



Modélisation du zooplancton et du micronecton marins

Anna Conchon

► To cite this version:

Anna Conchon. Modélisation du zooplancton et du micronecton marins. Sciences agricoles. Université de La Rochelle, 2016. Français. NNT : 2016LAROS011 . tel-01661554

HAL Id: tel-01661554

<https://theses.hal.science/tel-01661554>

Submitted on 12 Dec 2017

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

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

Ecole doctorale 523
Sciences pour l'environnement – Gay Lussac

THÈSE

pour obtenir le grade de docteur délivré par

l'Université de La Rochelle

Ecologie

présentée et soutenue publiquement par

Anna Conchon

20 juin 2016

Modélisation du zooplancton et du micronecton marins
Modeling marine zooplankton and micronekton

Co-directeur de thèse : **Patrick Lehodey**

Co-directeur de thèse : **Christophe Guinet**

Jury

Paco Bustamante,	Président du jury
Marion Gehlen,	Rapporteur
Arnaud Bertrand,	Rapporteur
Vincent Ridoux,	Examineur
Pierre Brasseur,	Examineur

CLS
8-10 rue Hermès
31520 Ramonville
France

Résumé

Le zooplancton et le micronecton sont les deux premiers échelons animaux de la chaîne trophique marine. Bien que de tailles très différentes ($200\mu\text{m}$ à 2mm pour le zooplancton, 2 à 20cm pour le micronecton), ces deux groupes d'espèces variées partagent un comportement singulier : les migrations nyctémérales. Ces migrations journalières entre la profondeur de jour et la surface de nuit induisent des flux de matière organique très importants entre les différentes profondeurs de l'océan. L'étude des cycles biogéochimiques océaniques a une grande importance pour l'étude du changement climatique. Cette étude est notamment conduite à travers le développement de modèles globaux de circulation océanique et de biogéochimie. La suite logique de ces développements est donc la modélisation du zooplancton et du micronecton.

La gamme de modèles SEAPODYM modélise avec parcimonie la chaîne trophique depuis le zooplancton jusqu'aux prédateurs supérieurs à l'aide de trois modèles. Cette thèse présente le modèle de biomasse de zooplancton SEAPODYM-LTL (pour lower trophic level, niveau trophique bas), ainsi qu'une analyse de sa sensibilité aux forçages. En effet, la particularité de ces modèles est leur forçage offline par des champs de courants, température et production primaire produits par d'autres modèles. Le modèle SEAPODYM-LTL est également comparé au modèle PISCES (NPZD), et présente des performances similaires à ce dernier dans le cas testé.

Afin d'améliorer les prédictions du modèle SEAPODYM-MTL (mid-trophic level, i.e. le modèle de biomasse de micronecton), une méthodologie d'assimilation de données a été mise en place pour affiner la paramétrisation utilisée. Des données d'acoustique active (38kHz) sont donc utilisées pour enrichir le modèle. Cette méthodologie a été conçue autour d'un cas test présenté dans cette thèse. L'extension du jeu de données acoustiques assimilées au modèle a permis de mettre en évidence le besoin de mieux modéliser les profondeurs des couches verticales de SEAPODYM. Cela qui a été réalisé également à l'aide du jeu de données acoustiques évoqué précédemment. Cette étude est également présentée dans cette thèse.

Mot-clés : modélisation, zooplancton, micronecton, observations, assimilation de données, acoustique active

Abstract

Zooplankton and micronekton are the first marine trophic levels. Different by their size ($200\mu\text{m}$ to 2mm for zooplankton, 2 to 20cm for micronekton), these two groups undergo diel vertical migration from depth by day to the surface during the night. These migrations create major organic matter fluxes between the deep ocean and the surface. Biogeochemical cycles are of great importance for climate change studies. These studies are conducted with ocean global circulation model and biogeochemical model. The way to go is develop low and mid-trophic level modelling approaches.

SEAPODYM ensemble of models are three parsimonious model of biomass at diverse level of the trophic chain, from zooplankton to top predators. This thesis introduce the zooplankton biomass model SEAPODYM-LTL (lower trophic level) and a forcing fields sensitivity analysis. Indeed, these model are forced off line by currents, temperature and primary production fields produced by other models. SEAPODYM-LTL has also been compared to PISCES (NPZD) and both have similar performance score in this study.

In order to improve SEAPODYM-MTL (mid trophic level) predictions, a data assimilation framework has been developed to find a better parameterisation. 38kHz active acoustic data have been used to improved the model. This methodology has been develop thanks to a test case that we present in this thesis. The gathered acoustic dataset permitted to show the need of a better definition of vertical layer depths. It has been developed using the acoustic dataset. The related study is presented in this thesis.

Keywords : modelisation, zooplankton, micronekton, observations, data assimilation, active acoustics

Table des matières

	Page
Chapitre 1 Introduction	1
I Zooplancton et Micronecton : le peuple migrateur	1
II Estimation de la biomasse et de la distribution spatiale de zooplancton et de micronecton	10
III SEAPODYM, un modèle de biomasse marine	15
Chapitre 2 Application simple du modèle avec une seule composante : modélisation de la distribution spatiale du mesozooplancton et validation	30
I SEAPODYM-LTL : un modèle parcimonieux	30
II SEAPODYM-LTL : a parsimonious zooplankton dynamic biomass model	32
III Perspectives	54
Chapitre 3 Modélisation de la distribution spatiale du micronecton	57
I Paramétrisation de SEAPODYM-MTL	57
II Données assimilées dans SEAPODYM-MTL	59
III Optimization of a micronekton model with acoustic data (Lehodey et al. (2015))	60
IV Un nouveau jeu de données plus étendu : résultats de l'assimilation de données en global, nouveaux défis	75
Chapitre 4 Définir les limites des couches biologiques dans SEAPODYM	81
I Développement d'un Average Micronektonic Layer Model	81
II Modeling depths of ocean bio-acoustic layers from space	84
III Une nouvelle définition des couches	105
Chapitre 5 Perspectives	108
I Paramétrisations optimales	108
II SEAPODYM-MTL et les différentes chaînes trophiques	110
III Établir un plan d'échantillonnage idéal du micronecton	117
IV Étudier l'accumulation de polluants dans la chaîne trophique	118
Chapitre 6 Conclusion	122

Table des figures

	Page
Chapitre 1 Introduction	
Fig. 1.1 Positions estimées des animaux suivis par le programme TOPP	2
Fig. 1.2 Nombre d'études publiées en fonction du niveau trophique	3
Fig. 1.3 Position des échantillonnages de COPEPOD	3
Fig. 1.4 Planche naturaliste de Giesbrecht (1892) représentant des copépodes	4
Fig. 1.5 Exemples d'organismes micronectoniques	6
Fig. 1.6 Biomasse saisonnière du micronecton de la mer de Tasman	7
Fig. 1.7 Type de migration verticale de trois espèces de poissons au Nord du Japon	8
Fig. 1.8 Pompe biologique	10
Fig. 1.9 Variation saisonnière des migrations nycthémerales	10
Fig. 1.10 Evitement d'un filet	12
Fig. 1.11 Proportion des études calculant des statistiques de performance des modèles biogéochimiques	15
Fig. 1.12 Chaîne trophique modélisée par la suite de modèles SEAPODYM	16
Fig. 1.13 Modélisations des migrations dans SEAPODYM-MTL	18
Fig. 1.14 Schéma du groupe mesozooplancton (SEAPODYM-LTL)	21
Fig. 1.15 Schéma des groupes fonctionnels micronectoniques (SEAPODYM-LTL)	22
Chapitre 2 Application simple du modèle avec une seule composante : modélisation de la distribution spatiale du mesozooplancton et validation	
Fig. 2.1 Schéma des relations entre les variables d'états de modèles classiques NPZD	31
Fig. 2.2 Schéma des relations entre les variables d'états de SEAPODYM-LTL	32
Fig. 2.3 Champs de production primaire climatologique annuel	55
Fig. 2.4 Champs de de biomasse de zooplancton climatologique annuel	55
Chapitre 3 Modélisation de la distribution spatiale du micronecton	
Fig. 3.1 Temps de maturité	58
Fig. 3.2 Temps de recrutement	58
Fig. 3.3 Positions des données utilisées dans les expériences.	75
Fig. 3.4 Positions des observations colorées en fonction de leur zone.	76
Fig. 3.5 Expérience n°1 : Diagramme de distribution de l'énergie entre les groupes fonctionnels	77

Chapitre 4 Définir les limites des couches biologiques dans SEAPODYM

Fig. 4.1	Exemple d'un transect acoustique entre la Nouvelle-Zélande et la Tasmanie	82
Fig. 4.2	Représentation d'un profil acoustique et des profondeurs l_i^{obs} des limites des différentes couches i	83
Fig. 4.3	Exemple du fonctionnement de la méthode de détection de la limite des couches . . .	83
Fig. 4.4	Exemple du transect acoustique entre la Nouvelle-Zélande et la Tasmanie de la figure 4.1 avec les limites observées et prédites par l'ancien et le nouveau modèle.	105
Fig. 4.5	Ratios de densité acoustique déterminés avec le nouvelle et l'ancienne définition. . . .	106

Chapitre 5 Perspectives

Fig. 5.1	Relation entre temps de développement et température du milieu possible à partir d'un ensemble d'espèces	109
Fig. 5.2	Position d'une tortue olivâtre superposée à la biomasse épipélagique de micronecton . .	111
Fig. 5.3	Chaines trophiques modélisées par la suite de modèles SEAPODYM	112
Fig. 5.4	Groupes phytoplanctoniques majoritaires selon l'algorithme PHYSAT (avril 2004) . .	113
Fig. 5.5	Distribution spatiale du micronecton issue des diatomées en (gWW/m2)	114
Fig. 5.6	Distribution spatiale du micronecton issue du petit phytoplancton en (gWW/m2) . .	115
Fig. 5.7	Pourcentage du micronecton issue des diatomées	116
Fig. 5.8	Différents dispositifs d'observations du micronecton par acoustique.	118
Fig. 5.9	Premier test de simulation du transfert de Cs137 au micronecton du modèle SEAPODYM-MTL	119

Liste des tableaux

Page

Chapitre 1 Introduction

Tab. 1.1	Micronecton dans l'Océan Pacifique	5
Tab. 1.2	Réponse acoustique rapportée à la masse.	13

Chapitre 3 Modélisation de la distribution spatiale du micronecton

Tab. 3.1	Récapitulatif des expériences et de leurs caractéristiques.	76
Tab. 3.2	Expérience n°1 : nombre d'observations, total et par pas de temps, nombre de pas de temps, et erreurs associées aux prédictions.	77
Tab. 3.3	Expérience n°2 : nombre d'observations total et par pas de temps, nombre de pas de temps, et erreurs associées aux prédictions.	78

Chapitre 1 Introduction

Sommaire

I	Zooplankton et Micronecton : le peuple migrateur	1
A	Des maillons critiques et peu connus de la chaîne trophique	1
B	Migrations nycthémérales, caractéristiques et conséquences	7
II	Estimation de la biomasse et de la distribution spatiale de zooplankton et de micronecton	10
A	Collecte d'organismes : la méthode d'observation traditionnelle	11
B	L'acoustique active : une méthode d'observation innovante et sans collecte . .	12
C	Modélisation	13
III	SEAPOODYM, un modèle de biomasse marine	15
A	Un modèle de dynamique de population lié à l'écosystème	16
B	... et spatialisé	17
C	Améliorer la paramétrisation du modèle à l'aide de l'assimilation de données . .	20
D	Modéliser le mésozooplankton ou le micronecton à l'aide de SEAPOODYM . .	21

I Zooplankton et Micronecton : le peuple migrateur

A Des maillons critiques et peu connus de la chaîne trophique

Depuis les années 1990, le phytoplancton est observable par satellite (McClain (2009)) sur la majeure partie de la surface de l'océan (90% de sa surface) avec des erreurs d'observation relativement faibles (moins de 5% Bailey and Werdell (2006)). Les produits satellitaires de couleur de l'eau peuvent maintenant également inclure le type phytoplanctonique des organismes observés (Alvain et al. (2005), Alvain et al. (2008)). L'évolution du phytoplancton à travers les siècles futurs est également envisageable grâce aux nombreux modèles biogéochimiques couplés aux modèles climatiques du système Terre-Océan (Bopp et al. (2001), Stock et al. (2011), Steinacher et al. (2010)).

À l'autre bout de la chaîne trophique, les prédateurs supérieurs sont suivis de près grâce aux systèmes de positionnement satellite ARGOS et GPS. Leurs moindres mouvements sont enregistrés plusieurs fois par seconde par les balises qu'ils portent (Gallon et al. (2013), Viviant et al. (2010), Guinet et al. (2014)). Si l'animal remonte à la surface régulièrement pour respirer la balise peut à ce moment-là émettre un résumé des données en quasi-temps réel. Les prédateurs marins qui n'ont pas besoin de respirer en surface sont eux

équipés de balises dites archives qui mesurent et stockent les données physiques sur leur environnement dans une mémoire électronique. Le positionnement par satellite étant impossible dans les profondeurs de l'océan, les données de température et/ou de lumière peuvent être utilisées une fois la balise récupérée pour déduire la trajectoire de l'animal (Hammerschlag et al. (2011), Block (2005)). Cette technologie mise au point il y a une vingtaine d'années représente aujourd'hui une quantité impressionnante de données. Par exemple le programme américain Tagging Of Pacific Predators (Block et al. (2011), figure 1.1) a marqué 1700 animaux de 23 espèces différentes.

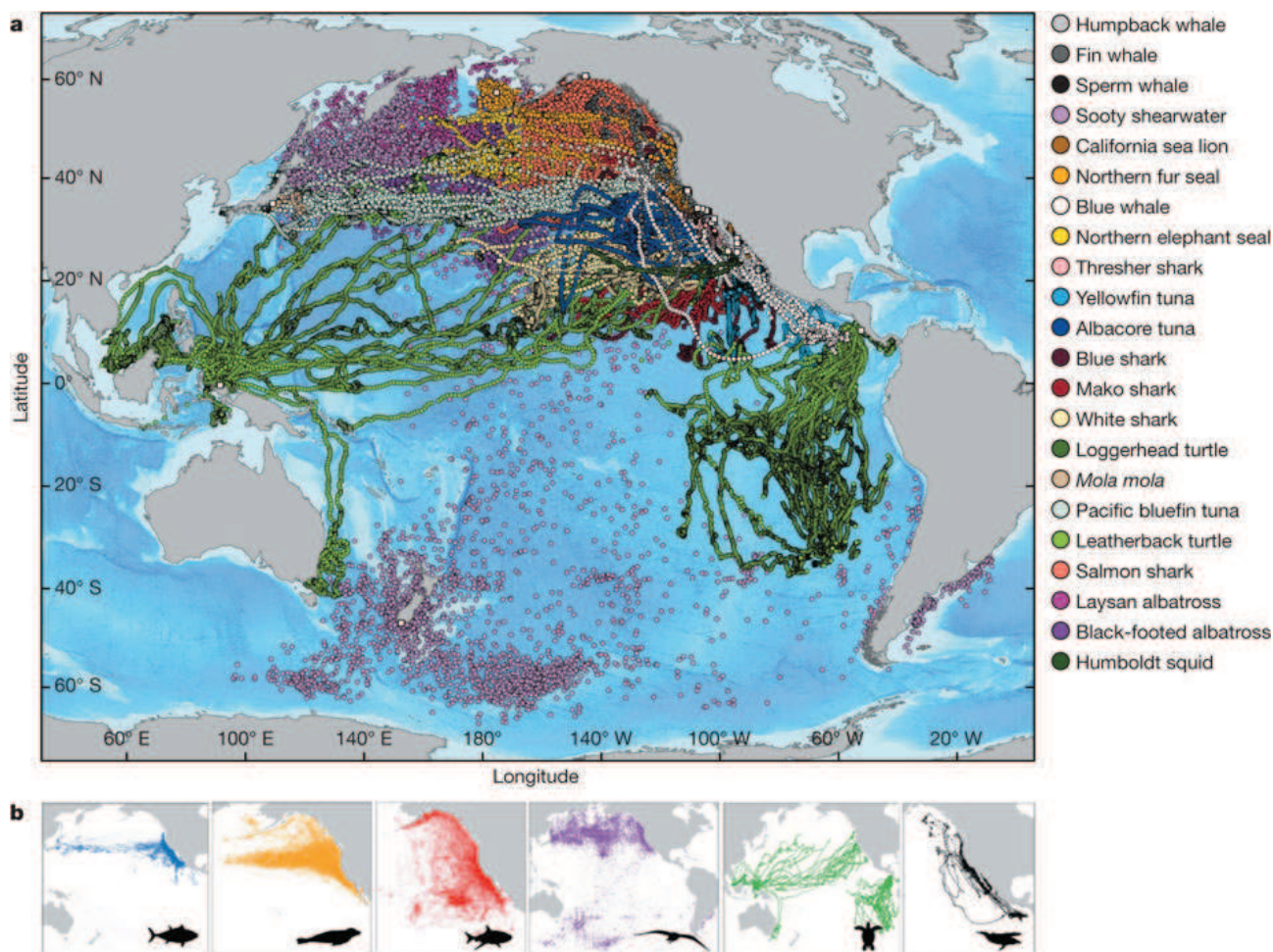


Figure 1.1 – Positions estimées des animaux suivis par le programme TOPP, (a) positions moyennes journalières (rond) de toutes les espèces marquées et lieu de déploiement médian (carré blanc : USA, Mexique, Nouvelle-Zélande, Japon, Papouasie), (b) positions moyennes journalières des groupes d'espèces principaux (thons, pinnipèdes, requins, oiseaux, tortues, cétacés). (D'après Block et al. (2011))

Entre ces deux types d'organismes bien différents se nichent deux groupes de niveau trophique intermédiaire : le zooplancton et le micronecton. Le moins connu et le moins étudié des deux, le micronecton (Figure 1.2), est d'une importance capitale dans la chaîne trophique : il constitue un échelon intermédiaire entre les producteurs primaires et les prédateurs supérieurs. La taille des organismes constituant ce groupe varie entre 2 et 20 cm. Ils se distribuent dans le domaine pélagique entre la surface et les grandes profondeurs. Leur spécificité éponyme est leur capacité à nager contre le courant sur de faibles distances (quelques dizaines de centimètres) malgré leur petite taille (Brodeur and Yamamura (2005)). Bien qu'ils soient présents dans toutes les mers du globe, la biomasse totale de ces organismes est sujette à discussion. La dernière estimation en date (Irigoien et al. (2014)), effectuée à l'aide de méthode acoustique, l'évalue à 10

milliards de tonnes soit un ordre de grandeur supérieur à l'estimation de Gjøsæter and Kawaguchi calculée à l'aide de données de chalutage (Gjøsæter and Kawaguchi (1980)).

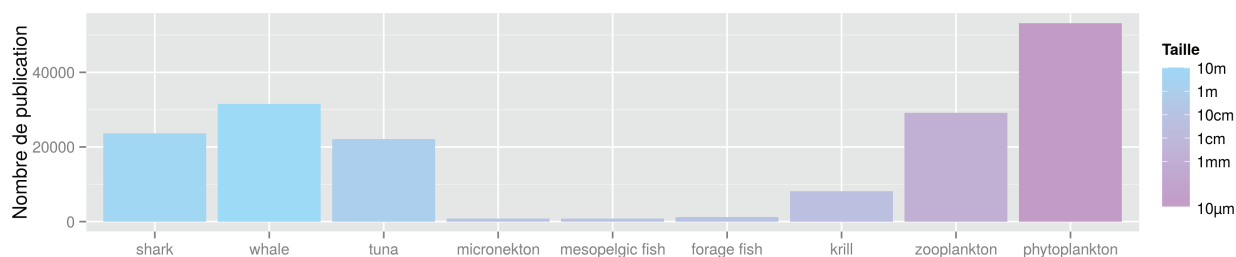


Figure 1.2 – Nombre d'études disponibles sur sciencedirect.com, en fonction du mot clé choisi. Les groupes d'organismes sont classés par niveau trophique du plus élevé au plus faible, la couleur des barres représente la taille moyenne du groupe considéré.

Le zooplancton est quant à lui un peu moins secret et beaucoup plus étudié, notamment via les grandes quantités d'observations disponibles (exemple de la base de données COPEPOD, Figure 1.3). Nous définissons ce groupe par sa position dans le réseau trophique (consommateur primaire) et sa taille (0.2 à 2 mm, Sieburth et al. (1978)). Ce groupe est également un assemblage de taxons très différents occupant une même fonction trophique et fenêtre de taille. On y retrouve des organismes passant toute leur vie dans ce groupe (holoplancton) et d'autres où seuls les stades larvaires sont représentés (méroplancton : cnidaires, mollusques, poissons, crustacés). Le taxon dominant du zooplancton permanent est la sous-classe des copépodes (voir Figure 1.4). Ces petits crustacés, de par à leur grande diversité spécifique, sont adaptés à de nombreux écosystèmes et par conséquent présents dans tous les océans du globe.

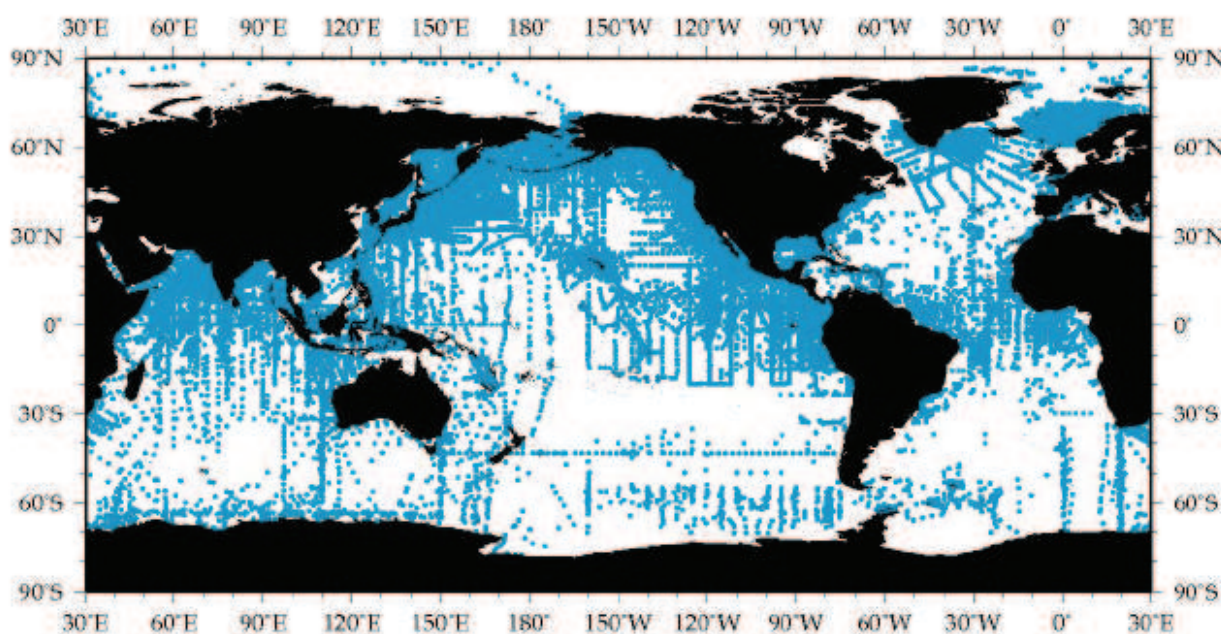


Figure 1.3 – Position des échantillonnages regroupés dans la base de données COPEPODD'après Moriarty and O'Brien (2013))

Comme pour le zooplancton, le micronekton inclut un grand nombre d'organismes de façon transitoire lors de leur développement larvaire et juvénile. C'est le cas de la plupart des juvéniles de poissons. En

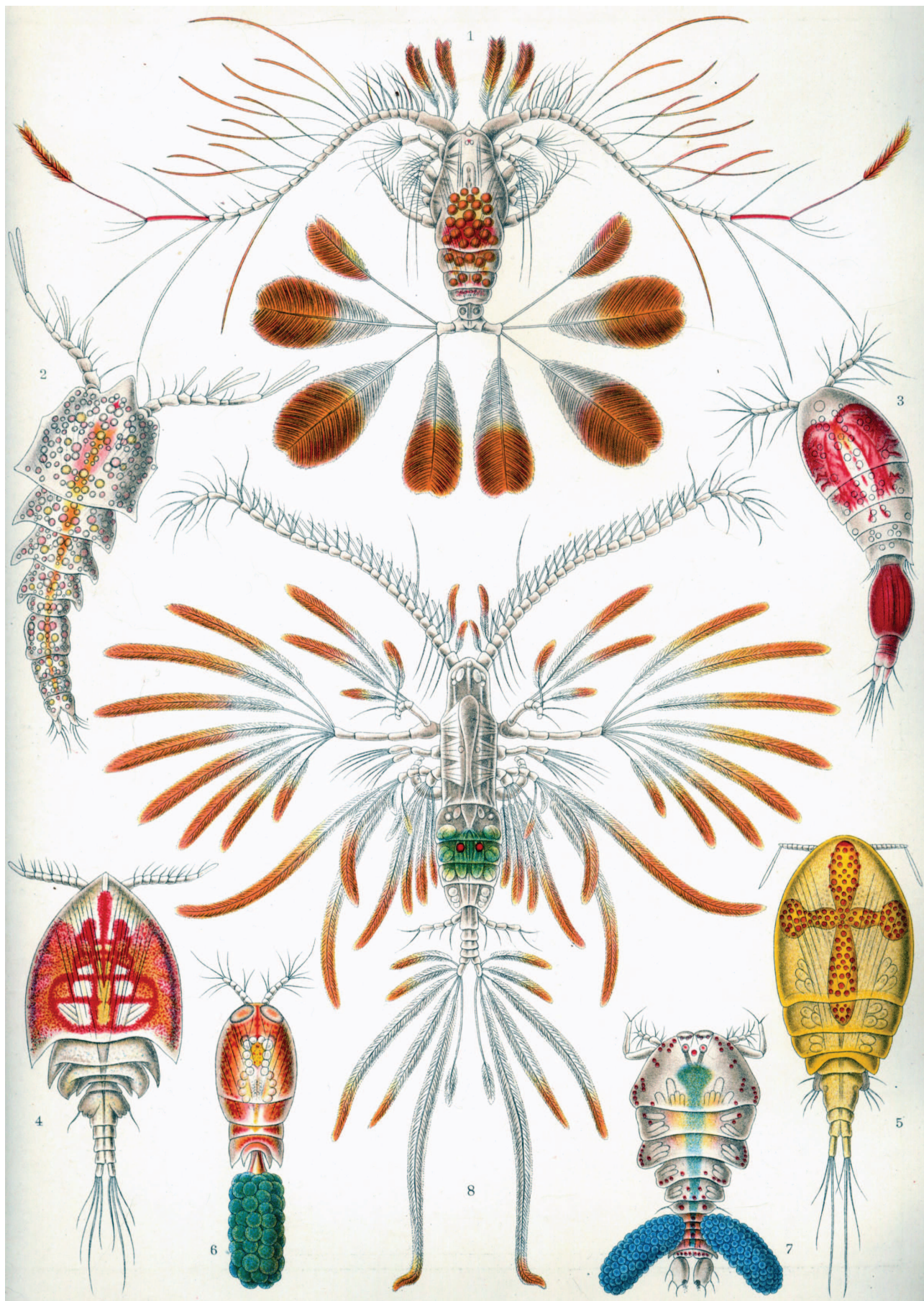


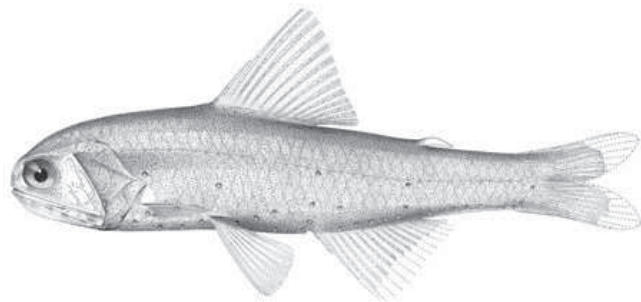
Figure 1.4 – Planche naturaliste de Wilhelm Giesbrecht (1892) représentant des espèces de copépodes. (1) *Calanus pavo* - (2) *Clytemnestra scutellata* - (3) *Oncaea venusta* - (4) *Cryptopontius thorelli* - (5) *Acontiophorus scutatus* - (6) *Corycaeus cenustus* - (7) *Sapphirina darwinii* - (8) *Augaptilus filigerus*

plus des poissons, le micronecton comporte de nombreux autres taxons parmi les grands groupes céphalopodes, crustacés et organismes gélatineux. Les poissons mésopélagiques, vivant entre la couche de surface et 1000 m de profondeur, sont majoritairement représentés par les myctophidés (Brodeur and Yamamura (2005)). Ces petits poissons lanternes se retrouvent dans tous les océans des plus froids (*Electrona Antarctica*, Figure 1.5b) aux tropiques (*Diaphus Garmani*, Figure 1.5a). Du fait de leur importante biomasse et de l'absence d'utilisation directe pour la consommation humaine, une étude s'est attachée à établir la faisabilité de transformer ces poissons fourrages en surimi (Shaviklo and Rafipour (2013)). Aussi abondantes que les myctophidés en terme de biomasse (Blackburn (1968), Blackburn et al. (1970), Nishikawa et al. (2001) et Brodeur and Yamamura (2005)), les euphausiacés sont la composante principale des crustacés micronectoniques (Figure 1.5c et 1.5d). Bien que considérés pendant longtemps comme du zooplancton, Nicol et Brodeur and Yamamura plaident en faveur de leur regroupement dans le micronecton du fait de leur comportement et de leur taille. Les céphalopodes constituent un groupe beaucoup moins abondant : Passarella and Hopkins estimèrent leur abondance à moins d'un individu par 10 km² dans le Golfe du Mexique (Passarella and Hopkins (1991)). Toutefois, les estimations réalisées à l'aide de traits de chalut sous-estiment leur abondance en raison de la capacité des ces animaux à éviter le filet. Enfin, le micronecton gélatineux est souvent oublié et non-inclus dans le micronecton. Les organismes gélatineux de grande taille sont principalement considérés comme du macrozooplancton. Cependant certains, comme les cténophores, ont des comportements (de migrations verticales notamment) et des tailles en accord avec notre définition du micronecton.

L'abondance relative de ces groupes taxonomiques varie d'un écosystème à l'autre (voir Figure 1.6 et tableau 1.1). Néanmoins, la faible abondance des céphalopodes semble universelle. Certains écosystèmes présentent une écrasante majorité de crustacés (mer de Wedell, Lancraft et al. (2004)), tandis que d'autres sont principalement poissonneux (Mer de Tasman, Figure 1.6).

Composante micronectonique	Concentration de micronecton (mL/10 ³ m ³)		
	<300miles	300 to 600 miles	>600 miles
Poissons	9.07	3.65	2.34
Crustacés	9.22	3.67	0.74
Céphalopodes	1.67	0.30	0.15
Total	19.96	7.62	3.23

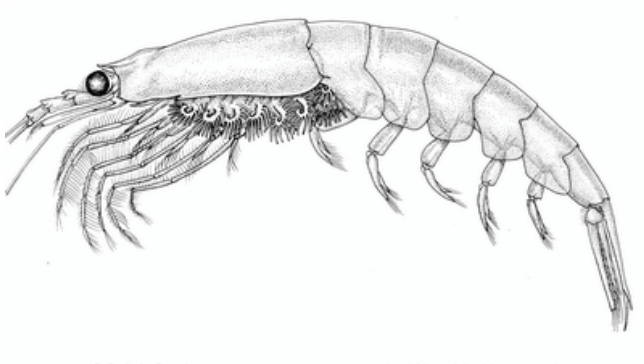
Table 1.1 – Moyennes des concentrations en micronecton de 69 stations groupées en fonction de leur distance à la côte (Adapté du tableau 6 de Blackburn (1968))



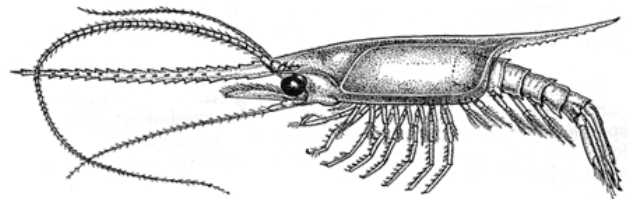
(a) *Diaphus garmani*



(b) *Electrona antarctica*



(c) *Euphausia superba*



(d) *Gnathophausiidae*



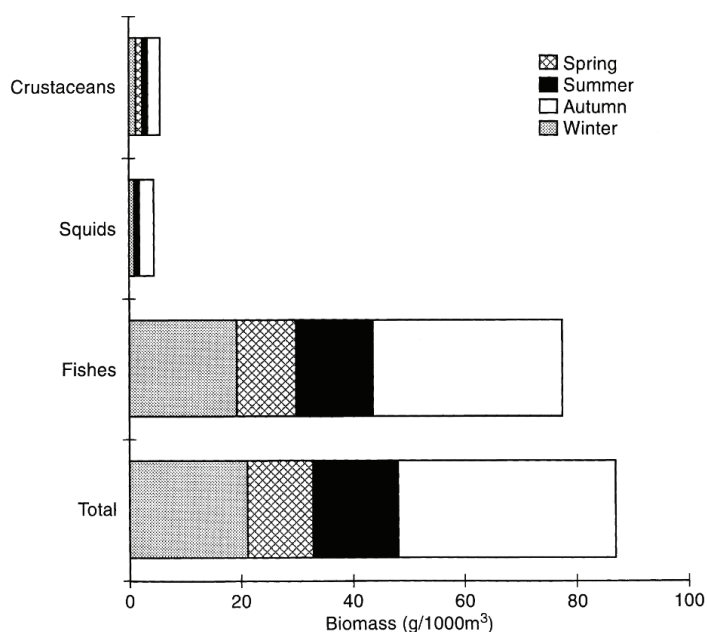
(e) *Pyroteuthis margaritifera*



(f) *Pterygoteuthis gemmata*

Figure 1.5 – Exemples d'organismes micronectoniques : poissons ((a) et (b)), crustacés ((c) et (d)) et céphalopodes ((e) et (f)).

Figure 1.6 – Biomasse saisonnière de micronecton pêchée dans les 900 premiers mètres de la colonne d'eau au sud de la Tasmanie (Figure 3 de Williams and Koslow (1997))



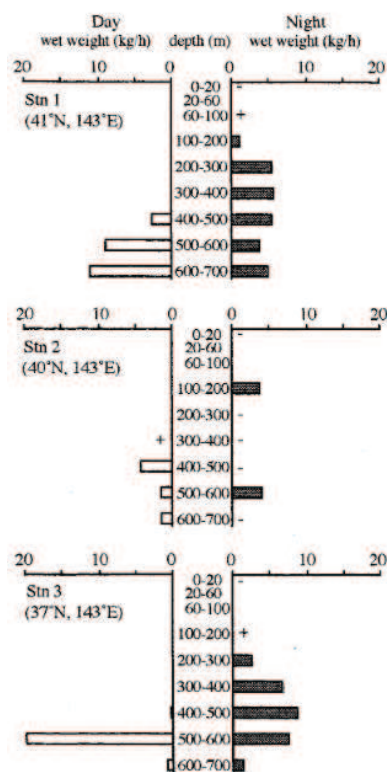
B Migrations nycthémerales, caractéristiques et conséquences

Le zooplancton et le micronecton partagent un comportement singulier : ils effectuent des migrations nycthémerales. Ce phénomène de migration verticale d'un habitat profond en journée vers un habitat plus en surface de nuit a été défini précisément par Lampert comme étant l'enfoncement de la profondeur moyenne d'une population pendant le jour (en opposition avec l'idée d'un mouvement vers le bas de tous les organismes, Lampert (1989)). C'est un comportement qui a été observé depuis les grandes expéditions océanographiques de la fin du 19ème siècle dont la plus célèbre est sans doute celle du Navire Challenger (1873-76). Barham les observa dans les années 60 à l'aide des tous premiers écho-sondeurs (Barham (1966)). La plupart de ces migrations se passent dans les mille premiers mètres de l'océan (Perry and Neilson (1988)). Néanmoins, l'existence de migration d'une plus grande amplitude est maintenant envisagée (Domokos (2009)).

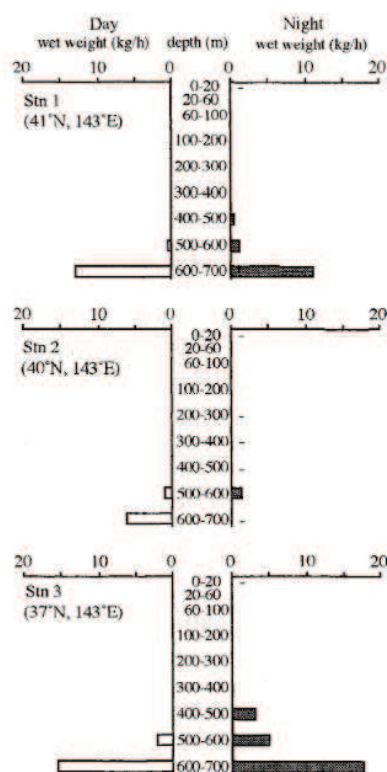
Ces migrations ont des caractéristiques très variées d'une espèce à l'autre dans un même écosystème mais aussi d'un écosystème à l'autre pour une même espèce. Watanabe et al. échantillonnèrent le micronecton présent à différentes profondeurs (de 0 à 700 m) de trois stations au Nord du Japon en Juillet 1995 (voir Figure 1.7 extraite de Watanabe et al. (1999)). On peut observer pour ces trois espèces des comportements de migration tout à fait différents dans la première station (41°N, 143°E) : *Stenobrachius nanochir* présente une absence de migration (biomasse maximale de jour et de nuit dans la couche 600-700 m), *Notoscopelus japonicus* une migration typique de 300-500 m le jour à 20-200 m la nuit, tandis que *Lampanyctus jordani* présente un étalement de sa distribution nocturne (200-700 m) par rapport à sa distribution diurne (500-700 m). Si l'on s'attarde sur *Notoscopelus japonicus*, on peut observer que cette espèce présente deux types de migrations : une migration typique à la station 1 et une absence de migration à la station 3. Il n'est donc pas possible d'attribuer un comportement à une espèce ou un écosystème.

À la vue de la remarquable synchronicité entre les levers/couchers de soleil et ces migrations, la lumière semble être le facteur déclenchant un tel comportement (Zaret and Suffern (1976), Hays (2003), Benoit-Bird et al. (2009)). La façon exacte dont la lumière influence ces migrations n'est pas connue. La théorie,

(a) *Lampanyctus jordani*



(b) *Stenobranchius nannochir*



(c) *Notoscopelus japonicus*

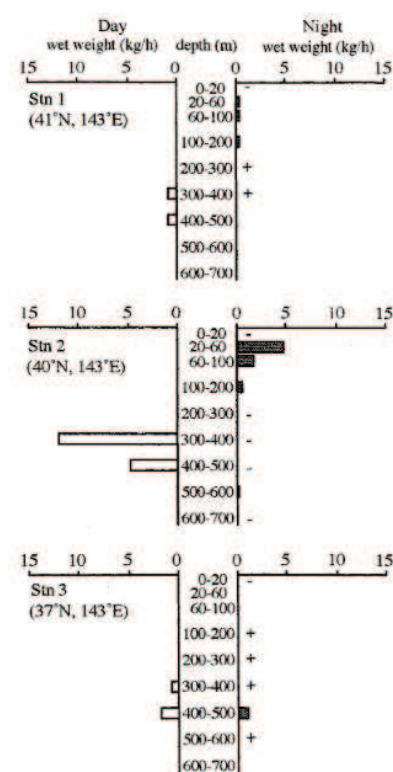


Figure 1.7 – Type de migration verticale de (a) *Lampanyctus jordani*, (b) *Stenobranchius nannochir* et (c) *Notoscopelus japonicus* au Nord du Japon en Juillet 1995. Un signe 'plus' (+) indique une prise de moins de 0.1 kg.h⁻¹, un moins (-) une couche non échantillonnée et un blanc une prise vide. (De Watanabe et al. (1999))

dite de l'isolume, propose que les organismes se positionnent dans la colonne d'eau à une profondeur telle qu'ils reçoivent en permanence la même irradiance lumineuse. Toutefois, cette théorie est difficilement vérifiable puisque les organes sensoriels des poissons sont des capteurs plus précis que ceux dont nous disposons à l'heure actuelle (Benoit-Bird et al. (2009)). Benoit-Bird et al. a notamment essayé sans succès de tester cette hypothèse au Sud d'Hawaï. Dans son étude, un proxy de la lumière ayant une échelle temporelle plus lente (à savoir le cycle lunaire) était plus corrélé aux profondeurs observées qu'un quelconque isolume.

S'il existe un consensus autour de la lumière comme élément déclencheur de ces migrations, il n'en est pas de même pour les facteurs pouvant expliquer ces migrations. Dans une synthèse bibliographique, Lampert dégagea deux causes théoriques probables : ces migrations apportent un avantage métabolique ou démographique (d'un point de vue énergétique) ou une diminution du risque de prédation dans la pénombre. De nombreuses études ont tout d'abord suggéré un avantage à migrer (McLaren (1963), McLaren (1974), Enright (1977) mais il est maintenant admis que la migration représente un coût énergétique associé à une diminution de croissance (Lampert (1989), Lampert et al. (1988), Aksnes and Giske (1990), Kremer and Kremer (1988)). Néanmoins ce coût semble être compensé par la diminution du risque de prédation pendant la journée, comme le résumait Kremer and Kremer : "il est préférable d'avoir faim mais d'être vivant !" ("better hungry than dead!", Kremer and Kremer (1988)).

Le seuil de lumière déclenchant les migrations est bas ou couplé à d'autres phénomènes. En effet, de telles migrations ont été observées en été sous un couvert de glace en Arctique comme en Antarctique à très hautes latitudes (Fortier et al. (2001), Tanimura et al. (2008)). Cisewski et al. mesura la vitesse verticale au cours de la journée (moyennée entre 100 et 200m) sur toute une année à trois stations antarctiques à hautes latitudes (entre 64 et 69°S), et observa une diminution de la vitesse moyenne durant l'été (voir Figure 1.9) extraite de Cisewski et al. (2010)).

Ces migrations ont un impact environnemental fort puisqu'elles participent au mécanisme appelé pompe biologique. Les migrations nycthémérales exportent de la matière organique vers les profondeurs dans des proportions importantes du fait de la respiration et de l'excrétion de pellets fécales en profondeur alors que l'ingestion de nourriture se produit en surface (Longhurst and Harrison (1988); Longhurst et al. (1990)). La respiration zooplanctonique peut représenter jusqu'à 30% du flux de carbone particulaire à travers la pycnocline (Longhurst et al. (1990), Hidaka et al. (2001)). Quant à l'excrétion de pelettes fécales par le zooplancton, elle représente 2 à 13% du flux de carbone dissout vers l'océan profond et 15 à 66% du flux d'azote (Steinberg et al. (2000); Steinberg et al. (2002)). L'export de carbone en dehors de la zone épipélagique par les migrations de poissons mésopélagiques est estimée entre 10 et 40% de l'export total de carbone vers le fond (Davison et al. (2013)).

Figure 1.8 – Pompe biologique : flux de carbone. Les différents processus représentés sont : la respiration (C_r), la défécation (C_d), l'excrétion de carbone organique (C_e) et la mort (C_m). D'après Davison et al. (2013)

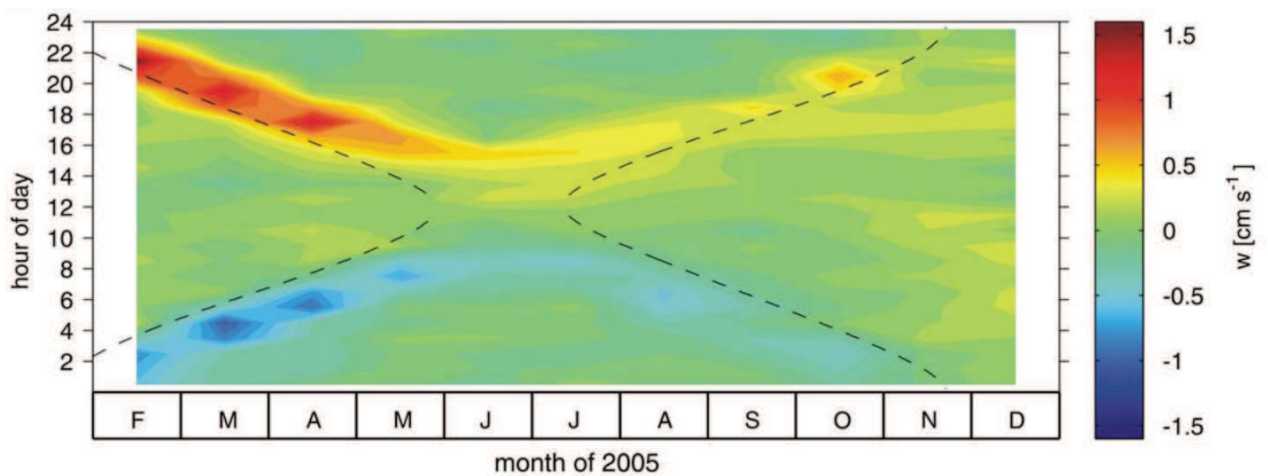
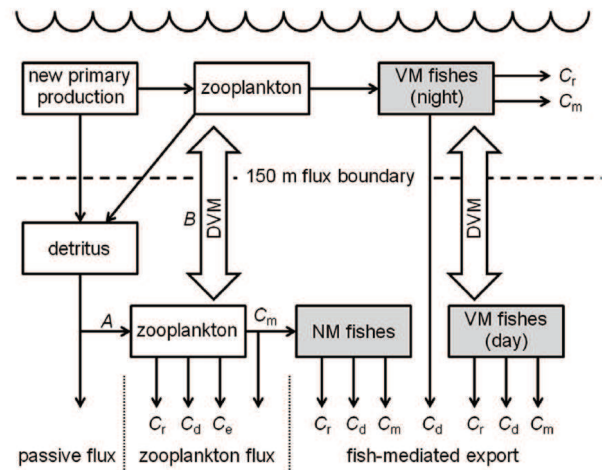


Figure 1.9 – Variation saisonnière des migrations nyctémérales à 69°S pendant l'année 2005. Moyenne des vitesses verticales entre 100 et 200m. Les lignes pointillées représentent les heures de lever et coucher de soleil. (Extrait de Cisewski et al. (2010))

II Estimation de la biomasse et de la distribution spatiale de zooplancton et de micronecton

Pour étudier un écosystème, deux méthodes s'offrent à nous : l'observer ou le modéliser. L'observation permet de voir le milieu aussi précisément que possible en un lieu et une date donnés. Ces observations sont limitées par la technique, le coût et le temps nécessaire aux observations. Néanmoins elles ne sont entachées que d'une erreur relativement faible. Pour observer les organismes marins qui nous intéressent on peut distinguer deux types de méthodes, avec ou sans collecte d'organismes.

La modélisation est la représentation d'un phénomène à l'aide d'outils mathématiques et souvent informatiques. La modélisation permet une approche multi-échelles, ainsi qu'une couverture importante des domaines spatio-temporels considérés. Néanmoins les erreurs de modélisations sont bien supérieures aux erreurs d'observations. Déterminer l'ensemble des processus affectant un organisme et les décrire mathématiquement est complexe.

Nous présentons ici les méthodes d'observations et de modélisations du zooplancton et du micronecton les

plus courantes.

A Collecte d'organismes : la méthode d'observation traditionnelle

La méthode la plus traditionnelle pour observer les organismes marins de petite taille est la collecte. Différentes méthodes peuvent être utilisées, spécialement pour le zooplancton dont la taille plus petite autorise de nombreuses méthodes de collecte.

Cas du zooplancton Utilisée depuis le XIX^{ème} siècle, la collecte avec un filet du zooplancton est la méthode la plus employée, même si les filets n'ont eu de cesse de s'améliorer. Ces filets font quelques mètres de long et sont constitués de mailles s'affinant en allant vers le bout de celui-ci, constitué par un tube solide fermé (non maillé). Ce filet n'a pas de dispositif de fermeture et d'ouverture, une contamination avec des organismes d'une profondeur non ciblée est donc possible. Néanmoins, il est peu coûteux et très facile à mettre en place (même sur des petits bateaux).

Il existe maintenant des filets avec des systèmes d'ouverture et de fermeture opérés depuis le bateau (le MULTINET par exemple) évitant les contaminations des profondeurs non-ciblées. Bien évidemment, ils sont plus coûteux et moins simple d'utilisation.

Une autre méthode consiste à utiliser une pompe et à filtrer l'eau de mer. Cette méthode garantit également une non-contamination des autres profondeurs et une estimation très précise du volume d'eau filtrée (contrairement à l'utilisation d'un filet). Néanmoins, il n'est pas possible d'échantillonner des profondeurs supérieures à 200m.

La plus ancienne méthode encore utilisée de nos jours fut créé en 1931 par Allister Hardy. Le Continuous Plankton Recorder (CPR) est inchangé depuis 1948, rendant les comparaisons aux données historiques fiables. Ce dispositif emprisonne le zooplancton entre deux gazes au fur et à mesure qu'il filtre l'eau de mer. Ce dispositif ne peut échantillonner que la surface. Plus d'un million de kilomètres ont été échantillonnés à l'aide du CPR entre 1948 et 2007.

Cas du micronecton En terme de collecte de micronecton, les méthodes se limitent aux filets. Néanmoins, du fait de leur taille, ceux-ci sont plus difficiles à manipuler que les filets à zooplancton. Une autre difficulté est liée à la spécificité des organismes ciblés par les filets : chaque filet a une sélectivité relative à chaque espèce, il est donc difficile de comparer des résultats issus d'échantillonnages réalisés avec différents filets (Kashkin and Parin (1983)). De plus, le micronecton a une capacité de nage beaucoup plus importante que le zooplancton et des phénomènes d'évitement du filet apparaissent (voir Kaartvedt et al. (2012), particulièrement la Figure 1 reproduit ici en Figure 1.10). Handegard and Tjøstheim (2005), Hjellvik et al. (2008), De Robertis and Handegard (2012)) ont également montré la capacité du micronecton à s'écarter en avant de la route des bateaux.

L'avantage majeur de la collecte d'organismes est la possibilité d'identification précise (possiblement jusqu'à l'espèce). Néanmoins, l'estimation de biomasse est entachée de nombreuses erreurs, que ce soit par les problèmes de sélectivité ou d'évitement (notamment pour le micronecton) mais aussi d'estimation de la biomasse précise dans le cas du zooplancton.

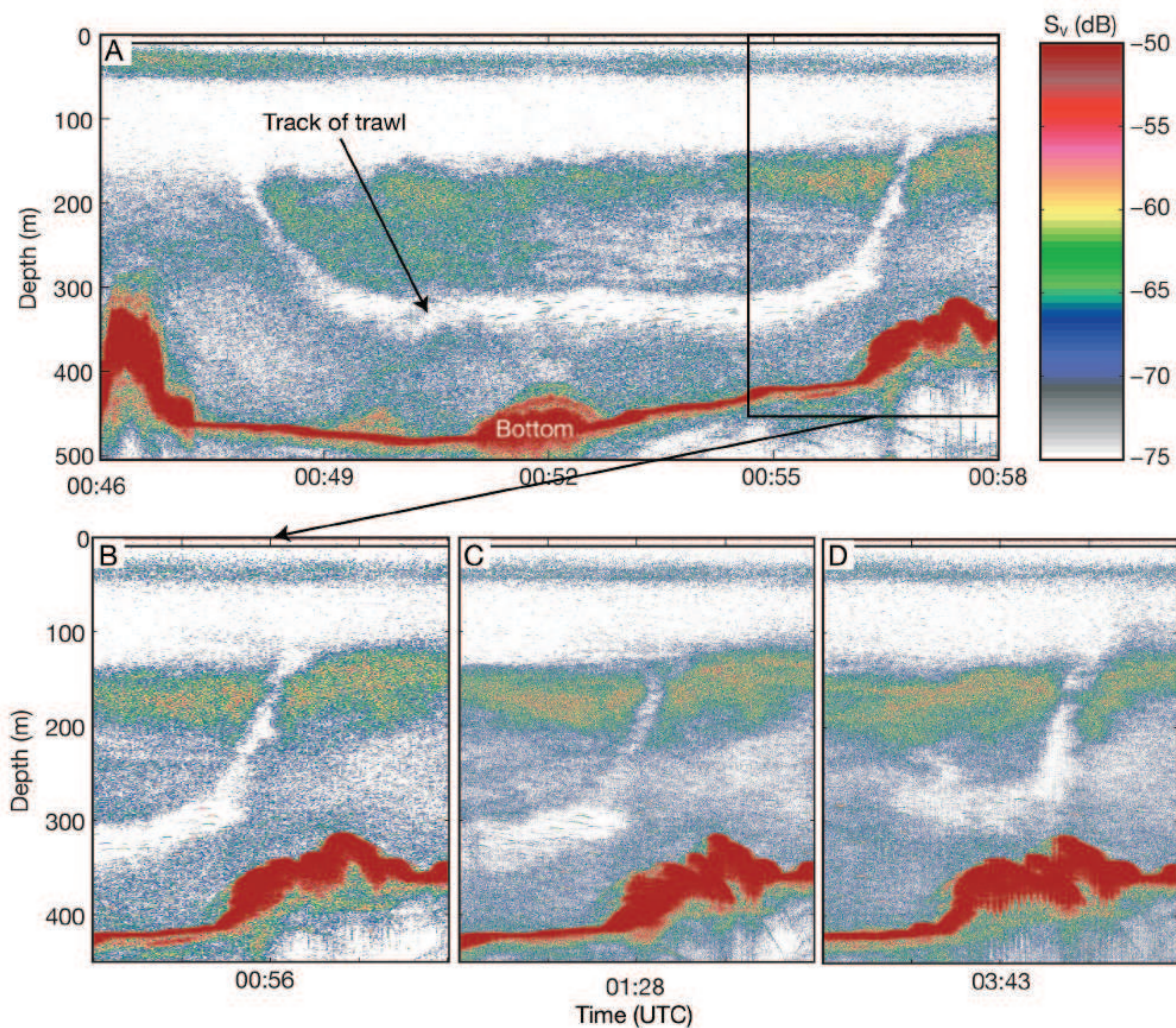


Figure 1.10 – Évitement d'un filet par le micronecton. (A) Echogramme lors du passage du filet. (B-D) Persistance du vide créé par le filet (déployé vers 23h30). Le filet n'avait pas de fond et n'était donc pas pêchant. Le vide ne peut donc pas être expliqué par les prises. La couleur de fond représente l'intensité de l'écho (volume backscattering strength, S_v). Figure extraite de Kaartvedt et al. (2012).

B L'acoustique active : une méthode d'observation innovante et sans collecte

L'acoustique active permet d'observer en profondeur ce qu'il y a dans le faisceau du sondeur. L'écho-sondeur émet une onde dans l'eau qui se propage jusqu'à heurter quelque chose et être réfléchi. Les réflexions sont traitées par l'écho-sondeur, ce qui permet une cartographie sous la trace du bateau. Les réflexions se produisant aux zones de discontinuité de densité, on observe des organismes vivants mais aussi des phénomènes physiques (comme les bulles d'air ou le plancher océanique). La fréquence de l'onde émise permet de cibler différentes tailles d'organismes (cette taille diminue quand la fréquence augmente), elle modifie également la portée de l'écho-sondeur (la portée diminue quand la fréquence augmente).

Cette méthode non-invasive (même pour les mammifères marins Lurton and DeRuiter (2011)) ne permet pas une identification taxonomique précise mais permet de distinguer des grands groupes en fonction de leur structure : absence ou présence de vessie natatoire, élasticité de la coquille, inclusion gazeuse, de densité semblable à l'eau... (voir Jech and Michaels (2006), Kloser et al. (2002))

La plupart de ces dispositifs sont placés sous la coque des navires et créent une zone aveugle de 15m en surface (tirant d'eau ajouté à la zone aveugle de l'écho-sondeur). O'Driscoll et al. (2009) a estimé que la biomasse cachée dans la zone aveugle pouvait représenter 20% de la réponse acoustique sur la colonne d'eau (11-29%, campagnes acoustique le long du Chatham Rise, Nouvelle-Zélande, durant les étés 2001 à 2007).

De plus, la conversion entre la réponse acoustique et la biomasse n'est pas évidente, puisque la réponse individuelle ("target strength") d'un individu est liée à son orientation (McGehee et al. (1998), Demer and Hewitt (1995)) et à sa structure. Par exemple, à 38kHz la réponse individuelle d'un poisson avec une vessie natatoire gazeuse est 10 fois plus importante qu'un poisson de la même taille n'ayant pas de vessie natatoire fonctionnelle (Davison (2011)). Le tableau 1.2 présente les réponses acoustiques rapportées à la masse de différents groupes d'animaux à la fréquence de 200kHz (extrait de Stanton et al. (1994)). Les modèles d'observation à partir de données acoustiques sont nécessaires mais restent imparfaits (Kloser et al. (2002), Davison et al. (2015)).

	Réponse acoustique Poids moyen		Caractéristiques physiques	
	Echelle	Echelle	Longueur	Poids humide
	linéaire	logarithmique (dB)	moyenne (mm)	moyen (mg)
Gastéropode (elastic-shelled)	19 000	43	2	4
Siphonophore (gas-inclusion)	260	24	37	1400
Décapode (fluid-like)	270	24	16	38
Poisson (swimbladder-bearing)	140	21	450	1 700 000
Salpe (fluid-like)	1	0	26	620

Table 1.2 – Réponse acoustique rapporté à la masse de l'organisme pour une fréquence de 200kHz. Les résultats sont similaires pour des fréquences plus élevées (extrait de Stanton et al. (1994)).

Les estimations de biomasse par acoustique active ou collecte d'organismes sont difficiles à comparer. Dans le cas du micronecton, l'estimation de la biomasse globale est complètement différente d'une méthode à l'autre (Koslow (1997), Irigoien et al. (2014)) et reste un objectif à long terme.

C Modélisation

Les observations sont "réelles" mais éparses et biaisées tandis que les modèles offrent une couverture extensive (à travers le temps ou le cycle de vie et le domaine spatial considéré) en proposant le meilleur résultat artificiel à partir de concepts mathématiques. Les modèles sont très divers et s'adaptent à la question posée.

Les modèles statistiques déduisent une quantité recherchée en fonction de variables explicatives à partir de relations empiriques. Il n'y a pas ici de recherche de liens de causalité mais la mise en évidence de corrélations. Par exemple, la distribution d'espèces peut être déduite de leur préférendum environnementaux : comme dans les études de Loots et al. (2007) et Koubbi et al. (2011) prédisant la distribution d'espèces de myctophidés respectivement sur le plateau de Kerguelen et l'océan Indien austral à partir de variables biotiques (concentration de chlorophylle *a*) et abiotiques (température, salinité, bathymétrie, concentrations

en nutriments, profondeur de la couche mélangée).

La biomasse et la distribution spatiale du zooplancton est classiquement décrite par des modèles biogéochimiques dits NPZD pour nutriment-phytoplancton-zooplancton-détritus. Ces modèles modélisent un ou plusieurs groupes de zooplancton à l'aide d'un système d'équations différentielles. Ils décrivent le cycle d'un ou plusieurs nutriments sur une partie du réseau trophique. Ils comptent un grand nombre de variables, d'équations et donc de paramètres. NEMURO (Kishi et al. (2007)) illustre ce fait : c'est un modèle NPZD non-spatialisé décrivant de fines interactions avec 11 variables, 2 cycles biogéochimiques et 75 paramètres (voir Chapitre 2).

Au niveau trophique du micronecton, on trouve peu d'approches de modélisation. La plus répandue est une approche centrée sur les espèces et qui décrit leur interactions, c'est le système Ecopath/Ecosim/Ecospace (Pauly et al. (2000), Christensen and Walters (2004)). Néanmoins, cet enchaînement de modèles nécessite un grand nombre de paramètres difficilement estimables et peut conduire à des conclusions discutables (Plaganyi and Butterworth (2004), Christensen and Walters (2004)).

Sur la base des travaux de Schaefer, Ryther et al. et Iverson, Jennings et al. développa une théorie liant les niveaux de production d'un groupe d'organismes à leur niveau trophique :

$$P_{TL} = PP \ TE^{TL-1} \quad (1.1)$$

où P_{TL} est la production d'un niveau trophique TL , PP la production primaire nette et TE le coefficient de transfert d'énergie. Strömberg et al. (2009) utilisent un modèle simple (sans advection) basé sur cette équation afin de prévoir la biomasse de zooplancton. Le modèle SEAPODYM utilise aussi ces principes au travers d'un modèle plus complexe incluant notamment de l'advection (Lehodey et al. (2010), Lehodey et al. (2015)). Cette thèse est centrée sur le modèle SEAPODYM et ces applications au micronecton et au zooplancton. Il est décrit en détail dans la section III de cette introduction.

La modélisation essaie de représenter une réalité, ici la distribution de groupes d'organismes dans l'océan, avec des variables telles que leur biomasse, sur un domaine spatial (océan discrétisé sur une grille plus ou moins fine, avec plus ou moins de couches) et une durée (avec un pas de temps défini plus ou moins long). Alors que les observations sont éparées, la modélisation couvre la totalité de son domaine spatio-temporel. Néanmoins, les prédictions de modèle doivent être confrontées aux observations. Arhonditsis et al. ont étudié l'ensemble des publications portant sur les modèles biogéochimiques aquatiques publiées entre 1990 et 2002 ($n=153$). Les résultats montrent une majorité d'études ne pratiquant pas d'analyse de sensibilité, d'optimisation, de calculs de qualité de l'ajustement ou même de validation (voir Figure 1.11, extraite de Arhonditsis et al. (2004)). Plus de 95% des études ne présentent pas de performance pour l'ensemble de leurs variables. Ils montrent également qu'il n'y a pas de corrélation entre le coefficient de détermination entre observation et prédiction du phytoplancton et le coefficient de détermination du zooplancton : la capacité d'un modèle à prédire un niveau trophique n'entraîne pas une bonne représentation du niveau trophique supérieur !

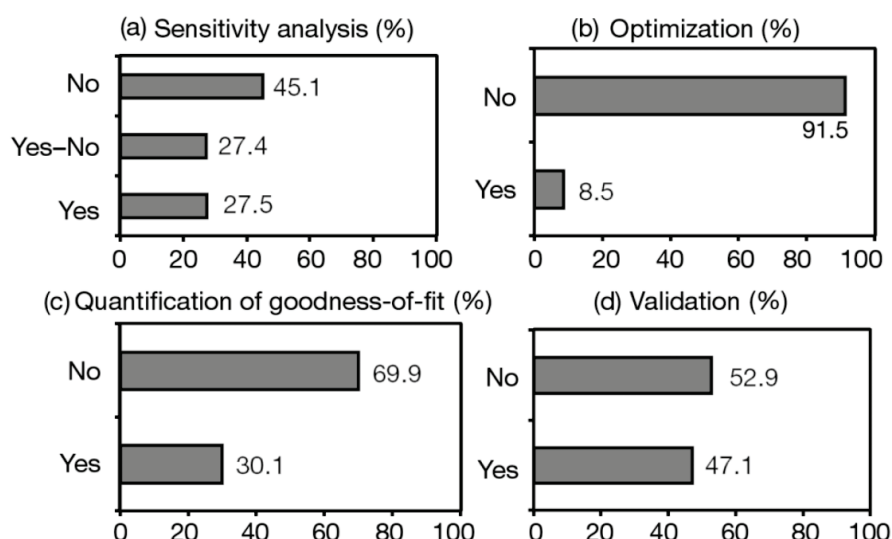


Figure 1.11 – Proportion des études calculant des statistiques de performance des modèles biogéochimiques (publiées entre 1990 et 2002) qui (a) effectuent des analyses de sensibilité, (b) utilisent des méthodes d'optimisation pour calibrer le modèle, (c) quantifient la qualité de l'ajustement prédictions/observations et (d) valident le modèle. Sur le panneau (a) 'YES-NO' indique la proportion d'approches qualitatives. (Extrait de Arhonditsis et al. (2004))

III SEAPODYM, un modèle de biomasse marine

SEAPODYM est un ensemble de modèles de biomasses d'organismes marins. Il comporte plusieurs composantes, chacune modélisant un échelon trophique particulier : le zooplancton (Chapitre 2), le micronecton (Lehodey et al. (2010); Lehodey et al. (2015)) ou les prédateurs supérieurs (principalement les thonidés mais aussi le chinchard ou les espadons, Lehodey et al. (2008); Hintzen et al. (2014); Dragon et al. (2015)).

C'est un ensemble de modèles eulériens, qui à l'instar des modèles de circulation océanique globale décrit des variables sur l'ensemble d'un domaine et non le long d'une trajectoire comme un modèle individu centré (donc lagrangien) pourrait le faire. Nous allons nous intéresser aux modèles décrivant le zooplancton (SEAPODYM-LTL pour "lower-trophic level", niveau trophique inférieur) et le micronecton (SEAPODYM-MTL pour "mid-trophic level", niveau trophique intermédiaire), qui sont en réalité deux configurations d'un même modèle.

SEAPODYM signifie "Spatial Ecosystem And POpulation DYnamics Model" (modèle spatial d'écosystème et de dynamique de population). C'est un modèle eulérien basé sur un système d'équation d'advection-diffusion-réaction. La partie III.A présente les aspects temporels, à savoir la modélisation des processus de dynamique de population, la mortalité et le recrutement ainsi que les liens avec l'écosystème. Quant à la partie III.B, elle s'appuie sur l'équation d'advection-diffusion pour souligner les hypothèses utilisées pour la simplification de celle-ci.

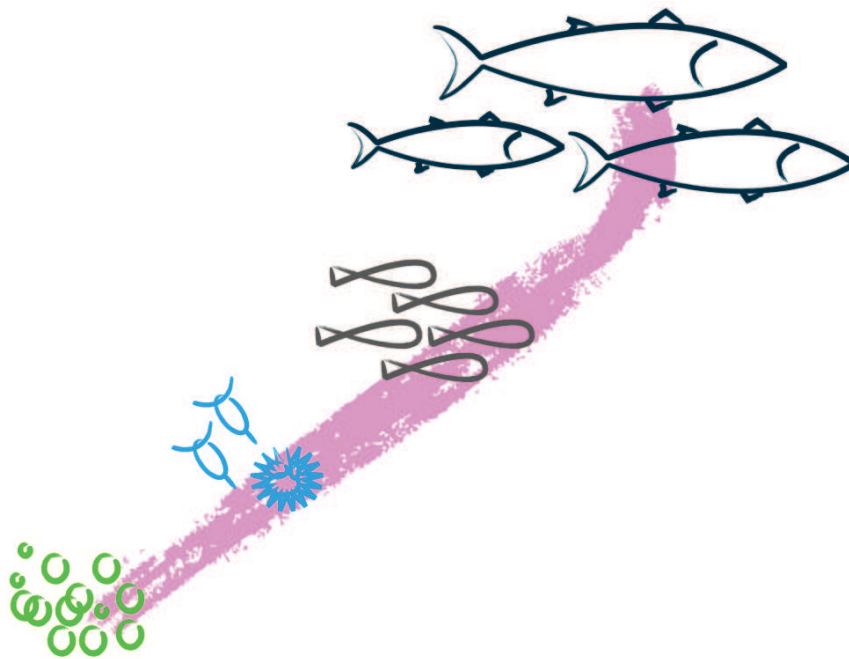


Figure 1.12 – La chaîne trophique simplifiée modélisée par la suite de modèles SEAPODYM. La ligne rose figure le flux d'énergie entre niveaux trophiques, les ronds vert le phytoplancton (non-modélisé dans SEAPODYM), le zooplancton en bleu, le micronecton en gris et les prédateurs (ici des thonidés) en bleu foncé.

A Un modèle de dynamique de population lié à l'écosystème ...

Commençons par considérer un groupe d'organismes marins de diverses espèces mais de taille similaire. Sans tenir compte des problématiques processus d'espace, la variation de la biomasse peut s'écrire de la façon suivante :

$$\frac{db}{dt} = \underbrace{-\lambda b}_{\text{Mortalité}} + \underbrace{p(\tau = \tau_r)}_{\text{Recrutement}} \quad (1.2)$$

où b est la concentration volumique de la biomasse (en gWW/m^3), p la concentration volumique de la production, t le temps, λ le taux de mortalité, τ l'âge et τ_r l'âge de recrutement. Pour l'instant ce modèle est un modèle de dynamique de population classique qui ne fait pas intervenir l'écosystème. Ce dernier va jouer un rôle à plusieurs niveaux.

Les travaux de Gillooly lient la température du milieu avec les temps de développement et de maturité (Gillooly (2000) ; Gillooly et al. (2001) ; Gillooly et al. (2002)). La maturité arrive plus rapidement dans un milieu chaud et les temps de développement y sont plus rapides. De ce fait, nous allons relier une variable environnementale, la température, au modèle de dynamique de population : les taux de mortalité et âge de recrutement sont liés à la température suivant une exponentielle décroissante (plus de détails dans la partie II du Chapitre 2 pour le zooplancton et la partie I du Chapitre 3 pour le micronecton, où sont entre autres discutées les paramétrisations de ces processus).

Néanmoins, dans les expériences de Gillooly, la température du milieu est constante tout au long de la vie des organismes ou du moins représentative de leur histoire de vie. Dans le cas d'un modèle eulérien, nous n'avons pas d'information individuelle sur l'histoire de vie des organismes comme c'est le cas dans un modèle lagrangien. On fait donc l'hypothèse suivante :

Hypothèse 1 (Température) *La température instantanée du milieu est proche de la température moyenne que les organismes marins ont expérimentée dans leur vie.*

Ainsi λ et τ_r varient avec la température de l'eau au pas de temps considéré.

La production (p) est modélisée à l'aide d'une équation d'avection-diffusion-réaction mais considérons ici uniquement les processus temporels :

$$\begin{cases} \frac{dp}{dt} = -\frac{\partial p}{\partial \tau} \\ p(\tau = 0) = E^{TL-1} \times PP \end{cases} \quad (1.3)$$

où E est le coefficient de transfert d'énergie, TL le niveau trophique et PP la production primaire nette. La condition initiale de ce problème de Cauchy est issue des théories de l'équation 1.1 (section II.C). Là aussi le modèle de dynamique de population est lié à l'écosystème, avec une production initiale directement liée à la production primaire nette. De plus, la résolution numérique de l'équation de production impose une discrétisation en âge.

B ... et spatialisé

Centrons nous maintenant sur les variations spatiales et développons les termes $\frac{db}{dt}$ et $\frac{dp}{dt}$:

$$\begin{cases} \frac{\partial b}{\partial t} = D \left(\frac{\partial^2 b}{\partial x^2} + \frac{\partial^2 b}{\partial y^2} + \frac{\partial^2 b}{\partial z^2} \right) - \frac{\partial}{\partial x}(ub) - \frac{\partial}{\partial y}(vb) - \frac{\partial}{\partial z}(wb) - \lambda(T)b + p(\tau_r(T)) \\ \frac{\partial p}{\partial t} = D \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \right) - \frac{\partial}{\partial x}(up) - \frac{\partial}{\partial y}(vp) - \frac{\partial}{\partial z}(wp) - \frac{\partial p}{\partial \tau} \end{cases}$$

Diffusion
Advection
Dynamique de population
Maturation

(1.4)

où $(x, y, z) \in \mathcal{D}$ est la position dans le domaine $\mathcal{D}$, D le coefficient de diffusion et (u, v, w) les vitesses totales de déplacement des organismes.

Ce système d'équation est également valable pour un groupe restreint d'organismes marins : considérons un groupe ayant un comportement de migration verticale homogène (par exemple effectuant une migration modérée ou restant à la surface jour et nuit). Comme ce groupe est défini par son comportement nous le nommerons groupe fonctionnel. Notons b_i la concentration volumique en biomasse du groupe fonctionnel i et p_i sa production. On a donc :

$$\begin{cases} \frac{\partial b_i}{\partial t} = D \left(\frac{\partial^2 b_i}{\partial x^2} + \frac{\partial^2 b_i}{\partial y^2} + \frac{\partial^2 b_i}{\partial z^2} \right) - \frac{\partial}{\partial x}(ub_i) - \frac{\partial}{\partial y}(vb_i) - \frac{\partial}{\partial z}(wb_i) - \lambda(T)b_i + p_i(\tau_r(T)) \\ \frac{\partial p_i}{\partial t} = D \left(\frac{\partial^2 p_i}{\partial x^2} + \frac{\partial^2 p_i}{\partial y^2} + \frac{\partial^2 p_i}{\partial z^2} \right) - \frac{\partial}{\partial x}(up_i) - \frac{\partial}{\partial y}(vp_i) - \frac{\partial}{\partial z}(wp_i) - \frac{\partial p_i}{\partial \tau} \end{cases} \quad (1.5)$$

où b_i et p_i sont respectivement la concentration volumique en biomasse et en production de ce groupe fonctionnel. La condition initiale en âge devient :

$$p_i(\tau = 0) = E'_i E^{TL-1} \times PP \quad (1.6)$$

où E'_i est le pourcentage de production alloué à ce groupe fonctionnel. Nous faisons l'hypothèse suivante sur les E'_i :

Hypothèse 2 *Le pourcentage de production alloué à un groupe fonctionnel est constant en temps et en espace.*

La diffusion verticale est très faible, on la néglige donc ($\frac{\partial^2 b_i}{\partial z^2} = 0$ et $\frac{\partial^2 p_i}{\partial z^2} = 0$). Les vitesses correspondent à la somme de la vitesse du courant (u_c, v_c, w_c) et de la vitesse de nage (u_s, v_s, w_s). La vitesse horizontale de nage propre des organismes est très faible ($u_s \ll u_c$ et $v_s \ll v_c$) et non dirigée, nous la modélisons donc comme un terme de diffusion, soit pour l'équation de biomasse :

$$\frac{\partial}{\partial x}(u_s b_i) + \frac{\partial}{\partial y}(v_s b_i) = D_s \left(\frac{\partial^2 b_i}{\partial x^2} + \frac{\partial^2 b_i}{\partial y^2} \right) \quad (1.7)$$

Hypothèse 3 (Diffusion) *Le terme de diffusion horizontale comprend la diffusion physique ainsi que les mouvements aléatoires de petite amplitude des organismes.*

Les vitesses verticales du courant sont très faibles devant les vitesses horizontales. Mais, comme nous nous intéressons à des groupes potentiellement migrants, les vitesses de déplacement vertical ne peuvent pas être considérées comme négligeables pendant ces phases de migration. Néanmoins, nous poserons l'hypothèse suivante afin de simplifier le modèle.

Hypothèse 4 (Migration) *Les migrations sont instantanées.*

Les migrations sont des processus complexes, nous préférons ne pas les modéliser en détail et nous concentrer sur les profondeurs où se trouvent les animaux le jour ou la nuit (Figure 1.13).

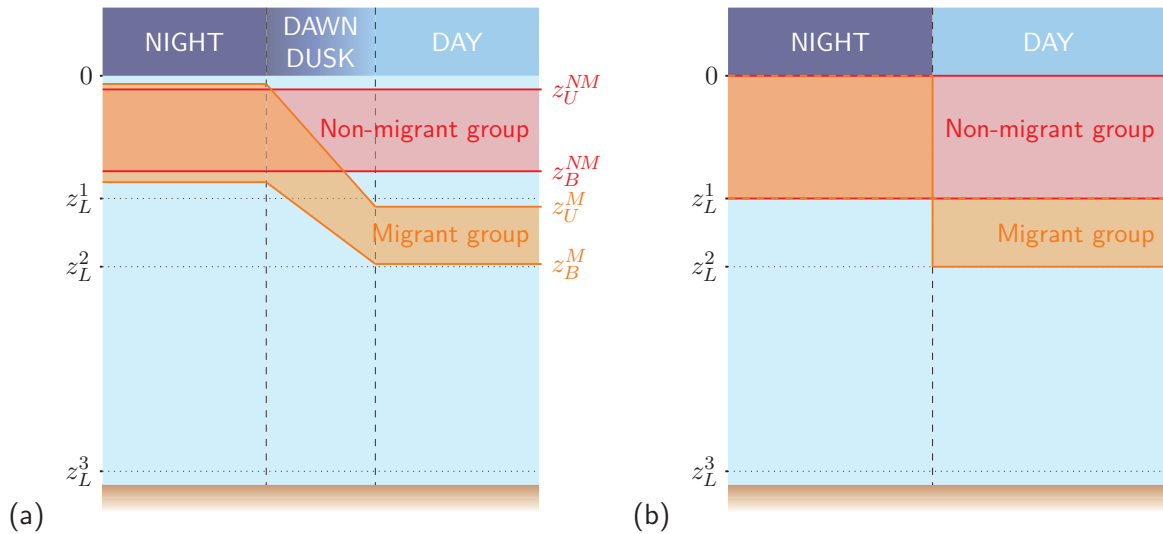


Figure 1.13 – Modélisations des migrations : (a) migrations non-instantanées et (b) migrations instantanées.

Ainsi, on obtient :

$$\begin{cases} \frac{\partial b_i}{\partial t} = D_h \left(\frac{\partial^2 b_i}{\partial x^2} + \frac{\partial^2 b_i}{\partial y^2} \right) - \frac{\partial}{\partial x}(u_c b_i) - \frac{\partial}{\partial y}(v_c b_i) - \frac{\partial}{\partial z}(w_s b_i) - \lambda(T)b_i + p_i(\tau_r(T)) \\ \frac{\partial p_i}{\partial t} = D_h \left(\frac{\partial^2 p_i}{\partial x^2} + \frac{\partial^2 p_i}{\partial y^2} \right) - \frac{\partial}{\partial x}(u_c p_i) - \frac{\partial}{\partial y}(v_c p_i) - \frac{\partial}{\partial z}(w_s p_i) - \frac{\partial p_i}{\partial \tau} \end{cases} \quad (1.8)$$

où D_h est le coefficient de diffusion horizontale comprenant les aspects physiques et la nage. Le but du modèle SEAPODYM est d'être un modèle parcimonieux des écosystèmes marins et donc de représenter la réalité de la façon la plus simple possible tout en restant réaliste. Ceci est d'autant plus important que les

observations et les connaissances sur ces organismes et leurs dynamiques sont très limitées. Nous délaissions donc l'aspect tridimensionnel fin et nous concentrerons sur les variations horizontales. Pour ce faire nous allons intégrer l'équation 1.8 sur la couche qu'occupe le groupe fonctionnel i (soit entre $z_i^{min}(x, y, t)$ et $z_i^{max}(x, y, t)$).

Les observations acoustiques de micronecton présentent trois couches dans la partie supérieure de l'océan. Chacun des groupes fonctionnel occupe une combinaison de couches spécifique. Ils cohabitent dans des couches de même profondeurs.

Pour la cohérence de nos champs de forçage on supposera (Figure 1.13) :

Hypothèse 5 *Les organismes se répartissent intégralement sur trois couches qu'ils occupent en fonction de leur comportement de migration.*

De plus,

Hypothèse 6 (Stabilité des couches) *Les variations des profondeurs occupées par les organismes sont lentes.*

de ce fait, on a par exemple, $\int_{z_i^{min}}^{z_i^{max}} \frac{\partial^2 b_i}{\partial x^2} dz = \frac{\partial^2 B_i}{\partial x^2}$, où on a posé :

$$\begin{cases} \int_{z_i^{min}}^{z_i^{max}} b_i dz = B_i \\ \int_{z_i^{min}}^{z_i^{max}} p_i dz = P_i \end{cases} \quad (1.9)$$

L'équation du modèle devient :

$$\begin{cases} \frac{\partial B_i}{\partial t} = D_h \left(\frac{\partial^2 B_i}{\partial x^2} + \frac{\partial^2 B_i}{\partial y^2} \right) - \int_{z_i^{min}}^{z_i^{max}} \left(\frac{\partial}{\partial x} (u_c b_i) + \frac{\partial}{\partial y} (v_c b_i) \right) dz - [w_s b_i]_{z_i^{min}}^{z_i^{max}} - \lambda(T) B_i + P_i(\tau_r(T)) \\ \frac{\partial P_i}{\partial t} = D_h \left(\frac{\partial^2 P_i}{\partial x^2} + \frac{\partial^2 P_i}{\partial y^2} \right) - \int_{z_i^{min}}^{z_i^{max}} \left(\frac{\partial}{\partial x} (u_c p_i) + \frac{\partial}{\partial y} (v_c p_i) \right) dz - [w_s p_i]_{z_i^{min}}^{z_i^{max}} - \frac{\partial P_i}{\partial \tau} \end{cases} \quad (1.10)$$

Les profils verticaux de proxys de biomasse (acoustiques par exemple), montrent des couches très définies. De ce fait aux limites de celles-ci les valeurs de biomasse sont faibles et donc :

$$[w_s b_i]_{z_i^{min}}^{z_i^{max}} \approx 0 \quad (1.11)$$

Décomposons le terme d'advection, avec $u_c = \bar{u}_c + u'_c$ et $b_i = \bar{b}_i + b'_i$ où $\bar{x}$ est la moyenne sur la verticale et x' la déviation par rapport à la moyenne :

$$\int_{z_i^{min}}^{z_i^{max}} u_c b_i dz = \bar{u}_c B_i + \underbrace{\int_{z_i^{min}}^{z_i^{max}} \bar{u}_c b'_i + u'_c \bar{b}_i dz}_{=0} + \int_{z_i^{min}}^{z_i^{max}} u'_c b'_i dz \quad (1.12)$$

Le terme d'ordre 2 est négligé car u'_c et b'_i sont considérés comme faibles et ainsi on obtient les équations utilisés dans SEAPODYM-LTL et MTL :

$$\frac{\partial B_i}{\partial t} = D_h \left(\frac{\partial^2 B_i}{\partial x^2} + \frac{\partial^2 B_i}{\partial y^2} \right) - \frac{\partial}{\partial x} \bar{u}_c B_i - \frac{\partial}{\partial y} \bar{v}_c B_i - \lambda B_i + P_i \quad (1.13)$$

$$\frac{\partial P_i}{\partial t} = D_h \left(\frac{\partial^2 P_i}{\partial x^2} + \frac{\partial^2 P_i}{\partial y^2} \right) - \frac{\partial}{\partial x} \bar{u}_c P_i - \frac{\partial}{\partial y} \bar{v}_c P_i - \frac{\partial P_i}{\partial \tau} \quad (1.14)$$

C Améliorer la paramérisation du modèle à l'aide de l'assimilation de données

Arhonditsis et al. souligne le nombre important d'étude (91.5%) n'utilisant pas de méthode d'optimisation afin d'améliorer les performances du modèle décrit. Dans le cadre de cette thèse nous utiliserons des méthodes d'assimilation de données. En voici un rapide aperçu.

L'assimilation de données est l'ensemble des méthodes permettant l'amélioration de la connaissance d'un système dynamique par la prise en compte simultanée d'informations théoriques (un modèle) et empiriques (des observations). On présente ici les grandes lignes de l'assimilation de données variationnelle, c'est à dire reposant sur la minimisation d'une fonctionnelle.

On définit $\mathbf{x}$ le vecteur d'état (dans notre cas la biomasse considérée), $\mathcal{M}$ le modèle (le membre de droite de l'équation 1.13) tels que :

$$\begin{cases} \frac{d\mathbf{x}}{dt} = \mathcal{M}(\mathbf{x}, \theta) \\ \mathbf{x}(0) = \mathbf{x}_0 \end{cases} \quad (1.15)$$

où θ est le vecteur des paramètres et $\mathbf{x}_0$ la condition initiale.

Nous allons expliciter dans les chapitres suivants l'opérateur d'observation $\mathcal{H}$ que nous utilisons afin de comparer le vecteur d'observations $\mathbf{y}^{obs}$ et le vecteur d'état $\mathbf{x}$. Cet opérateur projette le vecteur d'état dans l'espace des observations. Il est défini par :

$$\mathcal{H}\mathbf{x} = \mathbf{y}, \quad (1.16)$$

où $\mathbf{y}$ appartient à l'espace des observations et $\mathbf{x}$ à l'espace du vecteur d'état.

Le but de l'assimilation de données variationnelle est de minimiser la distance entre les observations et le vecteur d'état projeté, c'est à dire les prédictions du modèle transformées pour être comparable avec les observations. Cette distance est appelée fonction coût et peut s'écrire :

$$L(\mathbf{x}_0, \theta) = \frac{1}{2} \|\mathcal{H}\mathbf{x}(\mathbf{x}_0, \theta) - \mathbf{y}\| \quad (1.17)$$

où $\|\cdot\|$ est une norme et $\mathbf{y}^{obs}$ le vecteur d'observation.

Dans notre cas, cette méthode est utilisée pour définir un jeu de paramètres optimal, i.e. minimisant $L(\theta)$ à $\mathbf{x}_0$ fixé. On cherche :

$$\theta^a \text{ tel que } L(\theta^a) = \min_{\theta} (L(\theta)) \quad (1.18)$$

Ce qui implique que :

$$\nabla J|_{\theta=\theta^a} = 0 \quad (1.19)$$

Le code adjoint (de façon simpliste la dérivée du code numérique du modèle), permet de calculer à moindre coût les gradients de la fonction coût. La méthode de Quasi-Newton est une méthode de descente de gradient permettant de résoudre une équation du type $f(x) = 0$. Elle est utilisée ici pour résoudre le problème 1.18. Dans la recherche du minimum, la nullité du gradient est une condition nécessaire et non suffisante. Dans le cas d'une fonction coût non-quadratique, il est possible de converger vers un minimum local et non global. Il est donc indispensable de répéter le processus avec des paramétrisations initiales différentes.

D Modéliser le mésozooplancton ou le micronecton à l'aide de SEAPOODYM

Cas du mésozooplancton C'est la configuration la plus simple du modèle SEAPOODYM. On s'intéresse au mésozooplancton, l'ensemble des organismes phytoplanctophages d'une taille comprise entre $200\mu\text{m}$ et 2mm (Sieburth et al. (1978)). Dans ce cas, on considère un seul groupe non-migrant vivant dans la couche épipélagique (Figure 1.14). Le Chapitre 2 de cette thèse décrit en détail ce modèle au travers d'un brouillon d'article. Dans cet article, nous décrivons ce modèle, sa paramétrisation et sa capacité de prédiction. Il est également comparé à un modèle NPZD classique, PISCES. De plus, l'importance de la résolution des données de forçage du modèle est explicitée.

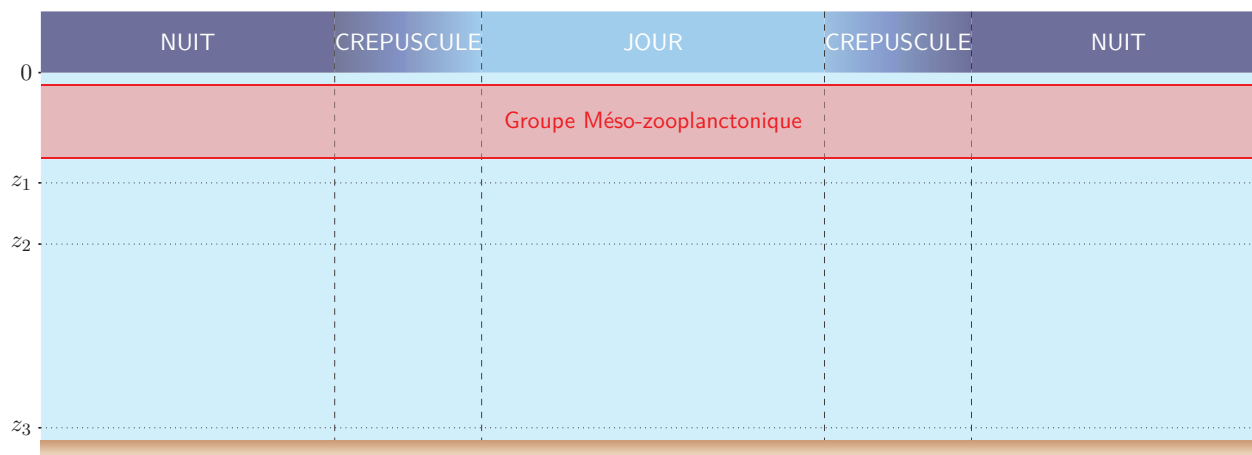


Figure 1.14 – Schéma du groupe fonctionnel de SEAPOODYM-LTL

Cas du micronecton Cette configuration est plus complexe : elle comprend trois couches et six groupes fonctionnels (représentés sur la Figure 1.15). La moitié de ces groupes sont migrants. Ils sont numérotés en fonction des couches qu'ils occupent le jour et la nuit : le groupe $n.m$ occupe la couche n la journée et m la nuit. Les migrations nycthémérales inversées ne sont pas prises en compte du fait de leur faible fréquence. La principale difficulté d'une telle modélisation est la paramétrisation de la distribution d'énergie entre les groupes, soit les paramètres E'_i . Pour obtenir une paramétrisation réaliste de cet ensemble de paramètres, nous avons développé une méthodologie pour assimiler des données acoustiques au modèle (méthode décrite dans le Chapitre 3 au travers de l'article Lehodey et al. (2015)). Néanmoins, avec une base de données acoustiques plus importante, les résultats des expériences d'optimisation des paramètres n'étaient pas satisfaisantes. Deux causes potentielles ont été isolées : la définition des profondeurs des couches du modèle et l'absence d'optimisation des paramètres liés à la températures (paramétrisation des fonctions $\lambda(T)$ et $\tau_r(T)$). Puisque l'optimisation de chacun des paramètres du modèle dépend de la définition des profondeurs de couche, une analyse plus poussée de la définition des couches a été réalisée avec un large jeu de données acoustiques (Chapitre 4).

De nombreuses perspectives s'offrent à ces modèles, notamment à travers la distinction en différentes chaînes trophiques, l'ajout de groupes dans le modèle micronecton (neuston, benthos) et bien sûr l'assimilation d'un grand nombre de données acoustiques avec la nouvelle paramétrisation des profondeurs de couches. Ces perspectives sont décrites dans le Chapitre 5.

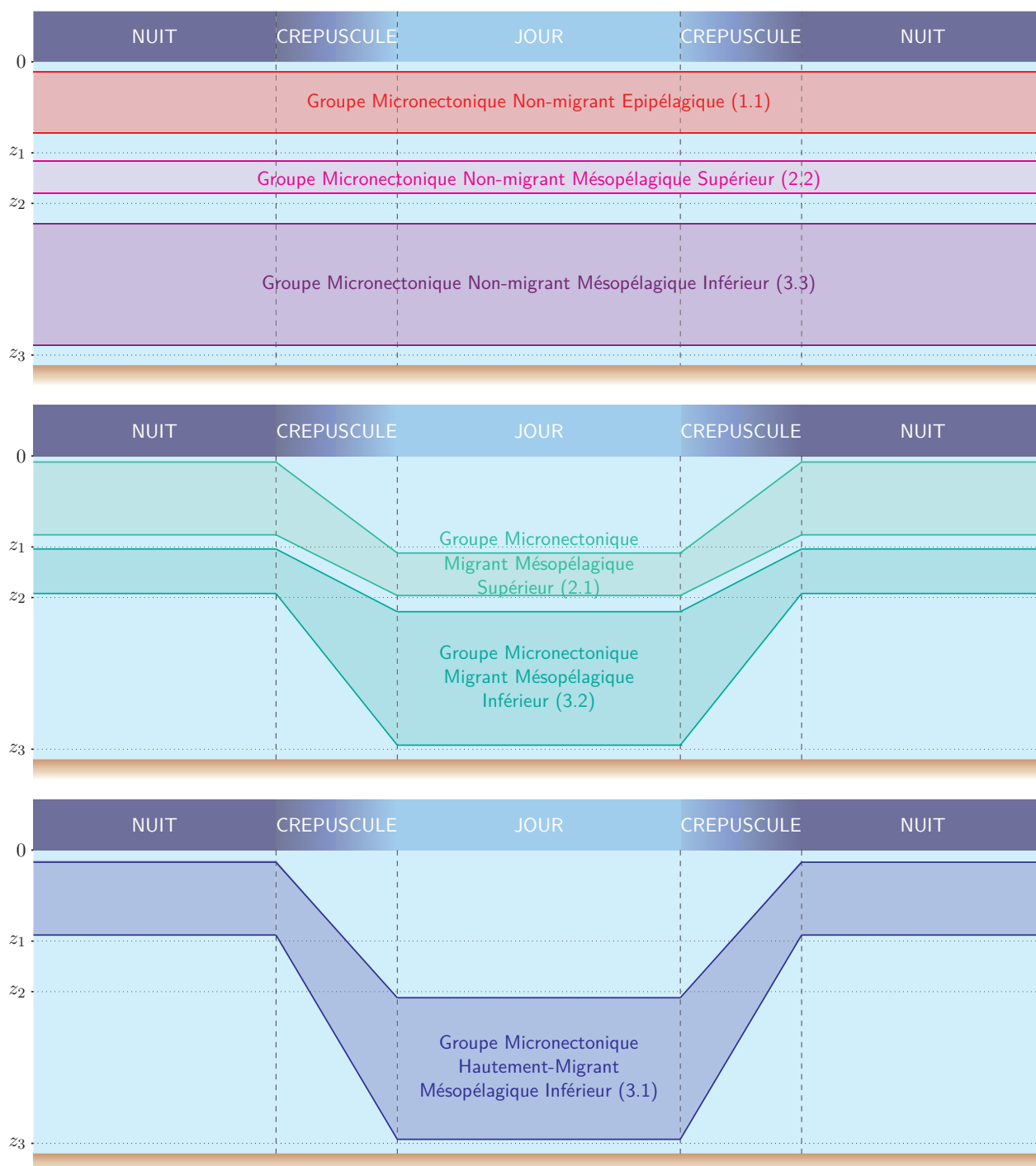


Figure 1.15 – Schéma des groupes fonctionnels de SEAPODYM-MTL définis par leur comportement de migration verticale

Bibliographie

- Aksnes, D. L. and Giske, J. (1990). Habitat profitability in pelagic environments. *Marine ecology progress series. Oldendorf*, 64(3) :209–215.
- Alvain, S., Moulin, C., Dandonneau, Y., and Bréon, F. (2005). Remote sensing of phytoplankton groups in case 1 waters from global seawifs imagery. *Deep Sea Research Part I : Oceanographic Research Papers*, 52(11) :1989 – 2004.
- Alvain, S., Moulin, C., Dandonneau, Y., and Loisel, H. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean : A satellite view. *Global Biogeochemical Cycles*, 22(3). GB3001.
- Arhonditsis, G. B., Brett, M. T., et al. (2004). Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Marine Ecology Progress Series*, 271(2004) :13–26.
- Bailey, S. W. and Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment*, 102(1-2) :12 – 23.
- Barham, E. G. (1966). Deep scattering layer migration and composition : observations from a diving saucer. *Science*, 151(3716) :1399–1403.
- Benoit-Bird, K. J., Au, W. W. L., and Wisdom, D. W. (2009). Nocturnal light and lunar cycle effects on diel migration of micronekton. *Limnology and Oceanography*, 54(5) :1789–1800.
- Blackburn, M. (1968). Micronekton of the eastern tropical pacific ocean : family composition, distribution, abundance, and relations to tuna. *Fish. Bull. US*, 67 :71–115.
- Blackburn, M., Laurs, R., Owen, R., and Zeitzschel, B. (1970). Seasonal and areal changes in standing stocks of phytoplankton, zooplankton and micronekton in the eastern tropical pacific. *Marine Biology*, 7(1) :14–31.
- Block, B. A. (2005). Physiological ecology in the 21st century : advancements in biologging science. *Integrative and Comparative Biology*, 45(2) :305–320.
- Block, B. A., Jonsen, I., Jorgensen, S., Winship, A., Shaffer, S. A., Bograd, S., Hazen, E., Foley, D., Breed, G., Harrison, A.-L., et al. (2011). Tracking apex marine predator movements in a dynamic ocean. *Nature*, 475(7354) :86–90.
- Bopp, L., Monfray, P., Aumont, O., Dufresne, J.-L., Le Treut, H., Madec, G., Terray, L., and Orr, J. C. (2001). Potential impact of climate change on marine export production. *Global Biogeochemical Cycles*, 15(1) :81–99.

- Brodeur, R. and Yamamura, O. (2005). Micronekton of the north pacific. *PICES Scientific Report*, 30 :1–115.
- Christensen, V. and Walters, C. J. (2004). Ecopath with ecosim : methods, capabilities and limitations. *Ecological Modelling*, 172(2-4) :109 – 139. Placing Fisheries in their Ecosystem Context.
- Cisewski, B., Strass, V. H., Rhein, M., and Krägefsky, S. (2010). Seasonal variation of diel vertical migration of zooplankton from adcp backscatter time series data in the lazarev sea, antarctica. *Deep Sea Research Part I : Oceanographic Research Papers*, 57(1) :78 – 94.
- Davison, P. (2011). The specific gravity of mesopelagic fish from the northeastern pacific ocean and its implications for acoustic backscatter. *ICES Journal of Marine Science : Journal du Conseil*, 68(10) :2064–2074.
- Davison, P., Checkley, D., Koslow, J., and Barlow, J. (2013). Carbon export mediated by mesopelagic fishes in the northeast pacific ocean. *Progress in Oceanography*, 116 :14–30.
- Davison, P., Lara-Lopez, A., and Koslow, J. A. (2015). Mesopelagic fish biomass in the southern california current ecosystem. *Deep Sea Research Part II : Topical Studies in Oceanography*, 112 :129–142.
- De Robertis, A. and Handegard, N. O. (2012). Fish avoidance of research vessels and the efficacy of noise-reduced vessels : a review. *ICES Journal of Marine Science : Journal du Conseil*, page fss155.
- Demer, D. A. and Hewitt, R. P. (1995). Bias in acoustic biomass estimates of euphausia superba due to diel vertical migration. *Deep Sea Research Part I : Oceanographic Research Papers*, 42(4) :455–475.
- Domokos, R. (2009). Environmental effects on forage and longline fishery performance for albacore (thunnus alalunga) in the american samoa exclusive economic zone. *Fisheries Oceanography*, 18(6) :419–438.
- Dragon, A.-C., Senina, I., Titaud, O., Calmettes, B., Conchon, A., Arrizabalaga, H., Lehodey, P., and Jacobson, L. (2015). An ecosystem-driven model for spatial dynamics and stock assessment of north atlantic albacore. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(999) :1–15.
- Enright, J. (1977). Diurnal vertical migration : Adaptive significance and timing. part 1. selective advantage : A metabolic model1. *Limnology and Oceanography*, 22(5) :856–872.
- Fortier, M., Fortier, L., Hattori, H., Saito, H., and Legendre, L. (2001). Visual predators and the diel vertical migration of copepods under arctic sea ice during the midnight sun. *Journal of Plankton Research*, 23(11) :1263–1278.
- Gallon, S., Bailleul, F., Charrassin, J.-B., Guinet, C., Bost, C.-A., Handrich, Y., and Hindell, M. (2013). Identifying foraging events in deep diving southern elephant seals, mirounga leonina, using acceleration data loggers. *Deep Sea Research Part II : Topical Studies in Oceanography*, 88-89(0) :14 – 22. Fourth International Symposium on Bio-logging Science.
- Gillooly, J. F. (2000). Effect of body size and temperature on generation time in zooplankton. *Journal of Plankton Research*, 22(2) :241–251.
- Gillooly, J. F., Brown, J. H., West, G. B., Savage, V. M., and Charnov, E. L. (2001). Effects of size and temperature on metabolic rate. *science*, 293(5538) :2248–2251.

- Gillooly, J. F., Charnov, E. L., West, G. B., Savage, V. M., and Brown, J. H. (2002). Effects of size and temperature on developmental time. *Nature*, 417(6884) :70–73.
- Gjøsaeter, J. and Kawaguchi, K. (1980). *A review of the world resources of mesopelagic fish*. Number 193-199. Food & Agriculture Org.
- Guinet, C., Vacquié-Garcia, J., Picard, B., Bessigneul, G., Lebras, Y., Dragon, A.-C., Viviant, M., Arnould, J. P., Bailleul, F., et al. (2014). Southern elephant seal foraging success in relation to temperature and light conditions : insight into prey distribution. *Marine Ecology Progress Series*, 499 :285–301.
- Hammerschlag, N., Gallagher, A., and Lazarre, D. (2011). A review of shark satellite tagging studies. *Journal of Experimental Marine Biology and Ecology*, 398(1–2) :1 – 8.
- Handegard, N. O. and Tjøstheim, D. (2005). When fish meet a trawling vessel : examining the behaviour of gadoids using a free-floating buoy and acoustic split-beam tracking. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(10) :2409–2422.
- Hays, G. (2003). A review of the adaptive significance and ecosystem consequences of zooplankton diel vertical migrations. In Jones, M., Ingólfsson, A., Ólafsson, E., Helgason, G., Gunnarsson, K., and Svavarsson, J., editors, *Migrations and Dispersal of Marine Organisms*, volume 174 of *Developments in Hydrobiology*, pages 163–170. Springer Netherlands.
- Hidaka, K., Kawaguchi, K., Murakami, M., and Takahashi, M. (2001). Downward transport of organic carbon by diel migratory micronekton in the western equatorial pacific : its quantitative and qualitative importance. *Deep Sea Research Part I : Oceanographic Research Papers*, 48(8) :1923 – 1939.
- Hintzen, N., Corten, A., Gerlotto, F., Habasque, J., Bertrand, A., Lehodey, P., Brunel, T., Dragon, A., and Senina, I. (2014). Hydrography and jack mackerel stock in the south pacific-final report. Technical report, IMARES.
- Hjellvik, V., Handegard, N. O., and Ona, E. (2008). Correcting for vessel avoidance in acoustic-abundance estimates for herring. *ICES Journal of Marine Science : Journal du Conseil*, 65(6) :1036–1045.
- Irigoin, X., Klevjer, T., Røstad, A., Martinez, U., Boyra, G., Acuña, J., Bode, A., Echevarria, F., Gonzalez-Gordillo, J., Hernandez-Leon, S., et al. (2014). Large mesopelagic fishes biomass and trophic efficiency in the open ocean. *Nature communications*, 5.
- Iverson, R. L. (1990). Control of marine fish production. *Limnology and Oceanography*, 35(7) :1593–1604.
- Jech, J. M. and Michaels, W. L. (2006). A multifrequency method to classify and evaluate fisheries acoustics data. *Canadian Journal of Fisheries and Aquatic Sciences*, 63(10) :2225–2235.
- Jennings, S., Warr, K. J., and Mackinson, S. (2002). Use of size-based production and stable isotope analyses to predict trophic transfer efficiencies and predator-prey body mass ratios in food webs. *Marine Ecology Progress Series*, 240(1).
- Kaartvedt, S., Staby, A., and Aksnes, D. L. (2012). Efficient trawl avoidance by mesopelagic fishes causes large underestimation of their biomass. *Marine Ecology Progress Series*, 456 :1–6.
- Kashkin, N. I. and Parin, N. V. (1983). Quantitative assessment of micronektonic fishes by nonclosing gear (a review). *Biological Oceanography*, 2(2-4) :263–287.

- Kishi, M. J., Kashiwai, M., Ware, D. M., Megrey, B. A., Eslinger, D. L., Werner, F. E., Noguchi-Aita, M., Azumaya, T., Fujii, M., Hashimoto, S., Huang, D., Iizumi, H., Ishida, Y., Kang, S., Kantakov, G. A., Cheol Kim, H., Komatsu, K., Navrotsky, V. V., Smith, S. L., Tadokoro, K., Tsuda, A., Yamamura, O., Yamanaka, Y., Yokouchi, K., Yoshie, N., Zhang, J., Zuenko, Y. I., and Zvalinsky, V. I. (2007). Nemuro - a lower trophic level model for the north pacific marine ecosystem. *Ecological Modelling*, 202(1-2) :12 – 25. Special Issue on {NEMURO} (North Pacific Ecosystem Model for Understanding Regional Oceanography) and NEMURO.FISH (NEMURO for Including Saury and Herring) Modeling of North Pacific Marine Ecosystems Special Issue on {NEMURO} and NEMURO.FISH.
- Kloser, R., Ryan, T., Sakov, P., Williams, A., and Koslow, J. (2002). Species identification in deep water using multiple acoustic frequencies. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(6) :1065–1077.
- Koslow, J. A. (1997). Seamounts and the ecology of deep-sea fisheries : The firm-bodied fishes that feed around seamounts are biologically distinct from their deepwater neighbors—and may be especially vulnerable to overfishing. *American Scientist*, pages 168–176.
- Koubbi, P., Moteki, M., Duhamel, G., Goarant, A., Hulley, P.-A., O'Driscoll, R., Ishimaru, T., Pruvost, P., Tavernier, E., and Hosie, G. (2011). Ecoregionalization of myctophid fish in the indian sector of the southern ocean : Results from generalized dissimilarity models. *Deep Sea Research Part II : Topical Studies in Oceanography*, 58(1-2) :170 – 180. Census of Antarctic Marine Life : Diversity and Change in the Southern Ocean Ecosystems.
- Kremer, P. and Kremer, J. N. (1988). Energetic and behavioral implications of pulsed food availability for zooplankton. *Bulletin of Marine Science*, 43(3) :797–809.
- Lampert, W. (1989). The adaptive significance of diel vertical migration of zooplankton. *Functional Ecology*, pages 21–27.
- Lampert, W., Schmitt, R.-D., and Muck, P. (1988). Vertical migration of freshwater zooplankton : test of some hypotheses predicting a metabolic advantage. *Bulletin of Marine Science*, 43(3) :620–640.
- Lancraft, T. M., Reisenbichler, K. R., Robison, B. H., Hopkins, T. L., and Torres, J. J. (2004). A krill-dominated micronekton and macrozooplankton community in croker passage, antarctica with an estimate of fish predation. *Deep Sea Research Part II : Topical Studies in Oceanography*, 51(17-19) :2247 – 2260. Integrated Ecosystem Studies of Western Antarctic Peninsula Continental Shelf Waters and Related Southern Ocean Regions.
- Lehodey, P., Conchon, A., Senina, I., Domokos, R., Calmettes, B., Jouanno, J., Hernandez, O., and Kloser, R. (2015). Optimization of a micronekton model with acoustic data. *Ices Journal of Marine Science*, 72(5) :1399–1412.
- Lehodey, P., Murtugudde, R., and Senina, I. (2010). Bridging the gap from ocean models to population dynamics of large marine predators : A model of mid-trophic functional groups. *Progress in Oceanography*, 54(1-2) :69 – 84. Special Issue : Parameterisation of Trophic Interactions in Ecosystem Modelling.
- Lehodey, P., Senina, I., and Murtugudde, R. (2008). A spatial ecosystem and populations dynamics model (seapodym)—modeling of tuna and tuna-like populations. *Progress in Oceanography*, 56(4) :304–318.

- Longhurst, A., Bedo, A., Harrison, W., Head, E., and Sameoto, D. (1990). Vertical flux of respiratory carbon by oceanic diel migrant biota. *Deep Sea Research Part A. Oceanographic Research Papers*, 37(4) :685 – 694.
- Longhurst, A. R. and Harrison, W. G. (1988). Vertical nitrogen flux from the oceanic photic zone by diel migrant zooplankton and nekton. *Deep Sea Research Part A. Oceanographic Research Papers*, 35(6) :881 – 889.
- Loots, C., Koubbi, P., and Duhamel, G. (2007). Habitat modelling of electrona antarctica (myctophidae, pisces) in kerguelen by generalized additive models and geographic information systems. *Polar Biology*, 30(8) :951–959.
- Lurton, X. and DeRuiter, S. (2011). Sound radiation of seafloor-mapping echosounders in the water column, in relation to the risks posed to marine mammals. *The International Hydrographic Review*, (6).
- McClain, C. R. (2009). A decade of satellite ocean color observations*. *Annual Review of Marine Science*, 1(1) :19–42. PMID : 21141028.
- McGehee, D., O'Driscoll, R., and Traykovski, L. (1998). Effects of orientation on acoustic scattering from antarctic krill at 120 khz. *Deep Sea Research Part II : Topical Studies in Oceanography*, 45(7) :1273 – 1294.
- McLaren, I. A. (1963). Effects of temperature on growth of zooplankton, and the adaptive value of vertical migration. *Journal of the Fisheries Board of Canada*, 20(3) :685–727.
- McLaren, I. A. (1974). Demographic strategy of vertical migration by a marine copepod. *American Naturalist*, pages 91–102.
- Moriarty, R. and O'Brien, T. (2013). Distribution of mesozooplankton biomass in the global ocean. *Earth System Science Data*, 5(1) :45.
- Nicol, S. (2003). Living krill, zooplankton and experimental investigations : a discourse on the role of krill and their experimental study in marine ecology. *Mar. Fresh. Behav. Physiol.*, 36(4) :191–205.
- Nishikawa, J., Nishida, S., Moku, M., Hidaka, K., and Kawaguchi, K. (2001). Biomass, abundance, and vertical distribution of micronekton and large gelatinous zooplankton in the subarctic pacific and the bering sea during the summer of 1997. *Journal of oceanography*, 57(3) :361–375.
- O'Driscoll, R. L., Gauthier, S., and Devine, J. A. (2009). Acoustic estimates of mesopelagic fish : as clear as day and night ? *ICES Journal of Marine Science : Journal du Conseil*.
- Passarella, K. C. and Hopkins, T. L. (1991). Species composition and food habits of the micronektonic cephalopod assemblage in the eastern gulf of mexico. *Bulletin of Marine Science*, 49(1-2) :638–659.
- Pauly, D., Christensen, V., and Walters, C. (2000). Ecopath, ecosim, and ecospace as tools for evaluating ecosystem impact of fisheries. *ICES Journal of Marine Science : Journal du Conseil*, 57(3) :697–706.
- Perry, R. I. and Neilson, J. D. (1988). Vertical distributions and trophic interactions of age-0 atlantic cod and haddock in mixed and stratified waters of georges bank. *Marine ecology progress series. Oldendorf*, 49(3) :199–214.

- Plaganyi, E. and Butterworth, D. (2004). A critical look at the potential of ecopath with ecosim to assist in practical fisheries management. *African Journal of Marine Science*, 26(1) :261–287.
- Ryther, J. H. et al. (1969). Photosynthesis and fish production in the sea. the production of organic matter and its conversion to higher forms of life vary throughout the world ocean. *American Association for the Advancement of Science. Science*, 166 :72–76.
- Schaefer, M. B. (1965). The potential harvest of the sea. *Transactions of the American Fisheries Society*, 94(2) :123–128.
- Shaviklo, A. R. and Rafipour, F. (2013). Surimi and surimi seafood from whole ungutted myctophid mince. *LWT - Food Science and Technology*, 54(2) :463 – 468.
- Sieburth, J. M., Smetacek, V., and Lenz, J. (1978). Pelagic ecosystem structure : Heterotrophic compartments of the plankton and their relationship to plankton size fractions 1. *Limnology and oceanography*, 23(6) :1256–1263.
- Stanton, T. K., Wiebe, P. H., Chu, D., Benfield, M. C., Scanlon, L., Martin, L., and Eastwood, R. L. (1994). On acoustic estimates of zooplankton biomass. *ICES Journal of Marine Science : Journal du Conseil*, 51(4) :505–512.
- Steinacher, M., Joos, F., Frolicher, T., Bopp, L., Cadule, P., Cocco, V., Doney, S. C., Gehlen, M., Lindsay, K., Moore, J. K., et al. (2010). Projected 21st century decrease in marine productivity : a multi-model analysis. *Biogeosciences*, 7(3).
- Steinberg, D. K., Carlson, C. A., Bates, N. R., Goldthwait, S. A., Madin, L. P., and Michaels, A. F. (2000). Zooplankton vertical migration and the active transport of dissolved organic and inorganic carbon in the sargasso sea. *Deep Sea Research Part I : Oceanographic Research Papers*, 47(1) :137 – 158.
- Steinberg, D. K., Goldthwait, S. A., and Hansell, D. A. (2002). Zooplankton vertical migration and the active transport of dissolved organic and inorganic nitrogen in the sargasso sea. *Deep Sea Research Part I : Oceanographic Research Papers*, 49(8) :1445 – 1461.
- Stock, C. A., Alexander, M. A., Bond, N. A., Brander, K. M., Cheung, W. W., Curchitser, E. N., Delworth, T. L., Dunne, J. P., Griffies, S. M., Haltuch, M. A., Hare, J. A., Hollowed, A. B., Lehodey, P., Levin, S. A., Link, J. S., Rose, K. A., Rykaczewski, R. R., Sarmiento, J. L., Stouffer, R. J., Schwing, F. B., Vecchi, G. A., and Werner, F. E. (2011). On the use of ipcc-class models to assess the impact of climate on living marine resources. *Progress in Oceanography*, 88(1-4) :1 – 27.
- Strömberg, K. P., Smyth, T. J., Allen, J. I., Pitois, S., and O'Brien, T. D. (2009). Estimation of global zooplankton biomass from satellite ocean colour. *Journal of Marine Systems*, 78(1) :18–27.
- Tanimura, A., Hattori, H., Miyamoto, Y., Hoshiai, T., and Fukuchi, M. (2008). Diel changes in vertical distribution of oithona similis (cyclopoida) and oncaea curvata (poecilostomatoida) under sea ice in mid-summer near syowa station, antarctica. *Polar Biology*, 31(5) :561–567.
- Viviant, M., Trites, A., Rosen, D., Monestiez, P., and Guinet, C. (2010). Prey capture attempts can be detected in steller sea lions and other marine predators using accelerometers. *Polar Biology*, 33(5) :713–719.

- Watanabe, H., Moku, M., Kawaguchi, K., Ishimaru, K., and Ohno, A. (1999). Diel vertical migration of myctophid fishes (family myctophidae) in the transitional waters of the western north pacific. *Fisheries Oceanography*, 8(2) :115–127.
- Williams, A. and Koslow, J. A. (1997). Species composition, biomass and vertical distribution of micro-nekton over the mid-slope region off southern tasmania, australia. *Marine Biology*, 130(2) :259–276.
- Zaret, T. M. and Suffern, J. S. (1976). Vertical migration in zooplankton as a predator avoidance mechanism¹. *Limnology and Oceanography*, 21(6) :804–813.

Chapitre 2 Application simple du modèle avec une seule composante : modélisation de la distribution spatiale du mesozooplancton et validation

Sommaire

I	SEAPODYM-LTL : un modèle parcimonieux	30
II	SEAPODYM-LTL : a parsimonious zooplankton dynamic biomass model	32
III	Perspectives	54
A	Paramétrisation optimale	54
B	Forçages biogéochimiques	54

I SEAPODYM-LTL : un modèle parcimonieux

La dynamique spatio-temporelle du zooplancton est depuis plusieurs décennies modélisée principalement par de nombreux modèles dit NPZ (pour Nutiment-Phyotplancton-Zooplancton) puis NPZD (pour Nutiment-Phyotplancton-Zooplancton-Détritus). Ces modèles expriment les flux de nutriments et de carbone entre ces quatre compartiments. La figure 2.1a (extraite de Lima et al. (2002)) décrit les liens entre les différentes composantes d'un modèle NPZD. Ces modèles, dans le cas où ils sont utilisés pour décrire la distribution de zooplancton, incluent au moins 4 variables d'état. Ce nombre de variables augmente avec la complexité du modèle (subdivision des composantes, voir figure 2.1b et 2.1c), ou avec le nombre de cycles biogéochimiques considérés. Leur nombre de paramètres peut monter jusqu'à une centaine.

Ces modèles complexes sont utilisés pour décrire des écosystèmes depuis les variations physico-chimiques jusqu'au deuxième niveau trophique. Le modèle SEAPODYM-LTL est une approche plus simple centrée uniquement sur le zooplancton. Il ne nécessite que 6 paramètres et utilise des champs de courant, température et production primaire en entrée. La figure 2.2 illustre les liens entre les deux variables d'état du modèle (la production et la biomasse de zooplancton) et avec les variables forçant le modèle. Nous développons ici une approche de calibration-validation du modèle SEAPODYM-LTL à plusieurs résolutions spatio-temporelles et avec différents forçages.

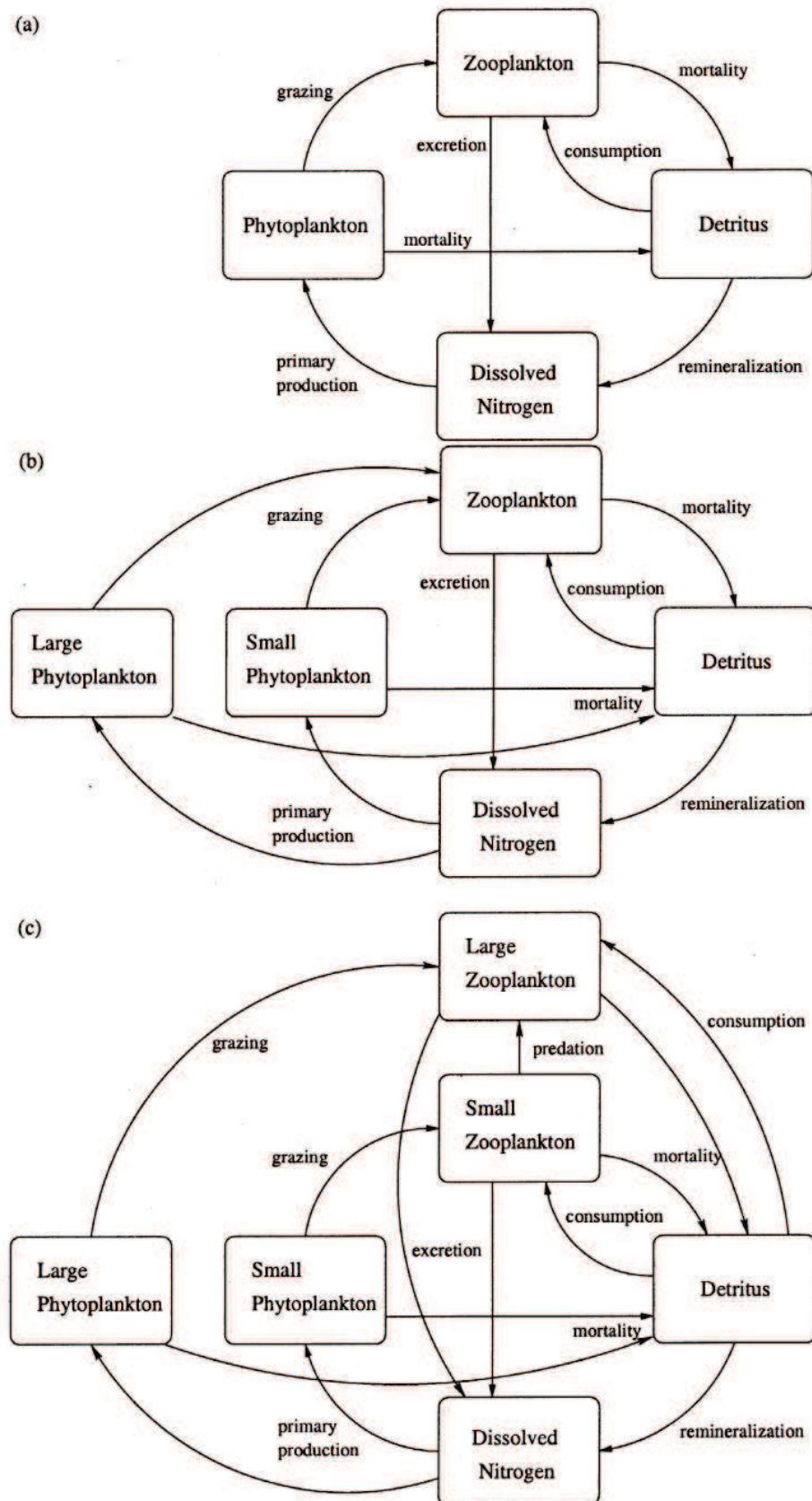


Figure 2.1 – Schéma des relations entre les variables d'états de modèles classiques NPZD. (a) modèle NPZD "classiques" à deux PFTs (phytoplancton et zooplancton), (b) modèle NPZD à 3 PFTs (distinction entre petit et grand phytoplancton) et (c) modèle NPZD à 4 PFTs (distinction entre petit et grand pour le phytoplancton et le zooplancton). (Figure extraite de Lima et al. (2002))

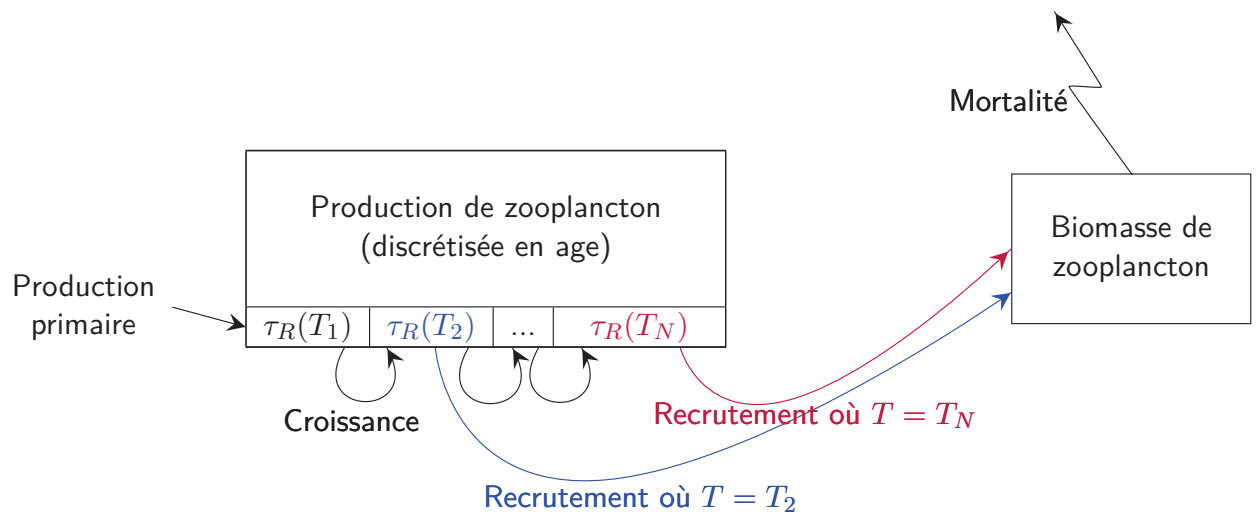


Figure 2.2 – Schéma des relations entre les deux variables d'états et les variables de forçage du modèle SEAPODYM-LTL.

II SEAPODYM-LTL : a parsimonious zooplankton dynamic biomass model

SEAPOODYM-LTL : a parsimonious zooplankton dynamic biomass model

April 21, 2016

Abstract

Mesozooplankton organisms are of critical importance for nutrient cycles in the ocean. With the increasing interest in climate change, zooplankton modelling has become a key issue. The traditional modelling methodology tends to describe the whole biogeochemical and plankton cycle in more and more complex models. We propose here a strictly different approach, much more simple and thus more cost efficient. SEAPOODYM-LTL is a parsimonious zooplankton dynamic biomass model. It use offline biogeochemical and physical inputs to create zooplankton biomass and production.

In this study, we run comparative analyses to assess the importance of resolution and primary inputs on model fit to observations. We also compare SEAPOODYM-LTL to PISCES a much more complex model.

We also illustrate this article with a regional application of the model to the Indonesian Seas at a $1/12^\circ$ x day resolution.

I. Introduction

Mesozooplankton are ubiquitous organisms distributed all over the global ocean. They have been defined by their size ranging from $200\ \mu\text{m}$ to 2 cm (Sieburth et al. (1978)), though the vast majority of mesozooplankton is constituted by crustacean plankton (copepods), typically 1 to 2 mm in size. Mesozooplankton organisms are of critical importance for nutrient cycles in the ocean. Many of them perform diel vertical migration between the productive surface during the night and the dark deeper ocean during the day. As they feed at the surface and excrete fecal pellets deeper in the ocean, they significantly contribute to the active transport of organic matter from surface to the deep ocean (nitrogen: Longhurst and Harrison (1988), Steinberg et al. (2002); carbon Longhurst et al. (1990), Steinberg et al. (2000))).

In addition they are of critical importance in the oceanic food web, both as primary consumers of phytoplankton and prey of larger marine organisms and especially all fish larvae. Therefore, variability of recruitment in many fish stocks has been linked to mesozooplankton biomass fluctuations (Beaugrand et al. (2003), Castonguay et al. (2008)).

Classical models of mesozooplankton biomass use the Nutrient-Phytoplankton-Zooplankton-Detritus (NPZD) approach computing fluxes of nutrient(s) and carbon between these biological compartments. Most of NPZD models have one or two zooplankton (micro- and mesozooplankton) functional group(s). In these models zooplankton interacts with phytoplankton (and if available detritus) through grazing (Gentleman et al. (2003)) and nutrient excretion. These mechanisms and their parameters seem critical for understanding carbon cycling and

lower levels ecosystem dynamics (Fasham (1993)). Zooplankton mortality is also a key issue (Steele and Henderson (1992); Ohman et al. (2004); Edwards and Yool (2000)) and is often used as a closure term, with e.g., a quadratic form linking the mortality rate proportionally to zooplankton biomass, that can strongly influence the dynamics of the models (Edwards and Yool (2000)). After three decades of development and use, NPZD models present a large variety of formulations for phytoplankton response to light, phytoplankton nutrient uptake, zooplankton grazing and mortality rates (Franks (2002), Gentleman et al. (2003), Heinle and Slawig (2013)). They have been widely used either at global (e.g. Bopp et al. (2005), Losa et al. (2006), Le Quéré et al. (2003), Gehlen et al. (2006)) or regional scale (e.g. Jiang et al. (2013), Liu and Chai (2009), Echevin et al. (2012)), as one dimensional idealised water column model (Denman and Pena (1999), Doney et al. (1996)) or three dimensional global ocean simulations at varying spatial resolutions.

Over the years, NPZD models became increasingly complex and evolved to representations of lower trophic level ecosystems. The concept of Plankton Functional Type (PFT) is central to most of these models. The concept of functional types was first developed for terrestrial vegetation models (Smith et al. (2001)) and adapted to marine biogeochemical models by Le Quéré et al. (2005). Their definition of PFT considers four criteria: "(a) the PFT should have an explicit biogeochemical role, (b) the PFT should be defined by a distinct set of physiological, environmental, or nutrient requirements controlling its biomass and productivity, (c) the behaviour of the PFT should have distinct effects on the performance of other PFTs, for instance, through selective depletion of nutrients or grazing, and (d) the PFT should be of quantitative importance in at least some region of the ocean." They propose 10 PFTs from which three are zooplankton: proto-zooplankton, meso-zooplankton and macro-zooplankton. The increase of PFTs goes along with that of state variables, which might correspond to up to 4 times the number of modelled nutrient cycles (Vogt et al. (2010)), as well as functional relationships and

associated parameters.

However, while increasing the number of state variables and parameters, PFT models require more data (and more precise data) to be parameterized leading Anderson (Anderson (2005); Anderson (2006)) to advocate that our knowledge is too immature to build accurate PFT models. Arhonditsis et al. (2004) tested all aquatic biogeochemical models published from 1990–2002 in order to assess their accuracy. Over the 153 models they analysed, only 30 had presented zooplankton predictive performances with a median coefficient of determination $R^2 = 0.24$ and a median relative error reaching 70%. Moreover, zooplankton coefficients of determination were not significantly correlated with phytoplankton coefficients of determination or nutrients coefficients of determination, meaning that the accuracy of the phytoplankton and nutrient predictions cannot be linked to those of zooplankton. Indeed, validation scores of NPZD models are mainly calculated for phytoplankton and nutrients and let zooplankton fields unvalidated.

Arhonditsis et al. demonstrated increasing complexity in NPZD models to be significantly correlated with increasing relative error (Arhonditsis et al. (2004)). Similarly, Friedrichs et al. (2009) tested 12 models of varying complexity with a data assimilation scheme, and demonstrated that "models with multiple zooplankton compartments did not necessarily outperform models with single zooplankton compartment". Therefore, while increasing complexity in models is needed to improve our representation of the real world, there is a need for objective assessment of model outputs and parameterization, with a variety of approaches including empirical ones (Anderson (2005)).

Here we propose one alternative consisting of a parsimonious approach to model mesozooplankton spatial dynamics. Because it is developed in the general SEAPODYM framework including mid-trophic level (MTL) (Lehodey et al. (2010), Lehodey et al. (2015)) and fish population dynamics (Lehodey et al. (2008), Senina et al. (2008)), this mesozooplankton model is called SEAPODYM-LTL

(Lower Trophic Level). In addition to providing an interesting comparison to the classical NPZD approach, the interest of this model is to be dynamically driven off line with physical variables (temperature and currents) from circulation models and primary production either predicted from NPZD models or derived from satellite-based ocean color data, facilitating very high resolution and operational configurations. While its major objective is to produce realistic high resolution distributions of mesozooplankton biomass for developing management applications of marine resources, it can be applied to test the sensitivity of NPZD zooplankton outputs to forcing variables. This model is presented together with its validation. Model results are compared to zooplankton biomass distributions from a state-of-the-art NPZD model generated with the same physical forcing at coarse resolution. Finally, an example of an operational high resolution application is provided for the Indonesian region.

II. SEAPODYM-LTL modelling approach

The modelling of the mesozooplankton functional group is a direct application of the approach followed for simulating the micronekton organisms at the microtrophic level (Lehodey et al. (2010), Lehodey et al. (2015)).

Here we consider a single functional group including all meso-zooplankton organisms. These organisms are defined by their time of development that is obviously strongly correlated to their size. Thus both holo- and meroplankton are considered in this group. The SEAPODYM-LTL model simply requires parameterizing i) the temperature-linked time development relationship defining the recruitment and turnover of the zooplankton population and ii) the energy transfer coefficient from primary production.

There are two state variables modeled in SEAPODYM-LTL: 1) the mesozooplankton production (gWW/m²/day) and biomass (gWW/m²) integrated over the epipelagic vertical layer of the ocean defined by 1.5 times the euphotic depth

(roughly between 100-200 m depth). A conversion factor ($c = 0.1$) is used to convert primary production unit expressed in mmolC into zooplankton grams of wet weight (Cushing et al. (1958)). The detailed model formulation is provided in appendix.

A. Mesozooplankton production (source term)

Mesozooplankton production in the model is the recruitment of new organisms entering in this functional group. The time of development to reach this minimum size is correlated to the ambient water temperature.

We choose to assimilate this recruitment time with the embryonic development time (time from egg to hatch). Gillooly et al. showed that embryonic development time is a function of the temperature (equation (1)).

$$\ln \left(\frac{t_{emb}}{m^{0.25}} \right) = 6.38 - 0.11 \bar{T} \quad (1)$$

where τ_{emb} is the embryonic development time and $\bar{T} = \frac{T}{1 + \frac{T}{273}}$ is the normalized temperature.

In order to express the embryonic development time only as a function of $\bar{T}$, we choose a averaged mass value of $m_0 = 10^{-7}$ g for a (Gillooly and Dodson (2000)). Leading to the following relation (equation (2)) for the recruitment time:

$$\tau_r = 10.5 \exp \left(-0.11 \bar{T} \right) \quad (2)$$

B. Mesozooplankton mortality rate (loss term)

In marine ectotherm organisms, there are clearly strong correlations between temperature and growth, time of development, age at maturation and mortality rate (Landry (1975), McLaren et al. (1989), Huntley and Lopez (1992), Gillooly (2000), Gillooly et al. (2002), Angilletta (2009)). The natural mortality rate (λ) of the functional mesozooplankton group is modelled as for the micronekton (Lehodey et al. (2010)) based on a relationship to the ambient temperature. We consider that this mortality rate

can be defined as the inverse of the time of development until age at maturity, often considered as a measure of the turnover of a population (Lehodey et al. (2010)). Gillooly et al. (2002) have shown that this time of development (scaled by the mass) can be described relatively to ambient temperature by the equations 3 below:

$$\ln \left(\frac{GT}{m^{0.25}} \right) = 7.2 - 0.11 \bar{T} \quad (3)$$

where GT is the generation time (sum of embryonic development time and post-embryonic development time, as defined in Gillooly (2000)).

Using a chosen mass value, m_0 , leads to an expression between GT and $\bar{T}$. Nevertheless, this expression is highly sensitive to this choice. Thus we choose to use Gillooly's dataset published in Gillooly (2000) to express directly GT as a function of $\bar{T}$.

The optimal regression model is :

$$\begin{aligned} \ln(GT) &= 4.48 - 0.11\bar{T} \\ R^2 &= 0.47 \end{aligned} \quad (4)$$

However, the prediction intervals are wide due to the data distribution (see Figure 1). The model uses the exponential of the expression (the function between GT and $\bar{T}$). Thus the predicted generation time at 0°C ranges between 27.9 to 276.9 days. We manually tuned this relationship in order to reduce the distance between zooplankton biomass observations and zooplankton biomasse predicted by the model. The function used in SEAPODYM-LTL, along with the optimal regression and prediction interval are plotted on figure 1.

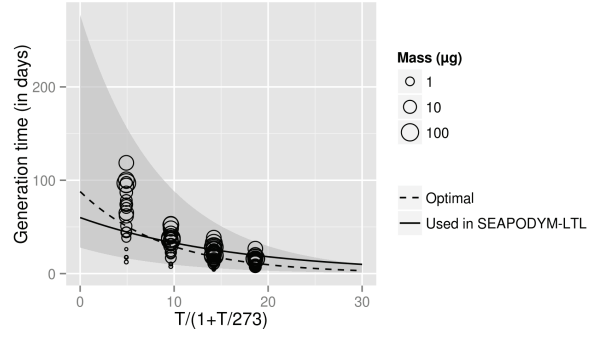


Figure 1: Generation time as a function of temperature. Data were extracted from Gillooly (2000). Shade area is the prediction interval of the linear regression between $\log(GT)$ and $\bar{T}$. Dashed line indicates the optimal regression line and the solid line is the function actually used in SEAPODYM-LTL.

Then, the mortality rate (λ) is parametrized in SEAPODYM-LTL with the relation :

$$\frac{1}{\lambda} = GT = 60 \exp(-0.06 \bar{T}) \quad (5)$$

C. Energy transfer from primary production

The energy transfer from primary production (a single coefficient $E=0.05$) is tuned for the model to fit zooplankton biomass observations available and described below. It is then compared to estimates or theoretical values provided in the literature.

Primary production is integrated over the water column. Simulations are produced in this study using either satellite derived Net Primary Production (NPP) or NPP provided by a NPZD model (cf section below).

III. NPZD model

PISCES (Pelagic Interaction Scheme for Carbon and Ecosystem Studies) is a model of intermediate complexity (Aumont et al. (2015)). The model includes four PFTs: two phytoplankton (nanophytoplankton and diatoms) and two zooplankton (micro- and mesozooplankton). It simulates the biogeochemical cycles of oxygen, carbon and nutrients (N (NO_3^- , NH_4^+), P, Si, Fe) using 24 state variables. Five detrital pools are included : small and large particulate organic carbon (POC), dissolved organic carbon, cal-

cium carbonate and biogenic silica. Within the context of this paper, the discussion of the model is limited to source and sink terms governing mesoplankton biomass. A detailed description of the model used in this study (PISCES-v1) is provided by Aumont and Bopp (2006), respectively in Aumont et al. (2015) for PISCES-v2.

A. Mesozooplankton growth (source)

Mesozooplankton graze on nanophytoplankton, diatoms, microzooplankton, small and large POC. Grazing intensity depends on prey concentration and size. Mesozooplankton graze on the preferred prey type with no switching (Moloney and Field (1991)). In addition to density dependent grazing, PISCES includes a formulation of flux-feeding on large POC. Flux-feeding depends on the flux and hence the concentration times the sinking speed. Growth efficiency is set equal to 0.35. Grazing intensity is a function of temperature with a Q10 of 2.14 (Buitenhuis et al. (2006)).

B. Mesozooplankton mortality (sink)

Sink terms in the equation governing mesozooplankton biomass include respiration and mortality. Both are temperature dependent with a Q10 of 2.14 (Buitenhuis et al. (2006)). Mortality is represented as a quadratic function on concentration and represents grazing by higher trophic levels, not resolved by PISCES.

IV. Zooplankton datasets used in validation

We use several observation datasets: inter-annual focuses on specific regions and a global monthly climatology. We choose to use volumetric data to integrate them on the first SEAPODYM-LTL layer (surface to 1.5 times the euphotic depth). This way, we can compare SEAPODYM-LTL biomass to observations.

A. Interannual datasets

All datasets were cut down to the 1998-2005 time series to match the simulations time period. Figure 2 represents the location and number of observations in the interannual datasets.

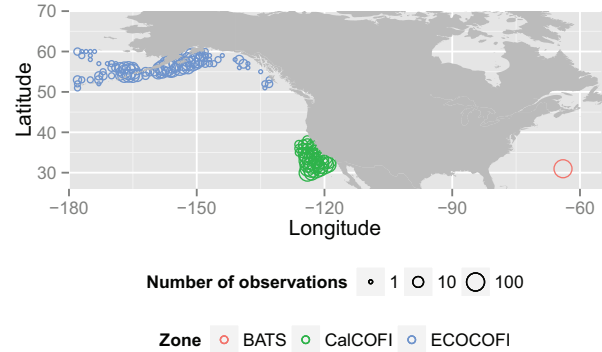


Figure 2: Location of observed data in the interannual datasets grouped in a 1° grid.

Bermuda Atlantic Time Series (BATS) BATS is a compilation of monthly cruises sampling multiple physical, chemical and biological properties of the site (31°40'N 64°10'W). The on-going zooplankton biomass monthly time series began in April, 1994. Biomasses are provided in volumetric concentration of weight masses. We converted it in the equivalent carbon mass ($\text{mgC} \cdot \text{mg}^{-1} \text{WW} = 0.12$, Cushing et al. (1958)).

CALifornia Cooperative Oceanic Fisheries Investigations (CalCOFI) dataset The CalCoFI dataset is a time-series of oceanographic and fisheries data sampling the coast of the US and Mexican California since 1949. The zooplankton biomass time series began in 1951. Zooplankton biomass are provided in displacement volume by volume. We convert them into equivalent carbon mass ($\text{mgC} \cdot \text{mL}^{-1} \text{DV} = 96$, Cushing et al. (1958)).

ECOsystems and Fisheries-Oceanography Cooperative Investigation (EcoFOCI) dataset The EcoFOCI dataset is a compilation of oceanographic and fisheries data sampling the Gulf of Alaska and the north of the US west coast. Since 1979, zooplankton biomass is sampled along the coast of Alaska. It is provided in displacement volume by

volume and converted into equivalent carbon mass ($\text{mgC.mL}^{-1}\text{DV} = 96$, Cushing et al. (1958)).

B. Monthly climatology

This monthly climatology was gathered by Moriarty and O'Brien. It compiles more than 150,000 observations in all oceanic basins. Data were converted into equivalent carbon mass with non-linear relations (Moriarty and O'Brien (2013)). Data were also converted to the equivalent mass that would have been sampled by a $333 \mu\text{m}$ mesh net. Figure 3 represents the location and number of observations in the monthly climatology. This climatology is use globally but we also extracted a regional dataset in the Indonesian Seas (90°E - 144°E - 20°S - 25°N , black rectangle in figure 3).

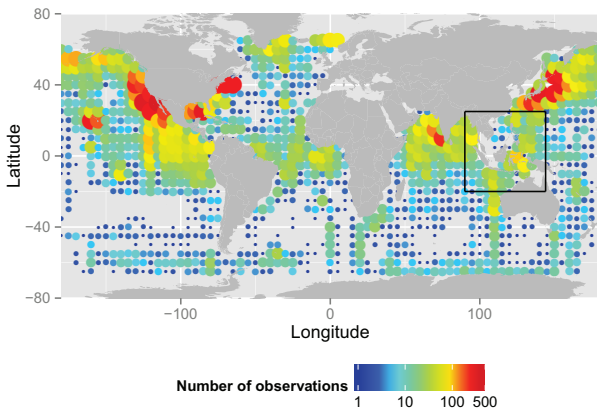


Figure 3: Location of observations in the monthly climatology grouped in a 2.5° grid.

V. Simulations of zooplankton biomass

In order to assess the skills of SEAPODYM-LTL, we ran multiple simulations with multiple forcing sets. It allows us to extract the importance of primary production, resolution and physical forcing sets. We also compare SEAPODYM to PISCES with the same forcing set (primary production, temperature and currents). We present here the physical and biogeochemical forcing sets used. The section C. sums up the simulations that are used in this paper.

A. Inputs : physical forcing sets

IPSL-CM5A-LR This simulation is based on simulations with the Earth System Model (ESM) developed at the Institut Pierre Simon Laplace (IPSL) for the CMIP5: IPSL- CM5A-LR (Séférián et al. (2014)). It is based on the ocean circulation model NEMO model (Madec (2008)). The simulation was interpolated over a $1^\circ \times 1^\circ$ regular grid to fit SEAPODYM-LTL numerical requirement. Hereafter this physical forcing is named NEMO- 1° in zooplankton simulation names.

Mercator - Glorys2 We use the ocean reanalysis GLORYS.v2 (Parent et al. (2013)) and the operational version of this ocean circulation model for the recent years (<http://www.mercator-ocean.fr/en/>), also based on the ocean circulation model NEMO model. Their resolution is $1/4^\circ \times$ week. We also used this dataset with a degraded resolution of $1^\circ \times$ week. Hereafter noted MERC- $1/4^\circ$ or MERC- 1° in the simulation names.

INDO12 INDO12 is used to drive the regional configuration of SEAPODYM-LTL (see appendix B.). INDO12 is $1/12^\circ$ regional version of the NEMO physical ocean model (Tranchant et al. (2015)). This model has a $1/12^\circ \times$ day resolution. It runs each week and provide real time fields. The open boundary conditions are derived from the operational version of Mercator model at a $1/4^\circ \times$ week resolution (see previous paragraph). Hereafter noted INDO12.

B. Inputs : biogeochemical forcing sets

Vertically Generalized Production Model (VGPM) The VGPM predicts net primary production as a function of chlorophyll, photosynthetically available radiation and the photosynthetic efficiency (Behrenfeld and Falkowski (1997)). The photosynthetic efficiency is a function of the temperature. It is obtained from the Ocean Productivity web site of the University of Oregon State (<http://www.science.oregonstate.edu/ocean.productivity/>). Hereafter called VGPM PP.

Eppley The Eppley version of the VGPM differs with the original in the description of the photosynthetic efficiency. It is based on an exponential function of the temperature (Eppley (1972), Morel (1991)). It is also obtained from the Ocean Productivity web site of the University of Oregon State <http://www.science.oregonstate.edu/ocean.productivity/>). Hereafter called Eppley PP.

PISCES This primary production field is the PP output, ie., the sum of small and large phytoplankton production, of the IPSL-CM5A-LR simulations (S  f  rian et al. (2014)). Hereafter called PISCES PP.

C. Simulations

Three comparative analyses are conducted (Table 8):

- CA1** The sensitivity of our zooplankton model to the type of primary production used as an input forcing is tested with the same physical forcing dataset (NEMO-1  ) using 3 different PP input fields: VGPM and EPPLEY satellite derived PP and PISCES PP.
- CA2** The differences between the zooplankton biomass distributions at global scale simulated by PISCES and SEAPODYM-LTL models are explored with simulation outputs based on the same physical forcing dataset (NEMO-1  ) and the PISCES generated primary production for SEAPODYM-LTL.
- CA3** The impact of the spatial resolution is investigated by comparison of SEAPODYM-LTL simulations using the same primary production input (VGPM) and physical forcing (temperature and currents) at resolution $1/4^\circ \times$ week and then interpolated at $1^\circ \times$ week. A higher resolution ($1/12^\circ \times$ day) is also used for comparison in the Indonesian region. It is issued from the regional configuration of SEAPODYM-LTL described in appendix.

D. Metrics

In all three CA, we compare the mean volumic concentration of biomass computed over the SEAPODYM layer, i.e., 1.5 the euphotic depth.

As observations, we used biomass fished standardized by the volume of water filtered by the net and assume that it is representative of SEAPODYM first layer. SEAPODYM predicts the integrated biomass over a layer and thus predictions are divided by the depth of this layer. As for PISCES, volumic concentration of biomass are averaged over SEAPODYM layer. In CA 2 and 3, the earth mask is not the same between configuration. The dataset is reduced to positions where observations and all predictions are available.

We use four metrics to characterise the fit between observations and model predictions (formula of each metric is presented in appendix B.) :

- the correlation (cor) between predicted and observed samples
- the ratio between predicted sample variance and observed sample variance (Var)
- the goodness of fit (GF) which compares the model and the mean of observations in terms of predictability
- the Root Mean Square Error ($RMSE$)

Exp.	Simulation name	Zooplankton model	Forcings	Resolution	Area
CA1	S-LTL_VGPM	SEAPODYM-LTL	NEMO/VGPM	1° x month	Global
	S-LTL_Eppley	SEAPODYM-LTL	NEMO/Eppley	1° x month	Global
	S-LTL_PISCES	SEAPODYM-LTL	NEMO/PISCES	1° x month	Global
CA2	S-LTL	SEAPODYM-LTL	NEMO/PISCES	1° x month	Global
	PISCES	PISCES	NEMO	1° x month	Global
CA3	MERC-1°	SEAPODYM-LTL	Mercator/VGPM	1° x week	Global
	MERC-1/4°	SEAPODYM-LTL	Mercator/VGPM	1/4° x week	Global
	INDO12	SEAPODYM-LTL	INDO12/VGPM	1/12° x day	Indonesian Seas

Table 1: Comparative analyses configurations

VI. Results

A. CA1 - Comparative analysis with different PP

The comparison between predicted and observed climatological biomass distributions of zooplankton (Figure 5) shows similar global patterns with higher biomass in the cold systems in temperate and subpolar regions and in the equatorial region. Lower biomass occur in the subtropical gyres. Predicted fields from the three simulations have lower variances than observations.

Estimated zooplankton biomass distributions differ between simulations. The results obtained with PISCES Primary Production (Figure 5b) give lower than observed zooplankton biomass in equatorial and tropical areas and in upwelling systems (California, Somalia), but higher biomass in the North Atlantic (latitude above 40°N). Satellite-derived primary productions produce higher zooplankton biomass in the tropics and coastal ecosystem regions. The simulation using VGPM generates the highest contrast between cold and warm areas (Figure 5c). The one base on EPPLEY (Figure 5d) seems closer to observations. This first visual inspection is confirmed by statistical metrics (Table 2).

RMSE values are all similar between simulations and close to the observed standard deviation (11.2 mgWW/m³). The correlation between observations and predictions is stronger when the primary

production is derived from satellite data (0.25 and 0.26) rather than based on modelled primary production (0.15). The range of variation is much wider when using VGPM as input compared to the two other simulations. Nevertheless the ratio between predicted and observed variance is below 1 for all three simulations. They have also a negative goodness of fit, the highest (close to 0) being obtained for the simulation using Eppley primary production.

Simulation name	<i>cor</i>	<i>Var</i>	<i>GF</i>	<i>RMSE</i>
S-LTL_VGPM	0.25	62%	-0.23	12
S-LTL_Eppley	0.26	26%	-0.01	11
S-LTL_PISCES	0.15	22%	-0.16	12

Table 2: Summary of statistical metrics of CA1 configurations : correlation (*cor*), percentage of variance (*Var*), goodnes of fit (*GF*) and error (*RMSE*) of each configurations.

Results based on regional time series of observations are presented with a Taylor diagram. As with climatological global dataset, the use of satellite derived primary production provide better correlations, (all remain below 0.5 however) and higher standard deviations closer to 1, i.e. the variability and dispersion of observations. The comparison between outputs produced with the two satellite derived products indicates slightly better correlations for the EPPLEY-based simulation and higher normalized standard deviations for VGPM-based simulation.

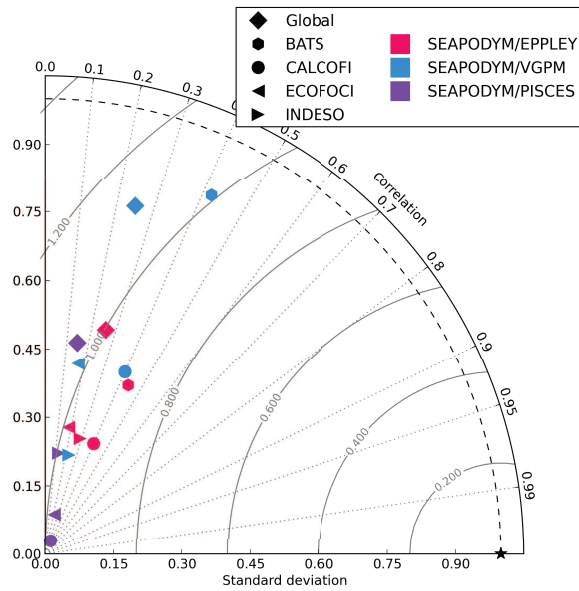
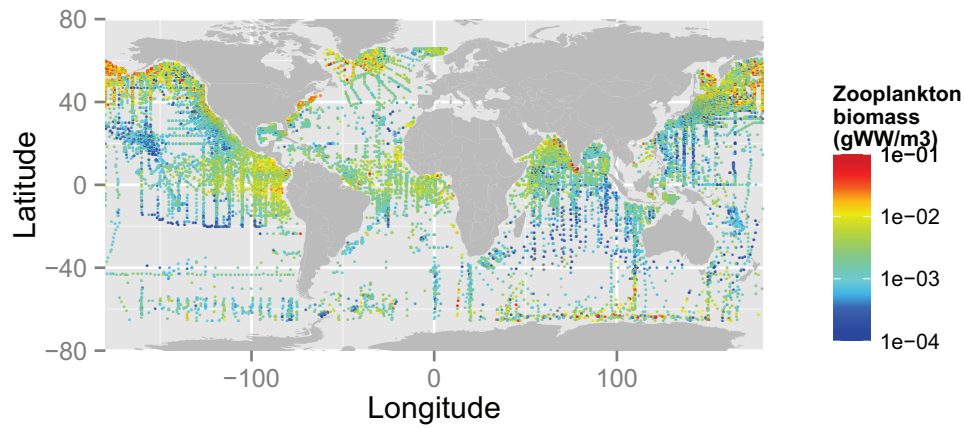
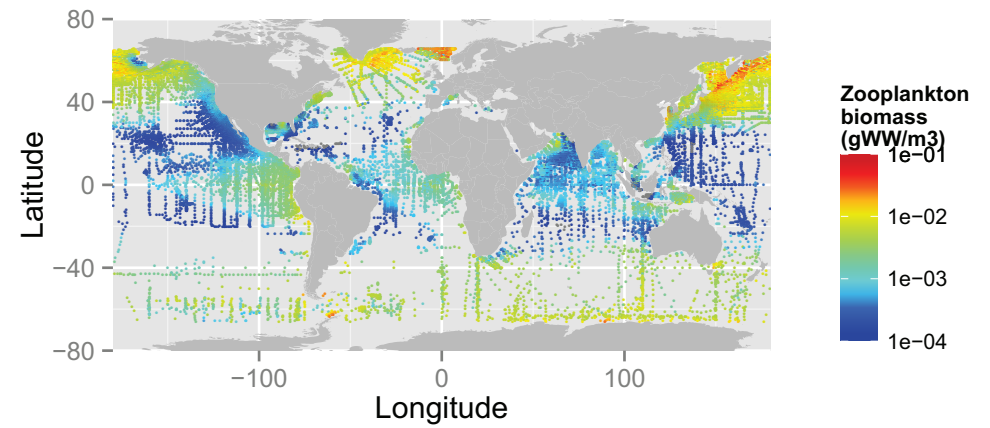


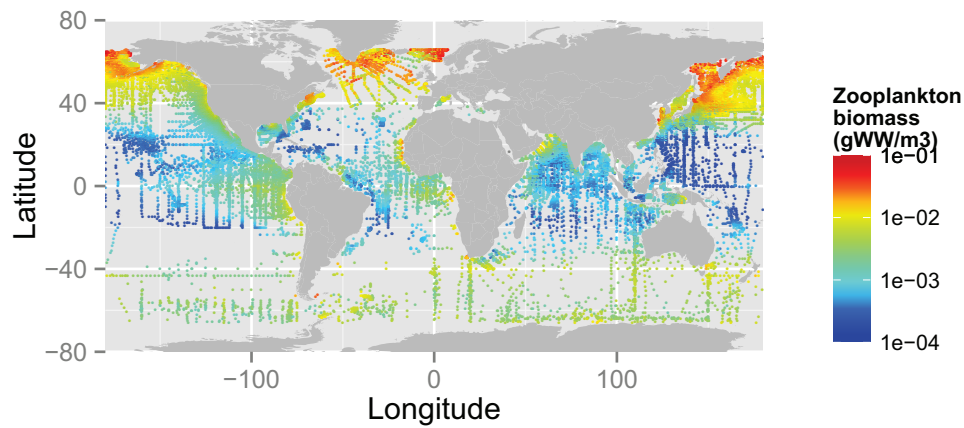
Figure 4: CA1 - The standard deviation is presented on the radial axis. A value of 1 indicates that model predictions have the same variability and dispersion than observations. The Pearson correlation coefficient is presented on the angular axis from 0 to 1. Values close to 1 indicate good covariation between data and model outputs. The normalized error is presented as concentric circles, whose center is the value 1 of the x axis (black star). The closer to the center the smaller the difference between predicted and observed values.



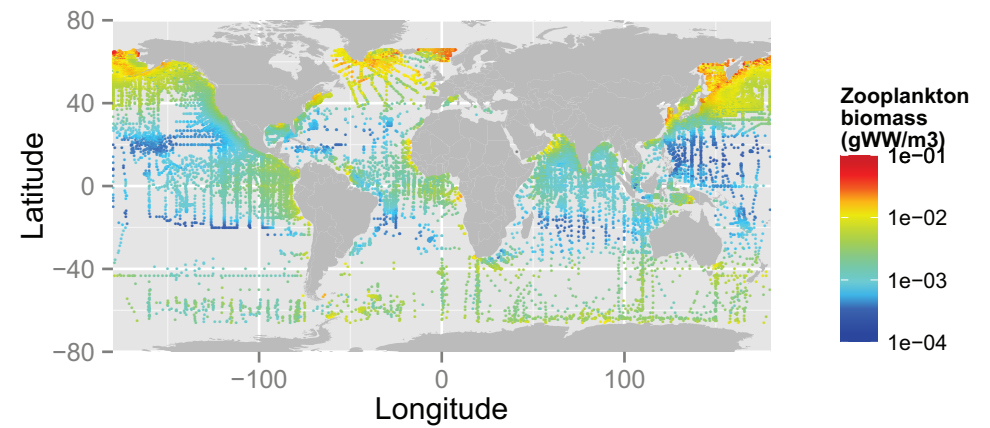
(a) Observations



(b) S-LTL_NEMO-1°_PISCES



(c) S-LTL_NEMO-1°_VGPM



(d) S-LTL_NEMO-1°_Eppley

Figure 5: CA1 - Spatial distribution of zooplankton biomass observed and predicted at observation positions : (a) observations and (b) to (d) predicted values issued from CA1. All data are aggregated regardless to their time of acquisition. Colors of scatter points represents the biomass. The scale is logarithmic and the same in all four pannels.

B. CA2 - Comparative analysis with different models

The zooplankton biomass distributions predicted by the two models SEAPODYM-LTL and PISCES using the same forcing (currents, temperature and PISCES primary production) are compared on figure 7. At global scale, predicted spatial distributions are in agreement with observed distributions but showing less variability. Both models underestimate biomass by one order of magnitude for PISCES and by a factor 0.5 to 1 (depending on the area) for SEAPODYM-LTL. SEAPODYM-LTL simulates some highly productive features (Japan sea, North Pacific) but none of the model simulations reproduce upwellings.

In both simulations, $RMSE$ values (Table 3) are similar to observed standard deviation (4.9 ± 8.5 mgWW/m³). Other metrics are also very close for both models (Table 3) with however, a higher value of Var for the zooplankton predicted with SEAPODYM-LTL. The ratio between predicted and observed variance is 0.37, three times lower than the variance of observations, but 0.02 for PISCES, i.e., 50 times lower than observations. The metrics computed for regional datasets (Figure 6) confirm these trends but with an interesting exception concerning the BATS time series, for which PISCES obtains a very good correlation (0.5) greater than SEAPODYM-LTL (0.41).

Simulation name	cor	Var	GF	$RMSE$
S-LTL	0.21	37%	-0.13	9
PISCES	0.23	2%	-0.14	9

Table 3: Summary of statistical metrics of CA2 configurations : correlation (cor), percentage of variance (Var), goodnes of fit (GF) and error ($RMSE$) of each configurations.

Taylor diagram of these configurations is presented on figure 6. For each testing dataset, correlations are similar for both models, at the exception of BATS for PISCES which exhibit a really strong correlation (0.5) greater than SEAPODYM (0.41). For every dataset, normalised standard deviation are higher for SEAPODYM.

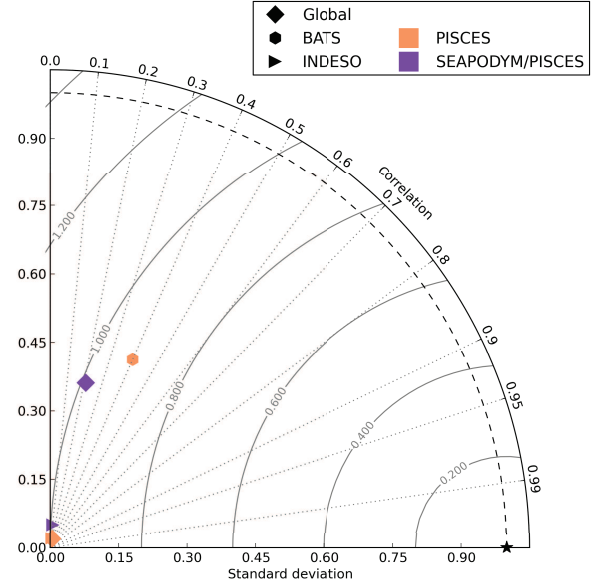
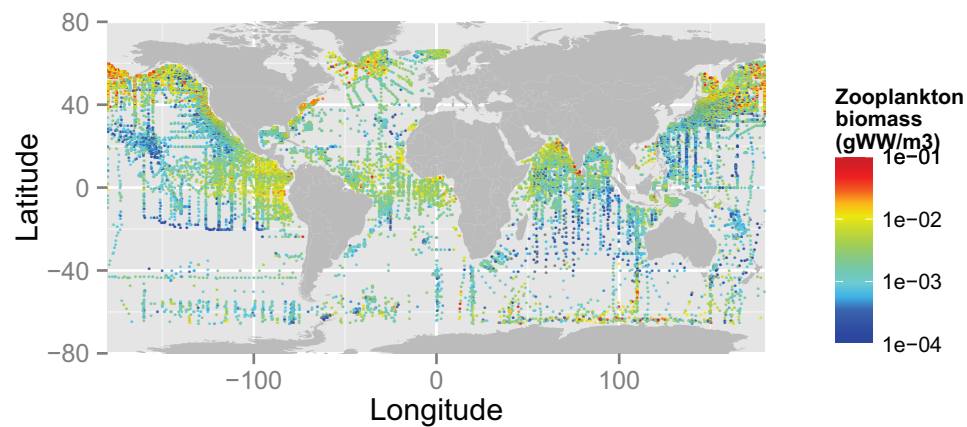
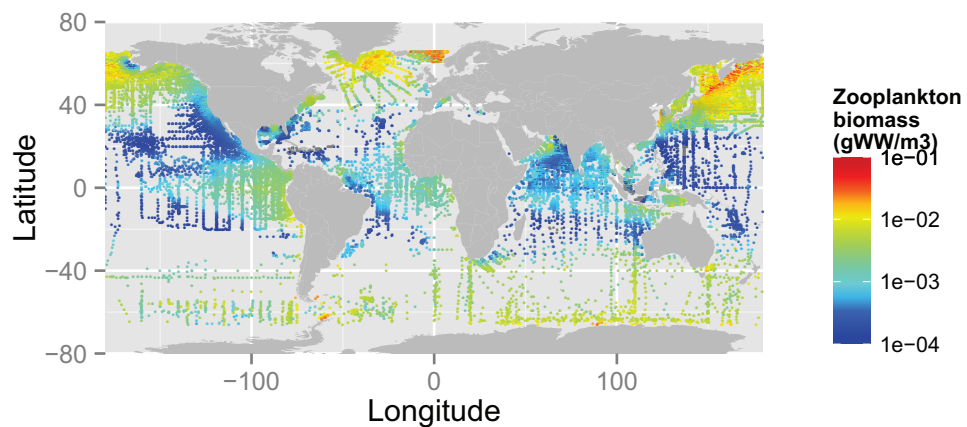


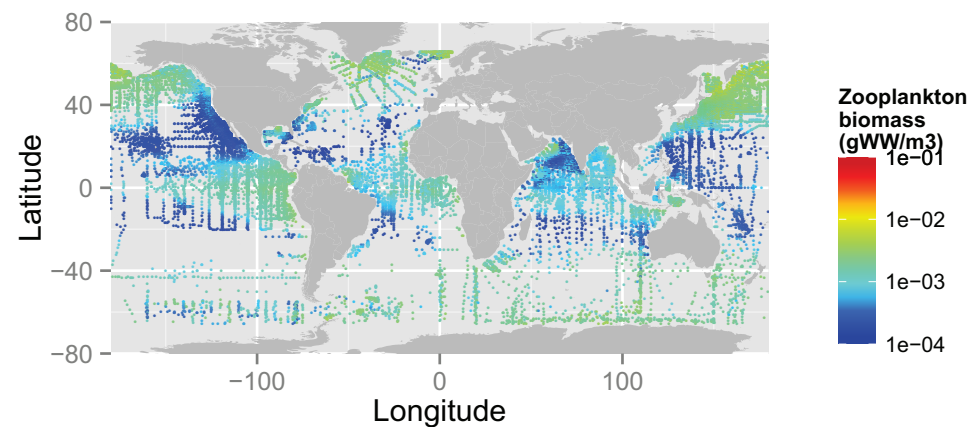
Figure 6: CA2 - The standard deviation is presented on the radial axe. A value of 1 indicates that model predictions have the same variability and dispersion than observations. The Pearson correlation coefficient is presented on the angular axe from 0 to 1. Values close to 1 indicate good covariation between data and model outputs. The normalized error is presented as concentric circles, whose center is the value 1 of the x axis (black star). The closer to the center the smaller the difference between predicted and observed values.



(a) Observations



(b) S-LTL_NEMO-1°_PISCES



(c) PISCES_NEMO-1°

Figure 7: CA2 - Spatial distribution of zooplankton biomass observed and predicted at observation positions : (a) observations and (b) to (d) predicted values issued from CA2. All data are aggregated regardless to their time of acquisition. Colors of scatter points represents the biomass. The scale is logarithmic and the same in all three pannels.

C. CA3 - Comparative analysis with different resolutions

The impact of resolution on predicted zooplankton biomass distribution is investigated using the same primary production input (VGPM) and the physical fields from the GLORYS reanalysis at the original resolution of $1/4^\circ \times \text{week}$ and then interpolated at $1^\circ \times \text{week}$ (Figure 9). The simulation at $1/4^\circ \times \text{week}$ gives similar overall distribution than the simulation using the $1^\circ \times \text{week}$ resolution $RMSE$ is bigger to the observed standard deviation ($6.9 \pm 10.9 \text{ mgWW/m}^3$). GF increases with resolution whereas Var is stable at 130%.

Simulation name	<i>cor</i>	<i>Var</i>	<i>GF</i>	<i>RMSE</i>
glo_MERC- 1°	0.36	130%	-0.66	14
glo_MERC- $1/4^\circ$	0.39	130%	-0.61	14
Indo_MERC- 1°	0.36	49%	-0.03	4.4
Indo_MERC- $1/4^\circ$	0.33	44%	-0.05	4.5
Indo_IND012	0.34	52%	-0.08	4.6

Table 4: Summary of statistical metrics of CA3 configurations : correlation (*cor*), percentage of variance (*Var*), goodnes of fit (*GF*) and error (*RMSE*) of each configurations.

The correlation and the goodness of fit slightly increases with the higher resolution. The Taylor diagram (Figure 8) for regional datasets does not show a general trend due to resolution but a grouping by dataset (Figure 8).

Spatial distributions of predicted values and observations for the month of June are displayed on figure 9. As for the other comparative analyses, predicted distributions are smoother than the observed one. Predicted global distributions are very similar between both configurations. When looking at a regional scale like Indonesian Seas (figure 10), one could notice a really good agreement between predictions through the three resolutions.

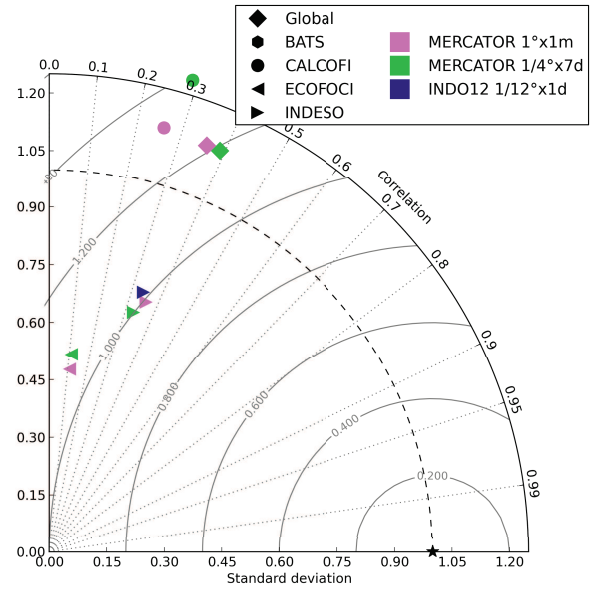
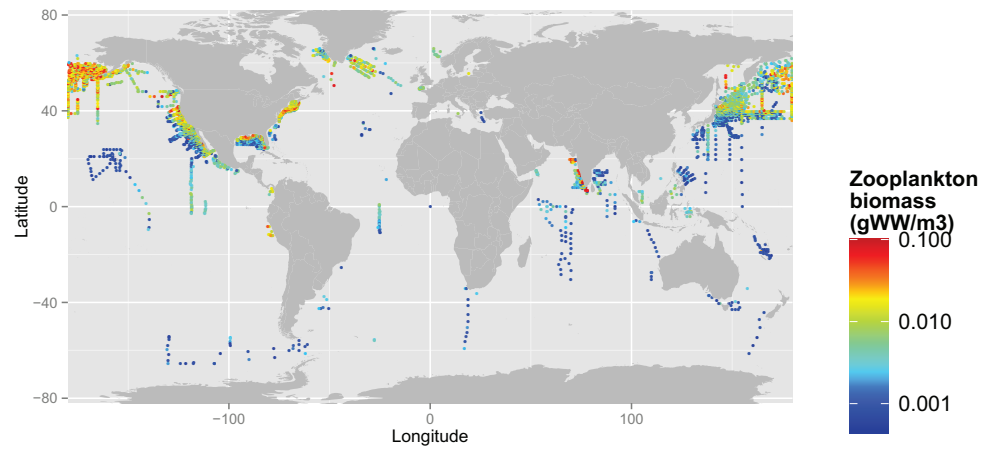
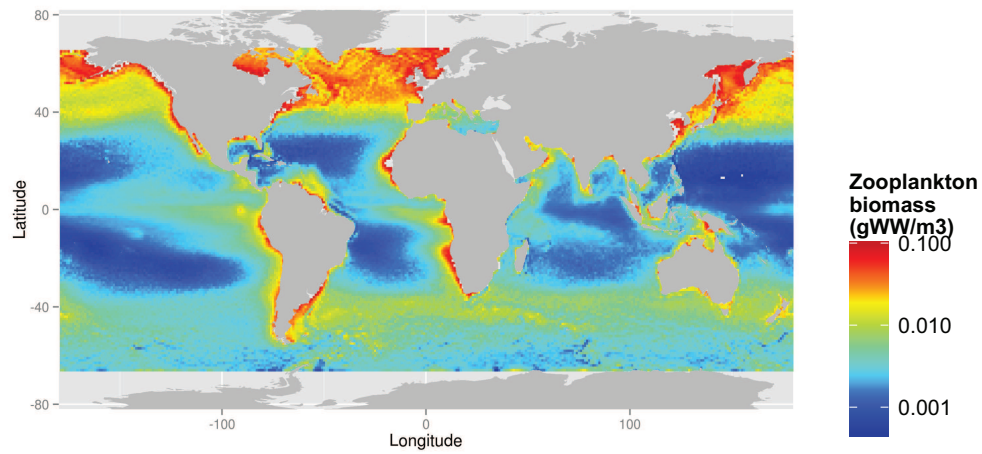


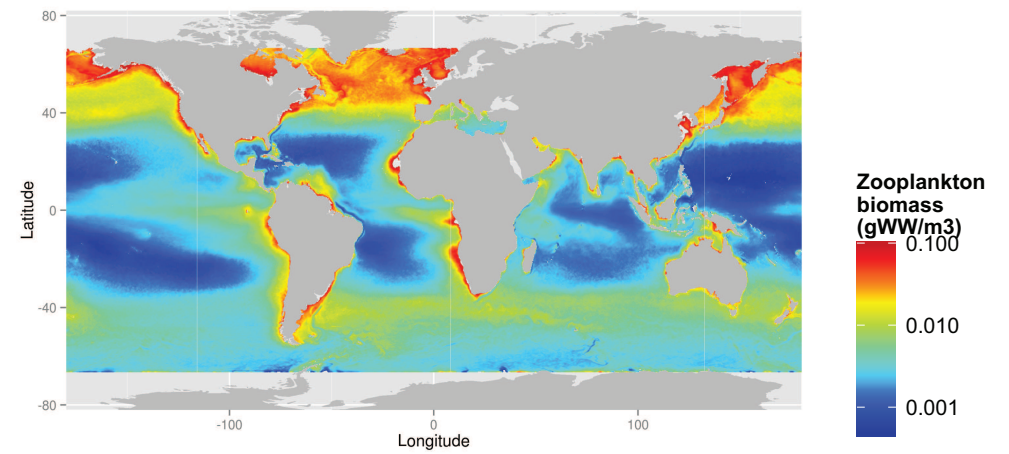
Figure 8: CA3 - The standard deviation is presented on the radial axe. A value of 1 indicates that model predictions have the same variability and dispersion than observations. The Pearson correlation coefficient is presented on the angular axe from 0 to 1. Values close to 1 indicate good covariation between data and model outputs. The normalized error is presented as concentric circles, whose center is the value 1 of the x axis (black star). The closer to the center the smaller the difference between predicted and observed values.



(a) Observations



(b) S-LTL_MERC-1°_VGPM



(c) S-LTL_MERC-1/4°_VGPM

Figure 9: CA3 - Spatial distribution of zooplankton biomass observed and predicted field at global scale for the month of June: (a) observations and (b) to (d) predicted values issued from CA3. Colors of scatter points (a) and background (b) to (d) represents the biomass. The scale is logarithmic and the same in all three panels.

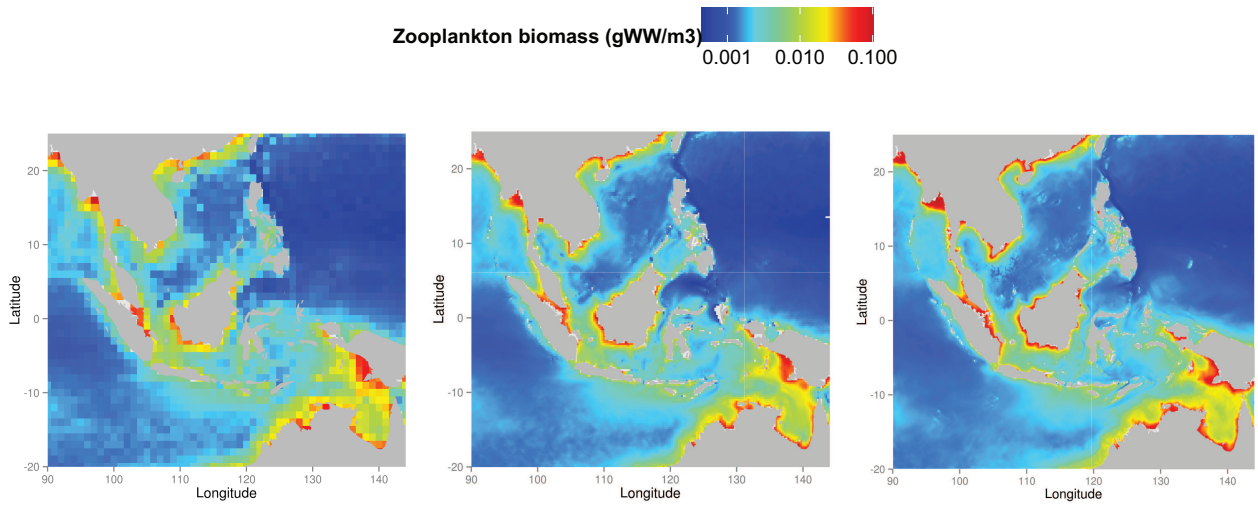


Figure 10: CA3 - Spatial distribution of zooplankton predicted at the regional scale for the month of June from CA3. Background color represents the biomass. The scale is logarithmic and the same in all three pannels.

VII. Discussion

SEAPODYM-LTL is a parsimonious model developed to simulate realistic zooplankton biomass from operational ocean circulation model (temperature and currents) and primary production derived from satellite ocean color data. It is also an interesting approach to compare and analyze the outputs with similar predictions from more complex biogeochemical models (e.g., PISCES) describing the full carbon cycle. This is all the more easier than SEAPODYM-LTL is coupled offline to its forcing variable (temperature, currents and primary production) and has a low computational cost with only two advection-diffusion-reaction equations to solve over a 2D grid. Since the adjoint-code is available, an optimization approach can be easily implemented to estimate the parameters from observations as it has been done already for other applications (Senina et al. (2008), Lehodey et al. (2015)). This is an essential step that will be rapidly developed. Since there are only a few parameters, the analysis should be greatly facilitated.

The model includes only a few parameters. First, one single energy transfer coefficient from primary production to zooplankton production was estimated by tuning its value to obtain a range of biomass

similar to the one observed in the ECOFOCI region. Its value was finally fixed to 0.05. This is within the range of values estimated by other approaches (Jennings et al. (2002b)). However, there is a trade-off between this parameter and the parameters that define the mortality of the zooplankton. These parameters are linked to the generation time (GT).

Currently, the first parameterization of the generation time (GT) is estimated from a linear regression using published data on the time of development of zooplankton species, measured in laboratory experiments (Gillooly et al. (2002)). The model of time of development proposed from these experiments includes the individual mass of zooplankton that is not a state variable of SEAPODYM-LTL, and thus cannot be used in its relation between GT and temperature. Therefore, rather than replacing this variable by a mean value, we re-estimated the linear regression that fit these data expressing directly GT as a function of temperature. There is also a similar relationship to parameterize between temperature and time of development until age of recruitment in the mesozooplankton functional group. This life stage is quite short however, and the outputs of the model, at least at the global scale investigated here, are insensitive to the parameter values used for this relationship.

While the skills of the models are measured from all available observations, there are various sources of uncertainty inherent to the different techniques used to collect zooplankton biomass data. These measures rely on trawling sampling by research vessels with different net sizes and sometimes mesh sizes. They are localized in space and time and several methods are used to quantify biomass, e.g., based on volume displacement, wet weight, dry weight, or ash weight. Various conversion factors have to be used to standardize the units for comparison with model outputs. Even with a careful approach as for the global climatological dataset prepared by Moriarty and O'Brien (2013) and based on more than 50 years of sampling, the dataset cover partially the global ocean, and unavoidably still includes various errors. In addition to this (partial) global climatology, three datasets provide information on interannual time scale in three different ecosystems: cold (ECOFOCI), temperate (BATS), and coastal upwelling (CalCOFI). They use common trawl mesh (200 μ m), thus reducing the sources of errors in their standardization. For comparison with standardized observations of zooplankton biomass, their equivalent mean volumic concentrations are extracted from the modelled biomass. However, even at the highest resolution used in this study, natural patchiness of zooplankton biomass cannot be reproduced.

Another source of uncertainty when fitting predictions and observations comes from the sensitivity of model outputs to their forcing variables that are temperature, currents and primary production. The first comparative analysis (CA1) gives a good illustration of the sensitivity to the primary production fields used. The model fit to observations is better when using primary production derived from satellite rather than predicted from PISCES. In each region and globally, at least one of the satellite inputs perform better than the model PP field. This was expected however, especially since PISCES cannot simulate coastal upwelling systems at the coarse resolution used (1° to 2°) while the higher productivity in these regions (e.g. CalCOFI) is captured by satellite ocean color data. Nevertheless, at global scale and coarse resolution,

the fit obtained using PISCES primary productivity is reasonably good. Interestingly, depending of the region the best score change between the simulations using either the VGPM or the Eppley satellite derived primary production, suggesting that an average of both products (or even several others) could be more appropriate. However, all primary production satellite products are known to be affected by significant uncertainties, particularly in coastal waters (Saba et al. (2011); Friedrichs et al. (2009)).

Whatever the product used for primary production, the SEAPODYM-LTL outputs show a constant bias with an overestimation of predicted zooplankton biomass in the arctic and subarctic waters, that is cold waters. This is a key result suggesting that parameters or the form of the functional relationship between temperature and time of development might be not well defined. The negative exponential form proposed as a theoretical relationship between temperature and time of development does not account for specific adaptation to very cold waters. It has been shown that enzymatic activity at low temperature is more efficient in ectotherms than in birds and marine mammals (Low et al. (1973)). A similar gradient exist within ectotherm species of warm and cold environment. For instance some fishes from the Notothenoid family inhabiting extremely cold water ($< 0^{\circ}\text{C}$) have higher Michaelis–Menten constant and catalysis rates than species from the same family in temperate waters (Fields and Somero (1998)). Therefore, the lack of data for cold water species can strongly influence the form of the relationship. And sigmoid rather than exponential function would likely provide a better fit to zooplankton data.

The second comparative analysis (CA2) uses the same physical and biogeochemical forcing to compare the zooplankton biomass distributions predicted by PISCES and SEAPODYM-LTL. In that case, both models perform equally in (Goodness of Fit and Pearson correlation) for the global dataset. PISCES get even the best score for the BATS time series, which indeed has been used to tune the model parameterization. However zooplankton

biomass predicted by PISCES has a much lower variance.

Finally, the increase in resolution (CA3) from $1^\circ \times$ week to $1/4^\circ$, using the same satellite derived primary production (VGPM), has a positive impact on the fit with the global climatology but not on the regional interannual datasets. The distribution of zooplankton biomass predicted by the operational global $1/4^\circ$ and regional $1/12^\circ$ models in the Indonesian region are very close, with their results intimately linked to the dynamics of primary production since the time of development is fast in these very warm waters. A similar comparison in a subarctic environment would likely provide more contrasted results.

While further developments are required to revise the form of the functional relationship between temperature and time of development, and to optimize the parameters, this study shows that a simple parsimonious approach can be proposed to model zooplankton biomass at global scale. It provides a complementary useful approach to NPZD models to investigate key mechanisms for the modeling of zooplankton. Finally it should provide useful operational products driven by satellite and operational ocean circulation models for marine resource applications.

References

- Anderson, T. R. (2005). Plankton functional type modelling: running before we can walk? *Journal of Plankton Research*, 27(11):1073–1081.
- Anderson, T. R. (2006). Confronting complexity: reply to le quere and flynn. *Journal of Plankton Research*, 28(9):877–878.
- Angilletta, M. J. (2009). *Thermal adaptation: a theoretical and empirical synthesis*. Oxford University Press.
- Arhonditsis, G. B., Brett, M. T., et al. (2004). Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Marine Ecology Progress Series*, 271(2004):13–26.
- Aumont, O. and Bopp, L. (2006). Globalizing results from ocean in situ iron fertilization studies. *Global Biogeochemical Cycles*, 20(2).
- Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., and Gehlen, M. (2015). PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. *Geoscientific Model Development Discussions*, 8(2):1375–1509.
- Beaugrand, G., Brander, K. M., Lindley, J. A., Souissi, S., and Reid, P. C. (2003). Plankton effect on cod recruitment in the north sea. *Nature*, 426(6967):661–664.
- Behrenfeld, M. J. and Falkowski, P. G. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and oceanography*, 42(1):1–20.
- Blayo, E. and Debreu, L. (2005). Revisiting open boundary conditions from the point of view of characteristic variables. *Ocean modelling*, 9(3):231–252.
- Bopp, L., Aumont, O., Cadule, P., Alvain, S., and Gehlen, M. (2005). Response of diatoms distribution to global warming and potential implications: A global model study. *Geophysical Research Letters*, 32(19).
- Buitenhuis, E., Le Quere, C., Aumont, O., Beaugrand, G., Bunker, A., Hirst, A., Ikeda, T., O'Brien, T., Piontkovski, S., and Straile, D. (2006). Biogeochemical fluxes through mesozooplankton. *Global Biogeochemical Cycles*, 20(2).
- Castonguay, M., Plourde, S., Robert, D., Runge, J. A., and Fortier, L. (2008). Copepod production drives recruitment in a marine fish. *Canadian Journal of Fisheries and Aquatic Sciences*, 65(8):1528–1531.
- Cushing, D., Humphrey, G., Banse, K., and Laevastu, T. (1958). Report of the committee on terms and equivalents. *Rapp. P.-V. Reun. Cons. Int. Explor. Mer*, 144:15–16.

- Denman, K. and Pena, M. (1999). A coupled 1-d biological/physical model of the northeast subarctic pacific ocean with iron limitation. *Deep Sea Research Part II: Topical Studies in Oceanography*, 46(11):2877–2908.
- Doney, S. C., Glover, D. M., and Najjar, R. G. (1996). A new coupled, one-dimensional biological-physical model for the upper ocean: Applications to the jgofs bermuda atlantic time-series study (bats) site. *Deep Sea Research Part II: Topical Studies in Oceanography*, 43(2):591–624.
- Echevin, V., Goubanova, K., Belmadani, A., and Dewitte, B. (2012). Sensitivity of the humboldt current system to global warming: a downscaling experiment of the ipsl-cm4 model. *Climate dynamics*, 38(3-4):761–774.
- Edwards, A. M. and Yool, A. (2000). The role of higher predation in plankton population models. *Journal of Plankton Research*, 22(6):1085–1112.
- Eppley, R. W. (1972). Temperature and phytoplankton growth in the sea. *Fish. Bull*, 70(4):1063–1085.
- Fasham, M. J. (1993). Modelling the marine biota. In *The global carbon cycle*, pages 457–504. Springer.
- Fields, P. A. and Somero, G. N. (1998). Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase a4 orthologs of antarctic notothenioid fishes. *Proceedings of the National Academy of Sciences*, 95(19):11476–11481.
- Franks, P. J. (2002). Npz models of plankton dynamics: their construction, coupling to physics, and application. *Journal of Oceanography*, 58(2):379–387.
- Friedrichs, M. A., Carr, M.-E., Barber, R. T., Scardi, M., Antoine, D., Armstrong, R. A., Asanuma, I., Behrenfeld, M. J., Buitenhuis, E. T., Chai, F., et al. (2009). Assessing the uncertainties of model estimates of primary productivity in the tropical pacific ocean. *Journal of Marine Systems*, 76(1):113–133.
- Gehlen, M., Bopp, L., Emprin, N., Aumont, O., Heinze, C., and Ragueneau, O. (2006). Reconciling surface ocean productivity, export fluxes and sediment composition in a global biogeochemical ocean model. *Biogeosciences*, 3(4):521–537.
- Gentleman, W., Leising, A., Frost, B., Strom, S., and Murray, J. (2003). Functional responses for zooplankton feeding on multiple resources: a review of assumptions and biological dynamics. *Deep Sea Research Part II: Topical Studies in Oceanography*, 50(22):2847–2875.
- Gillooly, J. F. (2000). Effect of body size and temperature on generation time in zooplankton. *Journal of Plankton Research*, 22(2):241–251.
- Gillooly, J. F., Charnov, E. L., West, G. B., Savage, V. M., and Brown, J. H. (2002). Effects of size and temperature on developmental time. *Nature*, 417(6884):70–73.
- Gillooly, J. F. and Dodson, S. I. (2000). The relationship of neonate mass and incubation temperature to embryonic development time in a range of animal taxa. *Journal of Zoology*, 251(3):369–375.
- Heinle, A. and Slawig, T. (2013). Internal dynamics of npzd type ecosystem models. *Ecological Modelling*, 254:33–42.
- Huntley, M. E. and Lopez, M. D. (1992). Temperature-dependent production of marine copepods: a global synthesis. *American Naturalist*, pages 201–242.
- Jennings, S., Pinnegar, J. K., Polunin, N. V., and Warr, K. J. (2002a). Linking size-based and trophic analyses of benthic community structure. *Marine Ecology Progress Series*, 226:77–85.
- Jennings, S., Warr, K. J., and Mackinson, S. (2002b). Use of size-based production and stable isotope analyses to predict trophic transfer efficiencies and predator-prey body mass ratios in food webs. *Marine Ecology Progress Series*, 240:11–20.
- Jiang, M., Barbeau, K. A., Selph, K. E., Measures, C. I., Buck, K. N., Azam, F., Mitchell, B. G.,

- and Zhou, M. (2013). The role of organic ligands in iron cycling and primary productivity in the antarctic peninsula: A modeling study. *Deep Sea Research Part II: Topical Studies in Oceanography*, 90:112–133.
- Landry, M. R. (1975). The relationship between temperature and the development of life stages of the marine copepod *acartia clausi* giesbr. *Limnology and Oceanography*, 20(5):854–857.
- Le Quéré, C., Aumont, O., Monfray, P., and Orr, J. (2003). Propagation of climatic events on ocean stratification, marine biology, and CO_2 : Case studies over the 1979–1999 period. *Journal of Geophysical Research: Oceans* (1978–2012), 108(C12).
- Le Quéré, C., Harrison, S. P., Colin Prentice, I., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Cotrim Da Cunha, L., Geider, R., Giraud, X., et al. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, 11(11):2016–2040.
- Lehodey, P., Conchon, A., Senina, I., Domokos, R., Calmettes, B., Jouanno, J., Hernandez, O., and Kloser, R. (2015). Optimization of a micronekton model with acoustic data. *ICES Journal of Marine Science: Journal du Conseil*, 72(5):1399–1412.
- Lehodey, P., Murtugudde, R., and Senina, I. (2010). Bridging the gap from ocean models to population dynamics of large marine predators: a model of mid-trophic functional groups. *Progress in Oceanography*, 84(1):69–84.
- Lehodey, P., Senina, I., and Murtugudde, R. (2008). A spatial ecosystem and populations dynamics model (seapodum)—modeling of tuna and tuna-like populations. *Progress in Oceanography*, 78(4):304–318.
- Liu, G. and Chai, F. (2009). Seasonal and interannual variability of primary and export production in the south china sea: a three-dimensional physical–biogeochemical model study. *ICES Journal of Marine Science: Journal du Conseil*, 66(2):420–431.
- Longhurst, A., Bedo, A., Harrison, W., Head, E., and Sameoto, D. (1990). Vertical flux of respiratory carbon by oceanic diel migrant biota. *Deep Sea Research Part A. Oceanographic Research Papers*, 37(4):685 – 694.
- Longhurst, A. R. and Harrison, W. G. (1988). Vertical nitrogen flux from the oceanic photic zone by diel migrant zooplankton and nekton. *Deep Sea Research Part A. Oceanographic Research Papers*, 35(6):881 – 889.
- Losa, S. N., Vézina, A., Wright, D., Lu, Y., Thompson, K., and Dowd, M. (2006). 3d ecosystem modelling in the north atlantic: Relative impacts of physical and biological parameterizations. *Journal of marine systems*, 61(3):230–245.
- Low, P. S., Bada, J. L., and Somero, G. N. (1973). Temperature adaptation of enzymes: roles of the free energy, the enthalpy, and the entropy of activation. *Proceedings of the National Academy of Sciences*, 70(2):430–432.
- Madec, G. (2008). Nemo ocean engine. Technical report.
- McLaren, I., Sevigny, J.-M., and Corkett, C. (1989). Temperature-dependent development in pseudocalanus species. *Canadian Journal of Zoology*, 67(3):559–564.
- Moloney, C. L. and Field, J. G. (1991). The size-based dynamics of plankton food webs. i. a simulation model of carbon and nitrogen flows. *Journal of Plankton Research*, 13(5):1003–1038.
- Morel, A. (1991). Light and marine photosynthesis: a spectral model with geochemical and climatological implications. *Progress in oceanography*, 26(3):263–306.
- Moriarty, R. and O'Brien, T. (2013). Distribution of mesozooplankton biomass in the global ocean. *Earth System Science Data*, 5(1):45.
- Ohman, M., Eiane, K., Durbin, E., Runge, J., and Hirche, H.-J. (2004). A comparative study of calanus finmarchicus mortality patterns at five localities in the north atlantic. *ICES Journal of Marine Science: Journal du Conseil*, 61(4):687–697.

- Parent, L., Ferry, N., Barnier, B., Garric, G., Bricaud, C., Testut, C.-E., Le Galloudec, O., Lellouche, J., Greiner, E., Drevillon, M., et al. (2013). Global eddy-permitting ocean reanalyses and simulations of the period 1992 to present. Technical report, Technical report, Mercator Ocean.
- Saba, V., Friedrichs, M., Antoine, D., Armstrong, R., Asanuma, I., Behrenfeld, M., Ciotti, A., Dowell, M., Hoepffner, N., Hyde, K., et al. (2011). An evaluation of ocean color model estimates of marine primary productivity in coastal and pelagic regions across the globe. *Biogeosciences*, pages 489–503.
- Séférián, R., Bopp, L., Gehlen, M., Swingedouw, D., Mignot, J., Guilyardi, E., and Servonnat, J. (2014). Multiyear predictability of tropical marine productivity. *Proceedings of the National Academy of Sciences*, 111(32):11646–11651.
- Senina, I., Sibert, J., and Lehodey, P. (2008). Parameter estimation for basin-scale ecosystem-linked population models of large pelagic predators: application to skipjack tuna. *Progress in Oceanography*, 78(4):319–335.
- Sieburth, J. M., Smetacek, V., and Lenz, J. (1978). Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions 1. *Limnology and oceanography*, 23(6):1256–1263.
- Smith, B., Prentice, I. C., and Sykes, M. T. (2001). Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within european climate space. *Global Ecology and Biogeography*, 10(6):621–637.
- Steele, J. H. and Henderson, E. W. (1992). The role of predation in plankton models. *Journal of Plankton Research*, 14(1):157–172.
- Steinberg, D. K., Carlson, C. A., Bates, N. R., Goldthwait, S. A., Madin, L. P., and Michaels, A. F. (2000). Zooplankton vertical migration and the active transport of dissolved organic and inorganic carbon in the sargasso sea. *Deep Sea Research Part I: Oceanographic Research Papers*, 47(1):137 – 158.
- Steinberg, D. K., Goldthwait, S. A., and Hansell, D. A. (2002). Zooplankton vertical migration and the active transport of dissolved organic and inorganic nitrogen in the sargasso sea. *Deep Sea Research Part I: Oceanographic Research Papers*, 49(8):1445 – 1461.
- Tranchant, B., Refray, G., Greiner, E., Nugroho, D., Koch-Larrouy, A., and Gaspar, P. (2015). Evaluation of an operational ocean model configuration at 1/12° spatial resolution for the indonesian seas – part 1: Ocean physics. *Geoscientific Model Development Discussions*, 8(8):6611–6668.
- Vogt, M., Vallina, S., Buitenhuis, E., Bopp, L., and Le Quéré, C. (2010). Simulating dimethylsulphide seasonality with the dynamic green ocean model planktom5. *Journal of Geophysical Research: Oceans (1978–2012)*, 115(C6).

Appendix 1 : Model formulation

A. Global configuration

SEAPODYM-LTL model describes two state variables:

- integrated zooplankton production P over the first biological layer of the ocean expressed in gWW/m²/day. $P = P(x, y, t, \tau)$ represents the biomass of new organisms at position $(x, y) \in \Omega \subset \mathbf{R}^2$, at time $t \geq 0$ and at age $\tau \geq 0$.
- integrated zooplankton biomass B over the first biological layer of the ocean expressed in gWW/m². $B = B(x, y, t)$ represents the biomass of all recruited organisms at position $(x, y) \in \Omega$ and at time $t \geq 0$.

Model formulation The dynamics of zooplankton production P and zooplankton biomass B are modelled by two one-way coupled advection-diffusion and advection-diffusion-reaction (ADR) equations (we omit dimensions for brevity):

$$\frac{\partial P}{\partial t} + \frac{\partial P}{\partial \tau} = D \left(\frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} \right) - \frac{\partial}{\partial x}(\bar{u}P) - \frac{\partial}{\partial y}(\bar{v}P) \quad (6)$$

$$\frac{\partial B}{\partial t} = D \left(\frac{\partial^2 B}{\partial x^2} + \frac{\partial^2 B}{\partial y^2} \right) - \frac{\partial}{\partial x}(\bar{u}B) - \frac{\partial}{\partial y}(\bar{v}B) - \lambda B + P(\tau_r) \quad (7)$$

where D is the diffusion coefficient modelling the non-directional movements of P and B , $\bar{u}$ and $\bar{v}$ are the averaged horizontal currents over the layers. $\lambda(T)$ is the zooplankton mortality rate and $\tau_r(T)$ is the zooplankton recruitment age, both depending on the ambient temperature T . In zooplankton biomass model (7), the reaction term $-\lambda(T)B$ describes the mortality of B (sink). The source term $P(x, y, t, \tau_r(T))$ represents the recruitment process (see below).

Boundary conditions and age initial conditions. Neuman condition is applied at domain boundary (coasts): for all time t and age τ we have $\mathbf{n} \cdot \nabla P|_{\partial\Omega}(t, \tau) = 0$ and $\mathbf{n} \cdot \nabla B|_{\partial\Omega}(t, \tau) = 0$, where $\mathbf{n}$ is the normal to the boundary $\partial\Omega$. First age cohort of zooplankton production $P(x, y, t, 0)$ is computed according to Jennings et al. (2002a) equation linking the primary production and the production of a specific trophic level ($P_{TL} = PP \cdot TE^{TL-1}$). In SEAPODYM-LTL the primary production is the sum over the first 0-400m depth layer of the ocean of the net primary production (NPP) calculated with Behrenfeld and Falkowski (1997) method. Thus, the age initial condition is expressed as follow :

$$P(\tau = 0) = c E \int_0^{z_{400}} NPP(z) dz \quad (8)$$

where c is the unit conversion coefficient between carbon molecular content and wet weight of zooplankton; E is the energy transfer coefficient between phytoplankton and zooplankton ($E = TE^{TL-1}$) and z_l is the depth limit of the considered layer.

Numerical resolution Numerical resolution is based on the discretization of the spatial, temporal and age dimensions (for production equation 6 only). Spatial domain Ω is uniformly discretized with same constant step for both directions. The age variable τ in production equation 6 is uniformly discretized into age classes (cohorts) of constant size. The number of age classes is defined by the maximal recruitment time divided by the time step of the simulation. Time integration is performed using a multi-steps splitting method: first ADR processes are solved using alternative direction *implicit*

method along the two spatial directions. For each direction a tridiagonal system is solved using a forward-backward Gauss elimination method. The ageing process $\frac{\partial P}{\partial \tau}$ in production equation 6 is resolved during an additional third step. Note that the age and time dimensions have the same discretization step, so the ageing is simply done by augmenting the age class indices.

B. Regional operational modeling

A regional model needs boundary conditions which must be consistent with the global model. SEAPODYM implements a relaxation method which is a widely-used class of Open Boundary Conditions (OBC) method (Blayo and Debreu (2005)). The computational regional domain Ω is extended by an additional buffer domain Ω_b . Equations (6) and (7) are solved on $\Omega \cup \Omega_b$ using Dirichlet boundary conditions leading to solution N (P or B depending on the considered equation). Regional solution N^{reg} is obtained by relaxing solution N in Ω_b by $(1 - \alpha)N + \alpha N^{glo}$, where N^{glo} is the solution of the global model (OBC) and α is a sigmoid function increasing from 0 on the interface between Ω and Ω_b to 1 on the boundary of Ω_b . The OBC N^{glo} are obtained from a global configuration with spatial and temporal coarser resolutions. During the global model integration, solution is saved in a domain that overlaps the buffer domain of the regional configuration. Record time stepping in the global configuration is set to temporal resolution of the regional model in order to avoid temporal interpolation. Solution of the global model is spatially interpolated to the regional resolution scale onto the buffer domain Ω_b . In the case of zooplankton production, we also interpolate the global solution along the age variable direction (cohorts) in order to fit with the regional age structure (which is more refined compared to the global model).

Appendix 2 : Metrics

Let o_i be the i^{th} observed zooplankton biomass and p_i the i^{th} predicted zooplankton biomass. n is the number of observations.

III Perspectives

A Paramétrisation optimale

La paramétrisation utilisée dans cette étude a été obtenue à partir de multiples essais. Le but était d'obtenir une biomasse prédite du même ordre de grandeur que la biomasse observée dans la zone ECOFOCI tout en conservant le coefficient de transfert d'énergie (E) entre 0.05 et 0.15 et la relation $GT(\bar{T})$ à l'intérieur de l'intervalle de prédiction (Figure 1 de l'article, section précédente). Néanmoins, cette méthode est loin d'être optimale.

L'assimilation de données (voir section C) est la méthode envisagée afin d'obtenir une meilleure paramétrisation de ce modèle. En effet, un grand nombre de données sont disponibles (plus de 42 000 en climatologie). Elles couvrent une grande partie de l'océan global (Figure 3 de l'article) et pourraient donc constituer deux jeux de données, un d'entraînement utilisé pour les expériences d'assimilation et un autre pour la validation.

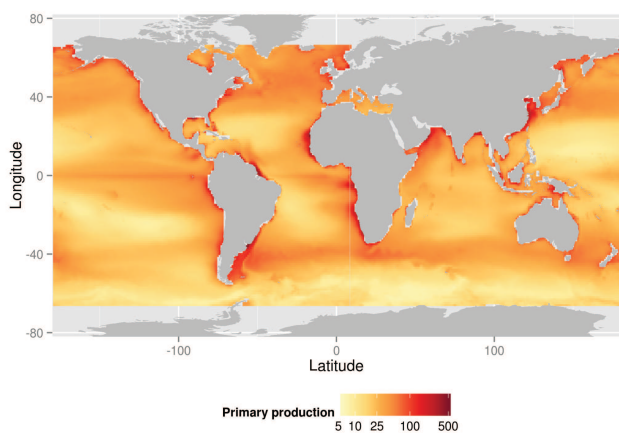
D'un point de vue technique, le code adjoint du modèle SEAPODYM-MTL a été développé et testé (voir Chapitre 3). L'adaptation de ce code adjoint à SEAPODYM-LTL demanderait très peu de travail de programmation. Toutefois, une attention doit être portée aux plans d'expériences utilisés afin de déterminer la paramétrisation optimale : la résolution du système d'équation (1.13 - 1.14) est réalisé en deux temps et seul l'adjoint de la résolution numérique de l'équation de biomasse a été écrit. Ce qui ne permet donc pas une optimisation des paramètres utilisés dans cette équation, à savoir les paramètres de recrutement. Néanmoins, des expériences de sensibilité à ces paramètres (non reproduites ici) ont montré une très faible sensibilité du modèle à ces paramètres. La méthodologie adaptée serait donc d'effectuer des expériences d'assimilation de données en fixant ces deux paramètres à différentes valeurs afin d'obtenir une meilleure paramétrisation de ce processus.

B Forçages biogéochimiques

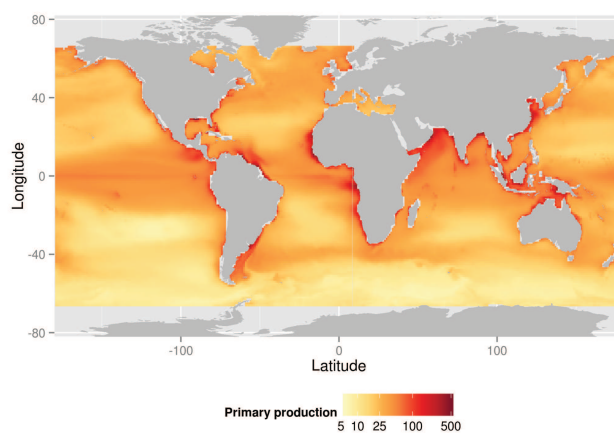
Cette étude illustre le problème des champs de forçages pour l'ensemble de la gamme de modèle SEAPODYM. Nous disposons de deux types de champs de production primaire issus tous les deux des mêmes données satellitaires (SST, PAR et Chl_a) à partir de deux modèles différents. Si la méso-échelle est bien reproduite dans les deux cas, la distribution globale de production primaire varie. Le modèle Eppley présente des valeurs plus faibles et un contraste moins marqué entre les zones productives et celles plus pauvres.

La distribution de sortie de la biomasse de zooplancton est donc elle aussi différente entre les deux cas. Le champ de biomasse issu du modèle forcé par les champs VGPM est plus contrasté et présente des valeurs nettement plus fortes dans les eaux froides.

Une étude de sensibilité sur l'ensemble de la gamme de modèle pourrait apporter une amélioration significative de leur prédiction en utilisant le modèle de production primaire le plus adapté.

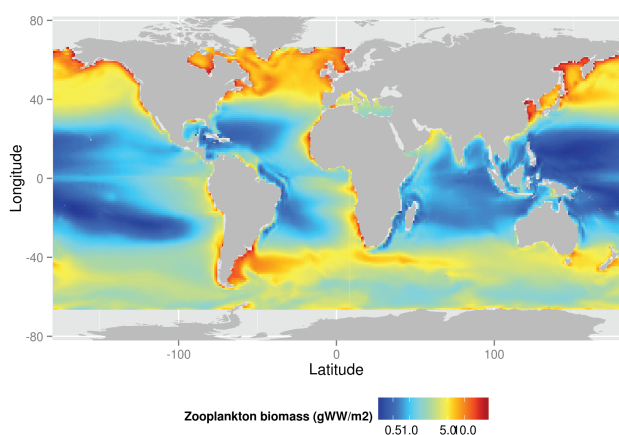


(a) VGPM

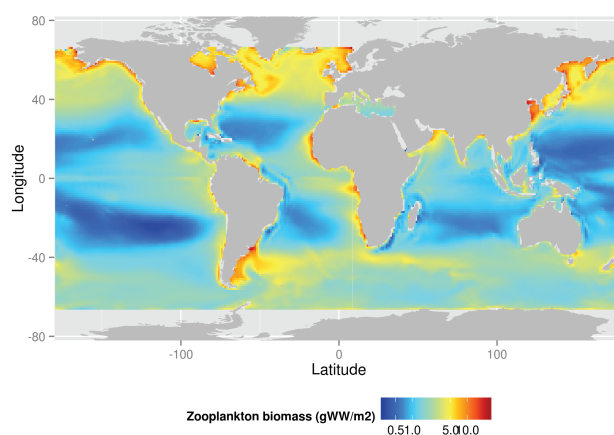


(b) Eppley

Figure 2.3 – Champs de production primaire climatologique annuel réalisé à partir des modèles (a) VGPM et (b) Eppley.



(a) VGPM



(b) Eppley

Figure 2.4 – Champs de de biomasse de zooplancton climatologique annuel réalisé en utilisant comme données d'entrée du modèle SEAPODYM-LTL les champs de production primaire (a) VGPM et (b) Eppley.

Bibliographie

Lima, I. D., Olson, D. B., and Doney, S. C. (2002). Intrinsic dynamics and stability properties of size-structured pelagic ecosystem models. *Journal of Plankton Research*, 24(6) :533–556.

Chapitre 3 Modélisation de la distribution spatiale du micronecton

Sommaire

I Paramétrisation de SEAPODYM-MTL	57
A Croissance	57
B Coefficients de transfert et de distribution d'énergie	59
II Données assimilées dans SEAPODYM-MTL	59
III Optimization of a micronekton model with acoustic data (Lehodey et al. (2015))	60
IV Un nouveau jeu de données plus étendu : résultats de l'assimilation de données en global, nouveaux défis	75
A Jeu de données global	75
B Expériences d'assimilation de données en global	76
C Conclusions	78

La structure du sous-modèle SEAPODYM-MTL, qui modélise les niveaux trophiques intermédiaires, a été longuement abordée dans le chapitre introductif de cette thèse. Nous nous centrons ici sur son paramétrage. Nous allons d'abord nous intéresser aux paramètres de croissance et de transfert d'énergie de la production primaire au micronecton. Afin de prendre en compte les différents comportements de migration nyctémérale, nous avons établi des groupes fonctionnels (voir partie III du Chapitre 1, particulièrement la figure 1.15). La paramétrisation de la distribution de l'énergie disponible à cet échelon trophique entre ces groupes fonctionnels sera abordée dans la section III au travers de l'article [Lehodey et al. \(2015\)](#). Enfin, nous appliquerons la méthodologie développée dans cet article à un jeu de données beaucoup plus vaste (voir section IV de ce chapitre). Ce qui soulèvera de nouvelles questions pour la modélisation du micronecton.

I Paramétrisation de SEAPODYM-MTL

A Croissance

Les processus de croissance sont modélisés dans SEAPODYM-MTL par un terme de mortalité et un terme de recrutement (équation (1.2) du Chapitre 1). Il est donc nécessaire de paramétrer le taux de mortalité et le temps de recrutement. Le taux de mortalité (λ) peut être calculé à partir du temps de recrutement et du

temps de maturité, plus disponible dans la littérature, avec la formule suivante (Lehodey et al. (2010)) :

$$\frac{1}{\lambda} = t_m - \frac{1}{3}t_r \quad (3.1)$$

Les travaux de Gillooly et al. proposent une décroissance exponentielle des temps de développement avec la température. Ils intègrent la variation de la masse dans leur équation et proposent un cadre théorique basé sur la cinétique biochimique et l'équation de Van't Hoff - Arrhénius. C'est cette équation qui permet d'exprimer la dépendance à la température des processus physiologiques (par exemple les vitesses de réactions enzymatiques). Elle a pour origine la thermodynamique et la description du comportement d'un système contenant un grand nombre de molécules à partir de considérations statistiques sur les interactions individuelles de ces molécules.

Inspiré des travaux de Gillooly et al., le taux de mortalité et le temps de recrutement sont fonction, dans SEAPODYM-MTL, de la température instantanée. Néanmoins, l'absence de données concomitantes de temps de développement, température et masse de l'animal conduit Lehodey et al. (2015) à chercher une relation uniquement entre le logarithme des temps de développement et la température ambiante. Une régression linéaire entre le logarithme du temps de maturité d'espèce de micronecton et la température permet d'estimer les paramètres de la relation avec un coefficient de corrélation élevé $R^2 = 0.89$. La figure 3.1 représente les données disponibles, l'exponentielle issue de la régression linéaire et l'intervalle de prédiction à 95% (en grisé). On peut noter qu'aucune tendance en fonction du taxon ne se dégage.

L'abondance extrêmement faible de données d'âge de recrutement associée à une température n'a pas permis d'obtenir une paramétrisation directe par régression. Il a été supposé que la forme de la relation était la même et les paramètres ont été obtenus par un ajustement manuel aux quelques rares observations disponibles. La figure 1.2 présente les données (avec leur incertitude matérialisée par des traits gris), l'exponentielle optimale (en trait plein) et l'intervalle de prédiction à 95% (en grisé). La pente de cette courbe a donc été fixée égale à celle du temps de maturité. Ce qui conduit à une paramétrisation très proche de la paramétrisation qui a été retenue soit 25% du temps de maturité (courbe en tirets sur la figure 3.2).

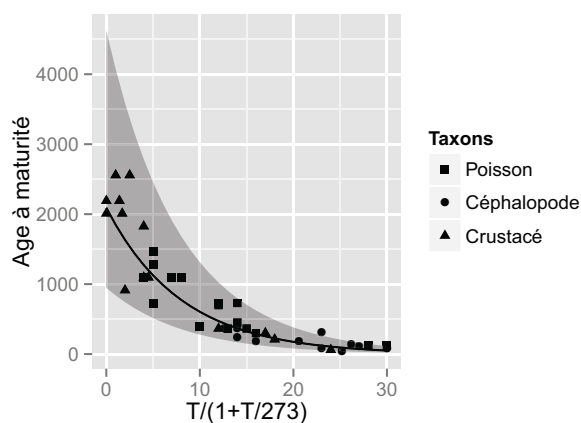


Figure 3.1 – Temps de maturité

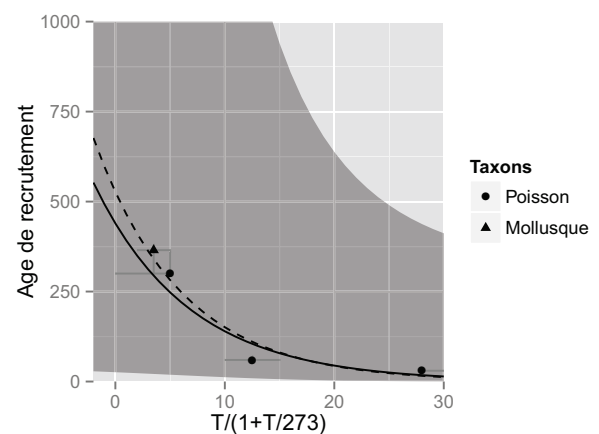


Figure 3.2 – Temps de recrutement

A l'aide de l'équation 3.1, on obtient donc les paramétrisations suivantes :

$$\frac{1}{\lambda} = 2109 \exp(-0.125 \bar{T}) \quad (3.2)$$

$$t_r = 527 \exp(-0.125 \bar{T}) \quad (3.3)$$

B Coefficients de transfert et de distribution d'énergie

La paramétrisation du transfert d'énergie entre la production primaire et chacun des groupes de micronecton peut se concevoir en deux étapes. D'abord un coefficient de transfert total (E) entre la production primaire et le micronecton, puis des coefficients de distribution (E'_i) relatifs à chacun des groupes fonctionnels. L'énergie disponible à cet échelon trophique est entièrement distribuée entre les groupes fonctionnels, la somme des coefficients est donc égal à 1.

D'après l'équation 1.1 dans l'introduction de cette thèse (Ryther et al. (1969), Schaefer (1965), Iverson (1990), Jennings et al. (2002)), le coefficient de transfert d'énergie entre la production primaire et le micronecton peut être calculé à partir du niveau trophique et d'un transfert d'énergie d'un niveau trophique à l'autre. Se basant sur un niveau trophique nul pour le phytoplancton, le niveau trophique du micronecton est de 2.5 (Iverson (1990); Lehodey et al. (2015)). La valeur du transfert d'énergie entre niveaux trophiques est estimée entre 0.1 et 0.2 (Iverson (1990); Jennings et al. (2002)). En zone côtière, ce transfert d'énergie se rapproche de 0.2, alors qu'il est plus proche de 0.1 en zone hauturière. Dans le cadre de SEAPODYM-MTL, la valeur retenue globalement est de 0.1.

Nous allons nous intéresser à la redistribution de cette énergie vers les différents groupes fonctionnels. En connaissant au préalable les populations présentes dans un écosystème, leur comportement de migration et en supposant la stabilité de ce dernier, on pourrait discriminer les organismes collectés en fonction de leur groupe fonctionnel et, donc, mesurer directement la distribution d'énergie entre les groupes fonctionnels. Cette mesure directe est difficile à mettre en place. Il n'est donc pas envisageable ici d'obtenir des mesures directes de cette répartition d'énergie. Afin d'estimer indirectement ces paramètres, nous utiliserons une méthode d'assimilation de données.

II Données assimilées dans SEAPODYM-MTL

Des observations sont comparées aux prédictions du modèle. Dans le cas de SEAPODYM-MTL, les observations doivent être comparables à la biomasse de micronecton. Nous pourrions utiliser des données de chalutage de micronecton, mais, du fait de la forte sélectivité des filets, les données ne seraient pas représentatives du micronecton. De plus, ces données sont très disparates et peu disponibles.

Nous allons donc utiliser des données acoustiques, plus précisément des densités acoustiques acquises à la fréquence de 38kHz. Cette fréquence est l'une des plus adaptées pour observer le micronecton (Bertrand et al. (1999); McClatchie and Dunford (2003); Kloser et al. (2009)). La portée des écho-sondeurs varie avec la fréquence. La fréquence de 38kHz est particulièrement intéressante pour nous puisque sa portée est similaire à la profondeur maximale du domaine du modèle (environ 1000 m).

De plus, cette méthode permet une acquisition en continue des observations, de jour comme de nuit, ce qui est crucial ici puisque nous modélisons un phénomène qui varie entre les différentes périodes du cycle jour-

nalier. Les observations issues de cette méthode sont très détaillées le long de la profondeur (intégrées sur 10 m de profondeur pour les moins précises dont nous disposons), ce qu'une méthode de collecte d'organismes ne pourrait pas atteindre. Cette méthode est en plein développement et la couverture des observations ne cesse d'augmenter. Néanmoins, ces observations présentent un inconvénient majeur : la densité acoustique n'est pas directement convertible en biomasse. Localement, la densité acoustique (REF) est proportionnelle à la biomasse. Cependant, ce coefficient de proportionnalité dépend notamment de l'assemblage d'espèces insonifiées. Nous ferons, ici, l'hypothèse que ce coefficient de proportionnalité ne varie pas le long de la colonne d'eau.

III Optimization of a micronekton model with acoustic data (Lehodey et al. (2015))

Cet article décrit la méthodologie résumée précédemment et l'applique à un jeu de données restreint, constitué d'un transect acoustique au Nord d'Hawaï. Des expériences dites jumelles ont également été réalisées afin d'établir la quantité de données minimales nécessaires. Ce type d'expérience utilise comme observations des prédictions du modèle obtenues avec des paramètres connus et permet donc d'étudier l'erreur faite sur la paramétrisation issue de l'assimilation de données.



Original Article

Optimization of a micronekton model with acoustic data

Patrick Lehodey^{1*}, Anna Conchon¹, Inna Senina¹, Réka Domokos², Beatriz Calmettes¹, Julien Jouanno³, Olga Hernandez⁴, and Rudy Kloser⁵

¹Space Oceanography Division, CLS, 8-10 rue Hermes, Ramonville 31520, France

²Ecosystems and Oceanography Division, Pacific Islands Fisheries Science Center, NMFS, NOAA, 2570 Dole Street, Honolulu, HI 96822, USA

³LEGOS, 18, av. Edouard Belin, Toulouse 31400, France

⁴LOCEAN, 4 Place Jussieu, Paris 75252, France

⁵CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart 7005, Australia

*Corresponding author: Tel: +33 561 393 770; Fax: +33 561 393 782; e-mail: plehodey@cls.fr

Lehodey, P., Conchon, A., Senina, I., Domokos, R., Calmettes, B., Jouanno, J., Hernandez, O., and Kloser, R. Optimization of a micronekton model with acoustic data. – ICES Journal of Marine Science, 72: 1399–1412.

Received 4 April 2014; accepted 30 November 2014; advance access publication 23 December 2014.

In the pelagic foodweb, micronekton at the mid-trophic level (MTL) are one of the lesser known components of the ocean ecosystem despite being a major driver of the spatial dynamics of their predators, of which many are exploited species (e.g. tunas). The Spatial Ecosystem and Population Dynamics Model is one modelling approach that includes a representation of the spatial dynamics of several epi- and mesopelagic MTL functional groups. The dynamics of these groups are driven by physical (temperature and currents) and biogeochemical (primary production, euphotic depth) variables. A key issue to address is the parameterization of the energy transfer from the primary production to these functional groups. We present a method using *in situ* acoustic data to estimate the parameters with a maximum likelihood estimation approach. A series of twin experiments conducted to test the behaviour of the model suggested that in the ideal case, that is, with an environmental forcing perfectly simulated and biomass estimates directly correlated with the acoustic signal, a minimum of 200 observations over several time steps at the resolution of the model is needed to estimate the parameter values with a minimum error. A transect of acoustic backscatter at 38 kHz collected during scientific cruises north of Hawaii allowed a first illustration of the approach with actual data. A discussion followed regarding the various sources of uncertainties associated with the use of acoustic data in micronekton biomass.

Keywords: acoustic, maximum likelihood estimation, micronekton, model optimization, modelling, Pacific ocean, SEAPODYM.

Introduction

The development of marine ecosystem models is providing new useful tools and products in investigating the combined effects of fishing, environmental variability, and the impacts of climate changes on species of interest; in particular, exploited or protected species (e.g. De Young *et al.*, 2004; Plagányi, 2007; Lehodey *et al.*, 2008; Senina *et al.*, 2008; Fulton, 2010; Dueri *et al.*, 2012; Sibert *et al.*, 2012). The first necessary step in the development of such ocean ecosystem models is to produce reliable predictions of the physical–biogeochemical oceanic environment. Physical ocean–circulation models are now providing realistic simulations of the ocean by assimilating large amounts of sea surface satellite and depth profile data (e.g. ARGO program) to produce reanalyses of the past and present ocean state at high resolution. The modelling

of lower trophic levels, that is, the phyto- and zooplankton, is also rapidly progressing (Brasseur *et al.*, 2009; Holt *et al.*, 2014) and should soon provide operationally realistic outputs to connect these ocean models to the higher biological levels. As an immediate alternative, primary production by phytoplanktonic groups can be inferred from satellite data (e.g. Behrenfeld and Falkowski, 1997), with the advantage of providing a realistic description of mesoscale activity.

While the modelling of these physical and biological variables is well advanced, this is not the case however, for the epi- and mesopelagic mid-trophic level (MTL) organisms (macrozooplankton and micronekton) that are a critical component in the oceanic ecosystem. By definition, micronekton are small organisms (~1–20 cm or g) that can swim; however, their small sizes are still strongly

impacted by oceanic circulation from large to mesoscale. They distribute between the surface and the deep layers with some groups of species migrating every night and day between these vertical layers. Predators have adapted their behaviour to chase these forage species by using their physical and physiological skills for accessing the different layers. These predators include many exploited stocks (e.g. tunas and swordfish) and protected species (e.g. turtles, seabirds, sharks, and marine mammals), and realistic simulations of micronektonic prey would provide a key explanatory variable that is currently missing to better understand individual behaviour and population dynamics. Additionally, micronektonic organisms are a major source of grazing of zooplankton and are predators of eggs and larvae drifting in the pelagic environment, including those of the exploited species. Micronekton also interact with the carbon cycle, including the CO₂ released in the atmosphere by human activity, as they contribute to the production of particular organic matter that is recycled either in the microbial loop or stored by sinking to the deep ocean (Holt et al., 2014). Thus, micronekton are central to the understanding and modelling of oceanic ecosystems and there is a pressing need for increased knowledge and for models pertaining to their biomass estimate and ecological role.

Several modelling approaches can be used to estimate micronekton biomass (e.g. Kitchell et al., 1999; Jennings et al., 2008; Maury, 2010; Lehodey et al., 2010a). The Spatial Ecosystem and Population Dynamics Model (SEAPODYM) in this study includes a representation of several MTL functional groups, that is, the epi- and mesopelagic micronekton, to describe the prey fields of tuna and other large pelagic predators (Lehodey et al., 2010a). This micronekton model has been used in several applications on tuna species (Lehodey et al., 2008; Senina et al., 2008; Lehodey et al., 2010b, 2013; Briand et al., 2011; Sibert et al., 2012), and its outputs used to investigate habitats and movements of other large oceanic species (Abecassis et al., 2013; Lambert et al., 2014).

A key issue in the SEAPODYM modelling approach is to parameterize the energy transfer from the primary production to the functional groups of micronekton. A preliminary parameterization (Lehodey et al., 2010a) was achieved based on a first compilation of existing data in the literature and a Pacific basin-scale simulation at coarse resolution (1° × month). A more rigorous approach requires the use of data assimilation methods to optimize the parameters using acoustic data that provide a synoptic view of micronekton biomass in the vertical layers of the ocean. Over the recent years, acoustic estimates of micronekton biomass have received increased attention, and many acoustic transects are collected at basin scales by various research and fishing vessels. An effort of standardization between these different data sources (e.g. IMOS project: <http://imos.org.au/anmnacous.html>) should fill the important gap in the understanding of basin-scale mid-trophic biomass and distribution, providing data to initialize and assimilate into ecological models.

Here, we present a methodology for a parameter optimization approach based on the assimilation of acoustic data in the MTL model. Assimilation experiments with the MTL model at high resolution in the Pacific Ocean are used to validate the method and to highlight the need for collecting and standardizing such data. After a series of twin experiments conducted to test the behaviour of the model, a transect of acoustic backscatter at 38 kHz collected during scientific cruises north of Hawaii illustrates the approach. We assumed that the intensity of the acoustic signal is proportional to the density of micronekton. This is not always true, for example,

due to resonant scattering diel shifts, and has been identified by the acoustic community as needing more research (e.g. Handegard et al., 2013).

Material and methods

The mid-trophic model

In its present version, the SEAPODYM-MTL submodel includes three vertical layers and six functional groups characterized by their vertical behaviour, that is, with or without occurrence of diel migration between three vertical layers. Recruitment, ageing, mortality, and passive transport with horizontal currents are modelled by a system of advection–diffusion–reaction equations, taking into account the vertical behaviour of organisms (Lehodey et al., 2010a). Since the dynamics are represented by a relationship of temperature-linked time development, only six parameters in the model have to be estimated. The first one (E) defines the total energy transfer between primary production and all the MTL groups. The other parameters are the relative coefficients (E'_n), redistributing this energy through the different components (the sum of which being 1). The parameterization of E requires absolute biomass estimates of MTL, while the matrix of E'_n coefficients can be estimated simply using relative day and night values integrated in the three vertical layers of the model. The six parameters of the MTL model have been initially tuned according to very limited *in situ* data and information from the literature (Lehodey et al., 2010a).

Terminology and definition of vertical layers

In its first version, the layers of the model were conveniently named as epi-, meso-, and bathypelagic (Lehodey et al., 2010a). However, the names of the functional groups were revised (Figure 1) to comply with the standard terminology that defines the bathypelagic layer < 1000 m. The three vertical layers are therefore renamed as epipelagic, upper- and lower- mesopelagic layers. They have been defined relatively to the euphotic depth (Z_{eu}), assuming the vertical migration of micronektonic organisms is primarily driven by light. When comparing acoustic profiles with Z_{eu} (Figure 1) in the region of study, the epipelagic layer appeared to extend deeper than the euphotic depth computed according to the VGPM model of Behrenfeld and Falkowski (1997). Thus, the first layer was computed using $1.5 \times Z_{eu}$. The second and third layer boundaries were defined as three and seven times the depth of the first layer (with a maximum set at 1000 m) based on visual inspection of acoustic transects (Figure 1). The definition of vertical layers, based on the euphotic depth and the few available acoustic transects for this study, could evolve while more data become available (see Discussion). Each MTL functional group is identified by the layer that is inhabited by its organisms during the day and night. With this terminology, the epipelagic group is called MTL1.1 since it is always in the first layer, while the highly migrant lower mesopelagic group becomes MTL3.1, that is, inhabiting the third layer during the day and moving up to the surface layer at night (Figure 1).

Model domain and forcing

Acoustic transect data are strongly influenced by the mesoscale activity and represent very detailed but also a very small portion of the ocean. A realistic high-resolution ocean physics and biogeochemistry feedback is needed to extract the largest information from data used in the optimization approach. We used the global ocean reanalysis GLORYS (glorysproducts@mercator-ocean.fr). With the project GLORYS, supported by the French Groupe

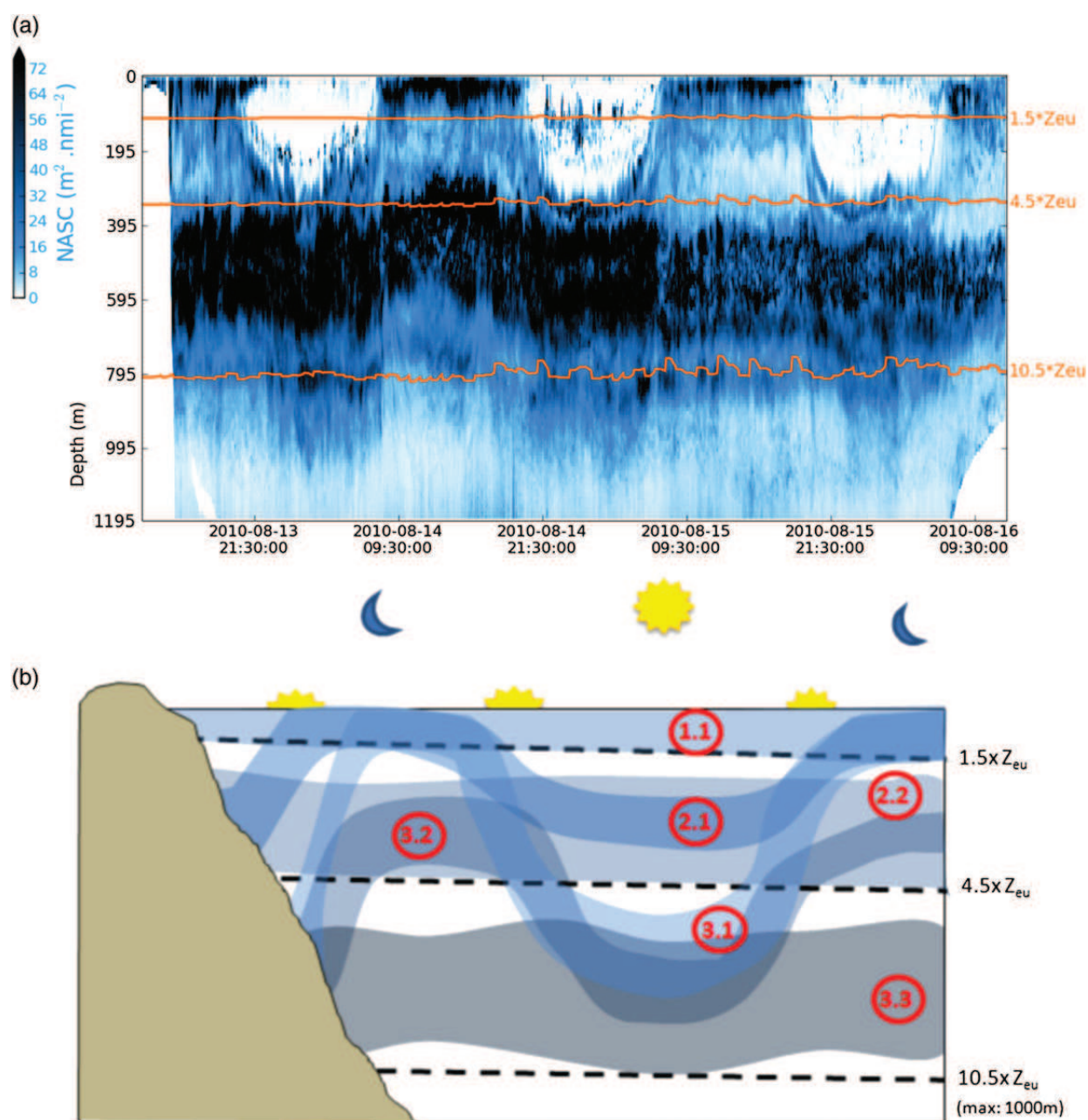


Figure 1. Identification of MTL functional groups on acoustic echogram (a) and conceptual model (b), revised from Lehodey *et al.* (2010a). 1.1: Epipelagic; 2.1: migrant upper mesopelagic; 2.2: upper mesopelagic; 3.1: highly migrant lower mesopelagic; 3.2: migrant lower mesopelagic; 3.3: lower mesopelagic.

Mission Mercator Coriolis, a first eddy resolving global ocean re-analysis has been produced for the 2002–2009 period with the ocean general circulation model configuration ORCA025 NEMO, that is, a spatial resolution of $1/4^\circ$ (Barnier *et al.*, 2006). The assimilation method is based on a reduced order Kalman filter (SEEK formulation, Pham *et al.*, 1998) adapted to eddy resolving global ocean model configuration (Tranchant *et al.*, 2008). The GLORYS.2v1 re-analysis was updated for the recent period with the operational Mercator-Ocean (<http://www.mercator-ocean.fr/eng>) global ocean model PSY3 that is using the same grid and assimilation

scheme than GLORYS. Because the ocean-circulation model assimilates satellites (SST and altimetry) and *in situ* data, predicted fields of temperature and currents are globally coherent with those of primary production derived from ocean colour data (Figure 2), using the VGPM model of Behrenfeld and Falkowski (1997).

The model computes spatial distribution of production and biomass of each functional group at the resolution of the physical model. The total biomass for each layer during day- and night-time was computed by adding the different components accordingly. For instance, the surface layer is inhabited only by the epipelagic

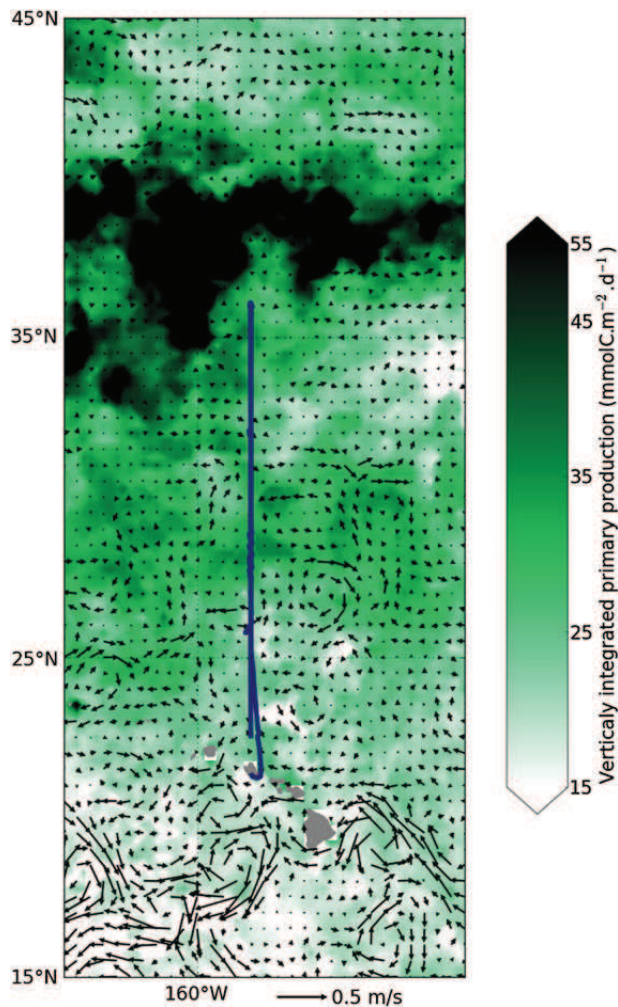


Figure 2. Snapshots showing the derived primary production computed following the VGPM model of Behrenfeld and Falkowski (1997) with superimposed surface currents (average in the euphotic layer) predicted in GLORYS2v1 reanalysis (glorysproducts@mercator-ocean.fr) and the Hawaiian transect.

group during the day but the sum of epipelagic, upper-migrant mesopelagic and highly migrant lower mesopelagic groups during the night. These predictions can be compared with acoustic biomass estimates by layer during nights and days. Though the original domain of the model is global between latitude 66.5°N and 66.5°S, a subdomain has been extracted for optimization in the North Pacific to save computational time, with geographical coordinates 66.5°N–0°N, 90°E–70°W and closed boundary conditions.

Acoustic data

A shipboard survey was conducted in March 2009 along the longitude 158°W from 22°N to 36°N (so-called Hawaiian transect in the following text) that has been used for optimization experiments (Figure 2). Bioacoustic signals were collected continuously down to 1200 m with a hull-mounted split-beam Kongsberg Maritime AS Simrad EK60 system (Horten, Norway), operating at the 38, 70, and 120 kHz frequencies. The EK60 system was calibrated before data collection with a 38.1-mm tungsten carbide sphere

according to standard methods (Foote et al., 1987). As the higher frequency signals attenuate shallower than the bottom of the lower mesopelagic layer, only 38-kHz full-resolution acoustic data were used. Volume backscattering coefficients (S_v , in dB re 1 m^{-1}) were converted into nautical area scattering coefficient (NASC, in m^2 nautical mile $^{-2}$) values and assumed proportional to micronekton biomass (Mac Lennan et al., 2002). Since the 38-kHz frequency signal is dominated by micronektonic organisms (Handegard et al., 2013), the integration of NASC is believed to be representative of the total biomass of the six functional groups of the model. However, a series of potential biases occur that are presented in Discussion. Also, the position of the transducer below the hull prevents any recording from the first 5 m, while data ~ 10 m from the transducer face are unreliable due to non-linearities in the nearfield. In addition to the loss of recording in the upper 15 m, water displaced in front of the vessel together with the possible escapement behaviour of organisms likely leads to a negative bias in biomass estimates in the surface layer. To compensate this bias, we decided to correct the signal by adding 5% of total biomass in the epipelagic functional group, considering that ~ 5 –10% of the total biomass should be in the first 100 m during day-time in the tropical and subtropical ocean (e.g. Grandperrin, 1975; Irigoien et al., 2014).

After a few tests, the Hawaiian transect showed a low signal-to-noise ratio (Figure 3), thereby making convergence in the model difficult. This noise is a result of the cavitations and bubble dropout, characterized by high-frequency localized extreme values and was filtered using a running geometric mean [Equation (1)] with a window range $(1 + 2f)$ of seven bins for each vertical layer at a resolution of 10 m.

$$\text{NASC}'(d, t) = \left(\prod_{t'=t-f}^{t'=t+f} \text{NASC}(d, t') \right)^{1/(1+2f)}. \quad (1)$$

Optimization approach

The simple modelling approach used to describe the MTL components with a limited number of parameters helps implementing a method of parameter estimation by using data assimilation. The matrix of E'_n coefficients can be estimated simply using relative day and night values of acoustic backscatter integrated in each of the three vertical layers of the model. The energy transfer coefficients are optimized to fit the relative ratios of micronekton biomass (or NASC) between layers changing during day and night periods. First, the NASC values are integrated at the resolution of the model in space (in each cell grids of the model and in each layer) and time (during night- and day-time, excluding transition periods). Based on the sum of values for the three layers, the signal of the first layer is increased by 5%. Then, the ratio (ρ) for each layer (K) and night or day period (Ω) is computed relatively to the corrected sum of the three layers [Equation (2)] defined by their upper (Z_u) and deeper (Z_d) vertical boundaries:

$$(\rho K)_{i,j,t,\Omega}^{\text{obs}} = \frac{\int_t \int_i \int_j \int_{Z_d(K)}^{Z_u(K)} (\text{NASC}(x, y, z, t) I_\Omega) dz dy dx dt}{\int_t \int_i \int_j \int_0^{Z_u(3)} (\text{NASC}(x, y, z, t) I_\Omega) dz dy dx dt}, \quad (2)$$

where (i, j) is the position (the cell) in the grid of the model domain, t is the time (in the model resolution), Ω is the part of the day (night or day), and I_Ω is the identity function (1 in Ω , 0 elsewhere).

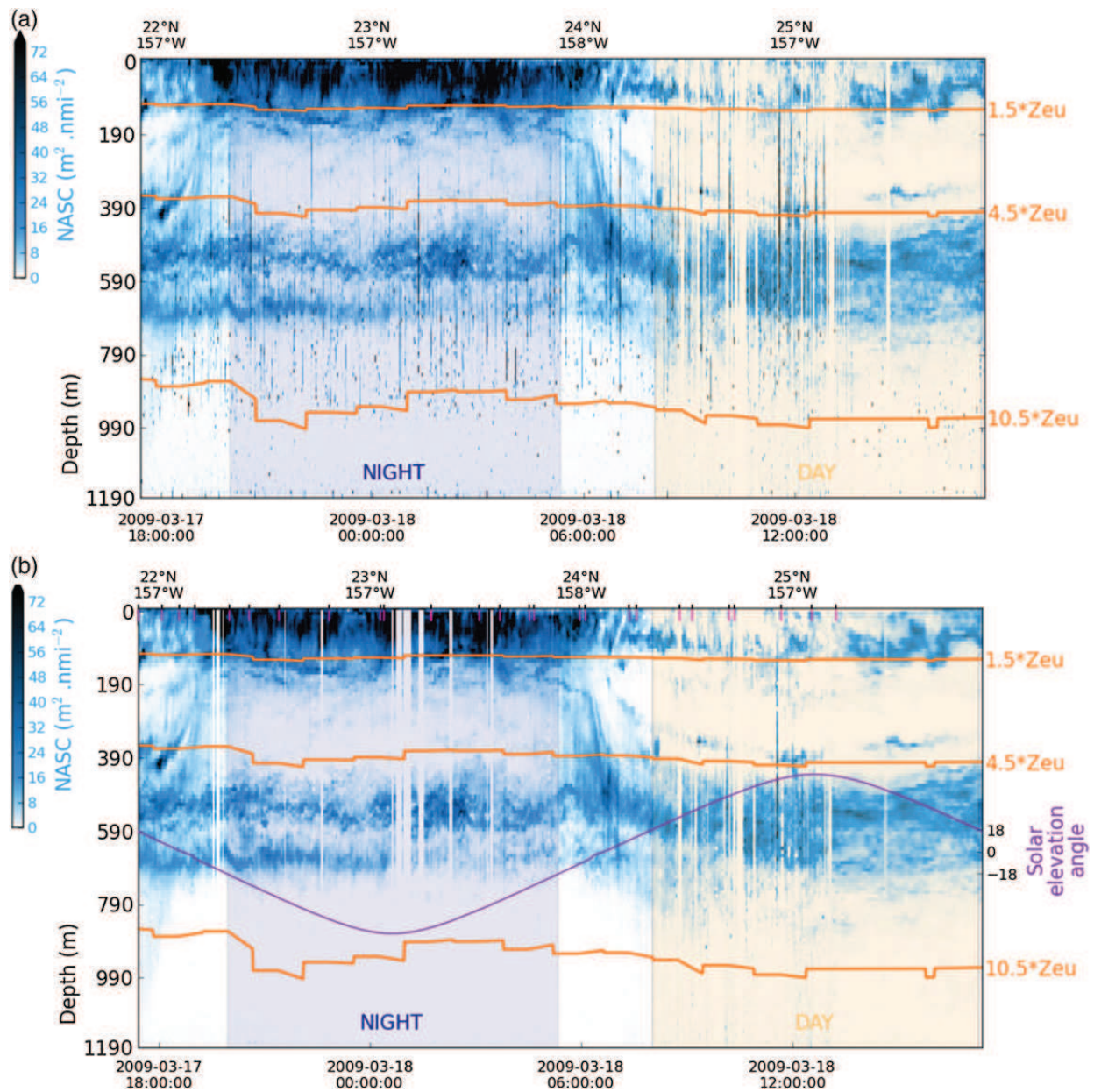


Figure 3. Example of a portion of acoustic transect before (a) and after (b) processing the data (both averaged over a minute). After the vertical layers have been defined, the signal strength is vertically integrated then averaged at the spatial resolution of the grid of the model ($1/4^\circ$) after excluding the sunset and sunrise periods. The orange lines delineate the vertical layer boundaries based on the euphotic depth. The purple line shows the variation of the solar elevation angle (which is used to discriminate between night and day) through the day. Small ticks (black and pink) on the horizontal axis indicate the position each five horizontal cells of the model grid.

Day and night periods are defined based on the solar elevation angle (Downey, 1990; see <https://pypi.python.org/pypi/pyephem/> for Python language scripts) with day when the altitude is $>18^\circ$ and night with altitude $\leq 18^\circ$ (Figure 3).

The integration of acoustical signal along a transect is illustrated in Figure 3. According to the local time of the day, these values can be compared with the relative distribution of predicted biomass in the same layers, accounting for the different combination of MTL components due to vertical migration. Sunset and sunrise periods are excluded by the definition using solar altitude.

Observed ratios (ρ^{obs}) for each layer K are compared with the model predictions (ρ^{pred}). The predicted ratios in grid cell (i, j) at time t , for the part of the day Ω , in a layer K are the sum of the biomass of each group inhabiting the layer K , at this position and time and divided by the total biomass in all layers [Equation (3)]:

$$(\rho_K)^{\text{pred}}_{i,j,t,\Omega} = \frac{\sum_{1 \leq n \leq 6} \sum_{K(n,\Omega)=K} B_{i,j,t}^n}{\sum_{1 \leq n \leq 6} B_{i,j,t}^n}, \quad (3)$$

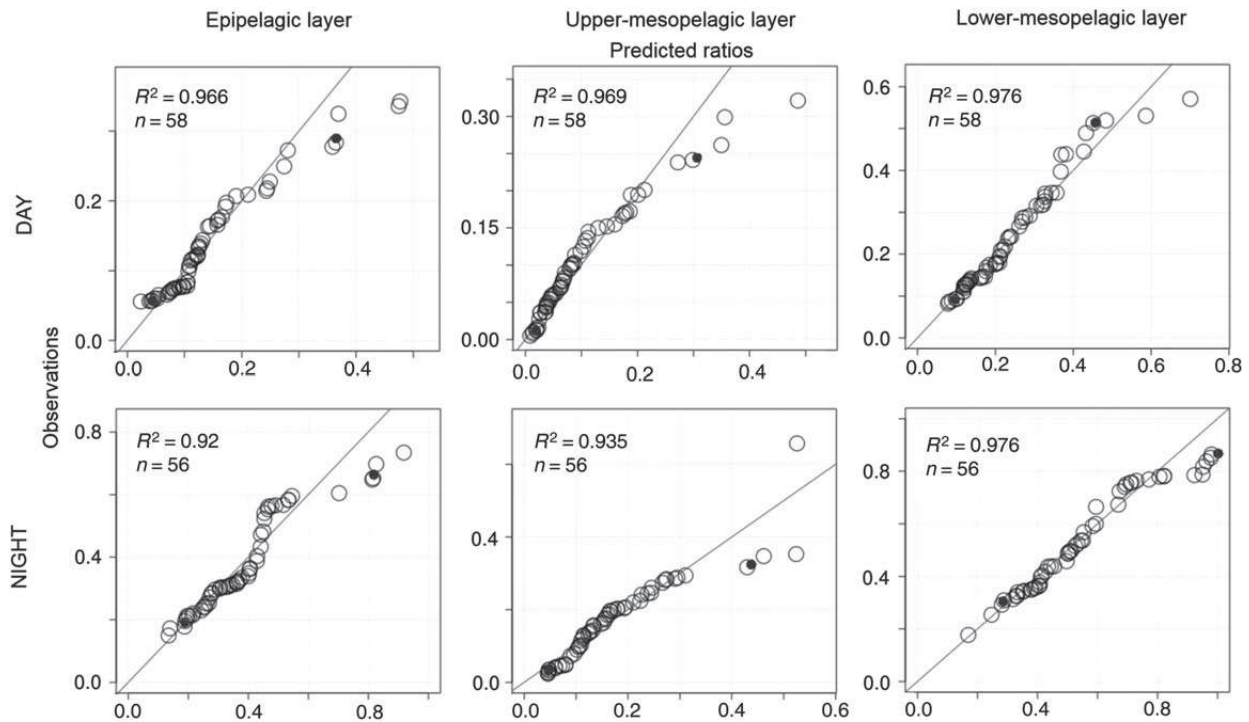


Figure 4. Q–Q plots of the observed ratios by layer and period of day using lognormal distribution for predicted ratios of epi- and upper mesopelagic layers and the complementary distribution ($1 - \rho_3^{\text{obs}}$) for lower mesopelagic layer. Black dots indicate the 5th and 95th percentile of the distribution.

where n is the functional group, B^n is the biomass of the group n , and $k(n, \Omega)$ is the layer inhabited by group n during the period of day Ω .

The optimization approach will search for the optimal set of parameters E_n that provides the best fit between observation and prediction. Since the approach uses relative ratio of integrated acoustic signal, it is possible to include data from various sources and without standardization. However, for the parameterization of E , that is, the total energy transfer that controls the level of absolute biomass, it would be necessary to use calibrated biomass estimates. They need to combine acoustic and micronekton net sampling, allowing to convert backscatter values in micronekton biomass after careful discrimination between various patches of records. This objective is not included in the present approach. Nevertheless, based on a detailed acoustic study (Kloser *et al.*, 2009), the previous estimate of biomass appeared largely underestimated and the value of E was increased (i.e. manually tuned to a value of 0.04) to fit these recent estimates.

The optimization approach uses the adjoint technique with the quasi-Newton gradient method to minimize the cost function (L^-) as described in Senina *et al.* (2008), but the new likelihood function required several modifications of the adjoint code.

First, the distribution of observed ratios suggests a lognormal distribution for data in the first 2 layers either during day or night, but not for the deepest (lower mesopelagic) layer. However, since the sum of ratios of the three layers is necessary at each given location and time, it was possible to use the complementary distribution to 1 (i.e. $1 - \rho_3^{\text{obs}}$) which in that case becomes much closer of the lognormal distribution. Quantile–quantile plots (Q–Q plots, Figure 4) presenting the fit between observations and theoretical

distributions confirmed that these distributions of observed ratios (ρ^{obs}) are close to the lognormal distribution ($R^2 > 0.9$).

Based on this distribution, the cost function is built as the sum of the negative log-likelihood computed for each layer [Equations (4–6)]:

$$L^- = \sum_{K=1}^3 L_K^-(\rho_K^{\text{obs}}) \quad (4)$$

with

$$L_K^-(\rho_K^{\text{obs}}) = \frac{1}{2} \sum_{i,j,t,\Omega} [\ln((\rho_K)_{i,j,t,\Omega}^{\text{obs}}) - \ln((\rho_K)_{i,j,t,\Omega}^{\text{pred}})]^2 \quad (5)$$

for $K = 1$ and 2, and since the sum of ratios is 1

$$L_K^-(\rho_K^{\text{obs}}) = \frac{1}{2} \sum_{i,j,t,\Omega} [\ln(1 - (\rho_K)_{i,j,t,\Omega}^{\text{obs}}) - \ln(1 - (\rho_K)_{i,j,t,\Omega}^{\text{pred}})]^2 \quad (6)$$

for $K = 3$.

Revised adjoint code was manually written then compared with results obtained with an automatic differentiation library (AUTODIFF, Otter Research Ltd., 1994) and validated by a derivative check, that is, we verify [Equation (7)] that the discrepancy between each gradient component [obtained by analytic differentiation (adjoint code)] and its finite difference approximation changes parabolically with step h varying from 10^{-8} to 10^{-2}

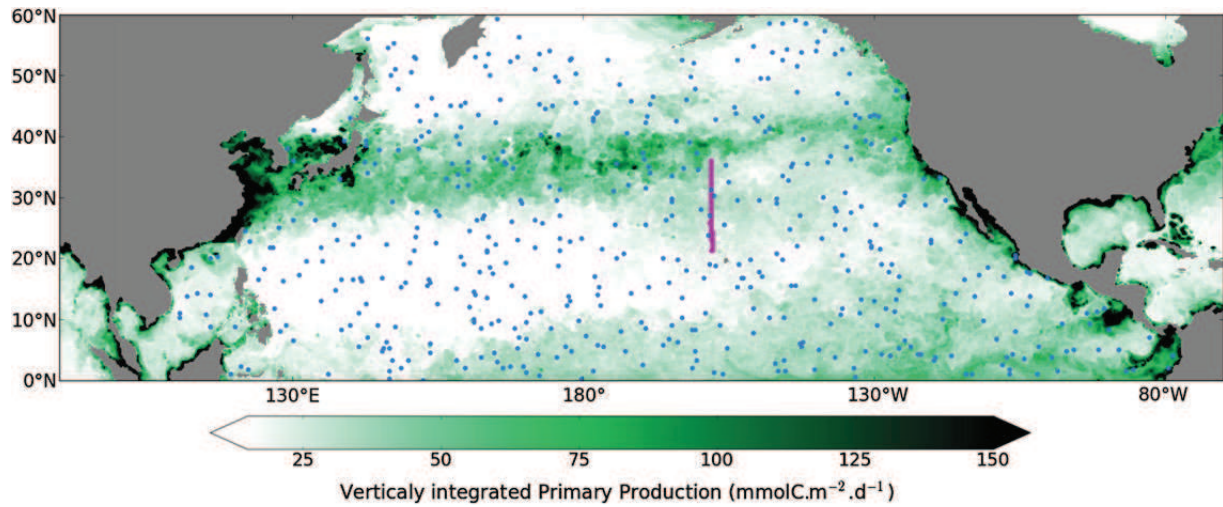


Figure 5. Distribution of pseudo-observations for twin experiments either along the Hawaiian transect (pink ticks) or distributed according a random sampling on position and period of day (blue dots), superimposed on primary production deduced from ocean colour data.

(Senina *et al.*, 2008).

$$\forall n, \frac{L^-(E'_n + h) - L^-(E'_n - h)}{2h} = \nabla_k L^- = O(h^2). \quad (7)$$

Each parameter E'_n is allowed to vary between 0 and 1 then rescaled to fulfil the condition of their sum being equal to 1 at the end of each likelihood function evaluation.

Finally, the general approach was validated with twin experiments. In such experiments, pseudo-observations are extracted from biomass outputs predicted from a run of the model (Figure 5) with an initial parameterization (P_{init}). They are used to verify that after changing the parameter values, the model can converge towards the original values of parameters. Once the convergence criterion is reached, the relative error to the exact value (e_n) is calculated by Equation (8):

$$e_n = \frac{|\hat{E}'_n - E'_n|}{E'}. \quad (8)$$

For a first series of five twin experiment simulations, we used pseudo-observations extracted along the Hawaiian transect. Then, 16 series of 5 replicate simulations were produced with pseudo-observations randomly distributed on 1, 3, 5, or 10 time steps of the simulation and 20, 50, 100, or 200 localizations by time step. Thus, a total of 85 twin experiment simulations were conducted with a range of data between a minimum of 20 and a maximum of 2000 observations. To compare the results, we first used the mean negative log-likelihood value of each group of five simulations. In addition, we also computed the relative standard deviation of the negative log-likelihood ($\sigma L^- / \bar{L}^-$) for each group of five replicate simulations to compare between these groups while accounting for their different averages ($\bar{L}^-$) and numbers of observations.

Results

Twin experiments

All twin experiments are run from an initial parameterization with all E'_n coefficients set to one-sixth, and with the same number of night and day pseudo-observation to avoid any potential bias.

Overall, the level of error on parameter estimates decreased with the number of pseudo-observations included in the experiment, from a maximum relative error of $\sim 10\%$ in the worst case with 20–50 pseudo-observations to $< 1\%$ in the best case with maximum spatial and temporal coverage. A minimum of 200 pseudo-observations appeared necessary to remain below the threshold of 5% error (Figure 6).

The response of the model showed a dissymmetry considering the number of pseudo-observations included either spatially or temporally. For the same number of pseudo-observations, it is better to have these data along the same transect but over several time steps than well dispersed over a single time step (Figure 6). Each group had its own response in the variance of estimated error, the most sensitive being the non-migrant upper mesopelagic group (2.2). The three groups that inhabit at least during nighttime the epipelagic layer, that is, the epipelagic (1.1), migrant upper mesopelagic (2.1), and highly migrant lower mesopelagic (3.1) groups were those that have the lowest level of error (Figure 6), followed by the non-migrant lower mesopelagic group (3.3) and finally the migrant lower mesopelagic (3.2) and the non-migrant upper mesopelagic (2.2) groups. These different responses could be related to more dynamic and contrasted signals in the epipelagic layer with higher temperature and stronger currents. This is certainly not the case for the non-migrant lower mesopelagic group, but this one has the advantage to stay alone in its layer at night with clear information on its biomass ratio. In the upper mesopelagic layer, the signal is less clear since it always results from a combination of two different groups and the turnover and dynamics are relatively slow. The result for the twin experiments imitating the actual Hawaiian transect, that is, with same number and localization of pseudo-observations, is shown in Figure 6 and confirms the general pattern with the largest error for the group 2.2 (7.14%) while it is $< 5\%$ for the others. The error on the coefficient of group 2.1 (4.64%) is slightly above what is expected.

The total negative log-likelihood increased with the number of pseudo-observations included in the twin experiments (Figure 7). This is not surprising since even with small errors, their sum increases with the number of observations. However, it increased homogeneously and symmetrically on both time and space axes.

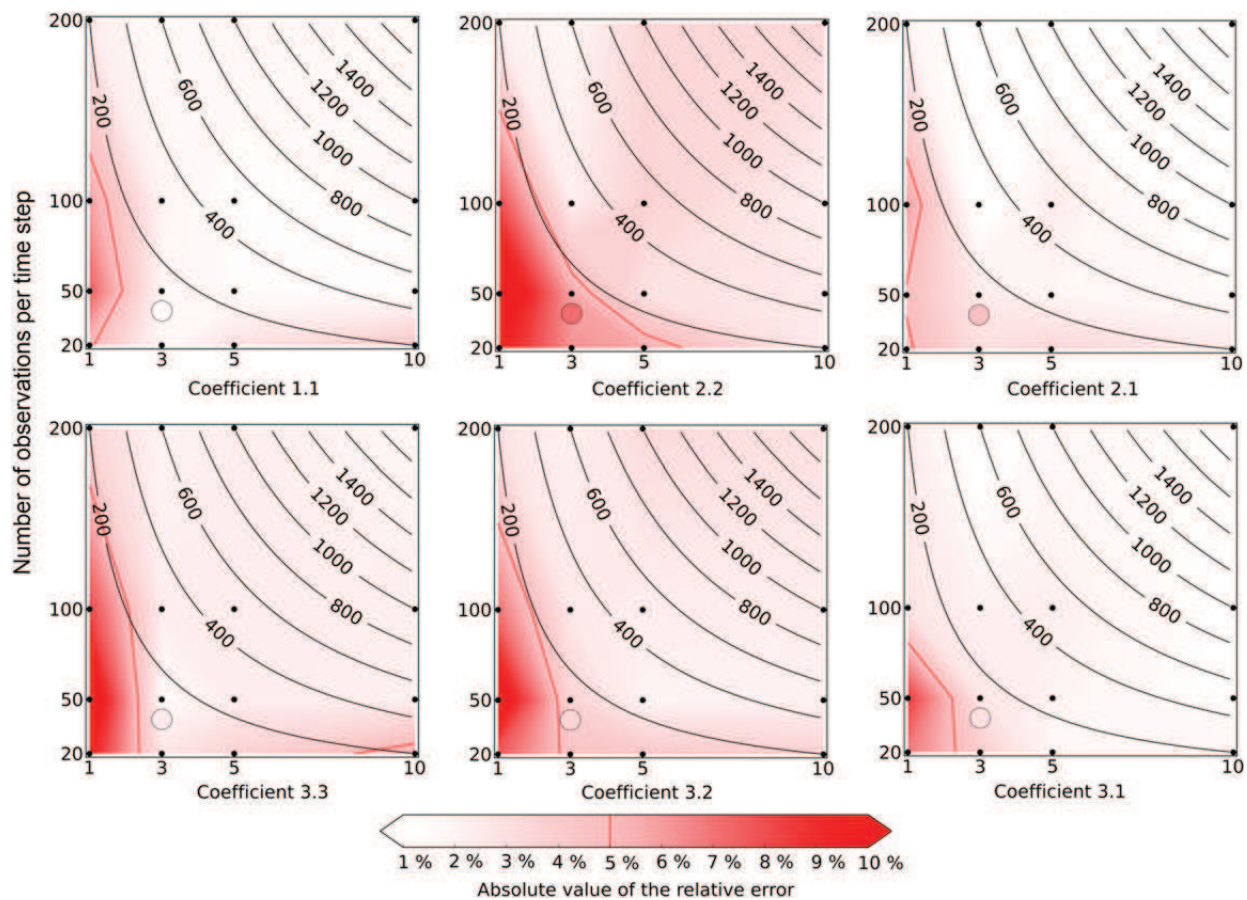


Figure 6. Mapping of the average absolute value of the relative error (e_n) for each coefficient E'_n in the retrieval of parameters in twin experiments (identified with black dots), over temporal (horizontal axis) and spatial resolution (vertical axis). Black curves show isolines with same numbers of pseudo-observations. Red lines delineate the 5% relative error area. Circles show the location of relative errors of twin experiments that mimic the actual Hawaiian transect.

The value for the samples that mimic the actual Hawaiian transect is below the expected level, possibly due to a favourable position of the transect crossing contrasted systems, that is, the tropical gyre and the Convergence Zone of Chlorophyll (Polovina *et al.* 2001).

The standard deviation of the negative log-likelihood once corrected to the total number of observations allowed for the analyses of the intra-samples variability independently from their size. This relative standard deviation showed the largest deviation ($>35\%$ of the log-likelihood) for small-sized samples (Figure 7). Again the amount of 200 observations appeared as a rough minimum threshold, a result also confirmed by the Bayesian information criterion (Akaike, 1974; Schwarz, 1978) that account for the number of observations and parameters (Figure 8), the latter being fixed in our case.

Estimated parameters with Hawaiian transect

At the resolution of the model, that is, $1/4^\circ \times \text{week}$, the transect used to run the real optimization experiment provided 116 observations over 3 consecutive time steps. The model successfully converged with a new parameterization with new coefficient values distributed over a larger range (0–0.435) than for the initial parameterization (Table 1). Coefficients of lower and highly migrant lower mesopelagic groups showed the largest increase with estimates

between boundary values. Conversely, estimates for the groups of the upper mesopelagic layer strongly decreased and even reached 0 for the non-migrant group (Table 1). The coefficient value of the non-migrant lower mesopelagic was also estimated at the boundary. Therefore, this first optimization is not entirely satisfying, possibly because of the small number of observations and the uncertainty on forcing variables (see also Discussion).

Nevertheless, although the new optimized parameterization should be considered very preliminary, it improved the average overall fit to observed ratios in the three layers (Figure 9). The low values of signal ratios in the upper mesopelagic layer and its weak variability may explain the difficulty to retrieve correct estimates of energy transfer coefficients for the groups inhabiting this layer. In addition, spatial and temporal resolution of the model still cannot predict the level of variability observed. The model also does not resolve the transition phase during twilight hours. Finally, other various sources of uncertainty that are inherent to the model and its forcing, as well as in the acoustic data, are detailed in Discussion.

However, one striking disagreement between observed and predicted ratios occurred at the beginning and end of the Hawaiian transect, that is, when leaving and returning to the Hawaiian Islands. The model could not predict the high values observed

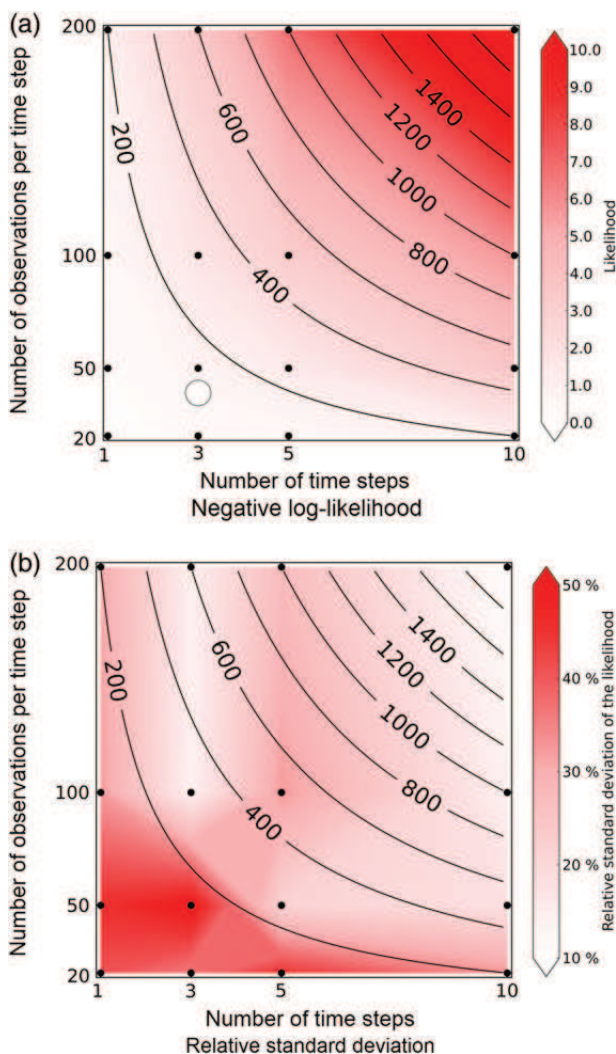


Figure 7. Mapping of the mean negative log-likelihood (a) and relative standard deviation of the negative log-likelihood (b), over temporal (horizontal axis), and spatial resolution (vertical axis). Black curves show isolines with same numbers of pseudo-observations and circle in (a) shows the location of the mean negative log-likelihood that mimics the actual Hawaiian transect.

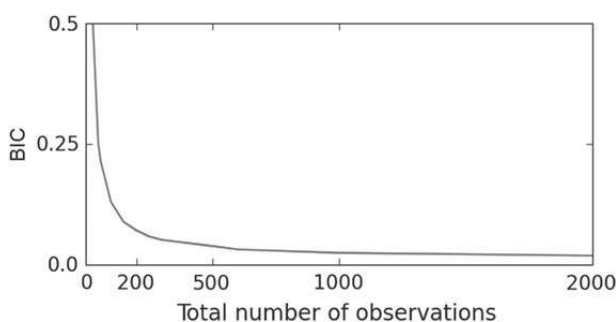


Figure 8. Bayesian information criterion computed for all twin experiments.

during the night in the epipelagic layer and, conversely, the low values in the lower mesopelagic layer. A more detailed investigation detected a potential problem associated with the mesoscale activity. Although the currents predicted with the GLORYS reanalysis compared fairly well with those deduced from altimetry data (Figure 10), the detailed mesoscale features might differ substantially locally. For example, the very high-ratio value in the epipelagic layer at night occurring at 25°N along the transect (Figure 9) seemed to coincide with a structure of two small eddies of opposite rotation that are visible on the altimetry map but not in the predicted currents of GLORYS (Figure 10). This type of structure associating one cyclonic and one anticyclonic mesoscale eddy is highly favourable to the concentration of organisms in the area of convergence of currents as predicted for instance in the southeast corner of the micronekton biomass distribution map on Figure 10d. The lack of such a structure at 25°N on the transect in predicted currents used to simulate the micronekton could explain the too low micronekton biomass predicted by the model in comparison with acoustic data.

Discussion

Acoustic sampling is the only available approach to collecting enough data at basin scales. At first glance, it shows micronekton distribution and density that is globally still missing today. However, even with a vast sampling effort, it could not offer the synoptic view that is required for understanding and eventually managing these huge oceanic ecosystems. Therefore, the data need to be complemented with ecological models with appropriate levels of detail and parameters tailored to match this type of data. This study provides a methodology for a simple and robust use of acoustic data in an original parsimonious model of micronekton functional groups.

Acoustic data initially allowed us to define the vertical boundaries of the three biological layers in the model. Such a definition is critical as physical variables that drive the dynamics of MTL groups are averaged within these layers. Ongoing work in other oceanic basins (e.g. the Southern Indian Ocean) suggests that this definition could be generalized at the ocean basin scale, although this question of vertical boundaries remains fully open and will require many other acoustic transects and research cruises from various oceanic regions and seasons to be confirmed.

Clearly, the initial optimal parameterization achieved here, with one single transect and single frequency is preliminary, and new experiments will have to be conducted with more data and acoustic interpretation. The approach based on relative signal ratios should help in combining different data sources, but this assumes no horizontal and vertical bias in the acoustic measurement and their biological interpretation. Reducing vertical and horizontal bias from the acoustic measurements can be achieved with modern calibrated echosounders from research and fishing vessels with the data available with appropriate metadata (Kloser *et al.*, 2009; www.imos.org.au; ICES, 2013). Thus, it is necessary to implement automatic data screening to remove obvious wrong signals and filtering so as to avoid spurious integration of noise (e.g. De Robertis and Higginbottom, 2007). It is also essential that the treatment of the signal along the vertical dimension be comparable with appropriate beam spreading compensation and absorption correction for the frequency used (Francois and Garrison, 1982). Adjustments need to be made for potential changes in transducer sensitivity and absorption in

Table 1. Matrix of the energy coefficients transfer used for the three-layer six-component MTL model according to the functional group and the number of vertical layers, and for different simulation experiments.

No of layer	Simulation experiment	Epi	Upper meso	Migrant upper meso	Lower meso	Migrant lower meso	Highly migrant lower meso
0 (land)	All	0	0	0	0	0	0
1	All	1	0	0	0	0	0
2	1	0.340	0.270	0.390	0	0	0
	Opt	0.463	0.310	0.227	0	0	0
3	1	0.170	0.100	0.220	0.180	0.130	0.200
	Opt	0.236	0.083	0.000	0.435	0.020	0.226

1, parameterization achieved from the literature (Lehodey et al., 2010a); Opt, optimization after data filtering and 5% of the biomass added in the epipelagic.

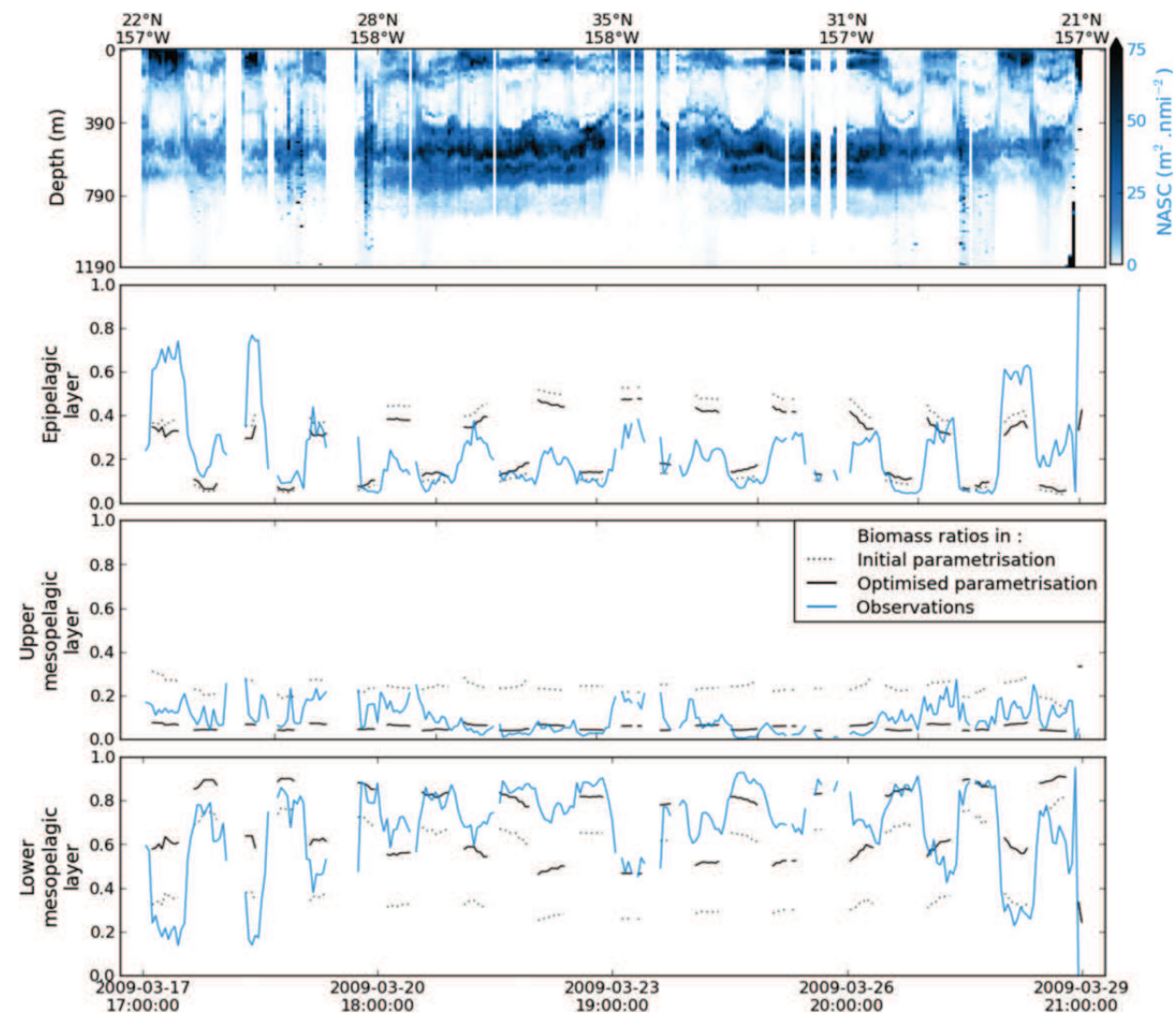


Figure 9. Comparison between observations and predictions. First panel shows the whole acoustic transect (averaged over a 1-h period) in NASC. The three other panels show biomass ratios in the corresponding layer (averaged over a 1-h period): blue lines for observed, black dotted lines for parameterization published in Lehodey et al. (2010a) and black solid lines for the new optimization.

the horizontal direction that are mainly a result of temperature (Demer and Renfree, 2008). This data processing method then enables a comparison of acoustic backscatter ratios between depths and regions. To compare datasets from different instruments and regions, the data metadata must describe the instrument, the calibration, and signal processing methods used (ICES, 2013).

This work has assumed that the biological interpretation of the acoustics contains no vertical or horizontal bias and that the acoustic signal is proportional to the micronekton biomass at all depths (Benoit-Bird and Au, 2002). This will not be the case at lower frequencies, such as 38 kHz used here, as a result of resonance scattering effects in the mesopelagic layer from fish and siphonophores

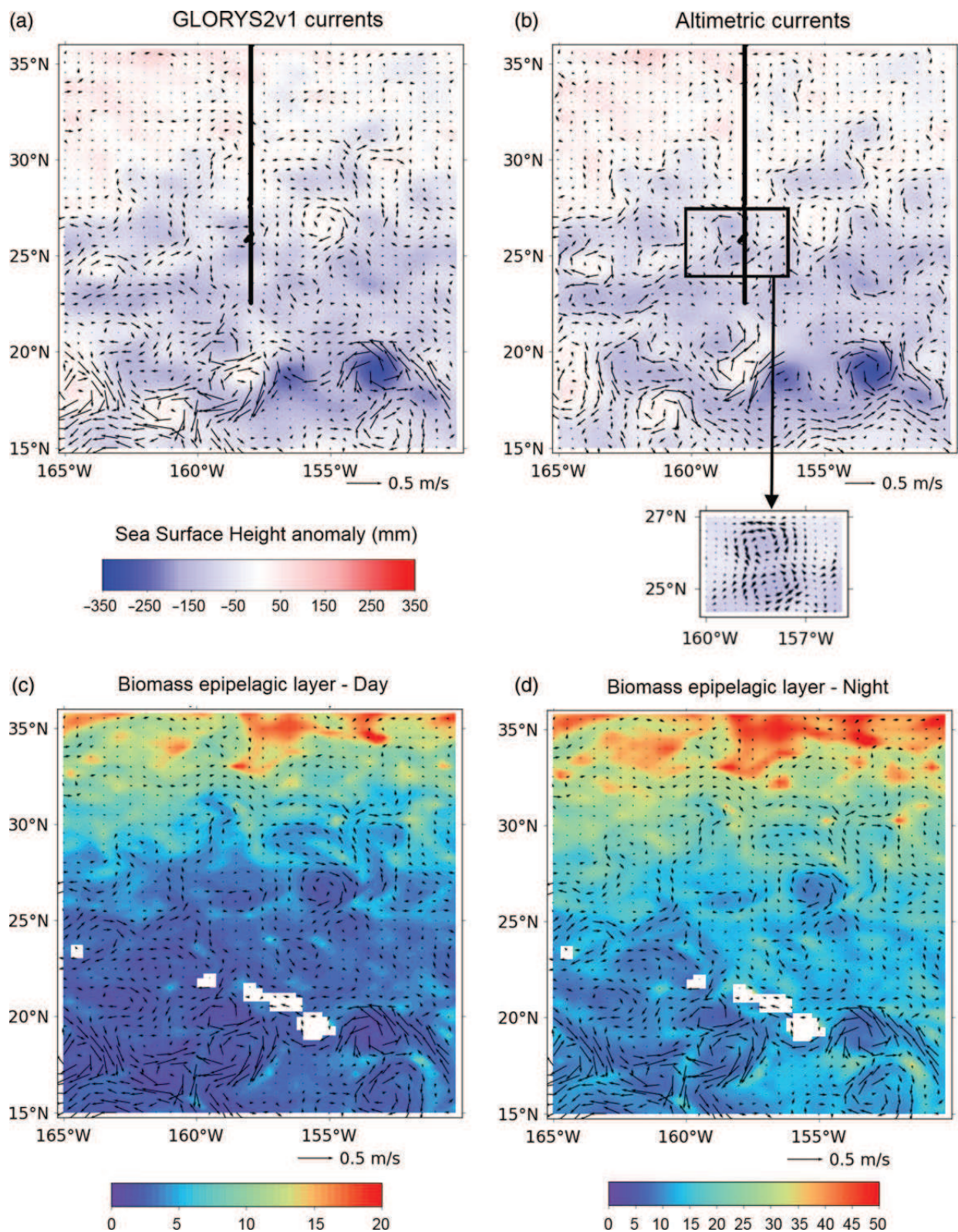


Figure 10. Mesoscale prediction and observation. (a) Current velocity (arrows) from GLORYS2v1 reanalysis (week of 16–22 March 2009) superimposed on observed sea surface height (SSH) anomaly (AVISO CNES/CLS, <http://www.aviso.oceanobs.com/>); (b) current velocity (arrows) deduced from altimetry for the same period and also superimposed on SSH. The acoustic transect (black line) is superimposed and a zoom is shown for the area identified with a rectangle. (c) Predicted biomass distribution in the epipelagic layer during day-time, and (d) during nighttime with superimposed GLORYS2v1 currents. Note the different ranges of values between colour bars.

with gas inclusions in particular. For example, at 600 m depth the backscatter at 38 kHz can be an order of magnitude >120 kHz frequency as used by Benoit-Bird and Au (2002) requiring a scaling adjustment (Kloser *et al.*, 2002). To accurately compare data from different layers and between different regions, the proper ratios of different types of micronektonic organisms and their sizes must be estimated, and adjustments must be made for differences in acoustic backscatter based on composition and size (Handegard *et al.*, 2013). This is true even if all micronektonic groups, that is, fish with and without gas bladder, crustaceans, cephalopods, and gelatinous organisms with and without gas inclusions (e.g. some siphonophores), are considered as a whole as in this study. Advances in multifrequency acoustics and detailed studies on the conversion of acoustic to biological units are required to narrow down the range of errors and biases based on resonance scattering. Finally, the calibration of diffusion parameter in the micronekton model needs also to be investigated through sensitivity analysis. Decreasing its value would certainly reinforce the contrast in the predicted biomass ratio but likely could make the model convergence more difficult to achieve.

Our estimate of epipelagic biomass may be biased low as a result of the near-surface region unsampled by the acoustics (~ 15 m depth) or affected by vessel avoidance (e.g. O'Driscoll *et al.*, 2009; De Robertis and Handegard, 2013). This bias may have a spatial and temporal structure and be region-specific depending on the species present. In this initial experiment, we arbitrarily fixed the missing near-surface biomass to 5% of the total signal, which is only a first-guess estimate. Based on 38 kHz acoustic data collected along a circumnavigation, Irigoien *et al.* (2014) give an estimate of 7% for the biomass in the upper 200 m after excluding five areas with exceptional high values. To explore the magnitude of this bias will require methodological and technological solutions. Using the existing echosounder data, it is possible to observe the migration behaviour of the species to determine the potential extent of the problem and the need for higher compensation or dedicated near-surface sampling technology (O'Driscoll *et al.*, 2009). Examples of available near-surface acoustic technology are upwards-looking echosounders (moored or mobile) and sideways-looking sonars (Handegard *et al.*, 2013). Estimating biomass from these devices is complicated by the need to know the target strength of the species at different orientations. At the least, these devices can detect the relative change of acoustic signal above and below the near-surface region to explore the potential magnitude of the bias. Sampling the near surface with acoustics also has some physical constraints due to the instruments' pulse length, waves, and bubbles. Uncertainty in biomass estimates from these issues will need to be addressed if precise epipelagic biomass estimates are to be obtained. Despite all these raised issues, observing the behaviour of the vertical migration of mesopelagic organisms could be the best method to highlight the potential magnitude of the problem to direct more research.

With the quality and coverage of acoustic data, forcing fields are another area for increasing the accuracy of the model in simulating the dynamics of micronekton functional groups. The twin experiments indicated that 200 observations could be sufficient to retrieve the correct parameter values. However, this is for the ideal and theoretical case of a perfect environmental forcing, that is, here the same used both for producing pseudo-observations and running optimization. In practice, there are many sources of uncertainties on the variables used to reproduce the oceanic environment.

The VGPM model of primary production that is used here is one among many satellite "chlorophyll-based" models with empirically determined functions that are generalized to basin scale. For instance, an alternative definition of the temperature-dependent photosynthetic efficiencies as suggested by Eppley (1972) leads to substantial differences in the estimates (see <http://www.science.oregonstate.edu/ocean.productivity/>). The accuracy of 21 ocean colour models was recently assessed by comparing their estimates with *in situ* measurements (Saba *et al.*, 2011). While on average, simple and more complex models had similar performances, their average accuracy was significantly higher at seabed depths >250 m, that is, for the Case 1 (pelagic) waters used in our study. Therefore, primary production estimates are likely degraded close to the coast and these errors can cascade downstream and propagate offshore.

Similarly, ocean-circulation models still have biases or drifts from observations despite the assimilation of satellite and available *in situ* data (Lellouche *et al.*, 2012; Ferry *et al.*, 2012). These sources of errors likely increase with depth in relation to the number of observations available for assimilation, and also near the coast due to the necessity of degrading the actual topography at the resolution of the model grid ($1/4^\circ$). The discrepancy observed in this study between mesoscale features predicted by the reanalysis or deduced from altimetry is a good illustration of these potential sources of errors with a direct consequence on the micronekton estimate. However, it is worth noting that the optimization procedure improved the average estimate independently of these local errors. This suggests that the method is robust and can account for these sources of errors.

The next step will consist in optimization experiments based on longer time series and with a larger amount of observations needed to compensate the various sources of uncertainties. Given the effort developed with the IMOS initiative (Kloser *et al.*, 2009; <http://www.imos.org.au/>) to collect 38 kHz and other acoustic frequency data routinely, it is envisaged to develop a configuration for the southern hemisphere and use all these available and standardized data. New fine-scale acoustic, optical, and trawling regional-based experiments are needed to interpret this broader spatial and temporal scale acoustics data to initialize and assimilate into the model (Handegard *et al.*, 2013). In particular, experiments to quantify the conversion of acoustic backscatter into biological units accounting for resonance scattering are needed. The optimal parameterization will be then used for a global hindcast simulation, allowing evaluation of the results in the northern hemisphere with historical data, for example, the MARECO platform. It is also worth noting that our twin experiments suggest that a long time series of acoustic data in a single point could provide more useful information for parameter optimization than the same amount of data spatially dispersed over a few time steps. It is essential to confirm or not this result given its potential consequences when designing ocean-monitoring networks.

Parallel with the progress of acoustic technology and associated validation experiments for quantifying and discriminating between the major groups of micronekton organisms, the model could be further developed easily to account for a more detailed description of the pelagic mid-trophic levels.

Acknowledgements

The authors thank the Groupe Mission Mercator Coriolis and the GLORYS members (MERCATOR-OCEAN, INSU-CNRS and the team "DRAKKAR") for providing the GLORYS-2v1 reanalysis. We also thank the Ocean Productivity team for providing us the

SeaWiFS-derived primary production. This study was partially supported by a grant from the US JIMAR Pelagic Fisheries Research Program and by the FP7 EU Euro-BASIN research project. We are also grateful to two anonymous reviewers who provided useful comments to improve a first version of the manuscript.

References

- Abecassis, M., Senina, I., Lehodey, P., Gaspar, P., Parker, D., Balazs, G., and Polovina, J. 2013. A Model of Loggerhead Sea Turtle (*Caretta caretta*) habitat and movement in the oceanic North Pacific. *PLoS ONE*, 8: e73274.
- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19: 716–723.
- Barnier, B., Madec, G., Penduff, T., Molines, J. -M., Treguier, A. -M., le Sommer, J., Beckmann, A., *et al.* 2006. Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy permitting resolution. *Ocean Dynamics*. doi:10.1007/s10236-006-0082-1
- Behrenfeld, M. J., and Falkowski, P. G. 1997. A consumer's guide to phytoplankton primary productivity models. *Limnology and Oceanography*, 42: 1479–1491.
- Benoit-Bird, K. J., and Au, W. W. L. 2002. Energy: converting from acoustic to biological resource units. *Journal of the Acoustical Society of America*, 111: 2070–2075.
- Brasseur, P., Gruber, N., Barciela, R., Brander, K., Doron, M., El Moussaoui, A., Hobday, A., *et al.* 2009. Integrating biogeochemistry and ecology into ocean data assimilation systems. *Oceanography*, 22: 206–215.
- Briand, K., Molony, B., and Lehodey, P. 2011. A study on the variability of albacore (*Thunnus alalunga*) longline catch rates in the south-west Pacific Ocean. *Fisheries Oceanography*, 20: 517–529.
- Demer, D. A., and Renfree, J. S. 2008. Variations in echosounder-transducer performance with water temperature. *ICES Journal of Marine Science*, 65: 1021–1035.
- De Robertis, A., and Higginbottom, I. 2007. A post-processing technique to estimate the signal-to-noise ratio and remove echosounder background noise. *ICES Journal of Marine Science*, 64: 1282–1291.
- De Robertis, A., and Handegard, N. O. 2013. Fish avoidance of research vessels and the efficacy of noise-reduced vessels: a review. *ICES Journal of Marine Science*, 70: 34–45.
- de Young, B., Heath, M., Werner, F., Chai, F., Megrey, B., and Monfray, P. 2004. Challenges of modeling ocean basin ecosystems. *Science*, 304: 1463–1466.
- Downey, E. C. 1990. XEphem Version 3.7.6 Reference Manual. <http://www.clearskyinstitute.com/xephem/> (last accessed 12 December 2014).
- Dueri, S., Faugeras, B., and Maury, O. 2012. Modelling the skipjack tuna dynamics in the Indian Ocean with APECOSM-E: Part 1. Model formulation. *Ecological Modelling*, 245: 41–54.
- Eppley, R. W. 1972. Temperature and phytoplankton growth in the sea. *Fishery Bulletin*, 70: 1063–1085.
- Ferry, N., Parent, L., Garric, G., Drevillon, M., Desportes, C., Bricaud, C., and Hernandez, F. 2012. Scientific validation report (ScVR) for reprocessed analysis and reanalysis. MyOcean project report, MYO-WP04-ScCV-rea-MERCATOR-V1.0. 66 pp. <http://www.mercator-ocean.fr/fre/Media/Files/GLORYS/MYO-WP4-ScVR-rea-MERCATOR-V1.0.pdf> (last accessed 19 December 2014).
- Foote, K. G., Knudsen, H. P., Vestnes, G., MacLennan, D. N., and Simmonds, E. J. 1987. Calibration of acoustic instruments for fish density estimation: a practical guide. *ICES Cooperative Research Report*, 144, 69 pp.
- Francois, R. E., and Garrison, G. R. 1982. Sound absorption based on ocean measurements. Part II: Boric acid contribution and equation for total absorption. *The Journal of the Acoustical Society of America*, 72, 1879, <http://dx.doi.org/10.1121/1.388673> (last accessed 19 December 2014).
- Fulton, E. A. 2010. Approaches to end-to-end ecosystem models. *Journal of Marine Systems*, 81: 171–183.
- Grandperrin, R. 1975. Structures trophiques aboutissant aux thons de longue ligne dans le Pacifique Sud-Ouest. Thèse de doctorat de l'Université d'Aix-Marseille II. 295 pp.
- Handegard, N. O., du Buisson, L., Brehmer, P., Chalmers, S. J., De Robertis, A., Huse, G., Kloser, R. *et al.* 2013. Towards an acoustic-based coupled observation and modelling system for monitoring and predicting ecosystem dynamics of the open ocean (2012). *Fish and Fisheries*, 14: 605–615.
- Holt, J., Allen, J. I., Anderson, T. R., Brewin, R., Butenschon, M., Harle, J., Huse, G., *et al.* 2014. Challenges in integrative approaches to modelling the marine ecosystems of the North Atlantic: physics to fish and coasts to Ocean. *Progress in Oceanography*. doi: 10.1016/j.pocean.2014.04.024.
- ICES 2013. A metadata convention for processed acoustic data from active acoustic systems, SISP 3 TG-AcMeta, ICES GFAST Topic Group, TG-AcMeta. 35pp.
- Irgoien, X., Klevjer, T. A., Rostad, A., Martinez, U., Boyra, G., Acuña, J. L., Bode, A., *et al.* 2014. Large mesopelagic fishes biomass and trophic efficiency in the open ocean. *Nature Communications*, 5. doi:10.1038/ncomms4271
- Jennings, S., Mélin, F., Blanchard, J. L., Forster, R. M., Dulvy, N. K., and Wilson, R. W. 2008. Global-scale predictions of community and ecosystem properties from simple ecological theory. *Proceedings of the Royal Society B*, 275: 1375–1383.
- Kitchell, J. F., Boggs, C. H., He, X., and Walters, C. J. 1999. Keystone predators in the Central Pacific. *Proceedings of the Symposium on Ecosystem Considerations in Fisheries Management*, Sep 30–Oct 3, Anchorage, Alaska, USA. Alaska Sea Grant College Program. AK-SG-99-01: pp. 665–68.
- Kloser, R. J., Ryan, T., Sakov, P., Williams, A., and Koslow, J. A. 2002. Species identification in deep water using multiple acoustic frequencies. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 1065–1107.
- Kloser, R. J., Ryan, T. E., Young, J. W., and Lewis, M. E. 2009. Ocean-basin scale acoustic observations of micronekton fishes: potential and challenges. *ICES Journal of Marine Science*, 66: 998–1006.
- Lambert, C., Mannocci, L., Lehodey, P., and Ridoux, V. 2014. Predicting cetacean habitats from their energetic needs and the distribution of their prey in two contrasted tropical regions. *PLoS ONE*, 9: e105958.
- Lehodey, P., Murtugudde, R., and Senina, I. 2010a. Bridging the gap from ocean models to population dynamics of large marine predators: a model of mid-trophic functional groups. *Progress in Oceanography*, 84: 69–84.
- Lehodey, P., Senina, I., Calmettes, B., Hampton, J., and Nicol, S. 2013. Modelling the impact of climate change on Pacific skipjack tuna population and fisheries. *Climatic Change*, 119: 95–109.
- Lehodey, P., Senina, I., and Murtugudde, R. 2008. A spatial ecosystem and populations dynamics model (SEAPODYM) – modelling of tuna and tuna-like populations. *Progress in Oceanography*, 78: 304–318.
- Lehodey, P., Senina, I., Sibert, J., Bopp, L., Calmettes, B., Hampton, J., and Murtugudde, R. 2010b. Preliminary forecasts of population trends for Pacific bigeye tuna under the A2 IPCC scenario. *Progress in Oceanography*, 86: 302–315.
- Lellouche, J.-M., Le Galloudec, O., Drévillon, M., Régnier, C., Greiner, E., Garric, G., Ferry, N., *et al.* 2012. Evaluation of real time and future global monitoring and forecasting systems at Mercator Océan. *Ocean Science Discussions*, 9: 1123–1185.
- Mac Lennan, D. N., Fernandes, P. G., and Dalen, J. 2002. A consistent approach to definitions and symbols in fisheries acoustics. *ICES Journal of Marine Science*, 59: 365–369.
- Maury, O. 2010. An overview of APECOSM, a spatialized mass balanced “apex predators ECOSystem model” to study physiologically structured tuna population dynamics in their ecosystem. In: John, M., Monfray, P. St. (Eds.), *Parameterisation of Trophic Interactions in Ecosystem Modelling*. *Progress in Oceanography*, 84: 113–117.

- O'Driscoll, R. L., Gauthier, S., and Devine, J. A. 2009. Acoustic estimates of mesopelagic fish: as clear as day and night?. *ICES Journal of Marine Science*, 66: 1310–1317.
- Pham, D. T., Verron, J., and Roubaud, M. C. 1998. A singular evolutive extended Kalman filter for data assimilation in oceanography. *Journal of Marine Systems*, 16: 323–340.
- Plagányi, É. E. 2007. Models for an Ecosystem Approach to Fisheries. FAO Fisheries Technical Paper, No. 477. FAO, Rome, 108 p.
- Polovina, J. J., Howell, E., Kobayashi, D. R., and Seki, M. P. 2001. The transition zone chlorophyll front, a dynamic global feature defining migration and forage habitat for marine resources. *Progress in Oceanography*, 49: 469–483.
- Saba, V. S., Friedrichs, M. A. M., Antoine, D., Armstrong, R. A., Asanuma, I., Behrenfeld, M. J., Ciotti, A. M., *et al.* 2011. An evaluation of ocean color model estimates of marine primary productivity in coastal and pelagic regions across the globe. *Biogeosciences*, 8: 489–503.
- Schwarz, G. E. 1978. Estimating the dimension of a model. *Annals of Statistics*, 6: 461–464.
- Senina, I., Sibert, J., and Lehodey, P. 2008. Parameter estimation for basin-scale ecosystem-linked population models of large pelagic predators: application to skipjack tuna. *Progress in Oceanography*, 78: 319–335.
- Sibert, J., Senina, I., Lehodey, P., and Hampton, J. 2012. Shifting from marine reserves to maritime zoning for conservation of Pacific bigeye tuna (*Thunnus obesus*). *Proceedings of the National Academy of Sciences*, 109: 18221–18225.
- Tranchant, B., Testut, C.-E., Renault, L., Ferry, N., Birol, F., and Brasseur, P. 2008. Expected impact of the future SMOS and Aquarius Ocean surface salinity missions in the Mercator Ocean operational systems: new perspectives to monitor the ocean circulation. *Remote Sensing of Environment*, 112: 1476–1487.

Handling editor: Richard O'Driscoll

IV Un nouveau jeu de données plus étendu : résultats de l'assimilation de données en global, nouveaux défis

A Jeu de données global

Diverses collaborations et la mise en place de bases de données acoustiques ont permis d'enrichir notre jeu de données de façon conséquente. La figure 3.3 représente la distribution spatiale des données. Les trois bassins océaniques sont échantillonnés, bien que de manière peu homogène. Deux zones présentent une densité importante d'observations, la mer de Tasman et l'océan Indien Austral. Quelques transects échantillonnent également le nord du Pacifique entre le nord d'Hawaii et les Iles Samoa Américaines. Quant à l'Atlantique, nous disposons d'une unique série temporelle de 11 mois, très détaillée, à moyenne latitude dans l'hémisphère nord.

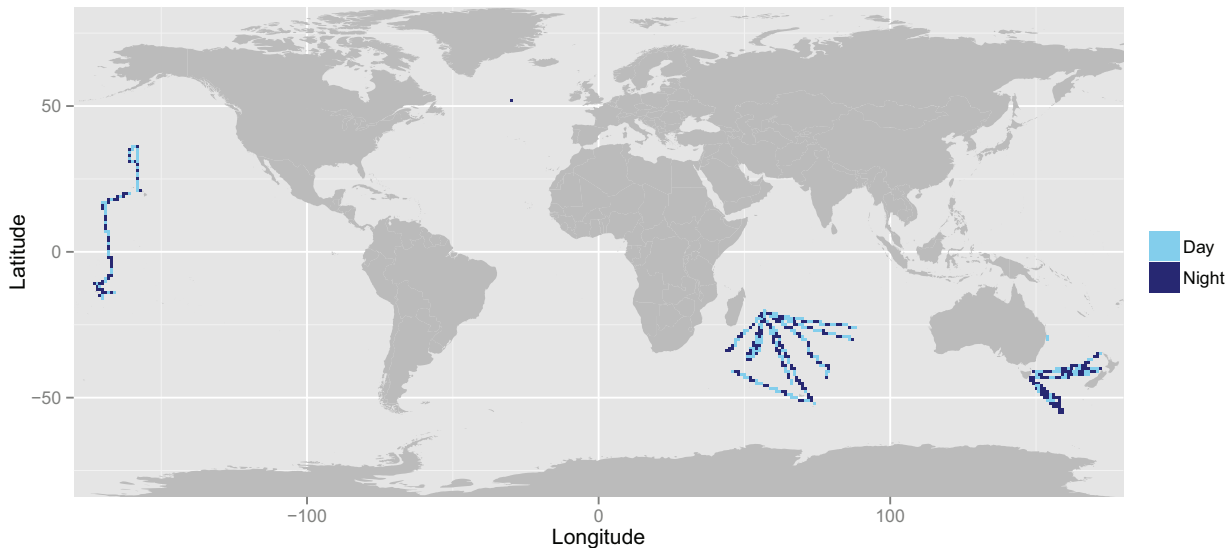


Figure 3.3 – Positions des données utilisées dans les expériences de cette section. Les positions sont moyennées à la résolution utilisée ($1^\circ \times 1$ mois) et colorées en fonction de la période du jour (jour ou nuit).

Ce jeu de données dispose d'un nombre suffisant d'observations d'après les expériences jumelles effectuées par Lehodey et al. (2015) ($n \gg 200$). Néanmoins il n'est pas réparti de façon homogène dans l'espace. De plus, les expériences jumelles conduites dans l'étude précédente utilisent le même forçage océanographique pour produire les observations synthétiques et pour estimer les paramètres. Dans la réalité, les modèles de circulation de l'océan et de production primaire basée sur les données satellites sont encore assez loin de fournir des prédictions identiques aux conditions réelles. Les nombreuses sources de bruits s'ajoutant au signal acoustique ont également été ignorées. Il est donc très probable que l'estimation optimale des paramètres demande au final beaucoup plus de données que le seuil établi par les expériences jumelles pour compenser l'accumulation de toutes ces sources d'incertitudes. Les valeurs d'erreur utilisées dans la section suivante sont donc des valeurs minimales.

A l'aide de ce jeu de données, nous souhaitons établir une paramétrisation de la distribution d'énergie entre les groupes fonctionnels (paramètres E'_i , $i \in [1, 6]$). On teste ici l'hypothèse que cette distribution est uniforme spatialement et temporellement.

B Expériences d'assimilation de données en global

On utilise ici les sorties du modèle de circulation océanique ECCO entre 1998 et 2012 interpolées sur une grille régulière de 1°x1mois comme champs de forçage du modèle (courants et températures). La production primaire et la profondeur euphotique sont issues du modèle VGPM à partir de données satellitaires de couleur de l'eau. La fonction coût et le pré-traitement des observations sont identiques au cas développé dans Lehodey et al. (2015).

Pour tester l'uniformité spatiale et temporelle des coefficients de distribution d'énergie, nous allons réaliser deux expériences afin d'évaluer les effets interannuels et la différence entre les bassins océaniques. Le tableau 3.1 résume les expériences effectuées.

Expérience n°	Zone	Saison	Année	Séparation
1	Tout	Tout	Tout	Zone
2	Mer de Tasman	Hiver austral	2009 à 2011	Année

Table 3.1 – Récapitulatif des expériences et de leurs caractéristiques.

Expérience n°1

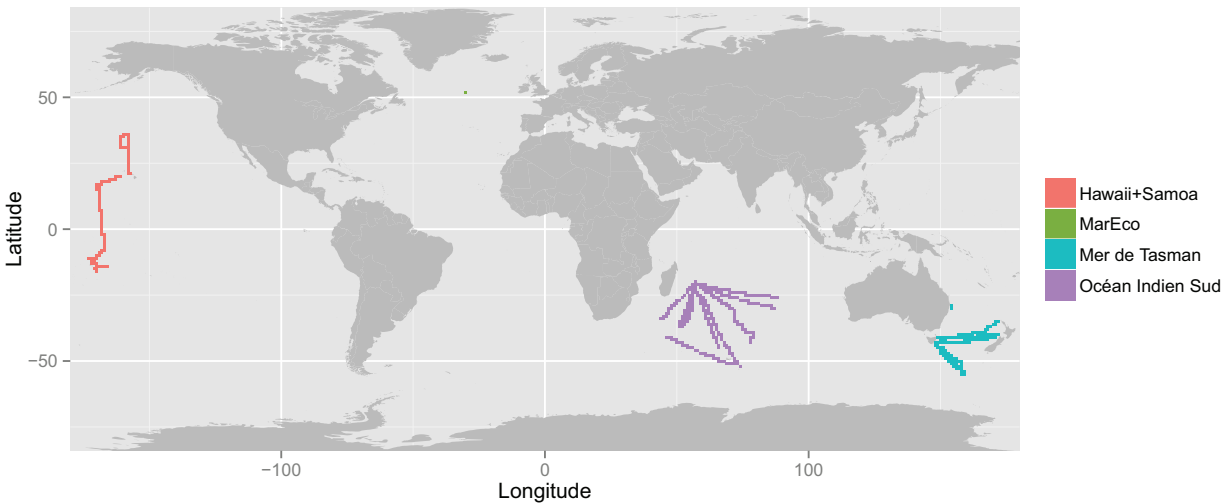


Figure 3.4 – Positions des observations colorées en fonction de leur zone.

Cette expérience vise à déterminer s'il y a une différence de prédiction lorsqu'on utilise uniquement les données d'un bassin océanique. Une expérience de référence est faite avec l'ensemble du jeu de données,

puis une prédiction par zone.

Le nombre d'observations varie d'une expérience à l'autre, le tableau 3.2 référence les nombres d'observations, de pas de temps et les coefficients dont la prédiction est associée à une forte erreur (à partir des expériences jumelles de [Lehodey et al. \(2015\)](#), et plus particulièrement la figure 6 de [Lehodey et al. \(2015\)](#)). Nous nous intéresserons donc aux zones 'Hawaii+Samoa', 'Mer de Tasman' et 'Océan Indien Sud' puisque les prédictions de ces zones sont suffisamment précises.

Zone	Nombre total d'observations	Nombre de pas de temps	Nombre d'observations moyen par pas de temps	Coefficient avec erreur sur la prédiction supérieure à 5%
Globe	906	27	34	Aucun
Hawaii+Samoa	164	32	5	Aucun
Mer de Tasman	336	42	8	Aucun
Océan Indien Sud	386	32	12	Aucun
MarEco	21	2	11	Tous

Table 3.2 – Expérience n°1 : nombre d'observations, total et par pas de temps, nombre de pas de temps, et erreurs associées aux prédictions.

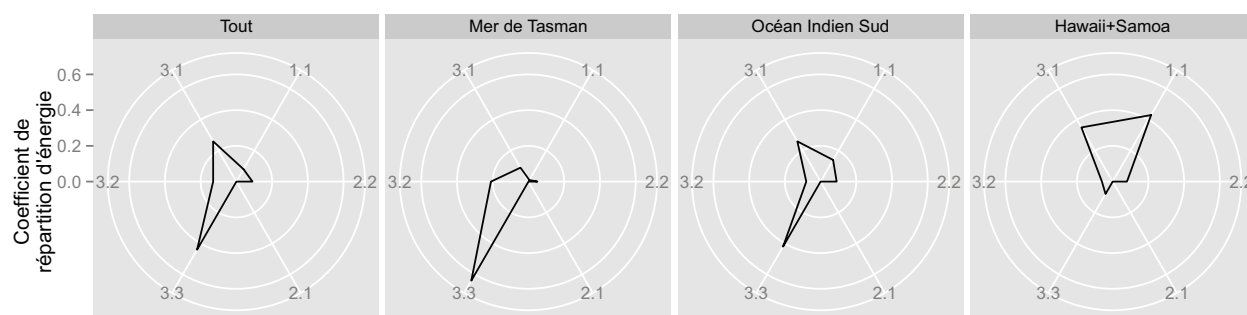


Figure 3.5 – Expérience n°1 : Diagramme de distribution de l'énergie entre les groupes fonctionnels (n.m où n est le numéro de la couche occupée durant la journée et m celui de celle occupée la nuit). Chaque graphique en radar présente les résultats d'une expérience. La valeur d'un coefficient est reportée sur la branche du graphique correspondant.

On observe une grande disparité dans les prédictions, notamment la zone 'Hawaii+Samoa' qui présente une distribution d'énergie axée sur le groupe non-migrant épipélagique (1.1) et le groupe hautement migrant mésopélagique inférieur (3.1), tandis que les zones 'Océan Indien Sud' et 'Mer de Tasman' proposent une distribution axée sur le groupe non-migrant mésopélagique inférieur (3.3). Ces deux dernières zones ne sont pas identiques non plus, l'une privilégiant en second plan le groupe hautement migrant mésopélagique inférieur (3.1) et l'autre le groupe migrant mésopélagique inférieur (3.2).

Expérience n°2

Dans cette expérience, on se focalise sur une zone restreinte, la Mer de Tasman. Les données y sont acquises principalement en hiver austral (et en automne). On se concentre donc ici sur les mois de mai à août. Le tableau 3.3 rassemble les nombres d'observations, de pas de temps et les coefficients plus entachés d'erreur des différents sous-jeux de données. On ne considèrera pas l'année 2012 puisqu'elle présente un nombre trop faible de données.

Année	Nombre total d'observations	Nombre de pas de temps	Nombre d'observations moyen par pas de temps	Coefficient avec erreur sur la prédiction supérieure à 10%
Tout	284	8	36	Aucun
2009	75	1	75	2.2 3.3 3.2
2010	73	3	24	Aucun
2011	128	2	64	2.2 3.3
2012	8	1	8	Tous

Table 3.3 – Expérience n°2 : nombre d'observations total et par pas de temps, nombre de pas de temps, et erreurs associées aux prédictions.

Les prédictions issues des différents jeux de données sont très proches. Les différences avec la prédiction de référence sont toutes inférieures aux erreurs trouvées avec les données de [Lehodey et al. \(2015\)](#).

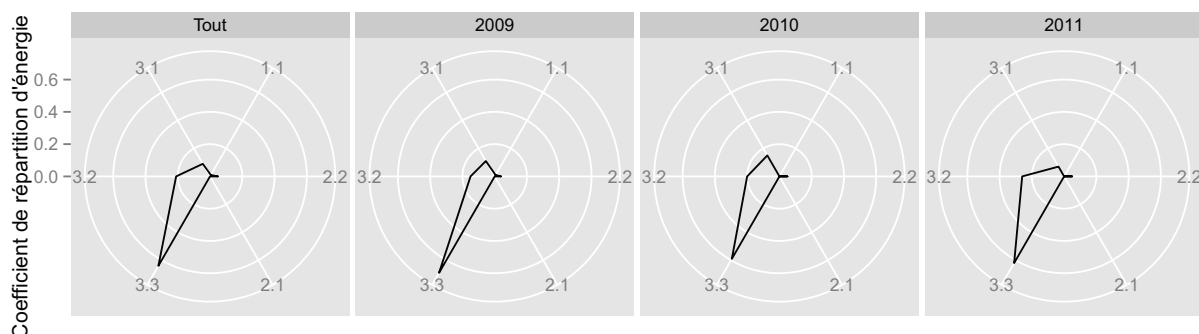


Figure 3.6 – Expérience n°2 : Diagramme de distribution de l'énergie entre les groupes fonctionnels (n.m où n est le numéro de la couche occupée durant la journée et m celui de celle occupée la nuit). Chaque graphique en radar présente les résultats d'une expérience. La valeur d'un coefficient est reportée sur la branche du graphique correspondant.

C Conclusions

On observe une différence forte dans les prédictions d'une zone à l'autre mais pas de variations interannuelles à l'intérieur d'une même zone. Notre jeu de données ne permet pas d'étudier les variations saisonnières à l'intérieur d'une même zone sur une même année tout en conservant de faibles valeurs d'erreurs sur les prédictions.

Chaque zone ne représente pas un écosystème mais une juxtaposition de différents écosystèmes. Nous ne pouvons pas, avec ces expériences, déterminer des paramétrisations par écosystème.

Ces résultats ne permettent pas de conclure sur l'uniformité spatio-temporelle de ces coefficients de distribution. Afin de conclure sur cette uniformité, il faut écarter les autres causes de déviations des coefficients prédits entre les zones. Nous pouvons identifier deux axes majeurs de réflexion : les valeurs des paramètres non-estimés et le pré-traitement des champs de forçages.

Les paramètres non-estimés sont au nombre de 5 : E le paramètre de transfert d'énergie depuis la production primaire, $1/\lambda_{curv}$ et $1/\lambda_0$ les paramètres de la fonction du taux de mortalité et τ_{curv} et τ_0 les paramètres de la fonction du temps de recrutement. Les prédictions ne sont pas sensibles aux variations de E du fait de l'expression de la fonction coût, ce qui n'est pas le cas des autres paramètres. Les figures 3.1 et 3.2 illustrent les incertitudes sur ces paramètres. De plus, il est possible d'optimiser ces paramètres à l'aide d'une méthodologie similaire à celle développée dans ce chapitre.

Lors du prétraitement des champs de forçages, les données d'entrée du modèle sont moyennées sur trois couches (eq intro). Ces couches représentent les profondeurs des couches de micronecton observées par acoustique. Leur définition est critique car elle peut fortement influencer les prédictions du modèle par le biais des champs de température et de courants moyennés verticalement. La définition utilisée ici a été paramétrée manuellement avec un nombre restreint de transects. Afin d'améliorer les prédictions des paramètres du modèle, cette définition doit être établie aussi précisément que possible avant d'optimiser les paramètres non-optimisés. Le Chapitre 4 s'y attelle.

Bibliographie

- Bertrand, A., Le Borgne, R., and Josse, E. (1999). Acoustic characterization of micronekton distribution in french polynesia. *Mar. Ecol. Prog. Ser.*, 191 :127–140.
- Gillooly, J. F., Charnov, E. L., West, G. B., Savage, V. M., and Brown, J. H. (2002). Effects of size and temperature on developmental time. *Nature*, 417(6884) :70–73.
- Iverson, R. L. (1990). Control of marine fish production. *Limnology and Oceanography*, 35(7) :1593–1604.
- Jennings, S., Warr, K. J., and Mackinson, S. (2002). Use of size-based production and stable isotope analyses to predict trophic transfer efficiencies and predator-prey body mass ratios in food webs. *Marine Ecology Progress Series*, 240(1).
- Kloser, R. J., Ryan, T. E., Young, J. W., and Lewis, M. E. (2009). Acoustic observations of micronekton fish on the scale of an ocean basin : potential and challenges. *ICES Journal of Marine Science : Journal du Conseil*, page fsp077.
- Lehodey, P., Conchon, A., Senina, I., Domokos, R., Calmettes, B., Jouanno, J., Hernandez, O., and Kloser, R. (2015). Optimization of a micronekton model with acoustic data. *Ices Journal of Marine Science*, 72(5) :1399–1412.
- Lehodey, P., Murtugudde, R., and Senina, I. (2010). Bridging the gap from ocean models to population dynamics of large marine predators : A model of mid-trophic functional groups. *Progress in Oceanography*, 84(1-2) :69 – 84. Special Issue : Parameterisation of Trophic Interactions in Ecosystem Modelling.
- McClatchie, S. and Dunford, A. (2003). Estimated biomass of vertically migrating mesopelagic fish off new zealand. *Deep Sea Research Part I : Oceanographic Research Papers*, 50(10) :1263–1281.
- Ryther, J. H. et al. (1969). Photosynthesis and fish production in the sea. the production of organic matter and its conversion to higher forms of life vary throughout the world ocean. *American Association for the Advancement of Science. Science*, 166 :72–76.
- Schaefer, M. B. (1965). The potential harvest of the sea. *Transactions of the American Fisheries Society*, 94(2) :123–128.

Chapitre 4 Définir les limites des couches biologiques dans SEAPOODYM

Sommaire

I	Développement d'un Average Micronektonic Layer Model	81
A	Les limites verticales des couches de SEAPOODYM	81
B	Principe de la méthode d'extraction des limites de couches	83
C	Définition d'un modèle pour les limites verticales des couches de SEAPOODYM-MTL	84
II	Modeling depths of ocean bio-acoustic layers from space	84
III	Une nouvelle définition des couches	105
A	Résultats de la méthode de détection des couches	105
B	Impact sur l'optimisation	106

I Développement d'un Average Micronektonic Layer Model

A Les limites verticales des couches de SEAPOODYM

Comme on l'a vu dans le chapitre précédent, les résultats des expériences d'assimilations de données sont sensibles à la définition de la profondeur des couches de SEAPOODYM-MTL. Afin de pouvoir conclure sur ces résultats, il est nécessaire d'étudier plus en détails cette définition. La définition des profondeurs des couches du modèle SEAPOODYM a évolué au cours de son développement. D'abord constantes (200 m, 400 m et 800 m), ces profondeurs ont ensuite été définies par des multiples de la profondeur euphotique (Z_{eu}). En effet, la lumière étant reconnue comme cause la plus probable des migrations nyctémérales, un proxy de cette dernière est un bon candidat pour qualifier la profondeur des organismes. Une première paramétrisation (Z_{eu} , $3 \times Z_{eu}$, $7 \times Z_{eu}$) fut affinée par la modification des multiples ($1.5 \times Z_{eu}$, $4.5 \times Z_{eu}$, $10.5 \times Z_{eu}$) par Lehodey et al. (2014). Néanmoins, cette définition fut obtenue manuellement sans analyse statistique. De plus, lors de ce choix nous ne disposions que de peu de données.

La figure 4.1 illustre l'adéquation entre un transect acoustiques et la définition des couches utilisées dans le chapitre précédent. Celle-ci ne semble pas idéale, notamment pour la 3ème couche, mais il n'est pas aisé de juger à partir de telles données. Nous avons donc décider de tout d'abord extraire des données acoustiques l'information pertinente, c'est-à-dire la limite de chaque couche observées, puis de quantifier

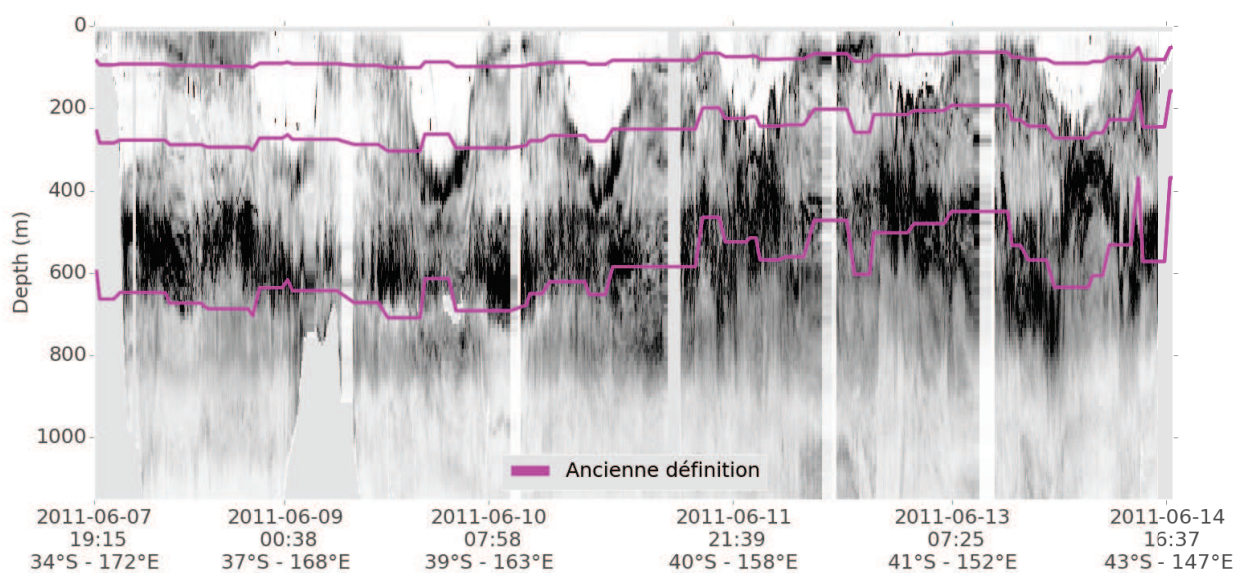


Figure 4.1 – Exemple d'un transect acoustique entre la Nouvelle-Zélande et la Tasmanie

l'écart à celle-ci de la définition actuelle et enfin de proposer une définition plus adaptée.

B Principe de la méthode d'extraction des limites de couches

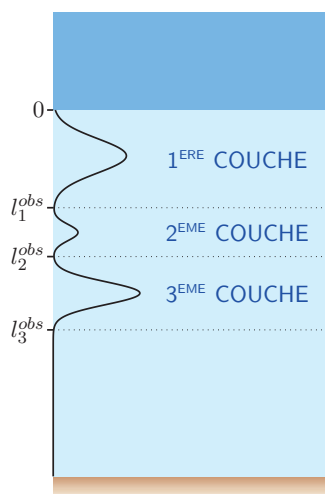


Figure 4.2 – Représentation d'un profil acoustique et des profondeurs l_i^{obs} des limites des différentes couches i

Notre jeu de données est constitué de 35 transects et d'une série temporelle. Comme nous souhaitons travailler sur des profils, nous séparons les transects en profils moyen sur une grille régulière à la résolution $1/4^\circ \times 1$ jour ($n = 2769$ après filtrage). Les différentes couches sont définies par leur profondeur maximale l_i^{obs} , comme on le voit sur la figure 4.2. La première couche s'étend de la surface à l_1^{obs} , la deuxième couche de l_1^{obs} à l_2^{obs} et la troisième de l_2^{obs} à l_3^{obs} .

Nous avons développé une méthode permettant d'extraire les limites de ces couches. L'étude visuelle d'un grand nombre de transect acoustique permet d'avoir une connaissance a priori de chaque l_i^{obs} . Par exemple, l_1^{obs} est inférieure à 250 m et l_3^{obs} entre 600 et 900 m. Nous allons nous focaliser sur chacune de ces zones de transition entre les couches en extrayant un intervalle du profil comportant le bas d'une couche et le haut de la suivante. Le but de la méthode que nous avons développée est de modéliser ces couches par une somme de fonction gaussiennes qui seront ensuite regroupées pour former le bas de cette couche et le haut de la suivante.

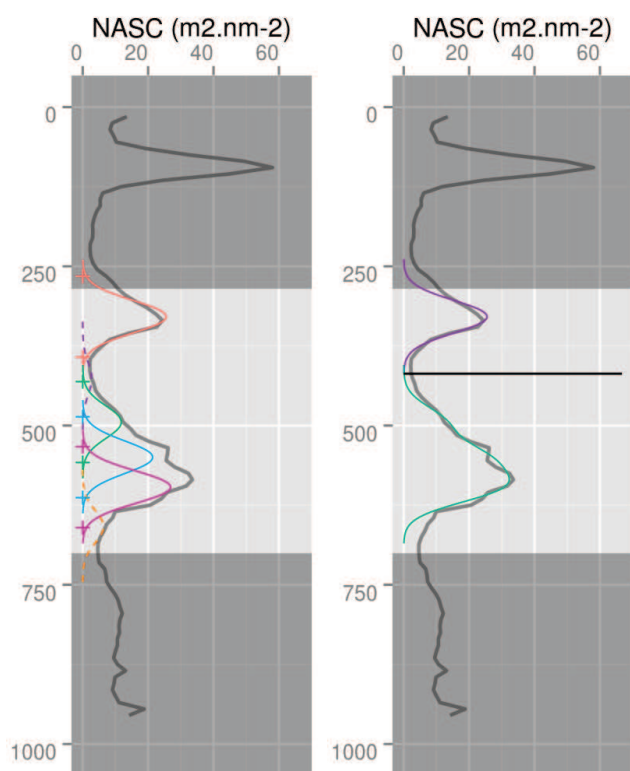


Figure 4.3 – Exemple du fonctionnement de la méthode de détection de la limite des couches

Prenons un exemple et déroulons les différentes étapes de notre méthode de détection. La figure 4.3 illustre cet exemple avec la détection de la limite de la deuxième couche. Sur les deux panneaux, le transect acoustique est tracé en gris et la zone à laquelle on s'intéresse est celle avec le fond gris clair.

Sur le panneau de gauche, on voit les différentes gaussiennes utilisées pour modéliser le bas de la seconde couche et le haut de la troisième. Certaines sont en tirets : elles ont été détectées comme étant du bruit (la jaune vers 600 m) ou parce qu'elles représentent de trop faibles variations (la violette vers 400 m). Les signes '+' représentent le haut et le bas de chaque gaussienne qui a été retenue. Les gaussiennes verte, bleue et rose se superposent : on les regroupe donc. La gaussienne orange ne se superpose avec aucune autre, elle est laissée seule.

Sur le panneau de droite, on peut voir en vert les sommes des gaussiennes des deux groupes : en violet la gaussienne orange du panneau de droite et en vert la somme des trois gaussiennes qui se superposaient. La courbe violette modélise le bas de la deuxième couche et la verte le haut de la troisième. La limite

basse de la deuxième couche est donc à l'intersection entre ces deux courbes. Le trait noir représente cette intersection et donc la profondeur de l_2^{obs} .

C Définition d'un modèle pour les limites verticales des couches de SEAPODYM-MTL

Pour chaque couche i , l'ensemble des valeurs de limites observées l_i^{obs} est ensuite corrélée à des variables environnementales. Les variables ayant les plus fortes corrélations ou anti-corrélations sont choisies pour créer des modèles linéaires. Le choix du modèle retenu se fait suivant les critères suivants :

- bonne corrélation
- faibles erreurs
- homosédasticité des variables

La partie II est la copie d'un article intitulé "Modeling average bio-acoustic layers of mesopelagic micronekton from space", soumis à Remote Sensing of Environment le 27 avril 2016.

II Modeling depths of ocean bio-acoustic layers from space

Modeling depths of ocean bio-acoustic layers from space

Anna Conchon, Olivier Titaud, Beatriz Calmettes, Nolwenn Behagle, Patrick Lehodey

April 19, 2016

Abstract

The management of marine resources and the conservation of marine biodiversity require a better understanding of the oceanic ecosystem, including the poorly known mesopelagic micronekton component, as well as the development of modeling tools. The interest for modeling the mesopelagic ecosystem has also increased recently due to the issue of climate change and the finding that organisms like mesopelagic micronekton performing Diel Vertical Migration (DVM) up and down in the water column over a daily cycle contribute to the biological pump transferring organic matter, and thus CO₂, in the deep ocean. DVM is a ubiquitous phenomenon in marine and freshwater pelagic ecosystems, observed using acoustic echo-sounders and showing consistent basin-scale patterns with three broad epi- and meso-pelagic zones, or bio-acoustic layers, between surface and ~1000m. Thus, mesopelagic micronekton modeling requires first to simulate the Average Micronektonic Layer Depths (AMLD) that are the depth boundaries between these observed bio-acoustic layers. This paper describes first a robust method to automatically detect micronekton layer depth boundaries from 38 kHz acoustic density data. A large dataset including acoustic transects from three Oceans was used. Then, detected bio-acoustic layer depth boundaries were used to calibrate AMLD models based on synoptic variables, especially those derived from satellite data. It included satellite derived euphotic depth, based on satellite surface irradiance and satellite surface ocean color, since change in light intensity is likely the strongest driver of DVM. The objective was to design simple models of average micronektonic layers that can provide a framework for modeling the dynamics of these organisms using a numerical transport scheme. Despite many sources of uncertainties and the high frequency variability of physical and biological features in the ocean, statistically significant relationships were obtained and support the idea that light intensity, modulated by phytoplankton concentration, is a strong driver of diel vertical migration of mesopelagic micronekton at global scale. Euphotic depth can be used to estimate large scale distribution of the first and second AMLD. The third AMLD was not properly defined by any variable tested here and was fixed to its average value (740m). More data, including data collected from mid-water acoustic platforms, are needed to define it correctly. This should be made easier with the automatic method proposed here that will allow processing massive acoustic datasets.

Keywords : Diel Vertical Migration, Euphotic depth, Irradiance, Mesopelagic layer, Micronekton, Ocean Colour, Sea Surface Temperature, Acoustic detection

I. Introduction

While biological phenomena are spatially structured by the physical environment, the ocean is characterized at basin scale by stronger vertical than horizontal physical gradients. Therefore, vertical gradients create layer boundaries with relatively homogeneous physical conditions between bound-

aries. It is not well understood how such physical structures interact with biological mechanisms, e.g. predator-prey interaction, to influence the vertical distributions of marine organisms. However, it is very clear that life history and behavior of pelagic organisms are strongly linked to the vertical dimension. The widespread use of echo-sounders after the 2nd World War showed that the Deep

Scattering Layer is a common attribute to all oceans. This mysterious acoustic signal was rapidly identified as due to the Diel Vertical Migrations (DVM) carried out by a huge biomass of small organisms inhabiting the lower mesopelagic layer (roughly below 400m) during the day and migrating at night in the epipelagic surface layer (above 200m) (Barham (1966); Gjøsaeter and Kawaguchi (1980)). They cross twice a day the upper mesopelagic layer (roughly 200-400m) which is usually poor, especially during day-time (at least based on acoustic data), and can be seen as a transition layer (Figure 1). These migrant organisms are micronekton (organisms within the size range 1-20 cm), a key component at the mid-trophic level of the pelagic ecosystem.

DVM are explained by adaptative evolutions to minimize the predation risk in the upper foraging zone (Kremer and Kremer (1988)). Observations support this idea that even very small organisms like zooplankton developed vertical migratory behavior as means of evading fish predators (Gliwicz (1986)), taking advantage that simple changes in buoyancy, e.g., by changing biochemical composition, can help to migrate up or down rapidly and over several hundred of meters (Visser and Jónasdóttir (1999)). More specialized adaptations in mesopelagic micronekton, e.g., with swimming bladder, make these vertical displacements very efficient in terms of energetic cost. Nevertheless, the predators of micronekton, essentially all larger marine species, have also evolved to chase them. This has become very obvious from the observations with electronic devices tracking the behavior of these animals, either fishes, squids, sea birds or marine mammals (e.g., Bailleul et al. (2007); Casey et al. (2010); Kooyman and Kooyman (1995); Nakano et al. (2003); Sims et al. (2005)).

The management of marine resources and the conservation of marine biodiversity require a better understanding of the oceanic ecosystem, including the poorly known mesopelagic micronekton component, as well as the development of modeling tools. Modeling approaches for micronekton are relatively rare. They simulate at least the diurnal

depth of the mesopelagic micronekton (the depth boundary between upper and lower mesopelagic zones on Figure 1). A few examples exist both for lake (Gabriel and Thomas (1988); Dodson (1990)) and marine (Bianchi et al. (2013)) systems. The interest for modeling the mesopelagic ecosystem and DVM has recently increased with the issue of climate change and the concept of biological pump (Volk and Hoffert (1985)). Indeed, a large part of anthropogenic CO₂ released in the atmosphere diffuse in the ocean and enters the marine food web through the mechanism of photosynthesis. Part of this carbon-based organic matter is transported and trapped in the deep ocean. Zooplankton and micronekton DVM are actively contributing to this vertical transfer. It has been estimated recently that between 10 to 40% of the total export in the region off the California coast could be attributed to mesopelagic micronekton DVM (Davison et al. (2013)).

Similarly, in order to model carbon fluxes from the surface to the depth, Bianchi et al. (2013) developed a model to predict the average diurnal depths of micronekton undergoing DVM. They characterized daytime DVM depth as the maximum daytime backscatter signal below 200 m depth observed in Acoustic Doppler Current Profiler data. Then a multiple linear regression is used to predict this depth from climatological series of chlorophyll a pigment concentration, mixed layer depth and vertical gradients of temperature and dissolved oxygen. The study showed that the depth of diel migration follows coherent large-scale patterns, with shallower migrations observed in regions characterized by lower subsurface oxygen, higher surface chlorophyll and shallower mixed layer depths. The authors propose that seawater oxygen concentration is the best single predictor of migration depth at the global scale.

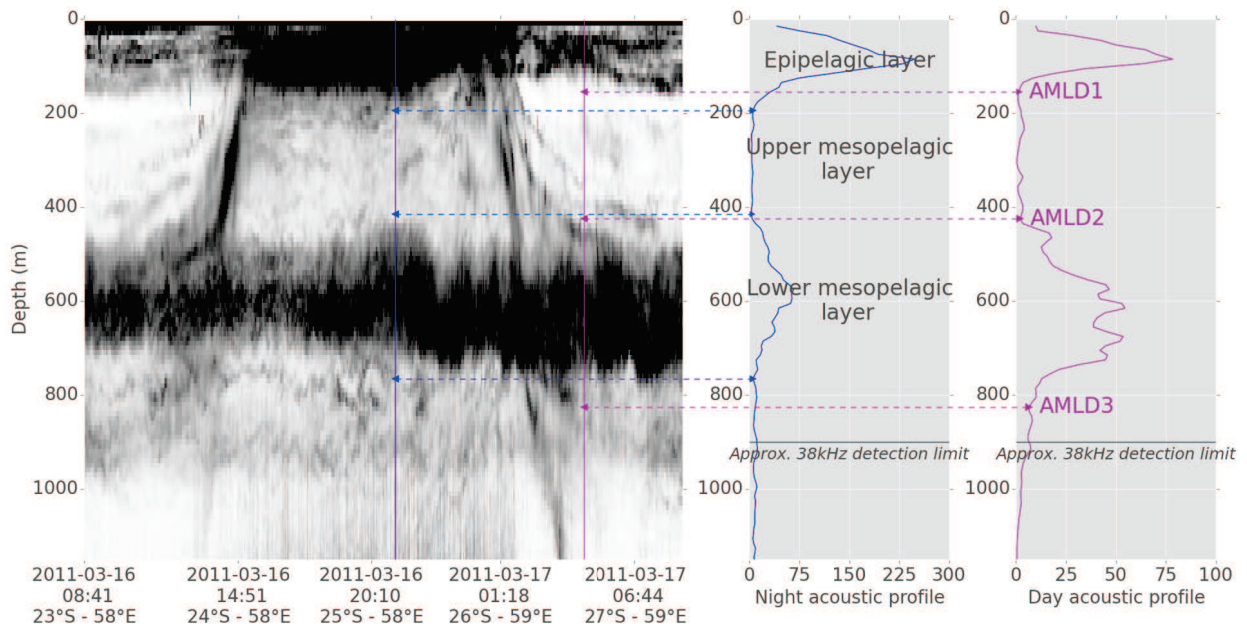


Figure 1: Bio-acoustic layers. Example of an acoustic transect with two vertical profiles illustrating the limits of vertical layers (AMLD) for which an automatic detection algorithm is proposed. Layers between AMLDs are defined. Note that scales of signal intensity are different between day and night profiles.

Though being essential step in understanding of the mesopelagic realm, DVM models do not provide micronekton biomass distribution. Further developments are necessary to achieve this objective. One tentative to derive a global biomass estimate by size range in marine environment is proposed by Jennings et al. (2008), using macroecological principles (size spectra) and ocean color data for primary production. They came to an estimate of 223 million metric tonnes (mt) for the global (90% coverage) ocean biomass of marine animals between 1 and 10 g, i.e., the size range of most of micronekton organisms. This estimation is well below a previous global estimation of 1000 million based on a compilation of trawling data (Gjøsaeter and Kawaguchi (1980)), itself very likely underestimated when considering acoustic rather than trawling estimates (Kloser et al. (2009); O'Driscoll et al. (2009); Irigoien et al. (2014); Davison et al. (2015)). Therefore, a very large uncertainty remains on the biomass estimate and distribution of this ocean ecosystem component.

We develop another modeling approach to simulate spatial and temporal dynamics of production and biomass of several oceanic micronekton functional groups (Lehodey et al. (2015)). The model

(SEAPODYM-MTL) considers a simplified ocean with the three vertical layers described above (epipelagic, upper and lower mesopelagic; Figure 1), assuming that organisms are homogeneously distributed in these layers with instantaneous vertical migration between day and night times. Thus it requires not only the average diurnal depths of micronekton undergoing DVM as defined by Bianchi et al., but also the depth boundary between epipelagic and upper mesopelagic layers, and the bottom depth of the lower mesopelagic layer (Figure 1). We are also interested to get a definition of these depths based on realistic synoptic variables, i.e. satellite-based, and not only on a climatological basis but with intra-seasonal to interannual and decadal variability. For the sake of simplicity, these three depth boundaries are called Average Micronektonic Layer Depths (AMLD) in the following text (Figure 1). AMLD definition is critical in our model since i) its system of advection-diffusion-reaction equations is solved over these three layers using associated average currents and temperatures, and ii) the model parameters (energy transfer coefficients from primary production and temperature related coefficients of mortality and recruitment) are optimized using a Maximum Likelihood Estimation approach

(Senina et al. (2008)) requiring aggregation of 38 kHz bio-acoustic densities following the definition of these vertical layers (Lehodey et al. (2015)).

Preliminary experiments have shown that optimization results are sensitive to the AMLD definition. In previous versions of the model, layers depths were first fixed (0-200m; 200-400m; 400-1000m, in Lehodey et al. (2010)). Then, based on a few 38 kHz acoustic transects, layer depths were linked to the euphotic depth Z_{eu} (0-1.5 Z_{eu} ; 1.5-4.5 Z_{eu} ; 4.5-10.5 Z_{eu} , in Lehodey et al. (2015), calculated either from biogeochemical models or derived from satellite data (Morel and Berthon (1989)). The use of multiples of the euphotic depth was consistent with the limited available dataset. However, we took the opportunity of new developing online bio-acoustic databases to expand our analysis to a much larger proportion of the global ocean.

Despite acoustic signal cannot be easily converted to biomass, it is the only observing system allowing large monitoring coverage of the surface to mesopelagic ocean (Kloser et al. (2002); Kloser et al. (2009); Davison et al. (2015)). With the development of these monitoring networks generating massive datasets, it is essential to seek for an automatic procedure rather than a visual examination to extract biological layer depth limits from 38 kHz acoustic density profiles. Then, we aim at developing simple linear models for each AMLD, testing the most synoptic and easy to access environmental variables as explanatory variables.

This paper describes a method to observe and automatically detect micronekton layer depth boundaries from 38 kHz acoustic density data. These observations are then used to calibrate Average Micronektonic Layer Depth (AMLD) linear models based on synoptic satellite-derived oceanic variables.

Here, we are looking for consistent large-scale patterns, i.e., the three broad epi- and meso-pelagic zones observed in ocean between surface and ~1000m (Brodeur and Yamamura (2005); Béhagel et al. (2015) and previous references). Obviously, much fine structures and multiple layers can be observed at higher resolution (e.g., Benoit-Bird

and Au (2006)), but they are not considered at the resolution of this global ocean model. Finally, results are discussed and further perspectives are presented to improve AMLD modeling.

II. Materials and methods

A Acoustic data and processing

To identify vertical boundaries of micronekton layers we use 38 kHz acoustic data that is the most appropriate frequency for micronekton (Bertrand et al. (1999); McClatchie and Dunford (2003); Kloser et al. (2009)). The acoustic density is used as a proxy of the micronekton biomass. Available acoustic data vary greatly in quality (type of acquisition platform, pre-processing, filtering procedure, etc, ...) and quantity (spatial and temporal coverage and number of platforms used) within the time period of the study and between ocean basin (Table 1; Figure 2). Two oceanic regions display a good spatial and temporal coverage of easy available datasets: the south-east Indian ocean and the Tasman Sea, thanks to the IMOS - Bio-Acoustic Ships of Opportunity sub-facility Data Set (IMOS SOOP-BA dataset; see Ryan (2011)). The coverage of the Indian Ocean is also improved by a series of transects provided by the MyctO-3D-MAP project. More sparse data are available in the Pacific Ocean (Domokos et al. (2007); Hazen and Johnston (2010); Lehodey et al. (2015)). In the Atlantic Ocean, one single dataset is available for this study but consists in a detailed one-year time series collected at the MarEco mooring (Ausubel et al. (2010); Handegard et al. (2013)).

Data processing needs to be tailored to the various sources of data (Table 1). To homogenize the datasets we average acoustic data using a regular grid at the $1/4^\circ$ spatial resolution and a daily time step, distinguishing between day, night and transition periods (see Lehodey et al. (2015)).

Data sources	Type of platform	Acquisition period	Number of profiles	References
IMOS BA-SOOP	Boats of opportunity	2009-2014	1492	Ryan (2011)
MyctO-3D-MAP	Research vessel	2013-2015	460	Béhagle et al. (2015)
R. Domokos	Research vessel	2009-2011	315	Lehodey et al. (2015)
R. Domokos	Research vessel	2002	132	Domokos et al. (2007)
E. L. Hazen	Research vessel	2006	150	Hazen and Johnston (2010)
MarEco	Bottom station	2004-2005	220	Ausubel et al. (2010) Handegard et al. (2013)

Table 1: Bio-acoustics data characteristics

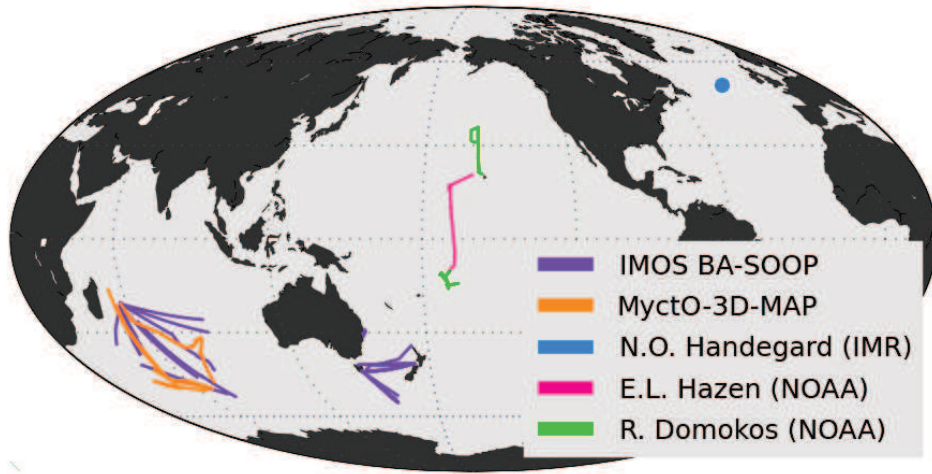


Figure 2: Distribution of available 38 kHz acoustic data and their sources.

B Detection of vertical boundaries

detailed in Appendix section D.

We develop an automatic procedure to define vertical boundaries of micronekton layers from 38 kHz acoustic profiles (see the first section of the Appendix). The method is constrained to find three layers using a priori robust information on the depth ranges of each layer. The actual limit is sought between values of minimum and maximum possible depths (Appendix section A). A Gaussian Mixture Model (a sum of Gaussian functions) is fitted to this selected depth window for each profile averaged at the resolution $1/4^\circ \times \text{day}$ with the R software mclust package (Fraley and Raftery (2002); Fraley et al. (2012)); see Figure 3 and Appendix section B). Gaussian functions that fit acoustic noise are filtered and discarded. Then, an aggregation procedure merges overlapping Gaussian functions (Appendix section C). This aggregation produces a series of sub-layers from which the actual boundary values need to be extracted using different criteria that are

Estimated layer boundary values may still show false detections or abrupt change between adjacent profiles (e.g., due to strong mesoscale activity). They are filtered in a final step by identifying outliers from adjacent cells and time steps.

C Outlier filtering

Once the vertical layer boundary values are detected, some of them still appear outside a realistic range relatively to the adjacent observations. These outliers are filtered by testing each value with those available in a five degree square centered on the position of tested value and at a monthly time step. If there is a minimum number of 10 observations within this square, the mean and standard error are computed. Tested values below or higher than the mean plus three standard errors are considered as

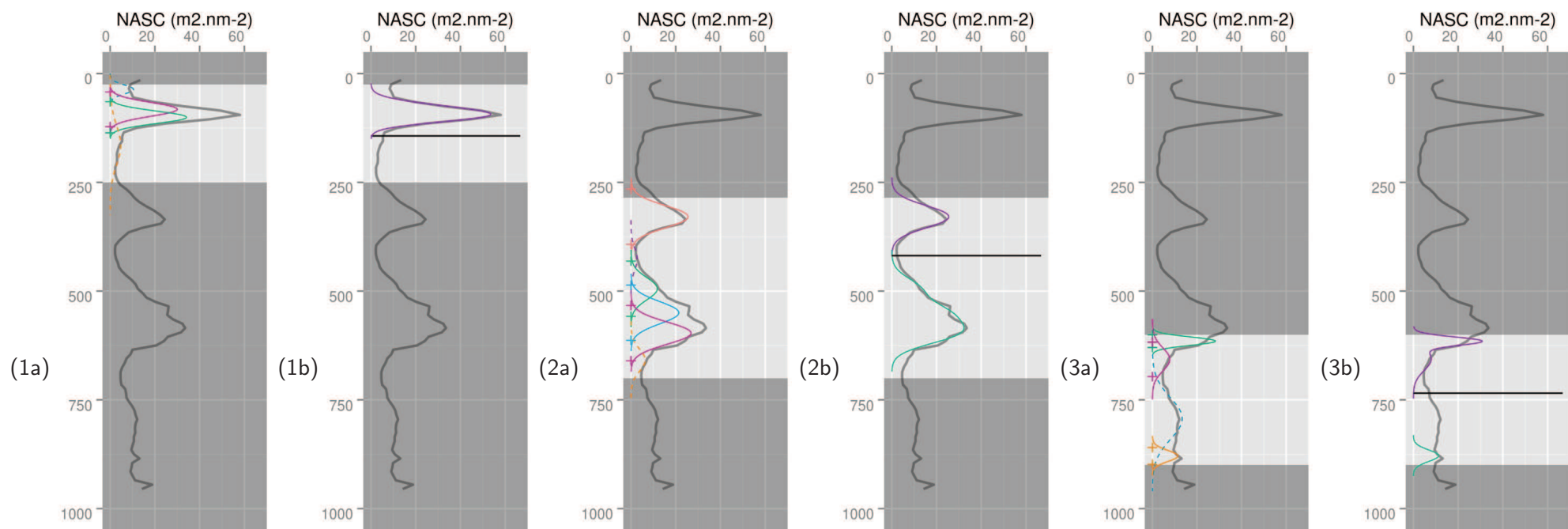


Figure 3: Layer boundary determination from a 38 kHz acoustic density vertical profile. In each plot, light grey color highlights the fixed depth window in which the layer boundary value is sought. Plot number indicates the layer with (a) the multiple Gaussian functions and their '+' mark limits and (b) the sum of Gaussian functions. The horizontal black line gives the selected boundary value. Dashed lines are used for rejected Gaussian functions.

outliers. Profiles with at least one AMLD tagged as outlier are rejected. If the number of observations in the 5° square is less than 10, all profiles in the square are also discarded.

D Synoptic variables tested in the models

The mechanisms that drive the vertical distribution of organisms in the ocean are still unclear even if there are several good potential hypotheses to explain them. We have selected a list of variables according to these hypotheses but also with the conditions of synoptic global coverage availability and realistic operational description of the ocean. Therefore, the list is restricted to either satellites derived data or global ocean reanalyses, i.e., hindcast simulations of the ocean circulation assimilating both satellites and in situ observations for a more realistic prediction. All selected variables have global coverage and are interpolated to the same spatio-temporal resolution used for acoustic data ($1/4^\circ \times$ day resolution). Then, values co-localized with acoustic profiles are extracted for linear model regressions.

Water masses physical criteria Temperature and salinity, either directly (thermocline and halocline) or indirectly through water density (pycnocline), are likely easily detected by marine organisms and in addition can influence the buoyancy in the water column and thus the distribution of small organisms (Lezama-Ochoa et al. (2014)). At global scale, sea surface temperature (available from satellite remote sensing, reflects latitudinal biomes and can be a good proxy of large biogeographical provinces (e.g., equatorial upwelling vs tropical gyre). We use the OSTIA satellite SST product provided by the European copernicus service (<http://marine.copernicus.eu/>) and extracted depths of isotherm and isohaline from the ocean reanalysis GLORYS.v2 (Parent et al. (2013)) and the operational version of this ocean circulation model for the recent years (<http://www.mercator-ocean.fr/en/>).

Chlorophyll-a At large scale, surface chlorophyll a concentration (Chl_a) is a proxy of ocean productivity that may impact DVM either directly since it is the primary source of the food chain, or indirectly by interacting with light attenuation. There are also strong correlations between physical characteristics of water masses and ocean productivity. This variable derived from ocean color satellite data is available from the Ocean Productivity web site of the university of Oregon State (<http://www.science.oregonstate.edu/ocean.productivity/index.php>).

Light and associated variables Since DVM occur at sun set and sun rise, the light is certainly a good candidate for defining the layer boundaries. Satellite remote sensing provides an estimate of surface solar irradiance (SSI). However, light penetration in the ocean is not simply correlated to the light measured in surface. Due to the attenuation by phytoplankton, it is also inversely related to chlorophyll concentration in the water that varies in space and time. This is why the variability in sunlight penetrating at the surface is included as a light-dependent term when estimating the euphotic depth from satellite data. Euphotic depth is defined as the depth where irradiance is reduced to 1% of the surface Photosynthetically Available Radiation (PAR), which is the light spectrum used by phytoplankton for photosynthesis. The PAR is computed from and used with satellite derived sea surface chlorophyll-a concentration to estimate the euphotic depth (Morel and Berthon (1989); Behrenfeld and Falkowski (1997)). Since the euphotic depth is a measure of light penetration combining both the intensity of light at the surface and the attenuation due to chlorophyll concentration, this variable seems well appropriate to correlate with DVM. We use the euphotic depth computed from the VGPM model (hereafter Z_{eu}) (Behrenfeld and Falkowski (1997); <http://www.science.oregonstate.edu/ocean.productivity/>). It is calculated using the Morel and Berthon Case I model. This model estimates Z_{eu} from satellite derived surface chlorophyll concentrations and is based on empirical equations

to fit field data.

E Model selection and predicted layers

Linear model parameters are estimated with a least square method (R software, package stats, function lm; Chambers et al. (1992) and Wilkinson and Rogers (1973)) using initial $1/4^\circ \times \text{day}$ resolution after profile filtering. Models are ranked according to three statistical metrics: the coefficient of determination (R^2), the Root Mean Square Error ($RMSE$) and the Normalized Root Mean Square Error ($RMSE$). We present and discuss model results with the highest correlation and lowest error values.

III. Results

A Automatic detection of boundary layers

The automatic procedure uses seven parameters for the second and third layers. For the first layer, there is one more parameter to detect the case of an empty layer. All parameters were manually tuned, based on visual inspection of all acoustic transects. One single parameterisation is then used for all profiles of the global ocean.

The method fails to estimate the actual boundary of a layer when:

- The signal to noise ratio is low. This may be the case especially below 500m. In that case, the procedure cannot detect the actual limits;
- There are still some upward or downward migrations included in the average profile that make the layer detection too difficult ;
- The profile presents two maxima in the third layer, leading to a false detection between these maxima.

Most of these false detections are rejected by the outlier filtering.

The method provides better result in the deep ocean than in shallow bathymetry. We observe that shallow bottom depth, i.e., lower than 800m, tends to play a role in micronekton layers structure. In our dataset only 3.7% of the total number of profiles (after all screenings and filtering) fall in that category, thus we did not modify the procedure for these profiles. With more shallow topography profiles, it would be necessary to modify the depth range in which actual layer boundaries are sought according to ocean topography.

From the 5,720 average profiles available at the resolution $1/4^\circ \times \text{day}$, 31.6% were acquired during migration transition phases and thus not usable. From the remaining data, 1.3% to 4.4% (depending on the layer) were discarded during the first data screening. Finally, the last filtering step detected 28.8% outliers. In most case, the profiles were obviously coming from a noisy transect or poorly sampled area. Finally a total of 2,769 average profiles (1,267 during the day and 1,502 during the night) over the 3890 available profiles were used in linear regression models.

There is no high correlation between observed layer boundaries. The first limit is not correlated with the second one ($r = 0.036$), whereas it is negatively correlated with the third one. Second and third layer boundaries are slightly positively correlated (Figure 4).

B Regression models

Here we only present the variables that provide the best results in term of correlation and errors with the observed layer boundaries (Table 2; Figure 4), i.e.: the sea surface temperature (SST), the sea surface chlorophyll a concentration (Chl_a) and the euphotic depth (Zeu). Overall, the correlations are weak due to very high dispersion of observations. Selected explanatory variables explain between 8 and 23% of the total variance of the observed layer boundary values. However, this is not too surprising given expected physical and biological very high variability at this resolution and the cumulative

sources of errors (see discussion). And despite this high variability, the results are highly significant statistically (Table 2).

Euphotic depth exhibits a significant correlation with the three layer limits: positive with the 1st and 2nd limits and negative with the deepest one (Figure 4). *SST* is significantly correlated with first and third layer limits but not the second one. As with *Zeu* the correlation between *SST* is also negative for the third layer depth. Finally, *Chl_a* concentration shows significant anti-correlation with the first two layer depth limits, a result that can be expected in the hypothesis of a light-driven mechanism of vertical migration, since light attenuation increases with phytoplankton biomass.

The previous definition proportional to *Zeu* currently used in SEAPODYM-MTL has always larger error values (Table 2). For the first layer (AMLD1), the most significant relationships are obtained with *SST* ($r = 0.359$) and *Zeu* ($r = 0.288$). Both relationships lead to a depth limit varying in maximum amplitude of 50 m for the range of observed values (108-146m for *SST* between 0-30°C, and 117-143 for *Zeu* between 40-110m). *Chl_a* and *Zeu* give similar correlations (0.438 and 0.431 respectively) for the second layer depth boundary (AMLD2) but with opposite sign. This is expected however, considering that the euphotic depth is inversely correlated to the concentration of phytoplankton. These relationships get the highest scores and also predict the largest amplitude in depth, varying between 328-487 m for *Chl_a* concentration between 1 and 0 mg m⁻³ and 400-514 m for *Zeu* (40-110m). Data are more evenly distributed over the whole range of observed euphotic depths than *Chl_a* concentration (Figure 4), suggesting a more robust relationship.

The results for the third layer (AMLD3) shows the highest correlation with *SST* (-0.474), but data distribution shows a marked shift near *SST* values of 14°C (Figure 4). This result is achieved with a substantial number of observations in cold waters (9.2% of the total, 255 observations, in latitudes higher than 50°). Therefore, the linear model is not likely appropriate in this case. *Zeu* provides also

a significant relationship with a lower correlation (-0.225) and less heteroscedasticity. But while this linear model predicts small amplitude (720-760m) in the depth variation over the range of observed *Zeu*, it has higher values of RMSE and than the average model.

In summary, at the broad scale and resolution observed here, *SST* (AMLD1) and *Zeu* (AMLD1 and AMLD2) seem sufficient to predict the average depths of the first two bio-acoustic layers and capture the first order of their variability. A larger dataset, the use of different platforms of observation and a better treatment of noise (see discussion) seem necessary to propose a model for the third depth limit. In the meantime, a fixed average depth of 740m would be a reasonable alternative (table 2). With this new definition, the predicted AMLD are substantially improved in comparison to previous definition (table 2), with an absolute error reduction of 35% for the first layer, 55% for the second one and 78% for the third one.

The figure 5 shows two long North-South acoustic transects collected between latitudes 22°S and 42°S in the Indian Ocean with the corresponding temperature and definition of AMDL. Two rich scattering layers are well visible: one at the surface and one around 600m depth. The upward and downward migration phases are also well visible. Outside these migration periods the second layer shows much less dense acoustic responses. In the second transect there is a second dense layer below 800m in the tropical region (Figure 5). Given the limitation of the echo-sounder at these depths, it is quite difficult to interpret this signal. It is possibly a true detection of mesopelagic organisms but it could be also largely overestimated. New mid-water platforms of observation are needed to investigate micronekton biomass at these depths.

Layer depth boundaries predicted by the new AMLD models are in good agreement with detected limits. These new AMLD definitions appear very steady, and while following large scale trends they are insensitive to local perturbations. For the first and second AMLD, the depth ranges vary between

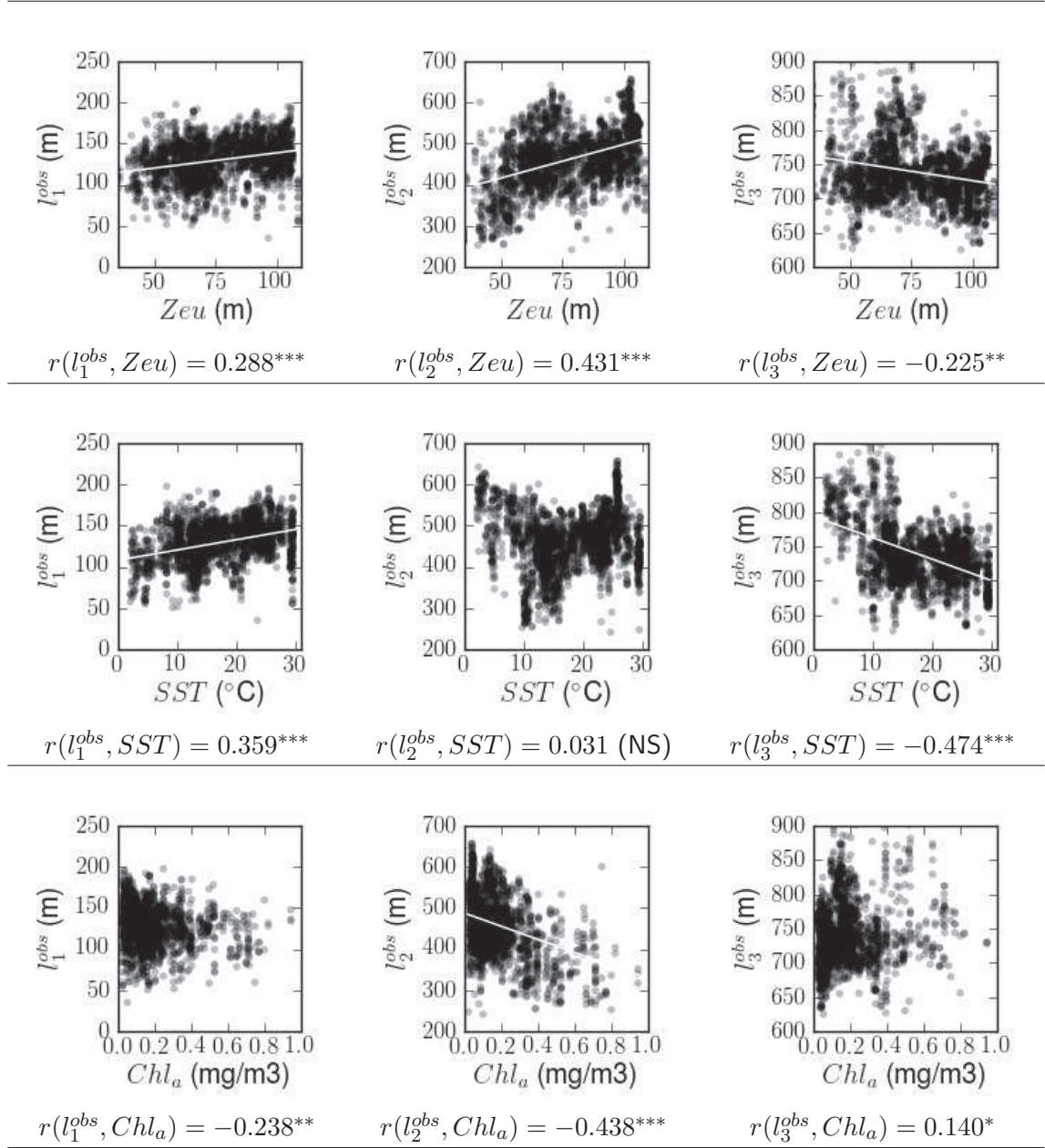


Figure 4: Relationships between detected layer boundaries and potential explanatory variables SST , Zeu and Chl_a . The Pearson's correlation score r is provided for each relationship. (NS) : non significant $P > 0.1$, * : $P < 0.1$, ** : $P < 0.01$ and *** : $P < 0.001$.

		1st layer boundary	2nd layer boundary	3rd layer boundary
AMLD model 2015	model	1.5 <i>Zeu</i>	4.5 <i>Zeu</i>	10.5 <i>Zeu</i>
	$RMSE$	34.2m	140m	199m
	$\overline{RMSE}$	28.5%	29.8%	27.3%
Mean model	model	129.8	460.0	738.9
	$RMSE$	23.7m	70.4m	45.1m
	$\overline{RMSE}$	24.3%	17.9%	5.8%
<i>Zeu</i>	model	101.7 + 0.37 <i>Zeu</i>	335 + 1.63 <i>Zeu</i>	780.9 – 0.55 <i>Zeu</i>
	R^2	0.08	0.19	0.05
	$RMSE$	22.7m	63.5m	43.9m
	$\overline{RMSE}$	23.5%	15.3%	5.7%
<i>SST</i>	model	108 + 1.2 <i>SST</i>		793.6 – 3.11 <i>SST</i>
	R^2	0.13	Non significant	0.23
	$RMSE$	22.2m		39.7m
	$\overline{RMSE}$	22.5%		5.3%
<i>Chl_a</i>	model	134.9 + 29.2 <i>Chl_a</i>	487 – 159 <i>Chl_a</i>	733.3 – 32.6 <i>Chl_a</i>
	R^2	0.06	0.19	0.02
	$RMSE$	23.0m	63.2m	47.7m
	$\overline{RMSE}$	23.8%	15.2%	6.0%

Table 2: Models of layer boundaries with respect to selected explanatory variables. Three statistical metrics are shown: R^2 , $RMSE$ and $\overline{RMSE}$. Results are compared to the previous definition (AMLD model 2015) proposed in Lehodey et al. (2015).

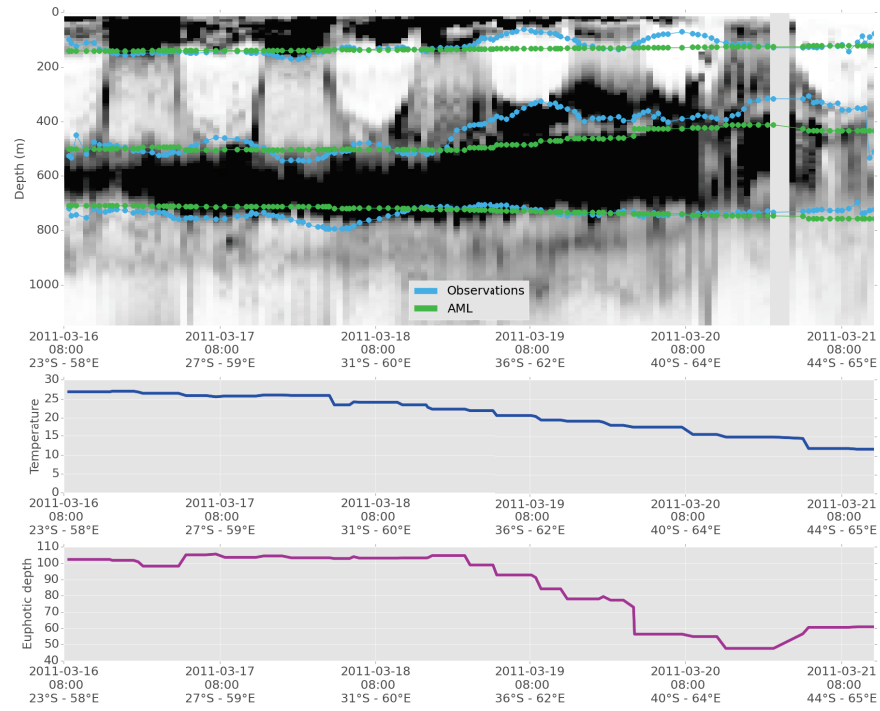
100-150 m and 400-500 m respectively. The third one is fixed at 740m. The new definition leads to a narrower range of variability than with the model based on multiples of *Zeu* but less error on the average depth value. When moving from tropical to temperate latitudes, both the first and second layer depth boundaries become shallower but with a stronger amplitude for the second, leading to an increasing thickness of the upper mesopelagic layer in latitudes lower than 30°, and a narrower deeper mesopelagic layer (Figure 5). This is in agreement with an increasing gradient of productivity observed between tropical and temperate waters and correlated with the deepening of euphotic depth and increasing phytoplankton concentrations.

IV. Discussion

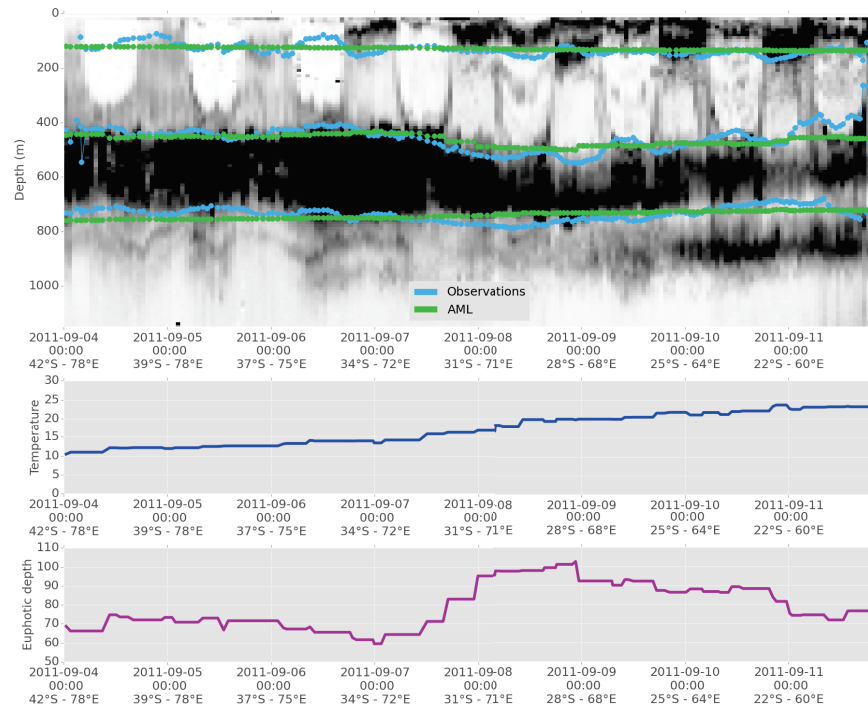
The acoustic sampling used in this study is built on an opportunity basis with linear transects across oceanic basins. It remains sparse over time even in

the most sampled regions. Therefore, data are not randomly distributed neither between ocean basins nor inside one specific basin. Moreover, some areas are sampled every year at the same season and thus do not include seasonal variability. This is obviously a drawback for this study and calls for further revisions of the results presented here, while data will continue to be collected in all oceanic regions.

These new data should also benefit of a better treatment with real-time noise filtering. Only in some cases the acquisition system used for the present dataset was specifically designed to reduce noise in acoustic data in real time. This may be a source of large bias since noise in acoustic data acquisition is known to increase with the distance to the transducers. For ship transects, the signal-noise ratio will be therefore the lowest in the deepest (3rd) layer. We have implemented a series of filtering techniques to reduce noise or simply remove layer detections from the worst profiles, but it does not prevent some specific profiles to be still noisy. Another potential issue using acoustic



(a)



(b)

Figure 5: Examples of two long latitudinal 38 kHz transects (IMOS database) in the Indian Ocean showing the automatically detected lower depth boundaries (blue dotted line) and the predicted depth from AMLD based on *Zeu* (AMLD2) and SST (AMLD1 and AMLD3). Second and third panels illustrate the variation of *Zeu* and SST along the transect. Subfigure (a) represents a summer transect between Mauritius Island to Kerguelen Island and subfigure (b) represents a winter transect between St Paul Island and Mauritius Island.

data to define vertical biological layers could come from the lack of acoustic response by some types of organisms, especially those that do not have gas in their body (Kloser et al. (2002); Davison et al. (2015)). This may lead to null or underestimated acoustic detection and then a lack of layer boundary identification using the algorithm developed in this study. However, in the future with the increasing interest in multi-frequency acoustic analyses (e.g. Kloser et al. (2002)), the link between actual biomass and acoustic response will be more accurate and lead to fewer bad layer boundary identification.

The proposed approach is constrained to detect three layers by restraining the search where the depth boundaries can be reasonably found. This method appears well adapted to the deep ocean but may need to be revised for shallower bathymetry (i.e., $< 800\text{m}$) where biological layers, based on visual inspection of acoustic data, seem to be impacted by the topography. However, very few profiles were concerned by this limitation in our dataset. Spurious detections can be also produced in the aggregation phase of multiple Gaussian functions (either real or generated from noise). Though it seems quite robust, the final outlier filtering may likely not discard all of them, but it is a minor weakness compared to the interest of processing automatically a huge volume of data as those generated by acoustic surveys.

While there are various sources of errors on acoustic data, there are also uncertainties on the oceanic variables selected to test a global model defining vertical layer boundaries. To reduce these sources of errors we decided to select only global synoptic variables available from satellites or strongly constrained by them. Nonetheless, these satellite measurements have themselves their own range of uncertainty (e.g., Mélin et al. (2016)). Fields derived from satellite measurements reproduce most scales of in situ processes but obviously only at the ocean surface. Therefore ocean general circulation models, even when assimilating satellite data, may have difficulty to simulate accurate subsurface circulation. While strongly linked to satellite data (solar irradiance and ocean color), the euphotic

depth is also derived from an empirical bio-optical model using a compilation of measurements of marine phytoplankton productivity (Behrenfeld and Falkowski (1997)). Both models of ocean general circulation and euphotic depth have the largest errors in coastal environment. However, our dataset has few profiles close to the coast where such errors are potentially higher.

There are certainly other explanatory variables that could be considered to correlate with DVM. We can assume that migrating mesopelagic species are driven by a threshold value in light intensity (Cohen and Forward (2009); Bianchi et al. (2013)). The depth of this value, i.e. an isolume, could be therefore a better criterion controlling the migrations. However it is difficult to estimate and the threshold may be different from one species to the other. Moreover, all proposed isolumes are related to the same drivers (*SSI* and attenuation by phytoplankton) that control the euphotic depth, and should be finally strongly correlated to it.

Dissolved oxygen concentration has been also proposed as a critical variable that controls the vertical distribution of marine organisms since they all need oxygen for respiration. The relationship has been particularly well demonstrated in regions where oxygen concentrations are reaching minimal values like in the Oxygen Minimum Zones (OMZ) of coastal upwelling systems (Bertrand et al. (2010)). Deep layers below the oxycline can serve as a refuge from predation (Parker-Stetter and Horne (2009)), and many organisms have developed vertical migration strategies to use, pass through and cope with the OMZ (Ekau et al. (2010)). These organisms need to come back in more oxygenated waters (i.e., the surface) for reducing their deficit in oxygen, and night time is certainly the period with the lowest predation risk. However, it is difficult to find a threshold value of oxygen concentration that fit layer boundaries both in OMZ and more oxygenated ocean regions. One more issue comes from the need to rely either on climatological files or on biogeochemical models to use fields of oxygen concentration at depth.

Linear regressions were chosen for testing AMLD models since they provide robust and easy to compare results, and also limit the possibility of abrupt changes in the layer depth estimates that could lead to numerical instability. The various sources of errors listed above and the high frequency variability of physical and biological features in the ocean can explain the relatively low values of correlation achieved with the different regression models tested in this study. Nevertheless, the dataset used is already sufficient and covering a large latitudinal range to provide statistically significant relationships. Those obtained for the deepest layer (AMLD3) have the lowest variability in depth amplitude and do not improve the results compared to a model simply using the average depth value. The acoustic signal at these depths collected from the surface becomes more and more noisy and the approach to process this noise has certainly a strong influence on the result. An increasing number of transects but also time series of data collected from towed echosounders or mid-water moorings and gliders should help to revise and improve this definition. The first and second layer depths (AMLD1 and AMLD2) show negative correlations with Chl_a concentration and positive correlations with Z_{eu} . These relationships are coherent and opposite correlations expected in the case of DVM piloted by light intensity. For a same level of surface irradiance, a lower concentration of chlorophyll means deeper euphotic depth.

Finally, thermocline and pycnocline do not provide significant relationships with detected layer depth boundaries. However, these variables are derived from an ocean reanalysis that may lack the accuracy needed in the subsurface circulation to obtain a good statistical relationship.

This study supports the idea that light intensity is a strong driver of diel vertical migration of mesopelagic micronekton at global scale. The satellite derived euphotic depth seems a good proxy to model at least the first two AMLD defined in this paper, the third one remaining apparently relatively constant in depth. Nevertheless, the dispersion of data is still wide and the statistical correlations are

low. Residuals might be correlated to other variables requiring multivariate linear models as for example in Bianchi et al. (2013). This may be particularly true in oceanic region with low dissolved oxygen concentration. In addition, considering that light is a driver of DVM in arctic latitudes, especially during polar night, or in the very deep ocean (van Haren and Compton (2013)), where light intensity is extremely low, remains a pending question and requires further studies and data collection. DVM have been observed under the ice in the Arctic sea, and Antarctica during summer (Fortier et al. (2001); Tanimura et al. (2008)), but also during the polar night (Berge et al. (2009)), while absence of such migrations are noted during summer by Cisewski et al. (2010) in the Lazarev Sea in Antarctica.

The present study objective was to design a simple model of average micronektonic layers that can be used for operational global modeling of the dynamics of these organisms using a numerical transport scheme requiring necessarily drastic simplifications. This approach has been designed to be robust in front of the multiple sources of errors both in the state variables (automatically detected layer depths) and the explanatory variables (cf. above). Multivariate models could be better solutions. However, before to move along this way, there is a need to process much more acoustic data collected from equatorial to polar systems. Some of them already exist – archived in unexploited databases – while others are currently collected from ongoing programs. Developing automatic and robust processing approaches as the one presented here is a critical step to use them.

Acknowledgments

The authors thank Rudy Kloser (CSIRO) and the team in charge of the IMOS dataset (<http://imos.org.au/basoop.html>), Yves Chérel (CNRS) and the French Institute Polaire Paul-Emile Victor for developing acoustic data acquisition with the R.V. M. Dufresne under the French ANR MyctO-3D-MAP, Réka Domokos (NOAA), Elliot

Lee Haven (NOAA) and Nils Olav Handegard (IMR) for providing acoustic data. We also thank the Ocean Productivity team for providing us SeaWiFS-derived euphotic depth and the Groupe Mission Mercator Coriolis for providing the GLORYS-2v1 reanalysis. This study was partially supported by the European Space Agency under the OSMOSIS research project, and the EU H2020 AtlantOS, GREENUP and French ANR AMORAD research projects.

References

- Ausubel, J., Trew Christ, D., Waggoner, P., et al. (2010). First census of marine life 2010: highlights of a decade of discovery.
- Bailleul, F., Charrassin, J.-B., Monestiez, P., Roquet, F., Biuw, M., and Guinet, C. (2007). Successful foraging zones of southern elephant seals from the kerguelen islands in relation to oceanographic conditions. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1487):2169–2181.
- Barham, E. G. (1966). Deep scattering layer migration and composition: observations from a diving saucer. *Science*, 151(3716):1399–1403.
- Béhagle, N., Cotté, C., Ryan, T. E., Gauthier, O., Roudaut, G., Brehmer, P., Josse, E., and Cherel, Y. (2015). Acoustic micronektonic distribution is structured by macroscale oceanographic processes across 20–50s latitudes in the south-western indian ocean. *Deep Sea Research Part I: Oceanographic Research Papers*.
- Behrenfeld, M. J. and Falkowski, P. G. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and oceanography*, 42(1):1–20.
- Benoit-Bird, K. J. and Au, W. W. (2006). Extreme diel horizontal migrations by a tropical nearshore resident micronekton community. *Marine Ecology Progress Series*, 319:1.
- Berge, J., Cottier, F., Last, K. S., Varpe, Ø., Leu, E., Søreide, J., Eiane, K., Falk-Petersen, S., Willis, K., Nygård, H., et al. (2009). Diel vertical migration of arctic zooplankton during the polar night. *Biology letters*, 5(1):69–72.
- Bertrand, A., Ballón, M., and Chaigneau, A. (2010). Acoustic observation of living organisms reveals the upper limit of the oxygen minimum zone. *PLoS ONE*, 5(4):e10330.
- Bertrand, A., Le Borgne, R., and Josse, E. (1999). Acoustic characterization of micronekton distribution in french polynesia. *Mar. Ecol. Prog. Ser.*, 191:127–140.
- Bianchi, D., Galbraith, E. D., Carozza, D. A., Mislan, K., and Stock, C. A. (2013). Intensification of open-ocean oxygen depletion by vertically migrating animals. *Nature Geoscience*, 6(7):545–548.
- Brodeur, R. and Yamamura, O. (2005). Micronekton of the north pacific. *PICES Scientific Report*, 30:1–115.
- Casey, J., Garner, J., Garner, S., and Williard, A. S. (2010). Diel foraging behavior of gravid leatherback sea turtles in deep waters of the caribbean sea. *The Journal of experimental biology*, 213(23):3961–3971.
- Chambers, J., Freeny, A., and Heiberger, R. (1992). Chapter 5 of statistical models in s. *Analysis of Variance; Designed Experiments*. Wadsworth & Brooks/Cole, Pacific Grove.
- Cisewski, B., Strass, V. H., Rhein, M., and Krägesky, S. (2010). Seasonal variation of diel vertical migration of zooplankton from adcp backscatter time series data in the lazarev sea, antarctica. *Deep Sea Research Part I: Oceanographic Research Papers*, 57(1):78 – 94.
- Cohen, J. H. and Forward, R. B. (2009). Zooplankton diel vertical migration - a review of proximate control. *Oceanography and marine biology*, 47:77–109.
- Davison, P., Checkley, D., Koslow, J., and Barlow, J. (2013). Carbon export mediated by mesopelagic fishes in the northeast pacific ocean. *Progress in Oceanography*, 116:14–30.

- Davison, P., Lara-Lopez, A., and Koslow, J. A. (2015). Mesopelagic fish biomass in the southern california current ecosystem. *Deep Sea Research Part II: Topical Studies in Oceanography*, 112:129–142.
- Dodson, S. (1990). Predicting diel vertical migration of zooplankton. *Limnology and Oceanography*, 35(5):1195–1200.
- Domokos, R., Seki, M. P., Polovina, J. J., and Hawn, D. R. (2007). Oceanographic investigation of the american samoa albacore (*thunnus alalunga*) habitat and longline fishing grounds. *Fisheries Oceanography*, 16(6):555–572.
- Ekau, W., Auel, H., Pörtner, H.-O., and Gilbert, D. (2010). Impacts of hypoxia on the structure and processes in pelagic communities (zooplankton, macro-invertebrates and fish). *Biogeosciences*, 7(5):1669–1699.
- Fortier, M., Fortier, L., Hattori, H., Saito, H., and Legendre, L. (2001). Visual predators and the diel vertical migration of copepods under arctic sea ice during the midnight sun. *Journal of Plankton Research*, 23(11):1263–1278.
- Fraley, C. and Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American statistical Association*, 97(458):611–631.
- Fraley, C., Raftery, A. E., Murphy, T. B., and Scrucca, L. (2012). mclust version 4 for r: normal mixture modeling for model-based clustering, classification. Technical report, and Density Estimation Technical Report No.
- Gabriel, W. and Thomas, B. (1988). Vertical migration of zooplankton as an evolutionarily stable strategy. *American Naturalist*, pages 199–216.
- Gjøsaeter, J. and Kawaguchi, K. (1980). *A review of the world resources of mesopelagic fish*. Number 193-199. Food & Agriculture Org.
- Gliwicz, M. Z. (1986). Predation and the evolution of vertical migration in zooplankton.
- Handegard, N. O., Buisson, L. d., Brehmer, P., Chalmers, S. J., Robertis, A., Huse, G., Kloser, R., Macaulay, G., Maury, O., Ressler, P. H., et al. (2013). Towards an acoustic-based coupled observation and modelling system for monitoring and predicting ecosystem dynamics of the open ocean. *Fish and Fisheries*, 14(4):605–615.
- Hazen, E. L. and Johnston, D. W. (2010). Meridional patterns in the deep scattering layers and top predator distribution in the central equatorial pacific. *Fisheries Oceanography*, 19(6):427–433.
- Irigoien, X., Klevjer, T. A., Røstad, A., Martinez, U., Boyra, G., Acuña, J., Bode, A., Echevarria, F., Gonzalez-Gordillo, J., Hernandez-Leon, S., et al. (2014). Large mesopelagic fishes biomass and trophic efficiency in the open ocean. *Nature communications*, 5.
- Jennings, S., Mélin, F., Blanchard, J. L., Forster, R. M., Dulvy, N. K., and Wilson, R. W. (2008). Global-scale predictions of community and ecosystem properties from simple ecological theory. *Proceedings of the Royal Society of London B: Biological Sciences*, 275(1641):1375–1383.
- Kloser, R., Ryan, T., Sakov, P., Williams, A., and Koslow, J. (2002). Species identification in deep water using multiple acoustic frequencies. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(6):1065–1077.
- Kloser, R. J., Ryan, T. E., Young, J. W., and Lewis, M. E. (2009). Acoustic observations of micronekton fish on the scale of an ocean basin: potential and challenges. *ICES Journal of Marine Science: Journal du Conseil*, 66(6):998–1006.
- Kooyman, G. and Kooyman, T. (1995). Diving behavior of emperor penguins nurturing chicks at coulman island, antarctica. *Condor*, pages 536–549.
- Kremer, P. and Kremer, J. N. (1988). Energetic and behavioral implications of pulsed food availability for zooplankton. *Bulletin of Marine Science*, 43(3):797–809.
- Lehodey, P., Conchon, A., Senina, I., Domokos, R., Calmettes, B., Jouanno, J., Hernandez, O., and Kloser, R. (2015). Optimization of a micronekton

- model with acoustic data. *ICES Journal of Marine Science: Journal du Conseil*, 72(5):1399–1412.
- Lehodey, P., Murtugudde, R., and Senina, I. (2010). Bridging the gap from ocean models to population dynamics of large marine predators: a model of mid-trophic functional groups. *Progress in Oceanography*, 84(1):69–84.
- Lezama-Ochoa, A., Irigoien, X., Chaigneau, A., Quiroz, Z., Lebourges-Dhaussy, A., and Bertrand, A. (2014). Acoustics reveals the presence of a macrozooplankton biocline in the bay of biscay in response to hydrological conditions and predator-prey relationships. *PLoS ONE*, 9:88054.
- McClatchie, S. and Dunford, A. (2003). Estimated biomass of vertically migrating mesopelagic fish off new zealand. *Deep Sea Research Part I: Oceanographic Research Papers*, 50(10):1263–1281.
- Mélin, F., Sclap, G., Jackson, T., and Sathyendranath, S. (2016). Uncertainty estimates of remote sensing reflectance derived from comparison of ocean color satellite data sets. *Remote Sensing of Environment*, 177:107–124.
- Morel, A. and Berthon, J.-F. (1989). Surface pigments, algal biomass profiles, and potential production of the euphotic layer: Relationships reinvestigated in view of remote-sensing applications. *Limnology and Oceanography*, 34(8):1545–1562.
- Nakano, H., Matsunaga, H., Okamoto, H., and Okazaki, M. (2003). Acoustic tracking of big-eye thresher shark *alopias superciliosus* in the eastern pacific ocean. *Marine Ecology Progress Series*, 265:255–261.
- O'Driscoll, R. L., Gauthier, S., and Devine, J. A. (2009). Acoustic estimates of mesopelagic fish: as clear as day and night? *ICES Journal of Marine Science/Journal du Conseil*, 66(6).
- Parent, L., Ferry, N., Barnier, B., Garric, G., Bricaud, C., Testut, C.-E., Le Galloudec, O., Lellouche, J., Greiner, E., Drevillon, M., et al. (2013). Global eddy-permitting ocean reanalyses and simulations of the period 1992 to present. Technical report, Technical report, Mercator Ocean.
- Parker-Stetter, S. L. and Horne, J. K. (2009). Nekton distribution and midwater hypoxia: A seasonal, diel prey refuge? *Estuarine, Coastal and Shelf Science*, 81(1):13–18.
- Ryan, T. (2011). Overview of data collection, management and processing procedures of underway acoustic data-imos basoop sub-facility. Technical report.
- Senina, I., Sibert, J., and Lehodey, P. (2008). Parameter estimation for basin-scale ecosystem-linked population models of large pelagic predators: application to skipjack tuna. *Progress in Oceanography*, 78(4):319–335.
- Sims, D. W., Southall, E. J., Tarling, G. A., and Metcalfe, J. D. (2005). Habitat-specific normal and reverse diel vertical migration in the plankton-feeding basking shark. *Journal of Animal Ecology*, 74(4):755–761.
- Tanimura, A., Hattori, H., Miyamoto, Y., Hoshiai, T., and Fukuchi, M. (2008). Diel changes in vertical distribution of *oithona similis* (cyclopoida) and *oncaea curvata* (poecilostomatoida) under sea ice in mid-summer near syowa station, antarctica. *Polar Biology*, 31(5):561–567.
- van Haren, H. and Compton, T. J. (2013). Diel vertical migration in deep sea plankton is finely tuned to latitudinal and seasonal day length. *PloS one*, 8(5):e64435.
- Visser, A. W. and Jónasdóttir, S. H. (1999). Lipids, buoyancy and the seasonal vertical migration of *calanus finmarchicus*. *Fisheries Oceanography*, 8(s1):100–106.
- Volk, T. and Hoffert, M. I. (1985). Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric co₂ changes. *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archean to Present*, pages 99–110.
- Wilkinson, G. and Rogers, C. (1973). Symbolic description of factorial models for analysis of variance. *Applied Statistics*, pages 392–399.

Multi-Gaussian model for boundary layer detection

A Pre-processing

We use vertical acoustic density profiles averaged on a daily period (night, day and migration times) over a $1/4^\circ \times$ day resolution grid with their native vertical resolution (10m for most of the profiles). Profiles collected during migration phases as defined in Lehodey et al. (2014) are not used. A first automatic data screening rejects the profiles if:

- missing values are too numerous (more than 90% of the profile is missing, due to filtering or shallow bathymetry);
- there is no variation in the profile;
- the mean NASC value is too high (2 order of magnitude higher than the mean of the other profiles).

Following this data screening, less than 1% of the profiles were rejected.

Then, the procedure is constrained to find three different layers in each profile. *A priori* information is used to limit the search in which each observed layer boundary can be found, i.e., within 25-250m, 285-700m and 600-900m for the 1st, 2nd and 3rd layer boundaries respectively (table 3). This depth range $[z_n^{min}, z_n^{max}]$ is fixed for the whole dataset (n stands for the layer, $n \in [1, 3]$).

B Gaussian Mixture Model

Let P be the average vertical acoustic density profile, then the restricted profile p_n is :

$$p_n(z) = \begin{cases} P(z), & \text{if } z \in [z_n^{min}, z_n^{max}], \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

A Gaussian Mixture Model (GMM) is fitted to the restricted profile p_n . We use the mclust package (Fraley and Raftery (2002) and Fraley et al. (2012)) of the R software.

The mclust package fits a probability distribution on an ensemble of realisations. This ensemble is created

to have a probability distribution fitting the restricted profile. The restricted profile is approximated with N_n^G Gaussian functions (table 3).

$$p_n(z) \approx \sum_{i=1}^{N_n^G} \frac{w_n^i}{\sqrt{2\pi} \sigma_n^i} \exp\left(-\frac{(z - \mu_n^i)^2}{2 \sigma_n^i{}^2}\right) \quad (2)$$

where, w_n^i , μ_n^i and σ_n^i are the weight, mean and standard deviation of the i^{th} Gaussian function approximating the p_n restricted profile. Gaussians functions are ranked according to their mean:

$$\forall i \in [1, N_n^G], \mu_n^i < \mu_n^{i+1} \quad (3)$$

It is necessary to add filtering for either small values of w_n^i or large values of σ_n^i . Small values of w_n^i can be the result of noise (associated to large sigma) or a very small localized variation of signal (associated to small sigma). Very large value of sigma is the signature of a Gaussian function that captures noise and should be filtered too. Finally, spurious detections of excessively deep or shallow mean depth values of μ_n^i can be generated by the GMM due to the use of restricted profiles. These outlier data are rejected with a filter based on minimum and maximum rational values. This filtering procedure is implemented using a weighting coefficient α_i :

$$\alpha_n^i = \begin{cases} 1, & \text{if } w_n^i < w_n^{min} \text{ or } \sigma_n^i > \sigma_n^{max} \text{ or } \mu_n^i > \mu_n^{max} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

and the following approximation of the restricted transect:

$$p_n(z) \approx \sum_{i=1}^{N_n^G} \alpha_n^i \frac{w_n^i}{\sqrt{2\pi} \sigma_n^i} \exp\left(-\frac{(z - \mu_n^i)^2}{2 \sigma_n^i{}^2}\right) \quad (5)$$

C Aggregation of Gaussian functions

Selected Gaussians functions (with $\alpha_n^i = 1$) are aggregated into sub-layers according to their extension defined by upper and lower values t_n^i and b_n^i :

$$t_n^i = \mu_n^i - \varpi_n \sigma_n^i \text{ and } b_n^i = \mu_n^i + \varpi_n \sigma_n^i \quad (6)$$

If the upper value of the next adjacent Gaussian function is above the lower limit of the previous function (if they overlap), then the two Gaussian functions are aggregated. More generally:

$$p_n(z) \approx \underbrace{\sum_{i \in G_n^1} \alpha_n^i \frac{w_n^i}{\sqrt{2\pi} \sigma_n^i} \exp\left(-\frac{(z - \mu_n^i)^2}{2 \sigma_n^i{}^2}\right)}_{1^{\text{st}} \text{ sublayer}} + \dots + \underbrace{\sum_{i \in G_n^m} \alpha_n^i \frac{w_n^i}{\sqrt{2\pi} \sigma_n^i} \exp\left(-\frac{(z - \mu_n^i)^2}{2 \sigma_n^i{}^2}\right)}_{m^{\text{th}} \text{ sublayer}} \quad (7)$$

where $\{G_n^m\}$ is a partition of $[1, N_n^G]$. We define $\{G_n^m\}$ recursively as :

$$G_n^m = [i_{min}^m, i_{max}^m], \text{ with } \begin{cases} i_{min}^m = \beta_{max}^{m-1} + 1, \\ i_{max}^m = \min_{i \in [i_{min}^m, N_n^G]} (i, t_n^i > b_n^{i-1}), \\ i_{min}^0 = 1, \\ i_{max}^{N_n^{SL}} = N_n^G. \end{cases} \quad (8)$$

where N_n^{SL} is the number of sublayers.

This way, elements of G_n^m follow:

$$\forall m \in [1, N_n^{SL}], \forall i \in G_n^m, t_n^i < b_n^{i-1} \quad (9)$$

D Limit determination

The aggregation of Gaussian functions produces a series of sub-layers from which the boundary layer values need to be identified. If an empty surface layer is detected, the first layer limit is the bottom of this empty surface layer. In all other cases, the limit is either the actual bottom of the first sublayer or the intersection between the first and the second ones if they are overlapping.

$$\max_{i \in G_n^1} (\mu_n^i + 3\sigma_n^i) < \min_{i \in G_n^2} (\mu_n^i - 3\sigma_n^i) \implies l_n^{obs} = \max_{i \in G_n^1} (\mu_n^i + 3\sigma_n^i) \quad (10)$$

$$\begin{aligned} & \max_{i \in G_n^1} (\mu_n^i + 3\sigma_n^i) > \min_{i \in G_n^2} (\mu_n^i - 3\sigma_n^i) \\ \implies l_n^{obs} &= \left\{ z, \sum_{i \in G_n^2} \alpha_n^i \frac{w_n^i}{\sqrt{2\pi} \sigma_n^i} \exp\left(-\frac{(z - \mu_n^i)^2}{2 \sigma_n^i{}^2}\right) - \sum_{i \in G_n^1} \alpha_n^i \frac{w_n^i}{\sqrt{2\pi} \sigma_n^i} \exp\left(-\frac{(z - \mu_n^i)^2}{2 \sigma_n^i{}^2}\right) = 0 \right\} \end{aligned} \quad (11)$$

If the observed limit is too deep ($l_n^{obs} > z_n^{max}$), the profile is rejected (less than 5% of profiles).

Parameter	Name	1rst limit	2nd limit	3rd limit
z_n^{min}	Minimal depth	25	285	600
z_n^{max}	Maximal depth	250	700	900
N_n^G	Number of gaussians	4	7	7
ϖ_n	Gaussian aggregation coefficient	2.5	2.0	2.0
μ_n^{max}	Mean limit	225	975	975
σ_n^{max}	Standard deviation limit	1000	2000	6000
w_n^{min}	Weight limit	0.1	0.05	0.1
$l_n^\emptyset$	Depth limit empty layer	50	285	600

Table 3: Gaussian mixture model and aggregation procedure parameters

III Une nouvelle définition des couches

A Résultats de la méthode de détection des couches

La méthode développée dans la partie II de ce chapitre permet de détecter de façon précise les limites inférieures des trois couches généralement observées par acoustique. Cette méthode est entièrement automatisée. De plus, la limite de chaque couche est cherchée indépendamment des autres.

La figure 4.4 illustre l'intérêt de ce travail. Elle reprend la figure 4.1 sur laquelle s'ajoute en bleu les limites détectées par la méthode décrite précédemment et en vert les profondeurs prédites par le modèle choisi.

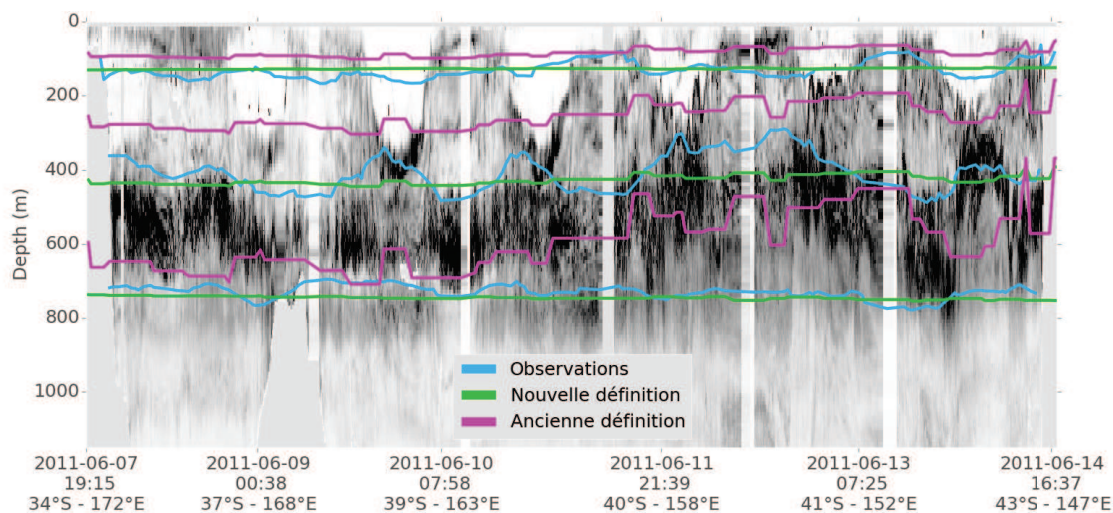


Figure 4.4 – Exemple du transect acoustique entre la Nouvelle-Zélande et la Tasmanie de la figure 4.1 avec les limites observées et prédites par l'ancien et le nouveau modèle.

Les définitions précédentes engendraient des discontinuités dans les profondeurs prédites (notamment pour la 3ème couche), un lissage était alors nécessaire lors de l'utilisation de telles profondeurs de couche dans le pré-traitement des champs d'entrée du modèle SEAPODYM. On observe une grande distance entre les limites observées et les limites prédites avec l'ancienne définition. Dans cet exemple là, les limites prédites avec l'ancienne définition sont systématiquement moins profondes que les limites observées et cette distance augmente vers la Tasmanie (partie droite de la figure). Les variations des limites prédites avec l'ancienne définition ne sont pas corrélées avec les limites observées. De plus les limites observées ne sont pas corrélées entre couches, une définition du type $(\alpha_1 \times Zeu, \alpha_2 \times Zeu, \alpha_3 \times Zeu)$ ne peut donc pas être adaptée à ce transect. L'article présenté dans la partie II de ce chapitre approfondit ces remarques et les élargit à l'ensemble de notre jeu de données.

Les profondeurs prédites avec la nouvelle définition sont continues et très lisses. Sur cette figure, on observe quasiment aucune variation de celle-ci. Ces profondeurs prédites ne reproduisent pas les variations à fine échelle des profondeurs observées. Pour ce faire, un modèle plus complexe aurait pu être choisi mais ce n'était pas le but ici. Nous cherchions à mettre en place un modèle simple, ne présentant pas de discontinuité, et s'obtenant avec des variables disponibles globalement. Nous avons obtenu un modèle linéaire d'une variable satellitaire pour les deux premières limites. Le modèle moyen (740m) est le modèle le plus

adapté, pour l'instant, pour la troisième limite.

B Impact sur l'optimisation

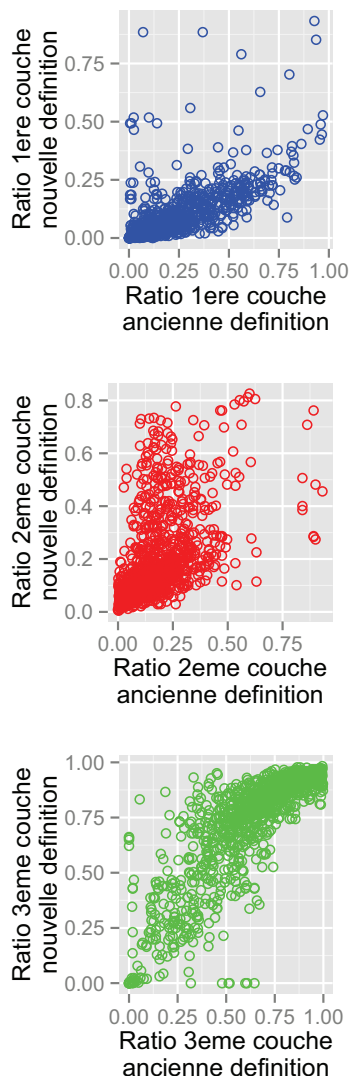


Figure 4.5 – Ratios de densité acoustique déterminés avec le nouvelle et l'ancienne définition.

Le but initial de la conception de cette méthode de détection et du choix d'une nouvelle définition des limites verticales du modèle SEAPODYM était d'éliminer ou de réduire une incertitude dans les expériences d'assimilation de données acoustiques afin de déterminer notamment l'uniformité de certains paramètres.

Afin de quantifier l'impact de ces nouvelles définitions, il faudra refaire ces expériences à l'identique avec la nouvelle définition. Dans un premier temps, nous avons calculé les ratios de densité acoustique observés avec ces nouvelles définitions. La figure 4.5 présente la valeur des ratios observés avec la nouvelle définition en fonction de leur valeur avec l'ancienne définition. On n'observe pas de tendance pour la seconde couche. Pour la première couche, les ratios avec la nouvelle définition sont plus faibles (72.2% inférieur à 25% contre 95.1% avec l'ancienne définition). Quant à la troisième couche : c'est l'inverse les ratios sont plus élevés (38.7% supérieur à 75% contre 65.4% avec l'ancienne définition).

Cette nouvelle définition va également impacter les valeurs des courants et de la température qui sont moyennés sur les couches du modèle. Il n'est donc pas possible de spéculer sur les futurs résultats des expériences d'assimilation.

Après avoir effectué les expériences d'assimilation identiques à celles présentés dans le Chapitre 3, il faudra inclure les paramètres de croissance afin d'obtenir des résultats sur l'uniformité des coefficients de distribution d'énergie.

Bibliographie

Lehodey, P., Conchon, A., Senina, I., Domokos, R., Calmettes, B., Jouanno, J., Hernandez, O., and Kloser, R. (2014). Optimization of a micronekton model with acoustic data. *ICES Journal of Marine Science : Journal du Conseil*, page fsu233.

Chapitre 5 Perspectives

Sommaire

I	Paramétrisations optimales	108
A	SEAPODYM-LTL	108
B	SEAPODYM-MTL	109
II	SEAPODYM-MTL et les différentes chaînes trophiques	110
A	Résultats préliminaires	112
B	Perspectives	117
III	Établir un plan d'échantillonnage idéal du micronecton	117
IV	Étudier l'accumulation de polluants dans la chaîne trophique	118

I Paramétrisations optimales

A SEAPODYM-LTL

Dans le Chapitre 2 les paramétrisations ont été établies d'après la littérature et par tâtonnement. Puisque les développements mathématiques et informatiques pour l'optimisation de paramètres (code adjoint, algorithme de descente de gradient...) ont été réalisés pour la configuration micronecton, il semble pertinent de les adapter à celle du zooplancton, d'autant plus qu'il existe un grand nombre de données de biomasse de zooplancton. L'approche peut être facilement étendue à l'estimation du coefficient de transfert d'énergie comme pour le micronecton, mais également aux paramètres de la relation liant le temps de développement des organismes à la température du milieu. Une fonction mathématique proposant plus de flexibilité dans sa forme que la relation exponentielle décroissante devra être testée. En effet, les résultats obtenus ici montrent que dans les eaux froides en dessous de 10°C les biomasses sont surestimées par le modèle. Les expériences en laboratoires n'ont pas ou très peu considéré les espèces d'eaux très froides ce qui peut conduire à une forte incertitude sur la forme de la relation générale. Une relation de type sigmoïde par exemple pourrait permettre certainement un meilleur ajustement aux observations (Figure 5.1).

D'un point de vue théorique, la justification d'une forme de fonction sigmoïde plutôt qu'exponentielle pourrait s'expliquer par l'adaptation des espèces à leur milieu. En effet les résultats obtenus en laboratoire et utilisés au Chapitre 2 concernent des temps de développements à différentes températures pour les mêmes espèces. Or, la relation cherchée doit émerger d'une moyenne sur une grande diversité d'espèces représentant l'ensemble du spectre de températures de leurs habitats. Il faut notamment prendre en compte l'adaptation de ces espèces aux températures froides pour lesquelles les organismes ectothermes ont développés des

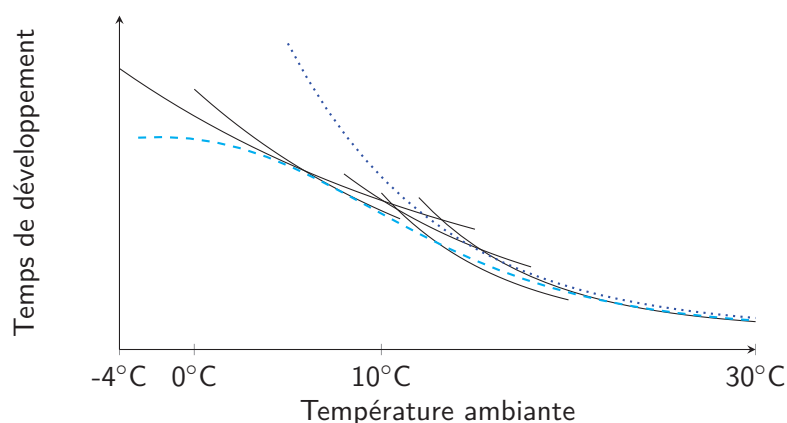


Figure 5.1 – Illustration de la relation entre temps de développement et température du milieu pouvant émerger à partir d'un ensemble d'espèces (traits fins continus pour chaque espèce et pointillées bleu pour l'ensemble) dont le métabolisme est adapté à la gamme de températures de leur milieu. L'absence de données de laboratoire pour les espèces d'eaux froides ($<10^{\circ}\text{C}$) peut fortement influencer la forme de la relation générale (tirets bleu clair).

évolutions remarquables. Ainsi, une plus grande efficacité des réactions enzymatiques à basse température chez les ectothermes comparés aux oiseaux et mammifères a été démontrée depuis longtemps (Low et al. (1973)). Les poissons de la famille des Notothenoidés dont certains vivent dans les eaux très froides proches du point de congélation (-1.86°C pour l'eau de mer) montrent également une adaptation des propriétés de cinétiques enzymatiques à ces conditions extrêmes, avec des constantes de Michaelis-Menten ou de taux de catalyse plus élevés que chez des espèces d'eaux tempérées-froides de la même famille (Fields and Somero (1998)).

B SEAPODYM-MTL

Dans le cadre de cette thèse, nous nous sommes attachés à améliorer la paramétrisation du modèle de micro-necton par la mise en place d'une méthodologie d'assimilation de données ainsi qu'une amélioration des définitions des profondeurs limites des trois couches verticales du modèle. Les premiers résultats d'optimisation (Chapitre 3) et ceux sur l'application au zooplancton suggèrent la nécessité d'une révision de la paramétrisation des temps de développements notamment pour les eaux froides, impliquant comme discuté ci-dessus l'utilisation d'une relation mathématique différente. Cela représente des changements relativement marginaux sur le code (direct et adjoint) et ne devrait pas poser de problème particulier. Néanmoins, il faut prêter attention au plan d'expérience. En effet, dans la version actuelle les champs de production sont calculées séparément (off-line), ce qui implique d'analyser la sensibilité des paramètres de mortalité (inversement proportionnel au temps de développement) estimés par optimisation à différentes valeurs de temps de recrutement. Quelques tests préliminaires indiquent néanmoins que cette sensibilité est faible et donc que la méthodologie est bien adaptée.

Dans le cadre de l'optimisation des paramètres du modèle à l'aide des données acoustique, un problème plus difficile se pose probablement avec l'interprétation de l'observation elle-même. Dans le Chapitre 3, il a été posé comme hypothèse que le coefficient de proportionnalité entre la densité acoustique et la biomasse

de micronecton ne varie pas avec la profondeur. Ainsi on peut utiliser des ratios de densités acoustiques observées et les comparer directement avec des ratios de biomasses prédites. Mathématiquement, on a donc :

$$NASC(z) = \alpha b(z) \implies \frac{\int_0^{lim1} NASC dz}{\int_0^{lim3} NASC dz} = \frac{\int_0^{lim1} b dz}{\int_0^{lim3} b dz}, \quad (5.1)$$

Toutefois la réalité est plus complexe. Ce coefficient dépend en effet des assemblages d'espèces présentes et de leurs caractéristiques (présence de vessie natatoire, nature de leur enveloppe...). Il n'est donc pas constant avec la profondeur. La mesure elle-même intègre plusieurs sources de bruits qu'il faut pouvoir traiter au mieux. Il est donc nécessaire d'envisager le développement d'un modèle d'observation qui puisse faire le lien entre la mesure acoustique et la prédiction de biomasse. Cette étape est prévue dans le cadre du projet européen MESOPP (Mesopelagic Southern Ocean Prey and Predators). Ce projet est basé sur un réseau de collaborations entre équipes européennes et australiennes travaillant sur les écosystèmes de l'Océan Austral et de l'Antarctique. Il a pour objectifs de développer une méthodologie commune pour le traitement de la donnée acoustique associée au micronecton, de partager les données collectées, de les utiliser pour la validation et l'optimisation du modèle micronecton SEAPODYM-MTL, et ensuite pour analyser le comportement des prédateurs (otaries, manchots, éléphants de mer, albatros...).

Finalement, un autre développement du modèle concerne la création d'un 7ème groupe fonctionnel représentant les organismes benthiques du même spectre de taille (ou plutôt de même temps de développement). L'introduction de ce nouveau groupe permettra une meilleure représentation des flux d'énergie dans les régions peu profondes, à l'approche des côtes et plateaux continentaux, là où la bathymétrie implique de réduire le nombre de couches verticales du modèle à 2 ou 1. Dans ce cas, la version actuelle du modèle redirige les flux d'énergie des groupes qui disparaissent proportionnellement vers ceux qui restent. Ainsi lorsqu'il ne reste plus qu'une seule couche verticale, le transfert est de 100% vers le groupe épipélagique qui est le seul restant. Cette simplification due à un modèle développé au départ pour l'océan du large entraîne très certainement une surestimation de la production et biomasse des groupes représentés dans ces cas particuliers, du fait de l'absence de prise en compte du flux d'énergie vers le fond et le système benthique. Il faudra bien entendu estimer la valeur de transfert d'énergie vers ce nouveau groupe à partir de la littérature, mais également à l'aide de données acoustiques.

II SEAPODYM-MTL et les différentes chaînes trophiques

En dehors de leur utilisation directe avec le modèle prédateur de SEAPODYM, les champs de biomasse du modèle de micronecton connaissent déjà de nombreuses applications : prédictions d'habitats pour des tortues marines (Abecassis et al. (2013)) ou des cétacés (Lambert et al. (2014); Roberts et al. (2016)), analyse des données de bio-logging de tortues marines (Chambault et al. (2016)), d'éléphants de mer (Yves Le Bras 2013, pers. Comm.) ou d'otaries (Laurianne Massardier, pers. Comm.). L'ensemble de ces applications est basée sur le lien proie-prédateur (Figure 5.2). Toutefois, les espèces cibles ne partagent pas les mêmes proies mais le même échelon trophique.

Par ailleurs, les groupes fonctionnels incluent des espèces très variées, et potentiellement du micronecton gé-

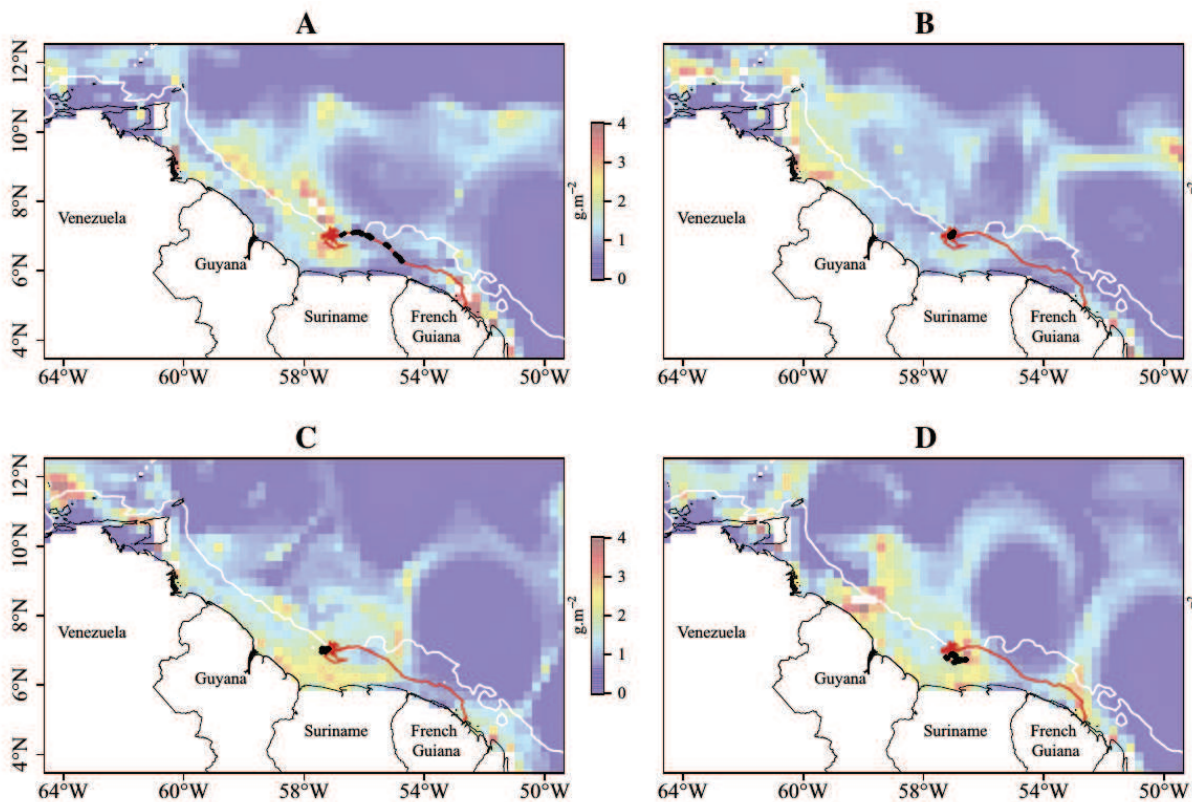


Figure 5.2 – Position d'une tortue olivâtre superposée à la biomasse épipélagique de micronecton dans la zone. Trajectoire totale en rouge, position de la tortue dans la semaine correspondant à la figure en noire. (A) 12-18 Aout 2013, (B) 9-15 Septembre 2013, (C) 14-20 Octobre 2013 et (D) 11-17 Novembre 2013. Extrait de Chambault et al. (2016)

latineux. Bien qu'on puisse ranger les salpes (tuniciers) dans le micronecton d'après leur taille (quelques cm à une dizaine de cm), elles se nourrissent essentiellement sur le phytoplancton et les bactéries (Sutherland et al. (2010)) et rentrent en compétition avec d'autres organismes brouteurs (zooplancton, micronecton). Ces organismes gélatineux peuvent servir de proies, mais sont souvent vus comme compétiteurs de la chaîne alimentaire plus classique phytoplancton (diatomées) => zooplancton => poissons. Plus généralement un grand nombre de questions se posent quand à la place et au rôle des organismes gélatineux dans l'écosystème marin. Leur utilisation comme proies a certainement été sous-estimée (Arai (2005)) du fait de leur digestion rapide, et ils semblent particulièrement bien adaptés pour profiter des perturbations de l'écosystème marin, naturelles (variabilité climatique) ou anthropiques (pêche, eutrophisation, changement climatique). Le débat sur l'accroissement de la biomasse de gélatineux en lien avec le changement climatique et la surpêche revient d'ailleurs régulièrement à la une des médias (Brotz et al. (2012) ; Purcell et al. (2007) ; Purcell (2012) ; Quiñones et al. (2013)).

L'intérêt est donc grand de pouvoir différencier au sein des groupes fonctionnels du micronecton des sous-groupes représentant plutôt des organismes gélatineux ou des organismes de types poissons-crustacés-calmars. Les blooms de gélatineux (par exemple les salpes) semblent associés à leurs facultés à filtrer des particules de très petites tailles. Il est possible également que les diatomées avec leurs spicules siliceuses encombrant rapidement les voies digestives de ces organismes. On peut donc schématiquement simplifier le système en deux chaînes trophiques (Figure 5.3) :

— une chaîne où le petit phytoplancton (haptophytes) et les bactéries dominent et conduisent à une

- production dominée par des organismes gélatineux,
- une chaîne où le gros phytoplancton (diatomées) domine la production primaire et conduit à des organismes de types poissons, crustacés, calmars.

Quelques simulations préliminaires ont donc été tentées avec le modèle SEAPODYM-MTL. Pour cela il est nécessaire d'avoir des champs de production primaire nette associés aux deux types ("petit" et "gros") phytoplanctons. Nous aurions pu choisir d'utiliser des sorties de modèle biogéochimiques de type NPZD disposant de ces variables mais nous avons préféré dans un souci de réalisme opter pour les données satellitaires dans un premier temps. Notamment, l'algorithme de classement PHYSAT (Alvain et al. (2005)) permet de classer les longueurs d'ondes reçues par le satellite en fonction de 5 types phytoplanctoniques (diatomées, nanoeucaryotes, phaeocystis-like, prochlorococcus et synechococcus), soit un groupe de gros phytoplancton (diatomées) et un autre de petits (haptophytes et microbes). Nous allons ici utiliser les distributions spatiales issues de cet algorithme dans sa version de 2008 (Figure 5.4).

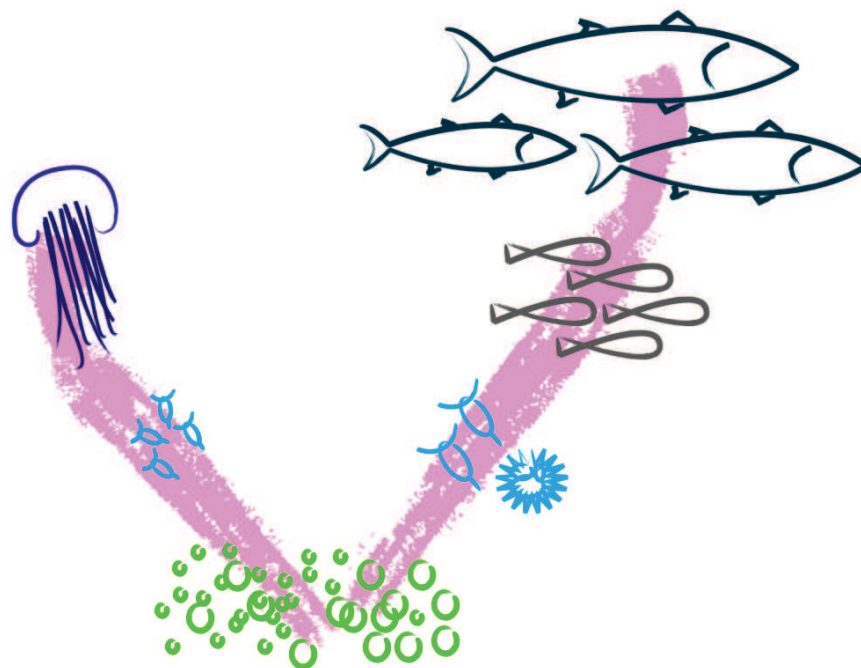


Figure 5.3 – Les chaînes trophiques simplifiées modélisables par la suite de modèles SEAPODYM. La ligne rose figure le flux d'énergie entre niveaux trophiques, les ronds verts le phytoplancton (petit et gros, non-modélisé dans SEAPODYM), le zooplancton en bleu (petit et gros), le micronecton en gris, les organismes gélatineux en violet et les poissons en bleu foncé.

A Résultats préliminaires

Les champs PHYSAT sont disponibles à la résolution $1/12^\circ \times 1$ mois entre 1998 et 2010. Cependant, comme on le voit sur la figure 5.4 de nombreuses longueurs d'onde ne sont pas attribuées à un groupe phytoplanctonique, créant des données manquantes. Afin de limiter leur nombre, nous avons choisi de diminuer la résolution spatiale au $1/4^\circ$ et de considérer uniquement une climatologie mensuelle. Dans chaque cellule au $1/4^\circ$, nous avons calculé le pourcentage de cellules initiales attribuées aux diatomées (en prenant en compte les données manquantes). Dans les cellules ne comptant aucune donnée valide, nous avons choisi d'attribuer la valeur 0%. Ces pourcentages ont ensuite été multipliés aux valeurs de production primaire

nette associée aux mêmes cellules.

Les autres forçages du modèle SEAPODYM-MTL (courants et températures) sont issus de la réanalyse GLORYS2 (Parent et al. (2013)) à la résolution $1/4^\circ \times$ semaine. Une climatologie de ces forçages a été réalisée sur la période 1998-2009. La paramétrisation des deux chaînes trophiques est, ici, identique à la paramétrisation utilisée par Lehodey et al. (2010).

La figure 5.5 (resp. fig. 5.6) représente la distribution spatiale du micronecton issue des diatomées (resp. du petit phytoplancton) en janvier. La figure 5.7 montre la proportion relative de chaque groupe. On observe des biomasses élevées de micronecton issues des diatomées dans des zones très localisées (upwelling du Pérou, upwelling du Benguela, mer du Japon, Géorgie du Sud, Atlantique nord). Le micronecton issu du petit phytoplancton est beaucoup plus abondant et domine quasiment partout, à l'exception de l'upwelling du Pérou où il y a codominance de ces deux types de micronecton. Le micronecton issu du petit phytoplancton est aussi moins localisé et présent dans l'ensemble de l'océan global à l'exception des centres de gyres.

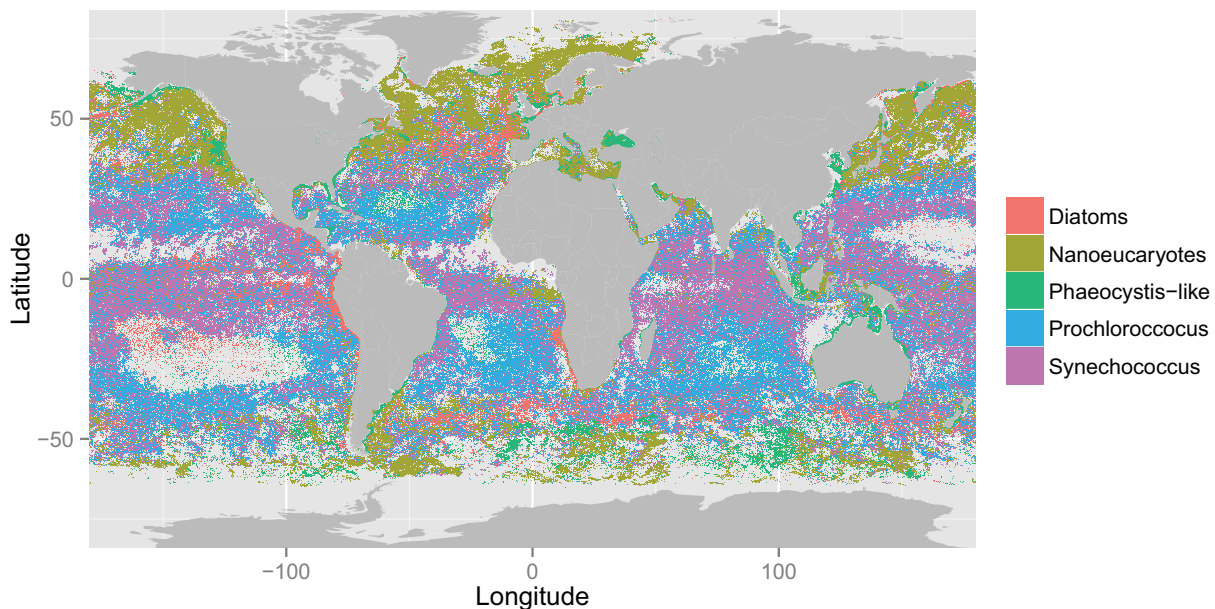


Figure 5.4 – Groupes phytoplanctoniques majoritaires selon l'algorithme PHYSAT (avril 2004)

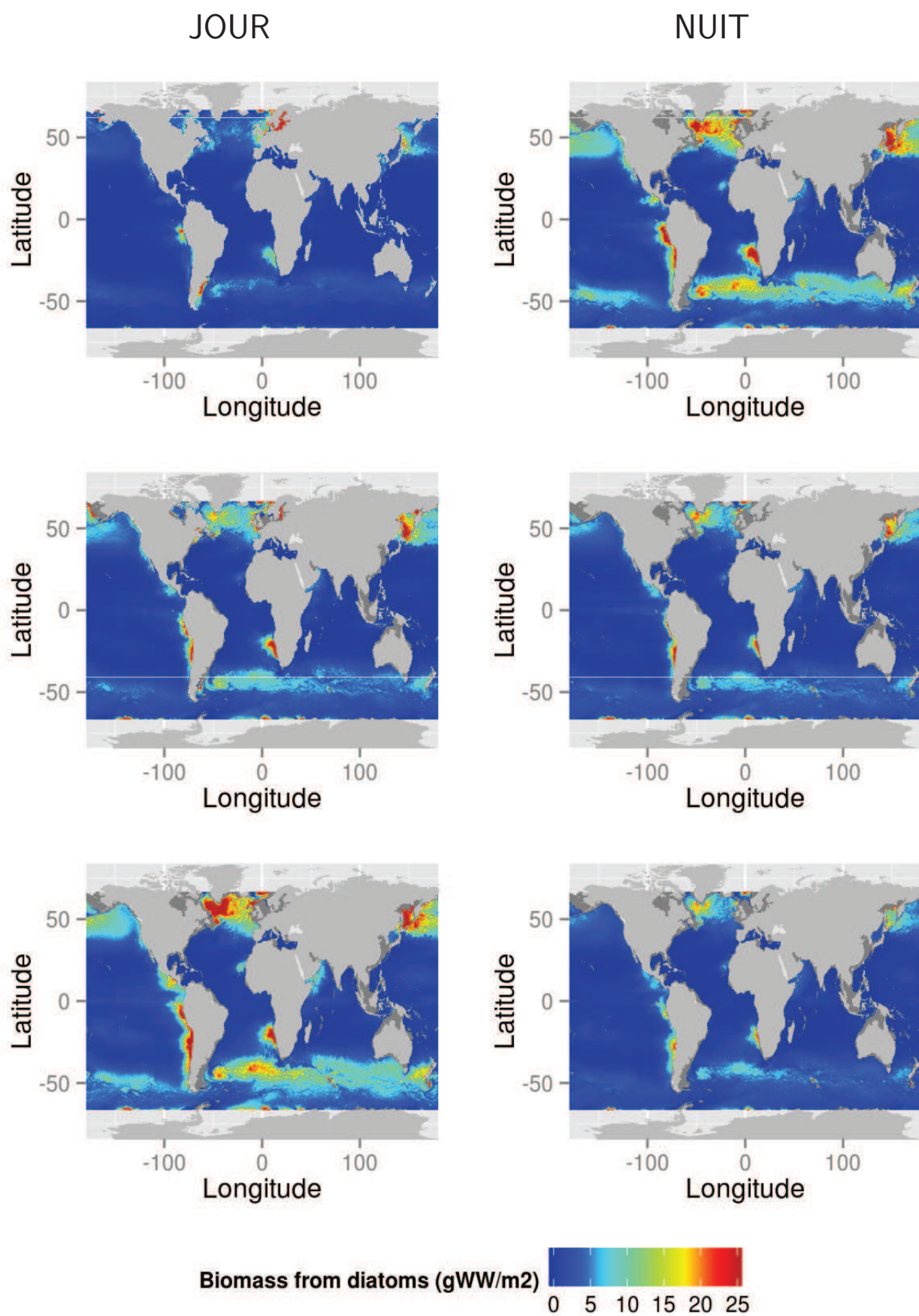


Figure 5.5 – Distribution spatiale du micronecton issue des diatomées en (gWW/m²). La colonne de gauche correspond au distribution de jour et celle de droite de nuit. La première ligne présente la première couche, la deuxième ligne la deuxième couche et la troisième ligne la troisième couche.

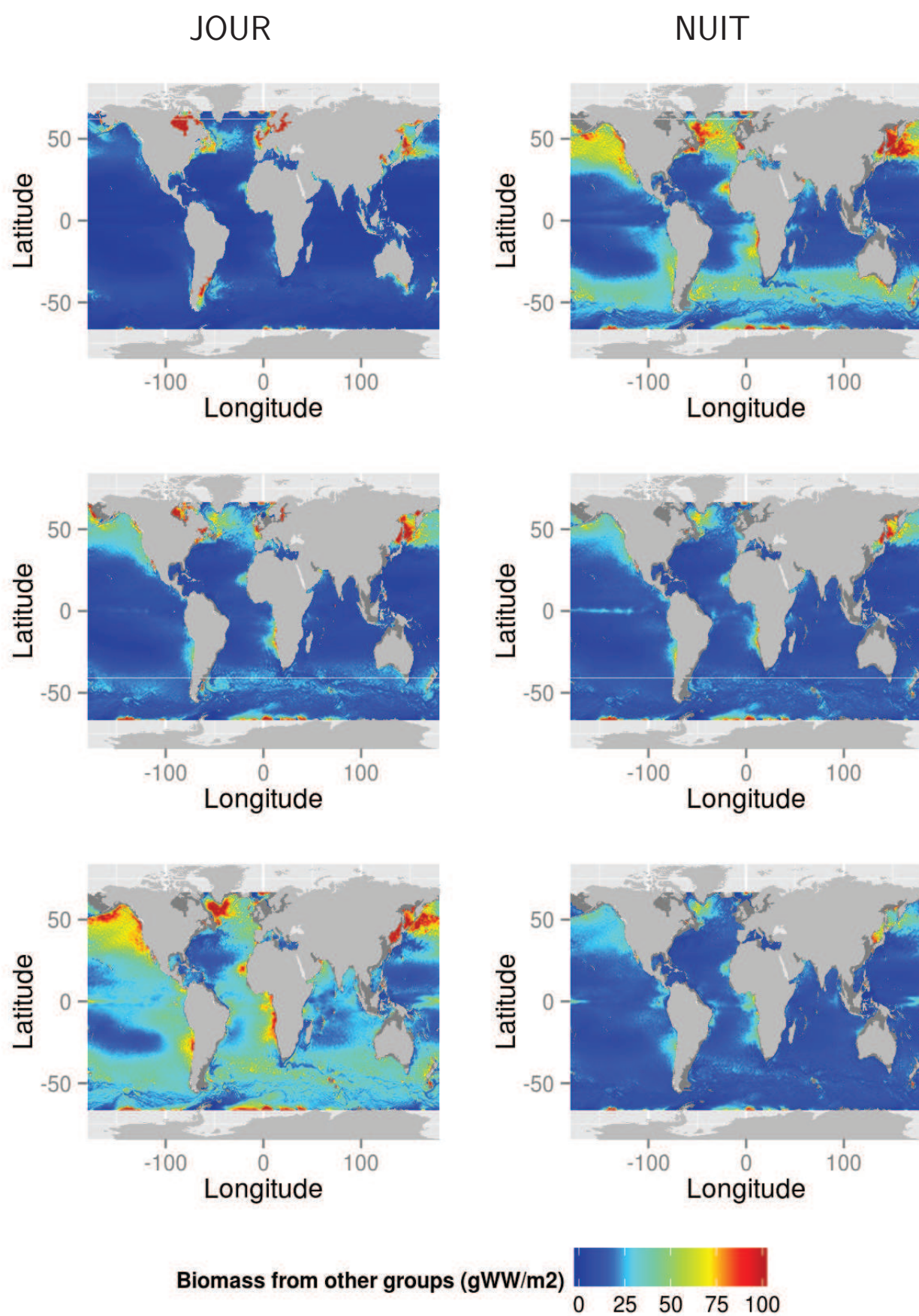


Figure 5.6 – Distribution spatiale du micronecton issue du petit phytoplancton en (gWW/m²). La colonne de gauche correspond au distribution de jour et celle de droite de nuit. La première ligne présente la première couche, la deuxième ligne la deuxième couche et la troisième ligne la troisième couche.

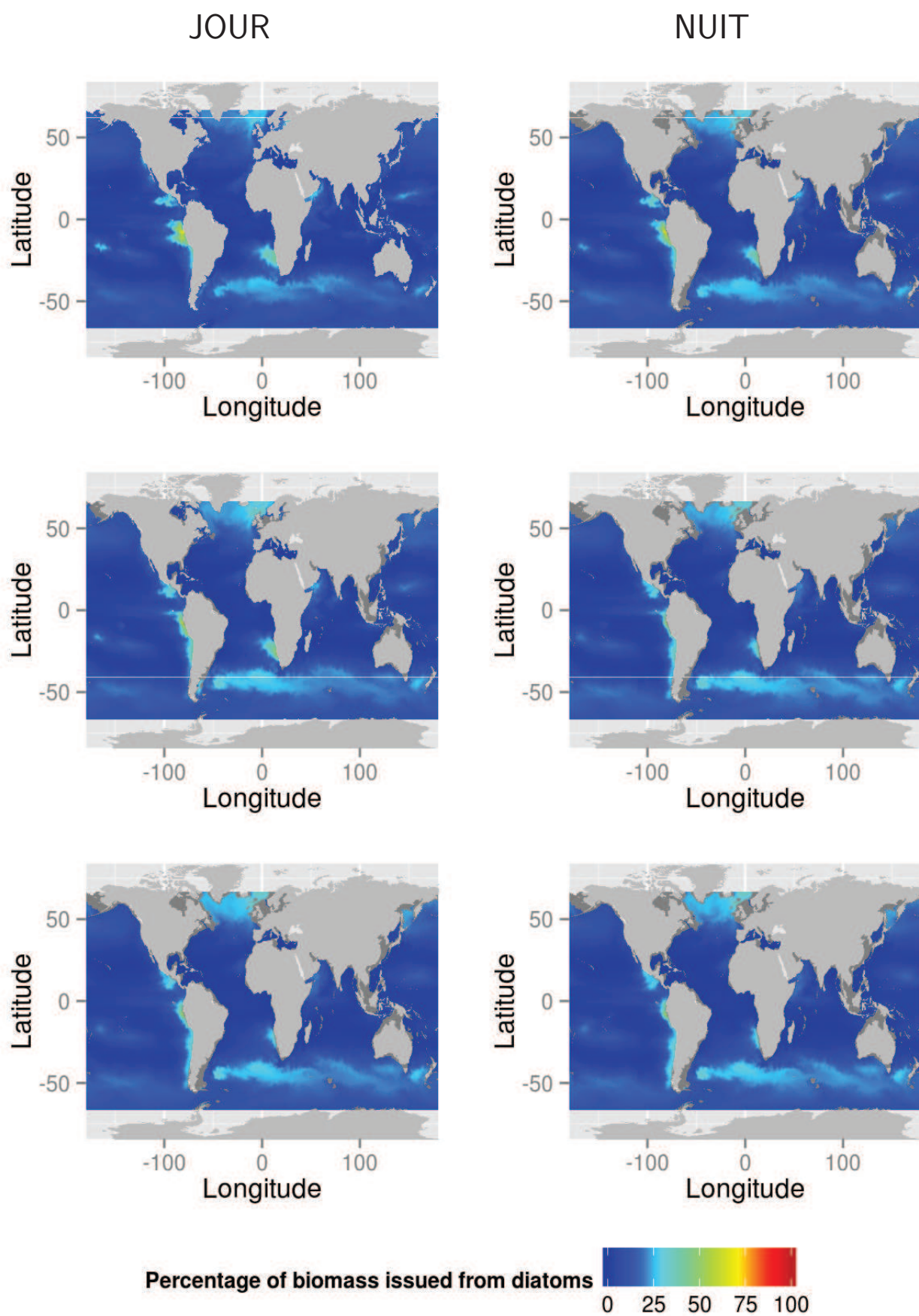


Figure 5.7 – Pourcentage du micronecton issue des diatomées. La colonne de gauche correspond au distribution de jour et celle de droite de nuit. La première ligne présente la première couche, la deuxième ligne la deuxième couche et la troisième ligne la troisième couche.

B Perspectives

Deux axes d'améliorations sont à envisager. Tout d'abord une meilleure estimation de l'apport de chaque groupe phytoplanctonique à la production primaire totale. Il sera intéressant d'utiliser ces nouveaux produits comme forçage des chaînes trophiques du modèle. Un traitement approprié des données manquantes par interpolation et rappel à la climatologie devra être mis en place. Ensuite un travail conséquent demeure pour paramétrer les coefficients de transfert d'énergie vers ces différents groupes et leurs relations temps de développement-température. Une fois l'information disponible dans la littérature passée en revue, il faudra certainement compter sur la mise en place de réseaux d'observation, de nouveaux développements technologiques comme les sondeurs acoustiques multifréquences et les scanners optiques, et des études en laboratoire pour préciser peu à peu ces paramétrisations.

III Établir un plan d'échantillonnage idéal du micronecton

A l'aide de la méthodologie développée dans le Chapitre 3 et [Lehodey et al. \(2015\)](#), il est possible de réaliser des expériences jumelles pour définir un plan d'échantillonnage acoustique idéal du micronecton. Le projet H2020 AtlantOS vise à améliorer les systèmes d'observations de l'océan Atlantique. C'est un projet multi-disciplinaire s'intéressant aussi bien à la physique, qu'à la biogéochimie ou à la biologie. Dans le cadre de ce projet, des expériences de simulations de systèmes d'observation (Observing System Simulation Experiment, OSSE) vont être effectuées pour différentes variables, notamment le micronecton.

Dans le cadre de ces OSSEs, on se focalisera sur l'approche acoustique. Trois dispositifs (représentés sur la figure 5.8) peuvent être envisagés : transect par bateau, mouillage ancré sur le plancher océanique et mouillage en eau intermédiaire. En dehors du fait que le type de données récoltées est différent (séries temporelles ou transect), ces moyens d'observation impliquent des erreurs différentes, notamment au niveau de leurs localisations, des coûts différents et des contraintes différentes. Il faudra donc choisir un dénominateur commun à chaque expérience, le coût total du dispositif pouvant être constant par exemple. L'étude de l'impact de la durée de l'acquisition des données peut être également envisagée.

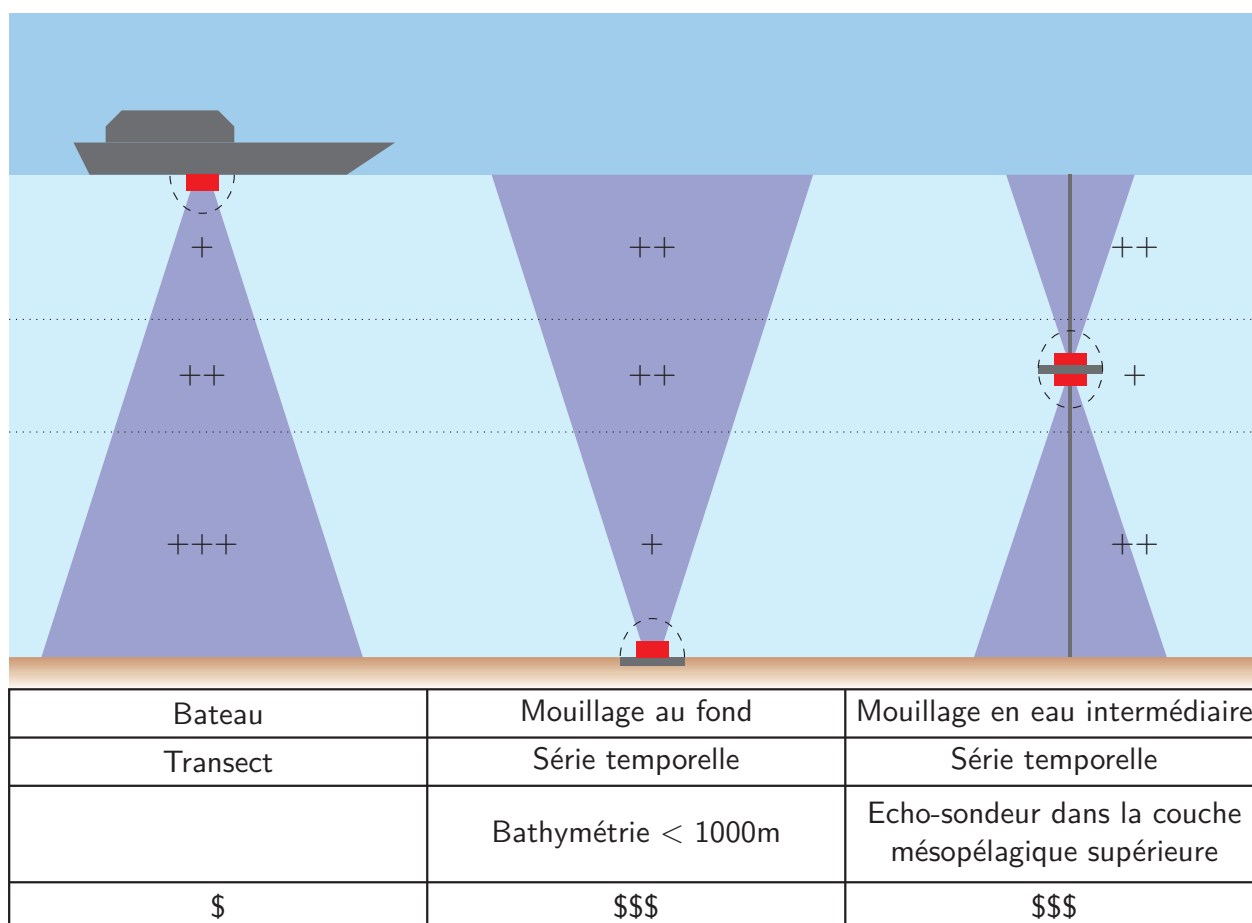


Figure 5.8 – Différents dispositifs d'observations du micronecton par acoustique, avec le type de données acquises, les contraintes liées au dispositif et un indice du coût.

IV Étudier l'accumulation de polluants dans la chaîne trophique

L'écotoxicologie marine est un domaine de recherche en soi. Les problématiques associées sont souvent assez cotières mais dans certains cas peuvent aussi concerner l'océan du large par exemple dans le cas d'accidents industriels majeurs. Il y a un intérêt majeur à pouvoir modéliser le transfert de produits toxiques tels que les métaux lourds, mais aussi dans certains cas des éléments radioactifs (Tchernobyl, Fukushima), dans la chaîne trophique marine. Ces modélisations ont potentiellement des applications directes pour la santé publique, la prévoyance et le suivi sanitaire.

L'idée d'utiliser le modèle de micronecton SEAPODYM-MTL pour simuler le transport et l'accumulation de radionucléides relargués par l'explosion de la centrale nucléaire de Fukushima au Japon en 2011 est donc venue assez naturellement et a été proposée avec succès dans le cadre d'un appel d'offre de l'Agence Nationale de la Recherche. Le projet (AMORAD) piloté par l'IRSN (Institut de Radioprotection et de Sureté Nucléaire) inclut donc la proposition de modéliser le transfert du Césium 137 aux groupes de micronecton, puis d'utiliser ces résultats pour prédire les zones de risques de capture de thons ayant potentiellement mangé du micronecton contaminé. L'objectif final étant de pouvoir développer un outil qui dans un cas similaire devrait être mis en place pour faciliter la surveillance et le contrôle des poissons importés.

Le césium est l'un des éléments largement rejetés lors de l'accident de Fukushima, et le seul élément ra-

Figure 5.9 – Premier test de simulation du transfert de Cs137 au micronecton du modèle SEAPODYM-MTL

bioactif pour lequel une légère biomagnification est observée (accroissement de la concentration à chaque niveau trophique). A partir d'un forçage de ce polluant fourni par les partenaires du projet, il s'agit d'intégrer et de calibrer les transferts qui s'opèrent via l'adsorption et l'ingestion du césium aux différents échelons trophiques. Les processus de transfert des radionucléides dans les organismes et les chaînes trophiques sont des processus complexes dépendants de très nombreux facteurs comme par exemple le temps d'exposition, la spéciation de l'élément, les caractéristiques physico-chimiques du milieu, le cycle de vie de l'organisme, les régimes alimentaires, etc. De ce fait des approches très pragmatiques sont utilisées. Un premier test avec un forçage historique fourni par IRSN a été produit (Figure 5.9) avec une hypothèse de transfert très simple, c'est-à dire proportionnel aux concentrations de césium $[Cs \times PP]$. Le projet se focalisera sur le développement d'une représentation plus explicite des mécanismes, notamment le taux d'ingestion, le rendement d'absorption et le taux d'excrétion pour chaque groupe, et sur leur paramétrisation.

Avec une approche très similaire, il est possible d'étendre ce type de développements aux métaux lourds. Ce type de pollution dans le milieu marin peut par exemple intervenir dans le cas d'exploitations minières dont certains rejets contaminés se font en mer par une buse partant au large. Il existe (malheureusement) d'autres cas de figures, notamment par le biais des grands fleuves.

Bibliographie

- Abecassis, M., Senina, I., Lehodey, P., Gaspar, P., Parker, D., Balazs, G., and Polovina, J. (2013). A model of loggerhead sea turtle (*Caretta caretta*) habitat and movement in the oceanic north pacific. *PLoS One*, 8(9) :e73274.
- Alvain, S., Moulin, C., Dandonneau, Y., and Bréon, F. (2005). Remote sensing of phytoplankton groups in case 1 waters from global seawifs imagery. *Deep Sea Research Part I : Oceanographic Research Papers*, 52(11) :1989 – 2004.
- Arai, M. N. (2005). Predation on pelagic coelenterates : a review. *Journal of the Marine Biological Association of the United Kingdom*, 85(03) :523–536.
- Brotz, L., Cheung, W. W., Kleisner, K., Pakhomov, E., and Pauly, D. (2012). Increasing jellyfish populations : trends in large marine ecosystems. *Hydrobiologia*, 690(1) :3–20.
- Chambault, P., de Thoisy, B., Heerah, K., Conchon, A., Barrioz, S., Dos Reis, V., Berzins, R., Kelle, L., Picard, B., Roquet, F., et al. (2016). The influence of oceanographic features on the foraging behavior of the olive ridley sea turtle *lepidochelys olivacea* along the guiana coast. *Progress in Oceanography*, 142 :58–71.
- Fields, P. A. and Somero, G. N. (1998). Hot spots in cold adaptation : localized increases in conformational flexibility in lactate dehydrogenase a4 orthologs of antarctic notothenioid fishes. *Proceedings of the National Academy of Sciences*, 95(19) :11476–11481.
- Lambert, C., Mannocci, L., Lehodey, P., and Ridoux, V. (2014). Predicting cetacean habitats from their energetic needs and the distribution of their prey in two contrasted tropical regions. *PloS one*, 9(8) :e105958.
- Lehodey, P., Conchon, A., Senina, I., Domokos, R., Calmettes, B., Jouanno, J., Hernandez, O., and Kloser, R. (2015). Optimization of a micronekton model with acoustic data. *Ices Journal of Marine Science*, 72(5) :1399–1412.
- Lehodey, P., Murtugudde, R., and Senina, I. (2010). Bridging the gap from ocean models to population dynamics of large marine predators : A model of mid-trophic functional groups. *Progress in Oceanography*, 54(1-2) :69 – 84. Special Issue : Parameterisation of Trophic Interactions in Ecosystem Modelling.
- Low, P. S., Bada, J. L., and Somero, G. N. (1973). Temperature adaptation of enzymes : roles of the free energy, the enthalpy, and the entropy of activation. *Proceedings of the National Academy of Sciences*, 70(2) :430–432.

- Parent, L., Ferry, N., Barnier, B., Garric, G., Bricaud, C., Testut, C.-E., Le Galloudec, O., Lellouche, J., Greiner, E., Drevillon, M., et al. (2013). Global eddy-permitting ocean reanalyses and simulations of the period 1992 to present. Technical report, Technical report, Mercator Ocean.
- Purcell, J. E. (2012). Jellyfish and ctenophore blooms coincide with human proliferations and environmental perturbations. *Annual Review of Marine Science*, 4 :209–235.
- Purcell, J. E., Uye, S.-i., and Lo, W.-T. (2007). Anthropogenic causes of jellyfish blooms and their direct consequences for humans : a review. *MARINE ECOLOGY-PROGRESS SERIES*-, 350 :153.
- Quiñones, J., Monroy, A., Acha, E. M., and Mianzan, H. (2013). Jellyfish bycatch diminishes profit in an anchovy fishery off peru. *Fisheries Research*, 139 :47–50.
- Roberts, J. J., Best, B. D., Mannocci, L., Fujioka, E., Halpin, P. N., Palka, D. L., Garrison, L. P., Mullin, K. D., Cole, T. V., Khan, C. B., et al. (2016). Habitat-based cetacean density models for the us atlantic and gulf of mexico. *Scientific reports*, 6.
- Sutherland, K. R., Madin, L. P., and Stocker, R. (2010). Filtration of submicrometer particles by pelagic tunicates. *Proceedings of the National Academy of Sciences*, 107(34) :15129–15134.

Chapitre 6 Conclusion

Comme souvent j'imagine, les résultats et les travaux d'une thèse s'écartent souvent du sujet et des objectifs initiaux du fait de nouvelles pistes intéressantes qui se dévoilent ou de difficultés techniques inattendues. Ma thèse a fait suite à un stage dans lequel j'avais commencé à explorer les possibilités du modèle SEAPODYM pour le micronecton et j'avais également entamé une application de ce modèle au zooplancton. Je pensais en démarrant la thèse pouvoir rapidement améliorer les premiers résultats à la fois sur le zooplancton et sur l'optimisation des paramètres du modèle. Puis dans le même temps, obtenir une simulation globale réaliste à partir de laquelle j'aurai développé et analysé des simulations régionales plus détaillées. En particulier, mes co-directeurs de thèse étaient très intéressés par la région Indonésienne et la zone Australe des Iles Kerguelen en lien avec leurs projets de recherche.

Mais j'ai appris que la recherche scientifique demandait beaucoup de patience et d'obstination ! Au fur et à mesure que ma thèse avançait, des questions imprévues se sont posées. L'application au zooplancton a démontré la sensibilité du modèle au forçage mais aussi suggère fortement la révision de la relation fonctionnelle utilisée pour modéliser l'effet de la température sur les temps de développement (Chapitre 2). La méthodologie de l'assimilation de données pour le modèle SEAPODYM-MTL a été développée autour d'un cas simplifié avec très peu de données à assimiler (Chapitre 3). Cette méthode a été étendue à l'ensemble d'un jeu de données rendu disponible au cours de la thèse. Mais, les résultats obtenus se sont avérés peu satisfaisants, et rapidement la question concernant la définition des couches verticales bio-acoustiques est apparue comme critique car les résultats du modèle y étaient très sensibles. Il n'y avait pas d'autre choix que de revenir sur cette problématique. L'utilisation des données acoustiques s'est imposée pour cette tâche et le modèle d'observation des AMLD par acoustique a vu le jour (Chapitre 4). Une nouvelle définition basée sur des variables synoptiques et réalistes dérivées de données satellites a été proposée qui améliore nettement la définition précédente. Cette définition devra certainement être revue à l'avenir avec la mise à disposition de plus en plus de données acoustiques. Mais les outils développés ici l'ont été dans l'esprit de pouvoir traiter rapidement ces prochains flux massifs de données. Tout est prêt maintenant pour relancer les expériences d'optimisation de paramètres, à la fois pour le zooplancton et le micronecton, et avec des jeux de données qui vont s'enrichir grâce à de nouveaux projets de recherche qui démarrent comme par exemple le projet européen MESOPP en coopération avec les collègues Australiens et Européens. Il reste néanmoins à prendre en compte le traitement du bruit dans le signal acoustique, ce qui pourrait bien nécessiter une autre thèse !

Néanmoins, des applications de ce modèle sont déjà développées, notamment avec le modèle opérationnel INDES0 sur la région indonésienne, mais aussi des simulations des dynamiques de populations des stocks de thons et de leurs pêcheries à l'échelle globale. De nouvelles perspectives se font jours également avec potentiellement des débouchés aussi variés que le suivi de polluants radioactifs, la modélisation de différentes chaînes trophiques pouvant aboutir aux blooms d'organismes gélatineux, et la prédiction d'habitats de baleines, otaries, manchots, tortues...

L'aventure ne fait donc que commencer et j'espère bien pouvoir contribuer à ses prochains développements !

Annexes

Annexe A : Formations

Durant ma thèse, j'ai eu la chance de participer à une école d'été à l'école de Physique des Houches intitulée 'Analyse de sensibilité, propagation d'incertitudes et exploration numérique de modèles en sciences de l'environnement' (<http://aspen.forge.imag.fr/>). Les organisateurs étaient Clémentine Prieur (UJF Grenoble), Hervé Monod (MIA - INRA, Centre de Jouy en Josas), Robert Faivre (MIA - INRA, Centre de Toulouse) et Laurence Viry (UJF Grenoble).

J'ai également validé des formations en ligne (MOOC, Massive Open Online Course). La première est intitulée 'R Programming' et est dispensé par Johns Hopkins University (<http://www.coursera.org/learn/r-programming/>). La seconde auquel j'ai participé est intitulée 'Writing in the Sciences' et est dispensé par Stanford (<https://www.coursera.org/course/sciwrite>).

Dans un autre ordre d'idée, j'ai également participé à une campagne en mer qui fut une source de formation non négligeable sur l'acquisition des données acoustiques. Cette campagne fut éclairante notamment sur la qualité des données, leurs répartition spatiale et la nécessité de filtrage de ces dernières. Cette campagne s'est déroulée sur le Marion Dufresne en Janvier-Février 2014 dans le cadre du projet MyctO-3D-MAP.

Annexe B : Communication scientifiques

J'ai pu présenté mes travaux à de nombreuses réunions des projets dans lesquels mon équipe était impliquée (MyctO-3D-MAP, Osmosis, INDESO, Atlantos). Endehors de ces occasions, j'ai participé aux conférences nationales et internationales suivantes :

2013 :

Primary production in the ocean : from the synoptic to the global scale (The 45th Liège Colloquium, Belgique)

Journée Eléphant de mer, (Chizé, France)

2014 :

IMBER Open Science Conference (Bergen, Norvège)

Journée Eléphant de mer, (Chizé, France)

Colloque National sur l'Assimilation de données (Toulouse, France)

2015 :

Marine Environmental Monitoring, Modelling and Prediction (The 47th Liège Colloquium, Belgique)

Marine Ecosystem Acoustics (Some Acoustics)– observing the ocean interior in support of integrated management (ICES Symposium, Nantes, France)

J'ai également collaboré avec d'autres personnes en dehors du cadre des projets cité si dessus. Ces collaborations ont notamment abouties à la publication suivante :

Chambault P, De Thoisy B, Heerah K, **Conchon A**, Barrioz S, Dos Reis V, Berzins R, Kelle L, Picard B, Roquet F, Le Maho Y. (2016). The influence of oceanographic features on the foraging behavior of the olive ridley sea turtle *lepidochelys olivacea* along the guiana coast. *Progress in Oceanography*, 142 :58–71.