandiraient moins vite et/ou qui pourraient avoir une taille asymptotique plus petite que les mâles. Cette hypothèse pourra être testée dans le futur et pourrait avoir des conséquences importantes sur les avis scientifiques dérivés des évaluations de stocks. En considérant une croissance unisexe, les modèles de population actuellement appliqués aux thons tropicaux expliquent généralement les faibles biomasses de certaines classes d'âge par une augmentation de la mortalité naturelle alors qu'un ralentissement de croissance des femelles pourrait être à l'origine de tels changements.

3.3. Choix de modélisation

Si les problèmes techniques sont résolus, une étape de hiérarchisation peut être intégrée visant à appréhender la variabilité de croissance inter-individuelle. Cette méthode consiste à mettre des priors sur les paramètres de lois de distribution *a priori* de certains paramètres prépondérants comme la taille asymptotique L_{inf} ou les taux de croissance K .

Une approche mécaniste de la croissance du thon pourrait aussi être envisagée par le biais de la théorie Dynamic Energy Budget (DEB ; Kooijman, 2001) qui a pour but de décrire quantitativement, par des modèles mathématiques réalistes, les grandes fonctions biologiques des organismes (croissance, nutrition, digestion, reproduction, maintenance et vieillissement).

3.4. Application à la gestion

La courbe ajustée à partir des données du RTTP-IO, peut être utilisée afin de convertir les matrices de captures disponibles à l'échelle des pêcheries en fréquences de taille en matrices de fréquences d'âge qui sont utilisées comme données d'entrée dans les modèles d'évaluation de stocks.

V Bibliographie

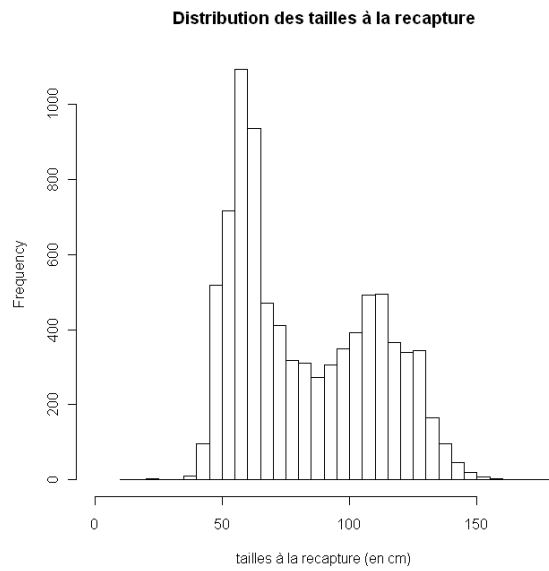
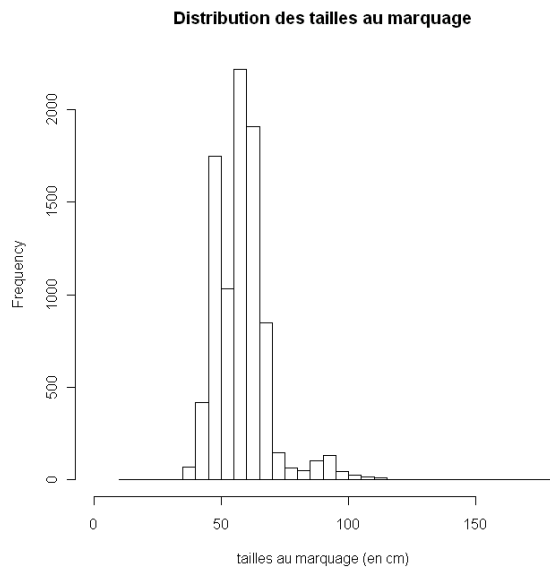
- Bard, F.X., 1984. Croissance de l'albacore (*Thunnus albacares*) Atlantique, d'après les données des marquages. Collect. Vol. Sci. Pap. ICCAT 20, 104–116.
- Brouard, F., Grandperrin, R., Cillaurren, E., 1984. Croissance des jeunes thons jaunes (*Thunnus albacares*) et des bonites (*Katsuwonus pelamis*) dans le Pacifique tropical occidental. ORSTOM de Port-Vila. Notes et Documents d'Océanographie 10, 23.
- Clark, J.S., 2005. Why environmental scientists are becoming Bayesians. *Ecology letters* 8, 2–14.
- Clark, J.S., 2003. Uncertainty and variability in demography and population growth: a hierarchical approach. *Ecology* 84, 1370-1381.
- Cressie, N., Calder, C.A., Clark, J.S., Hoef, J.M., Wikle, C.K., 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications* 19, 553–570.
- Efron, B., 1986. Why isn't everyone a Bayesian? *American Statistician* 40, 1–5.
- Eveson, J.P., Laslett, G.M., T, P., 2004. An integrated model for growth incorporating tag-recapture, length-frequency, and direct aging data. *Canadian Journal of Fisheries and Aquatic Sciences* 61, 292-306.
- Eveson, J.P., Million, J., 2008. Estimation of growth parameters for yellowfin, bigeye and skipjack tuna using tag-recapture data. Unpublished IOTC working paper IOTC-2008-WPTDA-07. Victoria, Mahe, Seychelles 33pp.
- Fonteneau, A., 1980. Croissance de l'albacore (*Thunnus albacares*) de l'Atlantique est. Collect. Vol. Sci. Pap. ICCAT 9, 152–168.
- Gascuel, D., Fonteneau, A., Capisano, C., 1992. Modélisation d'une croissance en deux stances chez l'albacore (*Thunnus albacares*) de l'Atlantique est. *Aquatic Living Resources* 5, 155–172.
- Fonteneau A., Gascuel D. (2008) Growth rates and apparent growth curves, for yellowfin, skipjack and bigeye tagged and recovered in the Indian Ocean during the IOTTP. Sci. Doc. IOTC-2008-WPTDA-08
- Gelman, A., 2004. Bayesian data analysis. CRC press.
- Gilks, W.R., Roberts, G.O., 1996. Strategies for improving MCMC. *Markov chain Monte Carlo in practice* 89–114.
- Gompertz, B., 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical transactions of the Royal Society of London* 513–583.
- Hallier, J., Million, J., 2009. The contribution of the Regional Tuna tagging Project. Indian Ocean to IOTC stock assessment IOTC-2009-WPTT-24.
- Hearn, W., Polachek, T., 2002. Estimating growth rate changes of southern bluefin tuna from tag-return data.
- Hilborn, R., Liermann, M., 1998. Standing on the shoulders of giants: learning from

- experience in fisheries. *Reviews in Fish Biology and Fisheries* 8, 273–283.
- Hillary, R., Million, J., 2008. Exploratory modelling of Indian Ocean tuna growth incorporating both mark-recapture data and otolith data. IOTC.
- Kooijman, S., 2001. Quantitative aspects of metabolic organization: a discussion of concepts. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 356, 331.
- Laslett, G.M., Eveson, J.P., Polacheck, T., 2002. A flexible maximum likelihood approach for fitting growth curves to tag-recapture data. *Can. J. Fish. Aquat. Sci.* 59, 976-986.
- Le Guen, J., Sakacawaj, G.T., 1973. Apparent growth of yellowfin tuna from the eastern Atlantic Ocean. *Fishery bulletin* 175.
- Lumineau, O., 2002. Study of the growth of yellowfin tuna (*Thunnus albacares*) in the Western Indian Ocean based on length frequency data. *IOTC Proc.* 5 316–327.
- Marcille, J., Stéquert, B., 1976. Croissance des jeunes albacores (*Thunnus albacares*) et patudos (*Thunnus obesus*) de la côte nord-ouest de Madagascar. *Cah. ORSTOM sér. Océanogr* 14, 153–162.
- Marsac, F., 1991. Growth of Indian Ocean yellowfin tuna estimated from size frequencies data collected on French purse seiners Coll. Vol. Work Doc., IPTP, TWS/91/17 :35-39.
- Marsac, F., Lablache, G., 1985. Preliminary study of the growth of yellowfin (*Thunnus albacares*) estimated from purse seine data in the western Indian Ocean.
- McAllister, M.K., Kirkwood, G.P., 1998. Bayesian stock assessment: a review and example application using the logistic model. *ICES Journal of Marine Science* 55, 1031.
- McCarthy, M.A., Masters, P., 2005. Profiting from prior information in Bayesian analyses of ecological data. *Ecology* 42, 1012–1019.
- Meyer, R., Millar, R.B., 1999. BUGS in Bayesian stock assessments. *Canadian Journal of Fisheries and Aquatic Sciences* 56, 1078–1087.
- Morize, E., Munaron, J.-., Hallier, J.-., Million, J., 2008. Preliminary growth studies of yellowfin and bigeye tuna (*THUNNUS ALBACARES* and *T. ONESUS*) in the Indian Ocean by otolith analysis. IOTC-2008-WPTT-30.
- Panfili, J., De Pontual, H., Troadec, H., Wright, P., 2002. *Manual of fish sclerochronology*, ifremer-Ird Editions. ed. Paris.
- Parent, É., Bernier, J., 2007. *Le raisonnement bayésien: modélisation et inférence*. Springer France Editions.
- Pilling, G.M., Kirkwood, G.P., Walker, S.G., 2002. An improved method for estimating individual growth variability in fish, and the correlation between von Bertalanffy growth parameters. *Canadian Journal of Fisheries and Aquatic Sciences* 59, 424–432.
- Polacheck, T., Eveson, J., Laslett, G., 2004. Increase in growth rates of southern bluefin tuna (*THUNNUS MACCOYII*) over four decades: 1960 to 2000. *Canadian Journal of Fisheries and Aquatic Sciences* 61, 307-322.
- Punt, A.E., Hilborn, R., 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. *Reviews in Fish Biology and Fisheries* 7, 35–63.
- Richards, F.J., 1959. A flexible growth function for empirical use. *J. exp. Bot* 10, 290–300.

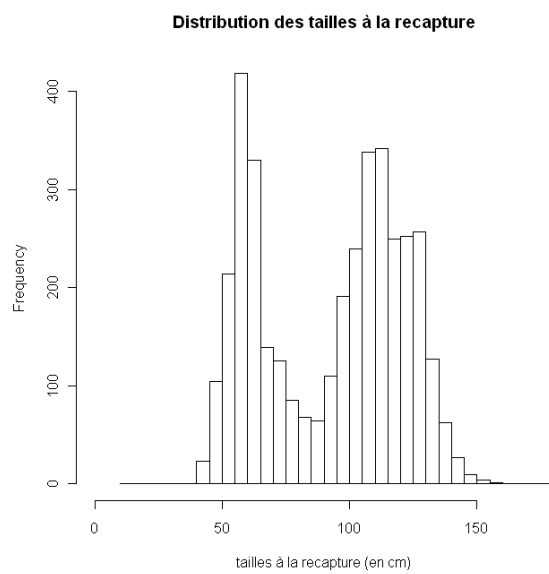
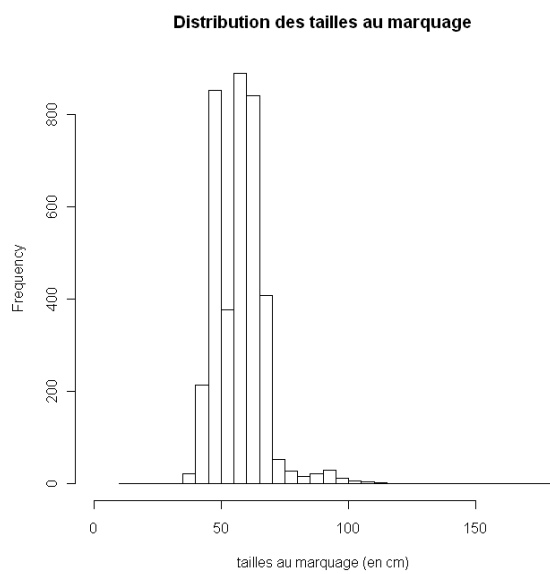
- Rivot, E., Prévost, E., Parent, E., Baglinière, J.L., 2004. A Bayesian state-space modelling framework for fitting a salmon stage-structured population dynamic model to multiple time series of field data. *Ecological Modelling* 179, 463–485.
- Romanov, E.V., Korotkova, L.P., 1988. Age and growth rates of the yellowfin tuna (*Thunnus albacares*)(Bonnaterre, 1788)(Pisces, Scombridae) in the northern part of the Indian Ocean, determined by counting the rings of vertebrae. *FAO/IPTP Coll. vol. of working doc 3*, 68–73.
- Schaefer, K.M., Fuller, D.W., 2006. Estimates of age and growth of bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean, based on otolith increments and tagging data. *Bulletin* 23.
- Schnute, J., 1981. A versatile growth model with statistically stable parameters. *Can. J. Fish. Aquat. Sci.* 38, 1128-1140.
- Secor, D.H., Dean, J.M., Laban, E.H., 1991. Manual for otolith removal and preparation for microstructural examination. Electric Power Research Institute.
- Steuert, B., Panfili, J., Dean, J.M., 1996. Age and growth of yellowfin tuna, *Thunnus albacares*, from the western Indian Ocean, based on otolith microstructure. *Fishery bulletin* 94, 124–134.
- Viera, A., 2005. Study of the growth of yellowfin tuna (*THUNNUS ALBACARES*) in the Indian Ocean based on length frequency data from 2000 to 2004. IOTC.
- Von Bertalanffy, L., 1938. A quantitative theory of organic growth. *Human biology* 10, 181–213.
- Wang, Y., 1998. Growth Curves with Explanatory Variables and Estimation of the Effect of Tagging. *Australian & New Zealand Journal of Statistics* 40, 299-304.
- Wild, A., Wexler, J.B., Foreman, T.J., 1995. Extended studies of increment deposition rates in otoliths of yellowfin and skipjack tunas. *Bulletin of Marine Science* 57, 555–562.

ANNEXES

a)

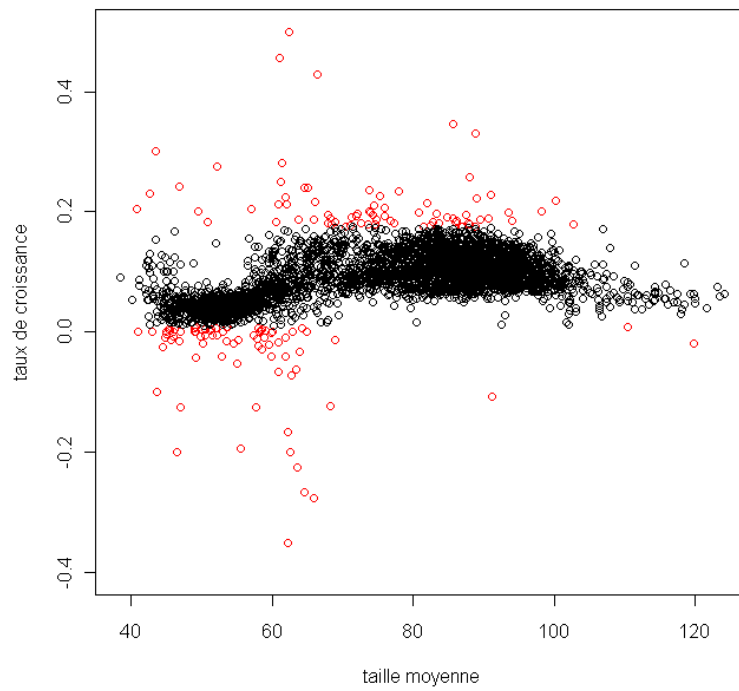


b)



Annexe 1 : a) Histogrammes des distributions des tailles des thons albacores du jeu de données original au marquage et à la recapture b) après sélection

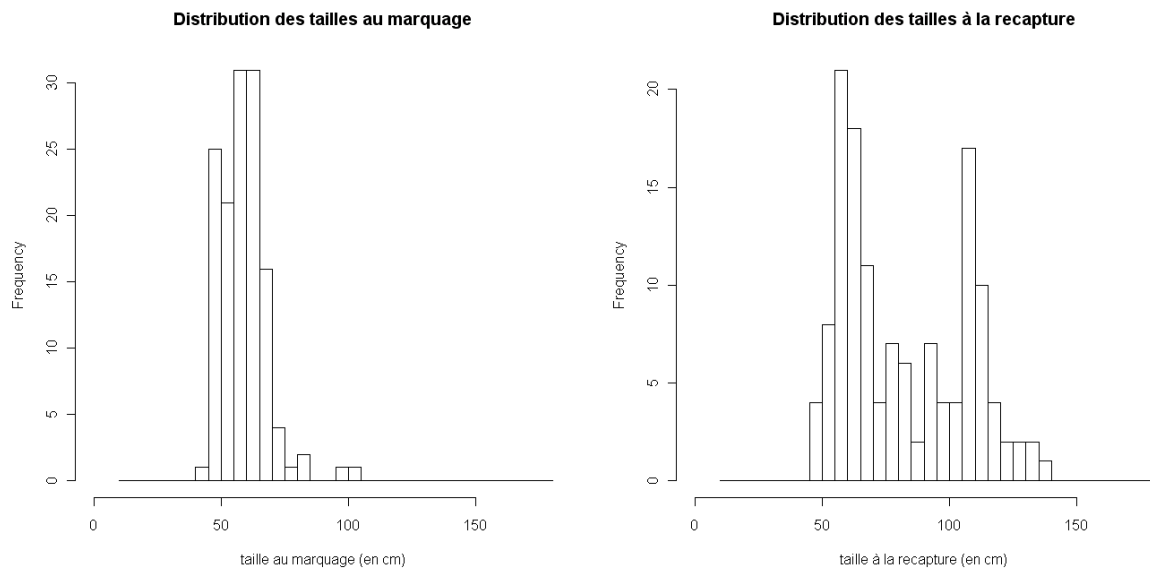
croissance en taille en fonction de la taille moyenne



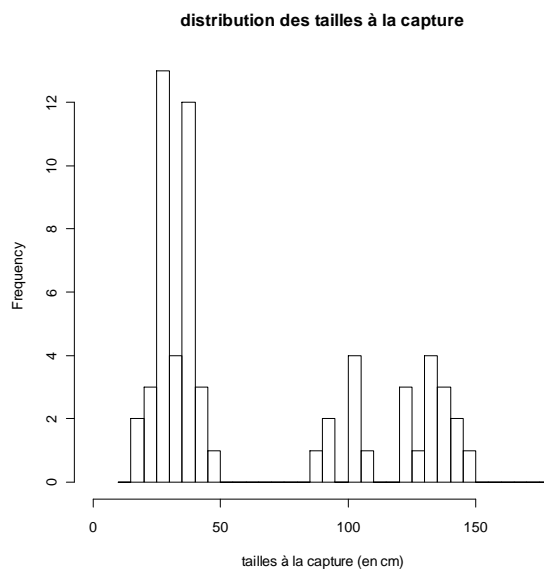
Annexe 2 : Graphique représentant la sélection sur le taux de croissance des thons albacores (en rouge : données exclues).

Pour un poisson marqué à une taille l_1 et recapturé à une taille l_2 après un délai $\Delta t_{1 \rightarrow 2}$, le taux de croissance est calculé comme $(l_2 - l_1) / \Delta t_{1 \rightarrow 2}$. Il est représenté en fonction de la taille moyenne $(l_1 + l_2) / 2$

a)



b)



Annexe 3 : a) Histogrammes des distributions de tailles des thons albacores du jeu de données d'otolithométrie au marquage et à la recapture b) Histogramme de la distribution de tailles à la capture des thons albacores issus de la campagne complémentaire.

EQM	L_{inf}	k_1	k_2	α	β	t_0
1 an	1135,2	0,0071	0,378	0,014	14,2	0,015
3 ans	13,7	0,0003	0,053	0,009	9,1	0,003
5 ans	2,2	0,0009	0,016	0,004	11,7	0,010
7 ans	2,5	0,0003	0,018	0,003	7,1	0,008

ERM (en %)

1 an	0,186	-0,27	-0,42	0,02	-0,14	0,30
3 ans	0,003	-0,05	-0,08	0,02	-0,12	0,11
5 ans	-0,004	-0,10	-0,06	0,01	0,13	0,23
7 ans	-0,008	-0,05	-0,06	0,00	0,03	0,20

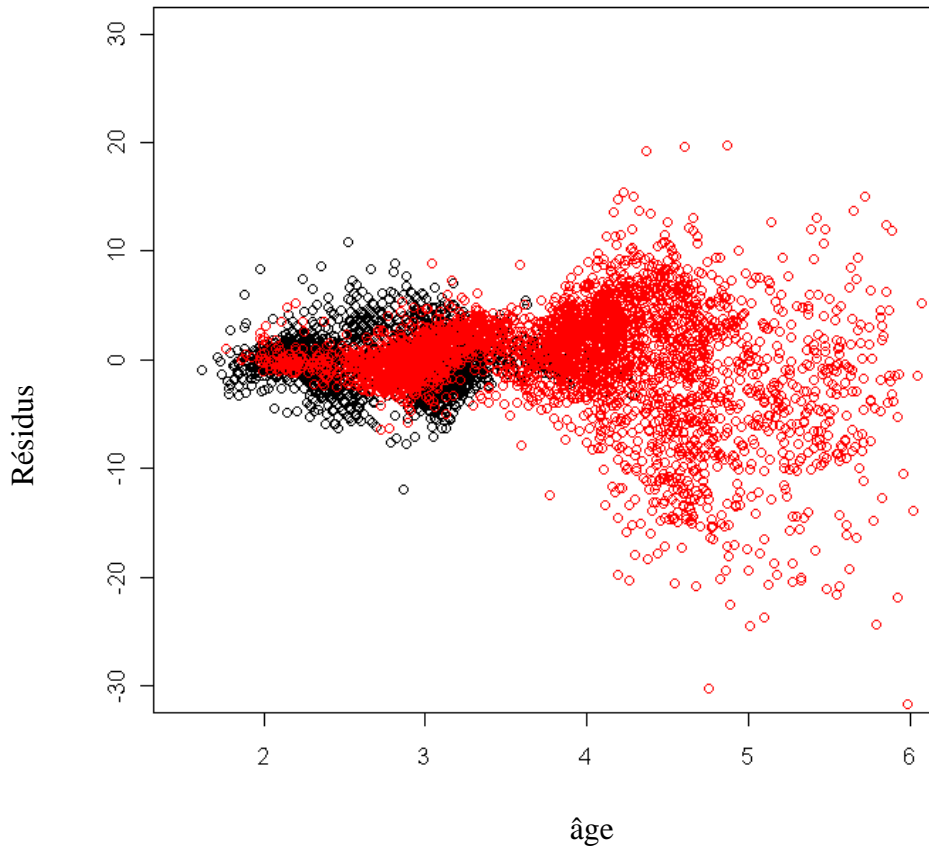
Annexe 4 : EQM et ERM des ajustements des paramètres de la fonction de croissance en fonction de l'âge maximum de marquage.

EQM	L_{inf}	k_1	k_2	α	β	t_0
CV0.05	33,8	0,0011	0,09	0,029	22,0	0,008
CV0.1	27,2	0,0007	0,19	0,092	26,0	0,007
CV0.2	104,9	0,0017	2,90	0,250	12,7	0,003
CV0.3	146,8	0,0014	2,96	0,542	34,7	0,044

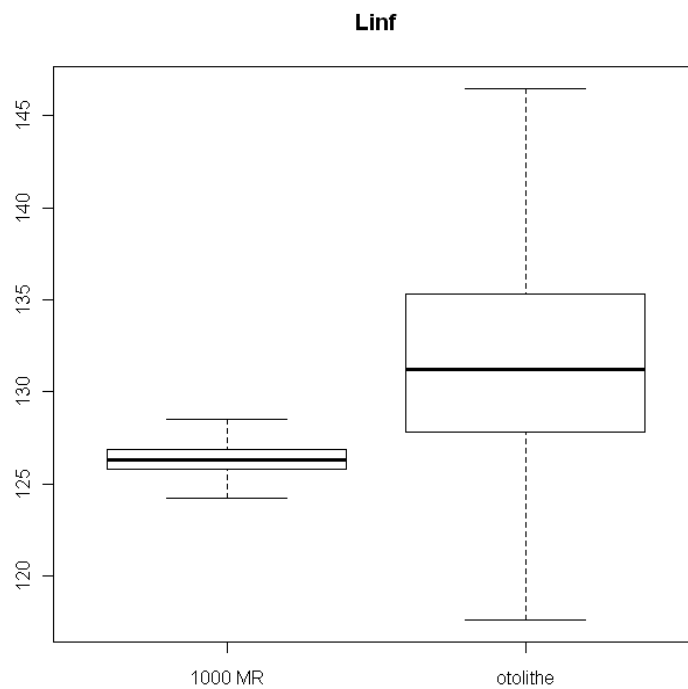
ERM (en %)

CV0.05	0,031	-0,11	-0,07	0,04	-0,43	0,20
CV0.1	-0,009	-0,04	-0,02	0,07	-0,43	0,14
CV0.2	-0,067	0,12	1,03	0,20	-0,10	0,06
CV0.3	-0,083	-0,04	1,02	0,28	-0,49	0,43

Annexe 5 : EQM et ERM des ajustements des paramètres de la fonction de croissance en fonction du coefficient de variation appliqué au bruit sur l'âge.



Annexe 6 : Résidus obtenus à l'issue de l'ajustement du modèle à l'intégralité des données de marquage-recapture (L_{inf} fixé)



Annexe 7 : Distributions *a posteriori* du paramètre L_{inf} dans le cas d'un modèle intégrant 1000 données de marquage-recapture et intégrant seulement les données d'otolithes.

Annexe 8: Code WinBUGS du modèle intégrant les données de marquage-recapture et les données de lectures de microstries d'otolithe.

Model

{

Priors

Linf

```
CVLinf<-0.05
MLinf<-144
Linf1<-1/pow(CVLinf,2)
Linf2<-(1/pow(CVLinf,2)-1)/MLinf
Linf~dgamma(Linf1,Linf2)
Linfp~dgamma(Linf1,Linf2)
```

Alpha

```
CValpha<-0.5
Malpha<-2.5
alpha1<-1/pow(CValpha,2)
alpha2<-(1/pow(CValpha,2)-1)/Malpha
alpha~dgamma(alpha1,alpha2)
alphap~dgamma(alpha1,alpha2)
```

beta

```
beta<-8
```

k1

```
CVk1<-0.4
Mk1<-0.5
ak1<-1/pow(CVk1,2)
bk1<-(1/pow(CVk1,2)-1)/Mk1
k1~dgamma(ak1,bk1)
k1p~dgamma(ak1,bk1)
```

k2

```
k2<-k1+epsilon
epsilon~dunif(0,2)
```

t0

```
t0~dunif(-2,0)
t0p~dunif(-2,0)
```

tau

```
tau~dgamma(0.01,0.01)
```

données de lectures de microstries d'otolithe (RTTP-IO)

```
for (i in 1:ntoto){
  Aoto2[i]<-Aot2[i]/365
  Aoto[i]<-Aot1[i]/365

  logELoto1[i]<-log(Linf*(1-exp(-k2*(Aoto[i]-t0))*pow(((1+exp(-beta*(Aoto[i]-t0-alpha)))/(1+exp(beta*alpha))),(-(k2-k1)/beta))))
  logLoto1[i] <- log(Loto1[i])
  logLoto1[i] ~ dnorm(logELoto1[i], tau)

  logELoto2[i]<-log(Linf*(1-exp(-k2*(Aoto2[i]-t0))*pow(((1+exp(-beta*(Aoto2[i]-t0-alpha)))/(1+exp(beta*alpha))),(-(k2-k1)/beta))))
  logLoto2[i] <- log(Loto2[i])
  logLoto2[i] ~ dnorm(logELoto2[i], tau)

}
```

Données de lectures de microstries d'otolithe (WSTTP)

```
for (i in 1:n_NM){
  Aotobis[i]<-Aoto_NM[i]/365

  logEL_NM[i]<-log(Linf*(1-exp(-k2*(Aotobis[i]-t0))*pow(((1+exp(-beta*(Aotobis[i]-t0-alpha)))/(1+exp(beta*alpha))),(-(k2-k1)/beta))))
  logL_NM[i] <- log(L_NM[i])
  logL_NM[i] ~ dnorm(logEL_NM[i], tau)

}
```

Données de marquage-recapture (WSTTP)

Priors sur les ages à la capture

```
CV<-0.8
M<-2
a1<-1/pow(CV,2)
a2<-(1/pow(CV,2)-1)/M

for (i in 1:n){
  A1[i]~dgamma(a1,a2)
  A2[i]<-A1[i]+(t2[i]-t1[i])/365
}
```

```
logEL1[i]<-log(Linf*(1-exp(-k2*(A1[i]))*pow(((1+exp(-beta*(A1[i]-alpha)))/(1+exp(beta*alpha))),(-(k2-  
k1)/beta))))
```

```
logL1[i] <- log(L1[i])
```

```
logL1[i] ~ dnorm(logEL1[i], tau)
```

```
logEL2[i]<-log(Linf*(1-exp(-k2*(A2[i]))*pow(((1+exp(-beta*(A2[i]-alpha)))/(1+exp(beta*alpha))),(-(k2-  
k1)/beta))))
```

```
logL2[i] <- log(L2[i])
```

```
logL2[i] ~ dnorm(logEL2[i], tau)
```

```
L11[i]~dnorm(L1[i],100000)
```

```
L12[i]~dnorm(L2[i],100000)
```

```
}
```

```
}
```