Evaluation de l'état des populations de raie bouclée

Florianne Marandel

UNIVERSITE SCIENCES BRETAGNE DE LA MER LOIRE ET DU LITTORAL

AGRO CAMPUS OUEST

THESE DE DOCTORAT DE

L'AGROCAMPUS OUEST
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 598
Sciences de la Mer et du littoral
Spécialité : *Halieutique*

Par

Florianne MARANDEL

Evaluation de l'état des populations de raie bouclée

Thèse présentée et soutenue à Nantes, le 19 septembre 2018
Unité de recherche : Unité Ecologie et Modèle pour l'Halieutique, IFREMER
Thèse N° : 2018-16 H-105

Rapporteurs avant soutenance :

Jennifer Ovenden - Professeur, Université de Brisbane
Nigel, Gilles Yoccoz - Professeur, Université Arctique de Norvège

Composition du Jury :

Nigel, Gilles, Yoccoz - Professeur, Université Arctique de Norvège
Nicolas Bez - Directeur de recherche, IRD
Olivier Lepais - Chercheur, INRA
Marie Savina-Rolland - Chercheuse, IFREMER

Président
Yannick Outreman - Professeur, Agrocampus-Ouest

Directeur de thèse
Verena M. Trenkel - Chercheuse, IFREMER

Invité
Pascal Lorance - Chercheur, IFREMER

THÈSE DE DOCTORAT

ÉVALUATION DE L'ÉTAT DES POPULATIONS DE RAIE BOUCLÉE

2015/2018

PRÉPARÉE PAR

FLORIANNE MARANDEL

SOUS LA DIRECTION DE

VERENA M. TRENKEL & PASCAL LORANCE

Ifremer

Remerciements

En tout premier lieu, je remercie chaleureusement mes deux directeurs de thèse de m'avoir fait confiance pendant ces 3,5 ans. Tout d'abord merci **Verena** pour ton implication, ton intérêt constant pour mon travail et ta droiture ! J'ai beaucoup appris avec toi tant sur le plan humain et que sur le travail. Travailler avec toi a été très exigeant mais ce fut un plaisir que je réitérerai bien ! Je pense pouvoir m'enorgueillir d'avoir amener un peu de violet, de rose et de paillettes dans tes figures et présentations ! **Pascal**, merci avant tout pour ta curiosité, ta rigueur et ta disponibilité. Tu es un peu la force tranquille du trio "raie bouclée" d'EMH ! Merci d'avoir su souligner les résultats intéressants mais aussi ceux qui l'étaient un peu moins... ! Merci à vous deux de m'avoir inclue dans bien d'autres choses que juste "la thèse" et de m'avoir permis de collaborer avec et sans vous sur bien d'autres projets.

Merci aux membres de mon jury de thèse (**Jennifer Ovenden, Nigel Yoccoz, Olivier Lepais, Yanick Outreman, Marie Savina-Rolland** et **Nicolas Bez**) d'avoir accepté d'évaluer mon travail ! En amont de ce travail, merci à mon comité de thèse (**David Causeur, Gregory Charrier, Jean-Baptiste Lamy, Marie Laure Pilet Nayel** et **Frédérique Viard**) pour leur aide précieuse et toujours très constructive !

Un remerciement spécial à **Jean-Baptiste Lamy** qui m'a vraiment beaucoup soutenue (sans peut être s'en rendre compte ?) dans la découverte de la génétique et tous les aléas humains et méthodologiques que ça a entrainé. Merci de ta patience, de tes nombreux retours et de ton enthousiasme ! Aussi, merci beaucoup **Gregory Charrier** de ton aide et de tes précisions au début de ma thèse.

For all the effective population size part, I am particularly indebted to **Robin Waples**. Thank you so much for welcoming me in Seattle ! This PhD thesis could not have been accomplished without your help. Moreover, thank you for your kindness, your patience but also for my first Thanksgiving diner !

Thank you **Mindy Rowse** for sharing your office with me. Thanks to you, I discover the Orcas Island and it was really a great trip !

The members of the **WGEF** have greatly contributed to my understanding of elasmobranchs ecology but also of fisheries. Thank you very much for all the nice time together. A special thanks for **Sophie McCully-Phillips**, **Graham Johnston** and the last but not the least, **Jim Ellis**. Thank you Jim for all the beers, the nice discussions but also all the nice memories in the "train" (and the flying popcorn !).

Merci à toute l'équipe du projet GenoPoptaille pour son aide, son implication mais aussi (surtout) pour tout le travail accompli ! Merci **Sabrina Le Cam**, **Adeline Bidault**, **Florence Cornette**, **Eric Stephen**, **Gérard Biais**, **Manuella Rabillier** et **Jean Laroche**.

Merci à l'**IFREMER**, à l'**Agence Nationale de la Recherche** et à la **Fondation Total** d'avoir financé ce travail ! Merci à **IFREMER** de m'avoir permis de partir deux mois aux Etats-Unis.

Un gros merci à tous ces collaborateurs d'un jour, d'une semaine, d'un mois qui m'ont permis de sortir un peu de ma thèse tout en faisant de la science marine ! Merci à **Stéphanie Mahévas** de m'avoir invitée à prendre part aux journées prospectives de l'IUML. Merci aussi pour tes encouragements cette dernière année. Un gros merci aux atlantes : **Sophie Pardo, Aurélien Babarit, Benoit Schoefs, Sandrine Jamet, Pierre Alexandre Mahieu, Justine Dumay, Hélène Callewaert, Jonas L'Haridon** et **Régis Baron**. On n'aura qu'à se retrouver tous ensemble dans 10 ans à l'Atlantide autour d'une bonne bière ? Merci de vos encouragements et d'avoir autant cru en moi ! Votre confiance a été grandement appréciée en cette fin de doctorat ! ! ! Merci à **Sophie Pilven** de m'avoir inclue dans de nombreux projets de vulgarisation scientifique, j'ai vraiment adoré tous ces moments !

EMH, EMH,... Merci à **toute l'équipe d'EMH** de m'avoir acceuillie pendant ces années ! J'ai beaucoup apprécié les pauses cafés, les gâteaux, les activités sportives, etc... Bien sûr, à mes yeux, il y a quelques VIP.... Alors merci **Olivier** pour tout le travail de programmation accompli ! Grâce à toi, j'ai pu optimiser le travail de simulation et gagner beaucoup de temps ! ! Enfin merci pour ta bonne humeur constante ! **Vincent**, merci beaucoup pour ton humanité, tes conseils avisés et ta capacité à relativiser ! J'ai beaucoup appris humainement grâce à toi. Je remercie aussi **Anne-Sophie** pour sa bonne humeur et sa compréhension. C'était toujours très agréable de parler et travailler avec toi. Merci aux autres thésards, post-doc et stagiaires pour leur aide, conseils, soutien, blagounettes (**Louise, Nans, Erwan, Fabien**, etc) !

Un franc merci à **Bérengère** et **Polo** pour leur accueil toujours très chaleureux à Brest ! J'ai adoré mes missions à Brest grâce à vous et je suis au taquet pour un Mysterium ou un Dice Forge quand vous voulez !

Un merci spécial à ces deux lascards qui étaient de mes très bons amis avant cette thèse, qui sont devenus mes collègues pendant et qui sont toujours de très bons amis après ! Merci **Arthur** et **Audric** pour votre amitié, votre humour, vos goûts décalés, les soirées jeux de rôles, jeux de société, jeux videos, etc !

Je remercie aussi mes copains qui n'ont rien à voir avec l'aspect scientifique de cette thèse ni dans mon quotidien ifremérien.... **Mandine**, avec toi, soyons sobre, soyons clair, soyons précis : merci pour ces 14 ans d'amitié, d'honneteté, merci de me prouver chaque jour qu'on peut être loin des yeux mais proche du coeur. Merci **Claire** pour ton amitié fidèle, nos coups de téléphone réguliers et ta bonne humeur ! Merci à cette jolie rencontre qu'est **Camille**, c'est top de pouvoir discuter et rigoler comme on le fait. Merci pour les 1001 messages quotidiens ! Merci à mon **Dadoo**, la **Picardo Team**, **Potchou**, etc...

(Au passage, je souhaite aussi remercier une actrice essentielle de cette thèse : la bande son ! Merci donc à Wicked, Rammstein, Aurora, Brigitte, Yoko Kanno, Therion, etc.... Sans vous, j'aurais mangé mon chapeau !)

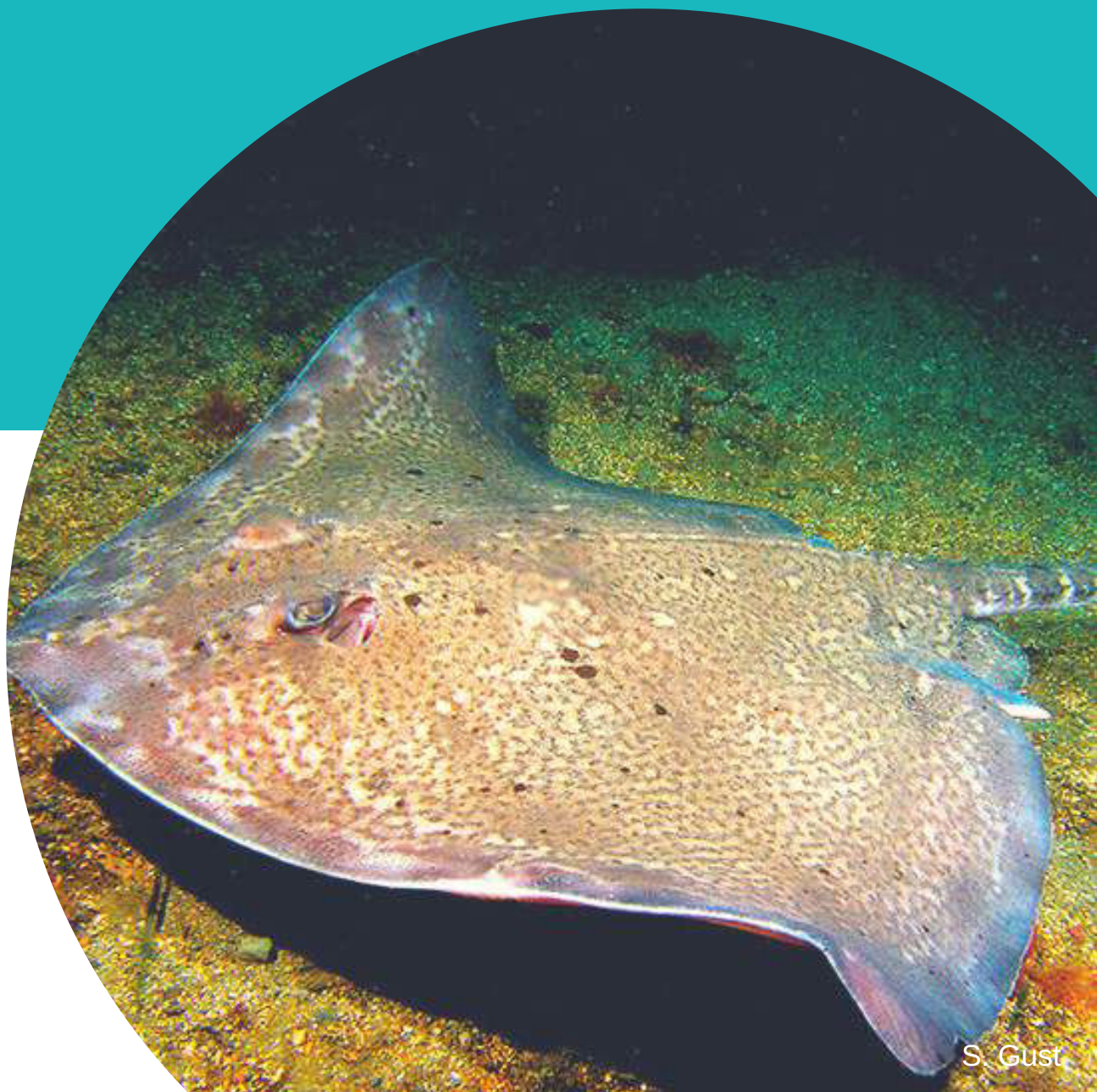
Merci à **Vivie** et toute sa "petite" famille pour les franches rigolades, les soirées arrosées et la bonne humeur. J'aurais carrément pu tomber pire comme belle-famille ! alors merci beaucoup de m'accepter comme je suis.

Merci à toute ma famille mais surtout merci **Papa** et **Maman** pour m'avoir permis d'écrire ceci aujoud'hui. Merci de m'avoir soutenue moralement et financièrement toutes ces années, c'est grâce à vous que j'en suis là et je ne vous en serais jamais assez reconnaissante. Cette thèse, c'est la fin de ma vie d'étudiante, c'est la fin de 9 années parfois difficiles alors merci pour tout ! Merci Papa pour tous les conseils avisés et merci Maman pour tes bras réconfortants (et tes bons petits plats !).

Enfin, Merci à **Ambroise**, pour ta présence et tout ce que tu m'apportes au quotidien. Merci d'avoir essayer de comprendre mon sujet et de t'être intéressé aux raies. Sans toi cette thèse n'aurait clairement pas été possible. Plus que tout. Plus qu'hier. Moins que demain.

CHAPITRE 1

Introduction générale



S. Gust

1. Contexte scientifique et conceptuel

L'exploitation des ressources marines remonte à plusieurs milliers d'années (Caddy and Cochrane, 2001). Ces ressources ont longtemps été considérées comme inépuisables bien que les activités de pêche aient très vite eu des effets significatifs sur les **populations** * et peuplements marins. Dans les années 1940, l'avénement de l'industrie a notamment marqué un tournant pour la pêche avec l'arrivée de nouvelles techniques de pêche mais aussi l'essor d'une économie de marché à la demande sans cesse croissante (Caddy and Cochrane, 2001). Rapidement des signes de surexploitation sont apparus dans de nombreuses **pêcheries** commerciales. Ainsi, en 1974, 10% des **stocks** mondiaux étaient considérés comme surexploités contre 31,4% en 2013 ; de même, en 1974, 40% étaient sous-exploités contre seulement 10% en 2013 (FAO, 2016a). Bien que moins discutés, des signes de surexploitation sont également apparus chez de nombreuses **espèces accessoires** comme les requins et les raies. Pour ces dernières, les captures commerciales déclarées en Atlantique Nord-Est ont été divisées par deux en 60 ans, avec une baisse constante depuis les années 2000 suites à des mesures de gestion, notamment l'introduction de quota de pêche (Fig. 1.1).

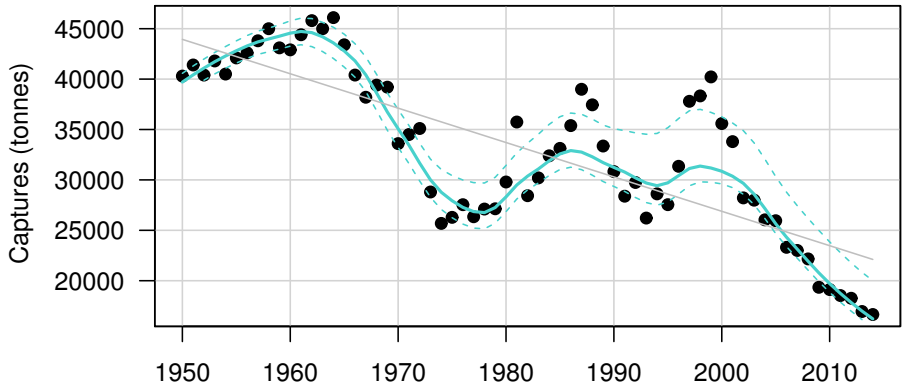


FIGURE 1.1 : En noir, tonnages pêchés en Atlantique Nord-Est de 1950 à 2014 pour l'ordre des Rajiformes. En gris, regression linéaire. En bleu, regression loess et intervalles de confiance à 95%. FAO (2016b)

La conservation des raies et requins (classe des élasmobranches) constitue un objectif majeur de la gestion des ressources marines (Dulvy et al., 2014; Davidson et al., 2016). Toutefois, l'estimation de l'**abondance** actuelle de ces espèces par les approches halieutiques usuelles est souvent impossible, en raison notamment d'un faible nombre d'individus observés. Il est donc primordial de disposer de techniques alternatives, éventuellement non-létales dans le cas d'espèces protégées, pour estimer leur abondance et améliorer la connaissance de leur biologie en lien avec leur gestion et conservation.

Pour améliorer la gestion des ces espèces, il est nécessaire d'améliorer notre connaissance de la structure des **stocks** et **populations**. En effet, une mauvaise identification des stocks peut entre autres conduire à une surexploitation voire à un épuisement des stocks les plus sensibles à la pression de pêche (Kerr et al., 2016). Génétiquement, ces phénomènes se traduisent généralement par une perte de **diversité génétique** et donc une réduction du potentiel adaptatif,

*Les mots en gras ont une entrée dans le glossaire en fin de thèse

pourtant essentiel à la survie à long terme d'une population soumise à des pressions environnementales changeantes (changement climatique, pollution, ...). Si la survie d'une population à court terme est essentiellement liée à des paramètres démographiques tels que les taux de survies par âge ; sa survie à long terme est liée au maintien de sa diversité génétique (Smith and Smith, 2001). Il est donc essentiel d'un point de vue gestion d'intégrer les deux aspects : démographie et génétique.

1.1 Les différentes unités de population : concepts et utilisation

1.1.1 La notion de stock

Le concept de stock est ancien et remonte à la fin du $XIX^{ème}$ siècle (Heinke, 1898). Il n'a pas cessé d'évoluer depuis mais reste néanmoins un concept controversé avec une multitude de définitions différentes usitée tant par les scientifiques que par les organes de gestion (Booke, 1981; Laurec and Le Guen, 1981; Ihssen et al., 1981; Carvalho and Hauser, 1994; Begg and Waldman, 1999; FAO, 2016a). L'ensemble de ces définitions présente des points communs : elles font référence à la fraction exploitée d'une espèce donnée dans une zone donnée. L'accent est ainsi mis sur l'aspect spatial, rendant la notion de stock facile à utiliser en gestion des pêches mais ignorant l'aspect biologique.

On préférera donc utiliser une définition plus détaillée où un stock est la fraction exploitée d'un groupe d'individus monospécifique se reproduisant entre eux au hasard dans un écosystème déterminé et présentant une certaine stabilité spatiale et temporelle (Laurec and Le Guen, 1981; Ihssen et al., 1981). Deux propriétés émergent de cette définition :

- un stock est consitué d'individus homogènes génétiquement,
- un stock est une entité relativement isolée (pas ou peu d'émigration ou d'immigration).

Ces propriétés impliquent que plusieurs stocks représentent des groupements d'individus relativement indépendants avec des flux migratoires faibles mais elles ne disent rien sur les différenciations démographique ou génétique. Au contraire, un stock génétique ou **dème** est défini comme une unité de reproduction isolée et génétiquement différenciée de ces voisins. Cette définition implique donc une limitation extrême des flux migratoires, une forte homogénéité génétique intra-dème et une forte hétérogénéité génétique inter-dèmes (Carvalho and Hauser, 1994).

1.1.2 La notion de population

De manière analogue, la notion de population est ancienne et ne fait pas l'unanimité entre les différentes disciplines qui manient ce concept (écologie et statistique notamment) mais aussi au sein d'une même discipline (Debouzie, 1999). En statistique, tout d'abord, une population peut être définie comme l'ensemble des individus étudiés (Upton and Cook, 2006). En écologie, une population peut se définir comme un groupe d'individus appartenant à la même espèce et occupant un même écosystème au même moment (Barbault, 1981; Ramade, 2003). C'est une entité écologique possédant ses caractéristiques propres comme la natalité, la mortalité, la distribution en classe d'âge, le **sex ratio**, etc... Une population au sens écologique possède ainsi des caractéristiques démographiques qui lui sont propres mais elle n'est pas strictement isolée des populations voisines. Ainsi en 1969, Levins introduit le terme de "métapopulation" pour définir une population constituée de plusieurs populations locales connectées par des individus se

dispersant (Hanski, 1991). En 1988, Pulliam applique le principe d'habitats "sources" et "puits" aux populations et insiste sur l'importance de ces sources et puits dans la dynamique d'une population donnée. Une population source est donc une population produisant plus d'individus que nécessaire à son maintien et une population puit, une population dépendante de l'immigration pour son maintien. Il apparait alors de ces concepts qu'étudier une population mais plus encore, délimiter une population afin d'en étudier l'abondance peut se réveler difficile.

1.2 Etudier l'abondance des populations marines

L'évaluation de l'abondance des populations marines constitue une démarche indispensable à une bonne gestion, notamment quantitative, des ressources halieutiques (Hilborn and Walters, 1992). De par la nature du milieu, mais aussi la grande taille des populations étudiées, il n'est pas possible d'effectuer des dénombrements absolus des populations marines. On réalise donc des estimations, les plus fidèles possibles, en échantillonnant seulement quelques individus.

De nombreuses techniques existent pour estimer l'abondance des populations marines à partir des échantillons prélevés, toutes souffrant de différents biais, imprécisions et hypothèses (Hilborn, 1992; Cadrin and Dickey-Collas, 2015; Maunder and Piner, 2015). Dans le cas des raies et requins, il est difficile d'évaluer l'abondance et la distribution des individus à cause des difficultés d'observation (Kuhnert et al., 2011). Des méthodes classiques comme la Capture-Marquage-Recapture sont ainsi difficilement applicables mais aussi coûteuses. L'utilisation de méthodes indirectes, par exemple inférant le nombre d'individus à partir de données de pêche (Hilborn and Walters, 1992) ou du génotype d'individus échantillonnés (Dudgeon and Ovenden, 2015; Bravington et al., 2016), peuvent ainsi être utilisées.

2. Cas d'étude : la raie bouclée (*Raja clavata*)

2.1 Biologie et écologie de la raie bouclée

La raie bouclée (*Raja clavata*, Fig. 1.2), Linné, 1758) appartient à la classe des élasmobranches (et plus précisément à la famille des Rajidae) qui regroupe des espèces aux traits d'histoire de vie particuliers chez les poissons (*e.g.* fertilisation interne, absence de phase larvaire, petite taille de portée chaque année, maturité sexuelle tardive, fort potentiel philopatrique, etc...) (Quéro and Vayne, 1997; Chevolot, 2006). C'est une espèce à cycle de vie relativement long avec un âge maximum estimé autour de 16 ans (Whittamore and McCarthy, 2005) et un âge de maturité sexuelle tardif autour de 6 ans (Whittamore and McCarthy, 2005; Serra-Pereira et al., 2011; Wiegand et al., 2011), correspondant à une taille de 67 cm pour les mâles et 74 cm pour les femelles. Sa taille maximale se situe autour de 105 cm (Le Quesne and Jennings, 2012).

FIGURE 1.2 : Raie bouclée (*Raja clavata*) péchée en Baie de Douarnenez en mars 2017

Le comportement reproducteur de la raie bouclée est assez méconnu. C'est une espèce ovipare présentant une fertilisation interne et un comportement polyandrique induisant des paternités multiples au sein d'une même portée (Chevolot et al., 2007), ainsi des frères et demi-frères sont observables au sein d'une population. Après l'accouplement, les femelles produisent des oeufs protégés par une enveloppe de kératine qui sont ensuite déposés sur le fond où ils s'accrochent au substrat à l'aide de quatre cornes (Chevolot, 2006; Holden, 1972). La figure 1.3 représente le cycle biologique de l'espèce. La fécondité des femelles se situe entre 48 et 140 oeufs par an (Ellis and Shackley, 1995; Serra-Pereira et al., 2011). Après 4 à 5 mois d'incubation, des raies complètement formées éclosent avec un sex ratio de 1 :1 (Delpiani, 2016), elles mesurent alors entre 10 et 14 cm. Il n'existe que peu d'information sur la mortalité des oeufs et des nouveaux-nés, néanmoins Ellis and Keable (2008) montrent qu'en milieu controlé (aquarium) la mortalité des oeufs avant éclosion est faible (environ 17%).

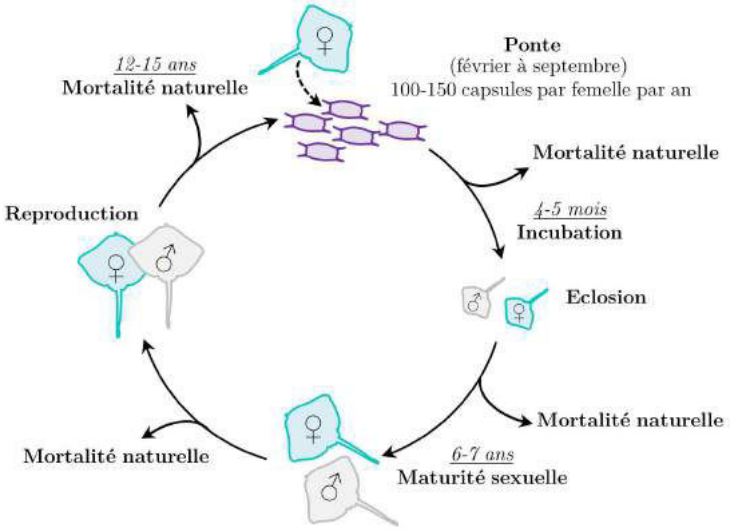


FIGURE 1.3 : Cycle biologique de la raie bouclée

De nombreuses études sur le comportement migratoire de la raie bouclée ont été conduites notamment par Capture-Marquage-Recapture dans la Mer du Nord (Walker, 1997; Hunter et al., 2005a,b). Considérées ensembles, ces études suggèrent que la raie bouclée pourrait présenter une forte différenciation des populations existantes. Les juvéniles tendent ainsi à demeurer dans des eaux de faibles profondeurs (10 à 30 m) pendant plusieurs années avec quelques mouvements saisonniers. Les adultes présentent des mouvements saisonniers entre des eaux de 30 m de profondeur et les hauts fonds, situés plus au large, où les accouplements ont lieu. Les individus recapturés ont été trouvés proches de leur point de libération, alimentant ainsi l'hypothèse d'un petit espace vital (autour de 75 à 130 km). Plus récemment, une modélisation statistique a confirmé la différence d'habitat préférenciel entre les juvéniles et les individus matures sexuellement (Maxwell et al., 2009; Martin et al., 2012).

2.2 Intérêt économique de l'espèce

Bien que peu de **débarquements** espèce par espèce soient disponibles, l'échantillonnage en criée indique que la raie bouclée est une des espèces de raies les plus fréquemment débarquées en Europe. La figure 1.4 présente les zones où la raie bouclée à été débarquée entre 2013 et 2015. Les captures commercielles débarquées ont été stantardisées afin de rendre compte de

l'hétérogénéïté présumée de l'abondance de l'espèce en Atlantique Nord-Est. Le pic de captures commerciales a ainsi lieu en Manche-Est et à l'ouest des côtes portugaises. La raie bouclée représente une des espèces de Rajidae les plus fréquentes au sein des débarquements Espagnols, Français, Irlandais et Anglais. Néanmoins, il est difficile d'évaluer les véritables débarquements de raie bouclée à cause du très fort degré de mauvaise identification de l'espèce, les raies étant très semblables dans leurs formes, couleurs et motifs (Bonfil, 1994; SharkTrust, 2009) et ayant la même valeur commerciale.

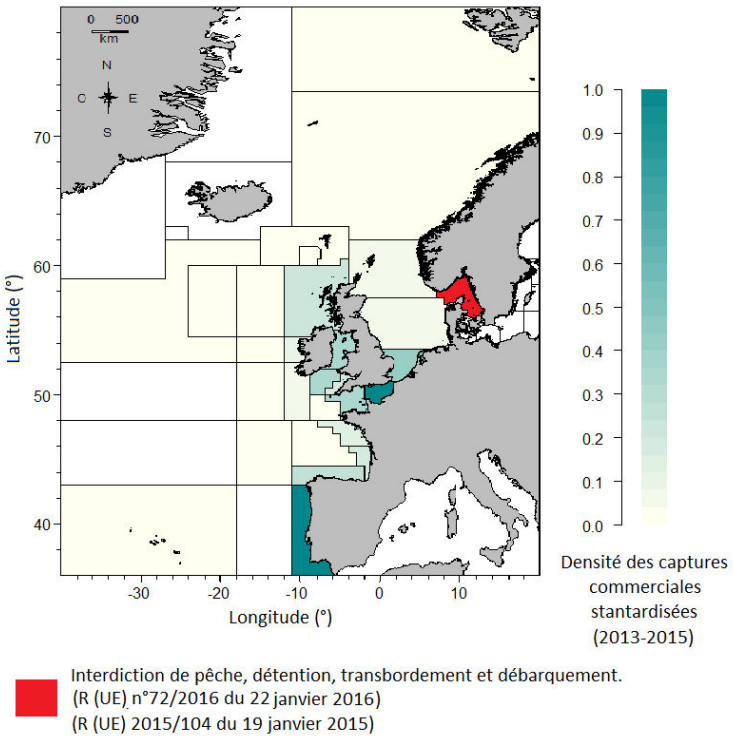


FIGURE 1.4 : Distribution des captures commerciales stantardisées entre zones (entre 2013 et 2015) et interdiction de pêche de la raie bouclée par division CIEM. (ICES, 2016b)

La figure 1.5 présente le nombre d'observations d'individus de raie bouclée entre 2010 et 2014 pour différents engins de pêche ciblant d'autres espèces. Ces données ont été récoltées par le programme scientifique d'observation OBSMER (Fauconnet et al., 2011; Dube et al., 2012; Cornou, A.S. et al., 2013, 2015a,b). Ce programme a pour but d'observer *in situ* les pécheries commerciales et l'ensemble de leurs captures en accordant une attention particulière à la **fraction rejetée**. On peut ainsi voir que la raie bouclée répresente une part significative des prises accessoires des pêcheries mixtes **démersales**. Il n'existe que peu de pécheries dirigées à la raie mais c'est par exemple le cas de quelques pécheries **palangrières** et au filet autour des Îles Britanniques, où la raie bouclée représente une des espèces cibles.

Actuellement il n'existe pas de gestion de la raie bouclée au niveau spécifique, seulement une limitation des captures par un quota plurispécifique et une interdiction de pêche à l'est des côtes norvégiennes (zone ICES 3a, en rouge sur la figure 1.4). Néanmoins, les études disponibles sur l'état des populations suggèrent une diminution de l'abondance des populations en l'Atlantique Nord-Est (à l'exception de la population située en Manche Est) (Dulvy et al., 2000, 2006; Figueiredo et al., 2007; ICES, 2016b).

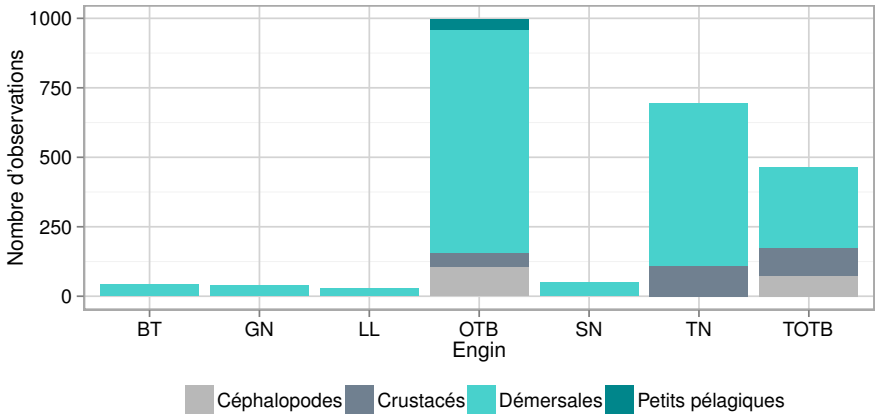


FIGURE 1.5 : Nombre d'observations de raies bouclée par engin et par espèce ciblée entre 2010 et 2014 en Atlantique Nord-Est. GN = filet maillant, TN = trémail, LL = palangres, OTB = chalut à panneaux, TOTB = chalut à panneaux jumeaux, SN = senne and BT = chalut à perche. (Fauconnet et al., 2011; Dube et al., 2012; Cornou, A.S. et al., 2013, 2015a,b)

2.3 Génétique des populations de raie bouclée

Assez peu d'études de génétique de populations d'élasmobranches sont disponibles et elles concernent presque toutes des espèces de requins (Lavery and Shaklee, 1989; Feldheim et al., 2001; Castro et al., 2007; Portnoy et al., 2009; Blower et al., 2012; Dudgeon and Ovenden, 2015; Andreotti et al., 2016). Néanmoins, depuis une quarantaine d'années, une attention particulière a été donnée à la raie bouclée d'un point de vue génétique et génomique (Nygren et al., 1971; Stingo, 1979; Stingo et al., 1980; Chevolot et al., 2005; Valsecchi et al., 2005; Chevolot et al., 2006, 2007; Pasolini et al., 2011). Ainsi dès 1971, Nygren et al. ont conduit des études cytologiques afin de déterminer le **caryogramme** de la raie bouclée et son nombre de chromosomes. Il a ainsi trouvé 49 paires de **chromosomes homologues** (Fig. 1.6). Le poids de son **génome** a été estimée entre 2,7 and 3,15 picogrammes (pg) (Gregory, 2016), ce qui équivaut à 2 640 et 3 080 Mb (1 Megabase = 1 000 000 de paires de bases). Des mesures sont disponibles pour des espèces voisines de Rajidae (Gregory, 2016) et semblent indiquer que la raie bouclée possède une taille de génome standard parmi les Rajidae. Pour comparaison, notamment avec d'autres taxons, d'autres tailles de génome sont référencées dans la table 1.1.

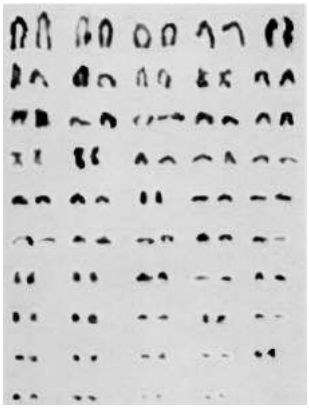


FIGURE 1.6 : Caryogramme de la raie bouclée avec $2n = 98$ chromosomes (Nygren et al., 1971)

TABLE 1.1 : Poids du génome et nombre de chromosomes de plusieurs espèces animales (Martinez-Lage et al., 1996; Gregory, 2016)

Nom commun	Nom latin	Poids du génôme (pg)	Nombre de chromosomes
Raie bouclée	*Raja clavata*	2.7 - 3.15	98
Raie douce	*Raja montagui*	3.45	
Raie chardon	*Leucoraja fullonica*	3.15	
Raie fleurie	*Leucoraja naevus*	3.15	
Requin peau bleue	*Prionace glauca*	4.30	86
Requin hâ	*Galeorhinus galeus*	8.65	
Saumon atlantique	*Salmo salar*	3.10	60
Maquereau commun	*Scomber scombrus*	0.97	
Moule commune	*Mytilus edulis*	1.60	28
Goéland argenté	*Larus argentatus*	1.50	66-70
Phasme	*Bacillus atticus atticus*	2.30	34
Homme	*Homo sapiens*	3.5	46
Girafe réticulée	*Giraffa camelopardalis*	2.85	30
Grand dauphin	*Tursiops truncatus*	3.03	44

En 2006, Chevolot réalisa sa thèse dont la problématique était d'évaluer la structure génétique des stocks de raie bouclée des îles Britanniques. Les principaux résultats concernent la structure des populations mais aussi l'estimation de la **taille de population efficace (N_e)** [†]. La taille de population efficace correspond au nombre d'individus se reproduisant effectivement au sein d'une population et donc transmettant leurs gènes à la génération suivante. Avec 5 **Microsatellites** et une méthode d'estimation de N_e dites "temporelle" (car basée sur le changement des fréquences alléliques dans le temps), la taille de population efficace de la raie bouclée dans le canal de Bristol et la mer d'Irlande est estimée à 283 individus (IC 95% : [145 ; 857]).

Le ratio de N_e avec la **taille de population absolue**, N est autour de $9e10^{-5}$ à $6e10^{-4}$ (Chevolot et al., 2008). Ce dernier résultat est surprenant compte tenu de la démographie de l'espèce. Les espèces présentant une **courbe de survie de type I** (*i.e.* espèce à durée de vie longue et à faible fécondité, à maturité tardive et avec une forte survie des adultes (Begon et al., 2006, l'équivalent des K-stratégistes) présentent souvent un ratio N_e/N approchant 1 (Portnoy et al., 2009). A contrario, les espèces avec une **courbe de survie de type III** (*i.e.*faible survie aux stade adulte, forte fécondité, maturation précoce, etc... (Begon et al., 2006, l'équivalent des r-stratégistes) présentent des ratios N_e/N très faibles : (10^{-3} - 10^{-5}) (Dudgeon and Ovenden, 2015). Les variations autour du ratio N_e/N des élasmobranches estimés dans la littérature mettent l'accent sur le besoin de mieux connaître cet indicateur avant d'utiliser la taille de population efficace dans la gestion et la conservation (Ovenden et al., 2016) . La variance dans le succès reproducteur est connue comme ayant une forte influence sur le ratio N_e/N (Frankham, 1995; Waples, 2016) mais ses effets sur le ratio N_e/N des élasmobranches sont toujours méconnus.

[†]Cette notion est détaillée dans le chapitre 5

Dans son étude de structure génétique des populations, Chevolot (2006) a mis en évidence une differentiation génétique forte et significative entre les populations continentales d'Atlantique Nord-Est(en bleu, Fig. 1.7), les populations Méditerranéennes (en vert, Fig. 1.7) et des Açores (en rose, Fig. 1.7). La différenciation entre l'Atlantique Nord-Est et la Méditerranée a également été mise en évidence par Pasolini et al. (2011). Des barrières ou goulot physiques telles que le détroit de Gilbraltar semblent ainsi participer à l'isolation génétique des populations mais au sein même des eaux Atlantiques ou Méditerranéennes, où il n'y a pas ou peu de barrières physiques, la divergence génétique des populations de raie bouclée est très réduite voir nulle.

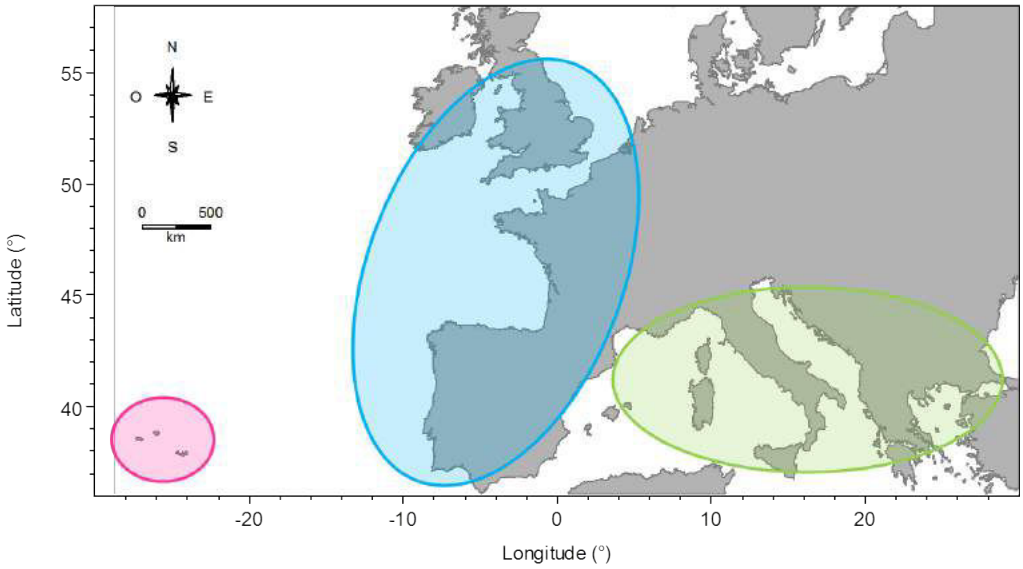


FIGURE 1.7 : Structure des populations mise en évidence par Chevolot (2006); Pasolini et al. (2011). En rose, l'unité génétique formée par les individus des Açores. En bleu, l'unité génétique formée par les individus des eaux Atlantiques. En vert, l'unité génétique formée par les individus des eaux Méditerranéennes.

Au sein de la métapopulation d'Atlantique Nord-Est, une différenciation génétique faible mais significative a également été mise en évidence entre les populations de Mer du Nord, de Manche Est et de Mer d'Irlande (Chevolot, 2006). De plus, les zones échantillonnées au printemps et en été étaient plus différenciées que celles échantillonnées en Automne et en Hiver. Ces résultats suggèrent fortement une dynamique spatiale et temporelle de la différenciation génétique de ces populations. Cette dynamique est surement liée à des attitudes comportementales précises de la raie bouclée. Une des conclusions principales de ce travail est la sous estimation de la capacité de dispersion de la raie bouclée. Elle serait alors de plusieurs centaines de kimomètres et non pas 75-130 km comme précédemment estimée par marquage (Walker, 1997; Hunter et al., 2005a,b).

Finalement, la thèse de Chevolot (2006) a montré quelques indices de paternités multiples pour la raie bouclée. Les avantages de la polyandrie sont multiples d'un point de vue évolutionnariste. Elle permet ainsi d'augmenter la diversité génétique des portées, de diminuer le risque d'incompatibilité génétique entre les partenaires et d'améliorer le fitness des femelles (Chevolot et al., 2007).

3. Le projet GenoPopTaille

Pendant ce travail de thèse, la raie bouclée a fait l'objet du projet GenoPopTaille (Fig. 1.8) financé par l'ANR (01/01/2015-30/09/2018). Ce projet regroupait plusieurs partenaires scientifiques : l'IFREMER, le LEMAR et l'APECS et se concentrait sur l'utilisation des méthodes génétiques pour l'évaluation des ressources halieutiques. Il se proposait d'estimer l'effectif absolu d'une population de géniteurs par l'identification génétique des paires parent-descendants. Pour cela, il fut nécessaire d'étudier préalablement la structure des populations afin de connaître les limites géographiques d'une population particulière. La méthode a ainsi été appliquée à la raie bouclée, dans le Golfe de Gascogne. Les paires parent-descendant ont été identifiées par génotypage d'un grand nombre de SNPs, **Single-Nucleotide Polymorphism**, sur un grand échantillon (environ 5 000) d'adultes et de juvéniles. Les défis relevés par le projet sont d'ordre pratique (échantillonnage) et statistique (modèles de dynamique population) avec une nouvelle utilisation des outils génétiques. Enfin, conjointement à cette thèse, le projet incluait une analyse de coût-bénéfice pour évaluer l'applicabilité de la méthode à d'autres populations de poissons, notamment des espèces en déclin ou menacées.

FIGURE 1.8 : Logo du projet GenoPopTaille. Création : F. Marandel and V.M. Trenkel

4. Problématiques et objectifs de la thèse

La problématique principale de cette thèse est partie du constat que l'évaluation de l'état mais aussi la gestion des populations des élasmobranches était une tâche difficile à cause du faible nombre d'observations de ces espèces. En conséquence, peu d'évaluations quantitatives ont été réalisées et l'évaluation de leurs populations repose principalement sur des savoirs experts.

Cette thèse s'articule donc autour de plusieurs questions :

- Peut-on évaluer l'état des populations d'élasmobranches à l'aide des méthodes et données disponibles actuellement ?
  - L'évaluation de l'état des populations par méthodes démographiques est-elle réellement impossible ou insuffisante pour ces espèces ?
  - L'évaluation de l'état des populations par méthodes génétiques est-elle une alternative pertinente ?
- Quel est l'état des populations de raie bouclée en Atlantique Nord-Est ?
- Quelle gestion possible pour la raie bouclée ?

Idéalement, le processus d'évaluation d'un stock halieutique est un processus chronologique nécessessitant tout d'abord une bonne connaissance de ce stock et de ses limites géographiques. Une fois ces limites connues, l'évaluation de son abondance peut alors être effectuée. Cette thèse reprend la chronologie d'une évaluation de stock en prenant la raie bouclée comme exemple.

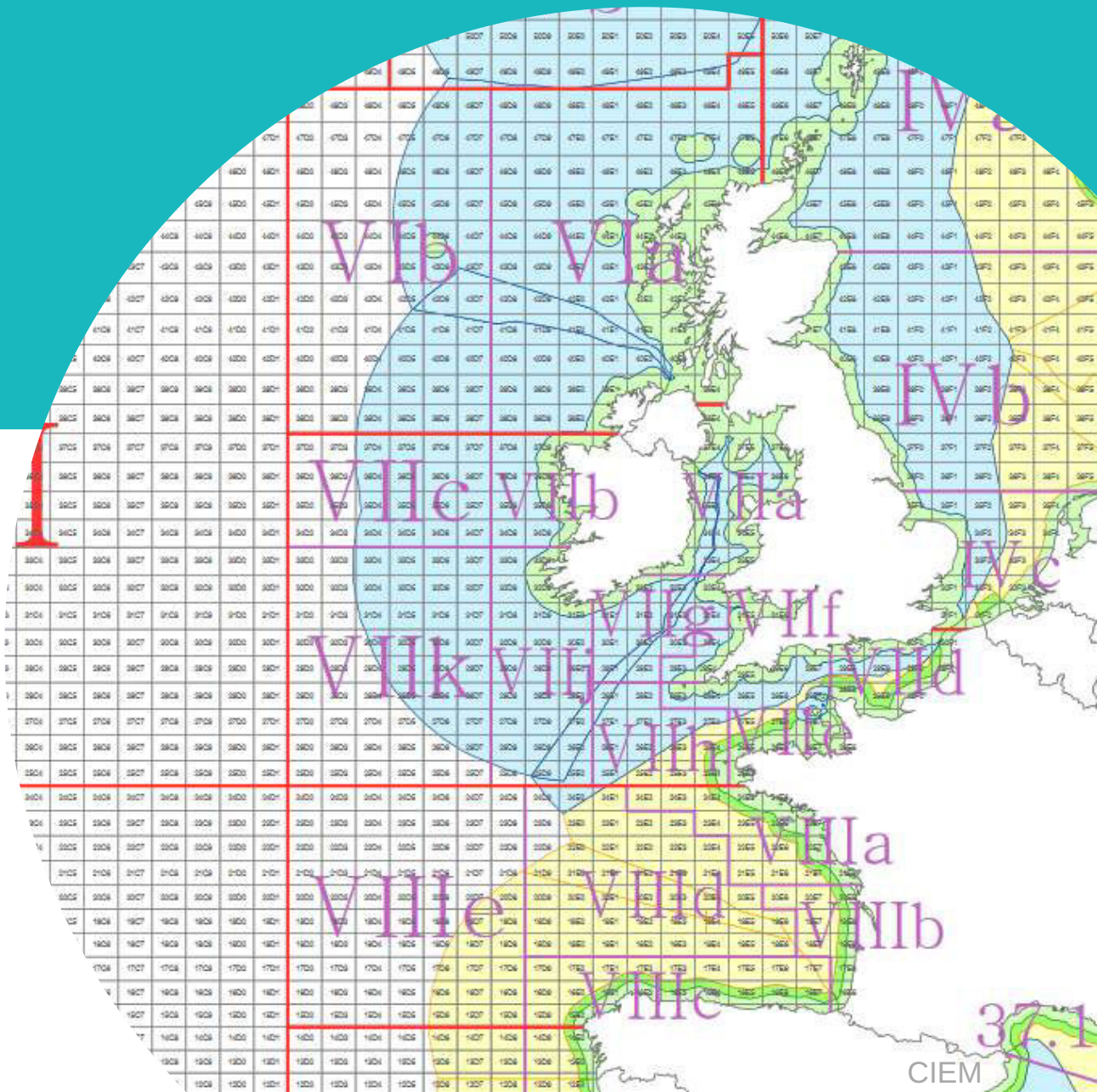
Ainsi la première étape de cette thèse se concentre sur la délimition d'unités de populations pour la gestion des raies en utilisant la raie bouclée comme exemple. Pour ce faire, le chapitre 2 s'intéresse à l'évaluation des connectivités génétique et démographique afin de délimiter des unités de gestion pertinentes. Cette partie repose sur des résultats issus de simulations de populationss de raie bouclée selon plusieurs scenarii de dispersion inter-populations.

La seconde étape de cette thèse s'articule autour de l'évaluation de l'abondance des espèces de raies en fonction des méthodes et données disponibles. Pour étudier cette problématique, des voies "classiques" d'estimation d'abondance (reposant sur des captures et indices d'abondances) sont éudiées dans les chapitres 3 et 4. En comparaison, des voies "alternatives" (reposant sur la génétique) sont investiguées dans les chapitres, 5 et 6. Dans ces parties, des données empiriques mais aussi des données simulées sont utilisées afin d'étudier les forces et faiblesses des différentes méthodes étudiées.

Les différentes parties de cette thèse s'appuient sur une comparaison entre les outils démographiques et génétiques disponibles. La partie génétique a demandé une bibliographie conséquente pour permettre une prise en main optimale des concepts et outils. Elle a servi de support afin d'orienter le sujet de thèse à ses débuts et faciliter le choix des méthodes. Cette revue a fait l'objet d'un rapport en anglais (Annexe A), il a été rédigé dans une optique didactique en début de thèse et ne comprends pas les résultats de ce travail.

CHAPITRE 2

Délimiter des unités de population : étude de la connectivité des populations de raie bouclée



Pour une gestion efficace des ressources et des écosystèmes marins, il est nécessaire de définir des unités de gestions appropriées. Ce premier chapitre étudie par simulations la délimitation des populations de raies par le biais des connectivités génétique et démographique. La connectivité génétique est calculée à l'aide d'un indice de différenciation génétique, nommé F_{ST}. Cette indice a été developpé au départ pour des **populations idéales** au sens d'Hardy Weinberg, c'est à dire présentant, entre autres, des traits d'histoires de vie extrêmement simplifiés. La première partie de ce chapitre compare donc les attendus théoriques et les valeurs après estimations des F_{ST} calculés afin d'étudier l'influence de la démographie des raies sur la différenciation génétique. La seconde partie s'interesse à la comparaison des connectivités génétique et démographique à proprement parler pour des populations de raies. Cette partie a fait l'objet d'une publication dans le Canadian Journal of Fisheries and Aquatic Sciences

1. Etude de la différenciation génétique des Rajidae

Les études de génétique des populations aident à définir des unités de population afin d'améliorer leur gestion et diminuer la perte de diversité génétique d'une espèce (Waples and Do, 2008; Spies et al., 2015). Ces unités sont habituellement identifiées par le degrée de connexion entre plusieurs populations putatives étudiées. En d'autres termes, par le degrée d'échange de matériel génétique entre deux (ou plusieurs) populations. Les composants génétiques de la structure des populations dépendent de nombreux facteurs comme les traits d'histoire de vie, la taille de population mais aussi le flux de gènes, la dérive génétique, la sélection et la mutation (Smith and Smith, 2001). Classiquement, les études génétiques ne considèrent que des marqueurs neutres pour inférer la structure des populations, ainsi seuls la dérive génétique et le flux de gènes sont considérés. Le premier tend à augmenter la différenciation globale entre plusieurs populations et le second tend à les homogénéiser (Beaumont et al., 2010).

A cause de leur phase larvaire, les poissons marins sont considérés comme ayant un fort potentiel de dispersion ainsi le flux de gènes est supposé élevé et la dérive génétique faible au sein de ces taxons, menant à une faible différenciation génétique des populations (Ward et al., 1994). Néanmoins,tous les poissons marins ne présentent pas cette phase larvaire et cette forte capacité de dispersion ne s'applique pas aux Rajidae ("skates and rays" dans la section 2.). Ces espèces présentent certes une grande variété de stratégies reproductives mais toutes basées sur la production de peu de petits déjà bien formés à la naissance (Quéro and Vayne, 2005). Cette biologie particulière tendrait ainsi à augmenter la differenciation génétique des populations. Par ailleurs, la grande taille des populations ainsi que la longévité des Rajidae tendraient à diminuer cette différenciation et à appliquer un effet tampon quant à son accroissement temporel.

Dans cette partie, l'effet de la biologie des Rajidae sur la différenciation génétique des populations est brièvement étudiée en comparaison avec la différenciation génétique de **populations idéales** au sens d'Hardy-Weinberg (Hamilton, 2009).

1.1 Métapopulation idéale *vs.* métapopulation Rajidae

Dans ce travail, la différenciation génétique de populations idéales et "Rajidae" simulées est comparée avec la valeur attendue par le modèle théorique en île associé (Hamilton, 2009).

Dans un premier temps, une métapopulation "idéale" ainsi qu'une métapopulation "Rajidae", constituées chacune de deux populations sont simulées. L'outil de simulation utilisé, le package R MetaPopGen (Andrello and Manel, 2015), a permis de simuler un **locus** biallélique

pendant 250 ans avec 100 replicas et huit différents taux de dispersion (0,2 ; 0,1 ; 0,05 ; 0,02 ; 0,01 ; 0,005 ; 0,002 ; 0,001). Il permet ainsi la simulation génétique d'une population selon des caractéristiques démographiques précises.

La démographie simulée de la métapopulation "Rajidae" reprend les principales caractéristiques démographiques décrites en introduction 2.1. Les populations sont à **générations chevauchantes** avec un âge maximum de 18 ans et un âge de maturité sexuelle à 6 ans. La fécondité des femelles est fixée à 150 oeufs, une valeur multiplée par le taux de survie de ces oeufs (0,3 valeur choisie afin d'avoir une démographie stable dans le modèle génétique du package R MetaPopGen). La fécondité des mâles est fixée à 100 000 spermatozoïdes, soit plus de 2 000 fois celle des femelles afin de correspondre à un tirage aléatoire des gamètes mâles. Chaque âge présente un taux de survie différent avec une survie augmentant progressivement jusqu'à 12 ans puis commençant à décroitre (Fig. 2.1). Le taux de mutation, fixé à 1e-06, est considéré négligeable. Le sex ratio est fixé à 1 :1 (Ellis and Shackley, 1995; Delpiani, 2016).

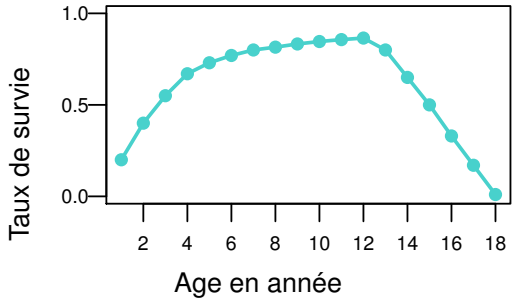


FIGURE 2.1 : Taux de survie par âge utilisés pour la métapopulation "Rajidae" (*Raja clavata*)

La valeur de différenciation génétique attendue (F_{ST} *attendu*) est calculée à l'aide d'un modèle en îles (Hamilton, 2009). Ce modèle suppose que les métapopulations ont atteint un équilibre et que la valeur de différenciation génétique ne change plus d'une année à l'autre. Il est admis que dans ce modèle la valeur de différenciation génétique attendue à l'équilibre est fonction de la **taille de population efficace** *, N_e, et du taux de dispersion, m, avec une correction dûe au nombre fini de populations considérées n (ici n=2) :

$$F_{ST}\ attendu = \frac{1}{4N_e m(\frac{n}{n-1})^2 + 1} \tag{2.1}$$

Trois tailles de populations efficaces ont été étudiées, chacune constituées respectivement de 500, 2 000 et 5 000 individus. Pour la métapopulation idéale, la **taille de population absolue** est égale à la taille de population efficace. Pour la métapopulation "Rajidae", l'ensemble de la fraction mature de chaque population a été considérée comme la taille de population efficace. La démographie décrite implique que la taille de population absolue est environ dix fois plus que grande que la taille de population efficace. Le tableau suivant donne les correspondances entre la taille absolue d'une population (N) et sa taille efficace (N_e) pour les deux types de métapopulations (Tab. 2.1).

*Cette notion est détaillée dans le chapitre 5

Métapopulation "idéale"		**Métapopulation "Rajidae"**	
N	N_e	N	N_e
500	500	5 000	500
2 000	2 000	20 000	2 000
5 000	5 000	50 000	5 000

TABLE 2.1 : Tailles de populations absolues simulées pour les deux types de métapopulation et tailles de populations efficaces correspondantes

1.2 Les générations chevauchantes : un effet tampon

Les premiers résultats, attendus, montrent une forte dépendance du degrée de différenciation génétique avec la taille de population efficace (non montrés). Plus cette dernière augmente, plus la valeur de F_{ST} diminue. De plus, plus la dispersion est faible, moins les populations échangent et donc plus la valeur de F_{ST} augmente (Fig. 2.2).

Pour de faibles taux de dispersion (de 0,001 à 0,005 environ), la différenciation génétique est plus forte pour la métapopulation "idéale" que pour la métapopulation "Rajidae" (environ 6 fois plus grande pour m=0,001) mais cette dernière présente une variabilité inter-replicas moindre (Fig. 2.2). Le même phénomène est visible pour les tailles de populations efficaces 2 000 et 5 000 individus (non montrées).

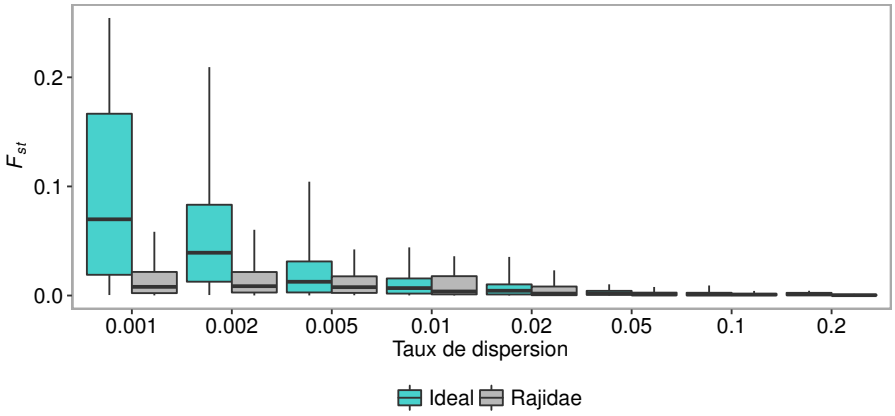


FIGURE 2.2 : Boîte à moustaches du F_{ST} simulé obtenu pour la métapopulation "idéale" (en bleu) et "Rajidae" (en gris) avec N_e=500 individus (Moyenne, 2,5%, 25%, 75%, 97.5%)

La comparaison avec les F_{ST} attendus (calculés avec l'eq. 2.1) montre que pour des taux de dispersion forts (0,01 à 0,2 environ), les deux metapopulations se comportent de manière similaire et correspondent relativement bien aux F_{ST} attendus (Fig. 2.3). Par contre, pour des taux de dispersion faibles (de 0,001 à 0,005 environ), la métapopulation "Rajidae" s'écarte du schéma attendu et les valeurs de F_{ST} sont bien inférieures à celle attendues (7 fois inférieures pour m=0,001). Les distributions de F_{ST} obtenues pour les 100 replicas "Rajidae" sont ainsi similaires à ceux "idéale" pour des taux de dispersion forts mais pour des taux de dispersion faibles, les F_{ST} "Rajidae" sont plus faible qu'attendus. Là encore, ce phénomène est visible pour les trois tailles de populations simulées.

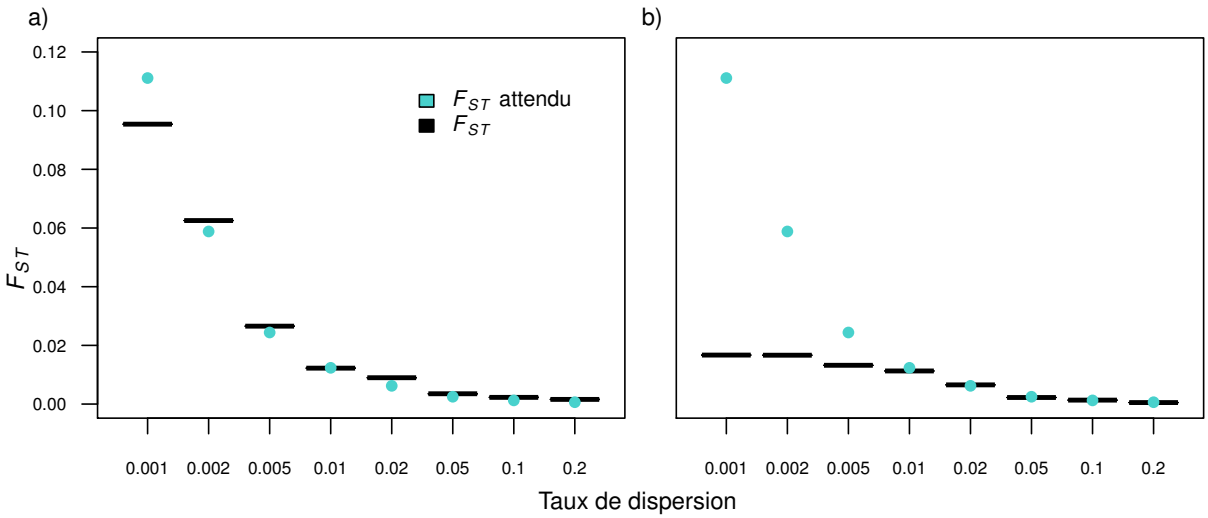


FIGURE 2.3 : Moyenne du F_{ST} obtenu à l'issue de 250 années de simulation pour 100 réplicas (tirets noirs) et comparaison avec la valeur attendue calculée par l'eq 2.1 (points bleu) avec N_e=500 individus. a) Métapopulation "idéale" b) Métapopulation "Rajidae".

Il semble que les générations chevauchantes aient un effet "tampon" et diminuent le taux d'accroissement du F_{ST} dans le temps. En effet, dans une métapopulation "idéale", il suffit d'un an pour avoir un total réarrangement du patrimoine génétique alors que dans une métapopulation à générations chevauchantes, comme la métapopulation "Rajidae", il faut plusieurs années pour voir ce réarrangement opérer. Ce même phénomène explique la variabilité inter-réplicas moindre observée pour la métapopulation "Rajidae".

De plus, il est intéressant de noter que pour des taux de dispersions très faibles (0,001 ; 0,002), la métapopulation "idéale" commence également à s'écarter des attendus théoriques (d'environ 7 à 15%). Ce phénomène est probablement dû à la non-stabilisation du F_{ST} qui continue d'augmenter même après 250 ans (et même après 10 000 ans). En effet, pour rappel, la formule utilisée pour calculer les F_{ST} attendus correspond à l'atteinte d'un équilibre de la métapopulation étudiée (eq. 2.1).

2. Article : Apports de la connectivité génétique et démographique pour la gestion des raies

Canadian Journal of Aquatic and Fisheries Sciences - 02 octobre 2017
DOI : 10.1139/cjfas-2017-0291

Florianne Marandel[1], Pascal Lorance[1], Marco Andrello[2], Grégory Charrier[3], Sabrina Le Cam[3,4] and Verena M. Trenkel[1]

RESUME - Au cours du 20ème siècle, de nombreuses espèces de raies et requins ont décliné en Atlantique Nord-Est à cause de la pêche. La définition d'unités de gestion est fondamentale pour éviter le déclin voire la disparition de ces populations. Les unités de gestion à prendre en compte pour exploiter **durablement** ces populations dépendent en grande partie de leur degré de connectivité. Dans ce travail, une étude simulatoire des connectivités démographique et génétique est réalisée afin d'évaluer leur apport pour la gestion des raies. L'étude se concentre sur 11 populations putatives de raie bouclée des eaux européennes (Atlantique Nord-Est, Açores et Méditerranée). Les simulations génétiques corroborent les résultats empiriques disponibles et mettent en avant trois grandes unités peu génétiquement différenciées (Atlantique Nord-Est, Açores et Méditerranée). L'étude démographique met en avant plus de différenciation au sein de l'unité Atlantique Nord-Est et suggère une gestion combinée pour plusieurs paires de populations putatives notamment en Manche Est et dans le Sud de la mer du Nord. Ces résultats sont sensibles aux hypothèses de taux de dispersion et d'hétérogénéité spatiale. Néanmoins, la connectivité démographique apparait comme fondamentale pour une gestion efficace, lorsque la connectivité génétique donne une perspective sur la conservation à long terme.

[1] Ifremer, rue de l'île d'Yeu, BP 21105, 44311 Nantes Cedex 3, France

[2] EPHE, CEFE UMR 5175, CNRS, Université de Montpellier, Biogéographie et Ecologie des Vertébrés, 1919 route de Mende, 34293 Montpellier, France

[3] Université de Bretagne Occidentale, Laboratoire des Sciences de l'Environnement Marin (LEMAR, UMR 6539 CNRS/IRD/UBO/Ifremer), Institut Universitaire Européen de la Mer (IUEM), Technopôle Brest-Iroise, 29280 Plouzané, France

[4] Ifremer, Génétique et pathologie des mollusques marins, La Tremblade sur mer, France

NRC Research Press

ARTICLE

Insights from genetic and demographic connectivity for the management of rays and skates

Florianne Marandel, Pascal Lorance, Marco Andrello, Grégory Charrier, Sabrina Le Cam, Sigrid Lehuta, and Verena M. Trenkel

Abstract: Studying demographic and genetic connectivity can help assess marine metapopulation structure. Rays and skates have no larval phase; hence, population connectivity can only result from active movement of individuals. Using thornback ray (*Raja clavata*) in European waters as a case study, demographic and genetic connectivity were studied for 11 putative populations with unequal population abundances and two hypotheses of dispersal rates. Genetic simulation results highlighted three large metapopulations: in the Mediterranean, around the Azores, and on the Northeast Atlantic shelf. Demographic results highlighted a finer population structure indicating that several pairs of putative populations might be demographically linked. Results were highly sensitive to dispersal assumptions and relative population abundances, which provided insights into the potential magnitude of genetic and demographic connectivity differences. Accounting for demographic connectivity appears to be crucial for managing and conserving rays and skates, while genetic connectivity provides a longer-term perspective and less subtle spatial structures. Moreover, accounting for heterogeneity in population abundances is a key factor for determining or interpreting metapopulation connectivity.

Résumé : L'étude de la connectivité démographique et génétique peut aider à évaluer la structure des métapopulations marines. Les raies n'ont pas de stade larvaire, de sorte que la connectivité des populations ne peut découler que de déplacements actifs d'individus. Utilisant le cas de la raie bouclée (*Raja clavata*) dans les eaux européennes, nous avons étudié la connectivité démographique et génétique pour 11 populations présumées en simulant des abondances inégales des populations ainsi que deux hypothèses de dispersion. Les résultats des simulations génétiques font ressortir trois grandes métapopulations, une dans la Méditerranée, une autour des Açores et une sur le plateau continental de l'Atlantique nord-est. Les résultats démographiques font ressortir une structure de populations plus fine qui indique que plusieurs paires de populations présumées pourraient être reliées du point de vue démographique. Les résultats sont très sensibles aux hypothèses concernant la dispersion et aux abondances relatives des populations, ce qui fournit de l'information sur l'ampleur possible des différences de connectivité génétique et démographique. La prise en considération de la connectivité démographique semble revêtir une importance clé pour la gestion et la conservation des raies, alors que la connectivité génétique fournit une perspective à plus long terme et des structures spatiales moins fines. La prise en considération de l'hétérogénéité de l'abondance des populations constitue en outre un facteur clé pour déterminer ou interpréter la connectivité de métapopulations. [Traduit par la Rédaction]

Introduction

Population connectivity is a crucial parameter to take into account when defining population units relevant for management and conservation purposes (Stearns and Hoekstra 2005; Sinclair et al. 2006; Schwartz et al. 2007). Population units relevant for management purposes are commonly defined using ecological and genetic information. For example, the degree of connectivity between populations can be assessed by the amount of exchanged individuals (Waples and Gaggiotti 2006). By reproducing in the population they joined, these individuals contribute to the local demography but also transfer their genetic material into the gene pool, thus inducing gene flow among populations. A metapopulation is a network of local populations that exchanges individuals but has somewhat independent dynamics (Levins 1969). Determining how many local or metapopulations exist and characterizing the relationships among them is a challenging task. Numerous definitions of the population concept exist (see Waples and Gaggiotti 2006). Almost all involve interbreeding individuals over a geographical area. However, they lack objective and quantitative criteria for delimiting distinct populations. In this context, different researchers might identify different population structures based on the same information (Waples and Gaggiotti 2006).

Received 17 July 2017. Accepted 2 October 2017.

F. Marandel, P. Lorance, S. Lehuta, and V.M. Trenkel.* Institut Français de Recherche pour l'Exploitation de la Mer (Ifremer), BP 21105, 44311 Nantes cedex 3, France.
M. Andrello. EPHE, PSL Research University, CEFE UMR 5175, CNRS, Université de Montpellier, Biogéographie et Ecologie des Vertébrés, 1919 route de Mende, 34293 Montpellier, France.
G. Charrier. Université de Bretagne Occidentale, Laboratoire des Sciences de l'Environnement Marin (LEMAR, UMR 6539 CNRS/IRD/UBO/Ifremer), Technopôle Brest-Iroise, 29280 Plouzané, France.
S. Le Cam. Université de Bretagne Occidentale, Laboratoire des Sciences de l'Environnement Marin (LEMAR, UMR 6539 CNRS/IRD/UBO/Ifremer), Technopôle Brest-Iroise, 29280 Plouzané, France; Institut Français de Recherche pour l'Exploitation de la Mer (Ifremer), Avenue de Mus de Loup, Ronce les Bains, 17390 La Tremblade, France.
Corresponding author: Florianne Marandel (email: florianne.marandel@ifremer.fr).
*Verena M. Trenkel currently serves as an Associate Editor; peer review and editorial decisions regarding this manuscript were handled by Eric Taylor.

Population connectivity has two components: genetic and demographic. Genetic connectivity is defined as the degree to which gene flow affects evolutionary processes within populations, and demographic connectivity is the relative contribution of dispersal to population dynamics (Lowe and Allendorf 2010). The two components inform on population connectivity at evolutionary and ecological time scales, respectively. For example, high genetic connectivity does not necessarily imply high demographic connectivity and that a single unit should be considered for management (Hawkins et al. 2016).

On one hand, genetic connectivity is often derived from the measure of genetic differentiation between populations shaped by the interplay of the four evolutionary forces (gene flow, genetic drift, selection, and mutation; Hallerman 2003; Stearns and Hoekstra 2005). Their respective magnitude is greatly dependant on population abundance, and only gene flow is relevant as the genetic component of population connectivity. Therefore, depending on the population abundance, genetic differentiation can be a poor proxy of gene flow. For example, in the case of large populations, genetic differentiation can remain weak despite reduced levels of gene flow because of a very low genetic drift. In studies of genetic population structure, the absolute number of migrants is used preferentially to the migration rate, as it conveys information on population abundance, i.e., census population size (Palumbi 2003). However, knowledge on population abundance is often lacking, especially for marine populations.

On the other hand, demographic population structure greatly depends on life history traits. Indeed, following the definition above of demographic connectivity, demographically connected populations display intrinsic growth or survival rates that are reciprocally affected by immigration or emigration (Lowe and Allendorf 2010). Thus, evaluation of demographic connectivity requires information on the contribution of dispersal but also on the demographic rates and abundance of each population. This information is not only needed for assessing the effect of dispersal on population growth rates but also for defining threshold values for demographic connectivity. Despite the importance of demographic connectivity for defining management units, defining appropriate thresholds for demographically connected populations has received relatively little attention in the literature (Waples and Gaggiotti 2006; Waples et al. 2008). Importantly, for both genetic and demographic connectivity studies, dispersal is a central process that we define as an individual leaving the home range of its birth population to move to another population's home range, the movement being one way and not a round trip (Dingle 2014).

In this study, we evaluated the use of genetic and demographic connectivity for identifying management units of rays and skates. Bycaught in several fisheries, many populations of skates and rays have declined, sometimes strongly, in the Northeast Atlantic during the 20th century (Quéro and Cendrero 1996; Dulvy et al. 2014). Therefore, the conservation of these species has become a major objective for ensuring sustainable exploitation of marine resources (Dulvy et al 2014; Davidson et al. 2016). For reaching this objective, it is fundamental to delimit appropriate management units. However, for many species, available data are restricted to life history traits (though not always available for all populations), landings, and, only in some cases, survey time series informing on abundance changes (ICES 2016). In European waters, recent landings levels of most ray and skate species differ greatly among the southern North Sea, the western Mediterranean Sea, and the Azores, which can be considered indicative of differences in population abundances, as there are no species-specific catch quotas in place for these species (Fig. 1*a*) (ICES 2016).

In contrast with teleosts, rays and skates have a low potential for dispersal and thus gene flow. They produce few offspring that develop in egg capsules fixed to the seabed during several months (Hoening and Gruber 1990). As there is no pelagic larval stage that can be dispersed by marine currents, the dispersal of rays and skates is solely based on the movements of juveniles and (or) adults. Ovenden (2013) argued that due to dispersal happening at later life stages, elasmobranch species might present crinkled connectivity, which she defined as a situation where dispersal is large enough to make populations genetically similar but too small to matter for demographic connectivity.

For several medium-sized ray species, comparable degrees of movement between adjacent populations have been observed (Walker et al. 1997; Hunter et al. 2005*a*, 2005*b*; Stephan et al. 2015). However, for two ray species sharing similar life history traits and overlapping spatial ranges, genetics studies have indicated distinct structuring patterns. The thornback ray (*Raja clavata*) seems to display strong spatial genetic differentiation (Chevolot et al. 2006), while the thorny skate (*Amblyraja radiata*) presents only weak although statistically significant differentiation (Chevolot et al. 2007). These contrasting results could be due to differences in dispersal behavior, despite similar biology, along the continental shelf edge, which might not constitute a barrier for thorny skate but it could be one for thornback ray (Chevolot 2006). However, while dispersal is a necessary condition for connectivity, it is not sufficient because dispersed individuals need to reproduce in the receiving population. So the difference could also be caused by differential integration into the spawning components.

In this study, we evaluated the use of genetic and demographic connectivity for identifying management units of rays and skates. We use the term "putative population" for assumed populations occurring at discrete sampling locations. Demographic and genetic criteria are applied to these putative populations to evaluate their connectivity. We used a modelling approach with life history parameters resembling those of thornback ray, a typical widespread ray species in European waters. This species is the most studied ray in the Northeast Atlantic. Its biological parameters are representative for medium-bodied skates and rays (refer to online Supplementary Table S1[1]). Similar to other rays and skates, local abundances of thornback ray differ strongly, taking recent landings as an indication of population abundances (Fig. 1*a*). As no reliable dispersal rate estimates were available on a European scale (but see Walker et al. 1997 for regional dispersal values), we defined plausible scenarios based on expert knowledge. Genetic connectivity was evaluated by calculating the fixation index F_{ST} defined by Wright (1949). In contrast, we did not evaluate demographic connectivity sensus stricto, as the effect of dispersal on population growth rates was not investigated directly. Instead, we evaluated whether the number of dispersed individuals was likely to contribute to local population abundances. In addition, using a matrix model we identified the life history parameters and life stages to which the population growth rate was most sensitive. These results provided context for interpreting the importance of dispersal for population growth.

Materials and methods

Life history parameters and intrinsic population growth

Usher matrix model

To appraise the potential importance of dispersal for population dynamics of a thornback ray-like species, we studied the sensitivity and elasticity of intrinsic population growth rate to variations and uncertainty in life history traits. To this aim we used an Usher matrix model (Usher 1966). The model consisted of four life stages grouping ages with similar demographic parameters:

[1]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/cjfas-2017-0291.

Fig. 1. (*a*) Standardized mean landings 2013–2015 ((value – min.)/(max. – min.)) of ten skate and ray species from the Northeast Atlantic grouped by spatial distribution pattern (spatial units as in panel (*b*)). Spatial units considered as single stocks by the International Council for Exploration of the Sea are connected by solid lines. (*b*) Thornback ray putative populations with dispersal paths (black solid lines). 200 m isobaths are shown in grey. Figure created using R statistical software and the "PBSmapping" package (Schnute et al. 2017).

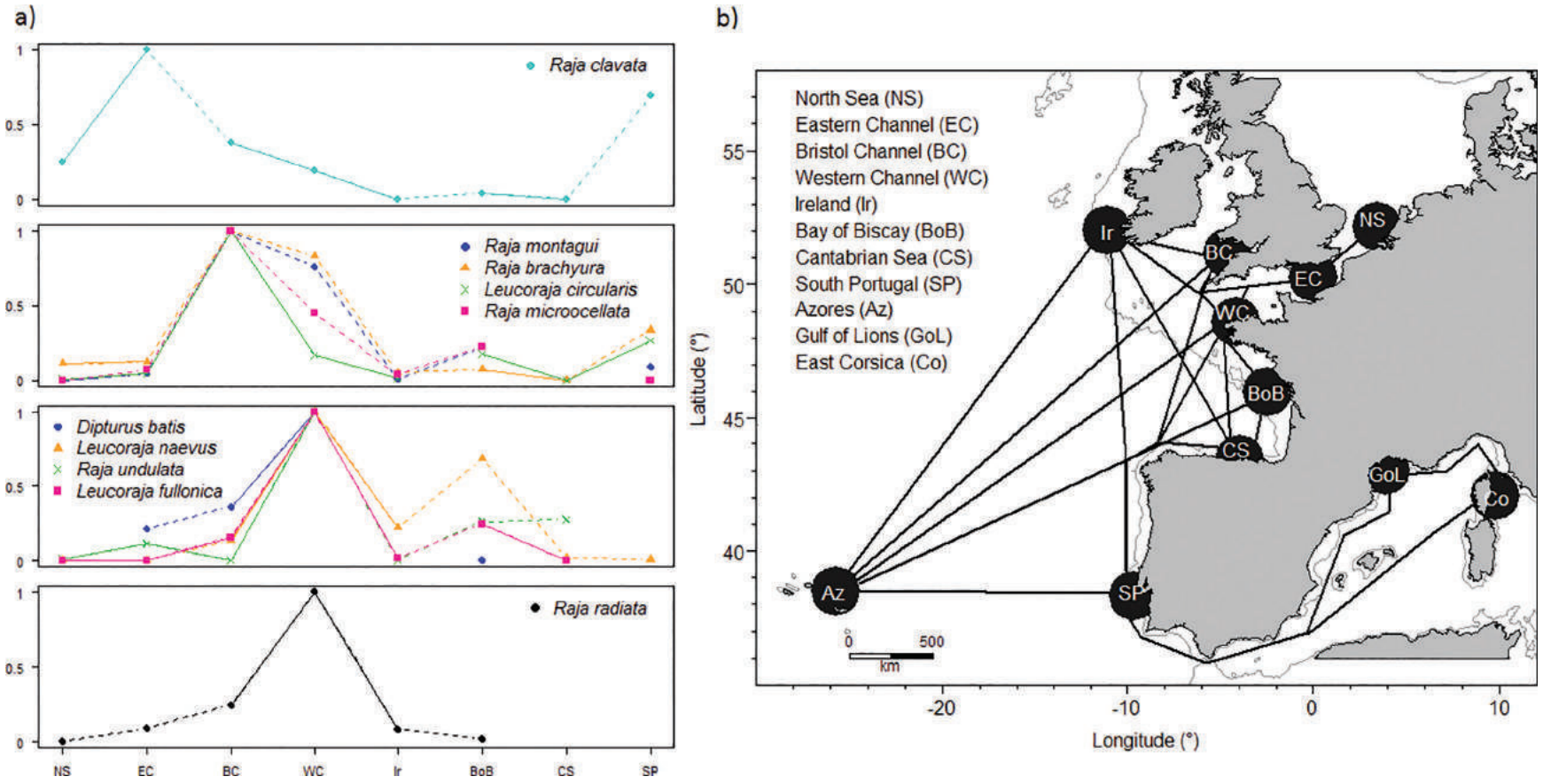


Table 1. Parameters values used in the Usher matrix model for the local elasticity as well as the minimum and maximum values used for the global sensitivity (Morris) analysis.

				Morris method	
Parameter	Description	Baseline value	Alternative value	Min.	Max.
Ω_4	No. of eggs stage 4	150[a]	48[b]	40	150
S_0	Egg and newborn survival	0.036[c]	0.01	0.01	0.5
S_1	Survival stage 1	0.20[c]	0.40	0.1	0.99
S_2	Survival stage 2	0.69[c]	0.40	0.1	0.99
S_3	Survival stage 3	0.81[c]	0.90	0.1	0.99
S_4	Survival stage 4	0.48[c]	0.90	0.1	0.99
H_2	Maturation rate stage 2	0.2[d]	0.3	0.05	1
H_3	Rate of fecundity increase stage 3	0.1[d]	0.2	0.05	1

[a]Holden et al. 1971.
[b]Ellis and Shackley 1995.
[c]Derived in this study.
[d]Educated guess.

stage 1 included age 1 individuals, stage 2 ages 2 to 6 (immature), stage 3 ages 7 to 11 (mature individuals), and stage 4 ages 12 years and older (highly fecund mature individuals).

$$(1)\qquad \mathbf{N}_{t+1} = \mathbf{U}\mathbf{N}_t,\quad \mathbf{U} = \begin{pmatrix} 0 & 0 & S_0 f_3 & S_0 f_4 \\ S_1 & S_2(1-H_2) & 0 & 0 \\ 0 & S_2 H_2 & S_3(1-H_3) & 0 \\ 0 & 0 & S_3 H_3 & S_4 \end{pmatrix}$$

where $\mathbf{N}_t$ is the vector of abundance by life stage, and $\mathbf{U}$ is the transition matrix. Each life stage s has a specific survival rate S_s. A maturation rate H_2 between stages 2 and 3 was derived from expert knowledge. S_0 is first year survival (from egg laying to 1 year). Fecundity was assumed to depend on the number of eggs Ω_s produced and the proportion of females P_{φ} ($f_s = \Omega_s P_{\varphi}$). Stage 3 was assumed less fecund than stage 4; this was achieved by setting the number of eggs to $\Omega_3 = 0.7\Omega_4$. So H_3 represents an increase in the rate of fecundity between stages 3 and 4. A sex ratio of 1:1 was assumed ($P_{\varphi} = 0.5$), which is typical for rays and skates (Steven 1933; Ellis and Shackley 1995; Delpiani 2016).

Survival rate estimates were obtained in the following way. First, size-at-age L_a was estimated using the von Bertalanffy equation with growth parameters t_0, K, and L_∞ from the literature (Serra-Pereira et al. 2008). Second, mortality-at-age M_a was estimated using the empiric relationship developed by Gislason et al. (2010):

$$(2)\qquad \ln(M_a) = 0.55 - 1.61\ln(L_a) + 1.44\ln(L_\infty) + \ln(K)$$

Third, mortality M_a was transformed to survival-at-age $S_a = \exp(-M_a)$. To account for senescence, survival was reduced from age 12 onwards (Fig. S1 in electronic Supplementary material[1]). Survival by life stage was calculated by averaging over the corresponding ages (except for S_1). As no information was available for first year survival (S_0), the value was chosen to lead to a stable population. All baseline parameter values are summarized in Table 1.

Elasticity and sensitivity analyses

At equilibrium, the first positive eigenvalue of $\mathbf{U}$ (eq. 1) corresponds to the intrinsic population growth rate λ. Local elasticity and global sensitivity analyses were conducted to identify the life history parameters and life stages the value of λ was most sensitive to.

Elasticity represents the relative change in population growth rate in response to a certain relative change in vital rate parameters $P_i \in (\Omega_4, S_0, S_1, S_2, S_3, S_4, H_2, H_3)$. Thus, elasticity can be compared for parameters of different unit, such as survival rates S_s ($0 \le S_s \le 1$) and the number of eggs ($0 \le \Omega_4$).

Elasticity E of population growth rate λ to changes in parameter P was defined as follows (Caswell 2001):

$$(3)\qquad E = \frac{P_i}{\lambda}\frac{\Delta\lambda}{\Delta P_i}$$

Two types of parameter value changes were tested one parameter at a time: alternative parameter values based on available data or expert guess and a ±10% change of all baseline parameter values (Table 1).

The global sensitivity analysis was carried out using the Morris (1991) method improved by Campolongo et al. (2007) and Pujol (2009). The method consists of calculating successively the so-called elementary effect, which corresponds to the change in the intrinsic population growth rate λ when the current value of P_i is changed successively by adding or subtracting Δ_i along trajectories, holding all other parameter values constant, divided by Δ_i:

$$(4)\qquad d_i(P) = \frac{\lambda(P_{\Delta_i}) - \lambda(P)}{\Delta_i}$$

The step size Δ_i for each parameter was defined as

$$(5)\qquad \Delta_i = \frac{\max(P_i) - \min(P_i)}{k}$$

These elementary effects are computed at various locations of the parameter space so interaction effects can be evidenced. Here, $k = 0$ was used and rather extreme maximum and minimum parameter values were chosen (Table 1). The method uses an efficient sampling design that led to 16 200 estimates of $d_i(P)$. These were then summarized by calculating for each parameter P_i the mean of absolute effects $\mu_i^* = \overline{|d_i(P)|}$, which quantifies sensitivity, and the variance σ^2, which informs on the strength of interactions. All calculations were carried out in R using the package "sensitivity" setting the number of elementary effect computed per factor to 1800 (V1.14.0; Pujol et al. 2014). Calculations were repeated 10 times to ensure convergence.

Evaluating connectivity

Dispersal between putative populations

For evaluating demographic and genetic connectivity, 11 putative populations of thornback ray were assumed, based mainly on

Table 2. Mean landings of thornback ray for the period 2013–2015, simulated population sizes for the 11 studied locations, ICES stock identity for nine locations (Mediterranean stocks are not covered by ICES), and membership of putative populations to genetic and demographic units derived from modeling results (see text) for dispersal scenarios 1 (SC1) and 2 (SC2).

					Genetic units (Gu)		Demographic units (Du)	
Code	Location (ICES division)	Landings (tonnes)	Simulated population size	ICES stock identity	SC1	SC2	SC1	SC2
NS	North Sea (4c)	427.3[a]	125 000	rjc-347d	Gu1	Gu1	Du1	Du1
EC	Eastern Channel (7d)	1112.9[a]	300 000	rjc-347d	Gu1	Gu1	Du1	Du1
BC	Bristol Channel (7fg)	549.4[a]	165 000	rjc-7afg	Gu1	Gu1	Du1	Du1
WC	Western Channel and Southern Celtic Sea (7he)	378.9[a]	90 000	rjc-echw	Gu1	Gu1	Du1	Du1
Ir	Ireland (7bcjk)	204.0[a]	50 000	Grouped with other Rajidae	Gu1	Gu1	Du1	Du1
BoB	Bay of Biscay (8ab)	241.4[a]	60 000	rjc-bisc	Gu1	Gu1	Du1	Du2
CS	Cantabrian Sea (8c)	204.9[a]	60 000	rjc-bisc	Gu1	Gu1	Du1	Du3
SP	South Portugal (9a)	840.4[a]	215 000	rjc-pore	Gu1	Gu1	Du1	Du3
Az	Azores (10)	180.0[b]	50 000	Grouped with other Rajidae	Gu2	Gu2	Du2	Du4
GoL	Gulf of Lion	15.2[c]	10 000	NA[d]	Gu3	Gu3	Du3	Du5
Co	Corsica	NA	10 000	NA[d]	Gu4	Gu4	Du4	Du6

[a]ICES 2016.
[b]Unpublished data.
[c]FAO 2016.
[d]Not covered by ICES.

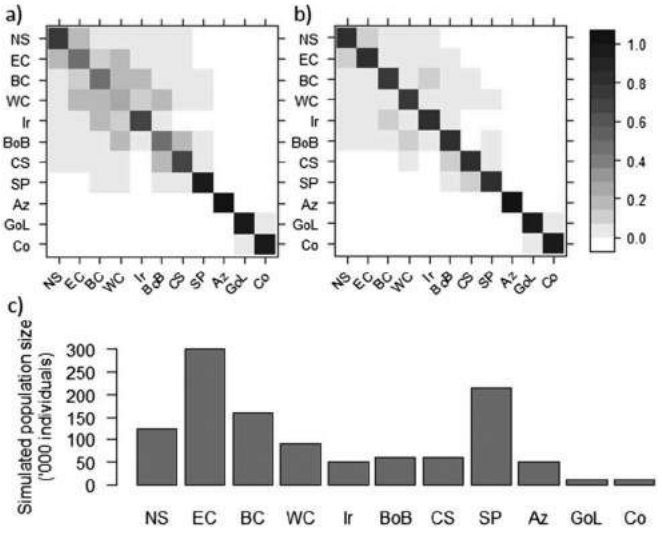
expert judgement and results from a population genetics study (Chevolot et al. 2006) (Fig. 1*b*). Only a few tagging studies were available for medium-sized skate and ray species, but all showed recapture or travelling distances less than 150 km (Supplementary Table S1[1]; Walker et al. 1997; Hunter et al. 2005*a*, 2005*b*). The International Council for the Exploration of the Sea (ICES) currently assesses several of these 11 putative populations together (Table 2).

To create the matrix of dispersal probabilities between the 11 putative populations, we first traced the pairwise shortest marine geographical distances (Fig. 1*b*) and set low dispersal rates between populations separated by depths greater than 200 m, as the modelled species is found in shallower waters (Quéro and Vayne 1997). Given the lack of information on potential migrations between the 11 putative populations, two somewhat arbitrary dispersal scenarios were tested. The first scenario corresponded to a strong potential for dispersal and the second to a more sedentary behavior. In the first scenario, the dispersal probability between adjacent populations i and j was set to $d_{ij} = 0.1$ between putative populations WC and Ir and BC and EC (see Fig. 1*b*) and to 0.2 for the other adjacent populations. Dispersal probabilities between more distant populations were obtained by accounting for the number of populations between the origin and arrival population along the dispersal tracks in Fig. 1*b*, e.g., dispersal between WB and CS populations via BoB was set to $0.2 \times 0.2 = 0.04$. The probability to remain in the same population was calculated as $d_{ii} = 1 - \sum_{j=1}^{10} d_{ij}$; it decreased with increasing number of direct neighbours. Note that in this scenario dispersal probabilities were symmetrical, i.e., $d_{ij} = d_{ji}$. The full dispersal matrix can be found in Supplementary Table S2[1].

In the second dispersal scenario, the probability d_{ii} to remain in the same population was fixed between 0.75 and 0.99 according to the size of the neighbourhood, with smaller values for geographically more central populations. Dispersal probabilities d_{ij} were a linear function of distance between populations, rescaled so that the sum of all probabilities summed to 1 (Supplementary Table S3[1]). This implied nonsymmetrical dispersal probabilities, i.e., $d_{ij} \neq d_{ji}$. The two dispersal matrices were used for studying demographic and genetic connectivity (Figs. 2*a*, 2*b*).

For genetic connectivity estimation (see below), dispersal was assumed to occur before the first birthday of egg laying only. This is equivalent to assuming that each individual reproduces in one population only during its life, either in its native population or in the one it dispersed to. This type of behavior has been observed for several fish species, with individuals breeding several years in the same area (Dittman and Quinn 1996; Feldheim et al. 2014; Bonanomi et al. 2016). For example, Bonanomi et al. (2016) found that Atlantic cod (*Gadus morhua*) can disperse more than 1000 km from their place of birth and spend several years growing in the place they dispersed to before returning as mature individuals to their place of birth for breeding. Natal philopatry is well-described for sharks (Feldheim et al. 2014) but not for skates and rays. However, for the thornback ray, a study using data storage tags indicated that most individuals were philopatric with a maximum travelling distance of 130 km (Hunter et al. 2005*a*, 2005*b*).

Fig. 2. Dispersal probabilities applied for dispersal scenarios 1 (*a*) and 2 (*b*) and population abundances assumed for genetic and demographic connectivity studies (*c*). For population abbreviations, see Fig. 1*b*.



Population abundance

To derive population abundances for the 11 putative populations, recent international landings were used (ICES 2016; FAO 2016). The Eastern (English) Channel (EC) population had the highest landings and the Gulf of Lion population the lowest (Table 2; Fig. 1*a*). As little regional information was available, landings in biomass were considered reflecting relative population abundances in numbers. For computational reasons, the maximum population abundance, corresponding to the population in the EC, was fixed to 300 000 individuals; this is of course much smaller than the actual population size. The abundance of the other populations, except for the two Mediterranean populations, was then set applying the ratio between their landings and those of the EC. For the two Mediterranean populations, instead of the corresponding number of individuals, 10 000 individuals were assumed to avoid too small numbers (Table 2; Fig. 2*c*). These population abundances will be referred to as N_0 below.

Genetic connectivity

Among the many simulators available for population genetics studies, only a few are designed to account for complex life histories and large population abundances. The R package MetaPopGen (v3.1.2; R Development Core Team 2008; v0.0.4; Andrello and Manel 2015) is such a simulator; it can simulate population genetic data for species with complex life history traits (overlapping generations, age-specific survival and fecundity, etc.) in a reasonable time but can simulate only one locus.

The simulations were set up to follow as much as possible the life cycle of medium-sized rays and skates taking the thornback ray as model. Yearly, each mature individual produced gametes according to its fecundity and sex. Fecundity was fixed to 140 gametes for females and 10 000 for males to represent the situation where female gametes are limiting reproduction. Gametes were subject to mutation, with a mutation rate of 1E-06, and fused into eggs with a particular sex and genotype. The dispersed newborns became part of a new population following a recruitment function, which was adapted to correspond to mortality from hatching to age 1. Survival-at-age was as in Supplementary Fig. S1[1].

For both dispersal scenarios, the 11 putative populations were simulated for 10 000 years, repeating the simulations 200 times. This time horizon was selected to study long-term effects of genetic differentiation since the last glacial maximum (Hewitt 2000). Population abundances remained constant at N_0 throughout the whole simulation period (Table 2). Dispersal probabilities varied between populations but were constant during the 10 000 years, as no information on changes in dispersal was available.

For each replicate, one neutral biallelic locus was simulated with an allelic frequency of 0.5 for all populations at the beginning of the simulation (year 0). As our aim was to identify management units based on genetic measures calculated for the whole population (observation errors were ignored) in a simulation framework, one locus was sufficient. Note that the 200 replicates cannot be considered as 200 independent loci, as each replicate had a different number of individuals in year t due to the stochasticity of population and genetic inheritance dynamics. The 11 putative populations were initially undifferentiated in all replicates.

Global and pairwise genetic differentiation between putative populations was estimated using the fixation index F_{ST} (Hamilton 2009) for each replicate and simulation year:

$$F_{ST} = (H_T - H_S)/H_T \quad (6)$$

where H_T is the heterozygosity in the pooled putative populations, and H_S is the mean heterozygosity in each population. Mean and median global F_{ST} were calculated over the 200 replicates.

To evaluate the contribution of the difference in abundance to genetic differentiation, we also simulated 11 putative populations with identical population abundances (10 000 individuals). From this simulation we selected population pairs leading to $F_{ST} > 0.001$; $F_{ST} < 0.001$ were considered to indicate absence of genetic differentiation, in which case differences in abundance cannot play a role. This threshold value was obtained from the relationship between pairwise F_{ST} values and the mean number of migrants between pairs (see Results).

The effect of differences in abundance on F_{ST} values was evaluated by calculating the ratio between the F_{ST} values of the selected pairs ($F_{ST} > 0.001$) for simulations with identical abundance and for simulations with different abundances as described above. These ratios of F_{ST} values were then linearly regressed on the absolute difference in abundance between pairs of exchanging populations.

Demographic connectivity

For studying global demographic connectivity C for each putative population i, the relative change in abundance after a single dispersal event was calculated as

$$C_i = \frac{-N_{0i}\sum_{j\neq i} d_{ij} + \sum_{j\neq i} d_{ji}N_{0j}}{N_{1i}} \quad (7)$$

where N_{0i} is the abundance before the dispersal event for population i (see Table 2), N_{1i} is the abundance after the dispersal event for population i, and d_{ij} is the dispersal rate from population i to population j as described above. This approach was chosen as life history parameters were not available for all putative populations, making it impossible to use a more detailed dynamic modelling approach. Further, considering a short-term perspective is in line with the requirement that exchanges between connected populations have to occur in most years for using them as basis for defining management units (Hawkins et al. 2016).

The origin of individuals in each population i after a single dispersal event was then investigated by calculating the proportion of individuals in population i that came from population j, $P_{i|j}$, as follows:

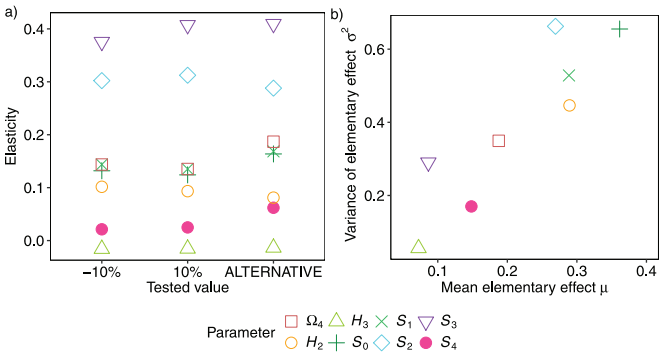
$$P_{i|j} = \frac{N_{0j}d_{ji}}{N_{1i}} \quad (8)$$

This metric $P_{i|j}$ is referred to as pairwise demographic connectivity.

Waples and Gaggiotti (2006) and Palsbøll et al. (2007), based on Hastings (1993), considered that populations linked by ≥10% migrants should be assigned to the same management unit. We therefore compared pairwise demographic connectivity estimates $P_{i|j}$ with the threshold value of 0.1. This choice is somewhat arbitrary and will be discussed below.

To evaluate the contribution of the difference in abundance to demographic connectivity, we also calculated the pairwise demographic connectivity between the 11 putative populations with identical population abundances (10 000 individuals). Then, we selected population pairs leading to $P_{i|j} > 0.1$. As for genetic connectivity, the effect of differences in abundance on the demographic connectivity values was evaluated by calculating the ratio between the $P_{i|j}$ values of the selected pairs with identical abundance and the $P_{i|j}$ values of the same pairs with different abundance. These ratios of $P_{i|j}$ values were then linearly regressed on the difference in abundance between pairs of exchanging populations.

Fig. 3. (*a*) Local parameter elasticity and (*b*) global Morris sensitivity analysis results for the Usher matrix model for thornback ray. For parameter definitions and values, see Table 1.



Results

Life history parameters and intrinsic population growth

Using the reference parameter values in the Usher matrix model, at equilibrium, 41% of individuals were 1 year old (stage 1), 39% immature (stage 2), and 20% mature (stages 3 and 4), among which 13% belonged to the stage 4 category representing mature individuals with a larger number of eggs.

Local elasticity analysis identified the survival of stage 3 individuals (S_3) as the parameter to which intrinsic population growth was the most reactive, while it was least reactive (and reacting negatively) to the rate of fecundity increase from stage 3 to stage 4 (H_3) (Fig. 3*a*); similar results were obtained for the set of alternative parameter values and when varying parameter values by ±10%. All elasticity values were <0.2 except for S_3 and S_2, thus indicating a degree of robustness to parameter value changes.

Results of the global sensitivity analysis (Morris method) differed from the local elasticity analysis (Fig. 3*b*). S_0 was the parameter with the highest influence on intrinsic population growth (λ) changes and the highest strength of interactions (large σ^2), directly followed by S_1, S_2, and H_2. Changes in parameters S_3, S_4, and H_3 had little influence on λ. These parameters interacted little, with the smallest value for σ^2 reached by H_3. The influence of the number of eggs Ω_4 was intermediate.

From these complementary analyses, it appeared that survival rates of immature individuals (S_0, S_1, and S_2) were the parameters to which the population growth rate was most sensitive. On the other hand, parameters linked to the mature stages (S_3, S_4, and H_3) appeared less important in terms of contribution to changes in λ and interactions with others parameters.

Genetic connectivity

Two hundred replicates were sufficient to capture stochastic variations (not shown). The median final global differentiation index (F_{ST}) between the 11 putative populations after 10 000 years of divergence was low, 0.014 (95% confidence interval (CI): 0.0010–0.044) and 0.013 (95% CI: 0.0012–0.042) for dispersal scenario 1 (strong dispersal) and scenario 2 (strong sedentary behavior), respectively (solid lines, Figs. 4*a* and 4*b*). For both scenarios, simulated F_{ST} trajectories varied strongly among replicates leading to a standard deviation between replicates of 0.12 in the final year. Note that even after 10 000 years of simulation, no equilibrium was reached.

Plotting mean pairwise F_{ST} values after 10 000 years against the mean number of migrants revealed that F_{ST} values were >0.001 for small number of migrants (<5 individuals) between population pairs (Fig. 5*a*). The threshold value 0.001 was therefore used for identifying genetically disconnected populations. Note that small pairwise F_{ST} values occurred even though there were very few pairwise migrants (pairs in bottom left corner of Fig. 5*a* with <1 migrant). This is due to the existence of multiple connections for certain populations.

Fig. 4. Density of simulated global genetic differentiation F_{ST} for 200 replicate trajectories for dispersal scenarios 1 (*a*) and 2 (*b*). Black solid line: median; black dotted line: mean; grey solid line: maximum. Index of pairwise genetic differentiation F_{ST} between all pairs of populations at the end of the simulations (10 000 years) averaged for 200 replicates for dispersal scenarios 1 (*c*) and 2 (*d*). For population abbreviations, see Fig. 1*b*.

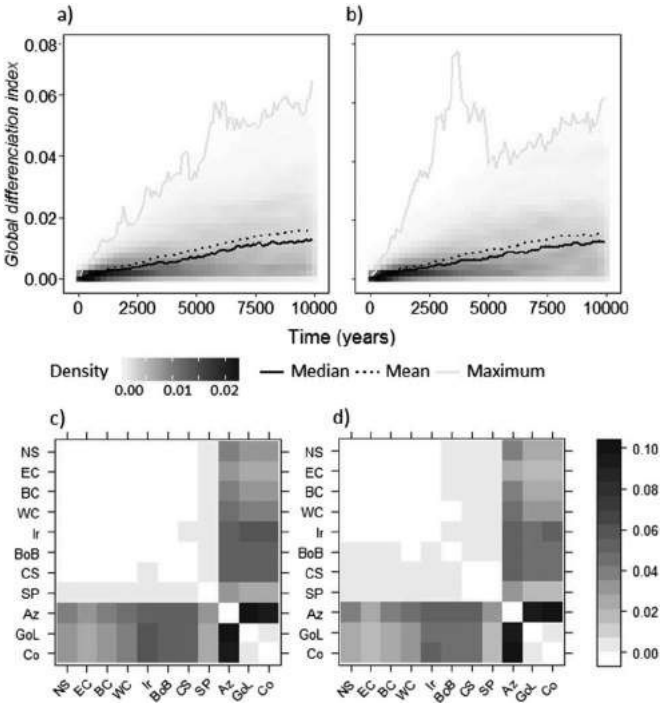
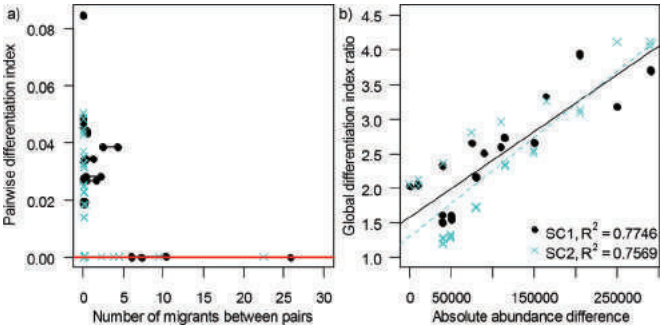


Fig. 5. (*a*) Mean pairwise F_{ST} of simulations plotted against the mean number of annual migrants from population *i* to population *j* and vice versa and (*b*) ratio between pairwise F_{ST} of simulations with identical abundance for all populations and pairwise F_{ST} > 0.001 of simulations with different abundances as a function of the pairwise difference of these abundances. In panel (*a*), the continuous horizontal line indicates the threshold value of 0.001 used in this study. Note that the *x* axis in panel (*a*) was cut at 30 migrants; all F_{ST} values beyond this value are below the threshold value. All population pairs have a single simulated F_{ST} value but two sets of number of migrants (one in each direction), which are linked by lines. SC; dispersal scenario.



For both dispersal scenarios, analysis of pairwise genetic differentiation between putative populations revealed high genetic differentiation between the Azores (Az) and the other populations (F_{ST} > 0.03 after 10 000 years; Figs. 4*c* and 4*d*). Mediterranean populations (GoL and Co) were also differentiated from the other

populations (mean F_{ST} > 0.035). This was not surprising given the assumed geographic isolation of both the Az and Mediterranean populations. Mediterranean populations were more differentiated (F_{ST} ≈ 0.1) from the Az population than from Atlantic populations mainly due to lower assumed dispersal rates but also due to smaller population sizes (Figs. 4*c* and 4*d*). The simulations with identical populations sizes (10 000 individuals) confirmed that in this case genetic differentiation patterns were driven by the assumed dispersal patterns, with Mediterranean populations being again more differentiated from Az compared with the Atlantic populations (not shown).

For dispersal scenario 1, after 10 000 years, Atlantic populations (except Az) were not genetically differentiated from each other, i.e., pairwise F_{ST} values were ≪0.001 (Fig. 4*c*; Table 2). Only South Portugal (SP) presented a weak genetic differentiation from more northern populations (F_{ST} = 0.0005 compared with 0.00002 between the northern populations). The two Mediterranean populations formed two units with low genetic heterogeneity (F_{ST} = 0.002).

For dispersal scenario 2, Atlantic populations (except Az) were more structured (globally higher F_{ST} value than for scenario 1) but not genetically differentiated from each other applying the threshold value of 0.001 (Fig. 4*d*; Table 2). Using a lower threshold of F_{ST} > 0.0001 for determining populations, Atlantic populations appeared more structured with two groups (Supplementary Table S5[1]). In this case the northern group was composed of the North Sea (NS), the EC, the Bristol Channel (BC), and the Ireland (Ir) populations. The southern group was composed of the Cantabrian Sea (CS) and SP. Finally, the Bay of Biscay (BoB) population was differentiated with all other populations except the Western Channel (WC) population.

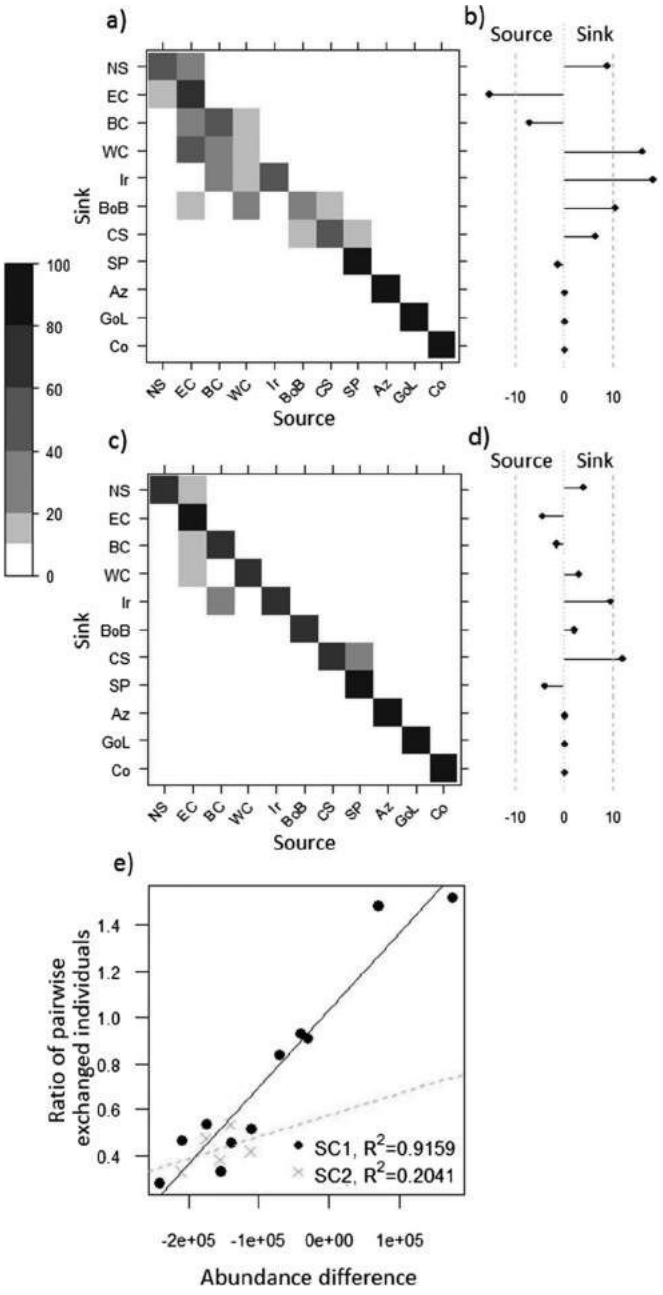
The two patterns of genetic differentiation between the 11 putative populations primarily but not exclusively reflected dispersal patterns, as relative differences in population sizes also played a role. For example, in scenario 1, differentiation between Az and EC was weaker than differentiation between Az and BoB despite an identical dispersal rate. To visualize this abundance effect, the ratio of F_{ST} values obtained assuming identical population abundances and population-specific abundances (N_0) was regressed against the absolute value of the difference in population-specific abundances. For both dispersal scenarios, this relationship was linear and rather similar (Fig. 5*b*). Thus, the difference found in the pairwise F_{ST} of the Az population with EC compared with BoB was due to the difference in abundance (Figs. 4*c* and 4*d*).

Demographic connectivity

The two dispersal scenarios led to qualitatively similar results in terms of global demographic connectivity, i.e., annual net contributions of dispersing individuals to the different putative populations (Figs. 6*b* and 6*d* and Supplementary Tables S6 and S7[1]). The biggest source population was the EC, which had an annual net loss of –15% and –4% for dispersal scenario 1 (strong dispersal) and scenario 2 (strong sedentary behavior), respectively. The annual net losses for the BC population were –7% and –2% for dispersal scenarios 1 and 2, respectively. For SP, these were –1% and –4%, respectively. Az, East Corsica (Co), and Gulf of Lion (GoL) populations had minor exchanges with the other populations, while all other populations experienced net gains for both dispersal scenarios. More precisely, the populations near Ir and in WC (including Southern Celtic Sea) gained most under dispersal scenario 1 (+18% and +16%), while for dispersal scenario 2 the biggest net gain was found for the CS population (+12%), followed by the Ir population (+10%).

Considering only population pairs linked by pairwise demographic connectivity of more than the 10% threshold value, 13 pairs of demographically closely linked populations emerged for dispersal scenario 1 (Fig. 6*a*). Three pairs exchanged reciprocally more than 10% of individuals: (*i*) the EC and the NS, (*ii*) the EC and the BC, and (*iii*) the BoB with the CS. For the two first pairs, if we look at net exchanges, the EC appeared to be a source while the NS and the BC were sinks. For the seven other pairs, only one of the two received more than 10% of individuals from the other (Fig. 6*a*): BC and EC, WC and EC, Ir and BC, Ir and WC, BoB and EC, BoB and WC, and CS and SP (the first populations is the sink and the second the source). In summary, for scenario 1 the English Channel was

Fig. 6. Pairwise percentage of individuals exchanged by putative population *i* (rows) with population *j* (columns), i.e., pairwise demographic connectivity for dispersal scenarios 1 (*a*) and 2 (*c*). Overall net percentage of individuals lost (negative value) or gained (positive value) by each putative population are also shown for dispersal scenarios 1 (*b*) and 2 (*d*). For geographic location of populations, see Fig. 1*b*. (*e*) Ratio between pairwise demographic connectivity of population with identical abundance and pairwise demographic connectivity of population with different abundances as a function of the pairwise difference of these abundances. SC: dispersal scenario.



the main source of individuals for the other putative Northeast Atlantic populations.

For dispersal scenario 2, only five pairs of demographically closely linked populations (>10%) emerged (Fig. 6*b*): NS and EC, BC and EC, WC and EC, Ir and BC, and SC and SP (the first population is the sink and the second the source). None of the links were reciprocal. Overall in this scenario, populations appeared to be less connected, and the main sources appeared to be SP and again EC, which were also populations with the highest assumed abundances.

Based on these results, a number of groups of linked populations emerged where each group could be considered to represent a single metapopulation (Table 2). For dispersal scenario 1, four metapopulations could be identified: the first one comprising the Northeast Atlantic populations from the NS to the SP, the second one being constituted by the Az, and the last two by the two Mediterranean populations (GoL and Co). Focusing only on reciprocal links (>10% for both receiving and donor population), the four previous metapopulations subdivided into eight: NS together with EC; BC together with WC, Ir, and BoB; CS together with S, Az, and GoL; and finally Co. In contrast, for scenario 2, six metapopulations emerged (Fig. 6*c*). The first one grouping populations from the NS to Ir, the second one including only BoB, the third one grouping CS and SP, while the fourth to the sixth metapopulations were similar to those found for scenario 1 (Az, GoL, Co). Again considering only pairs with reciprocal links, all 11 putative populations were demographically independent and hence should be managed separately under this dispersal scenario.

Abundance differences also played a role for demographic connectivity (Fig. 6*e*). To visualize this abundance effect, the ratio of $P_{i|j}$ values obtained assuming identical population abundances, and population-specific abundances (N_0) were regressed against the difference in population-specific abundances (not absolute difference as for genetic connectivity; Fig. 6*e*). For both dispersal scenarios, this relationship was linear and positive, though rather weak for scenario 2 in contrast with genetic connectivity for which no difference between scenarios was found (Fig. 5*b*).

Discussion

Population growth

Analyzing the elasticity and sensitivity of the intrinsic population growth rate in the Usher matrix model for a thornback ray-like species, we found that it was most sensitive to survival from egg laying to stage 1 (S_0), followed by the survival of stages 1 and 2 (S_1 and S_2), and least sensitive to the maturation rate (H_3), that is to the proportion of mature individuals moving to stage 4 with higher fecundity. This result likely applies to other ray and skate species that share the late age-at-maturity, and hence probably similar survival rates, and low fecundity as for thornback ray (Supplementary Table S1[1]).

Given survival during the early stages is such an important parameter for population dynamics, it would be important to obtain field estimates for each population. Unfortunately, we do not know of any in situ method for estimating first year survival. For older individuals, capture–mark–recapture methods might be feasible (e.g., Neat et al. 2015).

Population connectivity

Genetic simulation results indicated a strong influence of assumed dispersal rates on genetic population structure, though population abundance differences were found to substantially modify the effect of contrasting dispersal rates (Fig. 5). This result is not surprising, as the level of genetic differentiation among populations is directly related to the balance between gene flow (related to dispersal) and genetic drift (somewhat related to population abundance) (Wright 1949; Palumbi 2003; Waples and Gaggiotti 2006). Genetic connectivity simulations indicated four metapopulations (Table 2): the first along the Atlantic continental shelf, the second around the Azores (whose structure was not studied here), and the third and fourth in the Gulf of Lion and around Corsica, which globally agrees with the empirical genetic result found by Chevolot et al. (2006). However, it is important to remember that genetic connectivity studies provide information on an evolutionary time scale. A low genetic differentiation does not necessarily imply a contemporary high number of migrants because the time needed for genetic differentiation depends upon the number of breeders in each population. In large populations with low genetic drift, very low migration rates would maintain genetic similarity between populations that became physically separated (Reiss et al 2009).

Genetic connectivity has been studied for several ray and skate species, highlighting substantial differentiations among the studied locations (Chevolot et al. 2006; Frodella et al. 2016; Vargas-Caro et al. 2017). For example, Pasolini et al. (2011) found a genetic population structure in the eastern Atlantic for thornback ray and biscuit skate (*Raja straeleni*) and a significant correlation between genetic differentiation and coastal distance. However, for several species, including thornback ray, a major limitation to dispersal appeared to be bathymetry (Chevolot et al. 2006; Pasolini et al. 2011; Le Port and Lavery 2012).

Pairwise demographic connectivity indicated four (dispersal scenario 1) or six (dispersal scenario 2) metapopulations compared with only four for genetic connectivity (Table 2). Demographic-based metapopulations were defined as groups of putative populations for which, for sink populations, a single dispersal event resulted in at least 10% of individuals coming from another of the subpopulations (source population) within the metapopulation. The 10% threshold value is rather arbitrary. It was inspired by the value used for judging the importance of dispersal rates (Waples and Gaggiotti 2006; Palsbøll et al. 2007). Doubling the threshold value to 20% would increase the number of demographic metapopulations to six (scenario 1) and nine (scenario 2), respectively.

A strong effect of the difference in population abundances on genetic and demographic connectivity was found for dispersal scenario 1, but only for genetic connectivity for scenario 2. In scenario 2, most individuals stayed at their population of origin, which meant differences in population abundance played less a role for a single dispersal event considered for demographic connectivity.

The ICES currently considers six thornback ray stocks in the Northeast Atlantic, based upon ICES ecoregions, discontinuities in the species geographical distribution, and expert knowledge (Table 2). The connectivity-based metapopulation results disagreed with the current stock assessment units in several ways: the Bay of Biscay and Cantabrian Sea ICES stock was subdivided based on demographic results for dispersal scenario 2, while the Bristol Channel and Northern Celtic Sea stock as well as the Western Channel and Southern Celtic Sea stock were grouped with the Eastern Channel and the North Sea stocks based on genetic and demographic connectivity (both dispersal scenarios; Table 2). However, field studies are needed to confirm the identified metapopulation structure before any recommendations for changing stock assessment units can be made. As current management units for skates and rays are not aligned to stock assessment units, a complete revision of the management of skate and ray fisheries in European waters would be needed in a second step.

The results concerning potential metapopulation structures differed somewhat between demographic and genetic connectivity, with more populations being genetically connected than demographically. This phenomenon was empirically observed in shark and named crinkled connectivity (Ovenden 2013). It occurs when migration is above the threshold required to link populations genetically, but below the threshold for demographic links (Ovenden 2013). This difference between genetic and demographic connectivity is not surprising, as they provide informa-

tion on different temporal and geographical scales. Demographic data help define management units, provided that a specific threshold value for the exchange of individuals can be defined. Genetic studies provide large-scale differentiation information integrated over a longer time period, making results at a local scale and for shorter time periods less pertinent for management.

Connectivity is driven by the dispersal of individuals. Traditional stock assessment and fisheries management generally consider stocks as closed populations. However, if exploited stocks are not closed populations, the contributions of dispersals need to be considered in management (Frisk et al. 2014). This can be especially important for skates and rays where only juveniles and adults move. For example, winter skate (*Leucoraja ocellata*) abundance increased strongly on Georges Bank in the 1980s, which appeared biologically unrealistic. Frisk et al. (2008) suggested that this increase was due to movements among adjacent populations, thus connectivity.

Dispersal increases the number of individuals of the receiving population and thus is equivalent to an increase in survival. Assuming dispersal occurs only during the first year of life, we can express the calculated net contributions of immigrants to the 11 putative populations in terms of a change in first year survival rate. Assuming $S_0 = 0.036$ (baseline value in Table 1), an increase of 18% (maximum estimated net gain due to demographic connectivity) is equivalent to increasing the survival rate to $S_0 = 0.042$, while a 15% decrease (maximum estimated net loss) corresponds to $S_0 = 0.031$. This range of first year survival rates is small compared with the uncertainty surrounding realistic values. An increased survival rate of $S_0 = 0.042$ lead to an increase of the intrinsic growth rate of 2.26%, and a decreased survival rate of $S_0 = 0.031$ lead to a decrease of the intrinsic growth rate of 2.11%.

Methodology

Several assumptions were needed to overcome the lack of data and biological knowledge. In the Usher model, first year survival (S_0) was set conditional on other parameter values and assuming a stable population. The resulting survival rate might appear high compared with teleost species (0.036 in Table 1). This parameter combines the mortality of eggs and newly hatched individuals. Skates and rays do not provide any parental care, making the rigid keratin capsule the only protection against predators (Kormanik 1993), and so egg mortality is primarily caused by predation (Bunn et al. 2000; Cox et Koob 1993). Lucifora and Garcia (2004) reported gastropod predation rates around 0.24 for four ray species in the Southwest Atlantic. The sensitivity and elasticity analyses showed that this parameter was influential for the population dynamics of a thornback ray-like species; a similar importance can be assumed for other skates and rays in European waters, which share similar life history traits (Supplementary Table S1[1]).

Several assumptions were also necessary to study population connectivity. Absolute abundance estimates currently do not exist for any of the ray and skate populations in European waters (ICES 2016). Relative population abundances were derived by assuming that commercial landings represented relative population abundances. In quota-based fisheries management systems, landings do not directly inform on population abundance. However, under the current management in Europe, the quota is set for a pool of several rays and skates and there is no regulatory minimum landing size at species level, and this in turn makes quotas less restrictive at species-specific level. Thus, in this case, commercial landings might approximately reflect relative population biomass and abundance if mean mass is similar across populations. Changing the relative abundances of the 11 putative populations would modify estimated demographic and genetic connectivity patterns and absolute values; the degree of this would depend on by how much proportions were changed.

Population abundances were assumed fixed but are known to vary over time. For example, bottom trawl surveys have shown abundance variations in recent years, including a dramatic increase of thornback ray in the Eastern Channel over the last decade (ICES 2016). In contrast, a recent study of thornback ray population dynamics in the Bay of Biscay suggested no increase in recent years (Marandel et al. 2016). At an evolutionary time scale, over the 10 000 years simulated for genetic connectivity, sea level and temperature increases probably triggered changes in population distributions and abundances. Thus, the constant dispersal rates and population abundances that we used should not be taken as realistic either in genetic or demographic terms. The results should rather be regarded as indicative of the potential magnitude of the contrast between the genetic and demographic connectivity. However, if times series of abundance estimates were available for all putative populations, these might be used directly for studying demographic connectivity by analysing synchronism in interannual abundance variations (Östman et al. 2017).

In genetic simulations due to computational limits, simulated population abundances were much smaller than likely actual abundances. However, this did not affect the spatial pattern of genetic differentiation, only the absolute F_{ST} values, which might be higher in our simulations than in actual populations, as in reality their larger number is expected to reduce genetic drift.

In the absence of dispersal rate estimates for all putative populations, two scenarios, expected to reflect a plausible range of dispersal for the species, were investigated. The first one corresponded to a strong potential for dispersal and the second one to a stronger sedentary behavior. Simulating more extreme scenarios would basically reflect that with high dispersal rate panmixia is maintained, and with low dispersal rate, all populations differentiate. To carry out more realistic simulations, it would be necessary to estimate dispersal of all populations in the field. Several methods are available for this. Telemetry can provide estimates of individual movements between populations and of behavior patterns (Milner-Gulland and Rowcliffe 2007; Hawkins et al. 2016). For conventional tagging, due to often low reported recapture rate (6.9%; Stephan et al. 2015), a high number of individuals has to be tagged and released to be able to estimate dispersal rate with reasonable uncertainty. Electronic data storage tags allows for obtaining information from a higher proportion of tagged individuals (Hunter et al. 2006); however, cost may be prohibitive. This method was used for tracking basking sharks (*Cetorhinus maximus*; Sims et al. 2003, 2005) and studying thornback rays stock distribution in the southern North Sea (Hunter et al. 2006).

For genetic connectivity, dispersal was assumed to occur for newborns only. This is equivalent to assuming that each individual reproduces in one population only during its life, either in its native population or in the one it dispersed to before maturing. Thus, the assumed dispersal rates should be considered as the contribution of dispersal to each population and not as the individual contribution. If the assumed dispersal rates were applied to all ages, genetic connectivity could be modified, as the age of the disperser will affect the number of years during which it will reproduce in the receiving population and so will differently affect the gene pool.

Genetic connectivity was modelled by simulating a neutral marker. Neutral loci have been recommended for identifying management units (Funk et al. 2012). However, in the case of recently differentiated populations with large population size, markers under selection may be more efficient (Reiss et al. 2009; Gagnaire et al. 2015). Using non-neutral markers in simulations would require assumptions on the dynamics of selection. Local adaption might mean that immigrants have lower fitness, which would reduce their contribution to genetic connectivity, as their genotypes would be less integrated into the local gene pool. As a consequence, their contribution to population dynamics could be

modified and demographic connectivity reduced. Therefore, if adaptation to regional environment has occurred, both genetic and demographic connectivity could be lower than estimated in this study.

Demographic connectivity was studied applying a single dispersal event to the assumed population abundances. Given these abundances were derived from recent landings, their relative proportions might be expected to reflect the contemporary situation, in which case applying a single dispersal event would be informative for current management. Contrary to species for which dispersal primarily occurs during the larval phase and for which ocean current models can be used to predict connectivity, there is no reason to believe that thornback ray, or any other ray or skate for that matter, would follow bottom currents. Thus, to develop a more complex dynamic model, knowledge on factors determining individual dispersal, survival, and reproduction would be needed for all putative populations. None of these are currently available.

Conclusion

Demographic and genetic connectivity can provide complementary insights for medium-sized rays and skates as we demonstrated with the example of thornback ray. Genetic connectivity studies should be useful for determining long-term conservation units but will probably not be so helpful for defining management units, while demographic connectivity studies should be able to inform the definition of management units. Both types of analyses strongly depend on relative population abundances (Figs. 5 and 6*e*) but also on dispersal rates and patterns. To make progress towards a better estimation of connectivity and delimitation of management and conservation units, we encourage researchers to attempt obtaining local abundance and dispersal estimates for a range of ray and skate species.

Acknowledgements

This study received funding from the French "Agence Nationale de la Recherche" (ANR) and from the Fondation Total for the GenoPopTaille project. The authors thank David Causeur, Jean-Baptiste Lamy, and Frederique Viard for stimulating discussions on a preliminary version of this paper. FM thanks Ifremer for a Ph.D. studentship. We thank an anonymous reviewer and the associate editor for constructive comments that helped to improve the manuscript.

References

Andrello, M., and Manel, S. 2015. MetaPopGen: an r package to simulate population genetics in large size metapopulations. Mol. Ecol. Resour. **15**(5): 1153–1162. doi:10.1111/1755-0998.12371.

Bonanomi, S., Overgaard Therkildsen, N., Retzel, A., Berg Hedeholm, R., Pedersen, M.W., Meldrup, D., Pampoulie, C., Hemmer-Hansen, J., Grønkjaer, P., and Nielsen, E.E. 2016. Historical DNA documents long-distance natal homing in marine fish. Mol. Ecol. **25**(12): 2727–2734. doi:10.1111/mec.13580. PMID:26859133.

Bunn, N.A., Fox, C.J., and Webb, T. 2000. A literature review of studies on fish egg mortality: implications for the estimation of spawning stock biomass by the annual egg production method. Technical Report, Cefas Science Series.

Campolongo, F., Cariboni, J., and Saltelli, A. 2007. An effective screening design for sensitivity analysis of large models. Environ. Model. Softw. **22**(10): 1509–1518. doi:10.1016/j.envsoft.2006.10.004.

Caswell, H. 2001. Matrix population models: construction, analysis, and interpretation. 2nd ed. Sinauer Associates, Sunderland, Mass.

Chevolot, M. 2006. Assessing genetic structure of thornback ray, *Raja clavata*: a thorny situation? Thesis, University of Groningen.

Chevolot, M., Hoarau, G., Rijnsdorp, A.D., Stam, W.T., and Olsen, J.L. 2006. Phylogeography and population structure of thornback rays (*Raja clavata* L., Rajidae). Mol. Ecol. **15**(12): 3693–3705. doi:10.1111/j.1365-294X.2006.03043.x. PMID:17032267.

Chevolot, M., Wolfs, P.H.J., Pálsson, J., Rijnsdorp, A.D., Stam, W.T., and Olsen, J.L. 2007. Population structure and historical demography of the thorny skate (*Amblyraja radiata*, Rajidae) in the North Atlantic. Mar. Biol. **151**(4): 1275–1286. doi:10.1007/s00227-006-0556-1.

Cox, D.L., and Koob, T.J. 1993. Predation on elasmobranch eggs. Environ. Biol. Fishes, **38**(1–3): 117–125. doi:10.1007/BF00842908.

Davidson, L.N.K., Krawchuk, M.A., and Dulvy, N.K. 2016. Why have global shark and ray landings declined: improved management or overfishing? Fish Fish. **17**(2): 438–458. doi:10.1111/faf.12119.

Delpiani, G. 2016. Reproductive biology of the southern thorny skate *Amblyraja doellojuradoi* (Chondrichthyes, Rajidae): reproduction of *Amblyraja doellojuradoi*. J. Fish Biol. **88**(4): 1413–1429. doi:10.1111/jfb.12917. PMID:26923668.

Dingle, H. 2014. Migration: the biology of life on the move. 2nd ed. Oxford University Press, Oxford, New York.

Dittman, A., and Quinn, T. 1996. Homing in Pacific Salmon: mechanisms and ecological basis. J. Exp. Biol. **199**: 83–91. PMID:9317381.

Dulvy, N.K., Fowler, S.L., Musick, J.A., Cavanagh, R.D., Kyne, P.M., Harrison, L.R., Carlson, J.K., Davidson, L.N., Fordham, S.V., Francis, M.P., Pollock, C.M., Simpfendorfer, C.A., Burgess, G.H., Carpenter, K.E., Compagno, L.J., Ebert, D.A., Gibson, C., Heupel, M.R., Livingstone, S.R., Sanciangco, J.C., Stevens, J.D., Valenti, S., and White, W.T. 2014. Extinction risk and conservation of the world's sharks and rays. eLife, **3**: e00590. doi:10.7554/eLife.00590. PMID:24448405.

Ellis, J.R., and Shackley, S.E. 1995. Observations on egg-laying in the thornback ray. J. Fish Biol. **46**(5): 903–904. doi:10.1111/j.1095-8649.1995.tb01613.x.

FAO. 2016. Fisheries and Aquaculture Department, Statistics and Information Service. FishStat: Universal software for fishery statistical time series [online]. Available from http://www.fao.org/fishery/statistics/en. FAO, Rome.

Feldheim, K.A., Gruber, S.H., DiBattista, J.D., Babcock, E.A., Kessel, S.T., Hendry, A.P., Pikitch, E.K., Ashley, M.V., and Chapman, D.D. 2014. Two decades of genetic profiling yields first evidence of natal philopatry and long-term fidelity to parturition sites in sharks. Mol. Ecol. **23**(1): 110–117. doi:10.1111/mec.12583. PMID:24192204.

Frisk, M.G., Miller, T.J., Martell, S.J.D., and Sosebee, K. 2008. New hypothesis helps explain elasmobranch "outburst" on Georges Bank in the 1980s. Ecol. Appl. **18**(1): 234–245. doi:10.1890/06-1392.1. PMID:18372569.

Frisk, M.G., Jordaan, A., and Miller, T.J. 2014. Moving beyond the current paradigm in marine population connectivity: are adults the missing link? Fish Fish. **15**(2): 242–254. doi:10.1111/faf.12014.

Frodella, N., Cannas, R., Velonà, A., Carbonara, P., Farrell, E., Fiorentino, F., Follesa, M., Garofalo, G., Hemida, F., Mancusi, C., Stagioni, M., Ungaro, N., Serena, F., Tinti, F., and Cariani, A. 2016. Population connectivity and phylogeography of the Mediterranean endemic skate *Raja polystigma* and evidence of its hybridization with the parapatric sibling *R. montagui*. Mar. Ecol. Progr. Ser. **554**: 99–113. doi:10.3354/meps11799.

Funk, W.C., McKay, J.K., Hohenlohe, P.A., and Allendorf, F.W. 2012. Harnessing genomics for delineating conservation units. Trends Ecol. Evol. **27**(9): 489–496. doi:10.1016/j.tree.2012.05.012. PMID:22727017.

Gagnaire, P.-A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., Arnaud-Haond, S., and Bierne, N. 2015. Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. Evol. Appl. **8**(8): 769–786. doi:10.1111/eva.12288. PMID:26366195.

Gislason, H., Daan, N., Rice, J.C., and Pope, J.G. 2010. Size, growth, temperature and the natural mortality of marine fish: natural mortality and size. Fish Fish. **11**(2): 149–158. doi:10.1111/j.1467-2979.2009.00350.x.

Hallerman, E.M. (*Editor*). 2003. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Md.

Hamilton, M.B. 2009. Population genetics. Wiley-Blackwell, Chichester, UK; Hoboken, N.J.

Hastings, A. 1993. Complex interactions between dispersal and dynamics — lessons from coupled logistic equations. Ecology, **74**: 1362–1372. doi:10.2307/1940066.

Hawkins, S.J., Bohn, K., Sims, D.W., Ribeiro, P., Faria, J., Presa, P., Pita, A., Martins, G.M., Neto, A.I., Burrows, M.T., and Genner, M.J. 2016. Fisheries stocks from an ecological perspective: disentangling ecological connectivity from genetic interchange. Fish. Res. **179**: 333–341. doi:10.1016/j.fishres.2016.01.015.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature, **405**: 907–913. doi:10.1038/35016000. PMID:10879524.

Hoening, J.M., and Gruber, S.H. 1990. Life-history patterns in the elasmobranchs: implications for fisheries management. NOAA Technical Report.

Holden, M.J., Rout, D.W., and Humphreys, C.N. 1971. The rate of egg laying by three species of ray. ICES J. Mar. Sci. **33**: 335–339. doi:10.1093/icesjms/33.3.335.

Hunter, E., Buckley, A.A., Stewart, C., and Metcalfe, J.D. 2005*a*. Migratory behavior of the thornback ray, *Raja clavata*, in the southern North Sea. J. Mar. Biol. Assoc. U.K. **85**(5): 1095. doi:10.1017/S0025315405012142.

Hunter, E., Buckley, A.A., Stewart, C., and Metcalfe, J.D. 2005*b*. Repeated seasonal migration by a thornback ray in the southern North Sea. J. Mar. Biol. Assoc. U.K. **85**(5): 1199. doi:10.1017/S0025315405012300.

Hunter, E., Berry, F., Buckley, A.A., Stewart, C., and Metcalfe, J.D. 2006. Seasonal migration of thornback rays and implications for closure management: ray migration and closure management. J. Appl. Ecol. **43**(4): 710–720. doi:10.1111/j.1365-2664.2006.01194.x.

ICES. 2016. Report of the Working Group on Elasmobranchs Fishes. ICES, Lisbon, Portugal.

Kormanik, G.A. 1993. Ionic and osmotic environment of developing elasmobranch embryos. Environ. Biol. Fishes, **38**(1–3): 233–240. doi:10.1007/BF00842919.

Le Port, A., and Lavery, S. 2012. Population Structure and Phylogeography of the

Short-Tailed Stingray, *Dasyatis brevicaudata* (Hutton 1875), in the Southern Hemisphere. J. Hered. **103**(2): 174–185. doi:10.1093/jhered/esr131. PMID: 22174443.

Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bull. Entomol. Soc. Am. **71**: 237–240.

Lowe, W.H., and Allendorf, F.W. 2010. What can genetics tell us about population connectivity?: Genetic and demographic connectivity. Mol. Ecol. **19**(15): 3038–3051. doi:10.1111/j.1365-294X.2010.04688.x. PMID:20618697.

Lucifora, L.O., and Garcia, V.B. 2004. Gastropod predation on egg cases of skates (Chondrichthyes, Rajidae) in the southwestern Atlantic: quantification and life history implications. Mar. Biol. **145**(5): 917–922. doi:10.1007/s00227-004-1377-8.

Marandel, F., Lorance, P., and Trenkel, V.M. 2016. A Bayesian state-space model to estimate population biomass with catch and limited survey data: application to the thornback ray (*Raja clavata*) in the Bay of Biscay. Aquat. Living Resour. **29**(2): 209. doi:10.1051/alr/2016020.

Milner-Gulland, E.J., and Rowcliffe, J.M. 2007. Conservation and sustainable use: a handbook of techniques. Oxford University Press, Oxford.

Morris, M.D. 1991. Factorial sampling plans for preliminary computational experiments. Technometrics, **33**(2): 161–174. doi:10.1080/00401706.1991.10484804.

Neat, F., Pinto, C., Burrett, I., Cowie, L., Travis, J., Thorburn, J., Gibb, F., and Wright, P.J. 2015. Site fidelity, survival and conservation options for the threatened flapper skate (*Dipturus* cf. *intermedia*). Aquat. Conserv. Mar. Freshw. Ecosyst. **25**(1): 6–20. doi:10.1002/aqc.2472.

Östman, Ö., Olsson, J., Dannewitz, J., Palm, S., and Florin, A.B. 2017. Inferring spatial structure from population genetics and spatial synchrony in demography of Baltic Sea fishes: implications for management. Fish Fish. **18**: 324–339. doi:10.1111/faf.12182.

Ovenden, J.R. 2013. Crinkles in connectivity: combining genetics and other types of biological data to estimate movement and interbreeding between populations. Marine Freshw. Res. **64**: 201–207. doi:10.1071/MF12314.

Palsbøll, P., Berube, M., and Allendorf, F. 2007. Identification of management units using population genetic data. Trends Ecol. Evol. **22**(1): 11–16. doi:10.1016/j.tree.2006.09.003. PMID:16982114.

Palumbi, S.R. 2003. Populations genetics, demographic connectivity, and the design of marine reserves. Ecol. Appl. **13**(sp1): 146–158. doi:10.1890/1051-0761(2003)013[0146:PGDCAT]2.0.CO;2.

Pasolini, P., Ragazzini, C., Zaccaro, Z., Cariani, A., Ferrara, G., Gonzalez, E.G., Landi, M., Milano, I., Stagioni, M., Guarniero, I., and Tinti, F. 2011. Quaternary geographical sibling speciation and population structuring in the Eastern Atlantic skates (suborder Rajoidea) *Raja clavata* and *R. straeleni*. Mar. Biol. **158**(10): 2173–2186. doi:10.1007/s00227-011-1722-7.

Pujol, G. 2009. Simplex-based screening designs for estimating metamodels. Reliabil. Eng. Syst. Saf. **94**(7): 1156–1160. doi:10.1016/j.ress.2008.08.002.

Pujol, G., Iooss, B., Janon, A. with contributions from Lemaitre, P., Gilquin, L., Le Gratiet, L., Touati, T., Ramos, B., Fruth, J., and Da Veiga, S. 2014. Sensitivity: sensitivity analysis. R package version 1.10.1 [online]. Available from http://CRAN.R-project.org/package=sensitivity.

Quéro, J.C., and Cendrero, O. 1996. Incidence de la pêche sur la biodiversité ichtyologique marine: le Bassin d'Arcachon et le plateau continental Sud Gascogne. Cybium, **20**(4): 323–356.

Quéro, J.-C., and Vayne, J.-J. (*Editors*). 1997. Les poissons de mer des pêches françaises : identification, inventaire et répartition de 209 espèces. Delachaux et Niestlé, Lausanne.

R Development Core Team. 2008. R: a language and environment for statistical computing [online]. R Foundation for Statistical Computing, Austria, Vienna. Available from http://www.R-project.org.

Reiss, H., Hoarau, G., Dickey-Collas, M., and Wolff, W.J. 2009. Genetic population structure of marine fish: mismatch between biological and fisheries management units. Fish Fish. **10**(4): 361–395. doi:10.1111/j.1467-2979.2008.00324.x.

Schnute, J.T., Boers, N., and Haigh, B. 2017. PBSmapping: mapping fisheries data and spatial analysis tools. R package version 2.70.4 [online]. Available from https://CRAN.R-project.org/package=PBSmapping.

Schwartz, M., Luikart, G., and Waples, R. 2007. Genetic monitoring as a promising tool for conservation and management. Trends Ecol. Evol. **22**(1): 25–33. doi:10.1016/j.tree.2006.08.009. PMID:16962204.

Serra-Pereira, B., Figueiredo, I., Farias, I., Moura, T., and Gordo, L.S. 2008. Description of dermal denticles from the caudal region of *Raja clavata* and their use for the estimation of age and growth. ICES J. Mar. Sci. **65**(9): 1701–1709. doi:10.1093/icesjms/fsn167.

Sims, D., Southall, E., Richardson, A., Reid, P., and Metcalfe, J. 2003. Seasonal movements and behaviour of basking sharks from archival tagging: no evidence of winter hibernation. Mar. Ecol. Progr. Ser. **248**: 187–196. doi:10.3354/meps248187.

Sims, D., Southall, E., Metcalfe, J., and Pawson, M. 2005. Basking shark population assessment. Final project report. Global Wildlife Division. Department for Environment, Food and Rural Affairs, London.

Sinclair, A.R.E., Fryxell, J.M., Caughley, G., and Caughley, G. 2006. Wildlife ecology, conservation, and management. 2nd ed. Blackwell Publishing, Malden, Mass.; Oxford.

Stearns, S.C., and Hoekstra, R.F. 2005. Evolution: an introduction. 2nd ed. Oxford University Press, Oxford [England]; New York.

Stephan, E., Gadenne, H., Meheust, E., and Jung, L. 2015. Projet RECOAM : étude de cinq espèces de raies présentes dans les eaux côtières d'Atlantique et de Manche. Rapport final. Association Pour l'Etude et la Conservation des Sélaciens et Laboratoire BioGeMME, Brest, France.

Steven, G.A. 1933. Rays and skates of Devon and Cornwall. III. The proportions of the sexes in nature and in commercial landings and their significance to the fishery. J. Mar. Biol. Assoc. U.K. **18**: 611–625. doi:10.1017/S0025315400043939.

Usher, M.B. 1966. A matrix approach to the management of renewable resources, with special reference to selection forests. J. Appl. Ecol. **3**: 355–367. doi:10.2307/2401258.

Vargas-Caro, C., Bustamante, C., Bennett, M.B., and Ovenden, J.R. 2017. Towards sustainable fishery management for skates in South America: The genetic population structure of *Zearaja chilensis* and *Dipturus trachyderma* (Chondrichthyes, Rajiformes) in the south-east Pacific Ocean. PLoS ONE, **12**(2): e0172255. doi:10.1371/journal.pone.0172255. PMID:28207832.

Walker, P., Howlett, G., and Millner, R. 1997. Distribution, movement and stock structure of three ray species in the North Sea and eastern English Channel. ICES J. Mar. Sci. **54**(5): 797–808. doi:10.1006/jmsc.1997.0223.

Waples, R.S., and Gaggiotti, O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol. Ecol. **15**(6): 1419–1439. doi:10.1111/j.1365-294X.2006.02890.x. PMID:16629801.

Waples, R.S., Punt, A.E., and Cope, J.M. 2008. Integrating genetic data into management of marine resources: how can we do it better? Fish Fish. **9**(4): 423–449. doi:10.1111/j.1467-2979.2008.00303.x.

Wright, S. 1949. The genetical structure of populations. Ann. Eugenics, **15**(1): 323–354. doi:10.1111/j.1469-1809.1949.tb02451.x.

CHAPITRE 3

Évaluer l'état des populations de raie bouclée : une méthode démographique

Ifremer

Ce chapitre étudie par simulations et applications à des données empiriques, l'estimation de l'état des populations de raies par méthodes démographiques. Le cas d'étude choisi est la raie bouclée du Golfe de Gascogne. Cette partie a fait l'objet d'une publication dans Aquatic Living Ressources.

1. La modélisation au service de l'évaluation démographique de l'état des populations

En halieutique, mais plus généralement en écologie statistique, l'ajustement de modèles de dynamique de populations à des séries chronologiques d'observations (captures, indices d'abondance) joue un rôle central. Ces modèles permettent de réaliser des diagnostics sur l'état des populations étudiées mais aussi d'effectuer des prédictions quantifiées de leur évolution. Un modèle est par définition faux, c'est une représentation fonctionnelle de la réalité mise au point à partir de nos connaissances mais aussi d'hypothèses, bien souvent nécessaires. Ainsi aucun modèle ne peut prétendre être l'unique traduction de la réalité et chaque modèle est conçu afin de répondre à une problématique particulière. Le modèle présenté dans cette partie de la thèse est un modèle stochastique. Cette stochasticité met ainsi l'articulation entre les données observées et la modélisation développée au centre de notre problématique d'évaluation de stock à des fins de gestion et conservation. Dans cette approche, les incertitudes liées au processus d'observation et à la simplication des processus biologiques modélisés sont essentielles. Le modèle présenté a été conçu en gardant ces incertitudes à l'esprit afin d'évaluer au mieux l'état des populations d'une espèce à données limitées.

L'article suivant présente un modèle bayésien d'évaluation du stock de raie bouclée du golfe de Gascogne entre 1903 et 2014. Il constitue, à notre connaissance, la première approche bayésienne d'un modèle de biomasse sur une espèce de raie de l'Atlantique Nord-Est. Une première version du modèle utilise uniquement les **débarquements** disponibles ou reconstitués afin d'évaluer des trajectoires de biomasses ainsi que quelques indicateurs clés (comme la capacité biotique du milieu). Néanmoins, le manque de robustesses des trajectoires de biomasses estimées a conduit à un second modèle prenant en compte à la fois les débarquements et des indices d'abondances en biomasse. Ce second modèle a été testé par une approche de simulation-estimation. Sous réserve que les données (débarquements, indices de biomasse) sont suffisamment réalistes, le modèle permet d'estimer l'état actuel du stock ainsi que la **capacité biotique** du milieu. L'estimation du taux de croissance intrinsèque de la population est quant à elle très sensible à la distribution à priori de ce paramètre. Les indices d'abondances disponibles étant souvent peu fiables, un troisième modèle utilisant les débarquements disponibles ainsi qu'une estimation experte de la dépletion actuelle du stock a été mis en place. Ce dernier modèle permet de contourner le manque d'informations disponibles et permet d'obtenir une idée des trajectoires de biomasse de l'espèce ainsi que son état de dépletion actuelle, la distribution de ce paramètre étant relativement bien mise à jour.

2. Article : Un modèle Bayésien à état latent pour estimer la biomasse d'une population à l'aide de captures et de données de campagne limitées : application à la raie bouclée (*Raja clavata*) du golfe de Gascogne

Aquatic Living Resources - 11 Juin 2016
DOI :10.1051/alr/2016020

Florianne Marandel[1], Pascal Lorance [1] and Verena M. Trenkel[1]

RESUME - La raie bouclée du golfe de Gascogne est supposée en déclin depuis le début du $20^{ème}$ siècle. Afin d'évaluer ce déclin et estimer des trajectoires de biomasses, une série temporelle hypothétique de captures (de 1903 à 2013) a été élaborée. Un modèle bayésien de biomasse à état latent utilisant la fonction de production de Schaefer a été ajusté à cette série hypothétique de captures et à une série plus courte d'indices de biomasse issus de campagnes scientifiques (de 1973 à 2013, avec quelques années manquantes). Une approche de simulation-estimation a montré une forte sensibilité du modèle à la distribution *a priori* du taux de croissance intrinsèque. Le modèle fournit des trajectoires de biomasses corroborant le déclin de la population du golfe de Gascogne. La biomasse correspondant au Rendement Maximal Durable, B_{RMD}, est estimée à 32 000 tonnes soit 17 fois la biomasse estimée en 2014. Les biomasses estimées sans CPUE sont très incertaines. L'ajout d'une observation de déplétion actuelle améliore la précision, bien que les résultats y soient sensibles. Les résultats doivent être considérés avec précaution étant donné les nombreuses hypothèses nécessaires à la reconstitution de la longue série temporelle de captures ainsi qu'à l'élaboration des priors informatifs, notamment pour le taux de croissance intrinsèque. Néanmoins, les résultats obtenus confirment la déplétion de la raie bouclée du golfe de Gascogne avec une biomasse estimée en 2014 avoisinant les 3% de la capacité biotique du milieu.

[1] Ifremer, rue de l'île d'Yeu, BP 21105, 44311 Nantes Cedex 3, France

Aquat. Living Resour. 29, 209 (2016)

DOI: 10.1051/alr/2016020
www.alr-journal.org

Aquatic Living Resources

A Bayesian state-space model to estimate population biomass with catch and limited survey data: application to the thornback ray (*Raja clavata*) in the Bay of Biscay

Florianne MARANDEL[a], Pascal LORANCE and Verena M. TRENKEL

Ifremer, rue de l'île d'Yeu, BP 21105, 44311 Nantes Cedex 3, France

Received 26 November 2015; Accepted 11 June 2016

Abstract – The thornback ray (*Raja clavata*) in the Bay of Biscay is presumed to have declined during the 20th Century. To evaluate this decline and estimate biomass trajectories, a hypothetical catch time series was created for the period 1903–2013. A Bayesian state-space biomass production model with a Schaefer production function was fitted to the hypothetical catch time series and to a shorter research vessel Catch Per Unit Eeffort (CPUE) time series (1973–2013, with missing years). A censored likelihood made it possible to obtain biomass estimates without a CPUE time series or only with an estimate of biomass depletion. A simulation-estimation approach showed a high sensitivity of results to the prior for the intrinsic growth rate. The model provided biomass trajectories which corroborated and quantified the decline of the Bay of Biscay population. The estimated biomass corresponding to the maximum sustainable yield, B_{MSY}, was 32 000 tonnes, which is 17 times higher than the estimated biomass in 2014. The biomass estimates obtained without a CPUE time series were highly uncertain. Adding a current biomass depletion observation improved precision, though the biomass time trend was sensitive to this value. Results should be interpreted carefully as several assumptions were necessary to create the long catch time series and to define informative priors, notably for the intrinsic growth rate. Despite this, the results confirm the depleted state of the thornback ray in the Bay of Biscay with the estimated biomass in 2014 being around 3% of carrying capacity.

Keywords: Population dynamics / stock assessment / data poor / censored data / Bayes / thornback ray / state-space model

1 Introduction

Several marine fish stocks have strongly declined during the 20th Century as a consequence of overfishing, including certain rays and sharks (Quéro and Cendrero 1996; Dulvy et al. 2014). Global ray and shark landings peaked in 2003, but the recent decrease seems to be more driven by demand rather than being the result of a range of management measures (Davidson et al. 2015). The conservation of rays and shark populations has become a major management objective for ensuring sustainable exploitation of marine resources (Dulvy et al. 2014; Davidson et al. 2015). For numerous populations, available data are restricted to life history traits and landings at species level are often not available (Davidson et al. 2015). In Europe the situation has been improving in recent years with most landings now being declared at the species level (Silva et al. 2012; ICES 2014a). In the Bay of Biscay (ICES Subarea VIII), species-specific reporting of thornback ray has become mandatory since 2009 (EC 2009). However, skates and rays are morphologically similar and variable in their coloration and patterning (Steven 1931; Quéro and Guéguen 1981), regional usages of common names are confusing, and market values are moderate and the same for all species, which makes identification of landings to species level problematic (Silva et al. 2012). Whether catches are similar or much larger than landings due to discarding depends on the species, the gear and the period (Rochet et al. 2002; Silva et al. 2012). However, the effect of ignoring discards is reduced by the fact that skates and rays seem to survive discarding relatively well. Average short term skate survival under commercial fishing conditions was estimated around 55% (Enever et al. 2009).

In the Northeast Atlantic, stock abundance or biomass of ray and skates is usually not quantitatively estimated and their management relies on indicator trends (ICES 2015). The thornback ray, *Raja clavata* L. 1758, is one of the more widespread ray species in the Northeast Atlantic and a good example for this data-limited situation. As a consequence, the stock dynamics in the Bay of Biscay remains poorly understood. Nevertheless, thornback ray is currently classified as Near Threatened by the IUCN (2005), the largest threat coming from target and bycatch fisheries; in the Bay of Biscay

[a] Corresponding author: florianne.marandel@ifremer.fr

Article published by EDP Sciences

thornback ray is primarily by-caught in various fisheries (ICES 2015).

The thornback ray is a medium-sized ray which is found on sandy seabed throughout shelf areas in the Northeast Atlantic (Du Buit 1974; Quéro and Vayne 2005), from 62° N down to at least 18° N and in the Mediterranean and Black Seas (Quéro and Guéguen 1981; Chevolot et al. 2006). The longest observed thornback ray individual (female) was 107 cm total length and 14 years old (Holden 1972). Sexual maturity (L_{50}) of Northeast Atlantic populations occurs at 59 to 73 cm for males and 70 to 78 cm for females, depending on the area (Serra-Pereira et al. 2011; McCully et al. 2012). In the Irish Sea, this corresponds to ages at first maturity (A_{50}) of 3.9 and 5.3 years for males and females respectively (Whittamore and McCarthy 2005). Females spawn between 70 and 170 eggs from February to September with a moderate increase of the number of eggs per year with female size (Holden 1975) and regional differences in fecundity and spawning period (Serra-Pereira et al. 2011).

Although more biological information has become available in recent years, routine stock assessments are still not carried out for thornback ray (ICES 2015). Nevertheless, available studies suggest severe depletion or evidence of decrease of Northeast Atlantic populations (Dulvy et al. 2000, 2006; Figueiredo et al. 2007; ICES 2015). However, the age composition of catches is not routinely estimated therefore age-structured models cannot be used for stock assessment of the thornback ray in the Bay of Biscay. In contrast, biomass production models are good candidates in this context. Designed to describe population dynamics, they have been widely used for stock assessment and estimation of management reference points for species without age data (e.g., McAllister et al. 2001; Ono et al. 2012; Punt et al. 2015). Production models require only a time series of catches and a Catch Per Unit Effort (CPUE) time series. They are commonly considered as the simplest stock assessment models. Simulations have shown that unreliable reference point estimates were often due to the poor quality of the data rather than to the lack of age structure of the model (Hilborn and Walters 1992).

We implemented a biomass production model using a state-space model (SSM) formulation, which is not uncommon for fisheries models (e.g. Hammond and Trenkel 2005; Ono et al. 2012; Trenkel et al. 2012) and includes both process and observation errors. The process error represents random fluctuations in population size due to variations in recruitment or natural mortality. The observation error includes random sampling variability and catchability variations. Unlike the process error, the observation error can be reduced by improving sampling methods or by gathering more data (Parent and Rivot 2013; Gelman et al. 2014). To fit the model without a CPUE time-series or only a depletion estimate, we used a censored likelihood for the biomass process error. Previously a censored likelihood has been used to handle underreported catch data (Hammond and Trenkel 2005). By using a censored likelihood for the biomass production model and a depletion estimate, the model becomes equivalent to stock-reduction analysis (SRA, Kimura and Tagart 1982; Kimura et al. 1984).

The SSM was implemented in a Bayesian framework to draw inference on biomass trajectories and biological parameters such as the carrying capacity, and applied to the thornback ray population in the Bay of Biscay. The Bayesian framework provides flexibility for statistical modelling, inference and prediction and allows the integration of different types of information and multiple sources of uncertainty in data and models (Parent and Rivot 2013; Gelman et al. 2014). It differs from the frequentist framework in the way parameters are treated. The Bayesian approach considers parameters as random variables while the frequentist framework considers parameters as fixed values. Bayesian SSM have been widely used in fisheries science (e.g. McAllister and Ianelli 1997; Punt and Hilborn 1997; Robert et al. 2010). In our case, despite a general lack of data, the biology of thornback ray is known well enough to integrate it through informative priors. Several methods exist for obtaining posterior parameter distributions. Among them the Markov Chain Monte Carlo (MCMC) approach, which is commonly used for fisheries stock assessments (e.g. Hammond and Trenkel 2005), can be implemented easily using the freely available software BUGS (Bayesian inference Using Gibbs Sampling) (Thomas et al. 2006). The MCMC inference method is not detailed any further.

To assess the strengths and limitations of the proposed method, a simulation-estimation (SE) analysis was conducted. Widely used within the frequentist framework, SE approaches are less common in the Bayesian framework (Ono et al. 2012). The use of simulated data helps to determine the performance of the method by comparing the true parameter values used in the simulations to the posterior distributions. Three types of scenarios were investigated: scenarios with variation in biological parameters (such as the intrinsic growth rate), scenarios with variation in process and observation errors and scenarios exploring the use of only a final depletion estimate. The SSM was then fitted to thornback ray catch data from the Bay of Biscay for the period 1903 to 2013 together or without a research vessel CPUE time series for the years 1990 to 2013, or a biomass depletion estimate for 2014.

2 Material and methods

2.1 Data

2.1.1 Hypothetical landings

The longest time series of commercial ray and skate landings available for the Northeast Atlantic comes from the North Sea (Heessen 2003; Walker and Hislop 1998) while historic landings of rays and skates in the Bay of Biscay are unreliable with missing data for several countries in many years and unrealistic temporal patterns until the late 1990s. Therefore, a hypothetical time series of thornback ray landings for the Bay of Biscay was created for the period 1903 to 2013 by assuming that the overall trend between 1903 and 1995 followed that of total ray and skate landings in the North Sea and thereafter the landings collated by ICES were reliable (ICES 2014a). The North Sea landings time series is characterised by strong drops during the two world wars, followed by peaks in landings just after the wars (Fig. 1a). Although landings data for the Bay of Biscay are less reliable than for the North Sea, the effect of reduced fishing has been documented for the Second World War

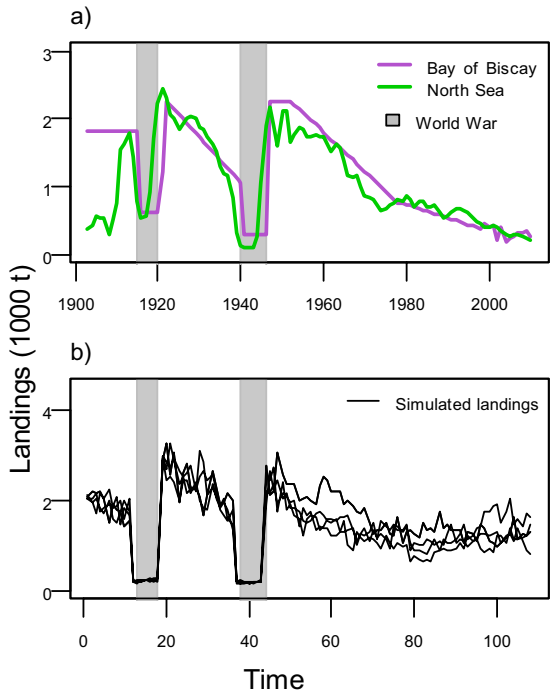


Fig. 1. (a) Hypothetical landings for thornback ray in the Bay of Biscay for the period 1903 to 2013 constructed to mimic total ray and skate landings in the North Sea. Landings peak after both World Wars (grey shaded areas) and decrease gradually thereafter. (b) Example simulated catch time series used for the simulation-estimation approach.

(WWII) with much higher landings after the war, including for rays and skates (Letaconnoux 1948). Landings during the period 1903 to 2000 were approximated by seven periods of stable or monotonously decreasing landings which were connected to create a complete time series. Stable landings were assumed from 1903 to 1913 like in Walker and Hislop (1998), for five years during WWI, for six years during WWII and just after it, but at a higher level. Decreasing landings were assumed between the two wars and after 1950s, with a changing slope in the late 1970s (Fig. 1a). The overall level was set so that landings in 1995 were about the mean of ICES landings in 1996-1999, that is 400 tonnes.

Discards were not included and landings were considered equivalent to catches because the rate of discarding of thornback ray in the Bay of Biscay is not quantified. Thornback ray has always been marketable. Thus historic discards might have been limited to small and damaged individuals. In recent years, discards may have increased as a consequence of restrictive quotas but such discards might at least partly survive (ICES 2014a).

Pauly and Zeller (2016) presented an alternative catch time series for total sharks and rays in the Bay of Biscay for the period 1950 to 2006. Again, several steps were needed to create an alternative hypothetical thornback ray catch time series from this. First, for years before 2006 average landing proportions of the main shark species (spurdog, porbeagle, smooth-hound, tope and catsharks) were calculated from the data held by ICES and then subtracted from the total shark and ray catches to obtain a rays only catch time series. Second, the average proportion of thornback ray in recorded landings was calculated from the ICES data and applied to these "rays" landings. From 2007 onwards species-specific thornback ray landings can be used (ICES 2015). In contrast to the time series created in this study, this alternative thornback ray catch time series shows no time trend since 1950 with catches varying around 500 tonnes (Appendix B). After 1980, both hypothetical catch time series are of the same order of magnitude but the alternative hypothetical catch time series presents higher inter-annual variations. The lack of a general time trend and the higher inter-annual variations make this alternative catch time series more unlikely. Consequently, it was not used in this study.

2.1.2 Research vessel CPUE

A research vessel CPUE was calculated using data from the EVHOE bottom trawl survey in the Bay of Biscay between 1987 and 2014 and from surveys carried out in 1973 and 1976. Each survey used the same trawl and sampling protocol but there were some differences in the area and depths covered (see Poulard and Blanchard 2005 and Poulard and Trenkel 2007 for EVHOE survey details). Post-stratification was used by first delineating the area occupied by thornback ray in each year and then calculating the swept area based total biomass in the occupied area. The post-stratified research vessel CPUE was well correlated to the index obtained using the full EVHOE stratification design, which was however not available for years before 1987. As the EVHOE survey was carried out in October-November, the research vessel CPUE of year t was compared to modelled biomass for year $t + 1$.

2.2 State-space model

2.2.1 Biomass dynamic

The biomass dynamics model was based on a discrete-time sequential equation that mimics the biomass dynamics of the population. The biomass at time $t + 1$ (B_{t+1}) depends on the biomass at time t (B_t), the production between times t and $t + 1$ and the cumulative catches during the same period. Production was modelled by the Schaefer production function, which integrates biological processes such as recruitment and growth. Following Agnew et al. (2000) the Schaefer function was chosen because of the shape of the stock-recruitment relationship of the species: rays and skates show a close relationship between mature population size and recruitment due to low fecundity of adults and late maturity (Whittamore and McCarthy 2005; ICES 2014b). It has previously been used for elasmobranchs (Walker 1992; Agnew et al. 2000).

The Schaefer production function has two biological parameters: intrinsic growth rate r and carrying capacity K, leading to the following state equation where C_t represents the catches during year t:

$$B_{t+1} = B_t + r * B_t \left(1 - \frac{B_t}{K}\right) - C_t. \quad (1)$$

Table 1. Model parameters and prior distributions used for the simulation-estimation study and the application to thornback ray in the Bay of Biscay. The BUGS aliases relate to the BUGS code provided in Appendix A.

Parameter	Description	Prior	BUGS Alias
r	Intrinsic population growth rate	Beta (34, 300) mean = 0.1, CV = 0.16	`r`
K	Carrying capacity	Uniform (20 000, 100 000)	`K`
$Y1903$	Initial relative biomass in 1903	Beta (17, 4) mean = 0.84, CV = 0.1	`Y0`
$Y2000$	Initial relative biomass in 2000	Beta (2,6) mean=0.16, CV=0.6	`Y0`
$1/\sigma^2$	Process error precision (inverse variance)	Gamma (400, 1) mean = 399, CV = 0.05	`ytau2`
q	Survey catchability	Uniform (0.01, 0.6)	`q`
$1/\tau^2$	Observation error precision (inverse variance)	Gamma (44,2) mean = 22, CV = 0.15	`itau2`
CV	Uncertainty of landings	0.2 (constant)	`CV`

To facilitate model fitting, the model was formulated for relative biomass $Y_t = B_t/K$ as in Hammond and Trenkel (2005). The relative biomass in the first year is denoted $Y0$. The process error was assumed to be normally distributed with constant variance σ^2.

$$Y_{t+1} \sim N\left((r+1)\,Y_t - rY_t^2 - \frac{C_t}{K}, \sigma^2\right). \quad (2)$$

The biomass distribution was truncated at both ends leading to a censored likelihood. Assuming that the mean biomass cannot be much larger than the carrying capacity, the upper bound for Y_t was set to $1 + 3\sigma$. We can also safely consider that biomass B_t was higher than the hypothetical landings, noted l_t, for a given time period. Moreover, the probability of catching more than half the population in a single year was considered low, leading to the constraints:

$$\frac{2l_t}{K} < Y_t < 1 + 3\sigma. \quad (3)$$

As the hypothetical landings l_t were uncertain but not necessarily biased, catches were modelled by a lognormal distribution with mean equal to the hypothetical landings and the variance corresponding to a constant coefficient of variation (CV) of 20%:

$$C_t \sim \log N\left(\log(l_t), \log\left(CV^2 + 1\right)\right). \quad (4)$$

Maximum Sustainable Yield (MSY) and corresponding B_{MSY} were calculated according to the following equations:

$$B_{\mathrm{MSY}} = \frac{K}{2}, \quad MSY = \frac{rK}{4}. \quad (10)$$

2.2.2 Observation model

The observation model links population biomass to the research vessel CPUE via a catchability constant q. The observation error of the research vessel CPUE i_t was modelled with a lognormal distribution and a constant variance τ^2, i.e. constant CV. It incorporated sampling variability and random variation in catchability.

$$i_t \sim \log N\left(\log\left(qY_tK\right), \tau^2\right). \quad (5a)$$

The observation model was replaced by a truncated normal distribution with variance ε^2 when, instead of a research vessel CPUE time series, only an estimate of the depletion level d_t was used in year t.

$$d_t \sim N\left(Y_t, \varepsilon^2\right) \quad 0 < d_t < 1. \quad (5b)$$

2.2.3 Prior distributions

Prior distributions and fixed values for the seven model parameters $\theta = (Y0, K, r, \sigma, q, \tau, CV)$ are summarized in Table 1. Thornback ray in the Bay of Biscay was exploited before 1903 but not overexploited (Quéro and Cendrero 1996). Therefore, an informative Beta distribution was selected as prior for $Y0$ (noted $Y1903$ for the case study) for which 95% of the probability mass were contained between 0.5 and 1 (Table 1). The informative prior distribution for the intrinsic growth rate r was derived using the Leslie method reviewed in McAllister et al. (2001). The method involved simulating an age-structured population model at equilibrium assuming 140 eggs per year for females 5 years and older, egg-to-juvenile natural mortality of 5 (corresponding to a survival rate of 0.0067) and adult natural mortality of 0.1. Numbers were transformed to biomass using growth parameters $L_\infty = 118$, $k = 0.155$ and $t_0 = 0.655$ (Wiegand et al. 2011) and weight-length parameters $a = 0.00000345$ and $b = 3.1807$ (Dorel et al. 1998). This provided an r estimate of around 0.105. To incorporate variability in intrinsic growth rates for elasmobranchs (Camhi et al. 2008; Frisk et al. 2005), an informative Beta distribution was selected for r for which 95% of the probability mass was contained between 0.05 and 0.15 with the mode at 0.105 (Table 1).

A uniform prior distribution was used for the carrying capacity K. As this parameter is very population-dependent, it would have been difficult to define an informative prior based on knowledge from other populations. The limits of the uniform distribution are somewhat arbitrary reflecting expert knowledge for the Bay of Biscay population (Table 1). An uniform distribution was also used as prior for catchability q with upper limit 0.5 and lower limit 0.01, both reflecting expert knowledge and results reported in the literature

(Fraser et al. 2007). An informative prior distribution was chosen for the precision of the process error ($1/\sigma^2$) as the data contains no information on this parameter (Table 1); similarly for the precision of the CPUE time series ($1/\tau^2$).

2.2.4 Bayesian inference

All computations were performed with the R platform (v3.1.2, R Foundation for Statistical Computing 2015). OpenBUGS (v3.2.3, Thomas et al. 2006) was used for Bayesian inference and was run within R using the *BRugs* package (Thomas et al. 2006). Results were calculated for three parallel MCMC chains, composed of 150 000 iterations with different initialization points. The burn-in for each MCMC chain was 40 000 iterations and autocorrelation among samples was limited by saving every 100th parameter value. Convergence was checked with several diagnostics including the Gelman-Rubin diagnostic (Gelman et al. 2014), the Geweke convergence diagnostic (Geweke 1992) and an expert appreciation of trace plots created using the package *Coda* (Plummer et al. 2006). The BUGS code is provided in Appendix A.

2.3 Simulation-estimation approach

For the simulation-estimation (SE) approach, time series of catches and biomass were simulated for 111 years as in the case study using equations (1) to (5) as operating model and different sets of parameter values. Model performance was then investigated for two cases. In the first case simulated catches and the research vessel CPUE were used for Bayesian inference. In the second case, catches and only a depletion index for the final year were used.

The SE approach focused on two main issues:

(i) Are the posterior parameter estimates sensitive to the prior distribution used for biological parameters(r, K)?
(ii) Does the model succeed in correctly estimating the depletion state of the simulated population in the last year? And does it succeed in estimating the underlying biomasses?

2.3.1 Simulation setup

Overall fourteen scenarios were investigated (Table 2). Catches used in all scenarios were created to mimic the temporal pattern of the hypothetical time series for thornback ray in the Bay of Biscay (Fig. 1a). Scenarios 1 to 4 aimed at testing the effects on model performance of the values of the biological parameters r and K. For this, two values were used for each parameter which were considered realistic for elasmobranchs and were towards the upper and lower end of the respective prior distributions. All other parameters (q, τ, σ, $Y0$ and CV) had the same value in all four scenarios, referred to as reference values and considered plausible for the Bay of Biscay thornback ray case study (Table 3).

Scenarios 5 to 12 aimed at testing the effects on model performance of the values of the variance parameters σ^2, τ^2 and CV; reference values were used for r, K, $Y0$ et q for these scenarios (Table 2). Two realistic values were tested for each parameter. The two values tested for the CV correspond to a maximum (0.4) and a realistic optimistic value (0.1). However, CV was always fixed at 0.2 for inference.

Table 2. Description of scenarios used in the simulation-estimation approach. Examples of variable catch time series are provided in Figure 1b. Reference values are summarised in Table 3.

Scenario	K	r	σ	τ	CV
Reference	60 000	0.105	0.05	0.2	0.2
1	90 000	0.13			
2	30 000	0.08			
3	90 000	0.08	reference values		
4	30 000	0.13			
5			0.03	0.2	0.1
6			0.1	0.2	0.1
7			0.1	0.3	0.1
8			0.03	0.3	0.1
9			0.03	0.2	0.4
10	reference values		0.1	0.2	0.4
11			0.1	0.3	0.4
12			0.03	0.3	0.4
I			reference values		
II					

Table 3. Reference parameter values used in the simulation-estimation approach and posterior mean estimates for thornback ray in the Bay of Biscay (95% credible interval) from FULL run (landings and research vessel CPUE for the years 1973–2013). In the FULL run, $Y0$ correspond to the initial relative biomass in 1903, noted $Y1903$ for the case study.

Parameter	Reference value	Thornback ray (CI 95%)
K	60 000	63 000 (42 000–94 000)
r	0.105	0.092 (0.065–0.12)
q	0.15	0.13 (0.07–0.19)
$Y0 \vert Y1903$	0.85	0.82 (0.63–0.94)
σ	0.05	0.05 (0.048–0.053)
τ	0.2	0.23 (0.19–0.27)
CV	0.2	0.2 (fixed)

The final two scenarios, denoted I and II, used reference values for all model parameters. The purpose of these two scenarios was to test model performance in the case where only a depletion estimate for the final year (year 111) was available instead of a research vessel CPUE time series. To create a depletion observation (d_{111}) a random draw from a truncated normal distribution with mean Y_{111} and variance ε^2 was carried out:

$$d_{111} \sim N\left(Y_{111}, \varepsilon^2\right) \quad 0 < d_{111} < 1. \tag{6}$$

To evaluate the impact of the observation error, two values were tested for ε: 0.05 (scenario I) and 0.2 (scenario II).

The model assumed catches were known. To simulate a time series of catches with lognormal observation error, variable exploitation rates (0.005–0.07) were applied to the simulated biomass of each year and a random draw was carried out from a lognormal distribution with the resulting values as means and coefficients of variation equal to the value of the parameter CV. The values of the exploitation rates were chosen by hand to achieve the desired temporal pattern

and comparable magnitude. For scenarios 1 to 12, a research vessel CPUE time series was simulated for the last 29 years to reflect the data available for the case study. For scenarios I and II a depletion index for the final year was simulated. For all scenarios, 100 replicates were generated. Initial analyses showed that 100 replicates were enough to reliably evaluate performance; performance measures stabilised at around 30 replicates (results not shown).

The simulated catches and research vessel CPUE time series (or depletion index) together with the prior distributions listed in Table 1 were used for Bayesian inference. For all simulations, the model used for the estimation was the same as the operating model used for simulations; this condition ensured that any differences between simulations and estimation were due to estimation performance only.

2.3.2 Performance assessment

The performance of the estimation method was assessed with three complementary criteria chosen to represent how well the estimated parameter values $\hat{\theta}_s$ (posterior means) of replicate s ($s = 1, \ldots, 100$) agreed with the true parameter values θ_{SC} of the scenario (Table 2). First, the Mean Relative Error (MRE) was calculated to quantify the bias of Bayesian estimation for each of the 14 scenarios (SC). A negative MRE value means that globally the model tended to underestimate the parameters value and a positive value means overestimation.

$$MRE\,(\theta_{SC}) = \frac{1}{100}\sum_{s=1}^{100}\left(\frac{\hat{\theta}_s - \theta_{SC}}{\theta_{SC}}\right). \quad (7)$$

Second, the Mean Squared Error (MSE) was calculated to measure precision of Bayesian posterior mean estimates for each parameter and scenario:

$$MSE\,(\theta_{SC}) = \frac{1}{100}\sum_{s=1}^{100}\left[(\hat{\theta}_s - \theta_{SC})^2\right]. \quad (8)$$

Finally, the Median of Absolute Relative Error was calculated which quantifies "average" model precision if the model is correct (Ono et al. 2012), where $\hat{\theta}'_s$ is the posterior median of replicate s ($s = 1, \ldots, 100$). The smaller the value, the more precise are the parameter estimates on "average".

$$MARE\,(\theta_{SC}) = \text{median}\left(\left|\frac{\hat{\theta}'_1 - \theta_{SC}}{\theta_{SC}}\right|, \ldots, \left|\frac{\hat{\theta}'_{100} - \theta_{SC}}{\theta_{SC}}\right|\right). \quad (9)$$

2.4 Application to thornback ray in the Bay of Biscay

The model was applied to the thornback ray in the Bay of Biscay using the priors defined in Table 1.

Four runs were made using different data combinations and time periods to explore the importance of the different data types. For the full run (FULL), the full hypothetical landings time series (1903–2013) and research vessel CPUE time series (1973, 1976, 1987–2013) were used in the model. To avoid having to make too many assumptions for reconstructing the catch time series a run (SHORT) restricted to the recent time period (2000–2013) was also carried out. For this run the prior $Y2000$ was used instead of that for $Y1903$ (see Table 1). The landings only run (LANDINGS) represented the case where no research vessel CPUE was available or where it was deemed unusable due to poor quality. The fourth run (DEPLETION) represented a situation where no research vessel CPUE but an estimate of the final depletion level d_{2014} was available. Given thornback ray in the Bay of Biscay is thought to be overexploited, a relatively small value was chosen ($d_{2014} = 0.1$) with a small standard deviation ($\varepsilon = 0.05$). These values are somewhat arbitrary but the aim was to compare the biomass trajectories obtained with a research vessel CPUE and with only information for the depletion level in the final year.

2.5 Posterior predictive check

For an overall assessment of the FULL run, we examined how well the fitted model could reproduce the available research vessel CPUE time series. We simulated 3000 CPUE time series using parameter values drawn from the joint posterior distribution, hypothetical catches and the model (Eqs. (1)–(6)). The distribution of simulated CPUEs was then compared with the observed research vessel CPUE time series.

3 Results

3.1 Simulation-estimation approach

For all replicates of all scenarios the population never crashed; moreover biomass trajectories never hit the lower bound ($2l_t/K$). Averaged across scenarios 1 to 12, estimates of parameter r had the smallest mean MRE and second smallest MARE (–0.013 and 0.105 respectively), followed by $Y0$ (MRE –0.05; MARE 0.04), τ (MRE –0.054; MARE 0.118), σ (MRE 0.055; MARE 0.389), K (MRE 0.15; MARE 0.17) and q (MRE 0.17; MARE 0.35). The parameter CV was not estimated though the true value was varied for certain scenarios. The MRE and MSE values for parameters K, r, $Y0$, σ and τ were negatively correlated, while they were positively correlated for q (not shown). For scenarios 1 to 12, results for performance measures MRE, MSE and MARE are detailed in Table 4.

The estimation quality of parameter K varied between scenarios depending on whether the true value of K was larger than the mean of the prior uniform distribution (MRE negative; scenarios 1 and 3), smaller (MRE positive; scenarios 2 and 4) or equal to the mean (small positive MRE; scenarios 5–12) (Table 4, Fig. 2a). The highest overestimation was found for scenario 4 (MRE = 0.5); this scenario combined low r with low K values. For this scenario the MRE corresponded exactly to the relative difference between the mean of the prior and the true value. The complementary scenario 1 (true r and K above mean of priors) had the smallest MRE of all tested scenarios. For all scenarios, except scenario 4, MRE values were smaller than the relative difference between true value and the mean of the prior for K, indicating that the data of these scenarios contained information on the parameter.

Table 4. Simulation-estimation results for 14 scenarios (see Table 2). Parameter estimation performance was measured by Mean Relative Error (MRE), Mean Square Error (MSE) and Median Absolute Relative Error (MARE). The best performing scenario is in green and the worst in red.

Measure	Parameter	1	2	3	4	5	6	7	8	9	10	11	12	I	II
MSE	*K*	2.11E+07	7.99E+07	1.86E+08	2.44E+08	1.61E+08	9.78E+07	1.11E+08	1.26E+08	1.76E+08	1.58E+08	1.20E+08	1.56E+08	2.37E+08	2.33E+08
	r	4.01E−04	3.38E−04	5.90E−04	8.50E−04	2.25E−06	4.98E−05	7.26E−05	1.04E−05	4.36E−06	4.41E−05	8.04E−05	1.17E−05	2.41E−05	2.36E−05
	MSY	3.19E+05	1.24E+05	1.08E+05	3.91E+04	1.10E+05	6.62E+04	6.56E+04	7.08E+04	1.19E+05	1.04E+05	5.92E+04	8.98E+04	1.37E−05	1.33E−05
	Y0	0.0017	0.0021	0.0015	0.0024	0.0021	0.0021	0.0020	0.0020	0.0020	0.0020	0.0020	0.0019	0.002	0.002
	q	0.0027	0.0013	0.0018	0.0014	0.0017	0.0110	0.0175	0.0025	0.0024	0.0074	0.0191	0.0024	–	–
	σ	2.49E−09	3.48E−09	3.33E−09	3.60E−09	3.98E−04	2.49E−03	2.49E−03	4.00E−04	3.99E−04	2.50E−03	2.49E−03	4.00E−04	1.20E−09	1.10E−09
	τ	4.25E−05	5.74E−05	6.75E−05	3.94E−05	1.17E−04	2.21E−04	6.05E−03	5.87E−03	1.44E−04	2.57E−04	6.18E−03	5.99E−03	–	–
MRE	*K*	−0.044	0.248	−0.134	0.500	0.204	0.082	0.099	0.178	0.212	0.134	0.120	0.200	0.234	0.233
	r	−0.153	0.225	0.298	−0.224	−0.005	−0.056	−0.070	−0.022	−0.007	−0.049	−0.075	−0.022	0.012	0.012
	MSY	−0.190	0.535	0.127	0.165	0.199	0.025	0.024	0.152	0.204	0.084	0.038	0.174	0.180	0.179
	Y0	−0.048	−0.054	−0.046	−0.057	−0.053	−0.053	−0.053	−0.053	−0.053	−0.052	−0.053	−0.051	−0.053	−0.053
	q	0.309	−0.146	0.120	−0.147	−0.234	0.504	0.720	0.021	−0.292	0.403	0.768	-0.058	–	–
	σ	−0.00073	−0.00016	−0.00018	−0.00097	0.66539	−0.49949	−0.49915	0.66706	0.66550	−0.49959	−0.49912	0.66707	0.00009	0.00003
	τ	0.029	0.034	0.037	0.028	0.050	0.071	−0.258	−0.254	0.055	0.076	−0.261	−0.257		
MARE	*K*	0.023	0.126	0.121	0.379	0.210	0.135	0.153	0.168	0.213	0.164	0.131	0.197	0.244	0.233
	r	0.157	0.229	0.303	0.229	0.009	0.057	0.073	0.025	0.013	0.059	0.077	0.026	0.038	0.038
	MSY	0.175	0.359	0.146	0.096	0.199	0.114	0.125	0.136	0.195	0.145	0.106	0.165	0.175	0.157
	Y0	0.037	0.043	0.035	0.046	0.043	0.042	0.042	0.041	0.041	0.041	0.041	0.040	0.043	0.042
	q	0.238	0.226	0.150	0.240	0.312	0.320	0.675	0.257	0.360	0.273	0.831	0.284	–	–
	σ	0.0016	0.0014	0.0014	0.0017	0.664	0.500	0.500	0.666	0.664	0.500	0.500	0.666	0.0009	0.0009
	τ	0.031	0.035	0.038	0.031	0.053	0.073	0.261	0.249	0.058	0.074	0.262	0.253		

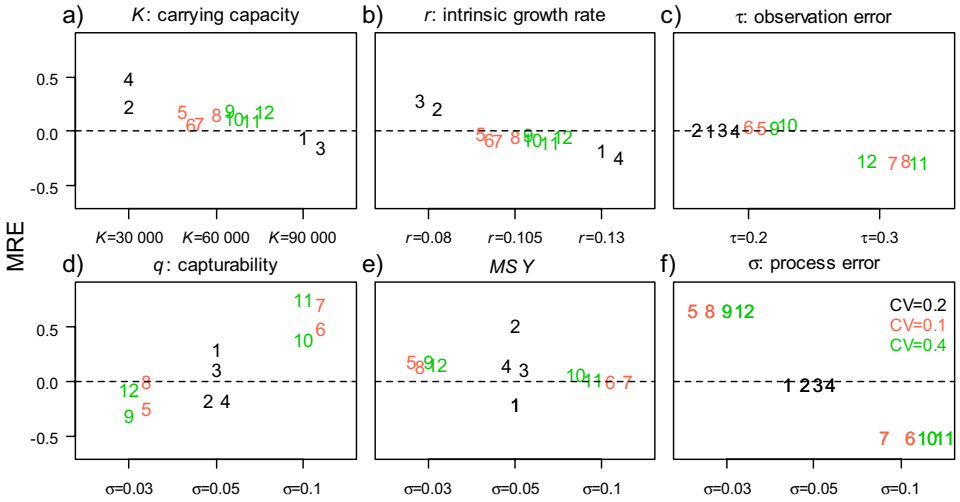


Fig. 2. Model performance criteria for scenarios 1 to 12 tested in the Simulation-Estimation study. Performance was measured by Mean Relative Error (y-axis) and plotted separately for different true model parameters values.

As for K, estimates of r can be separated into three categories according to the difference between the true value and the mean of the prior (Table 4, Fig. 2b): in scenarios 1 and 4 (true r value > mean of prior) the parameter was underestimated. In scenarios 2 and 3, with low true value, the parameter was overestimated. In general, estimates of r were driven by the prior and the parameter was hard to estimate though in all cases posteriors were shifted in the correct direction. As expected, MRE values for r and K were negatively correlated (Pearson, p-value = 0.06), except for scenarios 1 and 2. These scenarios correspond respectively to the scenario with high K/high r values and scenario with low K/low r values. The derived quantity MSY was estimated using the posterior estimates of r and K. It was generally overestimated, with the poorest performance obtained for scenario 2 (low r and low K). Estimates of MSY with smallest bias (MRE < 0.1) were obtained for scenarios (namely 6, 7, 10 and 11) which had high process errors σ (0.1) (Fig. 2e).

Estimation performance of parameter q varied between scenarios (Table 4). MRE increased with the process error σ, with strongly positively biased estimates for scenarios 7 and 11 (MRE > 0.5) (Fig. 2d). Thus larger process error made q more difficult to estimate. The initial relative biomass $Y0$ was generally well estimated with little differences between scenarios (Table 4). This is reassuring but not surprising given the true value was equal to the mean of the prior. Hence the simulations did not provide much insight into the identifiability of this parameter.

Globally, performance for parameter σ was driven by the difference between the prior and the true value (Fig. 2f). However, for small true values (0.03) the magnitude of MRE values was larger than the relative difference between the true value and the mean of the prior. This indicates that for small values of σ, posterior estimates were less biased. Parameter τ (Fig. 2c) was globally more estimable than σ and presented the lowest bias. In the scenarios (namely 7, 8, 11 and 12) in which the true value was 50% larger than the mean of the prior (0.3 and 0.2, respectively), the MRE was around 0.25, indicating a relative estimation bias of only 25%. Using larger values for CV (catch uncertainty) only slightly decreased model performance. MRE values were on average 3% larger for scenarios 9 to 12 with $CV = 0.4$ compared to scenarios 5 to 8 with $CV = 0.1$, with differences between parameters. The difference was 11% for MSE and 4% for MARE.

For scenarios I and II, using reference parameter values and a final year depletion index, posterior distributions presented updates of the prior distributions for all parameters except for the process error σ. Both scenarios performed similarly, with an overestimation of K, r and $Y0$ and a very small underestimation of σ. Performance metrics for K were comparable to those obtained for scenarios 1 to 12. The amount of observation error assumed for the final depletion level observation did not influence the quality of inference. Thus model performance achieved when using only a depletion estimate for the final year was comparable to that using a 24 years time series of biomass indices.

3.2 Application to thornback ray in the Bay of Biscay

3.2.1 FULL run: landings (1903–2013) and research vessel CPUE (1973–2013)

In this run, all available data were used for the estimation. The marginal posterior distributions of all parameters were updated compared to the prior distribution except for $Y1903$ (Fig. 3), σ and τ (not shown). The posterior mean of carrying capacity (K) was estimated to be around 63 000 tonnes (Table 3) and the intrinsic growth rate (r) 0.093, which is slightly lower than the prior derived from the Leslie method (0.105). The marginal posterior distribution of the relative initial biomass ($Y1903$) was almost identical to the prior distribution, with a mode at 0.82. The posterior for catchability (q) showed a strong update with a mean of 0.12. MSY was

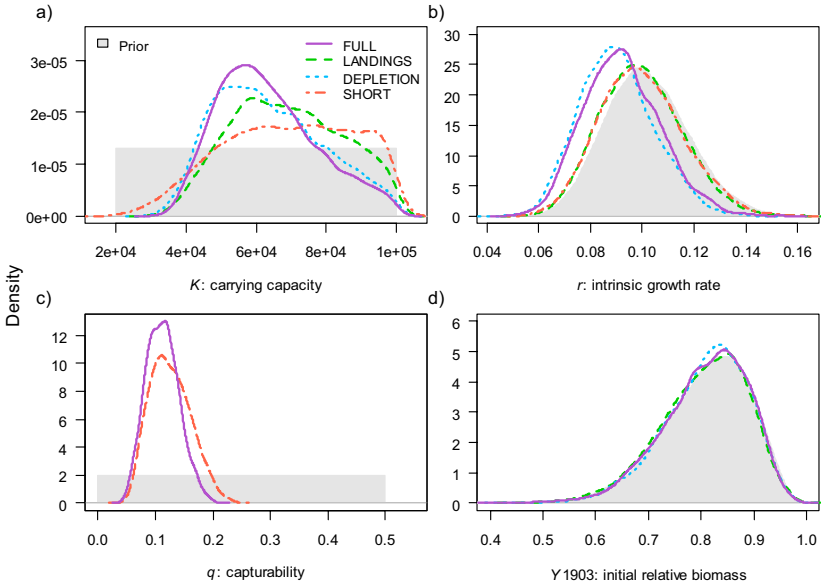


Fig. 3. Comparison of prior and marginal posterior parameter distributions for thornback ray in the Bay of Biscay for four model runs using different data combinations. FULL: landings and research vessel CPUE for the years 1973–2013; LANDINGS: landings only; DEPLETION: landings and final year depletion rate; SHORT: as in FULL but using data for years 2000–2013 only.

estimated at around 1 440 tonnes (CI 95%: 930–2 100). The analysis of the pairwise joint distributions of the four key parameters revealed significant negative correlation between K and r, K and q and $Y1903$ and K, and non-significant positive correlation between all other parameter pairs (Pearson product moment correlation: –0.6 to 0.03 with p-value > 0.05) (Fig. 4).

The influence of the catches and research vessel CPUE time series on the estimated biomass trajectory as well as the world wars' pattern is clearly visible in Figure 5a. The posterior distribution of biomass estimates was much tighter in recent years for which biomass indices were available. The 2014 biomass was estimated to be around 1 900 tonnes (CI 95%: 940–2 400) which represents 3% of the estimated carrying capacity.

3.2.2 SHORT run: landings (2000–2013) and research vessel CPUE (2000–2013)

In this run, only data for the last 14 years were used for estimation. The marginal posterior distributions of all parameters except σ and τ were updated compared to the prior distributions (orange lines, Fig. 3). The posterior mean of carrying capacity (K) was estimated to be around 68 000 tonnes (CI 95%: 34 000–98 000) and the intrinsic growth rate (r) 0.099 (CI 95%: 0.069–0.13), which is again lower than our preliminary estimate but higher than the value obtained in the FULL run. The marginal posterior distribution of the relative initial biomass ($Y2000$) presented a mode at 0.08 (CI 95%: 0.02–0.17). The posterior for catchability (q) showed a strong update with a mean of 0.13 (CI 95%: 0.07–0.2). MSY was estimated around 1 700 tonnes (CI 95%: 760–2 800) which is 300 tonnes higher than for the FULL run. The biomass in 2014 was estimated around 1 900 tonnes (CI 95%: 760–3 600) which represents 3% of the estimated carrying capacity as in the FULL run (Fig. 5b). The analysis of the pairwise joint distributions of the four key parameters revealed no significant correlations between parameters K and r, and r and q (Pearson product moment correlation: 0.007 and 0.03 with p-value > 0.05). Low significant correlation was found between all other parameter pairs (Pearson product moment correlation: 0.02 to –0.09 with p-value < 0.05) except for parameters K and q which were not correlated (Pearson product moment correlation: –0.76 with p-value > 0.05).

3.2.3 LANDINGS run

In this run, only the full time series of hypothetical landings was used but no CPUE. Analysis of the marginal posterior distributions showed lower update of prior distributions compared to the FULL run, in particular for K (green dashed line, Fig. 3). The posterior mean carrying capacity was estimated to be higher at around 70 000 tonnes (CI 95%: 41 000–98 000) while the intrinsic growth rate was around 0.1 (CI 95%: 0.07–0.13), which is similar to the mode of the prior. The posterior distribution of the relative initial biomass $Y1903$ was identical to its prior distribution. MSY was estimated to be around 1 700 tonnes (CI 95%: 1 000–2 600). Analysis of the pairwise joint posterior distributions showed relatively low correlations (–0.3 to 0.007, p-value < 0.05), with the largest negative value (–0.3, p-value < 0.001) between r and K. The estimated biomass trajectory first decreased and then increased after 1960 (green solid curve, Fig. 5). The biomass in 2014 was estimated to be around 55 000 tonnes (CI 95%: 22 000–87 000) which represents 79% of the estimated carrying capacity of this run.

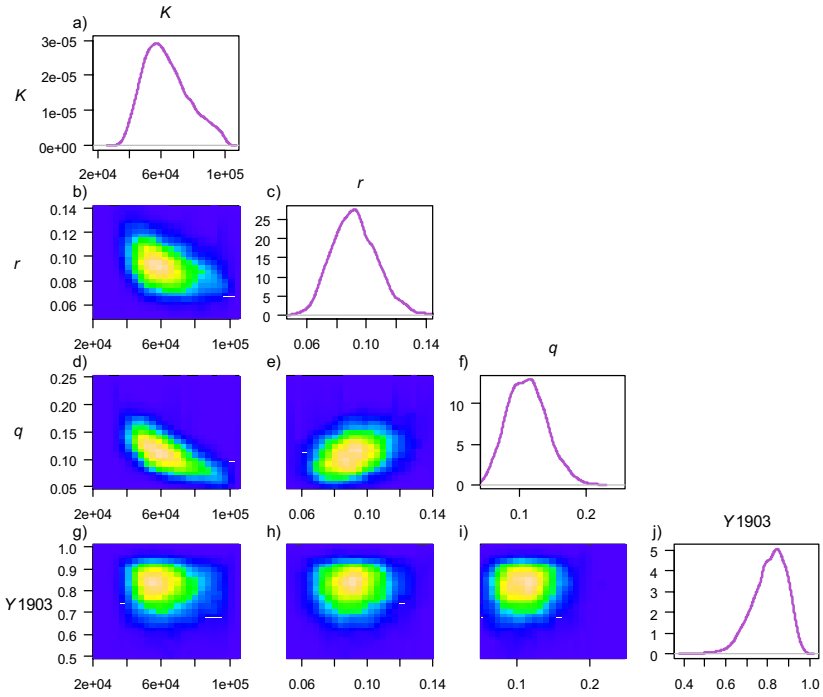


Fig. 4. Results for FULL run for thornback ray in the Bay of Biscay. Joint posterior (density surfaces) and marginal posterior distributions (lines) for main model parameters. *K* carrying capacity. *r* intrinsic growth rate. *q* survey catchability and *Y*1903 relative initial biomass. The colour scales goes from low values (blue) to high values (orange).

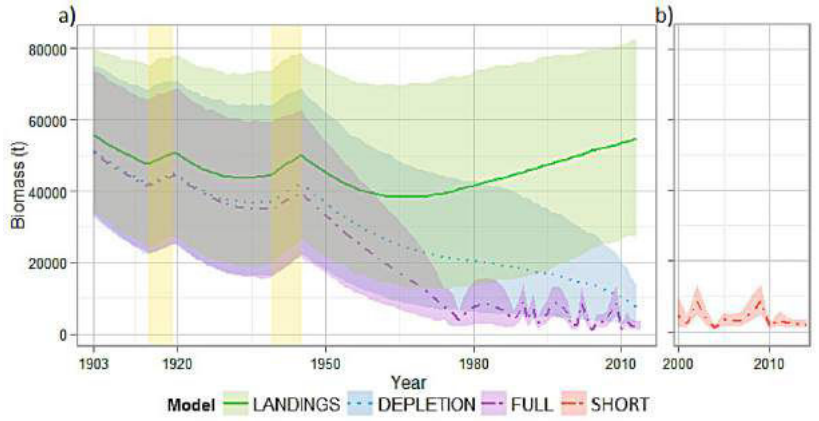


Fig. 5. (a) Estimated biomasses trajectories for thornback ray in the Bay of Biscay for model runs using different data series. LANDINGS: landings only; DEPLETION: landings and final year depletion rate; FULL: landings and research vessel CPUE for the years 1973–2013. Coloured areas: credible intervals between 2.5th and 97.5th percentiles. Vertical rectangles: World War I and II periods. (b) Estimated biomasses trajectories for thornback ray in the Bay of Biscay obtanied using only catches and research vessel CPUE time series from 2000 to 2013 (SHORT run).

3.2.4 DEPLETION run

In this run, an estimate of the final (2014) depletion level in addition to landings was used. Comparison of prior and posterior distributions showed a slightly better update for *K* compared to the LANDINGS run while the posterior for *r* was shifted to the left (blue dotted line, Fig. 3). The mean posterior carrying capacity was around 64 000 tonnes (CI 95%: 40 000–96 000) and the intrinsic growth rate 0.09 (CI 95%: 0.07–0.12). *MSY* was estimated around 1 400 tonnes (CI 95%: 900–2 100). Analysis of the pairwise joint posterior distributions showed again low correlations (~0.05, *p*-value > 0.05) except between *r* and *K* for which it was –0.45 (*p*-value < 0.001). The estimated biomass trajectory showed a continuous decline (blue dotted curve, Fig. 5). The 2014 biomass was estimated around

7 500 tonnes (CI 95%: 2 500–15 000) which corresponds to 11% of the estimated carrying capacity.

3.2.5 Comparison

The data used for the four model runs contained variable amounts of information which led to different degrees of update of prior parameter distributions (Fig. 3). The largest update was achieved when all data were used (FULL run), while the landings only (LANDINGS) run contained very little information. Posterior mean estimates of carrying capacity varied little between runs (63 000–70 000 tonnes), similar for estimates of the intrinsic growth rate (0.09 to 0.10). Posteriors of K and r were negatively correlated in all runs, with the highest negative correlation for the depletion (DEPLETION) run and the lowest for the FULL run. This shows that the identifiability of the two parameters increased as more data was used, i.e. the two parameter estimates became less confounded. Posterior distributions for $Y1903$ were relatively similar for the three runs which included this parameter, which is not surprising given the posteriors were identical to the prior distribution. The parameter q was present only in the FULL and SHORT runs and had identical posterior mean of 0.13.

The shape of the biomass trajectory was similar for the three long runs before 1950 but after this date each run lead to a different assessment of the dynamics of the thornback ray population in the Bay of Biscay. The LANDINGS run led to an unrealistic biomasses trajectory where recent biomass levels were similar to 1903 biomass levels. This run also had the largest credible intervals for biomass estimates (Fig. 5a). The FULL and SHORT runs provided more precise biomass estimates but implied the strongest depletion and hence the worst current state of the population.

Biomass trajectories for the FULL run (Fig. 5a) were similar to those for the SHORT run (Fig. 5b). Estimates for parameters MSY, K and q were also similar but with higher uncertainty for the SHORT run. The estimate for r was lower in the FULL run. Both models presented no update for parameters σ and τ. In the SHORT run, there was a large update of $Y2000$ with a posterior mean of 0.08 instead of 0.2 for the prior; there was no update for $Y1903$ in the FULL run.

3.3 Posterior predictive check

A posterior predictive check was carried out for the FULL run only. For this 3000 biomass trajectories were simulated using the joint posterior distribution of model parameters. The model reproduced the data for 7% of the simulated biomass trajectories which, gathered, shape the one estimated for thornback ray in the Bay of Biscay (Figs. 5a and 6a). However, the model also simulated very different biomass trajectories and among them, 22% lead to a final biomass above 60 000 tonnes i.e. above the estimated carrying capacity (green trajectories in Fig. 6).

Two patterns can be distinguished in the simulated biomass trajectories. The first one presented an increase in biomass after 1977 which corresponds to a change in the catch time series trend (green trajectories in Fig. 6a). They corresponded to relatively high carrying capacities around 80 000 tonnes (min: 54 000 tonnes and max: 107 000 tonnes). Values of r used in these trajectories varied between 0.04 and 0.14. The second pattern corresponded to trajectories which decreased over the whole period (purple trajectories in Fig. 6a). They corresponded to low and medium carrying capacities around 60 000 tonnes (min: 25 000 tonnes and max: 99 000 tonnes). Values of r creating these trajectories also varied between 0.04 and 0.13. The values for $Y0$ were similar for the two patterns (around 0.73).

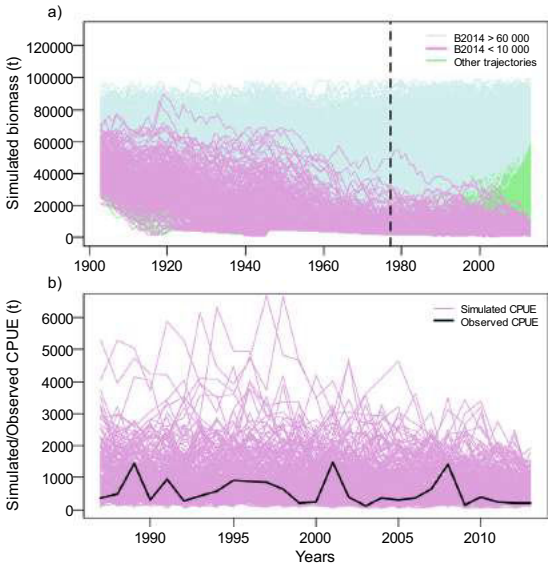


Fig. 6. Results for posterior predictive check for FULL run for thornback ray in the Bay of Biscay. (a) Simulated biomass trajectories. (b) Simulated CPUE trajectories for stock biomass B2014 < 10 000 (purple lines in (a) and research vessel CPUE time series (black line). In (a) the vertical line corresponds to the year 1977 where a shift in the blue trajectories is notable. In blue, biomass trajectories for which the biomass in 2014 (B2014) is above 60 000 tonnes. In purple, biomass trajectories for which the biomass in 2014 is below 10 000 tonnes. All others simulated trajectories are represented in green.

The simulated research vessel CPUE time series corresponding to biomass in 2014 below 10 000 tonnes were comparable in magnitude and time trend to the research vessel CPUE time series used in the estimation (Fig. 6b). Thus, the fitted model reproduced the observed data even though only around 7% of the trajectories showed the same pattern and magnitude.

4 Discussion

4.1 Bayesian state space model

The Bayesian SSM used in this study made it possible to include different sources of information while accounting for observation uncertainty and natural variability. The inclusion of expert knowledge can be divided into two types: censored

likelihood and informative priors. The censored likelihood improved the calculation time and the general performance of the model. It also allowed fitting the model without a research vessel CPUE. Biological knowledge was used for deriving an informative prior of the intrinsic growth rate.

For elasmobranchs the relationship between stock size and recruitment is rather linear, owing to their reproductive strategy of low fecundity and high per-offspring investment (Hoening and Gruber 1990). Despite being "the simplest production function" (Hilborn and Walters 1992), the Schaefer model is well suited for this case. It also fitted the data well and led to reasonable estimates for r and K.

The SE study with inferences based on catches and a research vessel CPUE (scenarios 1 to 12) showed the sensitivity of model results to prior distributions, i.e. the effect of differences between the mode of the prior and the true value. For scenarios using low K, the bias in estimates of K was positive and lower than 0.25 except for scenario 4 (low r low K) for which the bias was highly positive. This indicates that the estimates of K obtained for thornback ray in the Bay Biscay, obtained with the set of priors used in this study, could be biased to some extent. Further, posterior distributions for r were generally similar to prior distributions leading to an underestimation of r for scenarios where the true value was larger than the mode of the prior. This indicates that the data contained little information on this parameter. For the thornback ray case study this means that all inference is conditional on the chosen prior distribution for r. Parameters σ and τ were not really estimable but their values did not influence the estimates of other parameters. The SE protocol was created to mimic the available data. Hence, twenty nine years of biomass indices were used and the influence of the duration of the time series was not explored. SE results for the case with limited data (depletion index for the last year only, scenarios I and II) showed that this can change the overall trend of the estimated biomass trajectory, as was indeed found for the thornback ray case study. More work would be required to determine the minimum length of a research vessel CPUE time series needed to obtain reliable biomass estimates. However, model performance was comparable to the case of using no research vessel CPUE time series. This is encouraging for real world applications, if a depletion estimate is available with reasonable precision.

In summary, the SE approach provided general insights with respect to the applicability of the Bayesian SSM to thornback ray in the Bay of Biscay: possible overestimation of the carrying capacity and inference conditional on prior for intrinsic growth rate. Further, the model is not sensitive to the amount of process error (σ), observation error (τ) or uncertainty in landings (CV). However, the SE results should not be overinterpreted. First, the operating model was identical to the SSM used for estimation. Thus robustness to model misspecification was not tested. Second, for two parameters (q and $Y0$) only one value was used for the simulations.

4.2 Thornback ray in the Bay of Biscay

With the decline in catches of stocks traditionally targeted, elasmobranchs have become increasingly exploited by commercial fisheries leading to global catches peaking in the early 2000s (Davidson et al. 2015). Despite this, only few stock assessments of elasmobranchs stocks have been carried out (Bonfil 1994; Agnew et al. 2000). Some attempts to assess thornback ray stocks have been made (Heessen 2003), but none to the best of our knowledge for the Bay of Biscay mainly due to lack of data. In this study we tried to overcome this situation by creating a hypothetical landings time series and by using a Bayesian approach which made it possible to integrate biological knowledge in addition to the data. The hypothetical landings time series is of course highly uncertain but we partially accounted for this by modelling errors in catches (Eq. (4)). Further, the alternative catch time series was of similar magnitude for the last four decades, hence would have given similar results for that period.

Assuming the hypothetical landings were reasonable, a condition relaxed below, the results indicated that the Bay of Biscay population of thornback ray has steadily declined since the start of the 20th Century. Using all available data (FULL run), the current biomass was estimated as 7% of the carrying capacity and B_{MSY} as 32 000 tonnes (50% of K). Cope et al. (2015) derived an empirical relationship between species vulnerability to fishing and depletion rate of (more or less) unmanaged stocks for the US Pacific coast. For thornback ray in the Celtic Sea, McCully et al. (2015) estimated a vulnerability score of 1.61. For this level of vulnerability the empirical relationship of Cope et al. (2015) gives a depletion rate of 0.54. If our estimates are correct, the depletion of thornback ray in the Bay of Biscay is rather severe, suggesting either vulnerability to fishing is higher than in the Celtic Sea or that unmanaged fishing pressure was much higher than on US Pacific coast, or both. The posterior predictive check revealed that this scenario of high depletion is plausible but not the most likely to happen with the biological parameters estimated with the model. However, if the research vessel CPUE time series is trustworthy, the estimated biomass trajectory should be reliable. Further, the FULL and SHORT runs led to very similar biomass trajectories and parameter estimates which indicated that the estimated depletion state was primarily driven by the research vessel CPUE and not too dependent on uncertain historical landings. Hence, the hypothetical landings need only be reasonable for the period 2000 to 2013 for the estimates for thornback ray in the Bay of Biscay to be considered reliable. Trials showed that the estimation did not converge when the time series was reduced to the period for which no landings had to be reconstructed (2010–2013).

In all runs except the one using only a catch time series, the estimated biomass of thornback ray in the Bay of Biscay was below B_{MSY} during the whole second half of the 20th Century. Even the most optimistic estimates (run with landings only) showed exploitation above MSY during 35 years in the 20th Century. The FULL run assessment is in line with previous evaluations of a decline of the thornback ray population in the Bay of Biscay (Maurin 1994). There is historical evidence for the recovery potential of thornback ray (Letaconnoux 1948) but in the Bay of Biscay, the potential for a fast rate of recovery seems compromised by the low value of r.

In the Bay of Biscay, the only management of thornback ray is the Total Allowable Catch (TAC) set for all Rajiformes

combined for the Bay of Biscay and West Iberian waters (ICES subarea VIII and IX). This TAC was introduced in 2009 only. Our study of the biomass trajectory until 2014 therefore covers primarily the period where thornback ray was not subject to TAC management but only to fisheries levels management such as gear and effort regulations.

Survey catchability of thornback ray in the Bay of Biscay was estimated to be around 0.13. For the morphologically similar spotted ray (*Raja montagui*), Fraser et al. (2007) estimated a catchability of 0.15 in the International Bottom Trawl Survey in the North Sea which uses a similar bottom trawling as in the survey used here. Thus the estimated catchability seems reasonable.

The application of the SSM to thornback ray involved several strong assumptions for creating the landings time series but also for model formulation. By assuming that the time series of thornback ray landings followed that of all skates and rays, it was implied that the proportion of thornback ray in the landings of skates and rays remained constant. Larger skates such as white skate (*Rostroraja alba*) and those of the genus *Dipturus* are known to have severely declined (Dulvy et al. 2000; Dulvy and Reynolds 2002), which may have induced competitive release for medium-sized species such as thornback ray. However, at the same time the cuckoo ray (*Leucoraja naevus*), which represented about 3/4 of total landings of skates and rays in the Bay of Biscay in the last 10 years (ICES 2015), might have more benefited from this competitive release as it is a small bodied species, whose spatial distribution and depth range strongly overlaps with that of larger skates. As a result, the proportion of thornback ray in total ray landings could have been stable, as assumed here.

Carrying capacity K and intrinsic growth rate r were considered constant during the 20th and early 21st Century, that is during 111 years. This assumption was necessary in the absence of any information regarding how they might have varied. The intrinsic growth rate is the result of reproduction, individual body growth and natural mortality (Hoening and Gruber 1990), processes which are likely to vary with stock density and environmental conditions. Given their low and rather constant fecundity, elasmobranches may have low compensatory capacity (Kindsvater et al. 2016). Moreover, temperature in the Bay of Biscay has varied during the second half of the 20th Century leading to a change in environmental conditions in the species' habitat (Michel et al. 2009). However, the Bay of Biscay is central in the latitudinal range of the thornback ray, so that temperature changes should not have evolved beyond what is suitable for the species. In contrast, natural mortality may have changed in either direction as a consequence of changes in biotic interactions driven by the overexploitation of numerous fish species and the depletion of some apex predators (Lorance et al. 2009). The population growth rate and carrying capacity may also have been impacted by long-term changes in benthic communities driven by fishing (Hiddink et al. 2011). However, these aspects are unknown for benthic-feeding rays. Changes in species composition of the benthos in favor of scavengers (Rumohr and Kujawski 2000) such as gastropods may have increased predation rates on ray and skate eggs (Lucifora and Garcia 2004). Lastly habitats in the Bay of Biscay have been modified by fishing and other human pressures, such as eutrophication in coastal waters (Lorance et al. 2009) but the effect of these habitat changes on skates and rays is unknown. In summary, carrying capacity and intrinsic growth rate of thornback rays in the Bay of Biscay might have varied or might have remained stable during the last century, we simply don't know.

All models are restricted by data availability and represent a trade-off between accuracy and complexity. The Bayesian state space biomass production model provided a useful framework for data integration in the case of data-poor stocks. In recent years, the amount of biological information has increased on growth, migration, mortality and spatial distribution of a number of ray and shark stocks in the Northeast Atlantic (e.g. Walker 1997; Hunter et al. 2005; Maxwell et al. 2009; Ellis et al. 2011; McCully et al. 2012; Martin et al. 2012). These data could be used to attempt estimating stock status of other thornback ray stocks and those of other elasmobranchs. Further, the development of a multispecies version of the model to assess simultaneously several elasmobranch stocks in a given area could be considered. Rajids species are closely associated ecologically and it is impossible to target species completely separately.

Acknowledgements. We acknowledge funding from the French "Agence Nationale de la Recherche" (ANR) for the GenoPopTaille project. FM thanks Ifremer for a Ph.D. studentship. We would like to thank three anonymous reviewers for constructive comments which helped to improve the manuscript.

References

Agnew D.J., Nolan C.P., Beddington J.R., Baranowski R., 2000, Approaches to the assessment and management of multispecies skate and ray fisheries using the Falkland Islands fishery as an example. Can. J. Fish. Aquat. Sci. 57, 429–440.

Bonfil R., 1994, Overview of world elasmobranch fisheries, FAO fisheries technical paper. FAO, Rome.

Camhi M., Pikitch E.K., Babcock E.A. (Eds.), 2008, Sharks of the open ocean: biology, fisheries and conservation, Fish and aquatic resources series. Blackwell Science, Oxford, Ames, Iowa.

Chevolot M., Hoarau G., Rijnsdorp A.D., Stam W.T., Olsen J.L., 2006, Phylogeography and population structure of thornback rays (*Raja clavata* L., Rajidae). Mol. Ecol. 15, 3693–3705.

Cope J.M., Thorson J.T., Wetzel C.R., DeVore J., 2015, Evaluating a prior on relative stock status using simplified age-structured models. Fish. Res. 171, 101–109.

Davidson L.N.K., Krawchuk M.A., Dulvy N.K., 2015, Why have global shark and ray landings declined: improved management or overfishing? Fish Fish. 17, 2, 438–458, June 2016.

Dorel D., Cadiou Y., Porcher P., 1998, Poissons, crustacés et mollusques des mers communautaires. Paramètres biologiques et représentations graphiques.

Du Buit M.H., 1974, Contribution à l'étude des populations de raies du Nord-Ouest Atlantique des Faeroe au Portugal (Sciences Naturelles). Université de Paris VI, Paris.

Dulvy N.K., Reynolds J.D., 2002, Predicting extinction vulnerability in skates. Conserv. Biol. 16, 440–450.

Dulvy N.K., Metcalfe J.D., Glanville J., Pawson M.G., Reynolds J.D., 2000, Fishery stability, local extinctions, and shifts in community structure in skates. Conserv. Biol. 14, 283–293.

Dulvy N.K., Fowler S.L., Musick J.A., Cavanagh R.D., Kyne P.M., Harrison L.R., Carlson J.K., Davidson L.N., Fordham S.V., Francis M.P., Pollock C.M., Simpfendorfer C.A., Burgess G.H., Carpenter K.E., Compagno L.J., Ebert D.A., Gibson C., Heupel M.R., Livingstone S.R., Sanciangco J.C., Stevens J.D., Valenti S., White W.T., 2014, Extinction risk and conservation of the world's sharks and rays. Elife. eLife.00590 3.

EC, 2009, COUNCIL REGULATION (EC) No. 43/2009 of 16 January 2009 fixing for 2009 the fishing opportunities and associated conditions for certain fish stocks and groups of fish stocks, applicable in Community waters and, for Community vessels, in waters where catch limitations are required.

Ellis J.R., Morel G., Burt G., Bossy S., 2011, Preliminary observations on the life history and movements of skates (Rajidae) around the Island of Jersey, western English Channel. J. Mar. Biol. Assoc. UK 91, 1185–1192.

Enever R., Catchpole T.L., Ellis J.R., Grant A., 2009, The survival of skates (Rajidae) caught by demersal trawlers fishing in UK waters. Fish. Res. 97, 72–76.

Figueiredo I., Moura T., Bordalo-Machado P., Neves A., Rosa C., Serrano Gordo L., 2007, Evidence for temporal changes in ray and skate populations in the Portuguese coast (1998–2003) – its implications in the ecosystem. Aquat. Living Resour. 20, 85–93.

Fraser H.M., Greenstreet S.P.R., Piet G.J., 2007, Taking account of catchability in groundfish survey trawls: implications for estimating demersal fish biomass. ICES J. Mar. Sci. 64, 1800–1819.

Frisk M.G., Miller T.J., Dulvy N.K., 2005, Life histories and vulnerability to exploitation of elasmobranchs: inferences from elasticity, perturbation and phylogenetic analyses. J. Northwest Atl. Fish. Sci. 35, 27–45.

Gelman A., Carlin J.B., Stern H.S., Dunson D.B., 2014, Bayesian data analysis, Chapman & Hall/CRC texts in statistical science. 3rd edn., CRC Press, Boca Raton.

Geweke J., 1992, Statistics, probability and chaos: comment: inference and prediction in the presence of uncertainty and determinism. Stat. Sci. 7, 94–101.

Hammond T., Trenkel V., 2005, Censored catch data in fisheries stock assessment. ICES J.Mar. Sci.

Heessen H.J.L., 2003, Development of elasmobranchs stock assessment DELASS. Final report of DG Fish Study Contract 99/055.

Hiddink J.G., Johnson A.F., Kingham R., Hinz H., 2011, Could our fisheries be more productive? Indirect negative effects of bottom trawl fisheries on fish condition: Effects of bottom trawls on fish condition. J. App. Ecol. 48, 1441–1449.

Hilborn R., Walters C.J., 1992, Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York.

Hoening J.M., Gruber S.H., 1990, Life-history patterns in the elasmobranchs: implications for fisheries management (NOAA Technical Report No. NMFS 90), in Elasmobranchs as Living Resources: Advances in the Biology, Ecology, Systematics, and the Status of the Fisheries.

Holden M.J., 1972, The growth rates of *Raja brachyura*, *R. clavata* and *R. montagui* as determined from tagging data. ICES J. Mar. Sci. 34, 161–168.

Holden M.J., 1975, The fecundity of *Raja clavata* in British waters. ICES J. Mar. Sci. 36, 110–118.

Hunter E., Buckley A.A., Stewart C., Metcalfe J.D., 2005, Migratory behaviour of the thornback ray, *Raja clavata*, in the southern North Sea. J. Mar. Biol. Assoc. UK 85, 1095.

ICES, 2014a, Report of the Working Group on Elasmobranchs Fishes (WGEF) (No. ICES CM 2014/ACOM:19). Portugal, ICES, Lisbon.

ICES, 2014b, Thornback ray – ICES-FishMap.

ICES, 2015, Report of the Working Group on Elasmobranchs Fishes (WGEF) (No. ICES CM 2015/ACOM:19). Portugal, ICES, Lisbon.

IUCN, 2005, *Malacoraja clavata*: Ellis, J.: The IUCN Red List of Threatened Species 2005: e.T39399A10199527.

Kimura D.K., Tagart J.V., 1982, Stock Reduction Analysis, Another Solution to the Catch Equations. Can. J. Fish. Aquat. Sci. 39, 1467–1472.

Kimura D.K., Balsiger J.W., Ito D.H., 1984, Generalized stock reduction analysis. Canadian J. Fish. Aqua. Sci. 41, 1325–1333.

Kindsvater, H.K., Mangel, M., Reynolds, J.D., Dulvy, N.K., 2016. Ten principles from evolutionary ecology essential for effective marine conservation. Ecol. Evol. 6, 2125–2138.

Letaconnoux R., 1948, Effets de la guerre sur la constitution des stocks de poissons. Rapports et Procès-Verbaux des Réunions du Conseil Permanent International pour l'Exploration de la Mer 122, 55–62.

Lorance P., Bertrand J.A., Brind'Amour A., Rochet M.-J., Trenkel V.M., 2009, Assessment of impacts from human activities on ecosystem components in the Bay of Biscay in the early 1990s. Aquat. Living Resour. 22, 409–431.

Lucifora L.O., Garcia V.B., 2004, Gastropod predation on egg cases of skates (Chondrichthyes, Rajidae) in the southwestern Atlantic: quantification and life history implications. Mar. Biol. 145, 917–922.

Martin C.S., Vaz S., Ellis J.R., Lauria V., Coppin F., Carpentier A., 2012, Modelled distributions of ten demersal elasmobranchs of the eastern English Channel in relation to the environment. J. Exp. Mar. Biol. Ecol. 418-419, 91–103.

Maurin H., 1994, Inventaire de la faune menacée en France. Muséum national d'Histoire naturelle, Paris.

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

McAllister M.K., Ianelli J.N., 1997, Bayesian stock assessment using catch-age data and the sampling - importance resampling algorithm. Can. J. Fish. Aqua. Sci. 54, 284–300.

McAllister M.K., Pikitch E.K., Babcock E.A., 2001, Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. Can. J. Fish. Aqua. Sci. 58, 1871–1890.

McCully S.R., Scott F., Ellis J.R., 2012, Lengths at maturity and conversion factors for skates (Rajidae) around the British Isles, with an analysis of data in the literature. ICES J. Mar. Sci. 69, 1812–1822.

McCully Phillips S.R., Scott F., Ellis J.R., 2015, Having confidence in productivity susceptibility analyses: A method for underpinning scientific advice on skate stocks? Fish. Res. 171, 87–100.

Michel S., Vandermeirsch F., Lorance P., 2009, Evolution of upper layer temperature in the Bay of Biscay during the last 40 years. Aquat. Living Resour. 22, 447–461.

Ono K., Punt A.E., Rivot E., 2012, Model performance analysis for Bayesian biomass dynamics models using bias, precision and reliability metrics. Fish. Res. 125-126, 173–183.

Parent E., Rivot E., 2013, Introduction to hierarchical Bayesian modeling for ecological data, Chapman & Hall/CRC applied environmental statistics. CRC Press, Boca Raton.

Pauly D., Zeller D., 2016, Sea Around Us?: concepts, design and data (seaaroundus.org) Available at: http://www.seaaroundus.org (accessed 2.24.16).

Plummer M., Best N., Cowles K., Vines K., 2006, CODA: Convergence Diagnosis and Output Analysis for MCMC. R News 6, 7–11.

Poulard J., Blanchard F., 2005, The impact of climate change on the fish community structure of the eastern continental shelf of the Bay of Biscay. ICES J. Mar. Sci. 62, 1436–1443.

Poulard J.-C., Trenkel V.M., 2007, Do survey design and wind conditions influence survey indices? Can. J. Fish. Aqua. Sci. 64, 1551–1562.

Punt A.E., Hilborn R., 1997, Fisheries stock assessment and decision analysis: the Bayesian approach. Rev. Fish Biol. Fish. 7, 35–63.

Punt A.E., Su N.-J., Sun C.-L., 2015, Assessing billfish stocks: a review of current methods and some future directions. Fish. Res. 166, 103–118.

Quéro J.C., Cendrero O., 1996, Incidence de la pêche sur la biodiversité ichtyologique marine.

Quéro J.C., Guéguen J., 1981, Les raies de la mer celtique et du canal de Bristol. Abondance et distribution. Sci. Pêche. 318, 1–22.

Quéro J.C., Vayne J.J., 2005, Les poissons de mer des pêches françaises. Delachaux et Niestlé, Paris.

R Foundation for Statistical Computing, 2008, R Development Core Team (2008). R: A language and environment for statistical computing. Austria, Vienna.

Robert M., Faraj A., McAllister M.K., Rivot E., 2010, Bayesian state-space modelling of the De Lury depletion model: strengths and limitations of the method, and application to the Moroccan octopus fishery. ICES J. Mar. Sci. 67, 1272–1290.

Rochet M., Péronnet I., Trenkel V.M., 2002, An analysis of discards from the French trawler fleet in the Celtic Sea. ICES J. Mar. Sci. 59, 538–552.

Rumohr H., Kujawski, T., 2000, The impact of trawl fishery on the epifauna of the southern North Sea. ICES J. Mar. Sci. 57, 1389–1394.

Serra-Pereira B., Figueiredo I., Gordo L.S., 2011, Maturation, fecundity, and spawning strategy of the thornback ray, *Raja clavata*: do reproductive characteristics vary regionally? Mar. Biol. 158, 2187–2197.

Silva J.F., Ellis J.R., Catchpole T.L., 2012, Species composition of skates (Rajidae) in commercial fisheries around the British Isles and their discarding patterns. J. Fish Biol. 80, 1678–1703.

Steven G.A., 1931, Rays and Skates of Devon and Cornwall. Methods of Rapid Identification on the Fishmarket. J. Mar. Biol. Assoc. UK 17, 367.

Thomas A., O Hara B., Ligges U., Sturtz S., 2006, Making BUGS Open. R News 12–17.

Trenkel V.M., Bravington M.V., Lorance P., Walters C., 2012, A random effects population dynamics model based on proportions-at-age and removal data for estimating total mortality. Can. J. Fish. Aquat. Sci. 69, 1881–1893.

Walker P., 1997, Distribution, movement and stock structure of three ray species in the North Sea and eastern English Channel. ICES J. Mar. Sci. 54, 797–808.

Walker P., Hislop J.R.G., 1998, Sensitive skates or resilient rays? Spatial and temporal shifts in ray species composition in the central and north-western North Sea between 1930 and the present day. ICES J. Mar. Sci. 55, 392–402.

Walker T., 1992, Fishery simulation model for sharks applied to the Gummy shark, *Mustelus antarcticus* Gunther, from Southern Australian waters. Mar. Freshw. Res. 43, 195.

Whittamore J.M., McCarthy I.D., 2005, The population biology of the thornback ray, *Raja clavata* in Caernarfon Bay, North Wales. J. Mar. Biol. Assoc. UK 85, 1089.

Wiegand J., Hunter E., Dulvy N.K., 2011, Are spatial closures better than size limits for halting the decline of the North Sea thornback ray, *Raja clavata*? Mar. Freshw. Res. 62, 722.

CHAPITRE 4

Contourner l'absence de données démographiques : évaluation simultanée de l'état de plusieurs stocks de raies

Dans ce chapitre, l'évaluation de l'état des populatison d'élasmobranches par méthodes démographiques est poussé plus loin que dans le chapitre 3. Pour ce faire, un modèle Bayésien multispécifique d'estimations de l'abondance a été développé et appliqué aux principales espèces de raies du Golfe de Gascogne. Cette partie a fait l'objet d'un article soumis à l'ICES Journal of Marine Science.

1. Quelles espèces de raies en Atlantique Nord-Est ?

Plus de 500 espèces de raies sont répertoriées dans le monde (Dulvy et al., 2014) dont 44 en Atlantique Nord-Est (Quéro et al., 2003), et seulement 28 dans les eaux atlantiques françaises (Quero et al., 2017).

Dans le Golfe de Gascogne (zone ICES 8abd), 16 espèces de raies ont été débarquées entre 2005 et 2016 (ICES, 2017) et six d'entre elles répresentent la majorité des débarquements récents (voir annexe A) : la raie bouclée (*Raja clavata*), la raie chardon (*Leucoraja fullonica*), la raie circulaire (*Leucoraja circularis*) la raie douce (*Raja montagui*), la raie fleurie (*Leucoraja naevus*) et la raie lisse (*Raja brachyura*). Ces espèces partagent un cycle de vie similaire malgré quelques différences notamment dans leur longévité et âge de maturité. La figure 4.1 regroupe et compare quelques données biologiques disponibles pour ces six espèces. Elle met ainsi en avant le manque d'information pour la raie chardon et la raie circulaire. De plus, la fiabilité de ces données peut être discutée. Il est par exemple difficile d'échantillonner des individus de grandes tailles et donc de dériver des paramètres de croissance fiables.

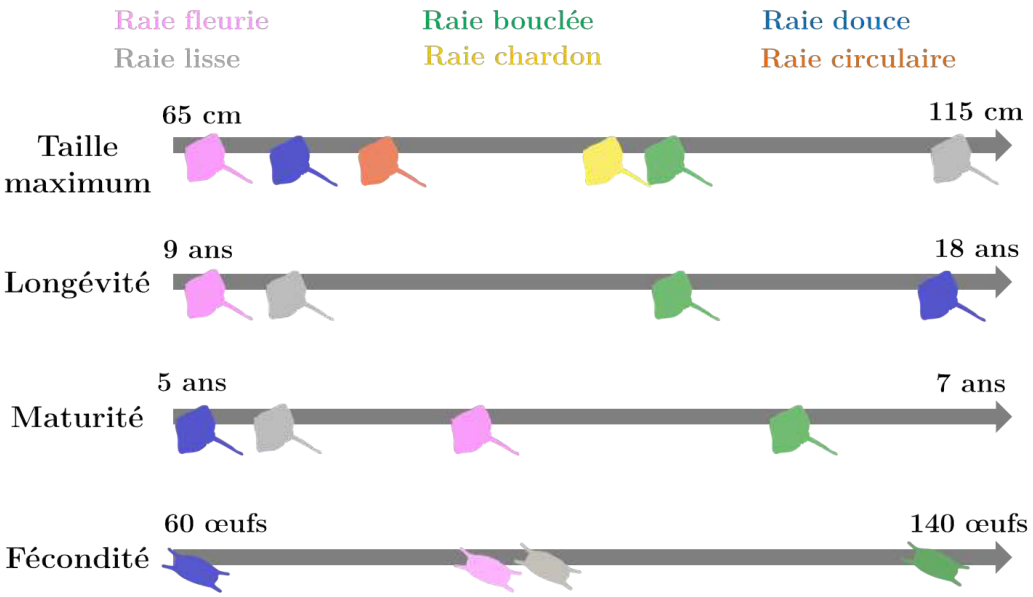


FIGURE 4.1 : Paramètres biologiques des six espèces principales débarquées dans le Golfe de Gascogne entre 2005 et 2016 (Quéro and Vayne, 1997)

Ainsi pour plusieurs espèces de raies du Golfe de Gascogne, il est impossible d'appliquer un modèle spécifique tel que celui développé dans le chapitre 3 à cause du manque de données biologiques mais aussi à cause du manque de données issues des pêches commerciales et scientifiques. Pour pallier à ce problème, un modèle Bayésien d'évaluation multispécifique d'abondance a été devéloppé. Appliqué aux six espèces de raies principales du Golfe de Gascogne, il permet d'estimer leurs trajectoires de biomasses et l'évolution de la composition spécifique du Golfe.

2. Article : Détermination des changements dans la communauté de raies du golfe de Gascogne via l'utilisation de données non spécifiques

Florianne Marandel[1], Pascal Lorance [1] and Verena M. Trenkel[1]

Soumis à ICES Journal of Marine Sciences - 13 Avril 2018
DOI : A venir

RESUME - Une gestion soutenable des ressources halieutiques nécessite la connaissance du statut des populations et communautés exploitées. Certaines espèces accessoires comme les raies sont gérées ensemble, les menant à la surexploitation voire même à des problèmes de conservation pour les espèces les plus vulnérables. Les méthodes traditionnelles d'évaluation de stock nécessitent des données spécifiques, donnéees disponibles seulement récemment pour les raies en Europe. Pour pallier à cette situation, un modèle Bayésien multispécifique de production de biomasse a été développé. Il utilise les debarquements non spécifiques disponibles ainsi que les indices de biomasses et débarquements spécifiques nouvellement disponibles. L'approche a été appliquée aux six espèces de raies principales du Golfe de Gascogne et a identifié des changements dans la composition spécifique du Golfe depuis les années 50. La composition spécifique en 1950 aurait été proche de celle avant exploitation. Depuis les années 90, *Leucoraja naevus* est devenue de plus en plus dominante, alors que les autres espèces déclinaient. Les abondances de *L. fullonica* et *L. circularis* ont beaucoup diminué suggérant une surexploitation chronique. Avec une analyse prospective basée sur ces résultats, nous avons mis en évidence que l'abondance de toutes les espèces sauf deux sont supposées augmenter avec les taux d'exploitation actuels. L'utilisation d'un modèle Bayésien multispécifique a permis d'estimer les changements historiques dans la communauté de raies du Golfe de Gascogne mais aussi d'estimer les trajectoires individuelles des espèces. Actuellement, les six espèces étudiées sont gérées sous un seul quota. Néanmoins, comme montré dans ce travail, les taux d'exploitation actuels ne semblent pas permettre aux espèces qui ont le plus décliné de récupérer dans les années à venir.

[1] Ifremer, rue de l'île d'Yeu, BP 21105, 44311 Nantes Cedex 3, France

Soumis à l'ICES Journal of Marine Sciences

Determining long-term ray community changes using grouped landings and limited species data

Florianne Marandel[1], Pascal Lorance[1] and Verena M. Trenkel[1]
[1] Ifremer, rue de l'île d'Yeu, BP 21105, 44311 Nantes Cedex 3, France

Abstract – Sustainable fisheries management requires the evaluation of the status of exploited populations and communities. Certain bycatch species such as rays are managed together which has led to overexploitation and conservations concerns for the most vulnerable ones. Traditional fisheries stock evaluation methods need species-specific input data while for ray species-specific landings have become available only recently in Europe. To overcome this limitation, we developed a Bayesian multispecies biomass production model. In addition to the grouped landings the input data are shorter time series with species-specific information (landings and biomass indices). Applying the approach to the six main ray species managed together in the Bay of Biscay (Northeast Atlantic) we identified long-term changes in community composition. The species composition in 1950 might have been close to what would be expected without fishing. Since the 1990s *Leucoraja naevus* became increasingly dominant, while the contributions of the other five species declined. The abundances of *Leucoraja fullonica* and *L. circularis* strongly decreased which suggests long-term overexploitation. In a prospective analysis using the model results, we found that all but two species are expected to increase over the next decade under the current harvest rates. Applying a Bayesian multispecies modelling approach made it possible to evaluate long-term changes in the Bay of Biscay ray community and gain insight into individual population trajectories. Currently the six studied ray species are managed by a single quota. However, as shown here the current harvest level is unlikely to permit the most depleted species to recover during the next decade.

Key words: Community changes; Multispecies model; Skates and rays; Species composition; Stock assessment;

1 Introduction

A major objective of ecosystem-based fisheries management is to maintain species and size diversity within exploited ecosystems (Livingston *et al.* 2005). Rajiformes, hereafter "rays", is a diversified order of the class Elasmobranchii (sharks and rays, here after elasmobranchs), which embody a large diversity of morphological and life history traits including 382 species out of 1209 elasmobranchs placing them in a key position for the functioning of marine food webs (WoRMS Editorial Board 2017; Dulvy *et al.* 2017). However, elasmobranchs present relatively low fecundity and high sensibility to overfishing (Stevens *et al.* 2000) as these traits are linked to the potential for population recovery (Hutchings & Kuparinen 2017).

Due to a lack of fishing limits, but also climate change and habitat loss, several elasmobranch populations have declined, sometimes strongly, around the world during the 20th century (Quéro & Cendrero 1996; Dulvy *et al.* 2014, 2017). For example, the European Union (EU) introduced catch quotas for Rajiformes in the North Sea in 2004 and these were generalized to all EU waters only in 2009 (CEC 2003, 2009). This lack of management has led to changes in the species composition. In the Southern North Sea, the elasmobranch community has shifted from an important occurrence of larger species to smaller, more productive species during the 20th century, though with some recent reversal of changes (Sguotti *et al.* 2016). In the Adriatic Sea, a change in elasmobranch community composition over the last half century has also been found (Ferretti *et al.* 2013), the main driver being overexploitation (Barausse *et al.* 2014). In both areas the large bodied *Squatina squatina* (angel shark) has been extirpated (Ferretti *et al.* 2013; Sguotti *et al.* 2016). This species was once also rather common in the Bay of Biscay (Quéro & Cendrero 1996). Elasmobranch community changes have also been observed in the Gulf of Mexico which Shepherd & Myers (2005) hypothesized were driven both by direct effects of fishing, mainly as bycatch, and indirect effects caused by predation release due to the decrease of larger sharks. Evidence for the effect of shark predation release on rays has also been found in the Western Atlantic (Myers *et al.* 2007). Globally, predation release seems to have been an important driver for increases in certain ray species, with reduced fishing mortality playing a smaller role (see review in Ward-Paige *et al.* (2012)).

Except for rare cases with long survey time series (e.g. North Sea, Sguotti et al. 2016), estimation of long-term time trends of rays has to rely on fishery-dependent data which

unfortunately is hampered by low abundances commonly resulting in insufficient observations (Kuhnert *et al.* 2011). Moreover, in European waters rays are subject to mandatory species-specific landings statistics only since 2009 (CEC 2009). Prior to this, landings were generally reported as "miscellaneous rays and skates" if not "miscellaneous sharks and rays". Aiming to investigate long term changes in the Bay of Biscay ray community in such a data limited situation we developed a multispecies biomass model that uses historic landings aggregated across species and more recent species-specific landing and survey time series to estimate long-term time trends. The results provide insights for fisheries management and the method has potential for being applied elsewhere.

In the Bay of Biscay during the past decade, four ray species represented the majority of commercial landings: *Raja brachyura*, blonde ray, *Raja clavata*, thornback ray, *Raja montagui*, spotted ray and *Leucoraja naevus,* cuckoo ray (ICES 2016). The three former species were also common in landings in the 19th century (Quéro & Cendrero 1996). In contrast, *L. naevus* which is distributed more offshore may not have been much exploited before the 20th century. *R. brachuyra* and *R. clavata* are currently classified as Near Threatened and the other two as Least Concern by the International Union for Conservation of Nature (IUCN 2005, 2007, 2008, 2014b), the major threat coming from target and bycatch fisheries (ICES 2016). *R. clavata* most likely severely declined during the 20th century (Marandel, Lorance & Trenkel 2016). Two other species are also present in recent landings with indications for higher past abundances at least in the northern part of the Bay of Biscay (Du Buit 1974), *Leucoraja circularis* (sandy ray) and *Leucoraja fullonica* (shagreen ray). *L. circularis* is considered Endangered and *L. fullonica* Vulnerable by the IUCN (IUCN 2014a; c). These six species were therefore considered in this study. Other Rajiformes occurring on the Bay of Biscay shelf and upper slope are *R. undulata* (undulate ray), *R. microocellata* (small-eyed ray), *Rostroraja alba* (white skate), *Dipturus batis* complex (blue skate composed of two species) and *D. oxyrinchus* (longnose skate). These other species were not considered because they have been subject to landing bans since 2009, occur at low abundance or data are even more uncertain than for the six main species which were studied here.

All six modelled species are small- to medium-sized rays, with *L. naevus* and *R. montagui* having maximum size around 80 cm and the other species 120-150 cm. *R. montagui* and *R. brachyura* are found in shallow waters, while *L. naevus* and *L. fullonica* are found offshore and *L. circularis* is more restricted to the shelf break and upper continental slope. Adults and juveniles of *R. clavata* occur in coastal waters including estuaries (Rousset 1990; Courrat *et al.* 2009); adults are also found on the shelf and down to the upper slope, where occurrence of juveniles is uncertain. In the remainder of this article the assemblage containing the six ray species will be referred to as ray community of the Bay of Biscay.

To account for uncertainty in landings and survey time series as well as life history information we used a Bayesian approach to fit the multispecies model. We explored the effect of different data combinations and variants of prior information on species and composition time trends. The main aim was to obtain robust estimates of time trends in a data limited situation which is rather common for species which are not the main target of fisheries and hence do not benefit from appropriate sampling and recording programs. Projecting populations' forwards under different harvest regimes we showed that unless species-specific management measures are adopted for the severely depleted *Leucoraja fullonica* and *circularis* and hence community structure cannot be expected to recover. The other species are expected to continue to increase under current harvest rates.

2 Material and methods

2.1 Data

All data are available in supporting information (Table A.3).

2.1.1. Commercial landings

Landings for *L. circularis, L. fullonica, L. naevus*, *R. brachyura, R. clavata*, and *R. montagui* as well as grouped landings recorded as Rajidae, Rajiformes, and *Raja* spp. were extracted for the Bay of Biscay (ICES divisions 8abd) for the period 1950 to 2016 (ICES 2011, 2016, 2017). Landings were missing for some countries and years. These were filled in by interpolation or extrapolation. Some recent landings were updated using landings corrected by experts (see Appendix A.1).

The species composition of grouped landings has changed since 1950. As only the currently most abundant six species were considered here, a rough estimate of historically important species such as *D. batis* were removed from the grouped landings (Figure 1a); no corrections were made for other species (see below).

The rough landings estimate for *D. batis* was obtained using the following information. Descriptions of historical landings and catch per unit of effort (CPUE) suggested that after WWII, landings of *Dipturus* species may have amounted to around 25% of grouped landings (Letaconnoux 1948). Further, sampling of landings in 1966-69 in one harbour suggested that *Dipturus* may have contributed around 10% of landings from the northern Bay of Biscay (Du Buit 1974). In the 2000s, landings of these species were minor, hence the main decline occurred probably well before, as documented for the Irish Sea (Brander 1981). *Dipturus* landings have been banned in EU waters since 2010 (CEC 2010).

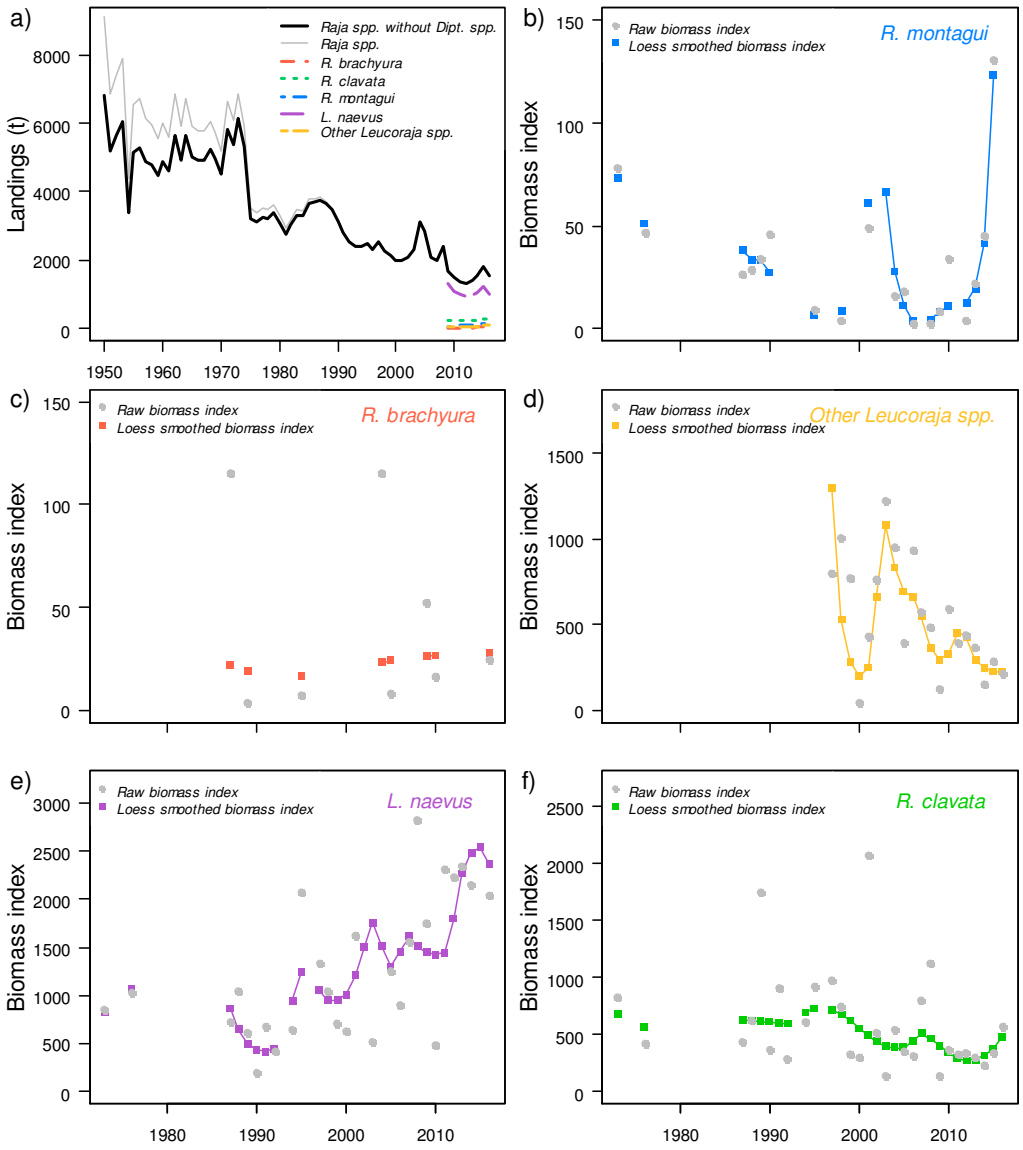


Figure 1: a) Total landings in tonnes for the Rajidae Species (Rajidae, Rajiformes, *Raja* spp, *L. naevus, L. circularis, L. fullonica, R. clavata, R. brachyura and R. montagui*) from (ICES 2011, 2016, 2017) and specific landings for *R. brachyura*, *R. clavata*, *R. montagui*, *L. naevus* and the other *Leucoraja spp.* from (ICES 2016). Survey biomass indices and smoothed indices for (b) *R. montagui*, (c) *R. brachyura*, (d) other *Leucoraja spp., (e) L. naevus* and (f) *R. clavata*.

Corrections for other species were not considered for the following reasons. Landings of *R. undulata* were banned from 2009 to 2014 and only small quotas have been allowed in more recent years (EC, 2009, 2016). Its contribution to landings prior to 2009 is unknown as it was never reported separately. The species occurs in shallow waters only and is exploited mostly by small-scale coastal fisheries. Thus it is unclear whether landings from the 1950s were properly included in official landings statistics or not, hence whether a correction might be warranted. *R. microocellata* occurs locally along the coast at low abundance, but it may have been locally "rather common" in the past (Quéro and Cendrero, 1996). *Rostroraja alba* was mentioned as having been "rather rare" in the past.

From 1950 to 2008 corrected total landings were used while after 2009 species-specific landings were considered reliable. Due to unreliable biomass indices for scientific surveys (see below), landings for *L. fullonica* and *L. circularis* were combined into other *Leucoraja spp.* landings.

2.1.2. Biomass indices

Biomass indices were calculated for each species using data from the scientific bottom trawl survey (EVHOE) in the Bay of Biscay from 1987 to 2016 (with missing years), for a survey description see Poulard & Trenkel (2007) (Figure 1b-f). For *L. naevus* and the three *Raja* species, the time series was

expanded using data from 1973 and 1976 collected with the same trawl and sampling plan (but limited to depths shallower than 200 m). For *L. fullonica* and *L. circularis*, as numbers caught in the survey were small, data were combined into a single other *Leucoraja spp.* index; further biomass indices were calculated using hauls from the Bay of Biscay and the adjacent Celtic Sea to cover the spatial distribution. As a consequence biomass indices of *L. naevus* and *Raja* species were available for the years 1973, 1976 and 1987-2016 and for other *Leucoraja spp.* for 1997-2016. Index calculation was based on first determining an area of occupancy (assuming a 100 km radius of influence for each haul) and then raising the average density to the occupied area. To remove observation variability indices were smoothed using a loess regression (the span was set to 0.4 for all species except for *R. brachyura* for which only 8 observations are available and so it was 2).

2.2 Modelling

2.2.1. Model

The state-space biomass production model is a multispecies extension of the discrete-time model developed for thornback ray (Marandel, Lorance & Trenkel 2016). The biomass dynamics of each species was assumed to follow

$$Y_{e,t+1} = (r_e + 1)Y_{e,t} - r_e Y_{e,t}^2 - \frac{C_{e,t}}{K_e} + \varepsilon_{e,t} \quad (1)$$

$$Y_{e,t} = \frac{B_{e,t}}{K_e} \qquad \varepsilon_{e,t} \sim N(0, \sigma^2)$$

where r_e is the intrinsic growth rate, K_e carrying capacity, $B_{e,t}$ biomass and $C_{e,t}$ are landings of species e in year t. The relative biomass in the first year is denoted $Y0_e$. To overcome the lack of species-specific landings before 2009, the probability of landings having been from species e at time t (t<2009) were assumed to correspond to its biomass proportion at time *t-1, i.e.* $p_{e,t} = \frac{B_{e,t-1}}{\sum_e B_{e,t-1}}$ and $\boldsymbol{p}_t=(p_{1,t}, \ldots p_{5,t})$. This assumption is supported by the fact that rays are generally bycaught by bottom trawlers, except the more coastal *R. brachyura* and *R. montagui* which are primarily landed by vessels using passive gears (V. Trenkel unpublished results). Landings by species were then drawn from a multinomial distribution using these probabilities:

$$\boldsymbol{C}_t \sim Multi(\boldsymbol{p}_t, C_{tot,t}) \quad (2)$$

where $C_{tot,t}$ are observed total landings of the six species and $\boldsymbol{C}_t$ is the vector of species-specific landings.

The observation model links the population biomass to biomass indices I

$$I_{e,t} = q_e B_{e,t} + \omega_{e,t} \quad (3)$$

$$\omega_{e,t} \sim N(0, {\tau^2}_e)$$

where q_e is a species-specific catchability coefficient.

Species-specific landings $l_{e,t}$ for 2009-2016 provided information on species biomass composition and were included in the likelihood function as

$$l_{e,t} \sim N(c_{e,t}, \rho^2) \quad t=2009,\ldots 2016 \quad (4)$$

where $c_{e,t} = \frac{B_{e,t-1}}{\sum_e B_{e,t-1}} C_{tot,t}$.

2.2.2 Model fitting and runs

The model was fitted using a Bayesian approach. Information on prior distributions and other technical issues can be found in the supporting information. In brief, informative priors were created for intrinsic growth rates r_e (equation 1) using life history parameters (McAllister, Pikitch & Babcock 2001) while uninformative priors were used for carrying capacities K_e. Values for the process variance σ^2 (equation 1) and the observation variances τ^2 (equation 3) and ρ^2 (equation 4) were assumed known (see Supplementary Material).

As no information was available concerning the initial biomass of each species in 1950, two sets of priors were tested for *Y0* parameters corresponding to two depletion hypotheses (BASE and EXPL). For the BASE run all species were assumed underexploited in 1950 (prior mode of *Y0*=0.8) (see Table A3). For the EXPL run all species were assumed fully exploited in 1950 (prior mode of *Y0*=0.5). In addition, to evaluate the impact of the *D. batis* total landings correction described above we ran a third run using a shorter time series during which this correction was less applied and biomass indices were available. Hence this SHORT run used landing from 1973 and assumed all species were fully exploited in 1973. Results are presented in details for run BASE assuming underexploitation in 1950, which was thought to be the most realistic. However, as posterior distributions for Y0 were identical to prior distributions (Fig A.9), we compared results from all runs but found no major differences; detailed results are provided in the supporting information (Supplementary material).

2.2.2. Analyses

Changes in community composition were studied in terms of the relative contribution of each species to total biomass. For each species relative biomass $Y_{e,t}$ trajectories were analysed. To evaluate the impact of fishing, harvest rates were calculated for each species by dividing estimated landings by estimated biomass in each year ($h_{e,t} = C_{e,t}/B_{e,t}$). Harvest rates where standardized by dividing them by the harvest rate leading to maximum sustainable yield (h_{MSY}=r/2).

To explore recovery potential under current (*status quo*) harvest rates, populations of all ray species except *L. naevus* (no recovery needed) were simulated forward during 10 years assuming constant harvest rates (mean for 2014-2016). As comparison, constant harvest rates set at 1/2 *status quo* and no harvest were simulated.

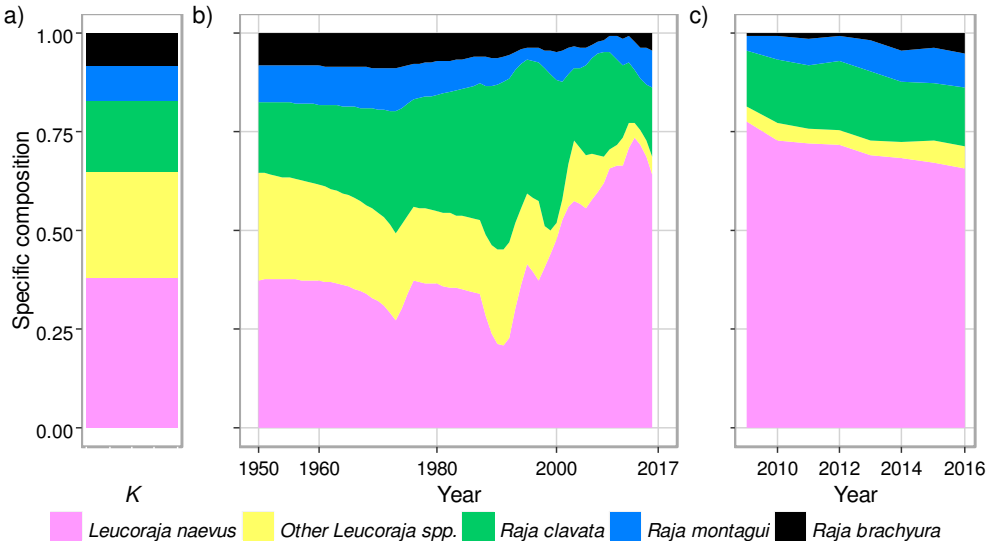


Figure 2: a) Species composition at carrying capacity for the five studied species; b) Mean estimated species composition between 1950 and 2017; c) Observed species composition in landings between 2009 and 2016.

3 Results

3.1 Community composition

Without exploitation populations will vary around their carrying capacity. In this situation the smallest species *L. naevus* would be the most abundant species (38% of total biomass), followed by the other *Leucoraja spp.* (27%) and *R. clavata* (17%) (Fig. 2a and Fig A.4). *R. brachyura* and *R. montagui* would present the smallest biomasses (8% and 9% respectively).

The mean estimated species composition in 1950 was similar to the unexploited situation (Fig. 2b). Between 1950 and 1991 the proportion of *L. naevus* remained globally constant while *R. clavata* increased to 41%, thus becoming the dominant species (in biomass). After this *L. naevus* increasingly dominated the species composition until reaching 63% of the community biomass in 2017. The proportions of *R. brachyura* and *R. montagui* continuously decreased until 2010 and increased thereafter. The proportion of other *Leucoraja spp.* continuously decreased during the studied period. The composition results were insensitive to assumptions regarding initial depletion and time series length (see Fig A.4 for results for the alternative runs).

In the landings, the proportion of *L. naevus* decreased slightly from 78% in 2009 to 66% in 2016, while those of all other species except *R. clavata* increased (Fig. 2c). The landings composition in 2009 was not identical to the community composition, with a higher proportion of other *Leucoraja spp*.

3.2 Population changes

The biomass of all species except *L. naevus* continuously decreased between 1950 and 2009 and increased during the most recent decade (Fig. 3; Fig A.2). The largest decrease was observed for other *Leucoraja spp.* (91%), followed by *R. brachyura* (69%), *R. montagui* and *R. clavata* (both 50%), while *L. naevus* only decreased by around 20% between 1950 and 2017. *L. naevus* first decreased between 1950 and 1991 (-2.1% per year) and increased thereafter (+18% per year). Its current biomass was estimated to be close to 64% of its carrying capacity (95 percentile interval: 46-83%). The decrease of *R. clavata* strongly accelerated after 1995 with the current biomass being at around 40% of carrying capacity (95 PI: 25-55%). *R. montagui* decreased until 1995 (-1.94% per year), initially strongly, and increased after 2008 (+12% per year); its current biomass is at 40% of carrying capacity (95 PI: 18-66%). *R. brachyura* followed a similar pattern with a strong decrease until 2009 (-1.65% per year) and increased thereafter; the current biomass is at 25% of carrying capacity (95 PI: 8-44%). Other *Leucoraja spp.* reached the lowest biomass in 2000 and increased to 8% (95 PI: 10-17%) in 2017.

Between 1950 and 1980 estimated landings for *R. clavata, L. naevus* and the other *Leucoraja spp.* were generally above 1 000 t, decreasing thereafter for all species (Fig 3). Until the early 1990s relative harvest rates indicated overexploitation for all species, in particular other *Leucoraja spp.* (values >1 in Fig. 3, Fig A.4). Harvest rates are at or below MSY values since 2011 except for the other *Leucoraja spp.*.

The highest carrying capacity was found for the small offshore *L. naevus*, followed by the other *Leucoraja spp.* and *R. clavata*. *R. brachyura* and *R. montagui* had similar estimated *K* (Fig A.6). However absolute values are uncertain as there was a negative correlation between estimates of *K* and survey catchability *q* as well as between *K* and *r* (Fig A.7).

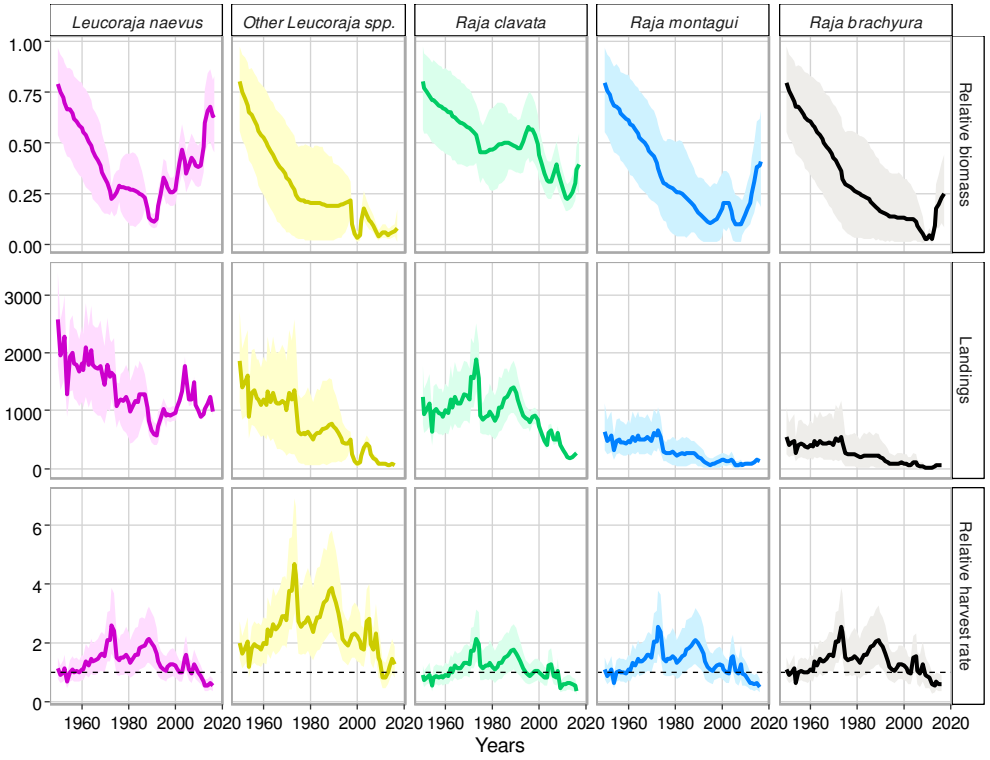


Figure 3: Estimated relative biomass between 1950 and 2017 (top row), landings between 1950 and 2016 (middle row) and relative harvest rates between 1950 and 2016 (bottom row) for the studied ray species (mean values with 95% percentile bands). Dotted horizontal lines for relative harvest rates correspond to maximum sustainable yield.

3.3 Management

Projecting the four depleted species (Y_{2017}<0.4) forwards indicated that under current harvest rates *R. clavata* and *R. montagui* are expected to increase above 50% carrying capacity (MSY biomass level) (Figure 4). Even without fishing other *Leucoraja spp.* are not expected to increase much by 2028 (from 8% in 2018 to 12% in 2028). Under *status quo* harvesting *R. brachyura* is predicted to increase to 41% (PI 16-66) of carrying capacity within ten years, halving the harvest rate will make it reach the same depletion level within around seven years.

4 Discussion

4.1 Community and population changes

Understanding changes in exploited communities is challenging for elasmobranchs because it requires disentangling the effects of exploitation, changes in predation and competition, and environmental changes, in a data limited situation (Ferretti *et al.* 2013). To overcome the data limitation, in particular the lack of species-specific landings for most of the time period, we developed a Bayesian approach simultaneously modelling all species. Applying this approach indicated that in the Bay of Biscay total biomass of the six most abundant ray species has declined since the 1950s. In 2017, only *L. naevus* was estimated to be above 50% of its carrying capacity; three species (*R. brachyura, R. montagui* and *R. clavata)* were estimated to be between 20 and 40% of their carrying capacity while the two other *Leucoraja spp.* were estimated to be fewer than 10%. The actual values are uncertain, but the relative order of species should be more robust. Comparable composition changes have been observed in the Irish Sea, where *L. naevus* strongly increased and *R. clavata* and *R. montagui* decreased between the early 1960s and the 1990s (Dulvy *et al.* 2000).

The species composition in 1950 might have been close to what would be expected without fishing, i.e. close to carrying capacity for all species. In contrast, since the 1990s the small offshore *L. naevus* increasingly dominated the ray community, while the contribution decreased for the populations of medium-sized species, the offshore other *Leucoraja spp.* and the coastal *R. brachyura*. For all species harvest rates throughout most of the 67 years were larger than those leading to maximum sustainable fisheries yield, pointing at overfishing by coastal fisheries and as well as fisheries on the continental slope as a major factor for their strong decline. While *L. naevus* is also found on the continental slope, its higher intrinsic growth rate compared to the other *Leucoraja spp.* might explains it sustained better the exploitation. For *R. clavata,* the wide bathymetric distribution, from coastal to the upper continental slope might have created refuges from

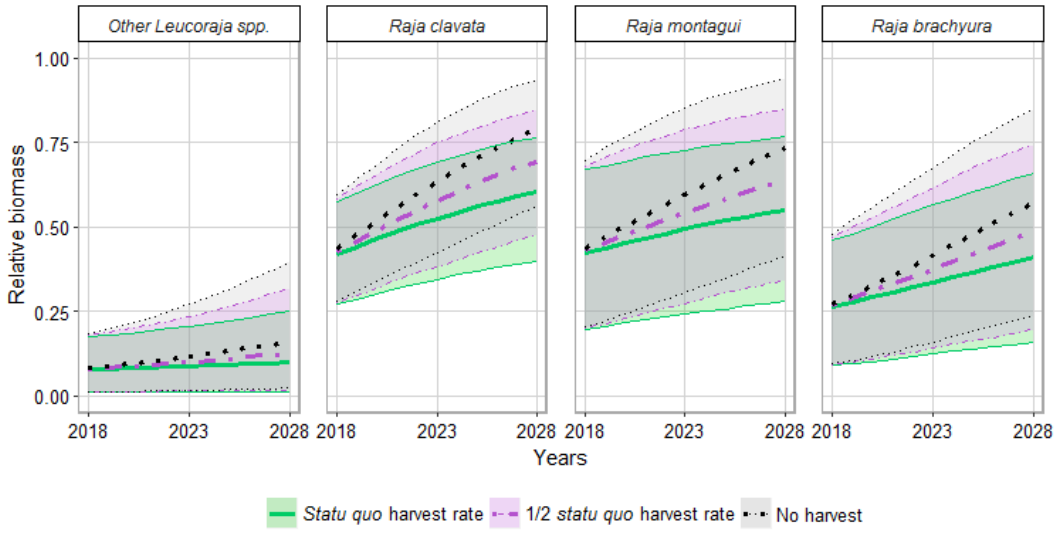


Figure 4: Projected relative biomass for four species under three harvest regimes (color and linetype) (mean values with 95% percentile bands).

fishing, as there is little targeted ray fishing, except in coastal areas. Analysing data from the Celtic Sea Shephard *et al.* (2012) concluded that the spatial heterogeneity of fishing created refuges for elasmobranchs in this area. To evaluate the potential refuge effect in the Bay of Biscay an analysis of the spatial overlap of fishing and species distributions would be needed. Unfortunately, the spatial coverage of the bottom trawl survey was insufficient in coastal waters for the data to be useable for such an analysis. The insufficient coverage of coastal waters also explains the lack of biomass indices for *R. brachyura* and *R. montagui* for several years and generally leads to be prudent regarding the results for these two species.

Prior to the 2010s, ray catches were mostly unmanaged in European waters. The total allowable catch (TAC) for Rajiformes, introduced in 2009 in the study area, may not have been constraining from the start but only as it was gradually reduced from ca. 6400 t in 2009 to ca. 3400 t in 2016. Note that this TAC covers the ecoregion including the Bay of Biscay and the west Iberian shelf. As a reaction to the limited catch opportunities fishers' producer organisations and governments have added various rules, such as a cap on landings per trip and minimum landing sizes. For example in France, rays smaller than 45 cm total length cannot be landed (JORF 2017) and are therefore discarded at sea. As rays have been estimated to be amongst the fish species that survive best being discarded (Depestele *et al.* 2014; Morfin *et al.* 2017; Knotek *et al.* 2018) these increased discards actually represented additional reduction of fishing mortality. If this effect is sufficiently large, it will lead to a biomass increase, though with some delay (Collie, Rochet & Bell 2013). Thus, the biomass increases observed in recent years for most species might at least partially be due to reduced fishing mortality caused by increased discarding in addition to reduced fishing fleet capacity (Mesnil 2008).

Predation release is another factor which can explain community changes (Shepherd & Myers 2005; Myers *et al.* 2007; Ward-Paige *et al.* 2012). Several large elasmobranch species have declined in the Bay of Biscay, including *S. squatina* (Quéro & Cendrero 1996), *D. batis* complex and *R. alba*. Stomach content information collected in the English Channel in the 1930s indicates that large (>120 cm) *D. batis* feed on both *Raja montagui* and *L. naevus* (Pinnegar 2014). Hence predation release might be a potential explanation for the increase in *L. naevus* while such release might have been compensated by fishing for *R. montagui*. Large and small rays may also have overlapping feeding niches (Smale & Cowley 1992), which could have led to competition release.

Differences in biogeographic affinity can contribute to ray species composition changes (Sguotti *et al.* 2016). Environmental conditions have changed in the Bay of Biscay over the last half century. Sea surface temperature has increased by 0.2°C on average per decade between 1964 and 2004, with a cooler period in the 1970s, while the increase was much less below 100m (Michel, Vandermeirsch & Lorance 2009). Thus, the four species distributed in shallow waters will have experienced warming waters. In the shallow waters of the North Sea, *R. clavata* and *R. montagui* were found to be associated with warmer sea surface temperatures (SST) (Sguotti *et al.* 2016). In the Eastern English Channel, SST also contributed to explain interannual variations in the spatial distribution of both species as well as for *R. brachyura* (Martin *et al.* 2012). Thus, similar impacts of temperature change on the spatial distribution might have occurred in the Bay of Biscay which could then have modified the biomass or the interaction with fisheries, but in the absence of dedicated studies this remains speculation. While temperature effects on teleost body growth in the wild are well documented (Brunel & Dickey-Collas 2010; Trenkel *et al.* 2015), there are no studies quantifying temperature effects on ray growth. If it occurred it

might have impacted intrinsic population growth rates during the study period. It is also likely that temperature affected incubation duration of egg-cases as observed in the ghost shark (Lyon, Francis & Francis 2011), and thus reduced natural mortality.

Assuming constant intrinsic population growth rates and carrying capacities over several decades is clearly a simplification. Changes in environmental conditions, habitat, competition and predation might have modified both parameters. Also, changes in the species composition of the benthos in favour of scavengers such as gastropods may modify predation rates on substrate-attached ray eggs (Lucifora & Garcia 2004). Further, the different factors are likely to have affected the six species in different ways.

The multi-species Bayesian model provided an effective way for integrating grouped landings with species-specific biomass indices and life history information (informative prior distributions for intrinsic growth rates) to study long-term changes for species which could not be investigated otherwise due to insufficient data, or only making very strong assumptions. For example, Marandel, Lorance & Trenkel (2016) attempting to study the dynamics of *R. clavata* assumed landings in the Bay of Biscay for this species followed the long term trend of the same species in the North Sea, an assumption whose validity is impossible to evaluate but clearly has strong impacts on the result.

Given the problem of grouped landings is widespread (Reuter *et al.* 2010), we believe the approach has potential for being applied in other areas. However, the suitability of the approach critically depends on the assumption of landings having been roughly proportional to biomasses in years with species-specific data. This assumption might be relaxed if proxies of species targeting are available. Unfortunately this was not the case here. Comparing the estimated species composition with that in recent landings (2009-2016), some differences were found. This seems to indicate that due to the biomass indices being available from 1973 for most species, the model was able to separate species targeting effects in landings from changes in the community composition. Hence the landings splitting assumption did not seem to have been determining for estimated species composition changes.

4.2 Insights for management

Species whose proportion in the ray community and biomass decreased strongly and for which current harvest rates were found too high might be considered the primary species of concern for management. In contrast, abundant species, in particular *L. naevus*, seem to be able to sustain the current level of harvesting and require no specific management action. To improve the species of concern, instead of setting a catch level for every species by splitting the current global TAC for Rajiformes, it might be more appropriate to implement dedicated management measures to minimise landings of the depleted *R. brachyura* and the two other *Leucoraja* species and allow higher landings for *L. naevus*. This might be achieved simply by prohibiting landings of the former, although a management approach leading to more stakeholder engagement would probably be more effective in the long-term. This is particularly important for rays, for which a number of conflicts have been created by too stringent regulations, decided without stakeholder consultation.

5 Funding

The study received funding from the French "Agence Nationale de la Recherche" (project GenoPopTaille, contract ANR-14-CE02-0006-01), the Fondation Total (project GenoPopTaille-Capsules) and the European Union (projet Pandora).

6 Acknowledgements

We thank Eloi Lezian for preliminary work on the model and the GDR EcoStat for constructive comments on preliminary models.

7 References

Barausse, A., Correale, V., Curkovic, A., Finotto, L., Riginella, E., Visentin, E. & Mazzoldi, C. (2014) The role of fisheries and the environment in driving the decline of elasmobranchs in the northern Adriatic Sea. *ICES Journal of Marine Science*, **71**, 1593–1603.

Brander, K. (1981) Disappearance of common skate *Raia batis* from Irish Sea. *Nature*, **290**, 48–49.

Brunel, T. & Dickey-Collas, M. (2010) Effects of temperature and population density on von Bertalanffy growth parameters in Atlantic herring: a macro-ecological analysis. *Marine Ecology Progress Series*, **405**, 15–28.

CEC. (2003) COUNCIL REGULATION (EC) No 2287/2003 of 19 December 2003 fixing for 2004 the fishing opportunities and associated conditions for certain fish stocks and groups of fish stocks, applicable in Community waters and, for Community vessels, in waters where catch limitations are required. Official Journal of the European Union, L 344 of 31.12.2003.

CEC. (2009) COUNCIL REGULATION (EC) No 43/2009 of 16 January 2009 fixing for 2009 the fishing opportunities and associated conditions for certain fish stocks and groups of fish stocks, applicable in Community waters and, for Community vessels, in waters where catch limitations are required.

CEC. (2010) COUNCIL REGULATION (EU) No 23/2010 of 14 January 2010 fixing for 2010 the fishing opportunities for certain fish stocks and groups of fish stocks, applicable in EU waters and, for EU vessels, in waters where catch limitations are required and amending Regulations (EC) No 1359/2008, (EC) No 754/2009, (EC) No 1226/2009 and (EC) No 1287/2009.

Collie, J., Rochet, M.-J. & Bell, R. (2013) Rebuilding fish communities: the ghost of fisheries past and the virtue of patience. *Ecological Applications*, **23**, 374–391.

Courrat, A., Lobry, J., Nicolas, D., Laffargue, P., Amara, R., Lepage, M., Girardin, M. & Le Pape, O. (2009) Anthropogenic disturbance on nursery function of estuarine areas for marine species. *Estuarine, Coastal and Shelf Science*, **81**, 179–190.

Depestele, J., Desender, M., Benoît, H.P., Polet, H. & Vincx, M. (2014) Short-term survival of discarded target fish and non-target invertebrate species in the "eurocutter" beam trawl fishery of the southern North Sea. *Fisheries Research*, **154**, 82–92.

Du Buit, M.H. (1974) *Contribution à l'étude Des Populations de Raies Du Nord-Ouest Atlantique Des Faeroe Au Portugal*. Sciences Naturelles, Université de Paris VI, Paris.

Dulvy, N.K., Fowler, S.L., Musick, J.A., Cavanagh, R.D., Kyne, P.M., Harrison, L.R., Carlson, J.K., Davidson, L.N., Fordham, S.V., Francis, M.P., Pollock, C.M., Simpfendorfer, C.A., Burgess, G.H., Carpenter, K.E., Compagno, L.J., Ebert, D.A., Gibson, C., Heupel, M.R., Livingstone, S.R., Sanciangco, J.C., Stevens, J.D., Valenti, S. & White, W.T. (2014) Extinction risk and conservation of the world's sharks and rays. *eLife.00590*, **3**.

Dulvy, N.K., Metcalfe, J.D., Glanville, J., Pawson, M.G. & Reynolds, J.D. (2000) Fishery stability, local extinctions, and shifts in community structure in skates. *Conservation Biology*, **14**, 283–293.

Dulvy, N.K., Simpfendorfer, C.A., Davidson, L.N.K., Fordham, S.V., Bräutigam, A., Sant, G. & Welch, D.J. (2017) Challenges and Priorities in Shark and Ray Conservation. *Current Biology*, **27**, R565–R572.

Ferretti, F., Osio, G.C., Jenkins, C.J., Rosenberg, A.A. & Lotze, H.K. (2013) Long-term change in a meso-predator community in response to prolonged and heterogeneous human impact. *Scientific Reports*, **3**.

Hutchings, J.A. & Kuparinen, A. (2017) Empirical links between natural mortality and recovery in marine fishes. *Proceedings of the Royal Society B: Biological Sciences*, **284**, 20170693.

ICES. (2011) Historical Nominal Catches 1950-2010. Catches in FAO area 27 by country, species, area and year. Source: Eurostat/ICES database on catch statistics.

ICES. (2016) *Report of the Working Group on Elasmobranchs Fishes (WGEF)*. ICES, Lisbon, Portugal.

ICES. (2017) Official Nominal Catches 2006-2015 Catches in FAO area 27 by country, species, area and year as provided by the national authorities. Source: Eurostat/ICES data compilation of catch statistics.

IUCN. (2005) Malacoraja Clavata *: Ellis, J.: The IUCN Red List of Threatened Species 2005: E.T39399A10199527.*

IUCN. (2007) Raja Montagui *: Ellis, J., Ungaro,N., Serena,F., Dulvy,N.,Tinti,F., Bertozzi, M., Pasolini, P., Mancusi,C.,Noarbartolo Di Sciara, G.: The IUCN Red List of Threatened Species 2007: E.T63146A12623141.*

IUCN. (2008) Raja Brachyura *: Ellis, J., Ungaro, N., Serena, F., Dulvy, N.K., Tinti, F., Bertozzi, M., Pasolini, P., Mancusi, C. , Noarbartolo Di Sciara, G.: The IUCN Red List of Threatened Species 2009: E.T161691A5481210.*

IUCN. (2014a) *Leucoraja circularis* : McCully, S., Ellis, J., Walls, R. & Fordham, S.: The IUCN Red List of Threatened Species 2015: e.T161464A48938919.

IUCN. (2014b) Leucoraja Naevus *: Ellis, J., Dulvy, N., Walls, R.: The IUCN Red List of Threatened Species 2015: E.T161626A48949434*.

IUCN. (2014c) *Leucoraja fullonica* : McCully, S. & Walls, R.: The IUCN Red List of Threatened Species 2015: e.T161461A48938639.

JORF. (2017) Texte n° 6, Journal Officiel de la République Française, n°0073 of 26 march 2017.

Knotek, R.J., Rudders, D.B., Mandelman, J.W., Benoît, H.P. & Sulikowski, J.A. (2018) The survival of rajids discarded in the New England scallop dredge fisheries. *Fisheries Research*, **198**, 50–62.

Kuhnert, P.M., Griffiths, S. & Brewer, D. (2011) Assessing population changes in bycatch species using fishery-dependent catch rate data. *Fisheries Research*, **108**, 15–21.

Letaconnoux, R. (1948) Effets de la guerre sur la constitution des stocks de poissons. Rapports et Procès-Verbaux des Réunions du Conseil Permanent International pour l'Exploration de la Mer. , **122**, 55–62.

Livingston, P., Aydin, K., Boldt, J., Ianelli, J. & Juradomolina, J. (2005) A framework for ecosystem impacts assessment using an indicator approach. *ICES Journal of Marine Science*, **62**, 592–597.

Lucifora, L.O. & Garcia, V.B. (2004) Gastropod predation on egg cases of skates (Chondrichthyes, Rajidae) in the southwestern Atlantic: quantification and life history implications. *Marine Biology*, **145**, 917–922.

Lyon, W.S., Francis, R. & Francis, M.P. (2011) Calculating incubation times and hatching dates for embryonic elephantfish (*Callorhinchus milii*). *New Zealand Journal of Marine and Freshwater Research*, **45**, 59–72.

Marandel, F., Lorance, P. & Trenkel, V.M. (2016) A Bayesian state-space model to estimate population biomass with catch and limited survey data: application to the thornback ray (*Raja clavata*) in the Bay of Biscay. *Aquatic Living Resources*, **29**, 209.

Martin, C.S., Vaz, S., Ellis, J.R., Lauria, V., Coppin, F. & Carpentier, A. (2012) Modelled distributions of ten demersal elasmobranchs of the eastern English Channel in relation to the environment. *Journal of Experimental Marine Biology and Ecology*, **418–419**, 91–103.

McAllister, M.K., Pikitch, E.K. & Babcock, E.A. (2001) Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 1871–1890.

Mesnil, B. (2008) Public-aided crises in the French fishing sector. *Ocean & Coastal Management*, **51**, 689–700.

Soumis à l'ICES Journal of Marine Sciences

Michel, S., Vandermeirsch, F. & Lorance, P. (2009) Evolution of upper layer temperature in the Bay of Biscay during the last 40 years. *Aquatic Living Resources*, **22**, 447–461.

Morfin, M., Méhault, S., Benoît, H.P. & Kopp, D. (2017) Narrowing down the number of species requiring detailed study as candidates for the EU Common Fisheries Policy discard ban. *Marine Policy*, **77**, 23–29.

Myers, R.A., Baum, J.K., Shepherd, T.D., Powers, S.P. & Peterson, C.H. (2007) Cascading Effects of the Loss of Apex Predatory Sharks from a Coastal Ocean. *Science*, **315**, 1846–1850.

Pinnegar, J.K. (2014) DAPSTOM - An Integrated Database & Portal for Fish Stomach Records.

Poulard, J.-C. & Trenkel, V.M. (2007) Do survey design and wind conditions influence survey indices? *Canadian Journal of Fisheries and Aquatic Sciences*, **64**, 1551–1562.

Quéro, J.C. & Cendrero, O. (1996) Incidence de la pêche sur la biodiversité ichtyologique marine.

Reuter, R.F., Conners, M.E., Dicosimo, J., Gaichas, S., Ormseth, O. & Tenbrink, T.T. (2010) Managing non-target, data-poor species using catch limits: lessons from the Alaskan groundfish fishery1: MANAGING DATA-POOR SPECIES WITH CATCH LIMITS. *Fisheries Management and Ecology*, 323–335.

Rousset, J. (1990) Population structure of thornback rays Raja clavata and their movements in the Bay of Douarnenez. *Journal of the Marine Biological Association of the United Kingdom*, **70**, 261.

Sguotti, C., Lynam, C.P., García-Carreras, B., Ellis, J.R. & Engelhard, G.H. (2016) Distribution of skates and sharks in the North Sea: 112 years of change. *Global Change Biology*, **22**, 2729–2743.

Shephard, S., Gerritsen, H., Kaiser, M.J. & Reid, D.G. (2012) Spatial Heterogeneity in Fishing Creates de facto Refugia for Endangered Celtic Sea Elasmobranchs ed H. Browman. *PLoS ONE*, **7**, e49307.

Shepherd, T.D. & Myers, R.A. (2005) Direct and indirect fishery effects on small coastal elasmobranchs in the northern Gulf of Mexico: Fishery effects on Gulf of Mexico elasmobranchs. *Ecology Letters*, **8**, 1095–1104.

Smale, M.J. & Cowley, P.D. (1992) The feeding ecology of skates (Batoidea: Rajidae) off the Cape south coast, South Africa. *South African Journal of Marine Science*, **12**, 823–834.

Stevens, J., Bonfil, R., Dulvy, N.K. & Walker, P.A. (2000) The effects of fishing on sharks, rays, and chimaeras (chondrichthyans), and the implications for marine ecosystems. *ICES Journal of Marine Science*, **57**, 476–494.

Trenkel, V.M., Lorance, P., Fässler, S.M.M. & Høines, å. S. (2015) Effects of density dependence, zooplankton and temperature on blue whiting *Micromesistius poutassou* growth: *micromesistius poutassou* growth variations. *Journal of Fish Biology*, **87**, 1019–1030.

Ward-Paige, C.A., Keith, D.M., Worm, B. & Lotze, H.K. (2012) Recovery potential and conservation options for elasmobranchs. *Journal of Fish Biology*, **80**, 1844–1869.

WoRMS Editorial Board. (2017) World Register of Marine Species. Available from http://www.marinespecies.org at VLIZ. Accessed 2018-02-18.

Appendix A1. Data preparation and modelling details

Correcting landings

For Spain landings were missing for several years. Linear interpolation was applied for 1970-1982 and 1988-1996. For the years 1950 to 1953, the mean ratio between Spanish and French landings during 1954-1958 was applied to French landings for 1950-1953 to estimate Spanish landings. Two years were missing for French landings (1957 and 1999) which were filled with the mean of the directly preceding and following year. Official landings for France for 2005 to 2016 were replaced by corrected French landings as derived by the Working Group of Elasmobranch Fisheries (WGEF) (ICES 2016). Similarly for official landings for Spain, United Kingdom, Ireland and Belgium for the period 2008 to 2016.

Prior distributions

To ensure parameter identifiability rather informative priors were required for the intrinsic growth rates r (eq 1). As it is the best known species, the mode for the prior distribution for *R. clavata* prior was derived using the matrix method described in McAllister et al. (2001). This involved simulating an age-structured population model at equilibrium using different demographic parameters: fecundity, adult natural mortality (from Wiegand et al. (2011)) and weight-length parameters (from Dorel (1986)). To derivate the modes of the priors for r for the other species, we used the threshold conservation fishing mortalities estimated by Le Quesne and Jennings (2012) to scale the value obtained for *R. clavata*. Parameter values used for calculating modes for r priors are summarized in Supplementary table 1.

For all species a uniform prior distribution was used for carrying capacity K_e. The limits of the uniform distributions were somewhat arbitrary but in all cases were wide enough not to be limiting for the estimation.

The prior for survey catchability q was the same for all species: Beta(1,3). The prior for $1/\sigma^2$ was also the same for all species: Gamma(400,1).

Fixed parameters

The fixed value for $1/\tau^2{}_e$ was set to one over the mean biomass index for *L. naevus* and *R. clavata*, thus assuming a variance equal to the mean. For the other species the variance $\tau^2{}_e$ was set to 10 times the mean as these species were less well covered by the survey. The observation variance ρ^2 for species-specific landings was assumed to have a CV of 0.2 for all species and years.

Bayesian inference

All computations were performed with the R platform (R Foundation for Statistical Computing 2008). JAGS (Plummer 2003) was used for Bayesian inference and was run within R using the *rjags* package (Plummer 2016). Results were calculated for three parallel MCMC chains, composed of 400 000 iterations with different initialization points. The burn-in for each MCMC chain was 50 000 iterations and autocorrelation among samples was limited by saving every 300th parameter value. Global convergence was checked with the Potential Scale Reduction Factor, *PSRF,* and the Multivariate Potential Scale Reduction Factor diagnostic, *MPSRF,* which summarizes individual *PSRF* (Brooks and Gelman 1998).

References

Brooks SP, Gelman A. 1998 General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat. 7, 434–455.

Dorel D. 1986 Poissons de l'Atlantique Nord-Est : Relations Taille - Poids.

ICES. 2016. Report of the Working Group on Elasmobranchs Fishes (WGEF). ICES, Lisbon, Portugal.

Le Quesne WJF, Jennings S. 2012 Predicting species vulnerability with minimal data to support rapid risk assessment of fishing impacts on biodiversity: Fishing impacts on fish biodiversity. J. Appl. Ecol. 49, 20–28. (doi:10.1111/j.1365-2664.2011.02087.x)

McAllister MK, Pikitch EK, Babcock EA. 2001 Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. Can. J. Fish. Aquat. Sci. 58, 1871–1890. (doi:10.1139/f01-114)

Plummer M. 2016 rjags: Bayesian Graphical Models using MCMC. R package version 4-6.

Plummer M. 2003 JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling In Proceedings of the 3rd International Workshop on Distributed Statistical Computing.

R Foundation for Statistical Computing. 2008 R Development Core Team (2008). R: A language and environment for statistical computing. Austria, Vienna. See http://www.R-project.org.

Wiegand J, Hunter E, Dulvy NK. 2011 Are spatial closures better than size limits for halting the decline of the North Sea thornback ray, Raja clavata ? Mar. Freshw. Res. 62, 722. (doi:10.1071/MF10141)

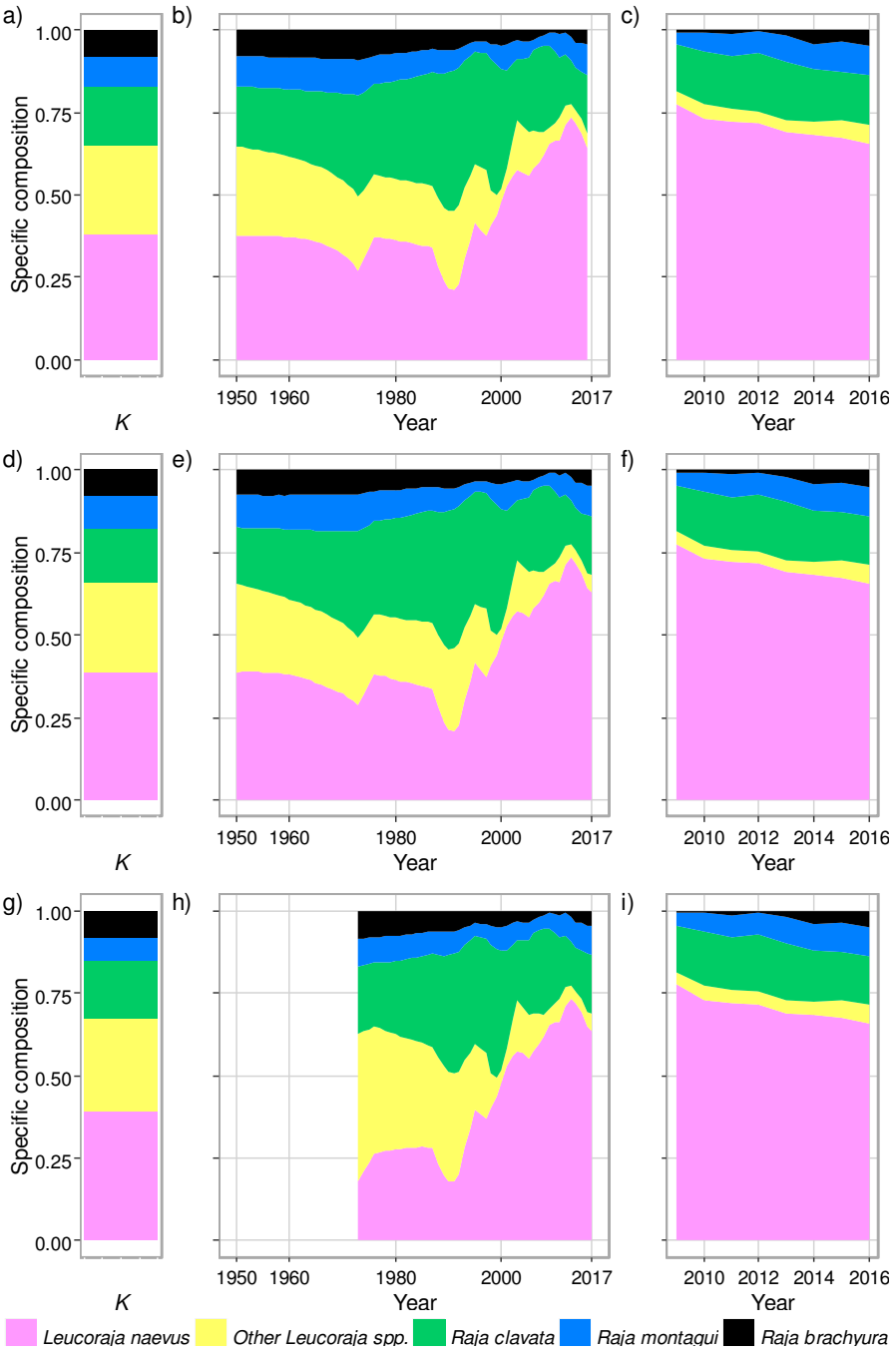


Figure A1: Relative biomass for the five studied species at carrying capacity for the BASE run (a), EXPL run (d) and the SHORT run (g), Estimated species composition between 1950 and 2017 for BASE run (b), EXPL run (e) and the SHORT run between 1973 and 2017 (h) ; Observed species composition in landings between 2009 and 2016 (c,f,i).

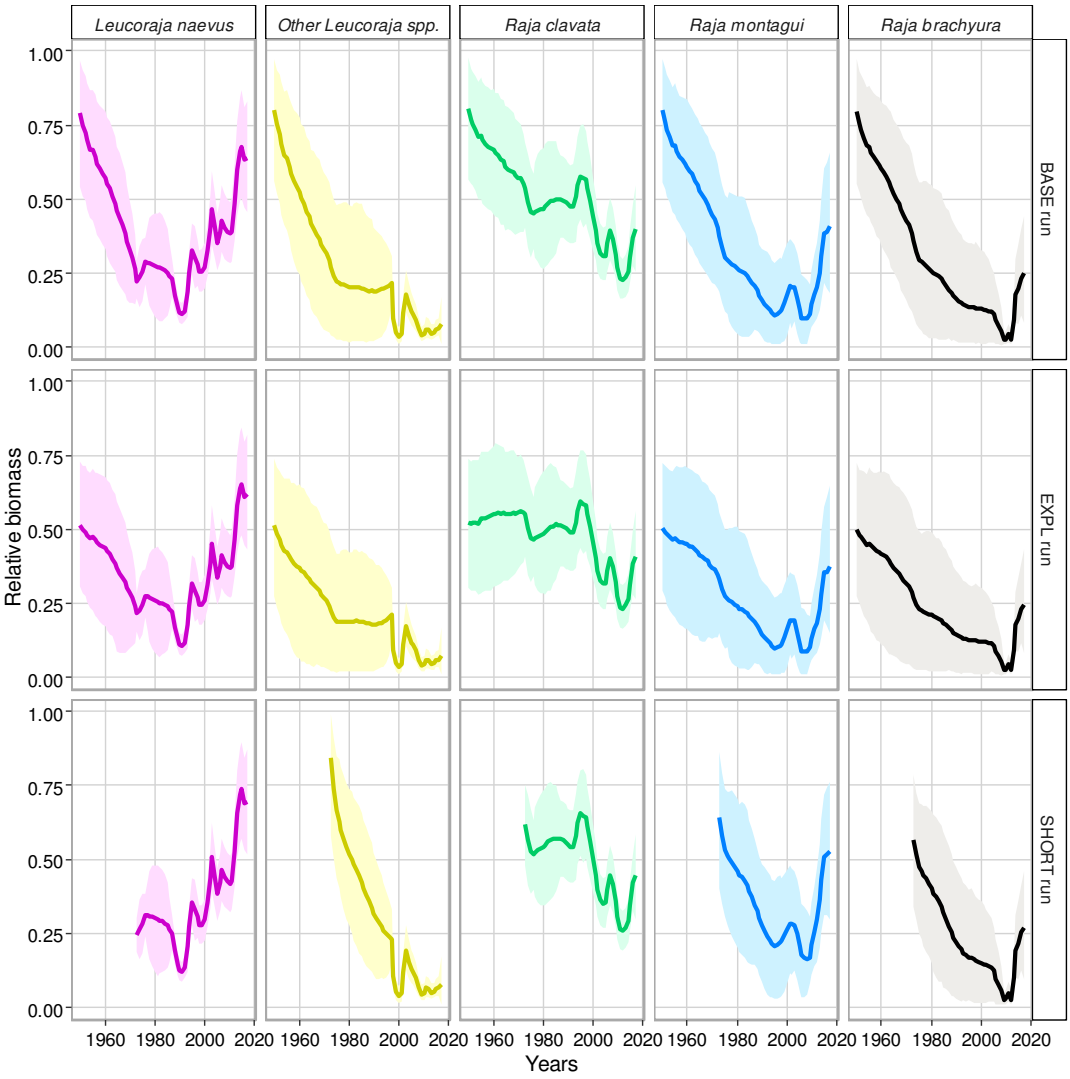


Figure A2: Relative biomass estimates between 1950 and 2017 for BASE run, assuming full exploitation in 1950 (EXPL run) and SHORT run (between 1973 and 2017) for the five studied species with 95% percentile bands.

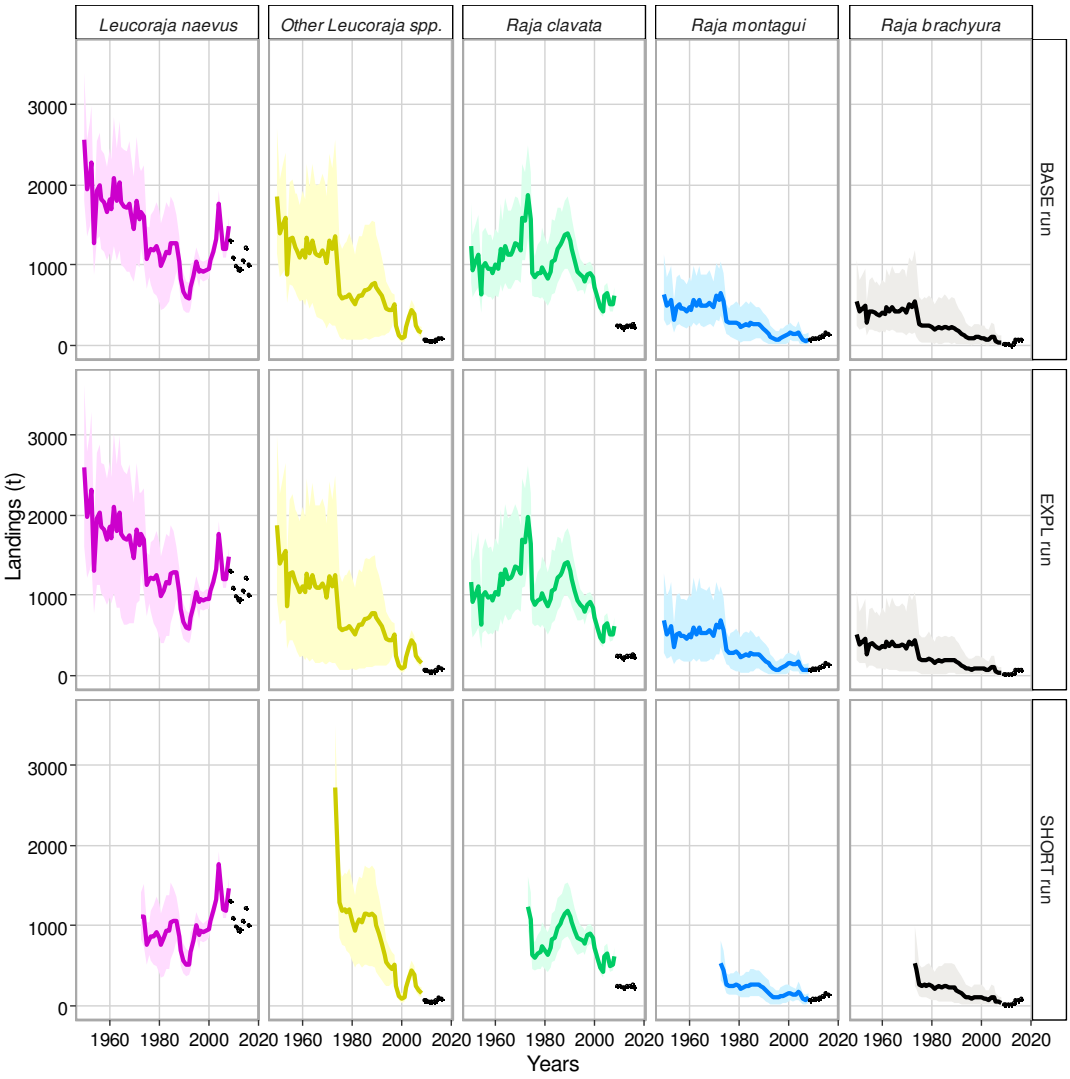


Figure A3: Estimated landings for BASE run, EXPL run and SHORT run for the five studied species with 95% percentile bands. Observed landings as dotted line for 2009-2016.

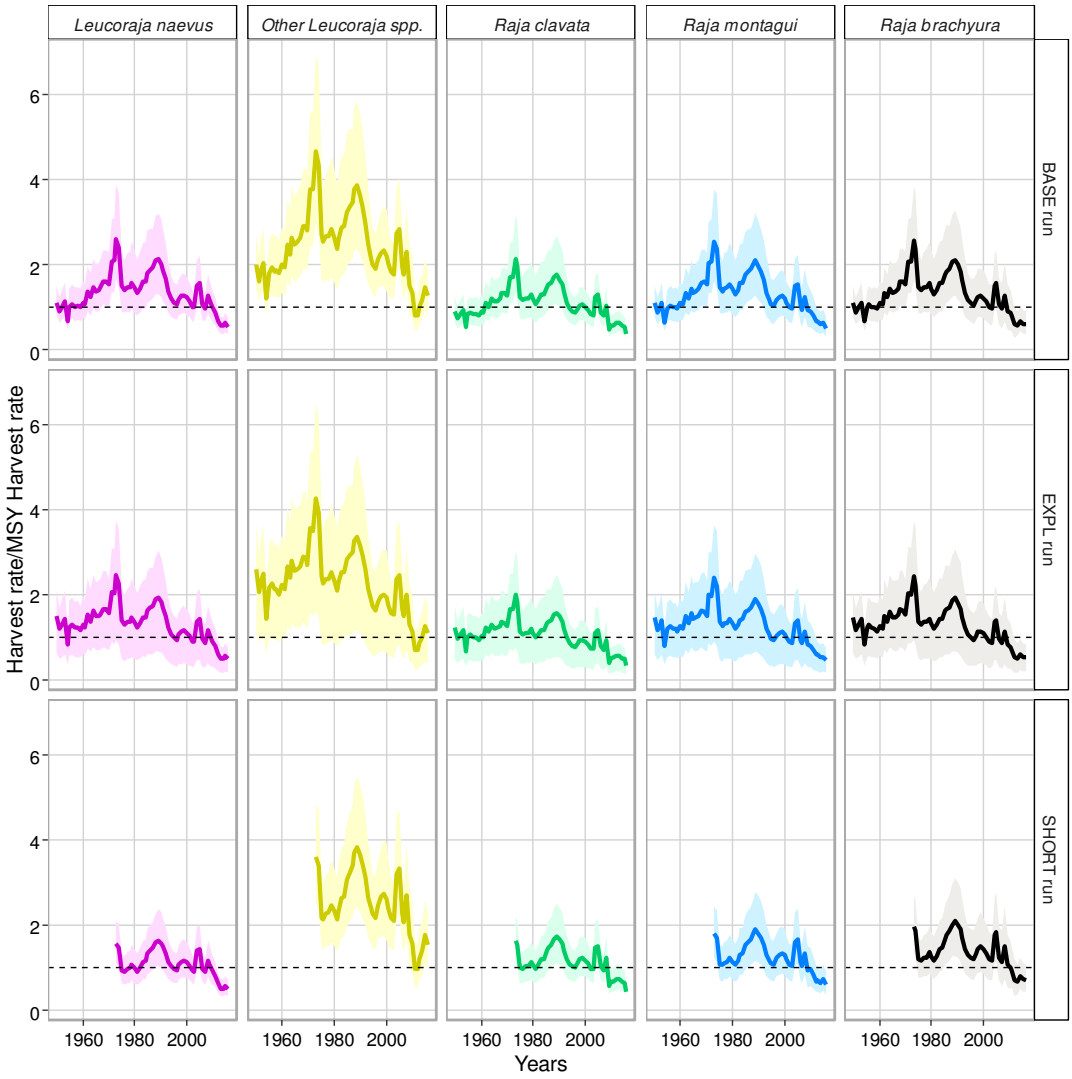


Figure A4: Estimated harvest rates divided by harvest rates at MSY for BASE run, EXPL run and for the SHORT run for the five studies species with 95% percentile bands.

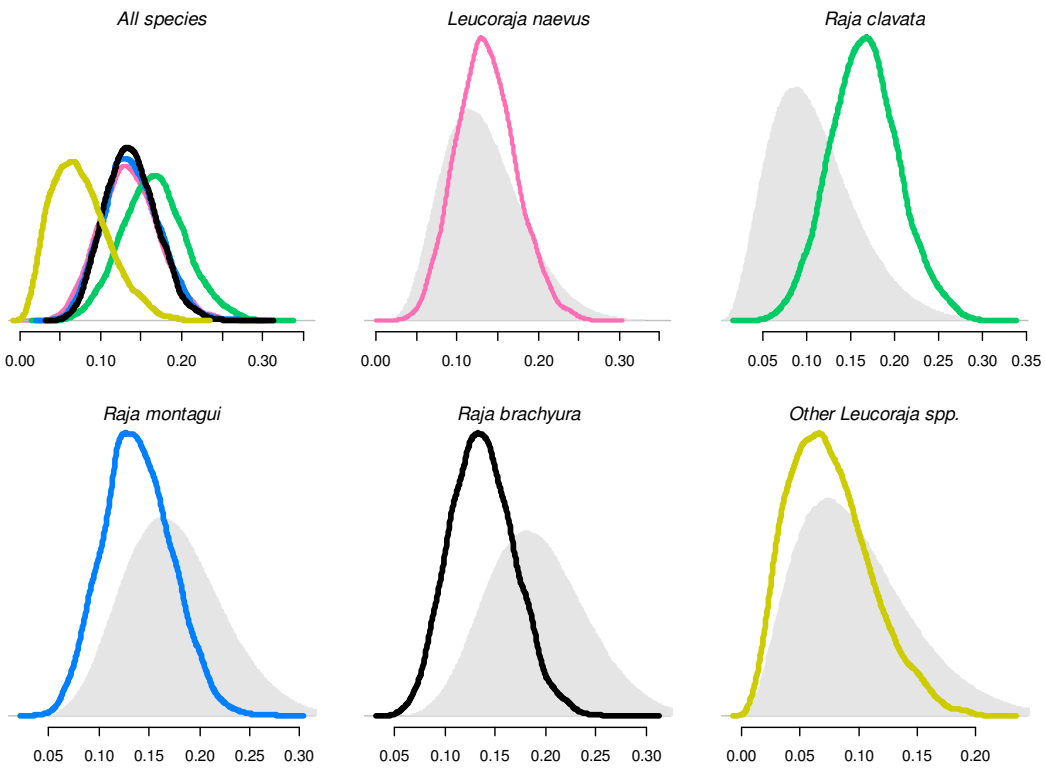


Figure A.5: Prior (grey surfaces) and posterior (lines) distributions for intrinsic growth rate *r* for BASE run.

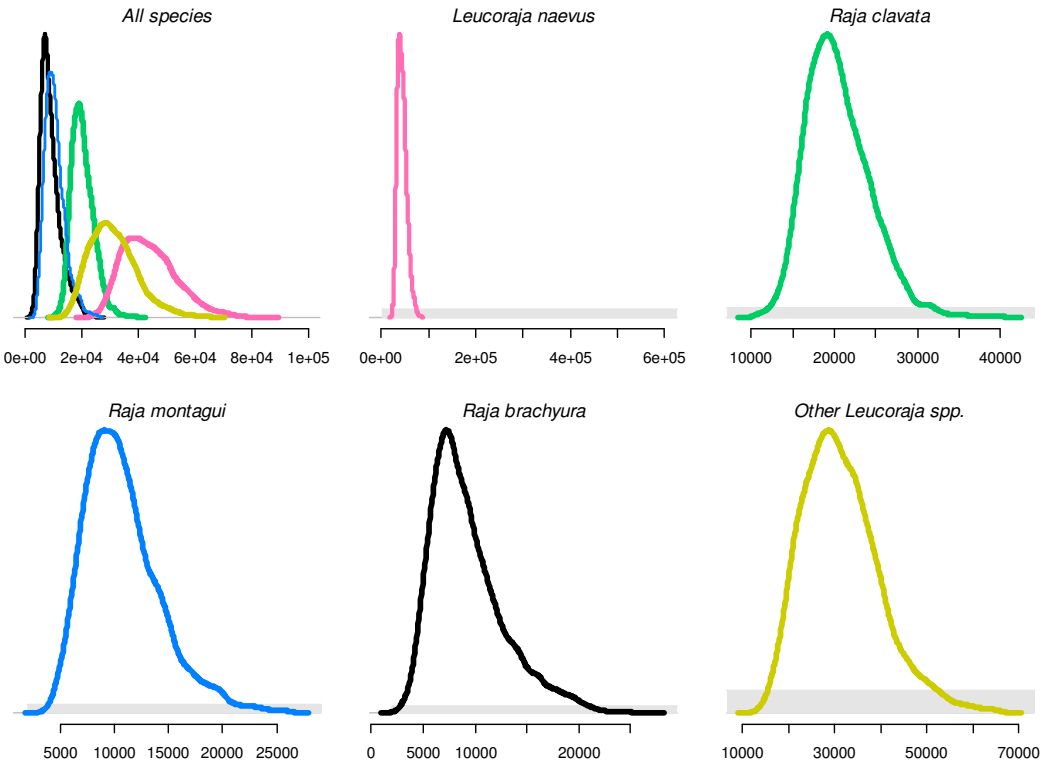


Figure A.6: Prior (grey surfaces) and posterior (lines) distributions for carrying capacity K for BASE run.

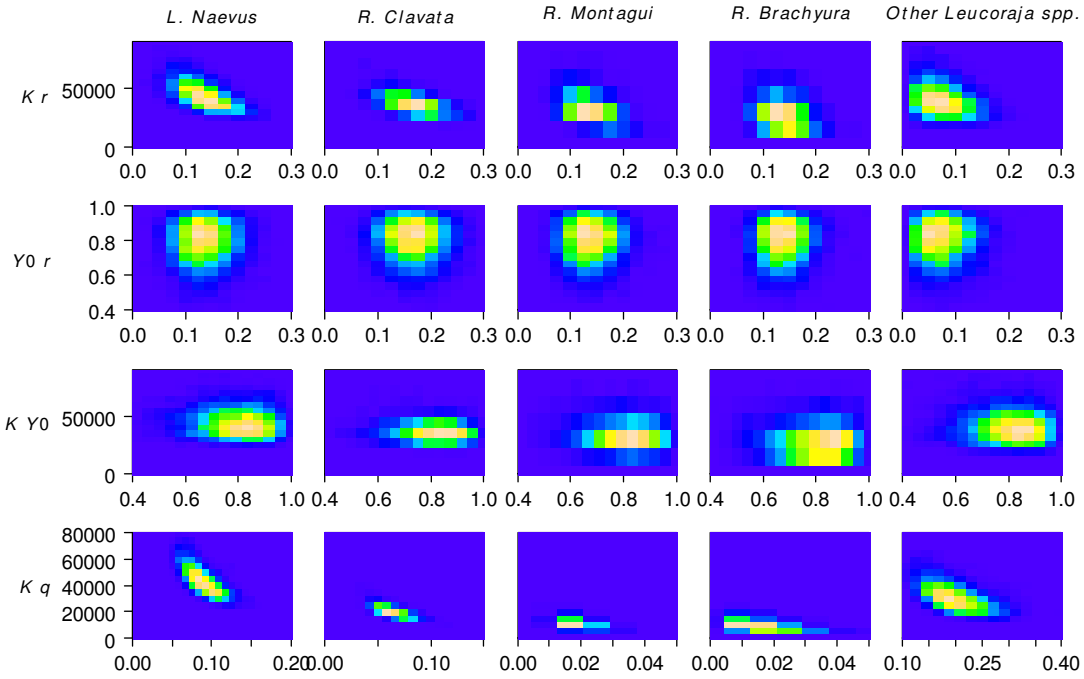


Figure A.7: Joint posterior distributions for model parameters r, K, $Y0$ and q in BASE run.

The colour scale goes from low values (blue) to high values (oranges).

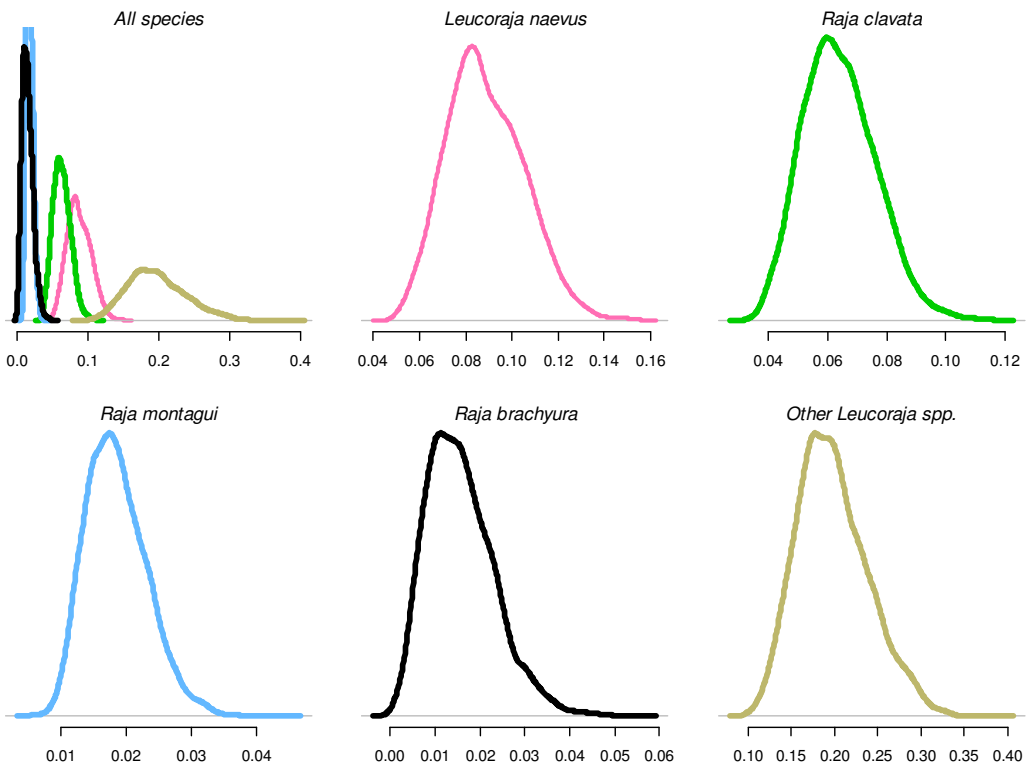


Figure A.8: Posterior (lines) distribution for survey catchability q in BASE run

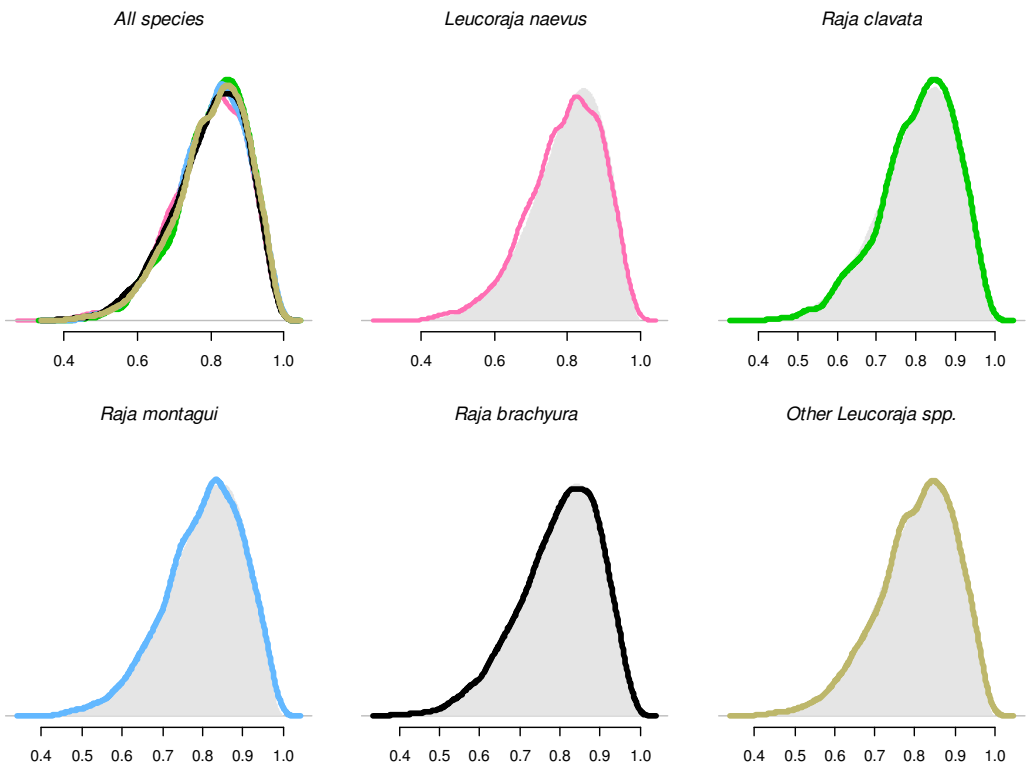


Figure A.9: Prior (grey surfaces) and posterior (lines) distributions for depletion rate *Y0* at the start of the time series for BASE run.

Table A.1: Parameter values used in the Leslie method described by (McAllister et al. 2001) for defining the prior for the intrinsic population growth rate of *Raja clavata* and threshold conservation fishery mortality values used for deriving intrinsic population growth priors for the other species.

Parameter	**Description**	***R. clavata***	***L. naevus***	***R. montagui***	***R. brachyura***	***L. fullonica***	***L. circularis***
L_∞	von Bertalanffy growth function	118 [a]	83.92 [c]	78.43 [c]	120.0 [h]		
k		0.155 [a]	0.197 [c]	0.256 [c]	0.19 [h]		
t_0		-0.665 [a]	-0.151 [c]	-0.968 [c]	-0.45 [h]		
a	coefficient length-weight relationship	0.00000345 [b]					
b	exponent length-weight relationship	3.1807 [b]					
A_{50}	age at 50% maturity	5					
f	number of eggs	140 [a, e]					
F_{cons}	threshold conservation fishing mortality	0.23 [g]	0.29 [g]	0.25 [g]	0.2 [g]	0.21 [g]	0.22 [g]

[a] Wiegand et al. 2011

[b] Dorel et al 1986

[c] Gallagher et al. 2005

[d] McCully et al 2012

[e] Garcia et al. 2008

[f] Du Buit 1972

[g] Le Quesne and Jennings 2011

h Fahy 1991

Table A.2: Prior distribution for model parameter

Scenario	Species	*Landings*	*r*~Beta mode, sd	*Y0*~Beta mode, sd	*K*~Uniform min,max	$1/\sigma^2$~Gamma mode, sd
1	*L. naevus*	1950-2016	0.132, 0.05	0.8, 0.10	2 000, 800 000	400,1
	R. clavata		0.105, 0.05	0.8, 0.10	2 000, 250 000	400,1
	R. montagui		0.114, 0.05	0.8, 0.10	2 000, 250 000	400,1
	R. brachyura		0.091, 0.05	0.8, 0.10	2 000, 250 000	400,1
	Other Leucoraja spp.		0.098, 0.05	0.8, 0.10	2 000, 250 000	400,1
2	*L. naevus*	1950-2016	/	0.5, 0.10	/	/
	R. clavata		/	0.5, 0.10	/	/
	R. montagui		/	0.5, 0.10	/	/
	R. brachyura		/	0.5, 0.10	/	/
	Other Leucoraja spp.		/	0.5, 0.10	/	/
3	*L. naevus*	1973-2016	/	0.5, 0.10	/	/
	R. clavata		/	0.5, 0.10	/	/
	R. montagui		/	0.5, 0.10	/	/
	R. brachyura		/	0.5, 0.10	/	/
	Other Leucoraja spp.		/	0.5, 0.10	/	/

Table A.3: Landings and loess smoothed biomass indices used in the multispecies Bayesian model

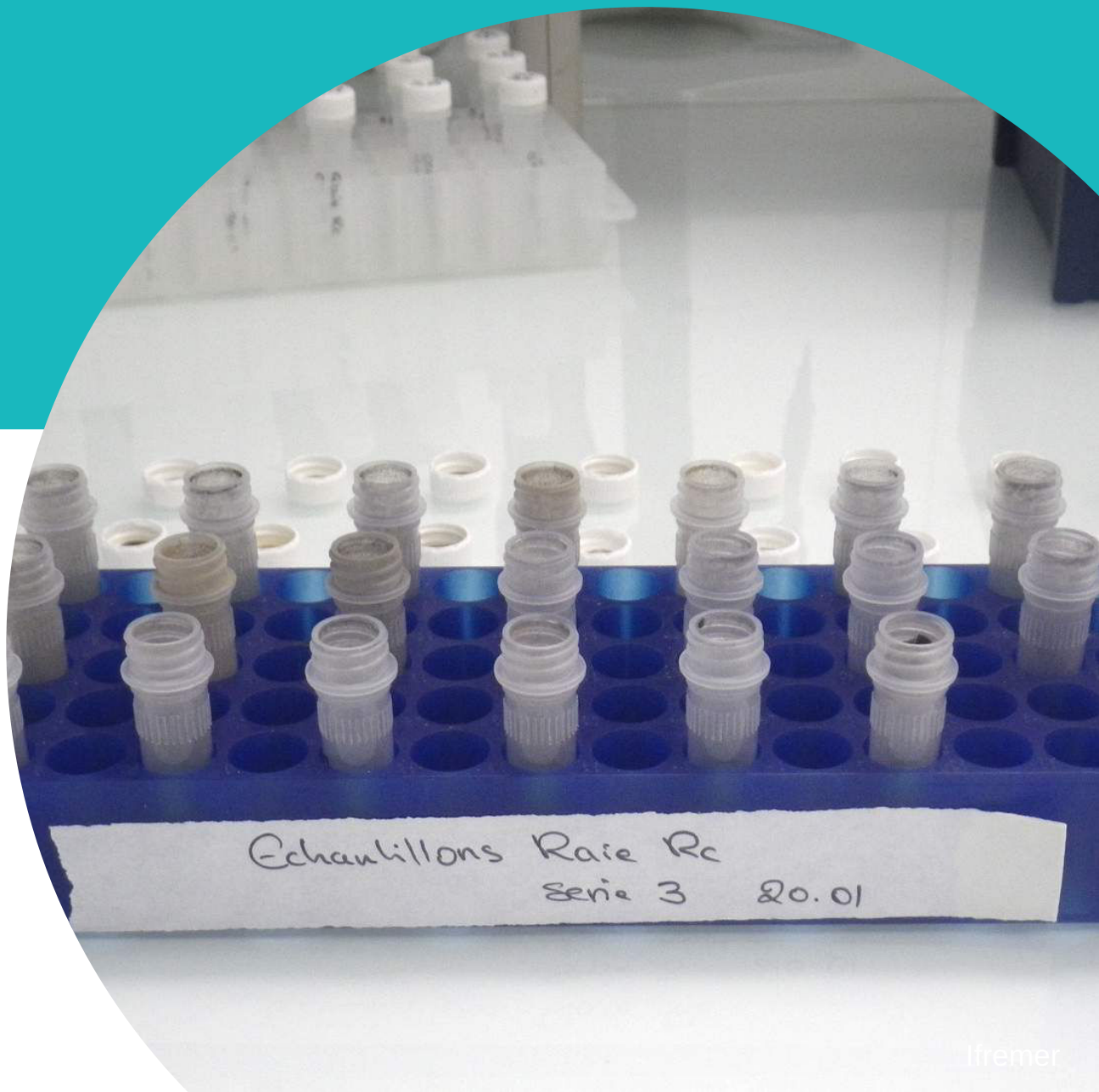
Years	Landings (t)						Biomass indices (Loess smoothed)				
	Total	*L. naevus*	*R.clavata*	*R. montagui*	*R. brachyura*	Other *Leucoraja spp.*	*L. naevus*	*R.clavata*	*R. montagui*	*R. brachyura*	Other *Leucoraja spp.*
1950	6834	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1951	5197	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1952	5648	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1953	6062	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1954	3408	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1955	5128	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1956	5299	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1957	4861	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1958	4782	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1959	4467	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1960	4892	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1961	4598	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1962	5661	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1963	4921	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1964	5638	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1965	4996	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1966	4919	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1967	4940	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1968	5220	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1969	4977	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1970	4537	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1971	5827	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1972	5386	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1973	6120	*NA*	*NA*	*NA*	*NA*	*NA*	826	682	73	*NA*	*NA*
1974	5316	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
1975	3202	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*

1976	3112	NA	NA	NA	NA	NA	1077	566	51	NA	NA
1977	3233	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1978	3201	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1979	3378	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1980	3082	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1981	2758	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1982	3010	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1983	3314	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1984	3286	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1985	3674	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1986	3716	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1987	3746	NA	NA	NA	NA	NA	868	630	38	22	NA
1988	3669	NA	NA	NA	NA	NA	648	625	34	NA	NA
1989	3459	NA	NA	NA	NA	NA	496	615	33	19	NA
1990	3098	NA	NA	NA	NA	NA	428	609	27	NA	NA
1991	2815	NA	NA	NA	NA	NA	411	600	NA	NA	NA
1992	2537	NA	NA	NA	NA	NA	448	596	NA	NA	NA
1993	2392	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1994	2392	NA	NA	NA	NA	NA	940	692	NA	NA	NA
1995	2494	NA	NA	NA	NA	NA	1239	733	6	17	NA
1996	2309	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1997	2527	NA	NA	NA	NA	NA	1047	712	NA	NA	1300
1998	2275	NA	NA	NA	NA	NA	950	679	9	NA	528
1999	2146	NA	NA	NA	NA	NA	954	622	NA	NA	284
2000	2009	NA	NA	NA	NA	NA	1009	556	NA	NA	199
2001	1998	NA	NA	NA	NA	NA	1210	495	62	NA	251
2002	2094	NA	NA	NA	NA	NA	1503	438	NA	NA	663
2003	2309	NA	NA	NA	NA	NA	1761	397	67	NA	1080
2004	3107	NA	NA	NA	NA	NA	1514	384	28	23	835
2005	2846	NA	NA	NA	NA	NA	1294	383	11	24	691

2006	2074	*NA*	*NA*	*NA*	*NA*	1455	443	3	*NA*	662
2007	1995	*NA*	*NA*	*NA*	*NA*	1610	507	*NA*	*NA*	550
2008	2384	*NA*	*NA*	*NA*	*NA*	1510	462	5	*NA*	363
2009	1690	1310	239	65	11	65	1450	404	8	26	291
2010	1510	1102	246	86	11	64	1422	341	11	27	328
2011	1362	982	217	92	18	52	1441	288	*NA*	*NA*	452
2012	1305	935	227	86	7	50	1797	274	12	*NA*	426
2013	1389	959	244	109	27	50	2260	280	20	*NA*	291
2014	1548	1057	241	121	67	63	2480	313	41	*NA*	245
2015	1802	1214	266	162	65	95	2541	375	123	*NA*	222
2016	1524	1001	226	134	76	87	2362	482	*NA*	28	222

CHAPITRE 5

Évaluer l'état des populations de raie bouclée : une méthode génétique



Ifremer

Ce chapitre étudie l'estimation de l'état d'une population de raie par méthode génétique. L'indicateur choisi est la taille de population efficace. La première partie du chapitre développe et explique le concept de taille de population efficace. La seconde partie étudie par simulation et par application à des données empiriques l'estimation de la taille de population efficace chez la raie bouclée.

1. La taille de population efficace, N_e

1.1 Définition

Le concept de taille de population efficace (N_e) a été introduit pour la première fois en 1931 (Wright, 1931). Il est utilisé depuis des dizaines d'années pour la conservation d'espèces terrestres (Schwartz et al., 2007) et plus récemment pour la conservation d'espèces marines (Dudgeon et al., 2012). Elle correspond au nombre d'individus participant réellement à la prochaine génération, génétiquement parlant (Hamilton, 2009). Autrement dit, les juvéniles, les individus trop âgés pour se reproduire ou ceux produisant des gamètes inutilisés ne sont pas comptabilisés dans cette taille de population. Elle renseigne ainsi sur la taille d'une population mais aussi sur sa santé génétique (Frankham, 1995; Portnoy et al., 2009).

Sa définition exacte est la taille d'une population idéale au sens de Wright-Fisher pour laquelle on aurait une dérive génétique équivalente à celle de la population naturelle étudiée (Lowe et al., 2004). Cette comparaison avec une population idéale standardise la mesure de la dérive génétique et permet une comparaison des tailles de populations efficaces d'espèces présentant des traits d'histoire de vie très différents. A noter que chez une population idéale, la taille de population efficace est égale à la taille de population absolue, N (Hamilton, 2009). Cette propriété a été exploitée dans les simulations introductives du chapitre 2. La définition induit également que, sauf bottleneck intense et récent, la taille de population efficace est plus petite que la taille de population absolue.

De nombreux facteurs influent sur la taille efficace d'une population. Ils sont brièvement discutés ci-dessous :

- **Variation dans la taille absolue** : les fluctuations dans la taille absolue d'une population entraine souvent des variations dans les fréquences alléliques entrainant à leur tour des variations dans la taille de population efficace.
- **Sex ratio** : un sex ratio déséquilibrée influence le taux de dérive génétique et donc la taille de population efficace. Par exemple, une population présentant un ratio d'un mâle reproducteur pour plusieurs femelles aura une dérive génétique bien supérieure à une population présentant un sex ratio de 1 :1 (Smith and Smith, 2001; Hamilton, 2009).
- **Succès reproducteur** : une population est stable dans le temps lorsque chaque pair d'individus produit une moyenne de deux descendants. Lorsque certains individus produisent un nombre de descendants bien supérieurs à d'autres, ils transmettent majoritairement leur allèles à la prochaine génération entrainant une baisse de population efficace (Hamilton, 2009).
- **Consanguinité** : la consanguinité induit une baisse de la variabilité génétique d'une population qui se traduit par une baisse de la taille de population efficace (Lowe et al., 2004; Hamilton, 2009).

– **Migration et dispersion** : lorsque des immigrants se reproduisent dans leur population d'acceuil, ils introduisent leur matériel génétique induisant une augmentation de la variabilité génétique et donc de N_e (Smith and Smith, 2001).
– **Générations chevauchantes** : ce facteur a une influence considérable sur la taille de population efficace, un individu pouvant se reproduire plusieurs fois au cours de son existence et potentiellement avec ses descendants. Néanmoins les effets de ce facteur sur N_e restent encore difficilement prévisibles (Waples et al., 2014).

1.2 Estimateurs de la taille de population efficace

Les méthodes génétiques d'estimation de la taille de population efficace sont devenues de plus en plus abordables grâce aux récentes avancées en génotypages, vitesses de calcul et disponibilité de logiciel user-friendly (Hare et al., 2011). Deux grands types d'estimateurs existent : ceux estimant la taille de population efficace historique et ceux estimant la taille de population efficace contemporaine. Dans cette thèse, le travail s'est concentré sur cette dernière catégorie. La taille de population efficace contemporaine peut être estimée génétiquement grâce à l'analyse de marqueurs génétiques précis (**Single-Nucleotide Polymorphism** et **Microsatellites** notamment). Là encore, deux grandes catégories existent :

– **Estimateurs basés sur un unique échantillonnage génétique** : ces estimateurs nécessitent un unique échantillonnage de la population étudiée. De par leur échantillonnage simplifié par rapport aux estimateurs multiples, ils sont très utilisés et ont vu leur développement exploser ces dernières années (Dudgeon et al., 2012). Le plus répandu de ces estimateurs est basé sur le déséquilibre de liaison entre loci (estimateur Linkage Disequilibrium), c'est à dire sur l'association non aléatoire d'**allèles** entre eux.
– **Estimateurs basés sur plusieurs échantillonnages génétiques séparés dans le temps** : autrement appellées, méthodes temporelles, ces méthodes nécessitent des échantillons séparés dans le temps soit par de plusieurs générations (Waples and Yokota, 2006), soit de plusieurs années si les âges des individus échantillonnés peuvent être déterminés (Jorde and Ryman, 1995). Le principe général de la méthode repose sur une comparaison de l'évolution des fréquences alléliques entre les échantillons, tout changement étant interprété comme résultant d'un changement de taille de population efficace.

Tous ces estimateurs reposent sur de nombreuses hypothèses explicites et implicites et sont donc sensibles à de nombreuses sources de biais. Selon les estimateurs, ces sources et leurs effets sont plus ou moins bien identifiés néanmoins chacun d'entre eux est sensible à la démographie de la population étudié et ce de manière tout à fait imprévisible car spécifique à l'espèce.

Dans cette thèse, l'estimation de la taille de population efficace pour la raie bouclée a été explorée par simulations afin d'étudier l'effet de sa démographie sur les estimations. Deux estimateurs ont été sélectionnés *a priori* en raison de leur popularité mais aussi de leurs hypothèses. En effet, la méthode "Linkage Disequilibrium" (Do et al., 2014) et la méthode temporelle corrigée pour la prise en compte des générations chevauchantes (Jorde and Ryman, 1995) sont les méthodes présentant le moins de violations d'hypothèses d'applicabilité à la raie bouclée. Dans un second temps, la taille de population efficace de la raie bouclée du Golfe de Gascogne est calculée à partir des données empiriques du projet GenoPoptaille afin d'être mise en relation avec les attendus théoriques mis en évidence par les simulations.

2. Article : Comparaison de données empiriques et simulées pour l'estimation de la taille de population efficace : application à la raie bouclée

Article en cours de soumission - 2018

Florianne Marandel[1], Verena M. Trenkel[1], Olivier Berthele[1], Sabrina Le Cam [2,3], Pascal Lorance [1], Robin S. Waples [4], Jean-Baptiste Lamy [3]

RESUME - La taille de population efficace (N_e) est un paramètre clé en évolution et en génétique des populations. Néanmoins N_e est difficile a estimer dans les populations naturelles car soumise à de nombreuses sources potientielles de bias. Ces facteurs incluent les biais dues à l'échantillonnage, venant des générations chevauchantes mais aussi venant des méthodes génétiques utilisées pour estimer N_e. Dans ce travail, nous avons simulé une population de poisson structurée en âge et imitant la démographie d'une raie bouclée afin d'étudier le bias et la précision des estimations de N_e. Deux méthodes ont été utilisées : la méthode Linkage Disequilibrium (LDM) et la méthode temporelle (TM). De plus, deux tailles de population ont été simulées 1 000 et 10 000 individus. Ensuite, N_e a été estimé à partir de données empiriques provenant de raies bouclées du Golfe de Gacsogne à l'aide de la méthode LDM. Les résultats des simulations montrent un bias de 30% avec la méthode LDM dès que la taille d'échantillon atteint 100 individus, et ce quelque soit la taille absolue de la population. La méthode TM performe mieux en terme de bias et précision dès lors que le délai entre les échantillon dépasse 7 ans. Les estimations empiriques de N_e sont très variables allant de 400 à 2 700 individus selon les filtres appliquées aux données. Augmenter le taux de données manquantes tend à augmenter N_e lors qu'augmenter le maximum de la fréquence allélique mineur tend à diminuer N_e. Ces estimations semblent indiquer une bonne santé génétique de l'espèce dans le Golfe de Gascogne. De plus, si une correction de 30% est effectuée sur ces estimations, toutes les N_e estimées sont supérieures à 50 ou 500 individus, ce qui garantit le maintien du potentiel adaptatif.

[1] Ifremer, rue de l'île d'Yeu, BP 21105, 44311 Nantes Cedex 3, France

[2] Université de Bretagne Occidentale, Laboratoire des Sciences de l'Environnement Marin (LEMAR, UMR 6539 CNRS/IRD/UBO/Ifremer), Institut Universitaire Européen de la Mer (IUEM), Technopôle Brest-Iroise, 29280 Plouzané, France

[3] Ifremer, Génétique et pathologie des mollusques marins, La Tremblade sur mer, France

[3] Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, Washington, United-States

In preparation

Estimating effective population size for non-model species: insights from simulations and empirical data of thornback ray

Florianne Marandel[1*], Robin S. Waples[3], Olivier Berthele[1], Jean-Baptiste Lamy[2], Sabrina Le Cam[2], Pascal Lorance[1], Verena M. Trenkel[1]

[1]Ifremer, Ecologie et Modèles pour l'Halieutique, Nantes, France
[2]Ifremer, Génétique et Pathologie des Mollusques Marin, La Tremblade, France
[3] Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, Washington

Abstract – Effective population size (N_e) is a key evolutionary parameter of population genetics. The number of studies estimating effective population size has greatly increased in recent years with the development of Next-generation sequencing technologies. However N_e remains difficult to estimate for natural populations as a variety of factors are likely to bias estimates. These factors include sampling design, violation of estimator assumptions, such as non-overlapping generations, as well as the sequencing method and bioinformatic treatment used. One issue inherent to the restriction site-associated DNA sequencing (RADseq) protocol is missing data. In this study, we simulated genetic markers for a non-model species by taking the thornback ray as example to evaluate the bias and precision of N_e estimates obtained with two estimators: the Linkage Disequilibrium method (LDM) and the Temporal method (TM) corrected for overlapping generations. Two population sizes were tested: 1 000 and 10 000 individuals. Simulation results revealed an average 30% underestimation using the LD method for sample sizes above 100 individuals. The corrected TM performed better in terms of bias and precision when the time lag between samples was larger than age-at-maturity. We also estimated N_e from thornback ray empirical RADseq-derived SNP data (sampled in the Bay of Biscay, Northeast Atlantic) with the LDM. Empirical N_e estimates were highly variable, ranging from 200 to 2700 individuals depending on the data filters used. Increasing the allowed percentage of missing data increases N_e estimates while increasing the maximum minor allelic frequency decreased estimates. Jointly simulation results and data sub-sampling provided insights into the expected bias and uncertainty of N_e estimates for a non-model species.

Key words – Effective population size, Simulation, Linkage-Disequilibrium, Temporal method, skates and rays, RADSeq, SimuPOP code

1. Introduction

Effective population size (N_e) is a key parameter of population genetics and conservation (Hamilton 2009). It is related to the number of individuals which actually participate to produce the next generation and thus informs on the genetic health of a population. Estimation of N_e is central for the prediction of evolutionary trajectories of natural population as it can help evaluating population viability. However estimating N_e can be challenging. Conceptually N_e is defined as the size of an idealized Wright-Fisher population that would experiment the same rate of genetic drift as the observed population (Hamilton 2009, Beaumont et al. 2010). This definition relies on strong assumptions, such as constant population size (census population size is equal to the effective population size in an ideal population), discrete generations and the absence of migration, all of which are all likely to be violated by natural populations (Waples et al. 2014).

There are two main approaches for estimating N_e: demographic methods based on life history traits and genetic methods mainly based on genetic markers. Demographic approaches estimate N_e as a function of classical demographic parameters such as survival-at-age and birth rate. They also identify key demographic processes that determine N_e for a given species. However, demographic methods often rely on strong assumptions such as discrete generations (Caballero 1994, Nomura 2002) or if overlapping generations are admitted, stable age structure (Waples et al. 2011). Recently, a new demographic method was developed which allows demographic stochasticity and heterogeneity at the expense of challenging data demands such as individual-level information (Engen et al. 2010). Thus the method has not been much used

*Corresponding author: florianne.marandel@ifremer.fr – (+33) 2 40 37 41 64

In preparation

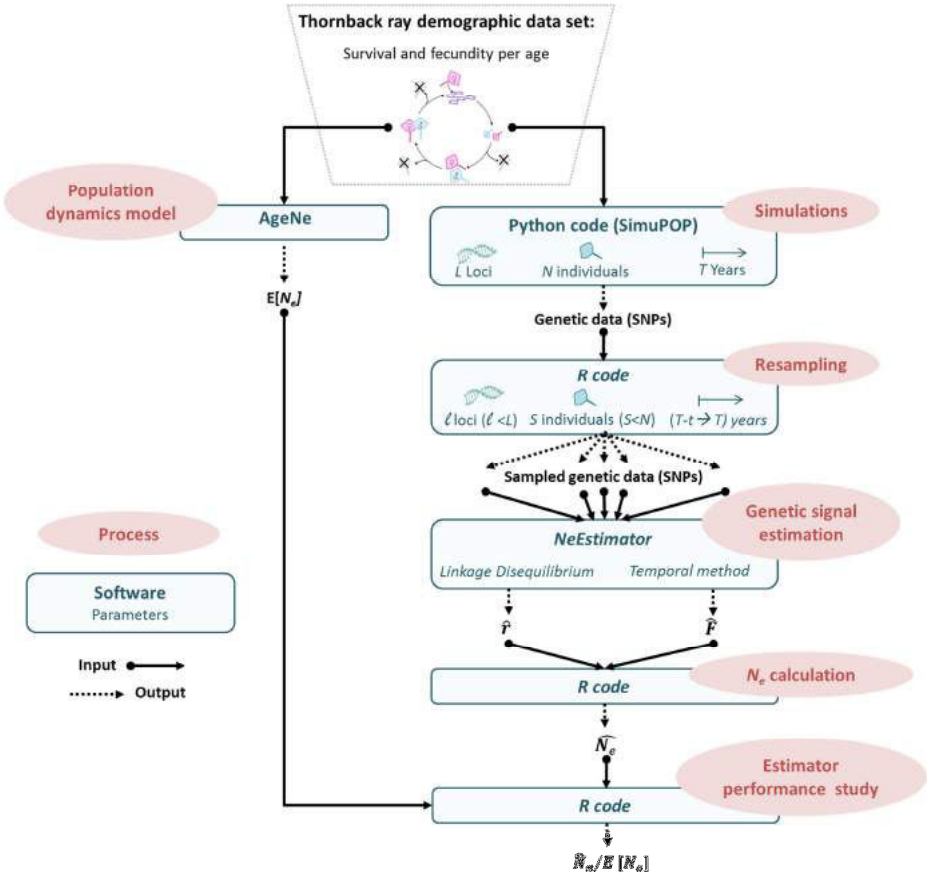


Figure 1: Flowchart of the methods used in the estimator performance simulation study. Software references: AgeNe (Waples et al. 2011), NeEstimator (V2, Do et al. 2014), R (R Foundation for Statistical Computing 2008) and simuPOP (Peng and Kimmel 2005).

so far (but see Trask et al. 2017). Genetic methods have gained in popularity and power due to recent advances in genotyping and sequencing technologies and computer processing speed. They rely on the extraction of various genetic signals which are theoretically known to be affected by population demography, mainly effective population size. Among genetic methods, single sample approaches are appealing since they require sampling only at one point in time. The most popular N_e estimator is based on a measure of linkage disequilibrium (LD) *i.e.* the nonrandom association of alleles at different loci. The LD method has been widely used during the last decade for a variety of organisms, including mammals (Cervantes et al. 2011, Juarez et al. 2015), insects (Francuski and Milankov 2015), reptiles (Bishop et al. 2009) and fishes (Wilson et al. 2014, Pilger et al. 2015). Other widely used genetic methods, called multiple samples estimation, require samples separated in time. The most widely multiple samples method is the temporal method (TM). It measures the observed amount of allelic frequency change between two temporally separated samples for estimating effective population size. Multiple samples methods are usually more accurate than single sample methods but they are more challenging to implement as several samples separated in time or one sample with knowledge of the age of the sampled individuals are needed. Despite these difficulties, TM is widely used (Waples and Yokota 2006, Serbezov et al. 2012).

Because all N_e estimation methods rely on strong assumptions, empirical estimates of N_e are often biased. Numerous recent genetic studies have documented how more realistic simulations or real data, which do not fulfil the methods' assumptions, have tremendous impacts on N_e estimates (Luikart et al. 2010, Waples and Do 2010, Hare et al. 2011, Robinson and Moyer 2013, Waples et al. 2014, Gilbert and Whitlock 2015, Marandel et al. submitted). These studies demonstrated that the amount and the direction of bias as well as the precision of estimates are species-specific and thus highly dependent on life history traits but also on the sampling fraction (Marandel et al. submitted). For non-model species, not only the genome is unknown leading to challenges for genetic marker development and data quality, but generally also the amount of expected species-specific bias for N_e estimates. The development of genetic markers for a non-model

In preparation

species typically involves the discovery and the assay development of each marker which is costly in time and research funding (Davey and Blaxter 2010). The next step consists then in genotyping individuals for these markers. A widely used method providing data in one step is restriction associated DNA sequencing (RADSeq), which provides thousands of sequenced markers across many individuals at reasonable costs (Davey and Blaxter 2010), but suffers from missing data for certain individuals for certain markers and unknown physical linkage between markers.

To determine the performance of N_e estimation for a non-model species, taking the thornback ray (*Raja clavata*) as an example, we carried out a genetic simulation study for two classical N_e estimators. We also evaluated N_e estimation for this species using empirical RADseq data. The simulations explored the effects of sample size, the number of loci used, the population size (LD and TM) and the time lag between sampling events (TM only) on estimation bias. In the simulation process, special attention was paid to the age structure of the simulated nonideal population. the For the empirical RADseq data, the effect on N_e estimates of the selected range of allelic frequencies and the proportion of missing data for a given SNP were explored. Together the simulation and the empirical results gave insights into the conditions for reliable estimation of the effective population size of a non-model species.

2. Material and methods

2.1. Simulation study of estimator performance

The main steps of the simulation study as well as the different software packages and tools used are detailed below and summarized in figure 1. The study was set up to represent thornback ray life history traits, i.e. low fecundity and relatively high survival rates. Survival and fecundity rates were used in two ways as in Waples et al. (2014): (1) to calculate the expected effective population size N_e using a hybrid Felsenstein-Hill method (AgeNe software, Waples et al. 2011) and (2) to perform simulations of age-structured genetic data (simuPOP python module, Peng and Kimmel 2005) (Fig. 1).

2.1.1 Expected effective population size

The expected N_e was calculated using the demographic approach implemented in the AgeNe software (Waples et al. 2011). This method is based on a discrete-time age-structured population dynamics model with sex- and age-specific survival and fecundity rates and assumes constant population size N. Outputs are adult population size (N_A) as well as generation length (G), expected effective population size E[N_e], and the ratio E[N_e]/N. The approach makes several simplifying assumptions, such as a constant census population size and constant population structure (for a full description see Waples et al. 2013).

For thornback ray fecundity was fixed at 140 eggs for ages 6 years and older (maximum age 18 years) and zero otherwise. Survival rate estimates were obtained in several steps. First, size-at-age L_a was estimated using the von Bertalanffy equation with growth parameters t_0, K and L_∞ from the literature (Serra-Pereira et al. 2008). Second, mortality-at-age M_a was estimated using the empiric relationship developed by Gislason et al. (2010). Third, mortality M_a was transformed to survival-at-age S_a=exp(-M_a). Finally, to account for senescence survival was reduced from age 12 onwards (Figure 2).

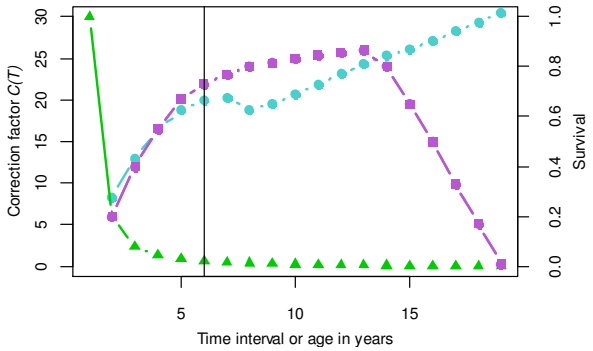


Figure 2: Blue dots: correction factor as a function of the time interval T between samples; green triangles: probability to survive to given age; purple squares: mean age-specific survival. Vertical line: age-at-maturity.

2.1.2 Genetic simulations

Simulations of genetic data were conducted with the individual-based program simuPOP (Peng and Kimmel 2005) to produce realistic genetic data for the modeled species (Fig. 1). The demographic characteristics were the same as used for estimating E[N_e]. Each simulation replicate (50 in total) was run during 100 years, after a burn in of 150 years. Two population sizes N were tested, 1000 and 10 000 individuals. Sex was randomly assigned to all individuals at the start of the simulations (t=0) assuming a 1:1 sex ratio. Newborn individuals were generated by randomly drawing one male and one female parent from the pool of potential parents and assigning the sex assuming a 1:1 sex ratio. All potential parents of the same sex and age had an equal opportunity to be parent. Individuals in the founder generation (t=0) were assumed unrelated and the genotype of each individual was generated for L genetic markers.

The simulated genetic markers correspond to SNPs (Single Nucleotid Markers) for which the initial allelic frequency was set to 0.5. Concerning the number of genetic markers, an initial trial using L=1600 SNPs and N=1000 individuals showed that a plateau in bias and precision of N_e estimates was rapidly reached. Hence the maximum number of simulated SNPs was L=500. All SNPs were assumed to be located on different chromosomes to avoid linkage disequilibrium due to physical linkage. As the simulations covered only a few generations, mutation processes were ignored. The simulation design is summarized in Table 1.

In preparation

Table 1: Summary of simulation design. In black, parameters used for simulation results reported in the main text. In grey, parameters for simulation results reported in supplementary material. *N* census population size (all individuals); *L* number of SNP; *S* sample size.

			Life stage sampled	
N	***L***	***S***	**LD Method**	**Temporal Method**
1000, 10000	500, 400, 300, 200, 100	50, 100, 200, 300	All, Mature, Immature, Newborn	Newborn

For validating the simulations, for each simulation replicate the observed loss of heterozygosity over time was compared to the expected loss where the expected loss was calculated using (Fox and Wolf 2006)

$$E\left[\frac{H_1}{H_0}\right] = 1 - \left(1 - \frac{1}{2E[N_e]}\right)^T \quad (1)$$

with H_x: the heterozygosity at time x, T the number of generations between t_0 and t_1 and E[N_e] the expected effective population size calculated above. For each simulation replicate, the ratio was calculated between the last time step (t_1=250) and the first time step after burn-in (t_0=150). Only the 30 replicates among the 50 carried out presenting the less divergence in heterozygosity loss from the expected ratio were selected for further genetic and statistical analyses. This ensured simulation results were comparable to the calculated expected E[N_e] value.

In terms of sampled ages, four samplings strategies were investigated for the genetic data: (1) only newborns (i.e. individuals <1 year), (2) only immature individuals (≥ 1 year and <6 years), (3) only mature individuals (≥ 7 years) (4) all individuals (Table 1). In each case, individuals were randomly sampled without replacement. Moreover, sampling was age-stratified with equal numbers for each age group included in the sample. Different sample sizes were investigated ranging from 50 to 300 individuals (Table 1), as we were interested in accuracy and precision assessment as well in practical sampling guidelines. For each simulation replicate and sample size 50 samples were drawn leading to overall 1500 estimates for each sample size (30 replicates with 50 samples each). Five sets of numbers of SNPs were tested $L \in (100, 200, 300, 400, 500)$. SNPs with minor allelic frequencies <0.05 were removed from the estimation as suggest by Waples and Do (2010) to minimize sampling bias, however they were rare as simulations were carried for 250 years only.

2.1.3 Estimators

Effective population size was estimated from the simulated genetic data using one single point estimation method (Linkage Disequilibrium) and a temporal method adapted for overlapping generations by Jorde and Ryman (1995), and improved by Jorde (2012):

$$N_e = \frac{C(T)}{2G\left(F - \frac{1}{S} + \frac{1}{N_1}\right)} \quad (2)$$

with $C(T)$ the correction factor for overlapping generations calculated from life history traits which depends on the time interval T between sampled cohorts (same birth date), S the harmonic mean of the two samples sizes, N_1 the number of newborns (assumed constant for each cohort), G generation length, and F a measure of the change in allelic frequencies between cohorts. $C(T)$ and G were calculated from age specific cumulative survival and birth rates using the software FactorC developed by Jorde and Ryman (1995) and improved by Jorde (2012). The demographic data used are in figure 2 and supplementary table 1.

2.1.4 Performance evaluation

For both estimation methods, accuracy and precision of N_e estimates was assessed by comparison with E[N_e] (Fig. 1). Moreover, we estimated the effects of sample size (S) and census population size (N) on log-transformed N_e estimates and quantified the variance due to simulation replication for the LD method with a liner mixed-effects model; the normality of the residuals was checked. N and S were treated as fixed effects and the variance due to simulation replicates as independent random effect nested in N.

2.2 Estimation of N_e from empirical data

2.2.1 Sampling

A total of 159 thornback ray were sampled in the Bay of Biscay (see Supplementary figure 1, for a map of sampling locations) between 2011 et 2016 (highest sampling in 2015 with 83 individuals). Sampled individuals were from various sources: 65 came from the French EVHOE survey (see Poulard and Blanchard 2005 and Poulard and Trenkel 2007 for EVHOE survey details), 44 were sampled by APECS, 27 were sampled directly in ports and 23 came from the dedicated French RaieJuve survey. The sex ratio of the sample was close to 1:1 (78 females, 81 males).

For all individuals, total length was reported; disc width was also reported for 138 individuals. Total length varied from 12.50 to 96 cm and disc width from 8 to 66 cm. Maturity occurs at around 73 cm (Serra-Pereira et al. 2008), implying that 66% of the sampled individuals were immature (Supplementary figure 2).

2.2.2 RAD-sequencing protocol and bioinformatic treatment

All individuals were genotyped by sequencing using a RADseq protocol to effectively subsample the genome of multiple individuals at homologous locations. The library construction followed the original protocol by Baird et al., (2008). Briefly, 1 µg of genomic DNA from each individual was digested with the restriction enzyme *Sbf*I-HF (New

England Biolabs), and then ligated to a P1 adapter labeled with a unique barcode. We used 16 barcodes of 5-bp and 16 barcodes of 6-bp long in our P1 adapters to build 32-plex libraries. Seven pools of 32 individuals were made by mixing individual DNA in equimolar proportions. Each pool was then sheared to a 350 pb average size using a Covaris S220 sonicator (KBiosciences), and size-selected on agarose gel to keep DNA fragments within the size range 300-700 pb. Each library was then submitted to end-repair, A-tailing and ligation to P2 adapter before PCR amplification for 18 cycles. Amplification products from six PCR replicates were pooled for each library, gel-purified after size selection and quantified on a 2100 Bioanalyzer using the High Sensitivity DNA kit (Agilent). Each library was sequenced on a separate lane of an Illumina HiSeq 2500 instrument by INTEGRAGEN (91000 Evry, France), using 100-bp single reads.

We used the program *Stacks* (Catchen et al., 2011, Catchen et al., 2013) before being trimmed to 95 bp. Both *de novo* and reference mapping pipelines (using the *Leucoraja erinacea* genome Wang et al. 2012, Wyffels et al. 2014) were applied to call individuals genotypes. Genotyped individuals were typed on 389 483 SNPs spread on 35 134 RAD loci (a sequence starting or ending with a restriction enzyme site). Given the randomness of restriction enzyme digestion DNA and the cumulated noise due to wet laboratory and sequencing protocols, we choose to retain loci with a calling rate percentage above 33% (CI 95 1.2 - 88). Initial SNP selection was made for a larger sample of individuals from different locations in the Northeast Atlantic. We therefore recalculated allelic frequencies for the selected individuals leading to some SNPs in the data set presenting a minor allelic frequency (MAF) of 0.

Finally, to reduce the dataset size without losing information (and reducing redundancy), we kept only one SNP per RAD loci. This was done by selecting the most polymorphic SNP for each available RAD locus. Thus the genetic data consisted of 35 134 SNPs for 159 individuals, with randomly missing data for certain SNPs for certain individuals (coded NA).

2.2.3 N_e estimation by Linkage Disequilibrium

The estimation of N_e was made with the Linkage Disequilibrium method implemented in the NeEstimator software. As the method is known to be sensitive to the polymorphism of the genetic markers, we tested six filters for the minimum MAF: 0.00, 0.05, 0.15, 0.25, 0.35 and 0.45 where a value of 0.05 means that MAF ∈ (0.05, 0.5). These filters were combined with five different filters for the maximum percentage of missing data (NA) for each included SNP: 30%, 25%, 20%, 15% and 10% leading to 30 empirical genetic datasets. Confidence intervals at 95% were directly computed by NeEstimator. Absolute value of N_e estimates will not be considered as the sample size (159 individuals) was insufficient for the Bay of Biscay population assumed >1 000 000 individuals (Marandel et al. 2016). Empirical data were therefore only informative concerning the effect of the MAF filter and the percentage of missing data on N_e estimates.

To test whether MAF or the percentage of missing data had the bigger effect on N_e estimates, a linear model was fitted to N_e estimates, followed by an ANOVA; residuals were checked for normality. We also tested the interaction between the two main effects (MAF and percentage missing data).

3. Results

3.1 Simulation study

3.1.1 Simulations validation

Generation length and age structure of simulations were consistent with the demographically estimated expected values. For thornback ray, the loss of heterozygosity in simulations was similar to the expected value (0.734), with the difference ranging from 0 to 17% depending on replicate (mean difference 7%). As detailed above, 30 replicates with less than 2% difference were selected for further analyses. This validation step ensured that the effective population sizes calculated with the simulated data were comparable to the expected value E[N_e] calculated with AgeNe.

3.1.2 Estimation with Linkage Disequilibrium method

Estimates of $\widehat{N}_e$ based on age-stratified samples were underestimated on average by around -30% for all combinations of sample size (S) and population size (N) except for S=50 and N=10 000 for which the mean underestimation was around -13%, but with high variability among replicates and samples (Figure 3). Only four N_e estimates were negative for L=500, S=50 and N=10 000 and none for N=1000.

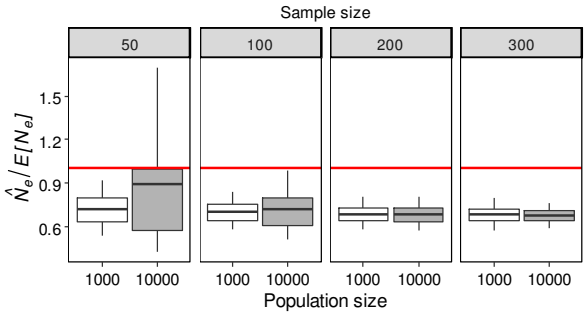


Figure 3: Ratio of $\widehat{N}_e$ calculated using Linkage Disequilibrium method for simulated genetic samples and E[N_e] calculated based on demography for different sample sizes. Colors: simulated population size. Boxplot: Percentile 95, 75, 25, 5 and mean (horizontal black line).

Sampled age: All age-stratified sampling types performed similarly in terms of accuracy and precision (Supplementary figure 3). Therefore, only results from sampling all ages are presented below.

Sample size: Figure 3 compares $\widehat{N}_e$ values for different sample sizes. The underestimation was similar for all sample and population sizes (except when S=50 and N=10 000); increasing sample size did not reduce the amount of underestimation. However, increasing sample size greatly

In preparation

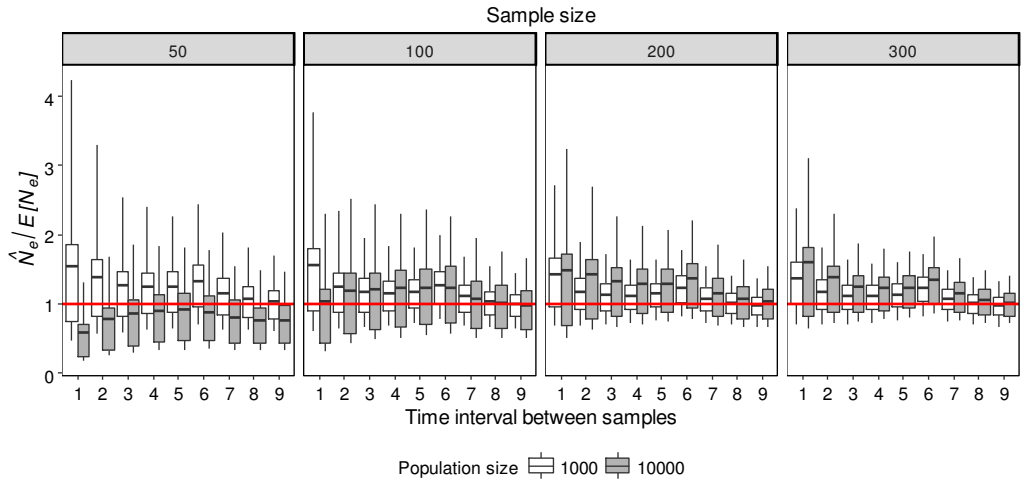


Figure 4: Ratio of $\widehat{N}_e$ calculated by the Temporal method using NeEstimator (V2, Do et al. 2014) and the correction factor of Jorde and Ryman (1995) and expected N_e calculated by AgeNe (Waples et al. 2011). Boxes: sample size (mean of the two sampling events) from 50 to 300. Colors: simulated population size. Boxplot: Percentile 95,75,25,5 and mean.

improved precision. For example, for population size *N*=1000, the 5%-95% interpercentile (90IP) range of $\widehat{N}_e$ was 0.37 for *S*=50 and 0.22 for *S*=300; for *N*=10 000, the 90IP was reduced from 1.54 to 0.17 when increasing sample size *S* from 50 to 300 individuals (Figure 3).

Number of SNP: The number of SNPs (tested range 100-500) had no effect on accuracy (Supplementary figure 4) but more SNPs slightly increased precision. For *N*=1000, the 90IP range was not reduced (0.22 for 100 and 500 SNPs) while the 25%-75% interquartile (IQ) range of $\widehat{N}_e$ was slightly reduced from 0.08 (100 SNP) to 0.07 (500 SNP).

Population size: Two population sizes were simulated and the same sampling scheme was applied to both for estimating N_e (see Table 1 for a summary of the simulation design). When sample size was above 100 individuals, the mean underestimation of N_e estimates was comparable between the two population sizes: -32% for *N*=1000 and -33% for *N*=10 000 (Figure 3). For $S \in (200, 300)$, precision was comparable between the two population sizes. For *S*= 300, IP range was 0.22 for *N*=1000 and 0.17 for *N*=10 000. However for smaller sample sizes, $S \in (50, 100)$, precision was higher for the smaller population size (for *S*=50, IP range was 0.37 for N=1000 and 1.27 for *N*=10 000) (Figure 3).

The results of the linear mixed-effects model showed that sampling variability (residual variance) was much larger than simulation replication variances, with simulation replication variance being larger for the smaller population size (Table 2). The results of the ANOVA for this linear mixed-effects model showed that *N* had the biggest effect on N_e estimates, compared to *S* and the interaction between the two.

Table 2: Simulation results. Summary of the linear mixed effects model for the effects of census population size and sample size on N_e estimates and variance estimates for replicate simulation runs and sampling.

Fixed effect	**df**	**MS**	**F**
Sample size (S)	3	4.82	133.03
Census size (N)	1	610.26	16842.77
S:N	3	5.96	54.84

Random effect	**Variance**
Replicates (*N*=1000)	0.008154
Replicates (*N*=10 000)	0.003235
Residuals	0.036233

3.1.3 Estimation with Temporal Method

Using the Temporal Method the mean estimation bias in $\widehat{N}_e$ for different time intervals between samples was rather variable, for example +25% for *S*=50 and *N*=1 000 and -20% for *S*=50 and *N*= 10 000 (Figure 4). All N_e estimates were positive.

Sample size: Figure 4 shows boxplots of $\widehat{N}_e$ for the four sample sizes for time intervals between the two sample events ranging from one to nine years. The effect of the sample size on $\widehat{N}_e$ was different for the two population sizes. For *N*=1000, mean bias across time intervals decreased between *S*=50

In preparation

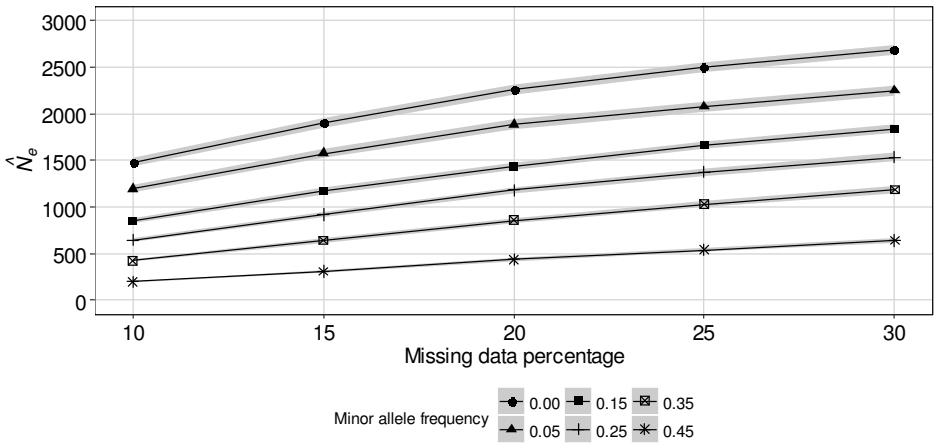


Figure 5: $\widehat{N_e}$ with 95% confidence bands for empirical thornback ray data using the Linkage Disequilibrium method (NeEstimator V2, Do et al. 2014) for different minor allele frequencies (MAF) and percent of missing data.

(+25%) and *S*=300 (+13%) while for *N*=10 000, the mean bias decreased between *S*=50 (-20%) and *S*=100 (+13%) and then increased for *S*=300 (+25%). However, for N=10 000, for sample sizes *S*≥100 the true value (E[N_e]) was included in the IP and IQ ranges. Concerning precision, the effect of the sample size differed for the two simulated population sizes. For *N*=1000, increasing *S* improved precision, the mean IP range across time intervals decreased from 1.92 (*S*=50) to 0.93 (*S*=300); for *N*=10 000, the IP range decreased from 1.33 (*S*=50) to 1.18 (*S*=300). Surprisingly, for N=10 000, the mean 90IP range increased between *S*=50 (1.33) and *S*= 200 (1.52) and decreased again for *S*=300 (1.18).

Time interval between samples: Globally, increasing the time interval between samples improved precision and reduced bias (Figure 4). The 90IP range decreased with increasing time interval from 3.77 to 1.09 (*N*=1000 and *S*=50). The reduction in bias with increasing time interval was more marked for *N*=10 000 than for N=1000; similarly for larger sample sizes. For all population and sample sizes, a decrease in $\widehat{N_e}$ was observed after interval 6, which corresponds to age-at-maturity (6 years for thornback ray, figure 2) coming to play in the correction faction *C(T)* in eq 2.

Number of SNP: The number of SNPs used had very little effect on bias (Supplementary figure 4) but slightly increased precision with the 90IP decreasing from 0.91 to 0.71 with increasing number of SNPs. Above *L*=200 there was no gain in precision or bias by increasing the number of SNPs.

Population size: Two population sizes were simulated and the same sampling scheme was applied to both. For *S*=50, the temporal method performed differently for the two population sizes with similar amount of bias but different directions (positive for N=1000 and negative for *N*=10 000) (Figure 3). For *S*≥200, the sign of the bias was the same for both population sizes, but bias was larger for *N*=10 000; precision was also lower.

3.2 Empirical data

The LD method was used for the 30 empirical data sets to test the effect of the amount of missing data and the MAF of included SNPs on $\widehat{N_e}$ values. The number of SNPs varied between 553 and 14 473. A complementary analysis (with *L*=1000 and 400 data sets for each tested MAF range and percent missing data) showed that the number of SNPs had no effect on $\widehat{N_e}$ estimates and little effect on their precision (results not shown).

For all empirical data sets, 95% confidence intervals (CI) of $\widehat{N_e}$ estimates were narrow. The widest CI were obtained for data sets using no filter on MAF (mean CI width 103.88) and the narrowest for data sets with the most stringent filter on MAF=0.45 (mean CI width 37.48). Independent of the percentage of missing data, the value of $\widehat{N_e}$ decreased as the MAF increased (Figure 5). For example, for the smallest percentage of missing data (10%), $\widehat{N_e}$ increased from 200.0 to 1481.0 individuals as the minimum MAF decreased from 0.45 to 0.00. Finally, $\widehat{N_e}$ decreased as the percentage of missing data decreased. No correlation was found between the amount of missing data and the MAF of a given SNP which if present might have explained the observed pattern (Figure 6).

Table 3: Analysis of variance of the linear model fitted to N_e estimates for testing the effects of the percent of missing data and minimum MAF. Empirical results for thornback ray. Adjusted R^2=0.97.

Name	df	MS	F	P-value
Missing data	1	2954288	178.71	<0.001
MAF	1	9902337	599.01	<0.001
Residuals	27	16531		

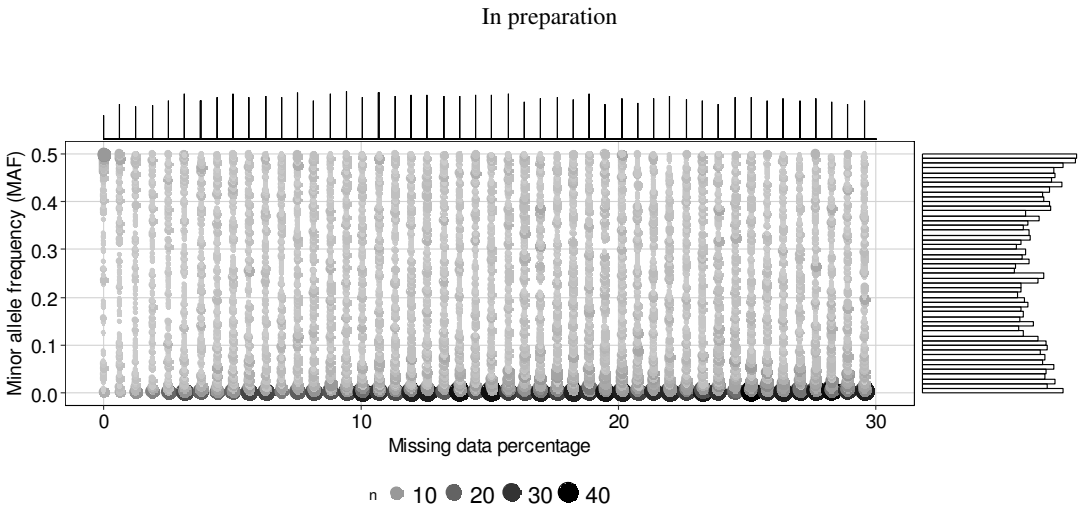


Figure 6: Number of SNPs with given minor allele frequency (MAF) and percent of missing data (limited to 30%) for empirical genetic data and associated histograms.

Fitting a linear model (R^2=0.97), we found that the empirical N_e estimates increased with the percentage of missing data (slope 44.38) and decreased with minimum MAF (slope -3 608.64). The ANOVA results showed that the variability in Ne estimates was mainly explained by the minimum MAF (Table 3).

4. Discussion

4.1 Factors affecting N_e estimates

Many marine fish populations have overlapping generations and large population sizes, whereas genetic effective population size estimators generally assume discrete generations and perform poorly for large populations (Marandel et al. Submitted). Using a simulation approach, we evaluated the performance of N_e estimators for a non-model species, taking the thornback ray as example, for the two most popular methods: the Linkage Disequilibrium and the Temporal method, the latter being corrected for overlapping generations.

Using the Linkage Disequilibrium method, N_e was globally underestimated for the two simulated population sizes (1000 and 10 000 individuals). For sample sizes above 100 individuals and both population sizes, N_e estimates stabilized at 30% underestimation with slightly increase in precision. Underestimation of N_e is a well-known property of the Linkage Disequilibrium method for species with overlapping generations (Waples et al. 2014). The reported amount of underestimation for random samples of adults lies between 50% (mosquito) and 10% (cod) (Waples et al. 2014) with the 30% found for a thornback ray like species lying in between. Assuming the simulated demography represents thornback ray demography well, a 30% correction of thornback ray N_e estimates obtained with the Linkage Disequilibrium method could therefore be carried out given sample size is sufficiently large.

A clear effect of sample size was found for both estimators. Here only relatively small population sizes could be simulated for computational reasons. However, the appropriate sample size depends on population size (Marandel et al. Submitted). Thornback ray populations in the Northeast Atlantic are thought to range from 500 000 individuals to more than 3 million of individuals (estimates derived from population biomass estimates in Marandel et al. (2016)). Thus simulations presented here do not inform on appropriate sample sizes for the thornback ray population in the Bay of Biscay.

Contrary to Waples and Do (2010) and Robinson and Moyer (2013), we found no effect of the age of sampled individuals on N_e estimates. Estimates based on newborn samples performed equally well to samples including a number of consecutives cohorts roughly equal to generation length as recommended by these authors. Using the Temporal method corrected for overlapping generations, N_e was globally overestimated for all tested sample sizes and both population sizes. The amount of overestimation depended on sample size, population size and the time interval between samples. The simulation results confirmed the sampling recommendation specifying that the time interval between samples and the sample size should be as large as possible to reduce bias and increase precision (Robinson and Moyer 2013). Moreover, time interval between samples should be larger than the age at maturity if the considered samples are cohorts.

For both estimation methods we found no effect of the number of SNPs on accuracy and little effects on precision, the latter contradicting previous simulation studies (Waples and Do 2010, Robinson and Moyer 2013, Waples et al. 2014).

Globally, for all tested sample sizes and time intervals above six years, the corrected Temporal method performed better in terms of accuracy than the Linkage Disequilibrium method but it performed worse in terms of precision as expected (Waples and Do 2010). Thus the Temporal method appears to be the better method for estimating N_e for thornback ray. However, it requires a high sampling effort with samples separated in time, which was not available for the empirical data set.

Empirical genetic data from thornback rays sampled in the Bay of Biscay were used to explore the estimation of N_e of a

non-model species. Genetic markers were obtained from RADseq thus individuals missing data for a given SNP were common (Nunziata and Weisrock 2018). We studied the effects of the amount of missing data and of the minor allelic frequency of selected SNPs on Linkage Disequilibrium N_e estimates. Depending on the filter combining proportion of missing data and minor allelic frequency, N_e estimates between 200.5 and 2681.0 individuals were obtained. This corresponds to a multiplier of 13, which is much larger than the range of values found in the simulation study. Thus, the effects of the proportion of missing data and the range of selected minor allelic frequencies had a larger effect than sample size (or time interval for the Temporal method). However, as discussed above, the effect of sample size depends on population size, thus the simulation results and the empirical results might not be completely comparable. The small sample size of the empirical data set might have increased the effects of missing data and selected allelic frequencies. Nevertheless, minor allelic frequencies in the simulated data were around 0.5, thus N_e estimates from empiric dataset with minimum MAF=0.45 were roughly comparable to the N_e estimates from simulated data in terms of selected SNPs. This selected MAF range led to the smallest empirical N_e estimates compared to the other MAF filters.

There exists only one empirical N_e estimate for a thornback ray population in the literature. Chevolot et al. (2008) estimated the effective population size of the Irish Sea population to be between 283 and 512 individuals depending on the estimation method used, which is comparable to our estimates for the Bay of Biscay.

4.2 Simulation assumptions

Thornback ray demography: In our simulations, thornback ray population maintains a constant size and age distribution over time which means that gene transmission does not vary in time. This is an ideal situation for applying the correction factor *C(T)* developed by Jorde and Ryman (1995). We also assumed that reproduction and survival of individuals in year *t* was independent of their reproduction status in year *t-1*. For ray species which do not provide any parental care, this assumption should not be violated.

Estimators: In this paper, we choose to use only two estimators: the Temporal method corrected for overlapping generations and the Linkage Disequilibirum method. Only one single point estimator was chosen from *a priori* knowledge as the other available single point methods (Molecular coanscestry, Heterozygosity excess) are known to perform poorly for population sizes above 1000 individuals (Pudovkin et al. 1996, Luikart and Cornuet 1999, Wang 2016).

Evolutionary forces: N_e estimators assume that of the four evolutionary forces (mutation, migration, selection and genetic drift), only genetic drift is responsible for the signal in the genetic data. This condition is fulfilled by our simulation set-up but is unlikely to occur in natural populations. Mutation probably has little effect on estimates (Waples and Do 2010). For the Linkage Disequilibrium method, migration will lead to an underestimation of N_e because the Linkage Disequilibrium caused by migration will be interpreted as genetic drift by the estimator (Wang 2005, Gilbert and Whitlock 2015). Selection can also cause non random association of loci which will again be interpreted as genetic drift by the Linkage Disequilibrium estimator (Waples and Do 2010). The Temporal method (with no correction factor applied) may also lead to underestimation in admixed populations under natural selection (Araki et al. 2007).

Unlinked loci: In the simulation study, we assumed unlinked SNPs while this assumption is far from biological reality, in particular if a large number of SNPs is used (Waples et al. 2016). Indeed, the slight improvement of precision found by increasing the number of SNPs in both methods will hardly be true in reality. The number of truly independent SNPs is equal to the number of chromosomes, which could be rather limited in some organisms. The finite number of chromosomes will create linkage disequilibrium (more precisely gametic linkage disequilibrium) only due to physical linkage between SNPs, rather than N_e changes (Waples et al. 2016). However, fortunately, thornback ray have a high number of chromosomes (Nygren et al. 1971) making this assumption less problematic.

4.3 Concluding remarks

In conclusion, for a non-model species, special attention should be paid to the interpretation of N_e estimates as large underestimation or overestimation can occur. Thus simulating a population mimicking the demography of the species of interest is needed to gain insights into the expected bias and precision. For thornback ray, the effects of demography were found by simulation to downwardly bias N_e estimates by around 30% for the Linkage Disequilibrium method. The effects of demography on N_e estimates using the Temporal method were less clear as the method is already corrected for overlapping generations. For time intervals above 6 years (equivalent to age at maturity) this method appeared to be the most appropriate for estimating N_e. For this method, a time interval between sampled cohorts larger than age-at-maturity should be preferred. For both methods, appropriate sample sizes depend on absolute population size.

Estimates of N_e based on empirical data were highly variable depending on the proportion of missing data and the range of allelic frequencies included. Our results showed that the amount of missing data had a bigger effect than the minimum allele frequency on N_e estimates. Further studies are needed to explore the effect of these two factors. Moreover, missing data is a typical problem for RADseq, thus it is an unavoidable problem which reinforces the need to better understand its effects on N_e estimates.

5 Acknowledgments

We acknowledge funding from the French "Agence Nationale de la Recherche" (ANR) for the GenoPopTaille project and from the Fondation Total for the GenoPopTaille-Capsules project. FM thanks Ifremer for a PhD studentship and

a grant who allowed her to work at the NOAA of Seattle. FM thanks Bo Peng for help with the simuPOP code. The authors thank Florence Cornette for help with the lab work. We acknowledge the surveys RAIECOAM and RAIEbeca for thornback ray sampling. The authors thank the UMR 8199 LIGAN-PM Genomics platform (Lille, France) which belongs to the 'Federation de Recherche' 3508 Labex EGID (European Genomics Institute for Diabetes; ANR-10-LABX-46) and was supported by the ANR Equipex 2010 session (ANR-10-EQPX-07-01; 'LIGAN-PM'). The LIGAN-PM Genomics platform (Lille, France) is also supported by the FEDER and the Region Nord-Pas-de-Calais-Picardie.

6 References

Araki, H., Waples, R.S., and Blouin, M.S. 2007. A potential bias in the temporal method for estimating N_e in admixed populations under natural selection. Mol. Ecol. **16**(11): 2261–2271. doi:10.1111/j.1365-294X.2007.03307.x.

Beaumont, A.R., Boudry, P., and Hoare, K. 2010. Biotechnology and genetics in fisheries and aquaculture. *In* 2nd ed. Blackwell, Chichester ; Ames, Iowa.

Bishop, J.M., Leslie, A.J., Bourquin, S.L., and O'Ryan, C. 2009. Reduced effective population size in an overexploited population of the Nile crocodile (*Crocodylus niloticus*). Biol. Conserv. **142**(10): 2335–2341. doi:10.1016/j.biocon.2009.05.016.

Caballero, A. 1994. Developments in the prediction of effective population size. Heredity **73**(6): 657–679. doi:10.1038/hdy.1994.174.

Cervantes, I., Pastor, J.M., Gutiérrez, J.P., Goyache, F., and Molina, A. 2011. Computing effective population size from molecular data: The case of three rare Spanish ruminant populations. Livest. Sci. **138**(1–3): 202–206. doi:10.1016/j.livsci.2010.12.027.

Chevolot, M., Ellis, J.R., Rijnsdorp, A.D., Stam, W.T., and Olsen, J.L. 2008. Temporal changes in allele frequencies but stable genetic diversity over the past 40 years in the Irish Sea population of thornback ray, *Raja clavata* . Heredity **101**(2): 120–126. doi:10.1038/hdy.2008.36.

Davey, J.W., and Blaxter, M.L. 2010. RADSeq: next-generation population genetics. Brief. Funct. Genomics **9**(5–6): 416–423. doi:10.1093/bfgp/elq031.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., and Ovenden, J.R. 2014. NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size Ne from genetic data. Mol. Ecol. Resour. **14**(1): 209–214. doi:10.1111/1755-0998.12157.

Engen, S., Lande, R., Saether, B.-E., and Gienapp, P. 2010. Estimating the ratio of effective to actual size of an age-structured population from individual demographic data: Effective population size of age-structured populations. J. Evol. Biol. **23**(6): 1148–1158. doi:10.1111/j.1420-9101.2010.01979.x.

Fox, C.W., and Wolf, J.B. (*Editors*). 2006. Evolutionary genetics: concepts and case studies. Oxford University Press, Oxford ; New York.

Francuski, L., and Milankov, V. 2015. Assessing spatial population structure and heterogeneity in the dronefly: spatial population structure in the dronefly. J. Zool.: n/a-n/a. doi:10.1111/jzo.12278.

Gilbert, K.J., and Whitlock, M.C. 2015. Evaluating methods for estimating local effective population size with and without migration: estimating N_e in the presence of migration. Evolution **69**(8): 2154–2166. doi:10.1111/evo.12713.

Gislason, H., Daan, N., Rice, J.C., and Pope, J.G. 2010. Size, growth, temperature and the natural mortality of marine fish: natural mortality and size. Fish Fish. **11**(2): 149–158. doi:10.1111/j.1467-2979.2009.00350.x.

Hamilton, M.B. 2009. Population genetics. Wiley-Blackwell, Chichester, UK ; Hoboken, NJ.

Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R.S., Ruegg, K., and Palstra, F. 2011. Understanding and estimating effective population size for practical application in marine species management: applying effective population size estimates to marine species management. Conserv. Biol. **25**(3): 438–449. doi:10.1111/j.1523-1739.2010.01637.x.

Jorde, P.E. 2012. Allele frequency covariance among cohorts and its use in estimating effective size of age-structured populations: ESTIMATING EFFECTIVE SIZE OF AGE-STRUCTURED POPULATIONS. Mol. Ecol. Resour. **12**(3): 476–480. doi:10.1111/j.1755-0998.2011.03111.x.

Jorde, P.E., and Ryman, N. 1995. Temporal allele frequency change and estimation of effective size in populations with overlapping generations. Genetics **139**(2): 1071090.

Juarez, R.L., Schwartz, M.K., Pilgrim, K.L., Thompson, D.J., Tucker, S.A., Smith, J.B., and Jenks, J.A. 2015. Assessing temporal genetic variation in a cougar population: influence of harvest and neighboring populations. Conserv. Genet. doi:10.1007/s10592-015-0790-5.

Luikart, G., and Cornuet, J.-M. 1999. Estimating the effective number of breeders from heterozygote excess in progeny. Genetics **151**(3): 1211–1216.

Luikart, G., Ryman, N., Tallmon, D.A., Schwartz, M.K., and Allendorf, F.W. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. Conserv. Genet. **11**(2): 355–373. doi:10.1007/s10592-010-0050-7.

Marandel, F., Lorance, P., Berthelé, O., Trenkel, V.M., Waples, R.S., and Lamy, J.-B. Submitted. Estimating effective population size of large populations, is it feasible? A study case on marine fish.

Marandel, F., Lorance, P., and Trenkel, V.M. 2016. A Bayesian state-space model to estimate population biomass with catch and limited survey data: application to the thornback ray (*Raja clavata*) in the Bay of Biscay. Aquat. Living Resour. **29**(2): 209. doi:10.1051/alr/2016020.

Nomura, T. 2002. Effective size of populations with unequal sex ratio and variation in mating success. J. Anim. Breed. Genet. **119**(5): 297–310. doi:10.1046/j.1439-0388.2002.00347.x.

Nunziata, S.O., and Weisrock, D.W. 2018. Estimation of contemporary effective population size and population declines using RAD sequence data. Heredity **120**(3): 196–207. doi:10.1038/s41437-017-0037-y.

Peng, B., and Kimmel, M. 2005. simuPOP: a forward-time population genetics simulation environment. Bioinformatics **21**(18): 3686–3687. doi:10.1093/bioinformatics/bti584.

Pilger, T.J., Gido, K.B., Propst, D.L., Whitney, J.E., and Turner, T.F. 2015. Comparative conservation genetics of protected endemic fishes in an arid-land riverscape. Conserv. Genet. **16**(4): 875–888. doi:10.1007/s10592-015-0707-3.

In preparation

Poulard, J., and Blanchard, F. 2005. The impact of climate change on the fish community structure of the eastern continental shelf of the Bay of Biscay. ICES J. Mar. Sci. **62**(7): 1436–1443. doi:10.1016/j.icesjms.2005.04.017.

Poulard, J.-C., and Trenkel, V.M. 2007. Do survey design and wind conditions influence survey indices? Can. J. Fish. Aquat. Sci. **64**(11): 1551–1562. doi:10.1139/f07-123.

Pudovkin, A.I., Zaykin, D.V., and Hedgecock, D. 1996. On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. Genetics **144**(1): 383–387.

R Foundation for Statistical Computing. 2008. R Development Core Team (2008). R: A language and environment for statistical computing. Austria, Vienna. Available from http://www.R-project.org.

Robinson, J.D., and Moyer, G.R. 2013. Linkage disequilibrium and effective population size when generations overlap. Evol. Appl. **6**(2): 290–302. doi:10.1111/j.1752-4571.2012.00289.x.

Serbezov, D., Jorde, P.E., Bernatchez, L., Olsen, E.M., and Vollestad, L.A. 2012. Short-Term Genetic Changes: Evaluating Effective Population Size Estimates in a Comprehensively Described Brown Trout (Salmo trutta) Population. Genetics **191**(2): 579–592. doi:10.1534/genetics.111.136580.

Serra-Pereira, B., Figueiredo, I., Farias, I., Moura, T., and Gordo, L.S. 2008. Description of dermal denticles from the caudal region of *Raja clavata* and their use for the estimation of age and growth. ICES J. Mar. Sci. **65**(9): 1701–1709. doi:10.1093/icesjms/fsn167.

Trask, A.E., Bignal, E.M., McCracken, D.I., Piertney, S.B., and Reid, J.M. 2017. Estimating demographic contributions to effective population size in an age-structured wild population experiencing environmental and demographic stochasticity. J. Anim. Ecol. **86**(5): 1082–1093. doi:10.1111/1365-2656.12703.

Wang, J. 2016. A comparison of single-sample estimators of effective population sizes from genetic marker data. Mol. Ecol. **25**(19): 4692–4711. doi:10.1111/mec.13725.

Wang, Q., Arighi, C.N., King, B.L., Polson, S.W., Vincent, J., Chen, C., Huang, H., Kingham, B.F., Page, S.T., Farnum Rendino, M., Thomas, W.K., Udwary, D.W., Wu, C.H., and the North East Bioinformatics Collaborative Curation Team. 2012. Community annotation and bioinformatics workforce development in concert--Little Skate Genome Annotation Workshops and Jamborees. Database **2012**(0): bar064–bar064. doi:10.1093/database/bar064.

Waples, R.K., Larson, W.A., and Waples, R.S. 2016. Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. Heredity **117**(4): 233–240. doi:10.1038/hdy.2016.60.

Waples, R.S., Antao, T., and Luikart, G. 2014. Effects of overlapping generations on Linkage Disequilibrium estimates of effective population size. Genetics **197**(2): 769–780. doi:10.1534/genetics.114.164822.

Waples, R.S., and Do, C. 2010. Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evol. Appl. **3**(3): 244–262. doi:10.1111/j.1752-4571.2009.00104.x.

Waples, R.S., Do, C., and Chopelet, J. 2011. Calculating N_e and N_e / N in age-structured populations: a hybrid Felsenstein-Hill approach. Ecology **92**(7): 1513–1522. doi:10.1890/10-1796.1.

Waples, R.S., Luikart, G., Faulkner, J.R., and Tallmon, D.A. 2013. Simple life-history traits explain key effective population size ratios across diverse taxa. Proc. R. Soc. B Biol. Sci. **280**(1768): 20131339–20131339. doi:10.1098/rspb.2013.1339.

Waples, R.S., and Yokota, M. 2006. Temporal estimates of effective population size in species with overlapping generations. Genetics **175**(1): 219–233. doi:10.1534/genetics.106.065300.

Wilson, C.C., McDermid, J.L., Wozney, K.M., Kjartanson, S., and Haxton, T.J. 2014. Genetic estimation of evolutionary and contemporary effective population size in lake sturgeon (*Acipenser fulvescens* Rafinesque, 1817) populations. J. Appl. Ichthyol. **30**(6): 1290–1299. doi:10.1111/jai.12615.

Wyffels, J., L. King, B., Vincent, J., Chen, C., Wu, C.H., and Polson, S.W. 2014. SkateBase, an elasmobranch genome project and collection of molecular resources for chondrichthyan fishes. F1000Research. doi:10.12688/f1000research.4996.1.

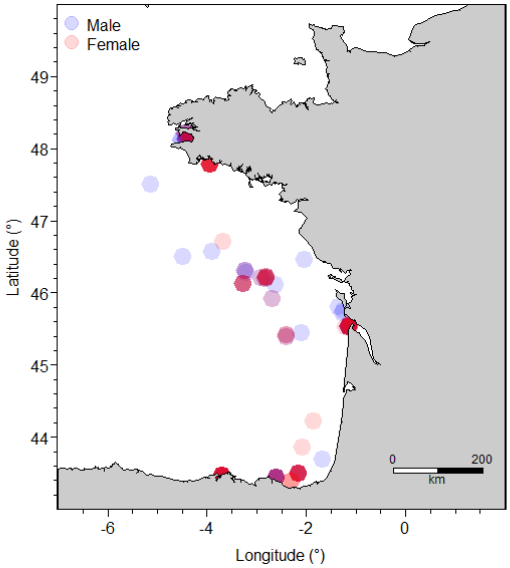
In preparation

7 Supplementary material

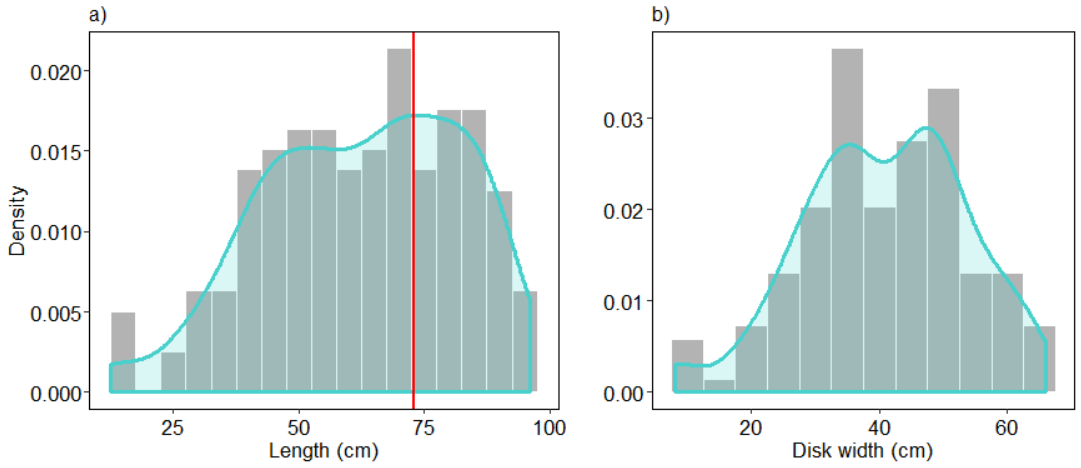
Supplementary Table 1: Demographic data used in AgeNe software (Waples et al. 2011) and FactorC software (Jorde and Ryman 2012). *S*: mean age-specific survival rate; *L*: probability to survive to age; B: mean number of offspring.

Age	***S***	***L***	***B***
0		1	0
1	0.2	0.2	0
2	0.4	0.08	0
3	0.55	0.044	0
4	0.67	0.02948	0
5	0.73	0.02152	0
6	0.77	0.016571	140
7	0.8	0.013257	140
8	0.815404	0.010809	140
9	0.832806	0.009002	140
10	0.846161	0.007617	140
11	0.856647	0.006525	140
12	0.865036	0.005645	140
13	0.8	0.004516	140
14	0.65	0.002935	140
15	0.5	0.001468	140
16	0.33	0.000484	140
17	0.17	0.000082	140
18	0.01	0.000082	140

In preparation



Supplementary figure 1: Map of the 159 sampled individuals in the Bay of Biscay. Locations are pointed with transparency: darker is the point, biggest is the density of points.



Supplementary figure 2: a) Size repartition of the sampled individuals, theoretical size at maturity is in red (Serra Pereira et al. 2008); b) Disk width of the sampled individuals

CHAPITRE 6

Évaluer la taille de population efficace des populations marines de grandes tailles : est-ce possible ?

IRD

Ce chapitre étudie l'estimation de la taille de population efficace pour des populations marines de grandes tailles. Pour ce faire des populations de grandes tailles (1 000 000 d'individus) ont été simulées, chacune présentant une démographie de raie bouclée. La taille de population efficace a ensuite été estimée grâce à la méthode Linkage Disequilibrium afin d'étudier la précision et les biais potentiels de ces estimations. Cette partie fait l'objet d'un article soumis à Fish and Fisheries.

1. Les populations marines : des populations de très grande tailles

Comme énoncé dans le chapitre 5, les estimateurs de la taille de population efficaces ont été développés et utilisés sur des populations terrestres depuis des décennies (Schwartz et al., 2007). Nombres d'entre eux sont explicitement non adaptés à l'estimation de N_e pour des populations de grandes tailles (méthode "Heterozygote Excess" Pudovkin et al., 1996, méthode "Molecular Coancestry" Nomura, 2008), alors que d'autres le sont de manière implicite (méthode Linkage Disequilibrium, Waples, 2016). Cette incapacité d'estimation résulte d'une dilution du signal génétique résultant de N_e dû aux grandes tailles de populations (Waples, 2016).

Pour la majorité des populations naturelles, N_e et N sont différentes. La relation entre N_e et N est une notion clé pour comprendre l'effet de l'exploitation sur les espèces marines. Pour les populations marines, les ratios N_e/N sont souvent estimés autour de valeurs très faibles comme 10^{-5} et non autour de 0.5, une valeur répandue pour les populations de vertébrés terrestres. La table ci-dessous regroupe quelques valeurs de ratios N_e/N estimés dans la littérature.

Population	**Lieu**	N_e/N	**Taille absolue considérée**	**Méthode utilisée pour estimée N_e**	**Référence**
Acoupa blanc (*Atractoscion nobilis*)	Ecloserie (Californie)	0,27 - 0,40	Adultes	Linkage Disequilibrium	(Bartley et al., 1992)
Truite Arc en Ciel (*Oncorhynchus mykiss*)	Ecloserie (Californie)	0,90	Adultes	Linkage Disequilibrium	(Bartley et al., 1992)
Saumon quinnat (*Oncorhyunchus tshawytscha*)	Rivière Sacramento	0,013 - 0,043	Adultes	Linkage Disequilibrium	(Bartley et al., 1992)
Huître du Pacifique (*Crassostrea gigas*)	Ecloserie (Etat de Washington)	$< 10^{-6}$	Adultes	Méthode temporelle	(Hedgecock et al., 1992)
Morue (*Gadus morhua*)	Mer du Nord	$3{,}9*10^{-5}$	Adultes	Méthode temporelle	(Hutchinson et al., 2003)
Dorade (*Pagrus auratus*)	Nouvelle Zélande	$1.8 - 2.8*10^{-5}$	Non précisée	Méthode temporelle	(Hauser et al., 2002)
Plie d'Europe (*Pleuronectes platessa*)	Islande et Mer du Nord	$2*10^{-5}$	Non précisée	Méthode temporelle	(Hoareau et al., 2005)

TABLE 6.1 : Ratios N_e/N pour quelques populations marines.

Hare et al. (2011) propose une explication à ces faibles ratios : les populations marines de grandes tailles seraient plus sensible à la dérive génétique et à la consanguinité induites par l'exploitation que ce que suggèrent leur taille de population absolue. Plus récemment, Waples (2016) démontre que ces faibles ratios s'expliquerait par l'incapacité des estimateurs disponibles d'estimer correctement N_e pour des populations idéales de grandes tailles (N=N_e=10^6). Le chapitre 6 explore cette incapacité d'estimation de N_e dans le cas de grandes populations de raie bouclée.

2. Article : Evaluer la taille de population efficace des populations de grandes tailles : est-ce possible ? Application à la raie bouclée

Soumis à Fish and fisheries (Ghoti) - 26 avril 2018

Florianne Marandel[1], Pascal Lorance[1], Olivier Berthele[1], Verena M. Trenkel[1], Robin S. Waples [2], Jean-Baptiste Lamy [3]

RESUME - L'exploitation durable des populations marines est une tâche difficile basée sur leurs abondances actuelles et passées. Les données issues de la pêche commerciale peuvent se réveler peu abondantes et peu fiables les rendant inappropriées pour la modélisation quantitative. Dans ce cas de figure, il est intéressant d'avoir des estimations d'abondance basées sur des données alternatives notamment génétiques. Une méthode intéressante et indépendante des données de pêche est l'estimation de la taille de population efficace (N_e). Dans cet article, un état de l'art sur l'estimation de N_e à partir de données empiriques a été effectué. Conjointement, un travail de simulation a été effectué afin d'étudier la faisabilité de l'estimation de N_e pour les populations de poissons de grandes tailles grâce aux méthodes d'estimations disponibles. L'état de l'art de 26 études a mis en évidence que les estimations publiées de N_e sont toutes très similaires malgré des différences dans les espèces étudiées. Une population de poissons structurée en âge a été simulée avec plusieurs tailles de population absolues (de 1 000 à 1 000 000 d'individus) et la taille de population efficace a été estimé par la méthode "Linkage Disequilibrium". Les résultats montrent que pour des tailles moyennes de populations (1 000 000 d'individus) et des tailles d'échantillons répandues (50 individus), la probabilité d'obtenir des estimations négatives de N_e est élevée (>50%). Ces valeurs négatives sont généralement interprétées comme une taille de population efficace infinie. De plus, les tailles de populations efficaces estimées sont toujours biaisées. Les simulations indiquent qu'environ 1% de la population devrait être échantillonnée afin d'obtenir des estimations précises. En conclusion, obtenir des estimations fiables et précises de la taille de population efficace de populations marines de grandes tailles semble hors de portée actuellement.

[1] Ifremer, rue de l'île d'Yeu, BP 21105, 44311 Nantes Cedex 3, France

[2] Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, Washington, United-States

[3] Ifremer, Génétique et pathologie des mollusques marins, La Tremblade sur mer, France

Accepté par Fish and Fisheries (Ghoti)

Estimating effective population size of large marine populations, is it feasible?

Florianne Marandel[1*], Pascal Lorance[1], Olivier Berthelé[1], Verena M. Trenkel[1], Robin S. Waples[2], Jean Baptiste Lamy[3]

[1]Ifremer, Ecologie et Modèles pour l'Halieutique, Nantes, France
[2] Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, Washington
[3]Ifremer, Génétique et Pathologie des Mollusques Marin, La Tremblade, France

Abstract – Sustainable exploitation of marine populations is a challenging task relying on information about their current and past abundance. Fisheries related data can be scarce and unreliable making them unsuitable for quantitative modeling. One fishery independent method that has attracted attention in this context consists in estimating the effective population size (N_e), a concept founded in population genetics. We reviewed recent empirical studies on N_e and carried out a simulation study to evaluate the feasibility of estimating N_e in large fish populations with the currently available methods. The detailed review of 26 studies found that published empirical N_e values were very similar despite differences in species and total population sizes (N). Genetic simulations for an age structured fish population were carried out for a range of population and samples sizes and N_e was estimated using the Linkage Disequilibrium method. The results showed that already for medium sized populations (1 million individuals) and common sample sizes (50 individuals), negative estimates were likely to occur which for real applications is commonly interpreted as indicating very large (infinite) N_e. Moreover, on average N_e estimates were negatively biased. The simulations further indicated that around 1% of the total number of individuals might have to be sampled to ensure sufficiently precise estimates of N_e. For large marine populations this implies rather large samples (several thousands to millions of individuals). If however such large samples were to be collected, many more population parameters than only N_e could be estimated.

Key words – Effective population size, Linkage-Disequilibrium, fish, management, census population size, simulation

1 N_E ESTIMATION FOR LARGE MARINE POPULATIONS

Fishery science is driven by the need to produce scientific advice for the management and conservation of marine resources and ecosystems (Dankel and Edwards 2016). This motivates the collection of information on population status and biology. Increasingly, attention is paid to the genetic state of marine populations (Ovenden et al. 2015) with numerous studies being published on genetic diversity (Bryan-Brown et al. 2017), genetic population connectivity (Bryan-Brown et al. 2017), and genetic population size (Luikart et al. 2010). For example, data from the Web of Science (WoS) show that between 2000 and 2017, the annual number of publications estimating effective population size (N_e) of marine species increased six fold (Fig. S1 in supporting information, Web of Sciences). Theoretically from a genetic point of view, N_e is defined as the size of an ideal population that is experiencing the same rate of change in allele frequencies or heterozygosity as the observed population (Luikart et al. 2010). Ideal populations are made of diploid organisms with sexual reproduction, non overlapping generations, random mating, no migration, no mutation, but also no natural selection (Wright 1931). Effective population size is considered a pertinent parameter for management as it relates to rates of genetic drift and loss in genetic variation (Hare et al. 2011). Moreover, N_e is a useful concept for evaluating the genetic future of marine populations (harvested or not) as reductions in N_e are positively correlated with reductions in population viability (Soulé 1987).

The use of N_e in scientific studies has increased (Wang 2005, Leberg 2005, Luikart et al. 2010, Supplementary figure 1) which can be linked to the increased availability of molecular markers but also the continual improvement of estimation methods (Luikart et al. 2010; Wang 2016; Waples et al. 2016). In the past, N_e was considered difficult to estimate but this situation has changed (Schwartz et al. 1998; Leberg 2005). As a consequence, N_e is nowadays commonly estimated for varied marine taxa: mammals (DeWoody et al. 2017), crustaceans (Watson et al. 2016), corals (Holland et al. 2017) and fishes (Laconcha et al. 2015; Zhivotovsky et al. 2016; Pita et al. 2017). Among commercial fish species, both target (Poulsen et al. 2005; Montes et al. 2016) and bycatch species (Chevolot et al. 2008) have been studied, representing a wide range of life history strategies, habitats, population structures but also census population sizes (i.e. total number of individuals in the population including immatures, denoted N), from hundreds to billions of individuals.

*Corresponding author: florianne.marandel@ifremer.fr – (+33) 2 40 37 41 64

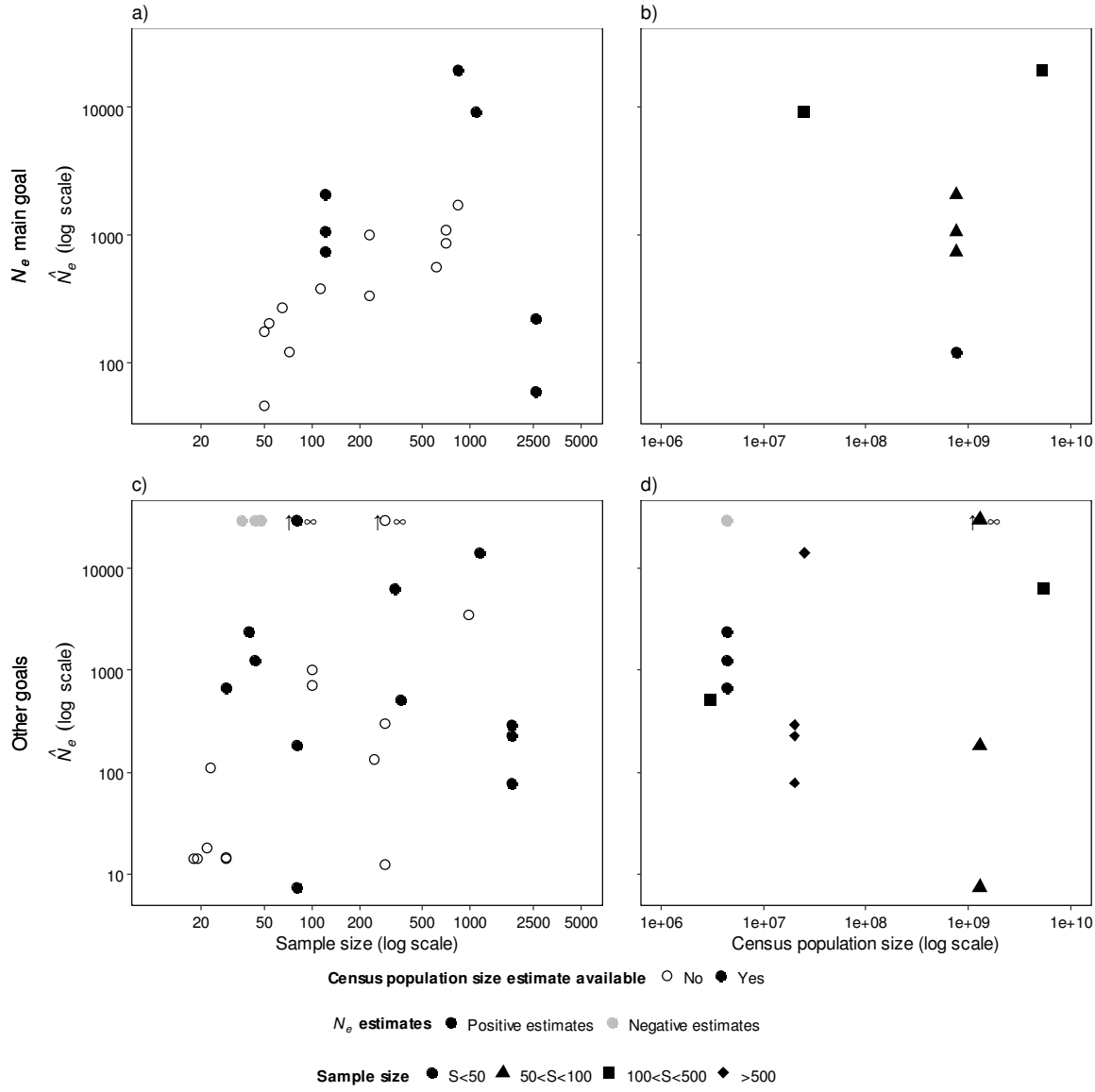


Figure 1: Meta-analysis of literature reported estimates of effective population size ($\widehat{N}_e$) in relation to sample size for a) studies with N_e estimation as main goal and c) studies with other goals, and in relation to census population size for b) studies with N_e estimation as main goal and d) studies with other goals. Infinite N_e estimates (∞) in the original publications were plotted at 30 000 while reported negative estimates were plotted in grey at 30 000. Sources for census population size estimates are provided in table S2. Points in common between panels a), b) and panels c),d) are filled in.

Many marine fish populations are very large compared to vertebrates but also present a large variety of reproductive strategies. In ideal populations as defined above all individuals have the same reproductive success making N_e equal to N. Natural populations do not have all properties of ideal populations, leading to variance in the reproductive success of individuals implying that some individuals can contribute genetically more to the next generation than others. Thus, in most natural populations N_e is smaller than N.

Genetic simulations for ideal populations indicated that N_e might not be reliably estimated for medium sized populations (N_e>10^6), independent of sample size (Waples 2016). In this seminal study Waples (2016) investigated two hypotheses which could lead to too small N_e estimates for large populations: unequal reproductive success (sweepstakes hypothesis, Hedgecock 1994) or biased estimation. He concluded that for the biological explanation to hold, few individuals would need to be responsible for most of the successful reproduction, *i.e.* the variance in reproductive success of same-age, same-sex individuals has to be orders of magnitude higher than the mean. Without ruling-out the sweepstake hypothesis, Waples (2016) suggests that biased estimation seemed to be a likely cause for creating small N_e estimates for large populations.

To evaluate the success in estimating N_e for natural marine populations, we analyzed 26 studies containing 55 empirical estimates of N_e for fish or crustaceans (tables S1 and S2 in supporting information). These studies correspond to all relevant studies published in 2016 or 2017 and the most cited studies for 2000 to 2015. Studies were separated into two categories according to the main goal of the study: estimating N_e (20 estimates in 14 studies) or other genetic questions (34 estimates in 12 studies). For studies estimating N_e as a side goal, sample sizes were smaller compared to studies estimating N_e as the main goal (313 mean & 19 - 1833 95% range for side goal; 3481 and 50-4063 95% range for main goal; Fig 1a & c). Few studies in either category used sample sizes larger than 500 individuals (25% side goal; 43% main goal; Fig 1. a & c). Only studies estimating N_e as a side goal reported negative or infinite N_e estimates (Fig 1.c & d). These negative or infinite N_e

estimates corresponded to low sample sizes (<50 individuals) or very large N (>1 billion individuals). In the reviewed studies N ranged from thousands of individuals (Zebra shark, southern Queensland Australia, Dudgeon and Ovenden 2015) to several billions (European anchovy in the Bay of Biscay, Montes et al. 2016) (Fig. 1b). No significant linear relationship was found between N_e and either N or sample size S. This was tested using a linear model with only main effects and data from the 11 studies for which N was available. The absence of relationship between N_e estimates and N seems to corroborate the simulation results obtained by Waples (2016), in particular the conclusion that N_e estimates for large populations can be biased to the point of becoming meaningless.

Several factors impact N_e estimates, while increasing sample size generally improves their accuracy and precision (Waples and Do 2010). However, for marine populations, obtaining a large number of samples (tissue, scales...) can be difficult and genotyping costs can also limit sample sizes. As a consequence most sample sizes were under 1 000 individuals in the reviewed studies (Fig. 1a & c). This led to sample sizes corresponding to less than 1% of the census population (for example, 8E-06%, for North Sea cod Poulsen et al. 2005; 2E-07% for plaice Hoarau et al. 2005, 2E-04% for European sardine Laurent and Planes 2007). Macbeth et al. (2013) showed by simulation that for the narrow-barred Spanish mackerel a sample size of 5000 individuals was necessary to estimate N_e of a population with census size N=10 000 using the Linkage Disequilibrium method (see below for details regarding this method). This result emphasizes the need for appropriate sampling designs for estimating N_e. Currently there are few recommendations available for appropriate sampling designs for estimating N_e as this is expected to be species dependent. For elasmobranchs, Dudgeon et al. 2012 advised that 50 individuals were sufficient for N_e <200 individuals while in this paper we show by simulation that, for a thornback ray like elasmobranch species assuming N_e <100 (N=1000 individuals), 300 sampled individuals would be needed for precise (though biased) estimation (see below). Other than the sampling design, the type (microsatellites or SNPs) and the number of markers can have a large effect on N_e estimates (Waples and Do 2010, F. Marandel unpublished results).

Numerous methods and estimators are available for estimating contemporary N_e. However, two approaches dominate the field: temporal estimation which requires temporally spaced samples from a population and single-point estimation which requires a sample from only a single point in time. Among the two approaches, the most popular method is the single-point Linkage Disequilibrium Method (LDM, Hill 1981, Waples et al. 2014). It was used for 22 estimates among the 55 estimates provided in the reviewed studies, while the Temporal Method (TM, Jorde and Ryman 1995) based on temporal changes in allele frequencies was used for 12 estimates and the Pseudo Likelihood Method (PLM, Wang 2001) also based on temporal changes in allele frequencies for seven estimates. Only 14 estimates used other methods. LDM, TM and PLM have been widely reviewed for various species (Schwartz et al. 1998; Wang 2005; Waples et al. 2014) with emphasis on the need for considering the life history of the studied species to obtain reliable N_e estimates or even to be able to interpret correctly N_e estimates. An example is the bias induced in N_e estimates by overlapping generations (which occurs in a natural population in contrast to an ideal population), i.e. where more than one breeding generation is present at any one time. There are several ways to minimize this bias in TM, notably using a long time lag between temporal samples (for example a generation length) or using a bias correction. Indeed, two decades ago, a correction factor for estimating N_e for species with overlapping generations was developed by Jorde and Ryman (1995) for TM. The calculation of this correction factor requires knowledge of life history traits, which might explain why it is not always used.

To further explore the (non-)feasibility of estimating effective population size for large populations using commonly used sample sizes, we present results from a simulation study in the next section. In contrast to Waples (2016) we simulated overlapping generations based on life history traits of thornback ray (*Raja clavata*), an elasmobranch widely distributed in European waters. Elasmobranchs are generally more vulnerable to fishing than teleosts and have smaller population size. Census population size of this species in the Northeast Atlantic might be millions of individuals (Marandel et al. 2016). Thus elasmobranchs are of interest for N_e estimation both in terms of conservation and technical applicability of the method. For N_e estimation we chose the Linkage Disequilibrium method as it is still the most widely-used method.

2 GENETIC SIMULATION OF A LARGE POPULATION

2.1 Method

Genetic simulations were set up mimicking thornback ray life history traits, i.e. a low fecundity with medium to high survival (Supplementary table S3). Populations of N individuals were simulated for 151 years but only the last year was used for estimating N_e. Life history traits were used in two ways as in Waples et al. (2014): (1) to calculate the expected (demographic) effective population size $E[N_e]$ (AgeNe software,Waples et al. 2011), (2) to carry out simulations to obtain age-structured genetic data (simuPOP module, Peng and Kimmel 2005) to which the LDM estimator of N_e was applied (Fig. S2).

The expected (demographic) effective population size per generation $E[N_e]$ was calculated using the AgeNe software based on life history traits (Felsenstein-Hill method, Waples et al. 2011). The method assumes a stable population (thus stable age structure) and constant survival and fecundities at age (Waples et al 2014, eq. 1):

$$E[N_e] = \frac{4\,N_1 G}{V_{k.}+2} \quad (1)$$

where N_1 is the number of age 1 individuals in the population and G is the generation length (= mean age of parents of newborns). Both depend on survival and fecundity rates, in addition N_1 depends on population size N. $V_{k.}$ is the inter-individual variance of lifetime reproductive success; the mean life time reproductive success for a stable population is 2, hence the 2 in the denominator of equation 1.

All modeled populations in simuPOP were simulated with a 1:1 sex ratio and random assignment of age at initialization (year 0). Newborn individuals were generated by drawing one male and one female from the pool of potential parents. All potential parents of the same sex and age had an equal probability to become a parent. Two hundred biallelic genetic markers corresponding to SNPs (Single Nucleotide Markers) were simulated with an initial allele frequency of 0.5.

Preliminary simulations were conducted with 1 600 biallelic genetic markers showing that a plateau in terms of precision and accuracy of N_e estimates was reached at around 200 markers.

Table 1: Simulation design. *N* is the simulated population size used in simuPOP; *E[N_e]* is the expected N_e estimated with AgeNe.

Simulated *N*	E[N_e]	Tested sample sizes (*S*)
1000	87	50, 100, 200, 300
10 000	870	50, 100, 200, 300, 500, 1000
100 000	8700	50, 100, 200, 300, 500, 1000, 1500
1 000 000	87 000	50, 100, 200, 300, 500, 1000, 1500, 5000, 10 000

Four populations sizes were simulated, $N \in$ (1 000, 10 000, 100 000, 1 000 000) individuals, to evaluate the performance of the LDM for different census sizes. Note that the largest simulated population size was smaller than many real fish populations due to computational constraints. As simulated populations contained immatures and overlapping generations with mature individuals reproducing several times, E[N_e] of each population was smaller than the simulated population size N (E[N_e]=0.087 N, Table 1). For each population size, 30 replicates were carried out to capture the stochasticity inherent in genetic simulations. For each population replicate nine sample sizes $S \in$ (50, 100, 200, 300, 500, 1000, 1500, 5000, 10 000), were investigated (Table 1); the larger sample sizes could only be explored for the largest population sizes. Sampled individuals were randomly drawn from newborns in the last year. For each population replicate and sample size, sampling was repeated 50 times, *i.e.* for each population and sample size there were 1500 simulated data sets. All 200 simulated loci were generally used for estimation, unless the minor allele frequency was <0.05 in which case it was removed as suggest by Waples and Do (2010) to minimize sampling bias.

The Linkage Disequilibrium (LD) is the non-random association of alleles at different gene loci, e.g. allele *A* at SNP locus 1 with allele *b* at SNP locus 2. When loci are inherited independently, the frequency of the *Ab* loci association is just the product of the two allele frequencies P_A and P_b in the population. In natural populations, overlapping generations, gene flow and linked loci will influence LD in addition to finite population size.

For applying the LDM, the LD is measured by the co-variance (D) and the squared correlation (r^2) between loci. The squared correlation r^2 is defined as:

$$r^2 = \frac{D^2}{P_A P_a P_B P_b} \qquad (2)$$

where *A* and *a* are the major et minor alleles (in frequency) at SNP locus 1 and *B* and *b* are the major et minor alleles at SNP locus 2, $D = P_{AB} - P_A P_B$ and P_A, P_a, P_B and P_b are the frequencies of alleles *A*, *a*, *B* and *b* respectively. P_{AB} is the haplotype (joint) frequency of the gamete/chromosome carrying the allele *A* at locus 1 and the allele *B* at the locus 2. Thus the calculation of LD is based on allele and haplotype frequencies. However in most fishery studies, haplotype frequencies are not available as the data does not contain information on which one of the pair of chromosomes holds which allele making the exact calculation of r^2 impossible. To circumvent this obstacle, a proxy is used, called the composite measure of linkage disequilibrium. The full explanation of this proxy is out of scope of this article and it reviewed in Hamilton and Cole (2004).

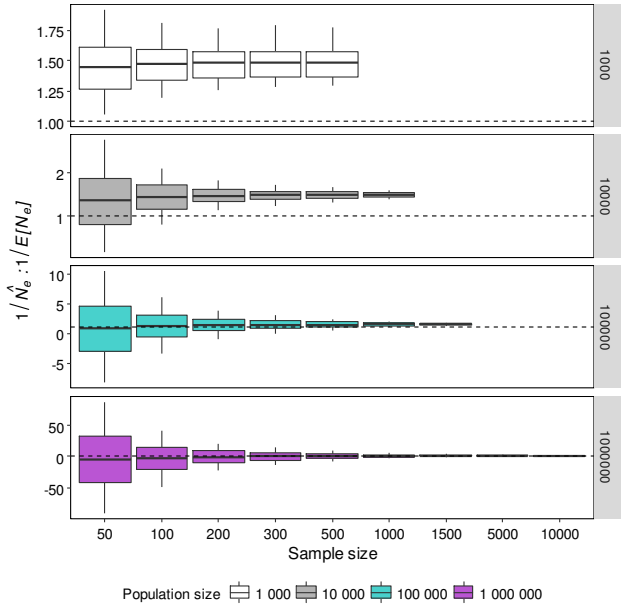


Figure 2: Ratio of inverse effective population size $1/\widehat{N_e}$ estimated with the Linkage Disequilibrium method for simulated genetic samples using NeEstimator (Do et al. 2014) and expected effective population size $1/E[N_e]$ for chosen simulation parameters calculated by AgeNe (Waples et al. 2011). Sample sizes go from 50 to 10 000 individuals (newborns). Colors: simulated population size. Box: 75 and 25 percentiles, vertical line: 95 and 5 percentiles; horizontal bar: mean estimates. Dashed line: $1/\widehat{N_e}=1/E[N_e]$

For estimating N_e based on the proxy estimate $\hat{r}^2$ adjusted for sample size S related sampling error according to Weir (1979), the following relationship was used (Waples 2006):

$$\widehat{N}_e = \frac{1/3 + \sqrt{1/9 - 2.76\hat{r}^{2\prime}}}{2\hat{r}^{2\prime}} \text{ with } \hat{r}^{2\prime} = \hat{r}^2 - 1/S - 3.19/S^2 \qquad (3)$$

Equation (3) shows that if $1/S$ is larger than $\hat{r}^2$ a negative estimate of $\widehat{N}_e$ is obtained. Thus negative estimates occur when sampling error is larger than the genetic signal (correlation between loci, eq 1), without invoking any genetic effect. The usual practitioner interpretation made is that negative N_e estimates indicate a very large effective population size, hence negative estimates are replaced by infinity (Laurie-Ahlberg and Weir 1979; Nei and Tajima 1981). In reality, negative estimates can also simply be caused by an insufficient sample size.

The estimator in eq 3 is implemented in NeEstimator V2 (Do et al. 2014) which was used for the simulated data sets. As this software does not account for overlapping generations N_e estimates will be biased to an unknown degree depending on the simulated life history (Waples et al. 2014).

Quantifying accuracy (or bias) and precision of estimates of effective population size is complicated because $\widehat{N}_e$ has a skewed distribution and can be arbitrarily large (or even negative as discussed above). Accordingly, we followed Wang (2001, 2009), who focused on bias and precision of the inverse $1/N_e$, which is proportional to the rate of genetic drift and is the signal for effective size that is detected by all genetic estimation methods. The estimates of $1/N_e$ were then compared to the inverse of the expected value $E[N_e]$ (eq 1).

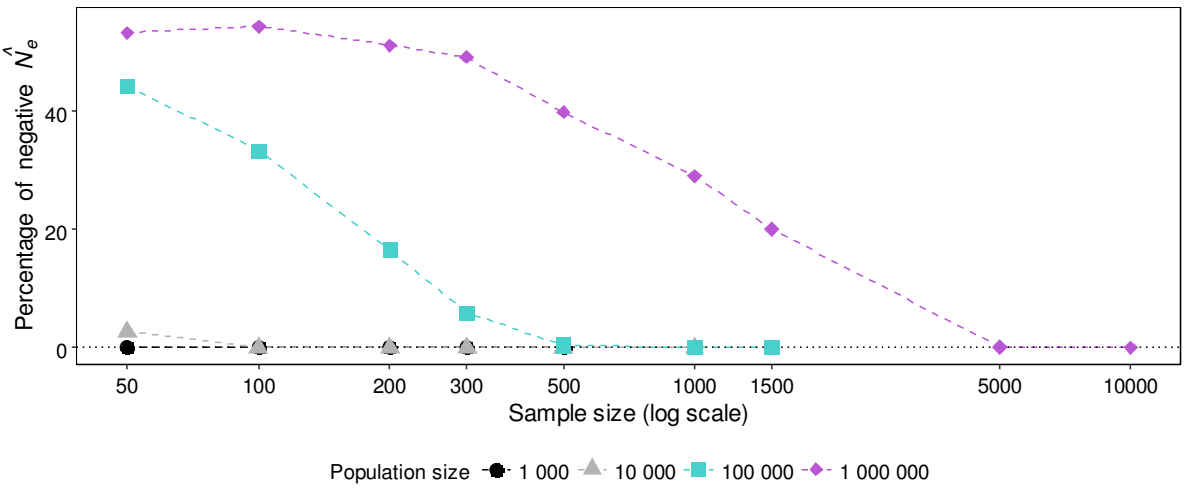


Figure 3: Percentage of negative effective population size estimates ($\hat{N}_e$) for a simulated thornback ray like population estimated with the Linkage Disequilibrium method (NeEstimator, Do et al. 2014). Colors and shapes: simulated population size.

Thus we analyzed the distribution of $\frac{1/\hat{N}_e}{1/E[N_e]} = \frac{E[N_e]}{\hat{N}_e}$. Note that the effect for $\hat{N}_e$ is then the inverse of that for $1/\hat{N}_e$, e.g. underestimation instead of overestimation. Relative bias and coefficient of variation (CV) of $1/N_e$ estimates were calculated as:

$$Relative\ bias = \frac{\mu - 1/E[N_e]}{1/E[N_e]} \quad and \quad CV = \frac{\sigma}{\mu}, \qquad (4)$$

where μ was the mean and σ the standard deviation of the 1500 $1/\hat{N}_e$ estimates.

2.2 Distribution of N_e estimates

For all simulated population sizes the interquartile range of relative estimates ($E[N_e]/\hat{N}_e$) decreased with sample size and for a given sample size was largest for the larger population sizes (Fig. 2, note different scales for y-axis). Estimates were generally positively biased though negative values occurred for the larger population sizes when sample size was small.

No negative N_e estimates for population size N=1000 were found, whatever the sample size S≥50 (Fig. 3). For N=10 000, only the smallest sample size (S=50) led to negative $\hat{N}_e$ estimates (3.5%). For N=100 000 and N=1 000 000, negative $\hat{N}_e$ estimates were absent when respectively at least 1000 or 10 000 individuals were sampled, which represents 1% of N. For sample sizes <100 individuals (the most common sample size found in the literature review above), the percentage of negative $\hat{N}_e$ reached a maximum of 53% for S=50 for N=1 000 000. Comparing results for N=10 000 with N=100 000, for sample size S=50, the number of negative N_e estimates increased by 1618%. The same comparison between N=100 000 and N=1 000 000 showed an increase of 120% of negative estimates. Thus with usual samples sizes (Fig. 1), a population of 1 000 000 individuals could easily be evaluated having an infinite N_e due to the high probability of obtaining a negative N_e estimate (>50%). Indeed, Zhivotovsky et al. (2016) attempted to estimate N_e for cod in the Barents Sea using a small sample (S= 43) and few microsatellites (13). As expected they found that all estimation methods gave negative N_e estimates.

2.3 Bias and precision of N_e estimates

In terms of relative bias, all simulated population sizes converged to a mean relative positive bias of around +50% (Fig. 4a). For all N, precision increased (CV decreased) with increasing sample size (Fig. 4b). As expected, the worst precision was obtained for N=1 000 000 and S=50 for which $1/\hat{N}_e$ was overestimated as much as 88 times for certain replicates and samples. Globally for all simulated population sizes, given a sufficient sample size, the CV for $1/\hat{N}_e$ was smaller than 0.2. Thus the sample sizes needed for stabilizing mean relative bias estimates and achieving a CV of less than 0.2 were around 1% of N for $N \in$ (10 000, 100 000 and 1 000 000). For N=1 000 it was S=50 as we did not test smaller sample sizes.

2.4 Discussion

Most marine fishes have overlapping generations and may have large population sizes (millions to billions of individuals), whereas genetic effective population size estimators generally assume discrete generations but also implicitly small population sizes. Using a simulation approach, we examined the feasibility of estimating the effective population size of a realistic fish species taking thornback ray as an example and using the popular LD method. For a given sample size, the results showed a large increase in the percentage of negative estimates with census population size. For example, in simulations for a population size of one million individuals, 200 SNPs and sample size 50 individuals, 53% of N_e estimates were negative. This means that a study attempting to estimate N_e for a real thornback ray population of one million individuals would have a 50% chance of producing a negative estimate, which could lead to the wrong conclusion that the effective population size was very large, i.e. infinite. Thornback ray populations in the Northeast Atlantic are thought to range from half a million to more than three millions individuals (F. Marandel unpublished results). Thus the percentage of negative estimates of N_e for a real thornback ray population can be expected to be even higher than what we found here if a sample of only 50 individuals is used. Waples (2016) simulated an ideal population of one million individuals and estimated N_e with 5000 sampled individuals and 100 SNPs. In this case, the percentage of negative estimates reached also 50%. Again this result for an

ideal population corroborates that estimating N_e for large real fish populations can be challenging already because of sampling difficulties. Moreover, in Waples (2016), even when N_e was estimated to be positive, the values were underestimated by as much as 99%. Thus, for real applications even when positive finite estimates of N_e are found, these estimates can still be hugely biased and imprecise (Fig 4). Note that the simulations assumed perfect genotyping, any genotyping errors will further decrease precision.

The probability of obtaining negative N_e estimates value can be reduced by increasing sample size. Our simulation study suggests that a sample size of around 1% of the census population size N might be sufficient to obtain precise (but biased) estimates using LDM, which at the same time avoids negative estimates. However, in the case of ray populations this means that appropriate samples sizes can reach several thousands of individuals. Much larger sample sizes might be necessary for teleost fish populations which obviously limits the economic and logistic feasibility of genetic effective population size studies.

A single sample method such as the LD method can easily be applied opportunistically in studies where N_e estimation is a side goal (for example, in population genetics studies), and thus, rely on small sample sizes that are not fit for this purpose. For example, Watson et al. (2016) studied the population genetic structure of the European lobster in the Irish Sea jointly with the estimation of N_e for nine sampling locations. For six locations using the LD method, N_e was estimated to be negative (with confidence intervals including infinity) and thus interpreted to be infinite. The sample sizes used in this study varied between 29 and 48 individuals which suggests that the negative N_e estimates were a consequence of the small sample sizes used rather than infinite effective population size.

In this study simulations were carried out for a thornback ray like species. While $1/\widehat{N}_e$ estimates were rather variable we found that for an appropriate sample size, the mean relative bias was around +50%. As overestimation of $1/\widehat{N}_e$ means that $\widehat{N}_e$ is underestimated, a 50% overestimation of $1/\widehat{N}_e$ corresponds to an underestimation of $\widehat{N}_e$ of around 31%. The existence of underestimation is a well-known property of the LD method for species with overlapping generations (Waples et al. 2014). The reported amount of underestimation for random samples of adults lies between 50% (mosquito) and 10% (cod) (Waples et al. 2014) with the 30% found for a ray like species for random samples of newborns lying in between. Assuming the simulations were sufficiently realistic for thornback ray, the correction of N_e estimates obtained with the LDM for a thornback ray like population might be attempted, but only if a sufficiently large sample size was used.

We now briefly discuss the assumptions made in the simulation study and their possible impacts on the results. Populations were simulated for 151 years and newborns were sampled in the final year only to estimate N_e with the LDM. The 150 first years can be considered a long burn-in to ensure reaching the equilibrium for population dynamics but also for the allele frequencies of the genetic markers. We used 200 SNPs with an allele frequency of 0.5 at the start. Using more SNPs might increase precision (Waples and Do 2010), though initial trials showed that the gain should be small, while using a different allele frequency, i.e. <0.5 minor allele frequency, would lead to more SNPs being excluded due to thresholding (SNPs with minor allele frequency <0.05 in the last year were excluded). No physical link between SNPs was assumed; technically this was achieved by coding each SNP on a different chromosome. This is an ideal situation which is not likely to happen when using empirical genetic markers. Physical linkage is expected to increase the downward bias of $\widehat{N}_e$ estimates (Waples et al. 2016). Further, we only used samples from newborns but results were similar using samples stratified by age for all ages or only mature ones (F. Marandel results not shown). We only studied the effect of sample size and its interaction with census population size and ignored other sources of errors such as genotyping errors, particular genomic or ecological features such as polyploidization, which will also impact real life estimates and probably imply that even larger samples are needed to stabilize bias and precision. Lastly, only the LDM was used for estimating N_e. Numerous other genetic estimators are available (see Wang 2016 for a complete review) but all are expected to perform poorly for small sample sizes (and several need corrections for overlapping generations).

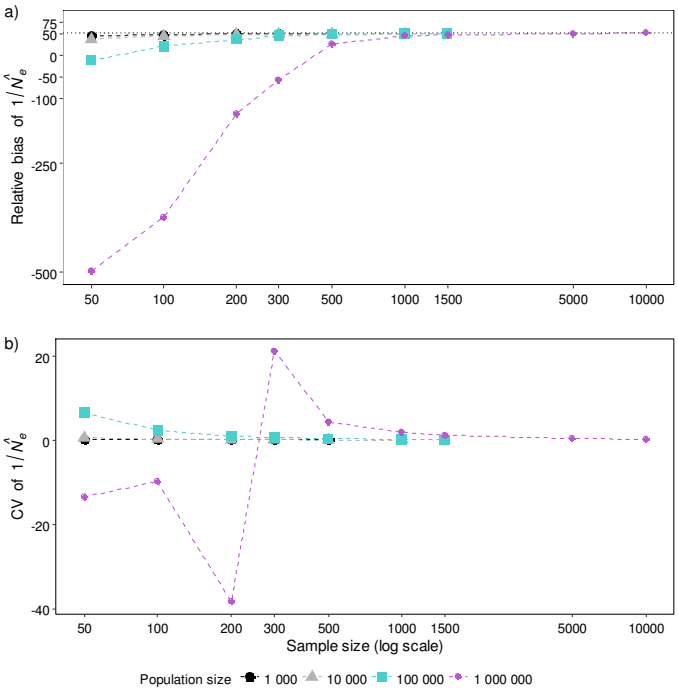


Figure 4: a) Relative bias and b) CV of inverse effective population size estimates ($1/\widehat{N}_e$) calculated by Linkage Disequilibrium method for simulated genetic samples using NeEstimator (Do et al. 2014). Relative bias is with respect to expected effective population size $1/E[N_e]$ for chosen simulation parameters calculated by AgeNe (Waples et al. 2011). Sample sizes go from 50 to 10 000 individuals (newborns). Colors: simulated population size.

3 CONCLUDING REMARKS

Numerous methods for estimating effective population size are available but they all suffer from different sources of bias and uncertainty. They also all demand high sampling effort, sometimes explicitly (*e.g.* the temporal method requires several samples separated in times) and sometimes implicitly (*e.g.* the Linkage Disequilibrium method requires a large number of individuals to be sampled). The amount of bias in genetic estimates of effective population size depends on the life history traits of the studied species (Waples et al. 2014). Thus, particular attention should be paid to the interpretation of positive finite $\widehat{N}_e$ estimates as large underestimation or overestimation can occur. Moreover, due to large population

sizes in the marine environment, negative N_e estimates are commonly found and should be interpreted with care as they might indicate insufficient sample sizes rather than infinite true N_e. In our simulations, for N=1 000 000 and S=50, half of all replicates led to negative N_e estimates suggesting sample size was insufficient. However, if by chance a positive finite N_e is estimated, it cannot be interpreted as a proof that the sample size is sufficient as half of the estimates were indeed positive for this sample size.

While theoretically it might be possible to correct N_e estimates, in practice at least two conditions need to be met. First, simulations reproducing the species life history sufficiently well will need to be carried out to estimate a species-specific bias correction factor. For a thornback ray like species we found NeEstimator underestimated N_e by 31%, while Waples et al. (2014) found a 10% bias for cod. Second, a sufficiently large number of individuals needs to be sampled, probably around 1% or more of census population size. While the first condition is time consuming, it remains feasible. However, given the large population sizes of many marine fishes, sampling 1% will require samples sizes which are often neither practical nor financially feasible. Thus, while the effective population size concept is suitable for evaluating the genetic status of marine populations, popular tools and sampling designs often miss the target (small sample size, too large population). If however precise bias-corrected estimates of effective population size can be obtained, declines in N_e track declines in N and thus, can be informative for management (Ovenden et al. 2016).

In conclusion, for large marine populations either appropriate sample sizes are used or N_e should not be estimated and reported. This study found that for a thornback ray like species sample size should be around 1% of absolute population size for the Linkage Disequilibrium method. If however such large samples are collected, other population quantities can be estimated using the same data. Absolute abundance and demographic parameters (fecundity, mortality) can be estimated with the close-kin mark-recapture (CKMR) method (Bravington et al. 2016a, 2016b). This method is based on the identification of pairs of close relatives (parents-offspring or half sibling pairs). Pairs of related individuals sampled at different locations can also on inform on migration (Feutry et al. 2017) and be used for estimating N_e (Waples et al. 2018). However, as these approaches have not been much used, further studies are needed to evaluate their merits and limits.

4 Acknowledgements

We acknowledge funding from the French "Agence Nationale de la Recherche" (ANR) for the GenoPopTaille project and the Fondation Total (project GenoPopTaille-Capsules). FM thanks Ifremer for a PhD studentship and a grant who allowed her to work at the NOAA of Seattle. FM tanks Bo Peng for his help with simuPOP code. We would like to thank two anonymous reviewers for constructive comments which helped to improve the manuscript.

5 References

Bravington, M.V., Grewe, P.M., and Davies, C.R. 2016a. Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. Nat. Commun. **7**: 13162. doi:10.1038/ncomms13162.

Bravington, M.V., Skaug, H.J., and Anderson, E.C. 2016b. Close-Kin Mark-Recapture. Stat. Sci. **31**(2): 259–274. doi:10.1214/16-STS552.

Bryan-Brown, D., Brown, C., Hughes, J., and Connolly, R. 2017. Patterns and trends in marine population connectivity research. Mar. Ecol. Prog. Ser. **585**: 243–256. doi:10.3354/meps12418.

Chevolot, M., Ellis, J.R., Rijnsdorp, A.D., Stam, W.T., and Olsen, J.L. 2008. Temporal changes in allele frequencies but stable genetic diversity over the past 40 years in the Irish Sea population of thornback ray, *Raja clavata* . Heredity **101**(2): 120–126. doi:10.1038/hdy.2008.36.

Dankel, D.J., and Edwards, C.T.T. 2016. Fishery systems and the role of management science. *In* Management science in fisheries: an introduction to simulation-based methods. Routledge, London : New York. pp. 3–15.

DeWoody, J.A., Fernandez, N.B., Brüniche-Olsen, A., Antonides, J.D., Doyle, J.M., San Miguel, P., Westerman, R., Vertyankin, V.V., Godard-Codding, C.A.J., and Bickham, J.W. 2017. Characterization of the Gray Whale *Eschrichtius robustus* Genome and a Genotyping Array Based on Single-Nucleotide Polymorphisms in Candidate Genes. Biol. Bull. **232**(3): 186–197. doi:10.1086/693483.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., and Ovenden, J.R. 2014. NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size Ne from genetic data. Mol. Ecol. Resour. **14**(1): 209–214. doi:10.1111/1755-0998.12157.

Dudgeon, C.L., Blower, D.C., Broderick, D., Giles, J.L., Holmes, B.J., Kashiwagi, T., Krück, N.C., Morgan, J.A.T., Tillett, B.J., and Ovenden, J.R. 2012. A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. J. Fish Biol. **80**(5): 1789–1843. doi:10.1111/j.1095-8649.2012.03265.x.

Dudgeon, C.L., and Ovenden, J.R. 2015. The relationship between abundance and genetic effective population size in elasmobranchs: an example from the globally threatened zebra shark *Stegostoma fasciatum* within its protected range. Conserv. Genet. **16**(6): 1443–1454. doi:10.1007/s10592-015-0752-y.

Feutry, P., Berry, O., Kyne, P.M., Pillans, R.D., Hillary, R.M., Grewe, P.M., Marthick, J.R., Johnson, G., Gunasekera, R.M., Bax, N.J., and Bravington, M. 2017. Inferring contemporary and historical genetic connectivity from juveniles. Mol. Ecol. **26**(2): 444–456. doi:10.1111/mec.13929.

Hamilton, D.C., and Cole, D.E.C. 2004. Standardizing a Composite Measure of Linkage Disequilibrium. Ann. Hum. Genet. **68**(3): 234–239. doi:10.1046/j.1529-8817.2004.00056.x.

Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R.S., Ruegg, K., and Palstra, F. 2011. Understanding and estimating effective population size for practical application in marine species management: applying effective population size estimates to marine species management. Conserv. Biol. **25**(3): 438–449. doi:10.1111/j.1523-1739.2010.01637.x.

Hedgecock, D. 1994. Does variance in reproductive success limit effective population size of marine organisms? *In* Genetics and evolution of aquatic organisms, Chapman and Hall. Beaumont M, London. pp. 122–134.

Hill, W.G. 1981. Estimation of effective population size from data on linkage disequilibrium. Genet. Res. **38**(03): 209–216.

Hoarau, G., Boon, E., Jongma, D.N., Ferber, S., Palsson, J., Van der Veer, H.W., Rijnsdorp, A.D., Stam, W.T., and Olsen, J.L. 2005. Low effective population size and evidence for inbreeding in an overexploited flatfish, plaice (*Pleuronectes platessa L.*). Proc. R. Soc. B Biol. Sci. **272**(1562): 497–503. doi:10.1098/rspb.2004.2963.

Holland, L.P., Jenkins, T.L., and Stevens, J.R. 2017. Contrasting patterns of population structure and gene flow facilitate exploration of connectivity in two widely distributed temperate octocorals. Heredity **119**(1): 35–48. doi:10.1038/hdy.2017.14.

Jorde, P.E., and Ryman, N. 1995. Temporal allele frequency change and estimation of effective size in populations with overlapping generations. Genetics **139**(2): 1071090.

Laconcha, U., Iriondo, M., Arrizabalaga, H., Manzano, C., Markaide, P., Montes, I., Zarraonaindia, I., Velado, I., Bilbao, E., Goñi, N., Santiago, J., Domingo, A., Karakulak, S., Oray, I., and Estonba, A. 2015. New Nuclear SNP Markers Unravel the Genetic Structure and Effective Population Size of Albacore Tuna (*Thunnus alalunga*). PLOS ONE **10**(6): e0128247. doi:10.1371/journal.pone.0128247.

Laurent, V., and Planes, S. 2007. Effective population size estimation on *Sardina pilchardus* in the Bay of Biscay using a temporal genetic approach: effective population sizes of sardines. Biol. J. Linn. Soc. **90**(4): 591–602. doi:10.1111/j.1095-8312.2007.00747.x.

Laurie-Ahlberg, C.C., and Weir, B.S. 1979. Allozymic Variation and Linkage Disequilibrium in Some Laboratory Populations of *Drosophila melanogaster*. Genetics **92**(4): 1295–1314.

Leberg, P. 2005. Genetic approaches for estimating the effective size of populations. J. Wildl. Manag. **69**(4): 1385–1399. doi:10.2193/0022-541X(2005)69[1385:GAFETE]2.0.CO;2.

Luikart, G., Ryman, N., Tallmon, D.A., Schwartz, M.K., and Allendorf, F.W. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. Conserv. Genet. **11**(2): 355–373. doi:10.1007/s10592-010-0050-7.

Macbeth, G.M., Broderick, D., Buckworth, R.C., and Ovenden, J.R. 2013. Linkage Disequilibrium Estimation of Effective Population Size with Immigrants from Divergent Populations: A Case Study on Spanish Mackerel (*Scomberomorus commerson*). G3amp58 GenesGenomesGenetics **3**(4): 709–717. doi:10.1534/g3.112.005124.

Marandel, F., Lorance, P., and Trenkel, V.M. 2016. A Bayesian state-space model to estimate population biomass with catch and limited survey data: application to the thornback ray (*Raja clavata*) in the Bay of Biscay. Aquat. Living Resour. **29**(2): 209. doi:10.1051/alr/2016020.

Montes, I., Iriondo, M., Manzano, C., Santos, M., Conklin, D., Carvalho, G.R., Irigoien, X., and Estonba, A. 2016. No loss of genetic diversity in the exploited and recently collapsed population of Bay of Biscay anchovy (*Engraulis encrasicolus, L.*). Mar. Biol. **163**(5). doi:10.1007/s00227-016-2866-2.

Nei, M., and Tajima, F. 1981. Genetic drift and estimation of effective population size. Genetics **98**(3): 625–640.

Ovenden, J.R., Berry, O., Welch, D.J., Buckworth, R.C., and Dichmont, C.M. 2015. Ocean's eleven: a critical evaluation of the role of population, evolutionary and molecular genetics in the management of wild fisheries. Fish Fish. **16**(1): 125–159. doi:10.1111/faf.12052.

Ovenden, J.R., Leigh, G.M., Blower, D.C., Jones, A.T., Moore, A., Bustamante, C., Buckworth, R.C., Bennett, M.B., and Dudgeon, C.L. 2016. Can estimates of genetic effective population size contribute to fisheries stock assessments? J. Fish Biol. **89**(6): 2505–2518. doi:10.1111/jfb.13129.

Peng, B., and Kimmel, M. 2005. simuPOP: a forward-time population genetics simulation environment. Bioinformatics **21**(18): 3686–3687. doi:10.1093/bioinformatics/bti584.

Pita, A., Pérez, M., Velasco, F., and Presa, P. 2017. Trends of the genetic effective population size in the Southern stock of the European hake. Fish. Res. **191**: 108–119. doi:10.1016/j.fishres.2017.02.022.

Poulsen, N.A., Nielsen, E.E., Schierup, M.H., Loeschcke, V., and GrøNkjaer, P. 2005. Long-term stability and effective population size in North Sea and Baltic Sea cod (*Gadus morhua*): effective population size in atlantic cod. Mol. Ecol. **15**(2): 321–331. doi:10.1111/j.1365-294X.2005.02777.x.

Schwartz, M.K., Tallmon, D.A., and Luikart, G. 1998. Review of DNA-based census and effective population size estimators. Anim. Conserv. **1**(4): 293–299. doi:10.1111/j.1469-1795.1998.tb00040.x.

Soulé, M. 1987. Viable Populations for Conservation. Cambridge University Press, Cambridge, UK.

Wang, J. 2001. A pseudo-likelihood method for estimating effective population size from temporally spaced samples. Genet. Res. **78**(3): 243–257.

Wang, J. 2005. Estimation of effective population sizes from data on genetic markers. Philos. Trans. R. Soc. B Biol. Sci. **360**(1459): 1395–1409. doi:10.1098/rstb.2005.1682.

Wang, J. 2009. A new method for estimating effective population sizes from a single sample of multilocus genotypes. Mol. Ecol. **18**(10): 2148–2164. doi:10.1111/j.1365-294X.2009.04175.x.

Wang, J. 2016. A comparison of single-sample estimators of effective population sizes from genetic marker data. Mol. Ecol. **25**(19): 4692–4711. doi:10.1111/mec.13725.

Waples, R.K., Larson, W.A., and Waples, R.S. 2016. Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. Heredity **117**(4): 233–240. doi:10.1038/hdy.2016.60.

Waples, R.S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci*. Conserv. Genet. **7**(2): 167–184. doi:10.1007/s10592-005-9100-y.

Waples, R.S. 2016. Tiny estimates of the N_e / N ratio in marine fishes: Are they real? J. Fish Biol. **89**(6): 2479–2504. doi:10.1111/jfb.13143.

Waples, R.S., Antao, T., and Luikart, G. 2014. Effects of overlapping generations on Linkage Disequilibrium estimates of effective population size. Genetics **197**(2): 769–780. doi:10.1534/genetics.114.164822.

Waples, R.S., and Do, C. 2010. Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evol. Appl. **3**(3): 244–262. doi:10.1111/j.1752-4571.2009.00104.x.

Waples, R.S., Do, C., and Chopelet, J. 2011. Calculating N_e and N_e / N in age-structured populations: a hybrid Felsenstein-Hill approach. Ecology **92**(7): 1513–1522. doi:10.1890/10-1796.1.

Waples, R.S., Grewe, P.M., Bravington, M.W., Hillary, R., and Feutry, P. 2018. Robust estimates of a high N_e / N ratio in a top marine predator, southern bluefin tuna. Sci. Adv. **4**(7): eaar7759. doi:10.1126/sciadv.aar7759.

Watson, H.V., McKeown, N.J., Coscia, I., Wootton, E., and Ironside, J.E. 2016. Population genetic structure of the European lobster (*Homarus gammarus*) in the Irish Sea and implications for the effectiveness of the first British marine protected area. Fish. Res. **183**: 287–293. doi:10.1016/j.fishres.2016.06.015.

Weir, B.S. 1979. Inferences about Linkage-Disequilibrium. Biometrics **35**: 235–254.

Zhivotovsky, L.A., Teterina, A.A., Mukhina, N.V., Stroganov, A.N., Rubtsova, G.A., and Afanasiev, K.I. 2016. Effects of genetic drift in a small population of Atlantic cod (*Gadus morhua* kildinensis Derjugin) landlocked in a meromictic lake: genetic variation and conservation measures. Conserv. Genet. **17**(1): 229–238. doi:10.1007/s10592-015-0774-5.

6 Competing interests

We have no competing interests.

7 Funding

The study received funding from the French “Agence Nationale de la Recherche” (project GenoPopTaille, contract ANR-14-CE02-0006-01), the Fondation Total (project GenoPopTaille-Capsules) and the European Union (projet Pandora).

Supporting information

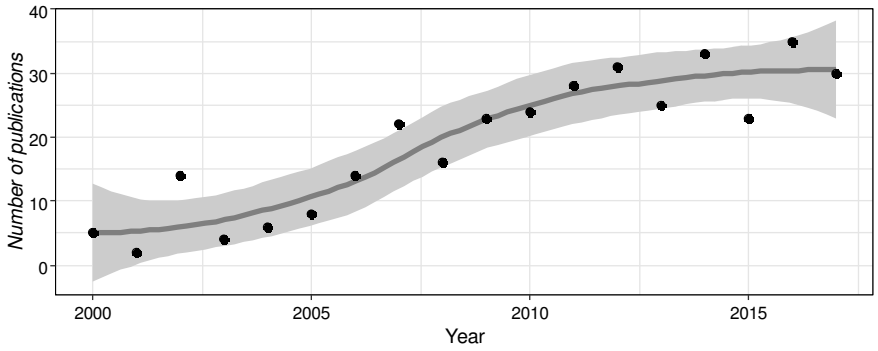


Figure S1: Number of peer-reviewed publications related to effective population size in marine population (from the Web of Science consulted on 23.01.2018, Search: TS = ("effective population size" AND marine) AND PY= (2000-2017)). Total number of publications: 343. Line: loess regression; shape: 95% intervals.

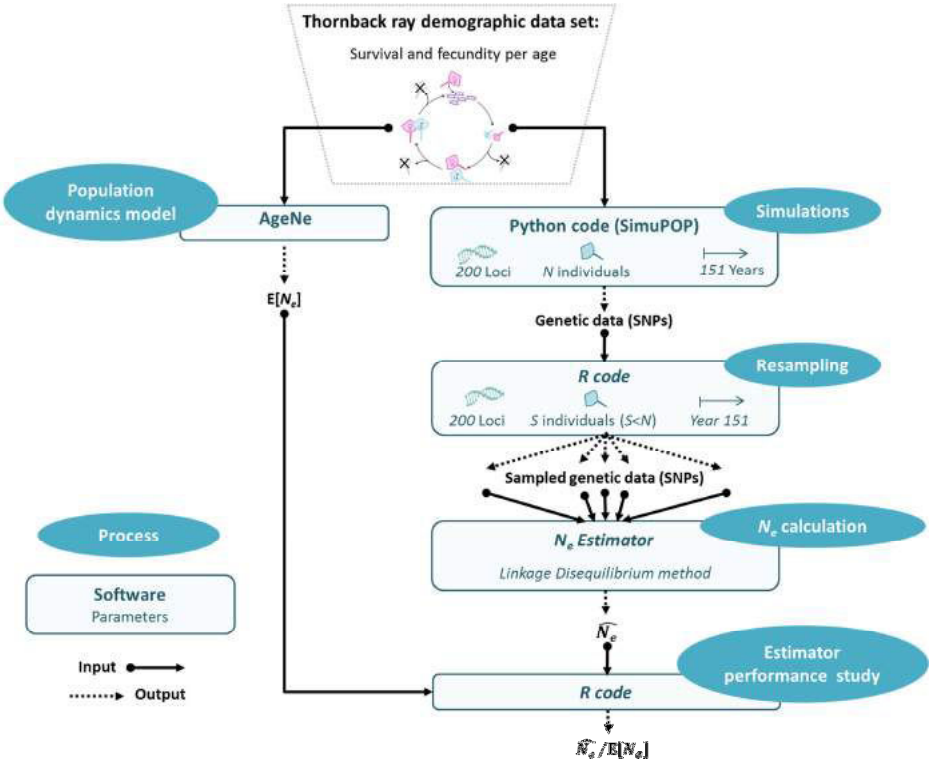


Figure S2: Flowchart for simulation study. References for software used: AgeNe (Waples et al. 2011), NeEstimator (V2, Do et al. 2014), R (R Foundation for Statistical Computing 2008) and simuPOP (Peng and Kimmel 2005).

Table S1: Sources for available census population size (*N*) of species with estimated empiric N_e in the review study

Species	Location or ICES zone if available	Year	*N*	Included ages	Source	Remarks
Albacore tuna	North Atlantic	2013	25 000 000	1-8+	ICCAT 2013	
Anchovy	Bay of Biscay	2017	5 466 000 000	All	ICES (2017a)	
Cod	North Sea, Eastern English Channel and Skagerrak	2017	764 631 000	1-6+	ICES (2017b)	
Cod	Barents and Norwegian seas	2016	1 296 598 000	3-12+	ICES (2017c)	
Darkblotched rockfish	Northeastern Pacific	2006	24 376 210	Not specified	Gomez-Uchida et al. (2006)	
European lobster		2011	4 299 999	Not specified	Butler et al. 2011 (IUCN)	Obtained by multiplying FAO landings on IUCN web site with a mean mass of 1kg
Hake	Southern Stock	2015	20 000 000	All	ICES (2017d)	
Plaice	North Sea	2016	5 238 382 000	All	ICES (2017b)	
Sardine	Bay of Biscay	2017	13 000 000 000	All	ICES (2017a)	
Thornback ray	Irish Sea	2008	3 000 000	All	Jim Ellis unpublished results cited in Chevolot et al. (2008)	

References

Butler, M., Cockcroft, A., MacDiarmid, A. & Wahle, R. 2011. Homarus gammarus. The IUCN Red List of Threatened Species 2011: e.T169955A69905303. http://dx.doi.org/10.2305/IUCN.UK.2011-1.RLTS.T169955A69905303.en. Downloaded on 20 March 2018.

Chevolot, M., Ellis, J.R., Rijnsdorp, A.D., Stam, W.T. & Olsen, J.L. (2008) Temporal changes in allele frequencies but stable genetic diversity over the past 40 years in the Irish Sea population of thornback ray, *Raja clavata* . *Heredity*, **101**, 120–126.

Gomez-Uchida, D. & Banks, M.A. 2006 Estimation of Effective Population Size for the Long-Lived Darkblotched Rockfish Sebastes crameri. *Journal of Heredity*, **97**, 603–606.

ICCAT. 2013. Report of the 2013 ICCAT North and South Atlantic albacore stock assessment meeting

ICES. 2017a. Report of the Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA), 24–29 June 2017, Bilbao, Spain. ICES CM 017/ACOM:17. 640 pp

ICES. 2017b. Report of the Working Group on Assessment of Demersal Stocks in the North Sea and Skagerrak (2017), 26 April–5 May 2017, ICES HQ. ICES M 2017/ACOM:21. 1234pp.

ICES. 2017c. Report of the Arctic Fisheries Working Group (AFWG), 19–25 April 2017, Copenhagen, Denmark. ICES CM 2017/ACOM:06. 493pp.

ICES. 2017d. Report of the Benchmark Workshop on Widely Distributed Stocks (WKWIDE), 4-11 May 2017, Copenhagen, Denmark. ICES CM 017/ACOM:36. 534 pp.

Table S2: Reviewed studies with N_e the estimated effective population size; S, the sample size and L, the number of genetic markers used.

Species	Location	N_e	S	L	Genetic marker type	Method	Sources
Albacore tuna	North Atlantic	14 398	1 131	58	SNP	Temporal Method (With correction)	Laconcha et al. 2015
Argentinian silverside	Central and southern Argentina	14	18	7	Microsattellite	Bayesian approximation	Valencia et al. 2017
Argentinian silverside	Central and southern Argentina	18	22	7	Microsattellite	Bayesian approximation	Valencia et al. 2017
Argentinian silverside	Central and southern Argentina	14	19	7	Microsattellite	Bayesian approximation	Valencia et al. 2017
Argentinian silverside	Central and southern Argentina	110	23	7	Microsattellite	Bayesian approximation	Valencia et al. 2017
Atlantic cod	Barents Sea	183.4	80	15	Microsattellite	Linkage Disequilibrium	Zhivotovsky et al. 2016
Atlantic cod	Barents Sea	INF	80	15	Microsattellite	Molecular coancestry	Zhivotovsky et al. 2016
Atlantic cod	Barents Sea	7.4	80	15	Microsattellite	Heterozygosity excess	Zhivotovsky et al. 2016
Atlantic cod	Barents Sea	INF	80	15	Microsattellite	Linkage Disequilibrium	Zhivotovsky et al. 2016
Atlantic cod	Barents Sea	INF	80	15	Microsattellite	Molecular coancestry	Zhivotovsky et al. 2016
Atlantic cod	Barents Sea	INF	80	15	Microsattellite	Heterozygosity excess	Zhivotovsky et al. 2016
Darkblotched rockfish	Northeastern Pacific	9 157	1 087	7	Microsattellite	Temporal Method (With correction)	Gomez-Uchida et al. 2006
Dusky grouper	Malta's Fisheries Management Zone	131.5	250	14	Microsattellite	Linkage Disequilibrium	Buchholz-Sorensen et al. 2016
Dusky kob	South Africa	999.7	230	18	Microsattellite	Linkage Disequilibrium	Mirimin et al. 2015
European anchovy	Bay of Biscay	6 342	330	349	SNP	Temporal Method (With correction)	Montes et al. 2016
European hake	Southern Stock	291	1 833	5	Microsattellite	Linkage Disequilibrium	Pita et al. 2017
European hake	Southern Stock	78	1 833	5	Microsattellite	Temporal Method (Without correction)	Pita et al. 2017
European hake	Southern Stock	226	1 833	5	Microsattellite	Pseudo-likelihood method	Pita et al. 2017
European lobster	British island	-272.7	36	12	Microsattellite	Linkage Disequilibrium	Victoria-Watson et al. 2016
European lobster	British island	677.7	29	12	Microsattellite	Linkage Disequilibrium	Victoria-Watson et al. 2016
European lobster	British island	-768.4	48	12	Microsattellite	Linkage Disequilibrium	Victoria-Watson et al. 2016
European lobster	British island	2 407.4	40	12	Microsattellite	Linkage Disequilibrium	Victoria-Watson et al. 2016

European lobster	British island	-1 234.9	44	12	Microsattellite	Linkage Disequilibrium		Victoria-Watson et al. 2016
European lobster	British island	-924.1	48	12	Microsattellite	Linkage Disequilibrium		Victoria-Watson et al. 2016
European lobster	British island	1 252.1	44	12	Microsattellite	Linkage Disequilibrium		Victoria-Watson et al. 2016
European lobster	British island	-406.8	48	12	Microsattellite	Linkage Disequilibrium		Victoria-Watson et al. 2016
European lobster	British island	-916	48	12	Microsattellite	Linkage Disequilibrium		Victoria-Watson et al. 2016
European sardine	8abd	60	2 571	27	No data	Temporal Method (Without correction)		Laurent et Planes 2007
European sardine	8abd	222	2 571	27	No data	Pseudo-likelihood method		Laurent et Planes 2007
Galapagos shark	Galápagos Marine Reserve	205	54	7934	SNP	Linkage Disequilibrium		Pazmiño et al. 2017
Gray whale	Sakhalin Island (Russia)	14.1	29	96	SNP	Linkage Disequilibrium		DeWoody et al. 2017
Gray whale	Sakhalin Island (Russia)	14.4	29	96	SNP	Heterozygosity excess		DeWoody et al. 2017
Kingklip	South Africa (CB West)	703	100	9	Microsattellite	No data		Henriques et al. 2017
Kingklip	South Africa (EC West)	1 023	100	9	Microsattellite	No data		Henriques et al. 2017
New Zealand snapper	New Zealand	176	50	7	Microsattellite	Temporal Method (Without correction)		Hauser et al. 2002
New Zealand snapper	New Zealand	46	50	7	Microsattellite	Decrease in heterozygosity		Hauser et al. 2002
North Sea cod	North and Baltic Seas	1 068	121	9	Microsattellite	Temporal Method (Without correction)		Poulsen et al. 2005
North Sea cod	North and Baltic Seas	2 067	121	9	Microsattellite	Pseudo-likelihood method		Poulsen et al. 2005
North Sea cod	North and Baltic Seas	491 497	121	9	Microsattellite	Bayesian approximation		Poulsen et al. 2005
North Sea cod	North and Baltic Seas	739	121	9	Microsattellite	Bayesian approximation		Poulsen et al. 2005
North Sea cod	North and Baltic Seas	121	72	3	Microsattellite	Temporal Method (With correction)		Hutchinson et Al. 2005
Oblique-banded snapper	Indian and Pacific Oceans	12.3	292	6	Microsattellite	Linkage Disequilibrium		Kennington et al. 2017
Oblique-banded snapper	Indian and Pacific Oceans	301.9	292	6	Microsattellite	Linkage Disequilibrium		Kennington et al. 2017
Oblique-banded snapper	Indian and Pacific Oceans	INF	292	6	Microsattellite	Linkage Disequilibrium		Kennington et al. 2017
Oblique-banded	Indian and Pacific	0	292	6	Microsattellite	Linkage Disequilibrium		Kennington et al. 2017

snapper	Oceans						
Plaice	North Sea	19 535	828	6	Microsattellite	Pseudo-likelihood method	Hoarau et al. 2005
Plaice	Iceland	1 733	828	6	Microsattellite	Pseudo-likelihood method	Hoarau et al. 2005
Red drum	Gulf of Mexico	3 516	967	8	Microsattellite	Temporal Method (With correction)	Turner et al. 2002
Red drum	Southeast United States	560	612	5	Microsattellite	Temporal Method (With correction)	Chapman et al. 2002
Spanish mackerel	Darwin (Australia)	-40 163	5 413	7	Microsattellite	Linkage Disequilibrium	Macbeth et al. 2013
Stripped bass	South Carolina (United States)	268	65	3	Microsattellite	Temporal Method (With correction)	Diaz et al. 2000
Thornback ray	Irish Sea and Bristol Channel	512	363	5	Microsattellite	Pseudo-likelihood method	Chevolot et al.2008
Tiger prawn	Australia	866.7	701	8	Microsattellite	Temporal Method (Without correction)	Ovenden et al. 2006
Tiger prawn	Australia	1 087.3	701	8	Microsattellite	Pseudo-likelihood method	Ovenden et al. 2006
White shark	South Africa	333	233	14	Microsattellite	Linkage Disequilibrium	Andreotti et al. 2016
Zebra shark	Queensland (Australia)	377	114	13	Microsattellite	Linkage Disequilibrium	Dudgeon and Ovenden 2015

References

Andreotti, S., Rutzen, M., van der Walt, S., Von der Heyden, S., Henriques, R., Meÿer, M., Oosthuizen, H. & Matthee, C. (2016) An integrated mark-recapture and genetic approach to estimate the population size of white sharks in South Africa. *Marine Ecology Progress Series*, **552**, 241–253.

Buchholz-Sørensen, M. & Vella, A. (2016) Population Structure, Genetic Diversity, Effective Population Size, Demographic History and Regional Connectivity Patterns of the Endangered Dusky Grouper, Epinephelus marginatus (Teleostei: Serranidae), within Malta's Fisheries Management Zone ed T.-Y. Chen. *PLOS ONE*, **11**, e0159864.

Chapman, R.W., Ball, A.O. & Mash, L.R. (2002) Spatial Homogeneity and Temporal Heterogeneity of Red Drum (Sciaenops ocellatus) Microsatellites: Effective Population Sizes and Management Implications. *Marine Biotechnology*, **4**, 589–603.

Chevolot, M., Ellis, J.R., Rijnsdorp, A.D., Stam, W.T. & Olsen, J.L. (2008) Temporal changes in allele frequencies but stable genetic diversity over the past 40 years in the Irish Sea population of thornback ray, *Raja clavata* . *Heredity*, **101**, 120–126.

DeWoody, J.A., Fernandez, N.B., Brüniche-Olsen, A., Antonides, J.D., Doyle, J.M., San Miguel, P., Westerman, R., Vertyankin, V.V., Godard-Codding, C.A.J. & Bickham, J.W. (2017) Characterization of the Gray Whale *Eschrichtius robustus* Genome and a Genotyping Array Based on Single-Nucleotide Polymorphisms in Candidate Genes. *The Biological Bulletin*, **232**, 186–197.

Diaz, M., Wethey, D., Bulak, J. & Ely, B. (2000) Effect of Harvest and Effective Population Size on Genetic Diversity in a Striped Bass Population. *Transactions of the American Fisheries Society*, **129**, 1367–1372.

Dudgeon, C.L. & Ovenden, J.R. (2015) The relationship between abundance and genetic effective population size in elasmobranchs: an example from the globally threatened zebra shark *Stegostoma fasciatum* within its protected range. *Conservation Genetics*, **16**, 1443–1454.

Gomez-Uchida, D. & Banks, M.A. (2006) Estimation of Effective Population Size for the Long-Lived Darkblotched Rockfish Sebastes crameri. *Journal of Heredity*, **97**, 603–606.

Hauser, L., Adcock, G.J., Smith, P.J., Bernal Ramirez, J.H. & Carvalho, G.R. (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (Pagrus auratus). *Proceedings of the National Academy of Sciences*, **99**, 11742–11747.

Hoarau, G., Boon, E., Jongma, D.N., Ferber, S., Palsson, J., Van der Veer, H.W., Rijnsdorp, A.D., Stam, W.T. & Olsen, J.L. (2005) Low effective population size and evidence for inbreeding in an overexploited flatfish, plaice (*Pleuronectes platessa L.*). *Proceedings of the Royal Society B: Biological Sciences*, **272**, 497–503.

Hutchinson, W.F., Oosterhout, C. v., Rogers, S.I. & Carvalho, G.R. (2003) Temporal analysis of archived samples indicates marked genetic changes in declining North Sea cod (Gadus morhua). *Proceedings of the Royal Society B: Biological Sciences*, **270**, 2125–2132.

Laconcha, U., Iriondo, M., Arrizabalaga, H., Manzano, C., Markaide, P., Montes, I., Zarraonaindia, I., Velado, I., Bilbao, E., Goñi, N., Santiago, J., Domingo, A., Karakulak, S., Oray, I. & Estonba, A. (2015) New Nuclear SNP Markers Unravel the Genetic Structure and Effective Population Size of Albacore Tuna (*Thunnus alalunga*) ed R. Cimmaruta. *PLOS ONE*, **10**, e0128247.

Laurent, V. & Planes, S. (2007) Effective population size estimation on *Sardina pilchardus* in the Bay of Biscay using a temporal genetic approach: effective population sizes of sardines. *Biological Journal of the Linnean Society*, **90**, 591–602.

Mirimin, L., Macey, B., Kerwath, S., Lamberth, S., Bester-van der Merwe, A., Cowley, P., Bloomer, P. & Roodt-Wilding, R. (2016) Genetic analyses reveal declining trends and low effective population size in an overfished South African sciaenid species, the dusky kob (Argyrosomus japonicus). *Marine and Freshwater Research*, **67**, 266.

Montes, I., Iriondo, M., Manzano, C., Santos, M., Conklin, D., Carvalho, G.R., Irigoien, X. & Estonba, A. (2016) No loss of genetic diversity in the exploited and recently collapsed population of Bay of Biscay anchovy (*Engraulis encrasicolus, L.*). *Marine Biology*, **163**.

Ovenden, J.R., Peel, D., Street, R., Courtney, A.J., Hoyle, S.D., Peel, S.L. & Podlich, H. (2006) The genetic effective and adult census size of an Australian population of tiger prawns (Penaeus esculentus): DRAUN EFFECTIVE POPULATION SIZE. *Molecular Ecology*, **16**, 127–138.

Pazmiño, D.A., Maes, G.E., Simpfendorfer, C.A., Salinas-de-León, P. & van Herwerden, L. (2017) Genome-wide SNPs reveal low effective population size within confined management units of the highly vagile Galapagos shark (Carcharhinus galapagensis). *Conservation Genetics*, **18**, 1151–1163.

Pita, A., Pérez, M., Velasco, F. & Presa, P. (2017) Trends of the genetic effective population size in the Southern stock of the European hake. *Fisheries Research*, **191**, 108–119.

Poulsen, N.A., Nielsen, E.E., Schierup, M.H., Loeschcke, V. & GrøNkjaer, P. (2005) Long-term stability and effective population size in North Sea and Baltic Sea cod (*Gadus morhua*): effective population size in atlantic cod. *Molecular Ecology*, **15**, 321–331.

Turner, T.F., Wares, J.P. & Gold, J.R. (2002) Genetic effective size is three orders of magnitude samaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics*, **162**, 1329–1339.

Valencia, E., Veliz, D., Tombari, A. & Vega, C. (2017) Genetic population structure and evidence of genetic homogeneity in populations of the Argentinian silverside Odontesthes bonariensis (Teleostei: Atherinopsidae) inhabiting central and northwestern Argentina. *Latin American Journal of Aquatic Research*, **45**, 708–716.

Watson, H.V., McKeown, N.J., Coscia, I., Wootton, E. & Ironside, J.E. (2016) Population genetic structure of the European lobster (*Homarus gammarus*) in the Irish Sea and implications for the effectiveness of the first British marine protected area. *Fisheries Research*, **183**, 287–293.

Zhivotovsky, L.A., Teterina, A.A., Mukhina, N.V., Stroganov, A.N., Rubtsova, G.A. & Afanasiev, K.I. (2016) Effects of genetic drift in a small population of Atlantic cod (*Gadus morhua* kildinensis Derjugin) landlocked in a meromictic lake: genetic variation and conservation measures. *Conservation Genetics*, **17**, 229–238.

CHAPITRE 7

Conclusions et perspectives

Ghibli - Ponyo sur la falaise

L'objectif principal de cette thèse était d'évaluer l'état des populations de raie bouclée pour leur gestion à l'aide des méthodes génétiques et démographiques disponibles. Cet objectif a été divisé en trois grandes parties :

1. l'évaluation des méthodes disponibles ;
2. l'évaluation de l'état des populations de raie bouclée ;
3. la gestion de la raie bouclée.

Les travaux de cette thèse offrent des réponses aux trois questions précédentes.

Dans les sections ci-dessous, une distinction est faite entre l'évaluation de l'état d'une population et sa gestion. L'état d'une population correspond à son statut lors de l'évaluation : bon état, tendance descendante, en danger, sont par exemple des états possibles. La gestion d'une population s'intéresse aux mesures quantitatives à mettre en oeuvre (type quota de pêche) afin de maintenir ou ramener une population à un état d'exploitation soutenable, par exemple à l'aide d'un quota de pêche.

1. Evaluation de l'état des populations d'élasmobranches : est-ce possible actuellement ?

1.1 Une première étape indispensable : la délimitation des populations

La première étape pour une gestion soutenable des populations animales consiste en l'identification de ces populations et leur délimitation. Différentes méthodes sont disponibles et cette thèse s'est intéressée à l'utilisation des connectivités génétique et démographique afin de délimiter les populations de raies (chapitre 2).

L'évaluation de la connectivité génétique s'effectue via la comparaison de marqueurs génétiques issus d'individus de plusieurs populations putatives. Cette méthode nécessite peu d'individus échantillonnés par population putative (une trentaine suffisent en général) et permet d'obtenir facilement des valeurs de différenciation génétique. Cependant ces dernières sont difficiles à interpréter pour plusieurs raisons. Tout d'abord, il n'existe pas de valeur seuil reconnue permettant de décider quand regrouper ou quand séparer les populations putatives. Ensuite, il existe une confusion entre l'effet de la taille de la population putative étudiée et le taux de dispersion qu'elle expérimente. Autrement dit, une même valeur de différenciation génétique peut s'expliquer par plusieurs combinaisons tailles de population/flux de gène différentes (chapitre 2).

Pour illustrer cette effet, deux populations de raie bouclée expérimentant six taux de dispersion différents ont été simulées pendant 1 000 ans. La taille d'une des populations a été fixée à 10 000 individus tandis que l'autre a varié entre 10 000 et 800 000 individus (voir la Fig. 7.1 pour la ratio entre les tailles de populations simulées). Sur cette figure, une valeur de différenciation génétique d'environ 0,01 peut correspondre à un taux de dispersion faible (0,001) appliquée à une population de grande taille (50 000) et de faible taille (10 000) mais elle peut aussi correspondre à un taux de dispersion fort (0,01) appliquée à des populations de petites tailles (10 000 pour les deux).

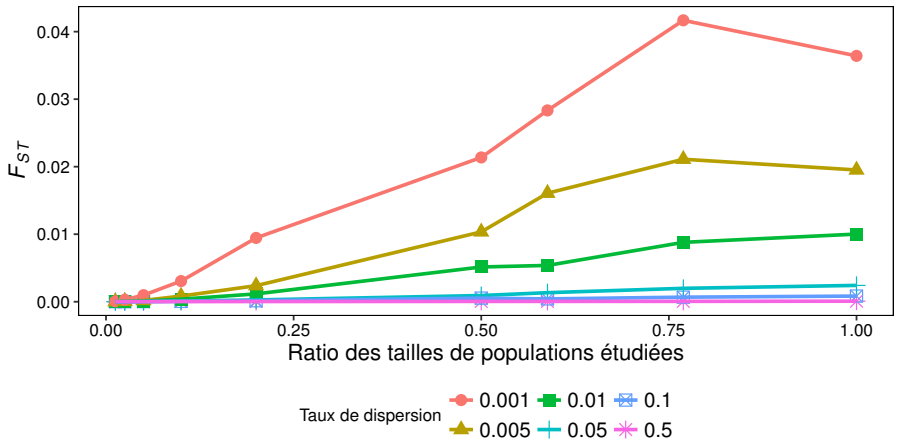


FIGURE 7.1 : Différenciation génétique observée (F_{ST}) entre deux populations de raie bouclée après 1 000 ans en fonction du taux de disperion et du ratio des tailles de populations simulées

Au delà de la difficulté d'interprétation, le signal génétique issu de la connectivité des populations est long à s'inscrire dans le génome et donc à être détectable. L'échelle des temps étudiée correspond à l'échelle de l'évolution, les valeurs de différenciation estimées au temps t correspondent souvent à la différenciation observée au temps $t - x$ années (x étant inconnu). La connectivité génétique renseigne donc sur des unités souvent passées mais aussi à une très large échelle géographique, elle trouve ainsi son utilité dans la conservation des espèces plus que dans leur exploitation. Ainsi si l'on se réfère à l'évaluation de connectivité génétique effectué par Chevolot et al. (2006), l'ensemble des populations de raie bouclée d'Atlantique Nord-Est (hormis les Açores) sont génétiquement connectées et devraient potentiellement être gérées ensemble. Encore une fois ces résultats sont à moduler en fonction du seuil de séparation/regroupement des populations choisi ainsi qu'en fonction du délai d'apparition du signal de connectivité dans le génome.

L'évaluation de la connectivité démographique s'effectue grâce à l'utilisation du taux de dispersion entre deux populations putatives et à leurs abondances relatives présumées. Ces deux variables permettent le calcul du nombre d'immigrants et leur participation à la démographie de la population puit. Cette dernière valeur permet de quantifier l'importance de l'immigration pour une population. Dans le cas de la raie bouclée, la population présumée de la Manche Est présente la plus grand abondance, la plaçant comme une source potentiellement indispensable à la survie de ses voisines et au contraire, tout à fait indépendante, démographiquement parlant, de ses voisines. L'échelle de temps évalué est ici contemporaine et les populations sont délimitée à une échelle plus fine et sans doute plus pertinente en gestion. Néanmoins, tout comme la connectivité génétique, la délimitation de population grâce à la connectivité démographique souffre du manque de valeur seuil reconnue permettant la séparation/regroupement des populations présumées étudiées.

La figure 7.2 regroupe sous la forme d'une matrice FFOM (Forces, Faiblesses, Opportunités et Menaces) les principaux coûts et bénéfices de la délimitation des populations via l'étude de la connnctivité des populations en comparant connectivité génétique et connectivité démographique. Elle reprend les principaux résultats du chapitre 2 ainsi que des aspects logistiques supplémentaires (coûts financiers, matériels, etc).

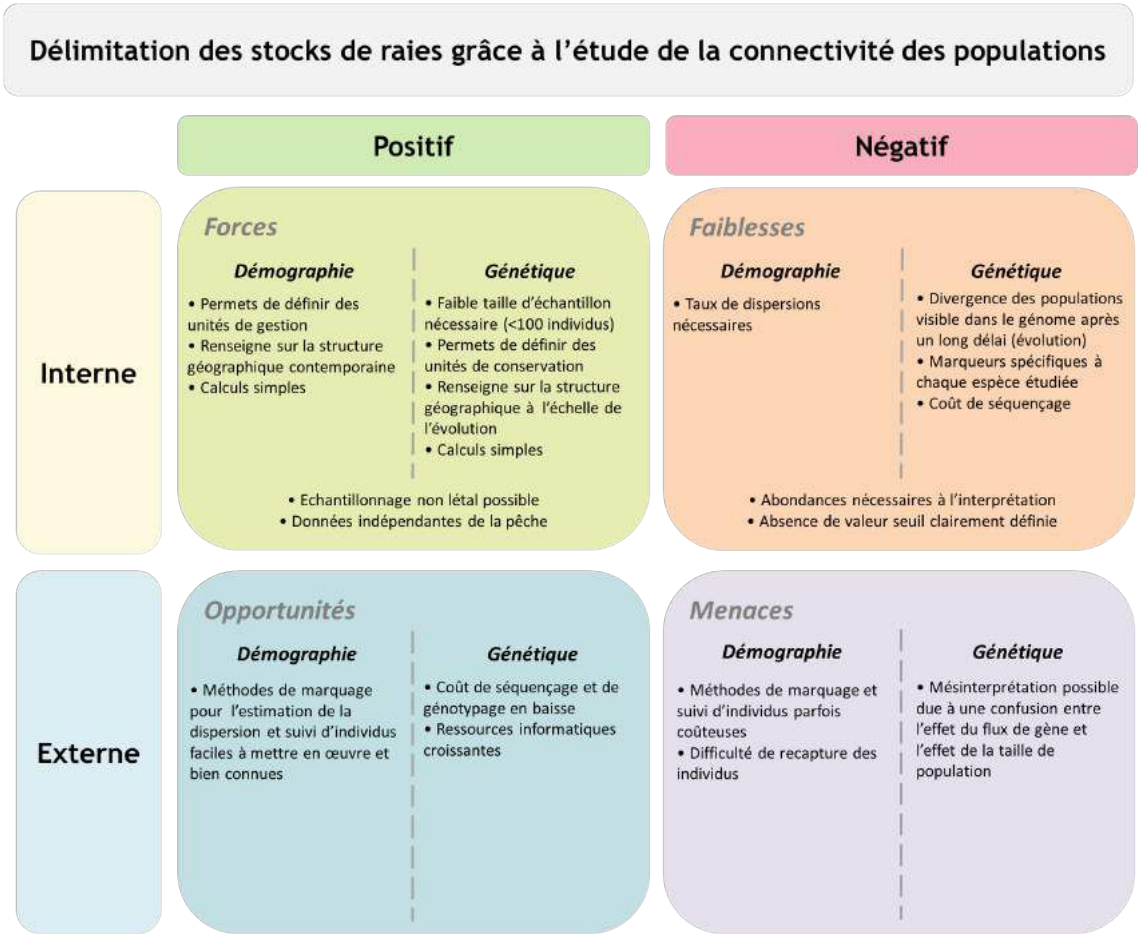


FIGURE 7.2 : Matrice FFOM (Forces, Faiblesses, Opportunités et Menaces) de la délimitation des populations des raies grâce à l'étude de la connectivité des populations

1.2 L'évaluation démographique de l'état d'une population : une tâche difficile mais plus impossible

L'évaluation démographique de l'état des populations d'élasmobranches a longtemps été considérée impossible à cause du manque de données biologiques mais surtout à cause du manque de données spécifiques issues de la pêche professionnelle (débarquements, rejets et effort et pêche) et issues des campagnes scientifiques. Aujourd'hui, la situation a changé. Après presque 10 ans d'obligation d'enregistrement des débarquements de manière spécifique, des séries de données exploitables issues de pêches professionnelles commencent à émerger. Ainsi certains modèles d'évaluation de stock halieutique deviennent utilisables sur presque toutes les espèces avec des degrés de précision des estimations très variables. Les stocks d'élasmobranches peuvent ainsi être séparés en 4 catégories selon les données disponibles pour évaluer leur état :

- **Données spécifiques stratifiées en âge** : par exemple pour l'Aiguillat commun (*Squalus acanthias*), les données disponibles sont stratifiées en âge permettant l'utilisation de modèles en âge/taille (De Oliveira et al., 2013) et donc l'estimation de son état mais aussi l'obtention d'estimations quantitatives.
- **Données specifiques non stratifiées en âge** : par exemple pour la raie bouclée, les

données commerciales et scientifiques disponibles sont spécifiques mais non stratifiées en âge. Des modèles quantitatifs monospécifiques peuvent donc être utilisés pour estimer son état et dans une moindre mesure pour obtenir des estimations quantitatives peu précises (chapitre 3 et 4).

– **Peu de données specifiques non stratifiées en âge** : par exemple pour la raie lisse (*Raja brachyura*), les données spécifiques restent rares. Elles sont néanmoins suffisantes pour être utilisées dans un modèle multispécifique quantitatif peu précis (chapitre 4) ou dans des approches de vulnérabilité (McCully Phillips et al., 2015).
– **Pas de données spécifiques disponibles** : pour certaines espèces, les données spécifiques sont absentes, ne permettant pas d'évaluer leur état. Néanmoins des approches génétiques semblent prometteuses pour ces espèces, notamment l'utilisation d'ADN environnemental (Weltz et al., 2017) qui permets d'estimer leur présence.

Grâce à la multiplication des données spécifiques, il est possible d'évaluer l'état des populations d'élasmobranches en adaptant les modèles halieutiques classiques à la quantité et à la qualité des données disponibles. Néanmoins, il est nécessaire de garder à l'esprit que la fiabilité et la précision des estimations sont conditionnelles à ces mêmes quantité et qualité. De plus, si ces estimations permettent d'améliorer les connaissances sur les différents stocks, elles ne permettent pas forcément de mettre en place une gestion appropriée (voir partie 2.2). La figure 7.3 regroupe les principaux coûts et bénéfices de l'évaluation démographique de l'état des stocks d'élasmobranches mis en évidence dans cette thèse.

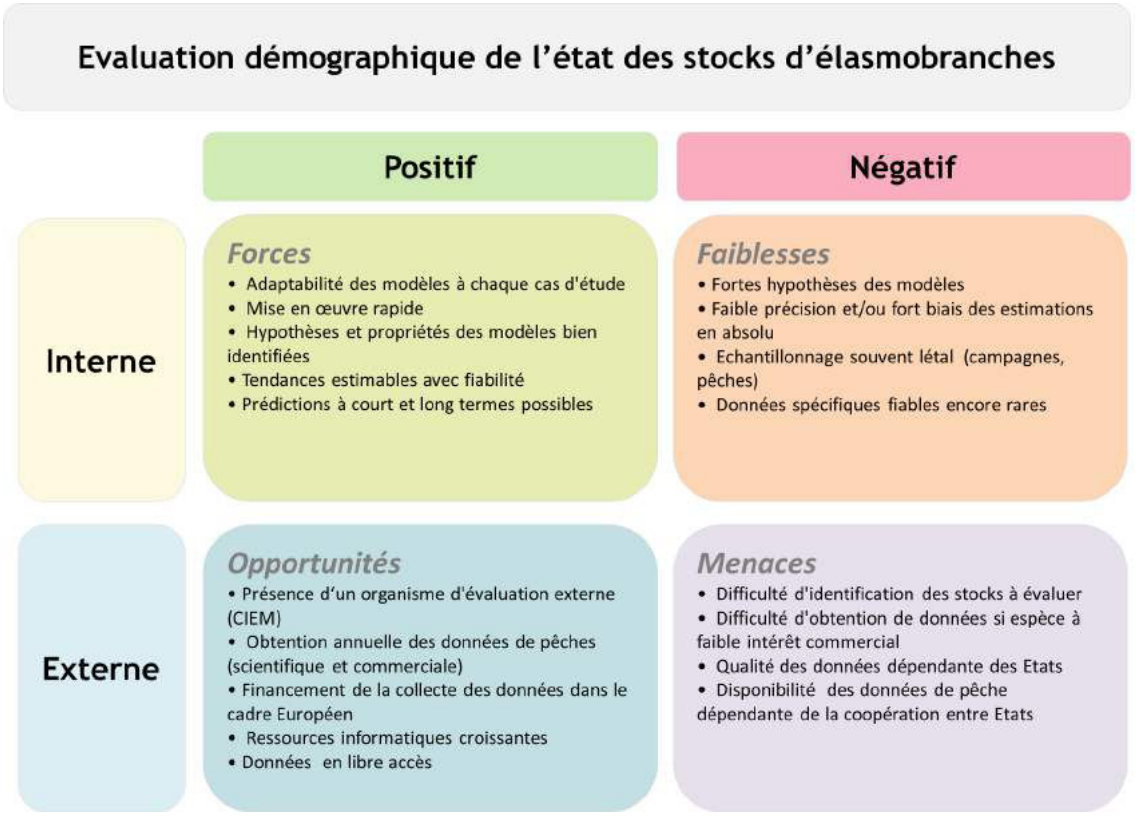


FIGURE 7.3 : Matrice FFOM (Forces, Faiblesses, Opportunités et Menaces) de l'évaluation démographique de l'état des stocks d'élasmobranches

1.3 L'évaluation génétique de l'état d'une population : un outil de conservation avant tout

Depuis une trentaine d'années, la génétique est devenue un outil répandu en écologie et gestion des populations naturelles, avec une utilisation accrue ces 15 dernières années. L'application d'outils génétiques sur les populations d'élasmobranches permet l'étude de leur diversité génétique, leur structure de populations mais aussi leur phylogénie (Dudgeon et al., 2012). Ces utilisations sont supportées par la possibilité qu'offre la génétique d'effectuer des échantillonnages non létaux mais aussi par le besoin d'échantillons de tailles souvent inférieures à celles des modèles démographiques classiques.

L'émergence de nouvelles technologies génétiques et la baisse des coûts de génotypage a également permis l'émergence de modèles et estimateurs permettant d'estimer les tailles des populations naturelles (absolue et efficace). Ces modèles et estimateurs sont appliqués depuis des décennies sur les populations terrestres (Schwartz et al., 2007) mais émergent seulement depuis une quinzaine d'années pour les populations marines. Ils présentent l'avantage d'utiliser des données indépendantes de la pêche mais reposent sur de fortes hypothèses de modélisations (voire section 1, du chapitre 5). Par exemple, la majorité des estimateurs de taille de population efficace ne prend pas en considération les générations chevauchantes, pourtant inhérentes à presque toutes les populations naturelles. La violation de ces conditions entraine de forts biais (chapitre 5) qui, couplé à la faible précision des estimations génétiques, rendent ces méthodes bien souvent inutilisables en gestion.

Dans cette thèse, les deux méthodes les plus populaires d'estimation de la taille de population efficace ont été discutées par le biais de simulations (chapitre 5). La première méthode, Linkage Disequilibrium, ne nécessite qu'un seul événement d'échantillonnage mais présente une forte sensibilité aux traits d'histoires de vie de l'espèce étudiée. Pour la raie bouclée, les estimations de N_e obtenues par cette méthode présente une sous-estimation constante de 30% laissant penser qu'une correction soit possible. Néanmoins l'intensité de cette correction est dépendante des traits d'histoire de vie et son estimation nécessite un travail de simulation conséquent. La seconde méthode, Temporelle corrigée pour prendre en compte les générations chevauchantes, est basée sur la comparaison d'échantillons séparés dans le temps. Elle performe mieux que la première méthode avec des biais réduits lorsque le délai entre les deux évènements d'échantillonnage dépassent sept annnées. Néanmoins, sa mise en oeuvre nécessite un échantillonnage précis : soit deux séries d'échantillonnage séparées dans le temps soit un seul événement d'échantillonnage et la comparaison d'individus d'âge différents. De plus, la correction spécifique nécessite de connaître précisément et quantitativement les traits d'histoire de vie de l'espèce étudiée. Ces deux conditions limitent ainsi son usage.

Au delà des propriétés et hypothèses des estimateurs génétiques, l'utilisation de la génétique pour l'évaluation de l'état des populations marines et leur gestion est limitée par les tailles d'échantillons nécessaires à sa mise en oeuvre. Par exemple, pour estimer la population efficace d'une espèce de poisson, une taille d'échantillon correspondant à 1% de sa population totale semble être nécessaire (chapitre 6), ce qui se traduit par plusieurs milliers d'individus dans le cas des populations marines. De telles tailles d'échantillons sont, dans un premier lieu, difficiles et coûteuses à obtenir et dans un second temps, coûteuses à séquencer et génotyper.

Dans l'hypothèse où de telles tailles d'échantillons puissent être échantillonné, un autre problème se pose : la grande taille des populations halieutiques (chapitre 6). L'ensemble des estimateurs de la taille de population efficace repose sur la détection d'un signal génétique dont

la variation reflèterait la variation de la taille de population efficace. Cependant, si une partie de ce signal correspond bien à la taille de population efficace, une autre partie est issue de la stochasticité inhérente à l'évolution génétique d'une population et cette dernière partie peut être estimée à tord comme reflètant N_e. Ainsi pour des populations de très grandes tailles ($<10^6$ individus), le signal génétique inhérent à N_e est noyé dans le bruit et N_e n'est souvent plus estimable. Ce phénomène se traduit entre autres par des estimations négatives.

L'utilisation des outils génétiques pour estimer l'abondance efficace d'une population naturelle ne semble pas adaptée aux populations marines exploitées de part la grande taille des populations étudiées, les hypothèses des estimateurs mais aussi le biais spécifique à l'espèce étudiée et la faible précision des estimations. Néanmoins, les outils génétiques permettant d'estimer la diversité génétique d'une population et donc son potentiel adaptatif semble être très intéressants pour les espèces marines entrant dans la catégorie "Conservation". En effet, une population d'assez grande taille mais avec une faible taille de population efficace signifie que peu d'individus participent génétiquement à la prochaine génération. Autrement dit, cette population présente une faible diversité génétique et donc une perte de son potentiel adaptatif. Une telle population a plus de chance de s'éteindre qu'une population de taille moyenne avec une grande diversité génétique. Ce type de constat ne peut être fait avec l'étude de la démographie d'une population. La figure 7.4 résume sous la forme d'une matrice FFOM (Forces, Faiblesses, Opportunités et Menaces) les principaux coûts et bénéfices de l'évaluation génétique de l'état des stocks d'élasmobranches via l'estimation de la taille de population efficace.

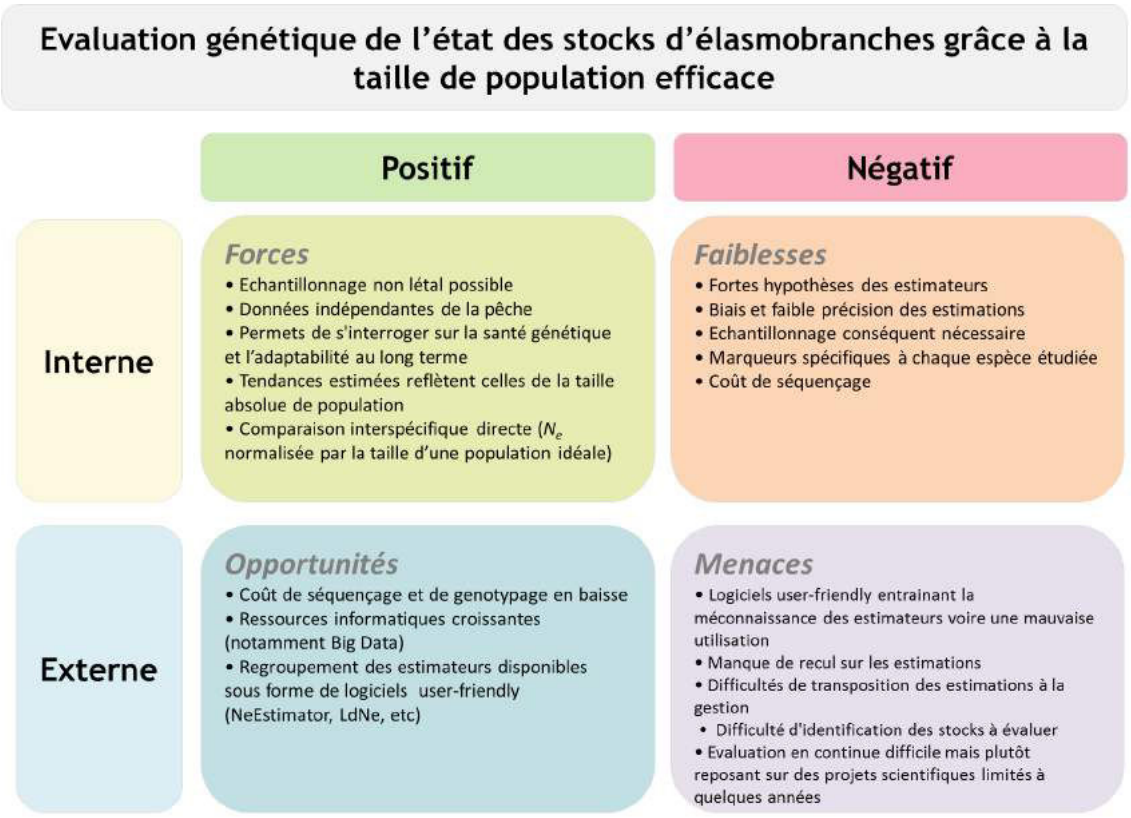


FIGURE 7.4 : Matrice FFOM (Forces, Faiblesses, Opportunités et Menaces) de l'évaluation géentique de l'état des stocks d'élasmobranches grâce à l'estimation de la taille de population efficace

1.4 Evaluation de l'état d'un stock de raie : bilan graphique

La figure 7.5 résume les différentes étapes nécessaires à l'étude de l'état d'un stock de raie. La première étape (engrenage 1) consiste en la délimitation des différents stocks d'une espèce, ceci pouvant être réalisé à l'aide de données génétiques et démographiques (chapitre 2). L'évaluation génétique de la connectivité renseigne à large échelle et sur une échelle temporelle écologique. L'évaluation démographique de la connectivité renseigne à une échelle fine et à une échelle temporelle plus immédiate.

La seconde étape consiste en l'évaluation de l'état du stock à proprement parler (engrenage 2). Là encore, l'évaluation peut se faire grâce aux deux outils, génétique et démographique. L'outil génétique apporte des informations sur la survie au long terme du stock et sa conservation, voir sur ses tendances en abondance (chapitres 5 et 6). L'outil démographique permet une évaluation à plus court terme via l'obtention de tendances en abondance voire des estimations en absolu si la qualité des données le permet (chapitre 3). Ces informations s'avèrent essentielles en gestion. Suite à ces évaluations, il est possible de mettre en place des mesures d'exploitation (engrenage 3A) ou de conservation (engrenage 3B) du stock étudié. Dans le cas des raies, l'exploitation est multispécifique et il est difficile voire impossible à l'heure actuelle de gérer en Atlantique Nord-Est une espèce de raie sans prendre les autres en considération (chaptitre 4).

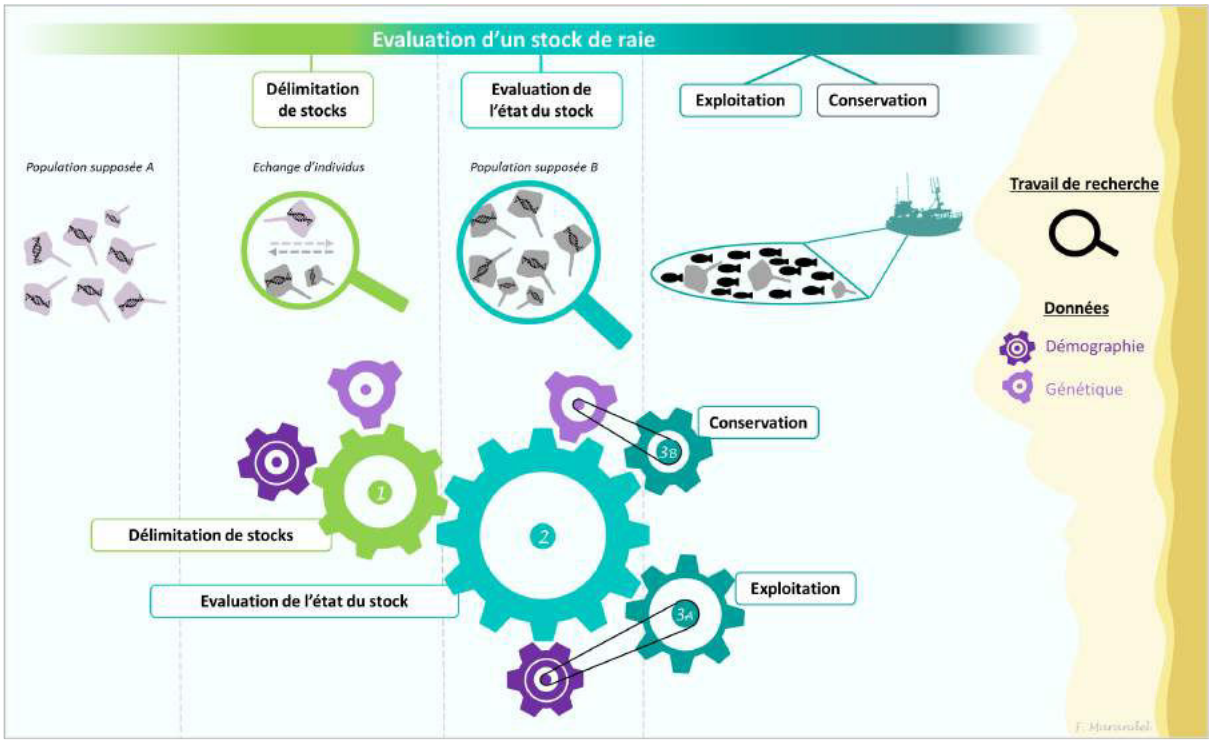


FIGURE 7.5 : Schématisation de l'évaluation de l'état d'un stock de raie et apport des différentes disciplines (démographie et génétique)

2. Bilan de l'état des populations de raie bouclée

2.1 Abondance de la raie bouclée dans le Golfe de Gascogne

Dans le chapitre 3, l'état de la raie bouclée du Golfe de Gascogne a été estimé à l'aide d'un modèle Bayésien monospécifique utilisant des débarquements et des indices de biomasses spécifiques (Fig. 7.6 a,b). Sous réserve que les données fournies au modèle soient fiables, le stock de raie bouclée du Golfe de Gascogne est estimé en mauvais état écologique autour de 3% de sa biomasse vierge, soit 2 046 tonnes en 2014 (Fig. 7.6 f), avec une stagnation autour de cette valeur depuis les années 90. La capacité biotique est estimée autour de 60 000 tonnes (K, Fig. 7.6 d) avec un taux de croissance intrinsèque de 0,105 (r, Fig. 7.6 c) et une capturabilité avoisinant 0,12 (q, Fig. 7.6 e).

Le modèle multispécifique utilisé dans le chapitre 4 estime les trajectoires de biomasse des six espèces de raies principales du Golfe de Gascogne. Il a été utilisé avec des débarquements (2009-2016) corrigés et retravaillés par le WGEF (groupe de travail européen travaillant sur les stocks d'élamsobranches) en 2016 et des indices de biomasse lissés (Fig. 7.6 a,b pour la raie bouclée). Les débarquements entre 1950 et 2009 sont estimés par le modèle. Le modèle multispécifique estime des biomasses de raie bouclée légèrement plus élevées que le modèle monospécifique mais un état totalement différent (Fig. 7.6 f). Le stock de raie bouclée du Golfe de Gascogne est estimé autour de 40% de sa biomasse vierge en 2017 (autour de 5 200 tonnes, Fig. 7.6 f) et donc, grossièrement, à sa biomasse au **Rendement Maximum Durable** (B_{RMD}).

Avec ces deux modèles, les biomasses récentes sont estimées faibles mais conjointement avec des capacités d'acceuil très différentes (K, Fig. 7.6 d) : presque 70 000 tonnes pour le monospécifique et seulement 19 000 tonnes pour le modèle multispécifique. De plus, le taux de croissance intrinsèque est estimé deux fois plus grand avec le modèle multispécifique (autour de 0,17, Fig. 7.6 c) et la capturabilité, deux fois plus faible (autour de 0,07, Fig. 7.6 e). Ces différences sont dues aux hypothèses des modèles mais surtout aux données fournies. Dans le modèle monospécifique, les débarquements spécifiques 2009-2014 ont été couplés à des débarquements reconstitués (1903-2008) afin d'améliorer la précision des estimations et ce, avec des effets minimes sur la moyenne des estimations (chapitre 3). Les débarquements spécifiques 2009-2016 utilisés dans le modèle multispécifique sont différents car retravaillés et corrigés par le WGEF en 2016 (Fig. 7.6 a). Pour les indices de biomasse, le choix a été fait dans le modèle multispécifique de lisser les indices disponibles alors qu'ils sont utilisés bruts dans le modèle monospécifiques (Fig. 7.6 b). Cette différence de lissage se retrouve dans les estimations des trajectoires de biomasses des deux modèles (Fig. 7.6 f).

Deux jeux de données, les plus vraisemblables et réalistes possibles, ont été utilisés à 2 années d'intervalles pour estimer l'abondance de la raie bouclée. Chacun mène a un état de la population très différent : bon état pour le modèle multispécifique et très mauvais état pour le modèle monospécifique. Dans le premier cas, la raie bouclée du Golfe de Gascogne est classée dans la catégorie "Gestion" alors que dans le second cas, elle est classée en "Conservation". Ces estimations très différentes renforcent le besoin de données spécifiques de bonne qualité et fiables quant à l'évaluation de l'état des populations de raie.

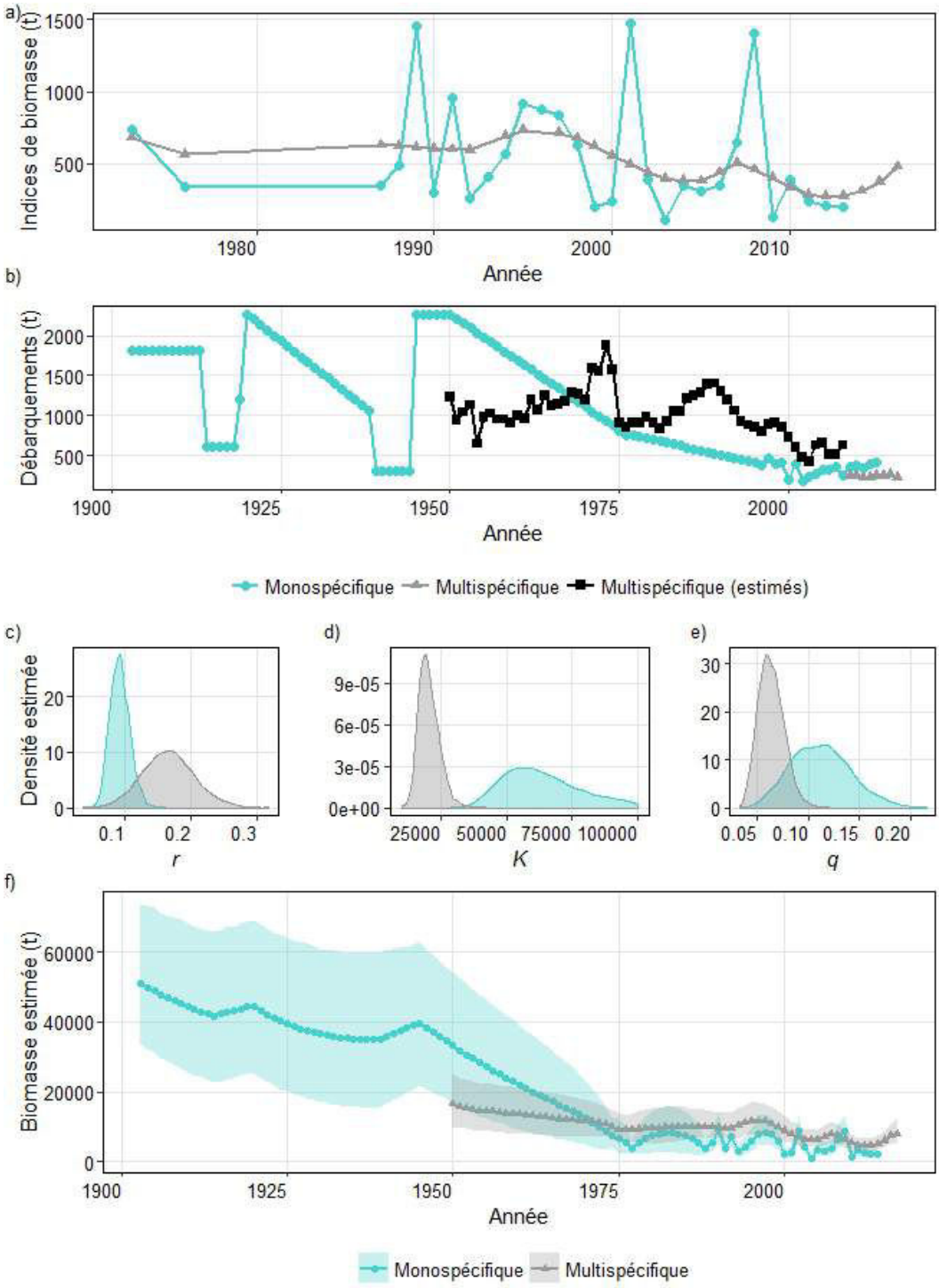


FIGURE 7.6 : Comparaison des données et estimations des deux modèles d'évaluation développés et utilisés sur la raie bouclée du Golfe de Gascogne. a) Indices de biomasse spécifiques utilisés ; b) Débarquements spécifiques utilisés et estimés par les différents modèles ; c) Distributions à posteriori du taux de croissance intrinsèque (r), d) de la capacité d'acceuil (K) et e) de la capturabilité (q) estimés, ; f) Trajectoires de biomasse estimées avec intervalles de confiance à 95%

Afin de valider l'hypothèse selon laquelle il est possible d'estimer l'état des populations de raie à l'aide des outils démographiques disponibles dès lors que l'on possède des données fiables, le modèle monospécifique a été utilisé avec les mêmes débarquements de 2009 à 2016 et les mêmes indices de biomasse que ceux utilisées dans le modèle multispécifique (Fig. 7.7 a,b).

Les distributions à posteriori estimées pour r et K sont identiques aux distributions obtenues précédemment (Fig. 7.6 et 7.7 c,d) mais la distribution à posteriori de q a été divisée par 6 (0,02 contre 0,12 précédemment, Fig. 7.6 et 7.7 e). Cette diminution de capturabilité se retrouve dans les estimations des trajectoires de biomasse estimées (non montrées). En effet, ces dernières sont nettement supérieures menant à un état relatif actuel estimé autour de 48% (Fig. 7.7 f) de K et donc similaire à celui estimé par le modèle multispécifique (40% de K, Fig. 7.7 f).

Bien que les deux modèles aient des hypothèses intrinsèques différentes, ils estiment tous les deux le même état pour la raie bouclée dans le Golfe de Gascogne. Ainsi la population de raie bouclée du Golfe de Gassogne semble être en bon état écologique, proche de sa biomasse au Rendement Maximum Durable. Cette espèce peut donc être classée dans la catégorie "Gestion" et non "Conservation". Néanmoins, les deux modèles ne permettent pas d'obtenir des estimations en valeur absolue pour la gestion à proprement parler. En effet, le modèle monospécifique estime des captures au RMD de 1 498 tonnes par an et le multispécifique, des captures au RMD de 828 tonnes par an. Cet écart très important ne permet donc pas l'utilisation de ces modèles en gestion. Néanmoins, les débarquements de raie bouclée de 2016 étant de 226 tonnes, ils sont probablement soutenables, cette interprétation étant confirmée par l'augmentation récente de l'indice de biomasse.

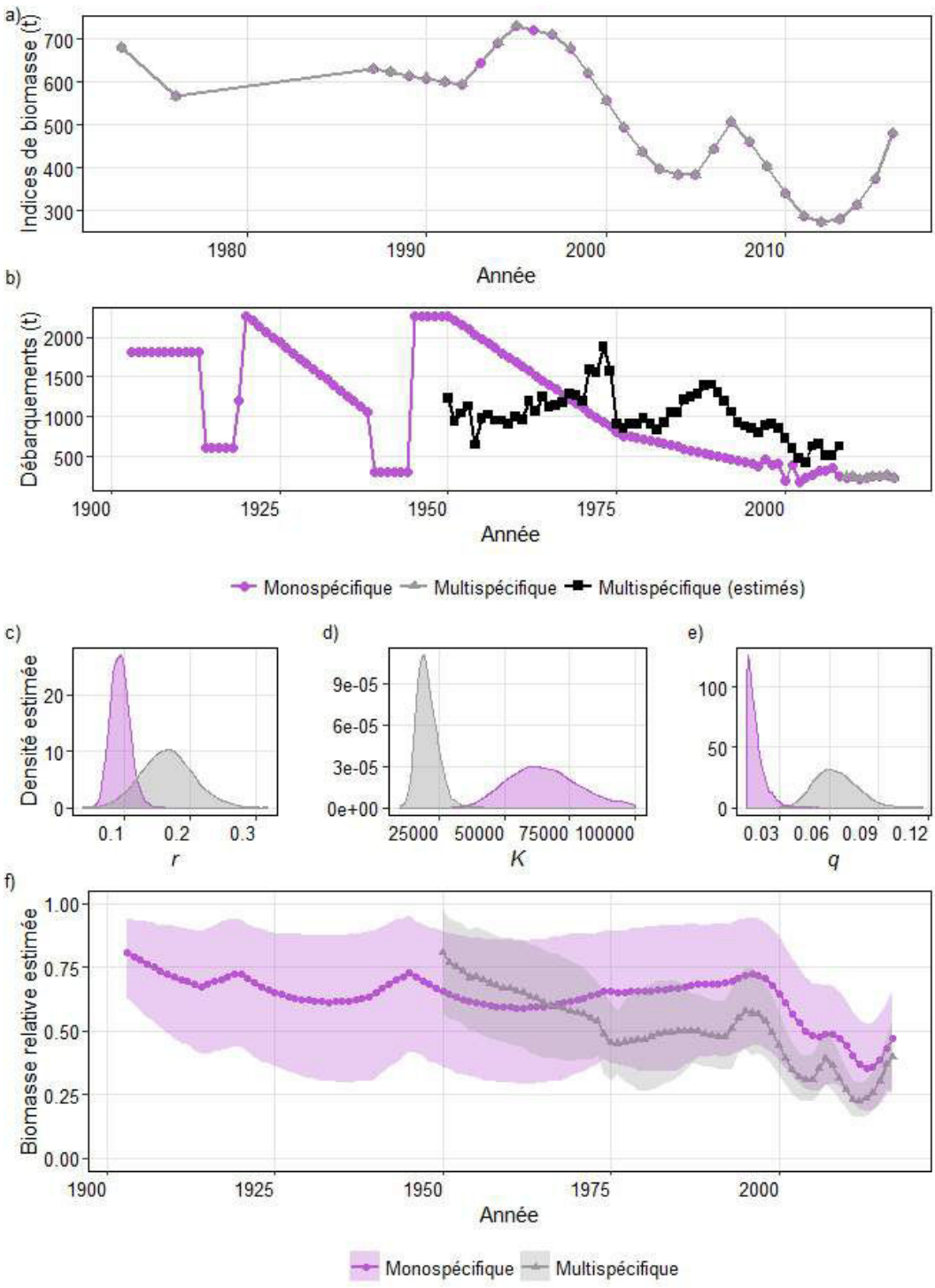


FIGURE 7.7 : Comparaison des estimations des deux modèles d'évaluation développés ajustés sur les mêmes données corrigés par le WGEF en 2016 de la raie bouclée du Golfe de Gascogne. a) Indices de biomasse utilisés ; b) Débarquements spécifiques utilisés et estimés par les différents modèles ; c) Densité de probabilité des taux de croissance intrinsèque (r), d) des capacités d'acceuil (K) et e) des capturabilités (q) estimés, ; f) Trajectoires de biomasse relative (B/K) estimées avec intervalles de confiance à 95%

2.2 Taille de population efficace de la raie bouclée dans le Golfe de Gascogne

Utilisant les données génétiques échantillonnées par le projet GenoPopTaille en 2015-2016, la taille de population efficace de la raie bouclée du Golfe de Gascogne a été estimée à l'aide de la méthode Linkage Disequilibrium. Les estimations varient entre 424,5 et 2 681 individus selon les filtres appliqués aux marqueurs génétiques (chapitre 5). L'étude de simulation nous a permis de mettre au point une correction spécifique aux traits d'histoire de vie de la raie bouclée et permettant de corriger la sous estimation de 30% systématiquement observée. En appliquant cette correction, les estimations varient entre 552 et 3 485 individus. Cette gamme de valeurs ne reflète probablement pas la réalité au vu des quantités de raie bouclée débarquées dans le Golfe de Gascogne (plusieurs centaines de tonnes).

L'interprétation de ces valeurs pour la gestion est encore aujourd'hui compliquée. Franklin (1980) propose la règle des 50/200 (très critiquée, voir Lande, 1995) préconisant que 50/500 individus efficaces sont un minimum pour éviter l'extinction à court terme/long terme. Si l'on applique ces règles à la raie bouclée, aucun risque d'extinction à court terme n'émerge mais la variabilité des estimations empêche de conclure à long terme. De plus, une estimation de la taille de population efficace ne permet pas d'inférer directement des mesures concrètes de gestion mais doit être reliée à la taille totale de la population (Ovenden et al., 2016). Ce lien peut être fait grâce au ratio N_e/N, estimable par génétique ou par démographie (AgeNe, Waples et al., 2011) avec encore une fois une faible précision. Dans le cas de la raie bouclée, le ratio N_e/N estimé par AgeNe (chapitres 5 et 6) est de 0,087. Appliqué aux N_e estimées et corrigées dans le chapitre 5, il conduit à des tailles de populations absolues de 6 343 et 40 060 individus, incompatibles avec tonnages pêchés chaque année. En appliquant un poids moyen de 1,42 kg (calculé par simulation de cohorte en utilisant les paramètres de croissance de Serra-Pereira et al., 2008), celà signifierait que la fraction participant génétiquement à la prochaine génération ne représenterait que 9 à 56 tonnes.

2.3 La gestion multispécifique des raies : une situation compliquée

L'évaluation des élasmobranches dans les eaux Européennes de l'Atlantique Nord-Est, mer du Nord et mer Baltique est effectuée par le CIEM (Conseil International pour l'Exploration de la Mer, ICES en anglais). Un groupe de travail leur est dédié : le WGEF (Working Group for Elasmobranch Fisheries). Il se réunit annuellement et rassemble les données disponibles pour ces espèces afin de décrire leur état dans les différentes zones CIEM et, quand c'est possible, émettre un avis quant à leur gestion.

Pour pouvoir gérer les stocks dits à données limitées (incluant des stocks de téléostéens), le CIEM a mis en place en 2012 une clasification en six catégories selon les estimations disponibles après évaluation (ICES, 2016a) :

- **Catégorie 1 -** Stocks disposant de données estimées suffisantes permettant une évaluation analytique quantitative complète (biomasse féconde, mortalité par pêche)
- **Catégorie 2 -** Stock disposant de données estimées suffisantes permettant une évaluation scientifique analytique (tendances de biomasse féconde, de mortalité par pêche mais pas de valeurs absolues)

- **Catégorie 3 -** Stock disposant uniquement de données permettant d'obtenir des tendances d'évolution de biomasse (sans évaluation analytique)
- **Catégorie 4 -** Stock disposant uniquement de données de capture fiables
- **Catégorie 5 -** Stock diposant uniquement de données de captures ou de débarquements
- **Catégorie 6 -** Stoks disposant uniquement de données de débarquements négligeables

Toutes les espèces d'élasmobranches gérées par le CIEM font parties des catégories 3, 5 et 6 à l'exception de l'Aiguillat commun (catégorie 1). L'ensemble des raies d'Atlantique Nord-Est font donc partie des catégories à données limitées. Ce manque se traduit par une gestion plurispécifique : depuis 2009 (2005 pour la Mer du Nord), la gestion des raies se fait via un TAC (Total de Captures Admissibles) plurispécifique. Ce TAC était non limitant au départ, il ne l'est devenu qu'au cours des années 2010 induisant de nouvelles mesures de gestion locale. Par exemple, en France, les raies de moins de 45 cm ne sont pas débarquées (JORF, 2017), leur survie aux rejets étant supposée élevée.

L'introduction du TAC plurispécifique a permis d'améliorer la gestion des raies néanmoins, comme vu dans le chapitre 4, les états des différentes espèces de raie dans le Golfe de Gascogne sont très variés. Les espèces de grande taille présente des états préocuppants alors que certaines espèces de plus petite taille semble être en bon voire très bon état écologique. Cette variété d'états ne se reflète pas dans le TAC plurispécifique bien qu'elle soit prise en compte en amont lors de l'évaluation des stocks par le WGEF. En effet, bien que des évaluations quantifiées de stock ne soit pas disponibles, des débarquements recommandés basés sur une approche de précaution sont émis chaque année par le groupe de travail, certaines espèces étant évaluées en années paires et les autres en années impaires (Fig. 7.8) (ICES, 2016b, 2017).

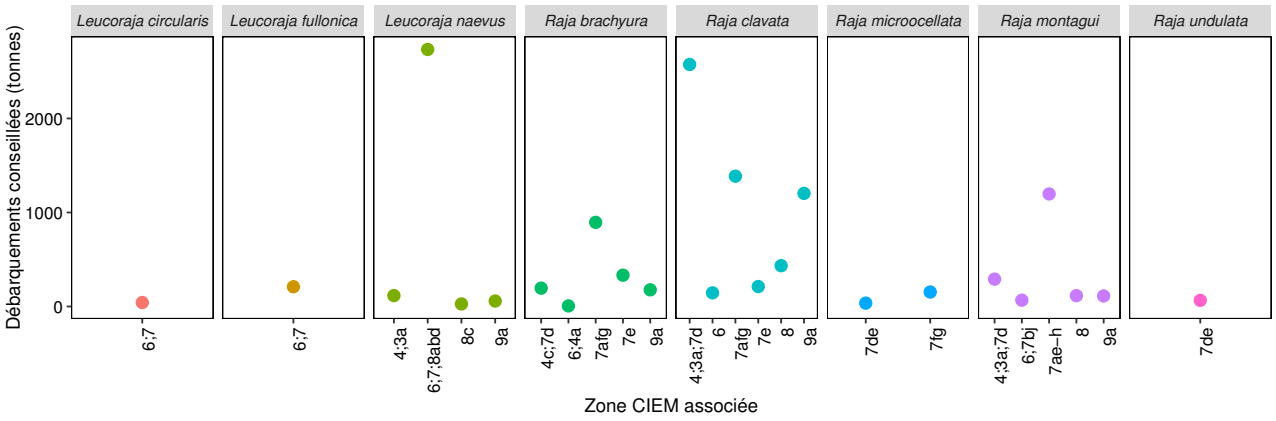


FIGURE 7.8 : Débarquements spécifiques conseillés par le CIEM en 2016-2017 pour les stocks de raies d'Atlantique Nord-Est (les débarquements spécifiques conseillés nuls ne sont pas affichés) (ICES, 2016b, 2017)

En se basant sur les débarquements recommandés du CIEM, il serait possible de passer à une gestion basée sur des TAC monospécifiques. Ce type de gestion présente de nombreux avantages quant à la conservation des espèces de raies mais aussi d'un point de vue économique pour les pêcheurs. En effet, le TAC plurispécifique alloué actuellement est en deça de la somme des débarquements conseillés par le CIEM. Passer à des TAC monospécifiques permettrait sans doute d'augmenter les débarquements de certaines espèces abondantes et donc d'induire un gain économique.

Néanmoins, le passage a une gestion monospécifique serait en pratique très compliqué. A cause de l'obligation de débarquement mise en place en 2016, toute espèce soumise à un TAC induit la fermeture des pêcheries la concernant si ce TAC est rempli (CEC, 2013). Dans le cas des raies, cela signifie que les TAC monospécifiques pourraient devenir limitants pour d'autres pécheries à fort interêt commercial et donc entrainer leur fermeture. Ce constat induit une remise en question dans la gestion des raies et plusieurs solutions sont envisagées dont l'abandon total des quotas sur les raies ou la mise en place d'une dérogation de l'obligation à débarquer.

3. Conclusion générale et perspectives

La première partie de ce travail de thèse a permis de mettre en évidence le besoin de définir des unités de populations adaptées à la gestion des espèces. L'étude de la connectivité démographie apparait comme le meilleur moyen d'y parvenir malgré sa forte exigence en données.

Les évaluations génétiques de la taille de population efficace effectuées dans cette thèse mettent en avant l'aspect conservation et santé à long terme d'une population exploitée mais ne permettent pas de mettre au point des mesures de gestion adaptées et proactives. De plus, l'effort d'échantillonnage nécessaire et plus que conséquent ne justifie probablement pas les estimations peu fiables et peu précises obtenues. Il existe néanmoins d'autres méthodes génétiques d'évaluation de l'abondance d'une population non étudiées dans cette thèse. La méthode "Close-Kin Mark-Recapture" est notamment très prometteuse bien qu'elle soit aussi gourmande en données que les estimations de N_e. Elle permet entre autres d'estimer la taille absolue d'une population via l'idendification génétique de paires d'individus apparentés (lien parent-descendant, frère/soeur et demi-frère/soeur) (Bravington et al., 2016).

Enfin, les évaluations démographiques effectuées démontrent qu'il est possible aujourd'hui d'évaluer l'état d'une population de raie afin d'en améliorer la gestion. La raie bouclée a ainsi été estimée dans un état proche du RMD dans le Golfe de Gascogne ainsi qu'en constante augmentation depuis les années 90 en Manche Est (résultats non montrés ici). Le modèle développé pourrait être appliqué à d'autres zones pour obtenir une image globale de l'état de la raie bouclée dans son aire de répartition. Il pourrait, de plus, être appliqué à d'autres espèces possedant des données suffisantes comme la raie fleurie. Pour les autres espèces, l'utilisation du modèle monospécifique permet d'estimer leur état, il pourrait ainsi être appliqué à d'autres zones et donc à d'autres espèces. Néanmoins, l'état des données actuelles ne permet pas encore d'estimer des valeurs de débarquements conseillés en absolu. Il y a eu depuis plusieurs années une prise de conscience quant à l'importance des espèces accessoires et le besoin de les gérer durablement, il est donc tout à fait probable que la situation s'améliore dans les années à venir.

BIBLIOGRAPHIE

Andrello, M. and Manel, S. (2015). MetaPopGen : an R package to simulate population genetics in large size metapopulations. *Molecular Ecology Resources*, 15(5) :1153–1162.

Andreotti, S., Rutzen, M., van der Walt, S., Von der Heyden, S., Henriques, R., Meÿer, M., Oosthuizen, H., and Matthee, C. (2016). An integrated mark-recapture and genetic approach to estimate the population size of white sharks in South Africa. *Marine Ecology Progress Series*, 552 :241–253.

Barbault, R. (1981). *Ecologie des populations et des peuplements*. Masson, Paris.

Bartley, D., Bagley, M., Gall, G., and Bentley, B. (1992). Use of linkage disequilibrium data to estimate effective size of hatchery and natural fish populations. *Conservation Biology*, 6 :365–375.

Beaumont, A. R., Boudry, P., and Hoare, K. (2010). *Biotechnology and genetics in fisheries and aquaculture*. Blackwell, Chichester ; Ames, Iowa, 2nd ed edition.

Begg, G. and Waldman, J. (1999). A holistic approach to fish stock identification. *Fisheries research*, 43 :37–46.

Begon, M., Townsend, C. R., and Harper, J. L. (2006). *Ecology : from individuals to ecosystems*. Blackwell Pub, Malden, MA, 4th ed edition.

Blower, D., Pandolfi, J., Bruce, B., Gomez-Cabrera, M., and Ovenden, J. (2012). Population genetics of Australian white sharks reveals fine-scale spatial structure, transoceanic dispersal events and low effective population sizes. *Marine Ecology Progress Series*, 455 :229–244.

Bonfil, R. (1994). *Overview of world elasmobranch fisheries*. Number 341 in FAO fisheries technical paper. FAO, Rome.

Booke, H. (1981). The conundrum of the stock concept - are nature and nurture definable in fisheries sciences ? *Canadian Journal of Fisheries and Aquatic Sciences*, 38 :1479–1480.

Bravington, M. V., Grewe, P. M., and Davies, C. R. (2016). Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. *Nature Communications*, 7 :13162.

Caddy, J. and Cochrane, K. (2001). A review of fisheries management past and present and some future perspectives for the third millennium. *Ocean & Coastal Management*, 44(9-10) :653–682.

Cadrin, S. X. and Dickey-Collas, M. (2015). Stock assessment methods for sustainable fisheries. *ICES Journal of Marine Science*, 72(1) :1–6.

Carvalho, G. and Hauser, L. (1994). Molecular genetics and the stock concept in fisheries. *Reviews in Fish Biology and Fisheries*, 4 :326–350.

Castro, A. L. F., Stewart, B. S., Wilson, S. G., Hueter, R. E., Meekan, M. G., Motta, P. J., Bowen, B. W., and Karl, S. A. (2007). Population genetic structure of Earth's largest fish, the whale shark (*Rhincodon typus*). *Molecular Ecology*, 16(24) :5183–5192.

CEC (2013). Regulation(EU) No 1380/2013 of the European Parliament and of Council of 11 December 2013 on the Common Fisheries Policy, amending Council Regulations (EC) No 1954/2003 and (EC) No 1224/2009 and repealing Council Regulations (EC) No 2371/2002 and (EC) No 639/2004 and Council Decision 2004/585/EC.

Chevolot, M. (2006). *Assessing genetic structure of thornback ray, Raja clavata : A thorny situation ?* PhD thesis, University of Groningen.

Chevolot, M., Ellis, J. R., Rijnsdorp, A. D., Stam, W. T., and Olsen, J. L. (2007). Multiple paternity analysis in the thornback ray, *Raja clavata* L. *Journal of Heredity*, 98(7) :712–715.

Chevolot, M., Ellis, J. R., Rijnsdorp, A. D., Stam, W. T., and Olsen, J. L. (2008). Temporal changes in allele frequencies but stable genetic diversity over the past 40 years in the Irish Sea population of thornback ray, *Raja clavata*. *Heredity*, 101(2) :120–126.

Chevolot, M., Hoarau, G., Rijnsdorp, A. D., Stam, W. T., and Olsen, J. L. (2006). Phylogeography and population structure of thornback rays (*Raja clavata* L., Rajidae). *Molecular Ecology*, 15(12) :3693–3705.

Chevolot, M., Reusch, T. B. H., Boele-Bos, S., Stam, W. T., and Olsen, J. L. (2005). Characterization and isolation of DNA microsatellite primers in *Raja clavata* L. (thornback ray, Rajidae). *Molecular Ecology Notes*, 5(2) :427–429.

Cornou, A.S., Dimeet, J, Tetard A., Gaudou, O., Dube, B., Fauconnet, L., and Rochet, M.J. (2013). Observations à bord des navires de pêche professionnelle. Bilan de l'échantillonnage 2012. Technical report, IFREMER - http ://dx.doi.org/10.13155/27787.

Cornou, A.S., Dimeet, J., Tetard, A., Gaudou, O., Quinio-Scavinner, M., Fauconnet, L., Dube, B., and Rochet, M.J. (2015a). Observations à bord des navires de pêche professionnelle. Bilan de l'échantillonnage 2013. Technical report, IFREMER - http ://dx.doi.org/10.13155/35856.

Cornou, A.S., Quinio-Scavinner, M., Delaunay, D., Dimeet, J., Goascoz, N., Dube, B., Fauconnet, L., and Rochet, M.J. (2015b). Observations à bord des navires de pêche professionnelle. Bilan de l'échantillonnage 2014. Technical report, IFREMER - http ://dx.doi.org/10.13155/39722.

Davidson, L. N. K., Krawchuk, M. A., and Dulvy, N. K. (2016). Why have global shark and ray landings declined : improved management or overfishing ? *Fish and Fisheries*, 17 :438–458.

De Oliveira, J. A. A., Ellis, J. R., and Dobby, H. (2013). Incorporating density dependence in pup production in a stock assessment of NE Atlantic spurdog Squalus acanthias. *ICES Journal of Marine Science*, 70(7) :1341–1353.

Debouzie, D. (1999). La notion de population en dynamique et génétique des populations. *Nature Sciences Sociétés*, 7(4) :19–26.

Delpiani, G. (2016). Reproductive biology of the southern thorny skate *Amblyraja doellojuradoi* (Chondrichthyes, Rajidae) : reproduction of *Amblyraja doellojuradoi*. *Journal of Fish Biology*, 88(4) :1413–1429.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., and Ovenden, J. R. (2014). NeEstimator V2 : re-implementation of software for the estimation of contemporary effective population size N_e from genetic data. *Molecular Ecology Resources*, 14(1) :209–214.

Dube, B., Dimeet, J., Rochet, M.-J., Tetard, A., Gaudou, O., Messannot, C., Fauconnet, L., Morizur, Y., Biseau, A., and Salaun, M. (2012). Observations à bord des navires de pêche professionnelle. Bilan de l'échantillonnage 2011. Technical report, OBSMER.

Dudgeon, C. L., Blower, D. C., Broderick, D., Giles, J. L., Holmes, B. J., Kashiwagi, T., Krück, N. C., Morgan, J. A. T., Tillett, B. J., and Ovenden, J. R. (2012). A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *Journal of Fish Biology*, 80(5) :1789–1843.

Dudgeon, C. L. and Ovenden, J. R. (2015). The relationship between abundance and genetic effective population size in elasmobranchs : an example from the globally threatened zebra shark *Stegostoma fasciatum* within its protected range. *Conservation Genetics*, 16(6) :1443–1454.

Dulvy, N. K., Fowler, S. L., Musick, J. A., Cavanagh, R. D., Kyne, P. M., Harrison, L. R., Carlson, J. K., Davidson, L. N., Fordham, S. V., Francis, M. P., Pollock, C. M., Simpfendorfer, C. A., Burgess, G. H., Carpenter, K. E., Compagno, L. J., Ebert, D. A., Gibson, C., Heupel, M. R., Livingstone, S. R., Sanciangco, J. C., Stevens, J. D., Valenti, S., and White, W. T. (2014). Extinction risk and conservation of the world's sharks and rays. *eLife.00590*, 3.

Dulvy, N. K., Jennings, S., Rogers, S. I., and Maxwell, D. L. (2006). Threat and decline in fishes : an indicator of marine biodiversity. *Canadian Journal of Fisheries and Aquatic Sciences*, 63(6) :1267–1275.

Dulvy, N. K., Metcalfe, J. D., Glanville, J., Pawson, M. G., and Reynolds, J. D. (2000). Fishery stability, local extinctions, and shifts in community structure in skates. *Conservation Biology*, 14(1) :283–293.

Ellis, J. R. and Keable, J. (2008). Fecundity of Northeast Atlantic spurdog (*Squalus acanthias*). *ICES Journal of Marine Science*, 65(6) :979–981.

Ellis, J. R. and Shackley, S. E. (1995). Observations on egg-laying in the thornback ray. *Journal of Fish Biology*, 46(5) :903–904.

FAO (2016a). La situation mondiale des pêches et de l'aquaculture 2016. contribuer à la sécurité alimentaire et à la nutrition de tous. Technical report, FAO.

FAO (2016b). Logiciel pour la pêche et de l'aquaculture. FishStatJ> – logiciel pour les séries chronologiques de données statistiques sur les pêches. in : Département des

pêches et de l'aquaculture de la FAO [en ligne]. Rome. [téléchargé le 17 january 2017] - http ://www.fao.org/fishery/statistics/software/fishstatj/fr.

Fauconnet, L., Badts, V., Biseau, A., Dimeet, J., Dinther, C., Dube, B., Gaudou, O., Lorance, P., Messannot, C., Nikolic, N., Peronnet, I., Reecht, Y., Rochet, M.-J., and Tetard, A. (2011). Observations à bord des navires de pêche professionnelle. Bilan de l'échantillonnage 2010. Technical report, IFREMER - http ://archimer.ifremer.fr/doc/00054/16490.

Feldheim, K. A., Gruber, S. H., and Ashley, M. V. (2001). Population genetic structure of the lemon shark (*Negaprion brevirostris*) in the western Atlantic : DNA microsatellite variation. *Molecular Ecology*, 10(2) :295–303.

Figueiredo, I., Moura, T., Bordalo-Machado, P., Neves, A., Rosa, C., and Serrano Gordo, L. (2007). Evidence for temporal changes in ray and skate populations in the Portuguese coast (1998–2003) – its implications in the ecosystem. *Aquatic Living Resources*, 20(1) :85–93.

Frankham, R. (1995). Effective population size/adult population size ratios in wildlife : a review. *Genetical Research*, 66(02) :95.

Franklin, I. (1980). Evolutionary change in small populations. In *Conservation biology : an evolutionary-ecological perspective*, pages 135–149. Sinauer Associates, Sunderland, Mass.

Gregory, T. (2016). Animal Genome Size Database : Species Record - http ://www.genomesize.com/index.php.

Hamilton, M. B. (2009). *Population genetics*. Wiley-Blackwell, Chichester, UK ; Hoboken, NJ (eds).

Hanski, I. (1991). Metapopulation dynamics : brief history and conceptual domain. *Biological Journal of the Linnean Society*, 42 :3–16.

Hare, M. P., Nunney, L., Schwartz, M. K., Ruzzante, D. E., Burford, M., Waples, R. S., Ruegg, K., and Palstra, F. (2011). Understanding and estimating effective population size for practical application in marine species management : applying effective population size estimates to marine species management. *Conservation Biology*, 25(3) :438–449.

Hauser, L., Adcock, G. J., Smith, P. J., Bernal Ramírez, J. H., and Carvalho, G. R. (2002). Loss of microsatellite diversity and low effective population size in an overexploited population of new zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences*, 99(18) :11742–11747.

Hedgecock, D., V, C., and Waples, R. (1992). Effective population nnumber of shellfish broodstocks estimated from temporal variance in allelic frequencies. *Aquaculture*, 108 :215–232.

Heinke, F. (1898). Naturgeschichte des herings. Deutscher Seefischerei Verein. Band III.

Hilborn, R. (1992). Current and future trends in fisheries stock assessment and management. *South African Journal of Marine Science*, 12(1) :975–988.

Hilborn, R. and Walters, C. J. (1992). *Quantitative fisheries stock assessment : choice, dynamics, and uncertainty*. Chapman and Hall, New York.

Hoareau, G., Boon, E., Jongma, D., Ferber, S., Palsson, J., Van der Veer, H., Rijinsdorp, A., Stam, W., and Olsen, J. (2005). Low effective population size and evidence for inbreeding in an overexploited flatfish, plaice (*Pleuronectes platessa L.*). *Proceedings of the Royal Society B : Biological Sciences*, 272(1562) :497–503.

Holden, M. J. (1972). The growth rates of *Raja brachyura*, *R. clavata* and *R. montagui* as determined from tagging data. *ICES Journal of Marine Science*, 34(2) :161–168.

Hunter, E., Buckley, A., Stewart, C., and Metcalfe, J. (2005a). Migratory behaviour of the thornback ray, *Raja clavata*, in the southern North Sea. *Journal of the Marine Biological Association of the UK*, 85(05) :1095.

Hunter, E., Buckley, A., Stewart, C., and Metcalfe, J. (2005b). Repeated seasonal migration by a thornback ray in the southern North Sea. *Journal of the Marine Biological Association of the UK*, 85(05) :1199.

Hutchinson, W. F., Oosterhout, C. v., Rogers, S. I., and Carvalho, G. R. (2003). Temporal analysis of archived samples indicates marked genetic changes in declining north sea cod (*Gadus morhua*). *Proceedings of the Royal Society of London B : Biological Sciences*, 270(1529) :2125–2132.

ICES (2016a). ICES Advice basis - Introduction to advice - http ://www.ices.dk/sites/pub/publication

ICES (2016b). Report of the Working Group on Elasmobranchs Fishes (WGEF). Technical Report ICES CM/ACOM :20, ICES, ICES Headquarters.

ICES (2017). Report of the Working Group on Elasmobranchs Fishes (WGEF). Technical Report ICES CM 2017/ACOM :16, ICES, ICES Headquarters.

Ihssen, P., Booke, H., Casselman, J., McGlade, J., Payne, N., and Utter, F. (1981). Stock identification : materials and methods. *Canadian Journal of Fisheries and Aquatic Sciences*, 38(12) :1838–1855.

Jorde, P. and Ryman, N. (1995). Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics*, 139(2) :1077–1090.

JORF (2017). Texte numéro 6, journal officiel de la république française, numéro 0073 du 26 mars 2017.

Kerr, L. A., Hintzen, N. T., Cadrin, S. X., Clausen, L. W., Dickey-Collas, M., Goethel, D. R., Hatfield, E. M., Kritzer, J. P., and Nash, R. D. (2016). Lessons learned from practical approaches to reconcile mismatches between biological population structure and stock units of marine fish. *ICES Journal of Marine Science : Journal du Conseil*, 74(6) :1708–1722.

Kuhnert, P. M., Griffiths, S., and Brewer, D. (2011). Assessing population changes in bycatch species using fishery-dependent catch rate data. *Fisheries Research*, 108(1) :15–21.

Lande, R. (1995). Mutation and conservation. *Conservation Biology*, 9(4) :782–791.

Laurec, A. and Le Guen, J. (1981). Dynamique des populations marines exploitées. Rapports Scientifiques et Techniques 45, CNEXO/Centre Océanologique de Bretagne.

Lavery, S. and Shaklee, J. (1989). Population genetics of two tropical sharks, *Carcharhinus tilstoni* and *C. sorrah* , in Northern Australia. *Marine and Freshwater Research*, 40(5) :541.

Le Quesne, W. J. and Jennings, S. (2012). Predicting species vulnerability with minimal data to support rapid risk assessment of fishing impacts on biodiversity : fishing impacts on fish biodiversity. *Journal of Applied Ecology*, 49(1) :20–28.

Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin Entomological Society of America*, 15 :237–240.

Linné, C. v. (1758). *Caroli Linnaei...Systema naturae per regna tria naturae :secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis.* Impensis Direct. Laurentii Salvii.

Lowe, A., Harris, S., and Ashton, P. (2004). *Ecological genetics : design, analysis, and application.* Blackwell Pub, Malden, MA, USA.

Martin, C., Vaz, S., Ellis, J., Lauria, V., Coppin, F., and Carpentier, A. (2012). Modelled distributions of ten demersal elasmobranchs of the eastern English Channel in relation to the environment. *Journal of Experimental Marine Biology and Ecology*, 418-419 :91–103.

Martinez-Lage, A., Gonzalez-Tizon, A., and Mendez, J. (1996). Chromosome diferences between european mussel populations (genus mytilus). *Caryologia*, 49(3-4) :343–355.

Maunder, M. N. and Piner, K. R. (2015). Contemporary fisheries stock assessment : many issues still remain. *ICES Journal of Marine Science*, 72(1) :7–18.

Maxwell, D., Stelzenmüller, V., Eastwood, P., and Rogers, S. (2009). Modelling the spatial distribution of plaice (*Pleuronectes platessa*), sole (*Solea solea*) and thornback ray (*Raja clavata*) in UK waters for marine management and planning. *Journal of Sea Research*, 61(4) :258–267.

McCully Phillips, S. R., Scott, F., and Ellis, J. R. (2015). Having confidence in productivity susceptibility analyses : a method for underpinning scientific advice on skate stocks ? *Fisheries Research*, 171 :87–100.

Nomura, T. (2008). Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evolutionary Applications*, 1(3) :462–474.

Nygren, A., Nilsson, B., and Jahnke, M. (1971). Cytological studies in Hypotremata and Pleurotremata (Pisces). *Hereditas*, (67) :275–282.

Ovenden, J. R., Leigh, G. M., Blower, D. C., Jones, A. T., Moore, A., Bustamante, C., Buckworth, R. C., Bennett, M. B., and Dudgeon, C. L. (2016). Can estimates of genetic effective population size contribute to fisheries stock assessments ? *Journal of Fish Biology*, 89(6) :2505–2518.

Pasolini, P., Ragazzini, C., Zaccaro, Z., Cariani, A., Ferrara, G., Gonzalez, E. G., Landi, M., Milano, I., Stagioni, M., Guarniero, I., and Tinti, F. (2011). Quaternary geographical sibling speciation and population structuring in the eastern atlantic skates (suborder rajoidea) *Raja clavata* and *Raja straeleni*. *Marine Biology*, 158(10) :2173–2186.

Portnoy, D. S., McDowell, J. R., McCandless, C. T., Musick, J. A., and Graves, J. E. (2009). Effective size closely approximates the census size in the heavily exploited western Atlantic population of the sandbar shark, *Carcharhinus plumbeus*. *Conservation Genetics*, 10(6) :1697–1705.

Pudovkin, A., Zaykin, D., and Hedgecock, D. (1996). On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics*, 144(1) :383–387.

Pulliam, H. (1988). Sources, sinks, and population regulation. *American Naturalist*, 132(5) :652–661.

Quero, J., Leaute, J., and Spitz, J. (2017). Faune francaise de l'atlantique. batoïdes : Torpediniformes, rajiformes & myliobatiformes (craniata : Elasmobranchii). *Société des Sciences Naturelles de la Charente-Maritime*, 10(8) :863–903.

Quéro, J., Porcher, P., and Vayne, J. (2003). *Guide des poissons de l'Atlantique européen*. Delachaux et Niestlé, Paris.

Quéro, J. and Vayne, J. (2005). *Les poissons de mer des pêches françaises*. Delachaux et Niestlé, Paris. p58-59.

Quéro, J.-C. and Vayne, J.-J., editors (1997). *Les poissons de mer des pêches françaises : identification, inventaire et répartition de 209 espèces*. Les encyclopédies du naturaliste. Delachaux et Niestlé, Lausanne.

Ramade, F. (2003). *Eléments d'écologie : écologie fondamentale*. Dunod, 2ème edition.

Schwartz, M., Luikart, G., and Waples, R. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology & Evolution*, 22(1) :25–33.

Serra-Pereira, B., Figueiredo, I., Farias, I., Moura, T., and Gordo, L. S. (2008). Description of dermal denticles from the caudal region of *Raja clavata* and their use for the estimation of age and growth. *ICES Journal of Marine Science*, 65(9) :1701–1709.

Serra-Pereira, B., Figueiredo, I., and Gordo, L. S. (2011). Maturation, fecundity, and spawning strategy of the thornback ray, *Raja clavata* : do reproductive characteristics vary regionally ? *Marine Biology*, 158(10) :2187–2197.

SharkTrust (2009). An illustrated compendium of sharks, skates, rays and chimaera. Chapter 1 : the British Isles. Part 1 : skates and rays - http ://www.vliz.be/en/catalogue ?module=ref&refid=247504.

Smith, R. L. and Smith, T. M. (2001). *Ecology & field biology*. Benjamin Cummings, San Francisco, 6th edition.

Spies, I., Punt, A. E., and Wilberg, M. (2015). The utility of genetics in marine fisheries management : a simulation study based on Pacific cod of Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(9) :1415–1432.

Stingo, V. (1979). New developments in vertebrate cytotaxonomy, II. The chromosomes of the cartilaginous fishes. *Genetica*, (50) :227–239.

Stingo, V., Du Buit, M.-H., and Odierna, G. (1980). Genome size of some selachian fishes. *Bolletino di zoologia*, 47(1-2) :129–137.

Upton, G. J. G. and Cook, I. (2006). *A dictionary of statistics*. Oxford paperback reference. Oxford University Press, Oxford ; New York, 2nd ed edition.

Valsecchi, E., Vacchi, M., and di Sciara, G. N. (2005). Characterization of a new molecular marker for investigating skate population genetics : analysis of three Mediterranean skate species (genus *Raja*) of commercial interest as a test case. *Journal of Northwest Atlantic Fishery Science*, 35 :225–231.

Walker, P. (1997). Distribution, movement and stock structure of three ray species in the North Sea and eastern English Channel. *ICES Journal of Marine Science*, 54(5) :797–808.

Waples, R. S. (2016). Tiny estimates of the N_e/N ratio in marine fishes : Are they real ? *Journal of Fish Biology*, 89(6) :2479–2504.

Waples, R. S., Antao, T., and Luikart, G. (2014). Effects of overlapping generations on Linkage Disequilibrium estimates of effective population size. *Genetics*, 197(2) :769–780.

Waples, R. S. and Do, C. (2008). Ldne : a program for estimating effective population size from data on linkage disequilibrium : Computer Programs. *Molecular Ecology Resources*, 8(4) :753–756.

Waples, R. S., Do, C., and Chopelet, J. (2011). Calculating N_e and N_e/N in age-structured populations : a hybrid Felsenstein-Hill approach. *Ecology*, 92(7) :1513–1522.

Waples, R. S. and Yokota, M. (2006). Temporal estimates of effective population size in species with overlapping generations. *Genetics*, 175(1) :219–233.

Ward, R., Woodwark, M., and Skibinski, D. (1994). A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. *Journal of Fish Biology*, 44 :213–232.

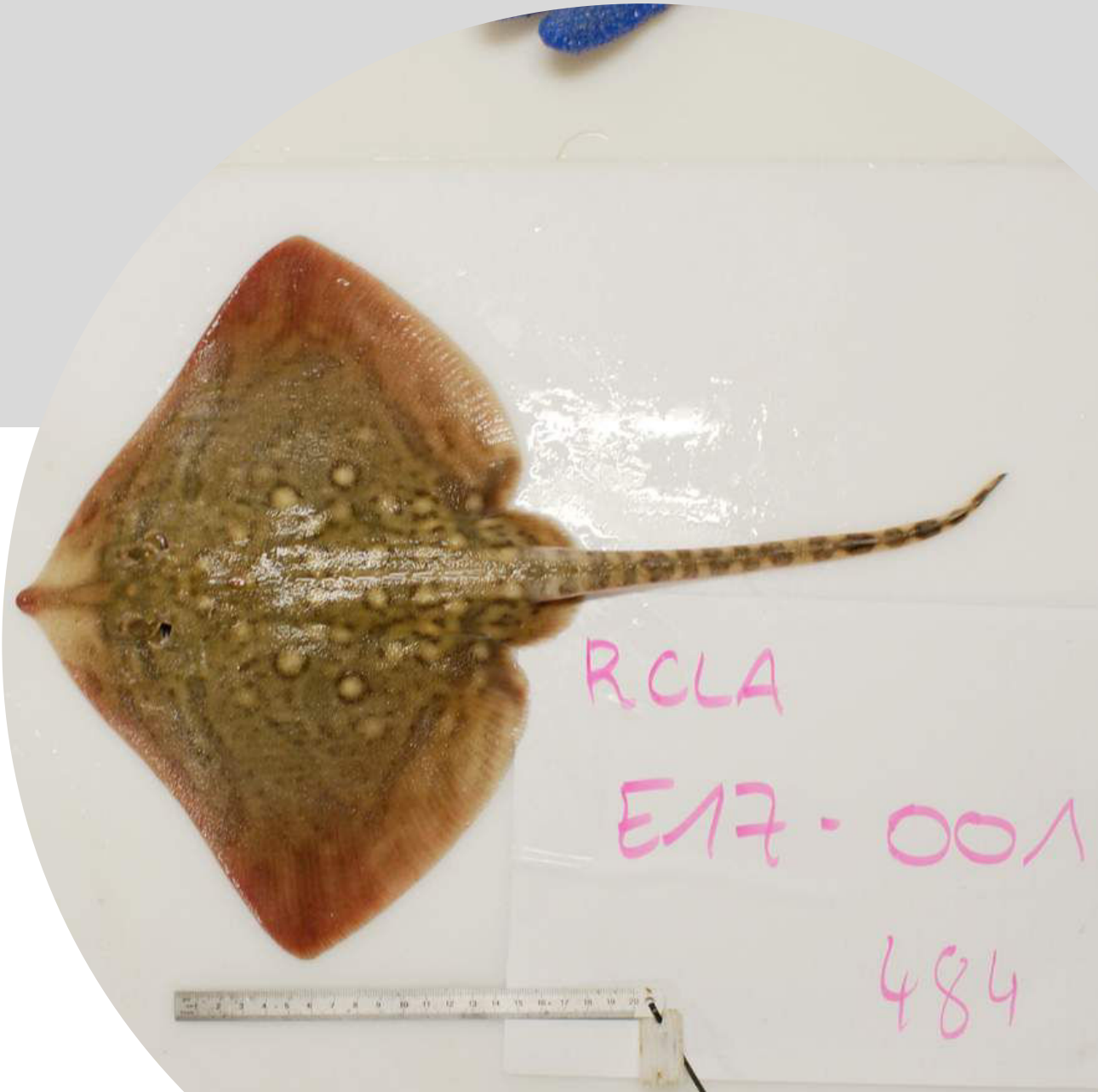
Weltz, K., Lyle, J. M., Ovenden, J., Morgan, J. A. T., Moreno, D. A., and Semmens, J. M. (2017). Application of environmental DNA to detect an endangered marine skate species in the wild. *PLOS ONE*, 12(6) :e0178124.

Whittamore, J. and McCarthy, I. (2005). The population biology of the thornback ray, *Raja clavata* in Caernarfon Bay, North Wales. *Journal of the Marine Biological Association of the UK*, 85(05) :1089.

Wiegand, J., Hunter, E., and Dulvy, N. K. (2011). Are spatial closures better than size limits for halting the decline of the North Sea thornback ray, *Raja clavata* ? *Marine and Freshwater Research*, 62(6) :722.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, (16) :97–159.

GLOSSAIRE



Abondance Nombre total d'individus d'un stock, toutes classes d'âge confondues. 1

Allèle Un allèle est une version variable d'un même gène ou d'un même locus génétique. 92

Capacité biotique Biomasse maximale de population que le milieu peut accueillir. 37

Caryogramme Arrangement standard de l'ensemble des chromosomes d'une cellule à partir d'une prise de vue microscopique. 6

Chromosomes homologues Chromosomes associés dans une même paire dans une cellule durant la méiose. 6

Courbe de survie de type I Ce type de survie est caractérisé par des probabilités de survie élevées aux premiers stades de vie, suivis d'un déclin rapide de ces probabilités pour les stades plus avancés. Il correspond à des espèces à longue durée de vie et avec de faibles fécondités. 7

Courbe de survie de type III Dans ce type de survie, la mortalité la plus élevée est expérimentée dans les premiers stades de vie, avec des taux de mortalité faibles pour le reste du cycle. Il caractérise des espèces produisant beaucoup de petits avec une maturité sexuelle précoce. 7

Débarquement Quantité de poissons capturés et mis à terre. 4, 37

Dème Groupe monospécifique d'individus, isolé et génétiquement différencié des autres groupes voisins. 2

Diversité génétique Degré de variété des gènes au sein d'une même espèce. 1

Espèce démersale Les espèces démersales sont constituées d'animaux vivant à proximité du fond. 5

Espèces accessoires Espèces involontairement pêchées mais présentant un intérêt commercial et donc débarquées. 1

Générations chevauchantes Se dit d'une espèce dont les individus vivent plusieurs années et où des croisements inter-générations sont possibles. 16

Génome Ensemble du matériel génétique d'un individu. 6

Locus Emplacement précis sur un chromosome. 15

Microsatellites Portion du génome constituée de petites séquences hautement répétitives de deux à dix nucléotides. 7, 92

Pêcherie Somme de toutes les activités halieutiques portant sur une ressource donnée. 1

Palangre Engin de capture se composant d'une ligne principale (grosse corde, filin d'acier ou de nylon) sur laquelle sont montés des avançons avec des hameçons appâtés. 5

Population Ensemble des individus se reproduisant entre eux et vivant dans un ecosystème donné et possédant des caractères communs transmissible par hérédité. 1

Population idéale Au sens d'Hardy-Weinberg, se dit d'une population présentant les caractéristiques suivantes : taille infinie, espèce diploïde à reproduction sexuée, rencontre aléatoire des individus et des gamètes, absence de migration, absence de mutation, absence de sélection naturelle et absence de générations chevauchantes. 15

Rejets Partie des captures qui n'est pas débarquée mais rejetée à la mer pour des raisons diverses (taille illégale, poisson endommagé, absence de marché ou dépassement des quotas). 5

Rendement Maximum Durable Captures maximales pouvant être maintenues à long terme. Abbréviation : RMD ; MSY en anglais. 140

Sex ratio Rapport du nombre mâles et de femelles au sein d'une espèce à reproduction sexuée. 2

Single-Nucleotide Polymorphism Variation (polymorphisme) d'une seule paire de bases du génome entre individus d'une même espèce. Abbréviation : SNP. 9, 92

Stock Ensemble des individus exploitables d'une espèce dans une zone donnée. Chaque stock est considéré n'avoir que des contacts limités avec le stock voisin ; il a également sa dynamique propre (croissance, reproduction). 1

Taille de population absolue Nombre total d'individus dans une population (incluant la fraction immature). 7, 16

Taille de population efficace Nombre d'individus dans une population idéale selon Wright-Fisher se reproduisant et contribuant génétiquement à la prochaine génération. 7, 16

Utilisation durable Utilisation des éléments constitutifs de la diversité biologique d'une manière et à un rythme qui n'entraînent pas leur appauvrissement à long terme, et sauvegardent ainsi leur potentiel pour satisfaire les besoins et les aspirations des générations présentes et futures. 19

ANNEXE A

État de l'art : la génétique des populations pour les halieutes

Table of contents

1. Introduction

The raw material of **evolution** * and adaptation to local environments is the genetic variability of individuals in local **population** (Smith and Smith, 2001) which ultimately leads to speciation. Genetic variability can be created by **mutation** and **recombination** but its distribution is defined by evolutive forces such as **migration**, **selection** and **genetic drift**. The pattern of genetic variation can be investigated to understand evolution and adaptation of a species or a population (Lowe et al., 2004; Hamilton, 2009). Such investigations are not recent. Ecological genetics were already a central element of research as soon as 1922 but over the last 30 years, the application of genetics in ecology has increased for a wide variety of biological problems (Dudgeon et al., 2012; Lowe et al., 2004; Portnoy and Heist, 2012). The development of molecular markers, such as SNP (Single Nucleotide Polymorphism) or microsatellites was also fundamental for the spread of genetic analyses. Moreover the increasing affordability of new analyses and the pressing need to address critical conservation and management issues led to the development of genetic analyses in numerous fields including investigations of stock structure and population demography in fisheries. A great virtue of genetic approaches is that a small tissue sample collected from the living or deceased animal at any age contains its complete nuclear and mitochondrial genomic information (Dudgeon et al., 2012). One famous example is the extraction of DNA from frozen woolly mammoths (Smith and Smith, 2001). DNA analysis in general allows the investigation of genetic relationships at several organisational levels (species, population, individual) and can be applied to numerous ecological problematic like natural selection, **mating system** or migration (Smith and Smith, 2001; Lowe et al., 2004; Dudgeon et al., 2012).

Contemporary ecological genetics investigate the origin and maintenance of the genetic variation within and between populations. Population size, **population structure**, the interactions between local selection and genetic drift are some of the main issues studied nowadays (Ryman and Utter, 1987; Lowe et al., 2004; Nikolic, 2009; Beaumont et al., 2010; Dudgeon et al., 2012; Portnoy and Heist, 2012). Among them, genetic variation is a great way to study the effect of loss in a population. Populations of many plants and animal species are being reduced to small, sometimes isolated, populations leading genetic deterioration. Genetic analysis allows us to quantify the loss of genetic diversity and the loss of adaptative potential and thus provide advice for the conservation and the management of such species (Smith and Smith, 2001). When the genetic population structure of a species is known, the distribution of sub-populations can be estimated and used for harvest regulation (Ryman and Utter, 1987).

Fisheries management requires understanding biological principles underlying resource dynamics. Management has focused for many years on ecology and population dynamics to the

*Words in bold are in glossary

detriment of understanding **population genetics**. Ecological and population dynamics can be seen as short-term focus and genetics as long-term and short-term focus. Genetics has not been forgotten in all studies, and its importance has maybe been more admitted for fishes than for other vertebrates. As early as the beginning of the XXe century, genetics was studied for fish sub-populations. In 1983, 15% of the concerns relating to the genetics of animals referred to fish (Ryman and Utter, 1987). In parallel, estimation of the **effective population size** (N_e) in fisheries management and marine conservation is quite recent and is emerging since the beginning of the XXI century (Dudgeon et al., 2012) while it has featured in terrestrial conservation efforts for decades (Schwartz et al., 2007). Genetic monitoring can estimate a population's effective size to evaluate abundance and genetic health, complementing conventional stock assessment methods (Hamilton, 2009). N_e is indeed a fundamental parameter in evolutionary biology (Hare et al., 2011) and conservation biology (Waples and Do, 2010). It is also a potential indicator for conservation and fisheries management and reference points exit to be compared with estimated N_e (Smith and Smith, 2001; Portnoy and Heist, 2012). Effective population size indicates a population's current and future viability (Hare et al., 2011). The aim is to preserve high genetic variability and a sufficient effective population size to maximise the adaptive potential in the face of new environmental conditions. However, using genetics can be a tough job, especially because of the specialised vocabulary and the many assumptions required for practical applications.

In this literature review, some fundamental aspects of population genetics are presented to introduce the notion of effective population size. Estimation methods, ecological processes affecting it, and its use in fishery research are also discussed. The definition of population used is the genetic one (but see glossaries).

2. Basics of genetics

Inherited characteristics of a species and variations in individuals are transmitted from parents to offspring. The sum of the hereditary information carried by the individual is the **genotypes** which directs the development of the individual and underlies the morphological, physiological and behavioral characteristics of the individual. The universal support of the hereditary information is the desoxyribonucleic acid, the DNA, present in every cell of the organism (Schleif, 1993; Hartl and Jones, 1998). It is a complex molecule with two strands in the shape of a double helix. Strands are linked by nitrogenous bases, which are paired: adenine (A) and thymine (T) (the purines bases), cytosine (C) and guanine (G) (the pyrimidine bases). The hereditary information is coded by the sequential pattern in which the base pairs occur. Each species is unique with its own base pairs arrangement and its own number of base pairs (Smith and Smith, 2001; Beaumont et al., 2010).

In eukaryotic cells, DNA is present in larger units called chromosomes going by pair, human possess 23 pairs of chromosomes for example. Each chromosome carries units of heredity, the **genes** which are also paired in the body cells. One version of a gene is inherited from the mother and the other one from the father, where each one forms a **haplotype**. Processes leading to a new eggs cell are summarized in Fig.1 which shows how maternal and paternal haplotypes are transmitted to the next generation during the fertilisation. Processes such as replication and meiosis are not explicated in details. The position of a gene on a chromosome is the **locus**; genes occupying the locus on a pair of chromosomes are called **alleles**. If each member of the allele's pair affects a given trait in the same manner, the two alleles are homozygous; if not, they are heterozygous (Hamilton, 2009). During the formation of germ cells, the pairs of chromosomes are split, so that each resulting cell nucleus receives only one-half of the full number of chromosomes (Beaumont et al., 2010).

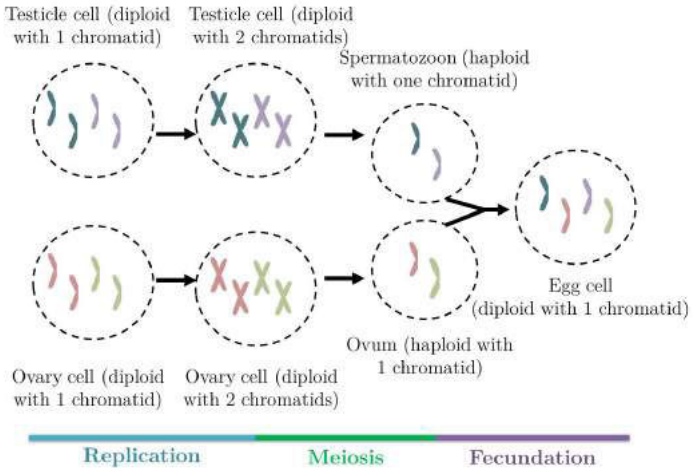


Figure 1: Fertilisation and inheritance of maternal and paternal haplotypes

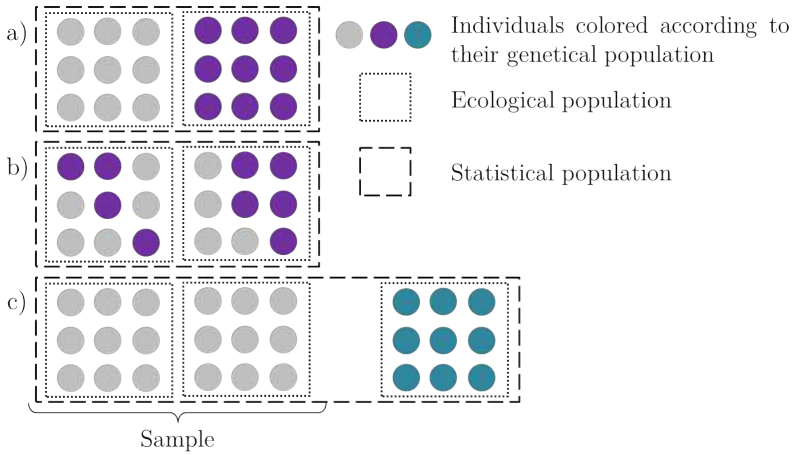
During the separation, several processes, such as recombination, occur which modify the genetic information and create variation. However, they are several others sources of variation of genetic information, like mutation (Smith and Smith, 2001). The genetic variation is a key notion in population genetics which inform of the genetic health of a population and its structure. The genetic variability of a population is a global measure of genetic differences among individuals. Those differences are the bases for the multiple observed phenotypes.The genetic variability is the basis of natural selection and so the driving force of evolution. Its estimate is directly linked to a population's capacity to respond to environmental changes (which is also determined by demographic features such as fecundity and natural mortality). For a natural population, its degree of genetic variability is directly associated to its viability. Several indicators are available to characterize the genetic variability and are explicated below.

3. Introduction to evolutionary genetics

The marine environment has often been considered as very dispersive and so, marine species have often been supposed as genetically unstructured. However, numerous marine species present a strong spatial genetic structure...

BOX 1. It's a population, isn't it ?

The word "population" has been used previously, imprecisely, to designate a group of organisms belonging to the same species. At this point, we need to clarify the concept. At least, three different population definitions can be recognized. In statistics, a population represents all the items under study. In ecology, a population is a group of individuals of the same species within the same habitat at the same time. However, in population genetics, a population designates a group of individuals of the same species which live in a geographical zone small enough that each member can reproduce with every other member. So a genetic population encompasses all individuals connected by gene flow (Hartl, 1994; Lowe et al., 2004). The three definitions may coincide, although more frequently they will not :



The figure represents examples for the relationships between statistical (dashed line), ecological (dotted lines) and genetical (colored points) populations. a) the sampled individuals comprise the entire statistical population and the ecological and genetical populations coincide. b) the sampled individuals comprise the entire statistical population but the ecological and genetical populations do not coincide. c) the sampled individuals do not comprise the entire statistical population but ecological and genetical populations coincide. Figure inspired from Lowe et al. (2004).

In this review, we use the genetical definition of a population.

3.1 Some principles of evolutionary biology

Evolutionary biology is a body of statements about the processes of evolution that are believed to have caused the history of evolutionary events (Pigliucci, 2009). Evolution is a change in the gene pool of a population. In order to understand evolution, it is necessary to view populations as a collection of individuals, each harboring a different set of traits. A single organism is never typical of an entire population unless there is no variation within that population. Individual organisms do not evolve, they retain the same genes throughout their life. The differentiation of a species into genetically different populations is a fundamental part of the process of evolution and requires genetic variation (Sober, 1994); several mechanisms exist to increase or create genetic variation and mechanisms to decrease it.

Genetic variation has two components: allelic diversity and non random associations of alleles. In most populations, there are enough loci and enough different alleles that every individual, identical twins excepted, has a unique combination of alleles.

The evolution of the genetic variability of natural populations is driven by four forces (Grant and Waples, 2000):

Mutation

Mutation represents the main source of genetic variability (Hartl, 1994). The cellular machinery that copies DNA sometimes makes mistakes which alter the sequence of a gene. This is called a mutation. There are many kinds of mutations. A point mutation is a mutation in which one "letter" of the genetic code is changed to another. Lengths of DNA can also be deleted or inserted in a gene. Finally, genes or parts of genes can become inverted or duplicated. Mutations are often known for being deleterious (cystic fibrosis, trisomy) but most mutations are thought to be neutral with regards to fitness. (but see Kimura, 1990) Mutations provide the raw material for evolution but they are rare: typical rates of mutation are between 10^{-10} and 10^{-12} mutations per base pair of DNA per generation. In contrat, rates of mutation in microsatellite loci are up to 10^{-6}.

Natural selection

Natural selection is the only mechanism of adaptive evolution. It is defined as differential reproductive success of pre-existing classes of genetic variants in the gene pool. Natural selection can maintain or deplete genetic variation depending on how it acts. When selection acts to weed out deleterious alleles, or causes an allele to sweep to **fixation,** it depletes genetic variation. When heterozygotes are more fit than either of the homozygotes, however, selection causes genetic variation to be maintained. Natural selection may not lead a population to have the optimal set of traits. In any population, there would be a certain combination of possible alleles that would

produce the optimal set of traits (the global optimum); but there are other sets of alleles that would yield a population almost as adapted (local optima). Transition from a local optimum to the global optimum may be hindered or forbidden because the population would have to pass through less adaptive states to make the transition. Natural selection only works to bring populations to the nearest optimal point. This idea is Sewall Wright's adaptive landscape (Wright, 1932). This is one of the most influential models that shapes how evolutionary biologists view evolution.

Natural selection can be broken down into many components, of which survival is only one. Sexual attractiveness is a very important component of selection, so much so that biologists use the term sexual selection when they talk about this subset of natural selection. Sexual selection is natural selection operating on factors that contribute to an organism's mating success.

Gene flow

New organisms may enter a population by **gene flow** from another population. If they mate within the population, they can bring new alleles to the local gene pool. In marine species, gene flow can result from **dispersal** at different life-history stages, including active movement as adults and juveniles and passive transport during pelagic egg and larval stages (Grant and Waples, 2000) but also by migration, ie round-trip movements of animals between regions or habitats. In genetics, migration is often considered as movements of individuals between sub-populations without any distinction between dispersal and round-trip movements. In this review, we chose to use the restrictive definition of **migration** with a distinction between dispersal, gene flow and migration.

Genetic drift

Allele frequencies can change due to chance alone. This is called genetic drift where the drift is due to stochastic sampling variability of the gene pool. Otherwise, the alleles that form the next generation's gene pool are a sample of the alleles from the current generation. When sampled from a population, the frequency of alleles differs slightly due to chance alone. The intensity of the genetic drift is a function of population size. In very small populations, genetic drift can lead to a high loss of diversity and to a genetic bottleneck. It can also lead to the apparition of a genetic differentiation between two populations separated in space, by a different trajectory of their allele frequencies. The impact of genetic drift is modulated by the sex ratio and the variance in reproductive success. Moreover, genetic drift can modulate the effects of migration and natural selection (Hamilton, 2009).

3.2 Linking genetics and evolutionary biology

Lamarck published a theory of evolution in 1809 (Lamarck, 1809), he thought that species arose continually from nonliving sources. These species were initially very primitive, but increased

in complexity over time due to some inherent tendency. Lamarck proposed that an organism's acclimatization to the environment could be passed on to descendants by inheritance of acquired characters. Fifty years later, Darwin's contributions include hypothesizing the pattern of common descent and proposing a mechanism for evolution – natural selection (Darwin, 1859). In Darwin's theory of natural selection, new variants arise continually within populations. A small percentage of these variants cause their bearers to produce more offspring than others. Darwin's theory did not accord with older theories of genetics. In Darwin's time, biologists subscribed to the theory of blending inheritance – an offspring was an average of its parents. We now know that the idea of blending inheritance is wrong. At the same time, Gregor Mendel, in his experiments on hybrid peas, showed that genes from a mother and father do not blend (Mendel, 1866). An offspring from a short and a tall parent may be medium sized; but it carries genes for shortness and tallness. The genes remain distinct and can be passed separately to descendants.

BOX 2. Mendel's laws

Between 1856 and 1863, the monk Gregor Mendel carried out experiments with pea plants that demonstrated the concept of particulate inheritance. He showed that phenotypes are determined by units that are inherited intact and unchanged through generations. Mendel used pea seed coat color as a phenotype easily tracked across generations.By establishing yellow and green "pure"-breeding lines of peas and by using them as parents, he crossed green-yellow "impure" lines. This work allowed him to reason on dominant traits (yellow seed coat here) and recessive trait (green seed coat here) in impure or heterozygous individuals.

His work is now well known for his two laws:

- Mendel's first law (law of segregation): Two members of a gene pair ($i.e$: the alleles) segregate separately into gametes, so the segregation of alleles at a single locus is independent.
- Mendel's second law (law of independent assortment): During gamete formation, the segregation of alleles of one gene is independent from the segregation of alleles of another gene.

Between 1900 and about 1925, Darwin's theory was regarded as outdated and replaced by Mendelian genetics - the two approaches were regarded as incompatible rivals. In 1908, Hardy (Hardy, 1908) and Weinberg (Weinberg, 1908) worked independently to formulate a relationship to predict allele frequencies given genotype frequencies. This relationship is well known as the Hardy-Weinberg equation: $p^2 + 2pq + q^2 = 1$ where p and q are allele frequencies for a genetic

locus with two alleles (Smith and Smith, 2001; Hamilton, 2009; Beaumont et al., 2010). A single generation of reproduction where a set of conditions are met will result in a population that meets the **Hardy-Weinberg** expected genotype frequencies (Ryman and Utter, 1987). The list of conditions is long and includes:

- Diploid organism
- Sexual reproduction
- **Non overlapping generations**
- Allele frequencies identical among both sexes
- Random mating
- Random union between gametes
- Infinite population size
- Negligible gene flow
- Negligible mutation
- Absence of natural selection

These conditions make sense when examined. For example, if natural selection acts within a single generation some genotypes will be more frequent than others, breaking the Hardy-Weinberg (HW) equilibrium. However, it's legitimate to wonder if a model with so many conditions is relevant and if all the assumptions are likely to be met in real life. The Hardy-Weinberg model was not meant to be an exact description of current populations. We must see it as a null model to which compare current populations (Hamilton, 2009). However, the majority of empirical studies on natural populations demonstrate equilibrium conditions even though population size is unlikely to be be infinite (Waples, 2015).

A very large population size, effectively infinite, is one of the HW conditions. Population size has profound effects on allele frequencies in biological populations and we accept that all biological populations are finite. Therefore, no current population ever exactly meets the population size condition of HW. Genotype and allele frequencies fluctuate from one generation to the next due to **genetic drift**. After some time, a small population will become homozygous for some alleles, and other alleles will be lost. Over time, allele frequencies spread out progressively as the proportion of fixed genes increases (see **fixation**), ultimately resulting in homozygous populations. The amount of genetic drift increases as the number of the individuals used to produce the next generation decreases (Ryman and Utter, 1987; Smith and Smith, 2001; Hamilton, 2009; Beaumont et al., 2010). A way to restate the population size condition of HW is to say instead that there is very little or no genetic drift occurring. This brings us to the **Wright-Fisher model** which introduced a schematic of the biological life cycle, with an infinite number of gametes in a finite population. The Wright-Fisher model is not biologically realistic but it allows the process of genetic drift to be modeled in a simple fashion (Hamilton, 2009). It makes assumptions identical to the conditions underlying the Hardy-Weinberg equation in addition to assuming that the **sex ratio** is equal and that each generation is founded by sampling $2N$ gametes from an infinite pool of gametes (Smith

and Smith, 2001). Here, we have $2N$ gametes which refers to the diploid state of the organism.

Several applications of the HW equilibrium exist. It is used to predict the expected frequency of a DNA profile, this application can often be seen in newspapers about crime scene and suspect identification. A second common application is the test of deviation from the HW equilibrium considered as null model. If a population has genotype frequencies which do not fit HW expectation, this is considered evidence that one or more of the evolutionary processes considered absent in the HW model are active (Holsinger, 2001; Hamilton, 2009).

Finally, many of the concepts of population genetics have been combined with ideas from population ecology to make up the field of evolutionary biology or evolutionary ecology.

3.3 Molecular markers for measuring genetic variation

Genetic variation can be measured and quantified at several levels: differences between sequences of DNA fragments, differences between proteins resulting from DNA coding sequence variation, etc... The first markers of genetic variation were phenotypic even though the DNA structure was discovered in 1953. Variations at protein levels were discovered during the 1960s and mitochondrial DNA was developed during the 1980s. Finally, the main emergence of the DNA molecular markers was after the 1990s and the introduction of the PCR technique (for Polymerase Chain Reaction) which allowed cheap and rapid amplification of DNA fragments (Hartl and Jones, 1998; Beaumont et al., 2010). Animal cells contain two types of DNA molecules with distinct characteristics: mitochondrial DNA and nuclear DNA. In review, we only explicit two types of nuclear **genetic markers**: the **microsatellites** and the **Single Nucleotide Polymorphisms (SNPs)**.

Microsatellites

Microsatellite loci contain repeated motifs of two to five bases and are scattered throughout the genomes of most eukaryotes Fig. 2, the number of alleles represents all the possible variants. The number of repeated units contained within a particular microsatellite can vary within a population, producing variation in the length of the locus. They are co-dominant, assumed to be neutral and hypervariable with a fast **mutation rate**.

GT pattern repeated 5 times: A A T G C A C G T G T G T G T G T T T C A T

GT pattern repeated 3 times: A A T G C A C G T G T G T T T C A T

Figure 2: Example of microsatellite sequences. The first sequence presents two alleles G and T repeated 5 times and the second one presents the same allelic pattern repeated 3 times.

Microsatellites are powerful markers for which two decades of experience have established the

advantages and limits. Indeed, the assumption of neutrality is coming into question and they are limited by the need of a mutation model (Chevolot, 2006). The mutation rate of microsatellites is around 10^{-4} (Brumfield et al., 2003).

Single Nucleotide Polymorphisms

SNPs are the most common form of genetic variation and their occurence throughout the entire genome makes them ideal for studying the inheritance of genomic regions (Baird et al., 2008). A SNP is a variation at a single nucleotide position in the genome between the maternal haplotype and the paternal haplotype, Fig.3, which occurs in at least 1% of the individuals within a species (Vignal et al., 2002). SNPs occur every 1000 bases or so in the human genome but can be much more frequent in many marine species. Their occurrence throughout the genome also makes them ideal for analyses of speciation and historical demography, especially in light of recent theory suggesting that many unlinked nuclear loci (meaning of unlinked loci is explained later) are needed to estimate population genetic parameters with statistical confidence (Brumfield et al., 2003; Beaumont et al., 2010). SNPs have relatively low mutation rates (10^{-8} to 10^{-9}) (Nielsen, 2000; Brumfield et al., 2003).

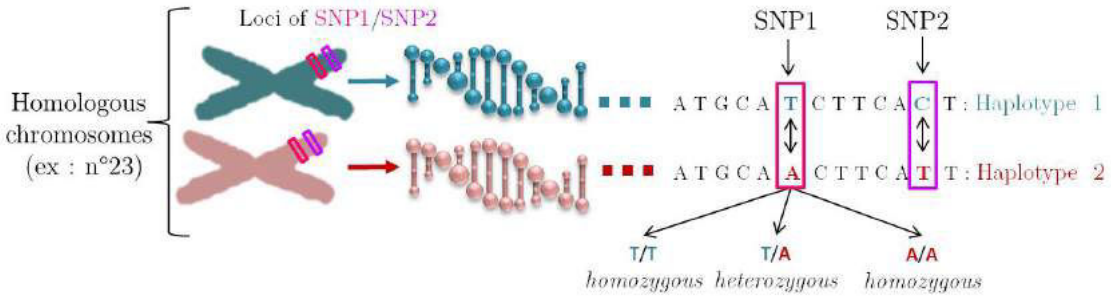


Figure 3: Example of two Single Nucleotide Polymorphisms (SNPs) in a **diploid** organism, heterozygous for the two SNPs are presented. For visibility only, chromosoms are represented at mitosis.

Two types of SNP are distinguished: transition, which is a change between two purine or two pyrimidine bases and transversion, which is a change between a purine and a pyrimidine base. In theory, one SNP can show up to the four allels but in most cases, a SNP is diallelic at each locus (Vignal et al., 2002; Beaumont et al., 2010). Two types of SNP data are available (Morin et al., 2009) phased data, if haplotypes are considered and unphased data if genotypes are considered. For haplotypic data the linkage phase is known, while for genotypic data the linkage phase is unknown (Schaid, 2004). In the Fig. 4, we have the example of two SNPs: the first one presents the possible allelic states T and A and the second one, the possible allelic states C and A. In the populations, there is numerous haplotypes possible and only four examples are listed in the Fig. 4.

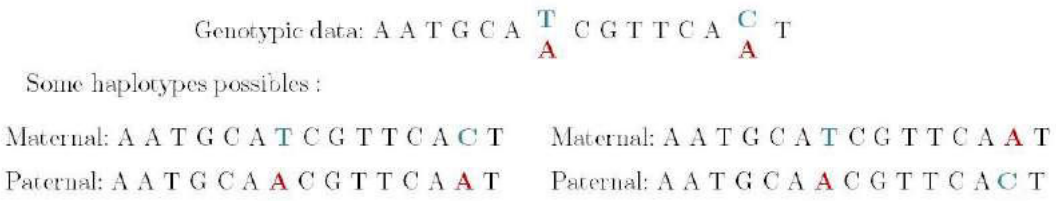


Figure 4: Example of haplotypic corresponding to genotypic SNP data. Here, we consider 2 SNPs which possess each two possible allelic states: T or A and C or A.

The haplotype can be inferred from the genotype in several ways :

- Using biochemistry methods
- Using family information : if the parental genotypes are known, the genotypes of progeny can generally be deduced.
- Using statistical tools and algorithms of genotypes such as implemented in the programm ARLEQUIN (Excoffier et al., 2005).

Several approaches have been developed to infer haplotypes from genotypes with statistical tools. The most known approach is by Clark (1990) which reduces the number of possible haplotypes with a parsimony approaches.

SNPs are simple (bi-allelic) and easy to model which make them powerful contributors to infer population histories. They also present low mutation rates and low scoring error rates. When all population genetic and analytical considerations are weighed, SNPs are superior to microsatellites (Tab. 1).

Finally, evolutionary genetics (which includes population genetics) is a central discipline in the study of evolutionary processes. It uses both molecular and classic genetic methods to understand the origin of variation. It describes patterns of genetic variation within and among populations and species, and employs both empirical studies and mathematical theories to discover how this variation is affected by processes such as genetic drift, gene flow, and natural selection.

Molecular markers	Pros	Cons
SNPs	– Low mutation rates – Bi-allelic – Variation easy to interpret – Deviation tests from neutrality available – Low scoring error rates	– High number of SNPs required – High **Ascertainment bias** – Challenging for computational treatment
Microsatellites	– Few microsatellites needed for population genetics – Display large allelic diversity – Deviation tests from neutrality available	– High mutation rates – Mutation rates variables across loci and across alleles within the same locus – Conducive to **Ascertainment bias**

Table 1: Pros and cons of two types of DNA markers for use in population genetics (Brumfield et al., 2003; Helyar et al., 2011)

3.4 Population structure and gene flow

The expectation that genotypes will be present at Hardy-Weinberg equilibrium frequencies depends on the assumption of random mating. Several processes in actual population make this assumption unlikely to hold for many populations (Waples, 2015). For example, within large populations the chances of mating are not uniform but depend on the location of the two mates within the population. This leads to what is called **population structure**. This phenomenon has profound implications for genotype and allele frequencies. Subdivision breaks up a population into smaller units that are genetically independent to some degree. One consequence is that each **subpopulation** has a smaller **effective population size** than the effective size of the entire population if there was random mating. Processes that cause population structure can be considered both as creative or constraining in evolutionary changes (Slatkin, 1987). Genetic isolation, for example, can prevent novel alleles from spreading but can also maintain unique alleles as required for genetic adaptation to local environments. As explained in above, a distinction is made between **migration**, **dispersal** and **gene flow**. As dispersal is simply the movement of individuals from one place to another, it may or may not result in gene flow. To confuse matters further, models do not make the distinction and the variable m (for migration rate) is almost universally used to indicate the rate of gene flow.

3.4.1 Studying population structure

In any population study, the ideal first step would be to collect samples of the species across its entire range to estimate genetic differentiation within the species as a whole. However, for economic or sampling constraints, most studies focus on limited sampling in specific areas with an economic or conservation interest. Depending on the markers employed, genotypes or haplotypes are scored for the individuals sampled and the data are analyzed in a variety of ways to quantify genetic variation between populations (Beaumont et al., 2010). There is no universal rule for the minimum number of individuals to sample per location. Waples and Do (2008) proposed that 50 individuals appears to be a good trade-off between sampling cost and the biais in estimating population structure.

Several indicators of genetic differentiation exist. One of the most applied indicators is the F-statistics, which was developed by Wright (1949, 1950).

3.4.2 Wright's F-statistics

Before introducing population structure measures, we need to define a key parameter in population genetic : the **fixation index** (Hamilton, 2009), sometimes called the **inbreeding** coefficient (Smith and Smith, 2001). A quantity, symbolized F, is commonly used to compare how much **heterozygosity** is present in an actual population relative to expected levels of heterozygosity under random mating (and other HW equilibrium conditions):

$$F = \frac{H_e - H_0}{H_e} \tag{1}$$

where H_e is the HW expected frequency of heterozygotes based on the population allele frequencies and H_0 is the observed frequency of heterozygotes.

Dividing by the expected heterozygosity puts F on a convenient scale of -1 and $+1$. Negative values indicate an excess of heterozygotes and positive values indicate an excess of homozygotes.

Accounting for divergence of sub-populations, studying the genetic differentiation among populations necessitates several new versions of the fixation index, the so called **F-statistics** (Wright, 1950). The concept of F-statistics was developed by Sewall Wright during the 1920s but the three parameters as we know them now were proposed in the 1950s (Wright, 1949, 1950). These indicators were designed to describe the population genetic structure of diploid organisms. Basic assumptions are that all populations are of the same size and that there are equal possibilities for any population to exchange individuals with any other population (Beaumont et al., 2010).

Firstly, heterozygosity is calculated for each diallelic loci and then averaged according to the scale considered (total population, subpopulation) (Beaumont et al., 2010). So, a series of

hierarchical measures of heterozygosity were defined:

- H_I : mean observed heterozygosity across subpopulations
- H_S : mean expected heterozygosity across subpopulations with random mating within each subpopulation
- H_T : expected heterozygosity with random mating within total population

where subscript T indicates the total population, S the subpopulation and I the individual level (Wright, 1965). Considering a diallelic loci, with the two alleles at equal frequencies, H_T and H_S have maximum values of 0.5 and H_I can vary between 0 (no observed heterozygotes) and 1 (all observed individuals are heterozygote).

Now, based on H_I, H_S and H_T, three hierarchical F-statistics are defined : F_{IS}, F_{ST}, F_{IT}.

F_{IS} : inbreeding coefficient.

$$F_{IS} = \frac{H_S - H_I}{H_S} \qquad -1 \leq F_{IS} \leq 1 \tag{2}$$

The F_{IS} coefficient represents the difference between the average observed and the HW expected heterozygosity due to non random mating, so it is a measure of the extent of genetic inbreeding within subpopulations. At the lower supbpopulation level, it is the correlation between homologous alleles within individuals with reference to the local population (Wright, 1949, 1950, 1965; Beaumont et al., 2010). In other words, it is the correlation between homologous alleles within individuals with reference to the local population (ie. the subpopulation under study). A F_{IS} value close to -1 means that all individuals are heterozygous or that there is an excess of heterozygotes compared to Hardy Weinberg expectations, 0 means that the subpopulations meet the HW assumptions and a value close to $+1$ means that there are no observed heterozygotes (Beaumont et al., 2010).

F_{ST} : fixation index.

$$F_{ST} = \frac{H_T - H_S}{H_T} \qquad 0 \leq F_{ST} \leq 1 \tag{3}$$

The F_{ST} coefficient represents the difference between the average expected heterozygosity of subpopulations and the expected heterozygosity of the total population, so it measures the reduction in heterozygosity due to subpopulation divergence in allele frequency (Lowe et al., 2004; Hamilton, 2009). At a lower level, it is the probability that two alleles sampled at random from a single subpopulation are identical by descent (Smith and Smith, 2001). A F_{ST} value close to 0 means there is no differentiation between subpopulations and a value close to $+1$ that there is complete differentiation between subpopulations.

Although F_{ST} has a theoretical range between 0 and 1, the observed maximum is usually much less than 1. Wright suggested the following qualitative guidelines for the interpretation of F_{ST} (Wright, 1984) :

- the range 0 to 0.05 may be considering as indicating little genetic differentiation
- the range 0.05 to 0.15 may be considering as indicating moderate genetic differentiation
- the range 0.15 to 0.25 may be considering as indicating great genetic differentiation
- values above 0.25 indicate very great genetic differentiation

F_{IT} : overall fixation index.

$$F_{IT} = \frac{H_T - H_I}{H_T} \qquad -1 \leq F_{IT} \leq 1 \tag{4}$$

The F_{IT} coefficient is the correlation between homologous alleles within individuals with reference to the total population (Beaumont et al., 2010). It describes the reduction of heterozygosity within individuals relative to the total population due to non-random mating within subpopulations and population subdivisions (Lowe et al., 2004).

The three F-statistics are not independent but interrelated according to the formula:

$$F_{ST} = \frac{(F_{IT} - F_{IS})}{(1 - F_{IS})} \tag{5}$$

Finally, Wright's F-statistics provide answers to two different questions:

- for the scored loci, are the genotypes in the proportions predicted by the HW model? (F_{IS} and F_{IT} provide answers)
- for the scored loci, are the allele frequencies different between various populations ? (F_{ST} provides answers)

The original formulation of F_{ST} by Wright considered only one biallelic locus. This was extended to accommodate multiple loci termed G_{ST} (Nei, 1973). θ or θ_{ST} (Weir and Cockerham, 1984; Weir, 1996) or ϕ_{ST} (Excoffier et al., 1992) are estimators based on analysis of variance of allele frequencies within and among sub-populations, etc... One of the biggest weaknesses of Wright's F-statistics is the ignorance of the bias due to the number of sampled individuals. The θ estimator by Weir and Cockerham (1984) takes this bias into account; it is often used as an alternative to the F-statistics.

BOX 3. Example of F-statistics calculation

The data shown in the following table are based upon three loci surveyed in three populations (Lowe et al., 2004). Allele frequencies have been calculated assuming HW equilibrium ; all means are arithmetic and it is assumed that sample sizes are equal.

For example, for the first locus :
$H_T = (0.4 + 0.4 + 0)/3 = 0.2667$
$H_S = (0.5 + 0.48 + 0)/3 = 0.33$
$H_T = 2 * 0.7 * 0.3$

	Phenotype frequency			Allele frequency		Expected number of heterozygotes in random mating total population
	a/a	a/b	b/b	a	b	
Locus 1						
Population 1	0.3	0.4	0.3	0.5	0.5	0.5
Population 2	0.4	0.4	0.2	0.6	0.4	0.48
Population 3	1.0	0	0	1.0	0	0
H_I		0.2667				
Mean population allele frequency				0.7	0.3	H_S=0.33
H_T				0.42		

Locus 2						
Population 1	0.3	0.1	0.6	0.35	0.65	0.455
Population 2	0.25	0.5	0.25	0.5	0.5	0.5
Population 3	0.65	0.2	0.15	0.75	0.25	0.375
H_I		0.2667				
Mean population allele frequency				0.53	0.47	H_S=0.44
H_T				0.4982		

Locus 3						
Population 1	1.0	0	0	1.0	0	0
Population 2	1.0	0	0	1.0	0	0
Population 3	1.0	0	0	1.0	0	0
H_I		0				
Mean population allele frequency				1.0	0	H_S=0
H_T				0		

BOX 3. Example of F-statistics calculation - cont.

So the averages are the following :

$H_T = (0.42 + 0.4982 + 0)/3 = 0.31$

$H_S = (0.33 + 0.44 + 0)/3 = 0.26$

$H_I = (0.2667 + 0.2667 + 0)/3 = 0.18$

Wright's F-statistics can be derived as :

$F_{IT} = \frac{H_T - H_I}{H_T} = \frac{0.31-0.18}{0.31} = 0.42$

$F_{ST} = \frac{H_T - H_S}{H_T} = \frac{0.31-0.26}{0.31} = 0.16$

$F_{IS} = \frac{H_S - H_I}{H_S} = \frac{0.26-0.18}{0.26} = 0.31$

Positive F_{IS} and F_{IT} indicate a deficit of heterozygotes with respect to HW expectation. F_{ST} around 0.16 means that 16% of the total genetic variation is between subpopulations, with 84% of the variations within subpopulations. It indicates a great genetic variation between subpopulations.

4. Genetic population size

The concept of effective population size appeared for the first time in 1931 proposed by the geneticist Sewall Wright (Wright, 1931). The definition of population size in population genetics relies on the dynamics of genetic variation in the population. It means that the size of a population is defined by the way genetic variation in the population behaves (Hamilton, 2009; Beaumont et al., 2010; Hare et al., 2011). There are two types of population sizes. One is the real count of individuals in a population (including immatures), called the **census size** N_c. The other one is the genetic size of the population, determined by comparing the genetic drift in a studied real population with the genetic drift in an ideal population, i.e. meeting the Wright–Fisher model which assumes Hardy-Weinberg conditions plus infinite number of gametes and equal sex ratio. The population size of an ideal population that produces the same rate of genetic drift as observed in the current population is the genetic size of the current population or its effective size (Hamilton, 2009). Thus the effective size, denoted N_e, is the size of an ideal population that experiences as much genetic drift as an actual population (regardless of its census size) (Lowe et al., 2004). This comparison with a theoretical ideal population standardizes the measurement of genetic drift and makes $N_e$comparable across populations with very different life histories (Hare et al., 2011). To summarize and give an ecological interpretation, the census size is the total number of individuals and the effective size is the number of breeding adults that currently contribute to the next generation (Hamilton, 2009). The effective size excludes juveniles, individuals too old to reproduce and those that provide non-contributory gametes. If N_e is very large, the variance in

allele frequencies between successive generations is very small (Beaumont et al., 2010). In an ideal population, N_e is equal to the census population size of a generation (Hamilton, 2009).

One of the major difficulties in discussing N_e is that there are two commonly estimated measures of N_e. Variance effective size is the size of an ideal population experiencing genetic drift at the same rate than the actual population and inbreeding effective size is the size of an ideal population losing heterozygosity due to increased relatedness. Typically both are not discussed in detail because for large, stable populations, they are similar. However, the sizes of many wildlife populations are not stable (Leberg, 2005).

4.1 Factors influencing effective population size

Population fluctuations, like bottleneck events have a high impact on effective size but there are several other aspects of biological populations (Tab.2) that have the same impact by increasing the sampling error in allele frequency across generations (Lowe et al., 2004).

Population size fluctuations can lead to changes in alleles represented in later generations which lead to changes in effective size.

An unequal sex ratio has a high influence on genetic drift and so on the population effective size of a population. For example, in a polygamous population with a ratio of one breeding male to several females, the offspring will be half or full **siblings**, leading to an increase in genetic drift compared to a case with breeding sex ratio of 1:1 where all offspring are less closely related (Smith and Smith, 2001; Hamilton, 2009).

A third factor is the degree to which adult individuals in the population contribute to the next generation. A population is stable in size over time when each pair of individuals produces on average two progeny. The variance in family size can be used to describe variation among individual reproduction. If the variance in family size increases, the alleles passed to the next generation come increasingly from those parents producing more offspring. When the variance in family size is equal to the average family size ($k = 2$), then the population size of parents is the effective population size. The Wright-Fisher model assumes that family sizes follow a Poisson distribution (i.e. mean equal to variance). An interesting fact is that in the case where the variance in family size is less than the mean family size, N_e can be larger than N (Hamilton, 2009).

A finite population size can be seen as a form of inbreeding. In small populations, the chance of mating with a relative is large since the number of mates is limited. Genetic drift also occurs due to finite population size. Both phenomenon increase homozygosity and decrease heterozygosity leading to a decrease in effective population size (Lowe et al., 2004; Hamilton, 2009).

Migration and dispersal are also phenomenons acting on the effective population size. If immigrants become part of the breeding population, they introduce a different genetic sample that tends to reduce genetic drift. Information from the field on genetic effects of migration is lacking

because of the difficulty to collect data (Smith and Smith, 2001).

Finally, overlapping generations have a great influence on the effective size of a population because age-structured species presents many co-evolutionary process (Waples et al., 2014). However, the exact effects of overlapping generations remain unclear.

Event	Consequences for genetic diversity
Population size fluctuations	– Changes alleles represented in later generations – Changes effective size
Unequal sex ratio	– Increases genetic drift – Reduces N_e (max if breeding sex ratio 1:1)
Variance in reproductive success	– Stable population : parents produce 2 offspring ($k = 2$) – If $k = 2$: population size of parents is the effective size
Inbreeding	– Increases homozygosity – Decreases population effective size
Migration	– Increases or decreases genetic drift – Increases or decreases population effective size
Overlapping generations	– Effects remain unclear

Table 2: Summary of effects of different factors influencing genetic drift and population effective size (Smith and Smith, 2001; Hamilton, 2009)

All factors contributing to genetic drift are confounded in the variable N_e, so a small value indicates strong drift but does not identify the causes. Estimating N_e is valuable for harvested population in a stock assessment context but can also provide insights into patterns of connectivity among populations.

4.2 Relationship between N_e and N_c

The first thing to understand is that, for most real populations, effective size and census size are not the same (Leberg, 2005). Commonly, due to genetic bottlenecks for example, the census size is larger than the effective size (Lowe et al., 2004). Many factors can cause $N_e \neq N_c$ including unequal sex ratios, high variation in reproductive success, non random mating, mating system, philopatry, gene flow, overlapping generations and temporal fluctuations in population size (Leberg, 2005).

BOX 4. Commonly N_e differs from N_c: example

Assume a population of 100 individuals at time $t-1$, which was reduced to 10 individuals at time t (by a **bottleneck** for example) and increased again to 100 individuals at time $t+1$ (each couple produced offspring)(Hamilton, 2009). The effective size of a population fluctuating over time is calculated as the reciprocal of the average of the sum of reciprocals of the effective number of each generation (Smith and Smith, 2001). This special sort of average is called the harmonic mean of N_e and is given by the following expression:

$$\frac{1}{N_e} = \frac{1}{t}\left[\frac{1}{N_{e(t=1)}} + \frac{1}{N_{e(t=2)}} + ... + \frac{1}{N_{e(t)}})\right]$$

In our example, N_e is equal to 25 and the mean census size is 70. Only those alleles that pass through the genetic bottleneck of 10 individuals are represented in later generations. Using the census size to predict the behavior of allele frequencies will underestimate the genetic drift. We expect allele frequencies in the current population to behave like the allele frequencies in an ideal Wright-Fisher population with a constant size of 25 individuals over three generations (Hamilton, 2009).

The relationship between N_e and N_c is a key to understanding the effects of harvesting: in marine species, N_e/N_c ratios are often as low as 10^{-5}. This very low ratio seems to indicate that enormous marine population may be more sensitive to genetic drift and inbreeding from intensive harvest than census size alone suggests (Hare et al., 2011). However these low ratios can also be a consequences to insufficient sample sizes (Waples, 2016).

Effective population statistics inform us about the genetic health of a population. To be able to handle the loss of genetic variability, inbreeding, a population must reach a critical size. The number of individuals that will ensure the persistence of a population in a viable state for a given interval of time is the **Minimum Viable Population (MVP)** (Soulé and Wilcox, 1980; Shaffer, 1981; Smith and Smith, 2001; Rai, 2006). It must be large enough to cope with environmental

changes, genetic drift or variation in individual birth and death. Different studies have been conducted to estimate threshold values for effective size. In 1980, M. Soulé addressed the question of MVP size in order to prevent extinction and concluded a minimum effective size of 50 individuals was needed (with a maximum of inbreeding at 1%). The same year, Franklin suggested that in the long term, genetic variability will be maintained only if population sizes are an order of magnitude higher than 50. This suggestion is based on the assumption that continued evolutionary change is necessary for populations and species survival, and that response to natural selection is limited by small population sizes. Due to this, Franklin proposed that for long term viability the effective size should be 500 in order to account for the expected rapid loss of genetic variance (Franklin, 1980). The two studies gave rise to the 50/500 rule which has been strongly criticized (Caughley, 1994; Lande, 1995).

Given loss of genetic variability can been expressed as the loss in heterozygosity, Meffe and Carroll (1997) argued that an effective size of 50 individuals was not enough for long term conservation (Fig. 5).

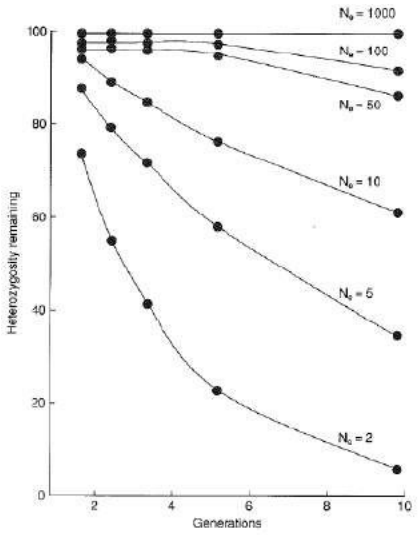


Figure 5: Simulated loss of genetic variability after 10 generations according to the effective size Ne. Populations with large Ne maintain their genetic variability while small ones loose it. (Meffe and Carroll, 1997)

Experiences with drosophila showed that a median extinction time of 47.5 generations was necessary to extinct a population of 50 effective individuals (Reed and Bryant, 2000). Such experiences have never been realized on rajids and it is hard to predict a minimal effective size for a given minimal time leading to the extinction of a population. Thus, MVP is only a general guideline for the genetic management of endangered species. Past studies showed that applying the 50/500 rule on wild populations can be a dangerous game for species survival. For example, a species with highly density-dependent reproduction, living in a more or less constant environment

may persist for a long time despite a decline in genetic diversity. There is no universal rule about MVP (Smith and Smith, 2001).

4.3 Estimators of effective population size

Genetic methods for estimating N_e have been increasingly practicable by recent advances in genotyping, software and computer processing speed (Hare et al., 2011). Just as there are several models based on different assumptions to describe how genetic variation over time, there are several ways to estimate the effective population size. Two main definitions of the effective population size exist (Hamilton, 2009):

- inbreeding effective population size which represents the size of an ideal population that would show the same probability of allele copies being identical by descent as an actual population.
- variance effective population size which represents the size of an ideal population that would show the same sampling variance in allele frequency as an actual population.

For the conservation and the management of wildlife populations, estimates of contemporary N_e are of interest. When estimates are based on single-cohort samples, they reflect the effective number of breeders in one reproductive cycle (N_b) (Dudgeon and Ovenden, 2015). For mixed-age sample, N_e estimates reflect the harmonic mean of the number of breeders over several generations (Waples et al., 2014).

Several estimates of N_e exist which can be summarized in two broad categories of estimates (Dudgeon et al., 2012) :

- Estimate of long-term or historical effective population size.
- Estimate of the contemporary or short term effective population size.

These two estimates of effective population sizes can be derived from samples collected at a single point in time or from temporally spaced samples. In receent years, several estimators based on single-sample method were developped (Waples et al., 2014).

4.3.1 Long-term N_e estimation methods

Estimates of long-term N_e are used if the goal is to understand the long-term N_e of a species prior to contemporary influences. The main method used is the coalescent method which requires only a single random sample of individuals (Hare et al., 2011). This method traces the history of genes in a population back to a common ancestor in order to describe the process of coalescence of allelic copies in the population back in time (Nikolic, 2009). Such models investigate patterns

among individuals from the past to the present and aim at reconstructing versions of events such as inbreeding, gene flow or natural selection in the past that could have lead to the current observed population. A coalescent event is defined as the merging back in time of two lineages into a single ancestral lineage. The method aims at connecting current lineages observed in the sample back to a single ancestor in the past, the most recent common ancestor (MRCA) (Hamilton, 2009). The coalescent model leads to definitions of the inbreeding effective population size that are identical to those obtain using autozigosity (Smith and Smith, 2001). For most management applications, this method targets the period just prior to human intervention (Alter et al., 2007).

Historical effective population size has been widely estimated for terrestrial species but also to elasmobranchs species such as *Raja typus* (Castro et al., 2007), using *ARLEQUIN* software (Excoffier et al., 2005), *Negaprion brevistoris* (Schultz et al., 2008), *Sphyrna lewini* (Nance et al., 2011), using the method developed by Beaumont (Beaumont, 1999) and implemented in the software *MSVAR* and sleeper sharks (Murray et al., 2008).

4.3.2 Contemporary N_e estimation methods

Contemporary N_e can be estimated using demographic methods or genetic method, here only the genetic methods are detailed. Short-term N_e estimation are usually based on allelic changes between two samples like loss of heterozygosity, loss of alleles or the increase of the inbreeding coefficient. Commonly, four methods are considered suitable for estimating N_e indirectly from genetic data. All four methods have limited power to estimate N_e when the true value is large, but since conservation is focused on small populations, these methods are increasingly used in the conservation context.

Multiple sample estimation

The temporal method requires population samples separated by at least two generation times (Dudgeon et al., 2012). The effective population size is estimated by relating the observed amount of allele frequencies change to that expected under pure genetic drift. This method estimates the variance effective population size using the expected variance in gene frequencies (Grant and Waples, 2000). Several factors may introduce bias into these estimates : sampling error, migration, natural selection, etc... Some developments made during the last decade address these aspects to minimize these bias. For example, several formulas are used for calculating N_e depending upon of the sampling scheme: random sampling of the entire population, only reproductive adults, single-cohort of newborns... (Waples and Yokota, 2006). Moreover, Jorde and Ryman (1995) developed a correction factor based on life history traits for accounting for overlapping generations.

Precision and accuracy of N_e estimates can be improved by increasing the number of loci, the number of alleles, the sample size or the amount of time between temporal samples. However,

the last option is the most difficult as it requires conducting long-term studies or that old samples are available. As a consequence, it is rarely applied (Leberg, 2005).

Single-sample estimation

Estimations of N_e using the single-time point estimate approach requires only a single sampling of the population. In the last decade, various estimation methods have been developped or improved (Dudgeon et al., 2012).

- **Linkage-Disequilibrium** (LD)
- **Heterozygosity Excess** (HE)
- Molecular Co-ancestry (MC)

4.4 Estimation of effective population with Linkage-Disequilibrium

4.4.1 Definition and estimation of the Linkage-Disequilibrium

In population genetics, Linkage-Disequilibrium (LD) is the non random association of alleles at different loci *i.e* the presence of statistical associations between alleles at different loci that are different from what would be expected if alleles were independent, based on their individual allele frequencies. When the genotype present at one locus is independent of the genotype present at another locus, they is a Linkage-Equilibrium.

The assumptions for using LD are the same than the Wright Fisher model, but also assume that loci are unlinked. Genes located close to each other on the same chromosome are said linked (Fig. 6, a). If two alleles from different genes on the same chromosome tend to be associated in different individuals at a greater frequency that expected due to random association, there is linkage disequilibrium between these genes. Two genes located on different chromosomes or at large distance on the same chromosome are unlinked (there is Linkage-Equilibrium).

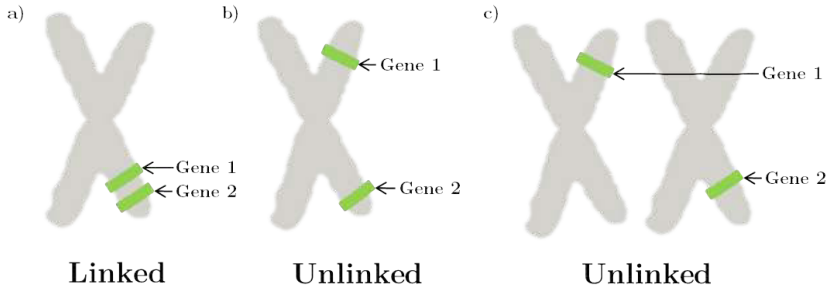


Figure 6: Chromosomes illustrating linkage between loci. Two genes are shown in each situation. As illustrated genes 1 and 2 are linked in the first situation and unlinked in the others.

When numerous loci are used, it is usually unavoidable that some will be linked. However, linkage becomes an issues only if genes are very close together or if the recombinaison rate is very low. The recombinaison rate can be seen as the proportion of gametes produced by an individual with haplotype differing from the two haplotypes of the individual. For example, an individual A1B1/A2B2 will produce 2 types of gametes: parental type (A1B1 and A2B2) and recombinant type (A1B2 and A2B1). r is the proportion of recombinant type and is bounded between 0.5 (genes are unlinked or very far from each other on the chromosome, Fig. 6 b and c) and 0 (genes are linked Fig. 6, a).

Linkage-Disequilibrium is measured by two statistics, D and r, which can be interpreted as the co-variance and the correlation between loci. These statistics are explicited in subsequent parts. The effective population size estimated by the LD method is a variance effective population size.

4.4.2 Origins of the Linkage-Disequilibrium

Several factors can create Linkage-Disequilibrium.

- **Genetic drift** : genetic drift conducts to a loss of variability over generations (by random disappearance of alleles or haplotype). genetic drift is much stronger when the population size is small.
- **Mutation** : mutation can occur in a haplotype and create a Linkage-Disequilibrium between the mutated locus and this haplotype. This disequilibrium normaly reduces over generations but can increase if genetic drift or selection occur.
- **Gene flow** : LD can also be created by the mix between populations or by migration. At the beginning, LD is proportional to the allelic differences between populations and is independent of the distance between markers. Over generations, LD decreases but the rate of decrease varies with the link between loci. For independents loci, LD tends to disappear and for linked loci, it persists much longer.
- **Selection** : as genetic drift, selection conducts to a decrease in the number of breeders and so a decrease in the number of haplotypes in the population. Alongside, this decrease leads to an increase of consanguinity.
- **Recombination** : the recombination rate is not constant along the genome. LD is strong in regions with an low recombination rate and high in not recombinate regions.

4.4.3 Assessing the population effective size with the Linkage-Disequilibrium method

LD can be illustrates by considering 2 SNPs markers A and B on the same chromosome. A possess the alleles A_1,A_2 and B the alleles B_1,B_2. Four haplotypes are possibles : A_1B_1, A_1B_2, A_2B_1 and A_2B_2 (Fig. 9).

Figure 7: Illustration of the four haplotypes possible by considering a heterozygous individual at two SNPs markers A and B.

If each allele possess a frequency of 0.5 and that the Hardy Weinberg conditions are met, each haplotype should have a frequency of 0.25. Any deviation to this haplotypic frequency reflects a Linkage-Disequilibrium. Several measures of the LD are available between alleles with 2 loci and all depend on the following equation where D is the coefficient of LD, p_{A_1} represents the frequency of the allele A_1 and $p_{A_1B_1}$ the frequency of the haplotype A_1B_1:

$$D = D_{A_1B_1} = p_{A_1B_1} - p_{A_1}p_{B_1} \tag{6}$$

Linkage-Disequilibrium between A_1B_2, A_2B_1 and A_2B_2 can also be calculated. When the genetic markers used are bi-allelic (like SNPs), results between all calculations are exchangeable ($D_{A_1B_1} = -D_{A_1B_2} = -D_{A_2B_1} = D_{A_2B_2}$). This does not apply to multi-allelic markers, such as microsatellites.

D is a non normalized equation. Other statistics have been derived and are more commonly used.

– Hill and Robertson (1968) defined the correlation coefficient (r^2) between alleles for bi-allelic loci.

$$r^2 = \frac{D^2}{p_{A_1}p_{A_2}p_{B_1}p_{B_2}}, \qquad 0 \leq r^2 \leq 1 \tag{7}$$

Under the null hypothesis ($H_0 : D_{ij} = 0$) this statistics follows a χ_2 distribution. Moreover, the correlation coefficient is less sensitive to the population size than other statistics.

– Lewontin (1964) proposed the D' (widely used in medicine):

$$D' = \frac{|D|}{D_{max}}, \qquad D_{max} = \begin{cases} min[pA_1oB_2, pA_2pB_1] & if D > 0 \\ max[pA_1pB_1, pA_2, pB_2] & if D < 0 \end{cases} \tag{8}$$

The r^2 and the D' as presented here are statistics of LD at 2 loci but it is more useful to consider LD on a chromosomal region. One of the classical approaches is to calculate a local average of LD for pairs of loci.

BOX 5. Linkage-Disequilibrium with unphased data

In diploid species, it is much easier to determine genotypes (unphased data) than haplotype (phased data). Consequently, which nucleotides, corresponding to SNP alleles in fig. 7, occur together on individual DNA molecule (chromosome) is usually unknown. This makes it difficult to estimate statistical associations such as LD among loci, as shown below. Let's consider two SNPs on the same pair of homologous chromosomes at two loci. A possess the alleles A_1, A_2 and B, B_1 and B_2. All the possible genotypes and the examplified number of sampled individuals are summarized in the following table:

	B_1B_1	B_1B_2	B_2B_2
A_1A_1	0	1	0
A_1A_2	0	1	0
A_2A_2	0	0	0

Genotypes of sampled individuals are:

Individual 1	A_1A_2	B_1B_2
Individual 2	A_1A_1	B_1B_2

Therefore, there are 4 possible haplotypes for individual 1: A_1B_1, A_1B_2, A_2B_1 and A_2B_2 and :

$$D_{A_1B_1} = fA_1B_1 - fA_1fB_1$$

$$f_{A_1} = 0.75 \quad f_{B_1} = 0.5$$
$$f_{A_2} = 0.25 \quad f_{B_2} = 0.5$$

But fA_1B_1 is not directly available.

As developed in the Box.5, estimates of Linkage-Disequilibrium in case of genotypic data is not a direct calculation. Methods currently used to estimate LD bypass this issue by inferring haplotype frequencies from a maximum likelihood method (Hill, 1974; Rogers and Huff, 2009) or by using the composite Burrows method (Weir, 1979, 1996). The first inference method developed by Hill (1974) requires that random union of gametes occurs in the population, in which case population genotypic frequencies can be replaced by products of gametic frequencies. Such cases are discussed and Weir (1979) recommended the use of the Burrows method which does not require random mating assumption and provides a straightforward calculation. Dr Peter Burrows (unpublished work but see Cockerham and Weir, 1977) considered a composite measure of LD : Δ. It performs better than an alternative maximum likelihood estimator (Weir, 1979; Schaid, 2004) and is estimated directly from genotype counts as follow (Russell and Fewster, 2009) :

$$\hat{\Delta}_{A_1B_1} = \frac{n_{A_1B_1}}{n} - 2\hat{p}_{A_1}\hat{p}_{B_1} \tag{9}$$

where $n_{A_1B_1}$ are the sample counts of possible genotypes (see Box 5.) and n the total number of counts (sampled individuals). $\hat{p}_{A_1}$ and $\hat{p}_{B_1}$ are proportions of alleles in the sample A_1 and B_1 in the n individuals genotyped at both loci. A small-sample correction factor of $n/(n-1)$ should be applied to $\hat{\Delta}_{A_1B_1}$ (Weir, 1979, 1996).

The corresponding correlation coefficient can be estimated as :

$$\hat{r}_{A_1B_1} = \frac{\hat{\Delta}_{A_1B_1}}{\{\hat{p}_{A_1}(1-\hat{p}_{A_1}) + (\hat{h}_{A_1A_1} - \hat{p_{A_1}}^2)\}\{\hat{p}_{B_1}(1-\hat{p}_{B_1}) + (\hat{h}_{B_1B_1} - \hat{p_{B_1}}^2)\}} \tag{10}$$

where $\hat{h}_{A_1A_1}$ and $\hat{h}_{B_1B_1}$ are the observed proportions of A_1A_1 and B_1B_1 homozygotes in the sample of size n .

The correlation coefficient r has $E(r) = 0$ for unlinked loci; in finite populations however, the correlation is likely to take non-zero values with small populations giving the largest values so the expectation of its square is non-zero and is a function of the effective population size, N_e (Russell and Fewster, 2009). The expression for $E(r^2)$ depends upon the mating structure and recombination fraction c in a population, and is also affected by sample size S (Waples, 2006; Russell and Fewster, 2009) The distribution of r^2 is not known but Weir and Hill (1980) showed that the expectation of squared disequilibrium coefficients is strongly affected by the mating system and recombination fraction (c) between the studied loci. For dioecious random mating :

$$E(\hat{r}^2) = \frac{(1-c)^2 + c^2}{2N_e c(2-1)} + \frac{1}{S} \tag{11}$$

If the loci are unlinked (c=0.5), this equation simplify to :

$$E(\hat{r}^2) = \frac{1}{3N_e} + \frac{1}{S} \tag{12}$$

$E(\hat{r}^2)$ can be expressed as the sum of a term due to finite population size, $E(\hat{r}^2_{drift})$, and a term due to sampling a finite number of individuals, $E(\hat{r}^2_{sample})$, (Waples, 2006; Russell and Fewster, 2009).

$$E(\hat{r}^2) = E(\hat{r}^2_{drift}) + E(\hat{r}^2_{sample}) \tag{13}$$

Waples (2006) suggested that if S varies among loci, the harmonic mean should be used. Replacing $E(\hat{r}^2)$ with its estimate, $\hat{r}^2$, and rearranging leads to an estimator for N_e (for dioecious random mating):

$$\hat{N}_e = \frac{1}{3(\hat{r}^2 - 1/S)} \tag{14}$$

Waples (2006) investigated the fact that equation 14 does not provide an unbiased estimate of effective population size for the range of S/N_e ratios likely to occur in the study of natural populations. $\hat{N}_e$ is downwardly biased if S is less than about $2N_e$, and the bias is substantial for $S < N_e$, particularly for samples of small census population size. By examination of data from samples taken only in generation 0 (which mimics samples from a population of infinite size) and so with a $E(\hat{r}^2_{drift}) = 0$, Waples found that $1/S + 3.19/S^2$ can be a good approximate of $E(\hat{r}^2_{sample})$ if S>30 and $0.0018 + 0.907/S + 4.44/S^2$ otherwise. As $E(\hat{r}^2_{sample})$ and r^2 can both be calculated directly from data, $\hat{r}^2_{drift}$ can be deduce and the effective population size can be estimated as follows (here for large samples $S > 30$) :

$$\hat{N}_e = \frac{1/3 + \sqrt{1/9 - 2.76\hat{r}^2_{drift}}}{2\hat{r}^2_{drift}} \tag{15}$$

For calculating confidence intervals, the distribution of r^2 is approximated by a χ^2 distribution with $M = L(L-1)/2$ degrees of freedom and L the number of loci (Hill, 1981; Waples, 1991). Confidence limits for r^2 are estimated with :

$$(1-\alpha)CI = (\hat{r}^2 M/\chi^2_{(\alpha/2),M}, \hat{r}^2 M/\chi^2_{(1-\alpha/2),M}) \tag{16}$$

As explained above, the method assumes that loci are neutral (non-selected) and physically unlinked ($c = 0.5$). Microsatellite loci are highly suitable for the Linkage-Disequilibrium method (Schwartz et al. 1998), because they are highly polymorphic and nearly selectively neutral, although this may be compromised by **genetic hitchhiking**. SNPs can also be used for

Linkage-Disequilibrium. Do et al. (2014a) showed that precision of the LD method is better with 200 SNPs compared with 20 microsatellites. The relationship between the estimated r^2 and N_e with LD method takes the form of a hyperbolic curve. When $\hat{r}^2$ is less than $1/n$, negative estimates of N_e are possible. In these cases, which are most likely to arise when the sample size is small, the contribution of genetic drift to Linkage-Disequilibrium is swamped by the contribution from statistical sampling. Because it is not possible for N_e to be negative, the conventional way of interpreting a negative N_e is to replace it with an estimate of infinity (Waples, 1991). In other words, negative estimates occur when the genetic results can be explained entirely by sampling error without invoking any genetic drift (Waples and Do, 2010).

4.4.4 Applm'ication of the Linkage-Disequilibrium method for calculating population effective size

The LD method is based on some simplifying assumptions: selective neutrality, closed populations, discrete generations that may not apply to many natural populations. LD method also assumes that of the four evolutionary forces (mutation, selection, migration and genetic drift), only drift is responsible for the signal in the data (Waples and Do, 2010). The consequences of violating these assumptions have been widely studied during the last 5 years. Effects of overlapping generations (Waples and Yokota, 2006; Robinson and Moyer, 2013; Waples et al., 2014), migration (Waples and England, 2011), Hardy-Weinberg assumptions (Waples, 2015), physical link between the used loci (Waples et al., 2016), the sample size (Waples, 2006; Waples and Do, 2010) and a comparison of the genetic markers used are synthesized in Tab. 3 and Fig.8 (Waples and Do, 2010; Do et al., 2014a). Nevertheless, the neutrality assumption needs to be evaluated more precisely for a use with a high number of SNP loci. Tenesa et al. (2007) studied the estimation of human effective population size with 1,000,000 of SNPs. In this study, for each chromosome, pairwise r^2 was calculated, only for SNP pairs between 5 kb and 100 kb apart from each ohter to avoid the influence of **gene conversion** on observed LD at SNPs that are closer and to minimize the effect of a very recent expansion of the human effective population size on LD.

Tested effects	Protocole	Used estimators	Reference
Sample size, number of loci, concurrent effects of both parameters.	Random subsamples of the total data set available.	NeEstimator	Dudgeon and Ovenden (2015)
Migration	Populations simulated with and without migration. Comparison of thes estimates and the simulated N_e values with the RMSE (Root Mean Square Error).	Colony; CoNe; MLNe; NeEstimator; ONeSAMP; TMVP	Gilbert and Whitlock (2015)
Overlapping generations	Simulations of age-structured genetic data which tracked demographic and genetic processes. Studies of the bias associated.	LDNe (old version of NeEstimator)	Waples et al. (2014)
Overlapping generations and age-structure.	Simulation of several life-history scenarios. Comparison of the estimates' CV and the simulated N_e values.	LDNe	Robinson and Moyer (2013)
Migration	Populations simulated with and without migration. Comparison of these estimates and the simulated N_e values.	LDNe	Waples and England (2011)
Comparison SNPs and microsatellites	Simulations of the two types of genetic data. Studies of the CV of N_e' estimates and of the MSE (Mean Square Error).	Unavailable program	Waples and Do (2010)
Sample size, number of loci, number of alleles per locus.	Simulations of genetic data. Studies of the CV of N_e' estimates and of the MSE (Mean Square Error).	Unavailable program	Waples and Do (2010)
Allele exclusion criteria	Simulations of genetic data. Studies of the CV of N_e' estimates and of the MSE (Mean Square Error)..	Unavailable program	Waples and Do (2010)
Sample size, alleles exclusion criteria	Simulations of genetic data. Comparison of the estimates and the true values.	LDNe	Macbeth et al. (2013)

Table 3: Summary of the main bias and sensitivity analysis conducted on the LD method. $N_eEstimator$: Do et al. (2014a), $ONeSAMP$: Tallmon et al. (2008), MLN_e: Wang and Whitlock (2003), $COLONY$: Jones and Wang (2010), CoN_e: Anderson (2005), $TMVP$: Anderson (2005), $LDNe$: Waples and Do (2008).

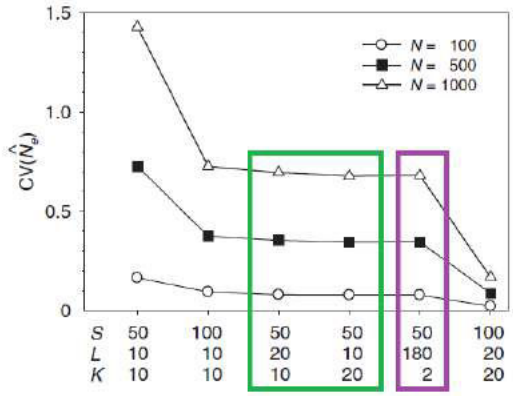


Figure 8: Effects of independently doubling sample size of individuals (S), number of loci (L), and number of alleles per locus (K) on the coefficient of variation (CV) of $\hat{N}_e$. Results are shown for three different ideal population sizes of N individuals (here N corresponds to N_e) (Waples and Do, 2010). Violet box can be interpreted as 180 SNPs and green box as 10 and 20 microsatellites with 20 and 10 alleles respectively.

Softwares

The most widely used software for estimating the population effective size is *NeEstimator* (Do et al., 2014a) which calculates contemporary unbiased estimates of the population effective size based on LD. Moreover, the softwares provides estimates of N_e based on others methods : Heterozygosity Excess, Molecular Co-ancestry, temporal method. The LD method with *NeEstimator* has been most widely used with microsatellite data but can be used with SNP data as well. Accurate estimates of N_e can be obtained with non overlapping generations by using 10-25 microsatellites, which correspond to approximately 180 SNPs and samples of at least 25-50 individuals. Precision can be improved by sampling more individuals or by sampling more genetic markers. However, precision can be improving more by doubling the number of individuals, especially for SNPs (Waples and Do, 2010).

In 2015, a new tool was developed to inferring historical N_e trends based on SNP data only : SNeP (Barbato et al., 2015). This software is a complementary tool for investigating demography through Linkage-Disequilibrium.

Applications

The Linkage-Disequilibrium method has been widely used for estimating effective population size and after temporal method, the Linkage-Disequilibrium (LD) method is the most widely used. As soon as 1991, Waples studied the feasibility of using LD method on cetacean populations for estimating population effective size and concluded that the LD method may provide meaningful information if a high number of polymorphic gene loci can be resolved (Waples, 1991). Beyond its application on marine mammals, LD method has been applied on several species including

terrestrial such as human (Tenesa et al., 2007), ruminants (Cervantes et al., 2011; Do et al., 2014b; Barbato et al., 2015) or cougar (Juarez et al., 2015) populations but also on insects (Francuski and Milankov, 2015), reptiles (Bishop et al., 2009; Monzón-Argüello et al., 2015) and of course, fishes (Macbeth et al., 2013; Van Doornik et al., 2013; Wilson et al., 2014; Dudgeon and Ovenden, 2015; Pilger et al., 2015; Perrier et al., 2015).

4.5 Estimation of effective population with Heterozygosity Excess

4.5.1 Definition and measures of the Heterozygosity Excess method

This method is based on the following principle: when the effective number of breeders (N_{eb}) in a population is small, the allele frequencies will (by chance) be different in males and females, which causes an excess of heterozygotes in the progeny with respect to Hardy-Weinberg equilibrium expectations (Pudovkin et al., 1996; Luikart and Cornuet, 1999; Leberg, 2005). In other words, when the number of breeders is small, the allele frequencies in males and females will be different due to binomial sampling error, which generates an excess of heterozygotes in the progeny relative to the HW expected proportions (Robertson, 1965; Rasmussen, 1979).

The proportion of heterozygotes expected in a progeny produced by a small and equal number of males and females can be calculated as followed :

$$H' = 2pq + pq/n = 2pq(1 + 1/(2n)) \tag{17}$$

where n is the number of haploid genomes in the mothers or fathers and p and q are the frequencies of alleles at a locus.

Pudovkin et al. (1996) called H' the proportion of heterozygotes expected to be observed (Hobs) in the progeny (given a limited number of parents), whereas the expected proportion of heterozygotes in the base population under Hardy-Weinberg proportions, $2pq$, was designated as H_{exp}. So :

$$\hat{N}_{eb} = \frac{H_{exp}}{2(H_{obs} - H_{exp})} \tag{18}$$

If the sample size of individuals is finite, H_{exp} must be estimated using the following unbiased estimator of $2N(2pq)/2N - 1$ where N is the number of progeny sampled (Luikart and Cornuet, 1999).

Pudovkin et al. (1996) proposed an other estimator of N_eb by using Selander (1970) index D, the excess of deficiency of heterozygotes, which is the reciprocal of the ratio $H_{exp}/(H_{obs} - H_{exp})$.

So the following more exact equation was derived :

$$\hat{N}_{eb} = \frac{1}{2D} + \frac{1}{2(D+1)} \quad with \quad D = \frac{H_{obs}}{H_{exp}} \tag{19}$$

4.5.2 Application of the Heterozygosity Excess method for calculating population effective size

Luikart and Cornuet (1999) conducted a simulation studies on the Heterozygosity Excess (HE). Results show that when the method is applied to natural populations, it often give estimates of N_{eb} equal to infinity with a poor precision. According to the authors, those results were not surprising in that only a few number of polymorphic loci were analyzed with a small number of progeny. They concluded that additional empirical evaluations are needed, but it is extremely difficult to find large data sets containing individuals produced from a known number of parents. Pudovkin et al. (2010) studied the sampling properties of the method. The larger is the number of loci surveyed or the larger is the sample size, the narrower is the confidence interval. The vast majority of simulations with large N_eb, the upper limits of the 95% confidence intervals often reach infinity. Authors concluded that the Heterozygosity Excess estimator is quite effective for very small numbers of parents, up to 30 individuals, if the sample sizes are larger than 200 and the cumulative numbers of alleles is more than 80. Larger numbers of alleles or loci can compensate for smaller sample sizes. If N_eb is large (50–100 individuals), samples size and the number of independent alleles should be much larger (500 – 1 000 individuals, 450 – 900 independent alleles). To summarize, Heterozygosity Excess (HE) method is a little biased estimator with a very low precision. The estimator is usefull only for very small random mating populations when many markers are genotyped from a large sample (Wang, 2005; Leberg, 2005). The assumptions of no mutation and no selection in the HE method are valid in general, because only one generation is concerned and markers are "neutral". The assumption of a single isolated population without immigration is violated in some natural populations. The strongest assumption made by the HE method seems to be random mating. When mating is not at random, then the HE generated by drift can easily be overwhelmed by that generated by non random mating (Wang, 1996, 2005). However, the method has some interesting properties : it requires only one single sample in time and is straightforward to compute (Balloux, 2004).

Software

As for the LD method, the most widely used software is $NeEstimator$, mostly used with microsatellite data (Do et al., 2014a). An other software is available Nb_HetEx (Zhdanova and Pudovkin, 2008).

Applications

The Heterozygosity Excess method for estimating effective size is less used than the two previous methods (Launey et al., 2001; Nomura, 2009; Zhivotovsky et al., 2015).

4.6 Estimation of effective population with Molecular Co-ancestry

The molecular co-ancestry between two individuals is the probability that two randomly sampled alleles from the same locus in two individuals are identical. It possess a straightforward relationship with genealogical co-ancestry and can be used to assess genetic diversity within and between populations (Nomura, 2008).

Let f_t be the coancestry among two randomly sampled individuals in generation t, and P be the probability that two randomly sampled alleles each from different individuals in generation t come from the same individual in generation $t-1$.

$$f_t = P\Big(\frac{1+F_{t-1}}{2}\Big) + (1-P)f_{t-1} \tag{20}$$

where F_{t-1} is the inbreeding coefficient of individuals in generation $t-1$.

The effective number of breeders, N_{eb} can be described.

$$N_{eb} = \frac{1}{P} \tag{21}$$

Like the Heterozygosity Excess method, the Molecular Co-ancestry method can estimate the population effective size of breeders from a simple formula if the actual N_{eb} is very small. It also suffers from a poor precision of the estimates (Luikart et al., 2010; Gilbert and Whitlock, 2015).

As for the previous methods, the most widely used software is $NeEstimator$, mostly used with microsatellite data (Do et al., 2014a). An other software is available $Molkin$ (Gutierrez et al., 2005).

4.7 Summary

Methods	Principle	Limits	Softwares & references
Multiple samples			
Temporal	Changes in alleles frequencies between samples separated in time from a population, which reflect genetic drift, are used to calculate the effective population size.	Requires population samples separated by at least two generations.	$N_e estimator$ Do et al. (2014a)
Single sample			
Linkage-Disequilibrium	Measure the non random association of alleles at different loci. That LD measure is related to a signal from the drift inside the population and to a signal related to the finite sample. Those signals are also related to the population effective size which allow its inference.	Potentially strongly biased by substructures, overlapping generations, migration and small samples.	$N_e estimator$ Do et al. (2014a), $OneSAMP$ Tallmon et al. (2008)
Heterozygosity Excess	The proportion of heterozygotes in the population is measured and compared to the expected proportion of heterozygotes As the proportion of heterozygotes is linked to the effective population size, it can be inferred.	Effective only for a very small number of parents. Very low precision.	$N_b HetEx$ Zhdanova and Pudovkin (2008), Balloux (2004)
Molecular co-ancestry	The estimator is obtained from a simple parameter (molecular co-ancestry) of allele sharing among sampled individuals. The molecular co-ancestry between two individuals is the probability that two randomly sampled alleles from the same locus in two individuals are identical by state.	Poor precision and applicable to small populations only.	Nomura (2008)

Table 4: Summary of the main methods available for estimating contemporary effective population size in natural populations

5. Synthesis of the tools available for genetic studies

The aim of this part is to review some of the available tools for studying populations genetics. Literature reviews, open source softwares developed by scientist and R-packages are described below. This review is non exhaustive.

5.1 Review articles available

The following table summarizes the main literature reviews on the use of genetic material in ecology, fisheries or for estimating population sizes.

Year	Title	Reference
2012	A review of the application of molecular genetics for fisheries management and conservation of sharks and rays	Dudgeon et al. (2012)
2012	Molecular markers: progress and prospects for understanding reproductive ecology in elasmobranchs	Portnoy and Heist (2012)
2011	Understanding and estimating effective population size for practical application in marine species management: applying effective population size estimates to marine species management	Hare et al. (2011)
2010	Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches	Luikart et al. (2010)
2005	Genetic approaches for estimating the effective size of populations	Leberg (2005)

Table 5: Summary of review articles available to our knowledge

5.2 Review of population genetic software and PYTHON modules available

5.2.1 Softwares and PYTHON modules

In the genetic field, numerous softwares are available for studying populations genetics. Almost all of them can be freely downloaded from the internet. The following list is non exhaustive

but aim to give a short review of the software available, their use, the input data type, the type of results provided, the last version available ...

▷ *ARLEQUIN*

ARLEQUIN (Excoffier and Lischer, 2010) is the powerful genetic package performing a wide variety of tests which can be divided into two categories : intra-population and inter-population methods. Here are some examples :

- Standard indices calculation
- Haplotype frequency estimation
- Test for Hardy-Weinberg equilibrium
- AMOVA (Analyses of MOlecular VAriance)
- ...

The current version available is the 3.5.2.2 at the following address: http://cmpg.unibe.ch/software/arlequin35/ and Arlequin can be launch from R with an R package. The graphical interface is designed to allow users to rapidly select the different analyses they want to perform on their data. It can handle several types of data either in haplotypic or genotypic form: DNA sequences, microsatellite data, allele frequency data, etc...

Several data formats are recognized (GENEPOP, Biosys, Phylip, Fstat, ...) but will be save in an Arlequin format *.arp*.

▷ *Bottleneck*

Bottleneck (Piry, 1999) is a program for detecting recent effective population size reductions from allele data frequencies. The principle as described in the internet page (http://www1.montpellier.inra.fr/CBGP/software/Bottleneck/pub.html) is the following. It computes for each population sample and for each locus the distribution of the heterozygosity expected from the observed number of alleles, given the sample size under the assumption of mutation-drift equilibrium. This distribution is obtained through simulating the coalescent process of n genes under three possible mutation models. This enables the computation of the average (H_{exp}) which is compared to the observed heterozygosity (Hobs, in the sense of Nei's gene diversity) to establish whether there is an Heterozygosity Excess or deficit at this locus. The distribution obtained through simulation enables also the computation of a P-value for the observed heterozygosity.

Bottleneck recognized several data formats, all are text files. The two main formats are the GENEPOP and GENETIX formats. The current version available is the 1.2.02 at the following address: http://www1.montpellier.inra.fr/CBGP/software/Bottleneck/.

▷ *BottleSim*

BottleSim (Kuo and Janzen, 2003) is a computer simulation program for simulating the process of population bottlenecks. It can implement an overlapping-generation model and simulate a wide range of scenarios regarding changes in population sizes. An option of generating raw genotypic data output is also available. The raw genotypic data output file contains the genotypic data from the last year of each iteration. The genotypic data output is in GENEPOP format.

The current version available is the 2.6 at the following address: `http://chkuo.name/software/BottleSim.html`.

▷ *CERVUS*

CERVUS (Kalinowski et al., 2007) is a computer program for assignment of parents to their offspring using genetic markers. It analyses genetic data from **codominant genetic markers** such as microsatellites (STRs) and SNPs. It assumes that the species is diploid and that markers are autosomal, although sex-linked markers can be used for some analyses. It also assumes that markers are inherited independently of each other, in other words that they are in Linkage-Equilibrium. The statistical method behind in a maximum likelihood one.

The last version available is the 3.0.7 from `http://www.fieldgenetics.com/pages/login.jsp`.

▷ *COLONY*

COLONY (Jones and Wang, 2010) implements a maximum likelihood method to assign sibship and parentage jointly, using individual multiloci genotypes at a number of co-dominant or dominant marker loci. It can be used in estimating full- and half-sib relationships, assigning parentage, inferring mating system (polygamous/monogamous) and reproductive skew in both diploid and haplo-diploid species.

The last update is COLONY2 available at `http://www.zsl.org/science/software/colony`.

▷ *CoNe*

CoNe (Anderson, 2005) computes the likelihood of Ne given data on two temporally spaced genetic samples. The statistical model used is based on the coalescent method using Markov chain Monte Carlo.

The last version available is the 1.0.1 available at `https://swfsc.noaa.gov/textblock.aspx?Division=FED&ParentMenuId=54&id=3436`.

▷ *DnaSP*

DNA sequence polymorphism (Librado and Rozas, 2009) is an interactive computer program for the analysis of DNA polymorphism from nucleotide sequence data. It calculates several

measures of DNA sequence variation within and between populations and also provides some neutrality tests. The program can also conduct computer simulations based on the coalescent process. *DnaSP* can read unphased data and can reconstruct the haplotype phases. *DnaSP* can automatically read several type of formats.

The last version is the 5.10.1, available at http://www.ub.edu/dnasp/.

▷ *DIY ABC*

DIYABC (Cornuet et al., 2014) is a software package for analyzing of population history using approximate Bayesian computation on DNA polymorphism data. It allows DNA sequence, microsatellites and SNPs. The program allows considering complex population histories including any combination of population divergence events, admixture events and changes in past population size (with population samples potentially collected at different times).

The last version is the 2.1.0 available at http://www1.montpellier.inra.fr/CBGP/diyabc/.

▷ *EASYPOP*

EASYPOP (Balloux, 2001) is a computer program allowing to simulate population genetics datasets. It allows generating genetic data for haploid, diploid, and haplodiploid organisms under a variety of mating systems. It includes various migration and mutation models. Output can be generated for the FSTAT, GENEPOP, and ARLEQUIN genetic analysis packages.

The lastest version available is the 2.0.1, on the following page http://www.unil.ch/dee/en/home/menuinst/softwares--dataset/softwares/easypop.html.

▷ *FSTAT*

FSTAT (Goudet, 2001) is a computer package for PCs which estimates and tests gene diversities and differentiation statistics from co dominant genetic markers. It can provide allelic richness per locus and sample, the F-statistics per locus ad sample, ... and can perform jackknifing and bootstrapping over loci. The current version is the 2.9.3.2,) available at http://www2.unil.ch/popgen/softwares/fstat.htm. The input file must be in a FSTAT format or GENEPOP. The format of the output file is a FSTAT format.

▷ *GENEPOP*

GENEPOP (Rousset, 2008) is a population genetics software package which computes exact tests fro Hardy-Weinberg equilibrium, population differentiation or F-statistics estimates by example. The input format is a GENEPOP format and the program allows the conversion into several widely used formats as FSTAT.

The current version is the 4.2, available at http://genepop.curtin.edu.au/.

▷ *GENETIX*

GENETIX (Belkhir et al., 1996) computes several basic parameters of population genetics such as Nei's D and H, Wright's F-statistics. It also proposed permutation-based statistical inference procedures, jackknifing and bootstrapping. The program handles only multiloci genotype on diploid organisms. Data can be written directly in the program or can be import from FSTAT, GENEPOP formats.

The current version is the 4.05 available at http://www.genetix.univ-montp2.fr/genetix/constr.htm#download.

▷ *LAMARC*

LAMARC (Likelihood Analysis with Metropolis Algorithm using Random Coalescence) (Kuhner, 2006) is a program which estimates population-genetic parameters such as population size, population growth rate, recombination rate, and migration rates. It approximates a summation over all possible genealogies that could explain the observed sample, which may be sequence, SNP, microsatellite, or electrophoretic data. All methods used in the program are derived from the coalescence theory. It requires random samples from each sub-population. A converter integrated to the program can convert PHYLIP file.

The latest version is the 2.1.0 available at http://evolution.genetics.washington.edu/lamarc/lamarc_download.html.

▷ *MLNE*

MLNE (Wang and Whitlock, 2003) is a program for calculating maximum likelihood estimates of effective population size (N_e) and migration rate from the observed temporal and spatial differences in marker allele frequencies.

The software is available at the following page http://www.zsl.org/science/software/mlne.

▷ *MolKin*

MolKin (Gutierrez et al., 2005) is a population genetics computer program that conducts several genetic analyses on multilocus information. Primary functions carried out by *MolKin* are the computation of the between individuals (and populations) Molecular Co-ancestry coefficients, the Kinship distance at individual and population levels. Additionally, users can compute with *MolKin* a set of among populations, genetic distances and F-statistics (Wright, 1950) from multilocus information.

The latest version is 3.0 available at https://pendientedemigracion.ucm.es/info/prodanim/html/JP_Web.htm.

▷ *ONeSAMP*

ONeSAMP (Tallmon et al., 2008) uses approximate Bayesian computation to estimate effective population size from a sample of microsatellite genotypes. It requires an input file of sampled individuals' microsatellite genotypes along with information about several sampling and biological parameters. The program provides an estimate of effective population size, along with 95% credible limits.

Not available at this moment (04/05/2018).

▷ *N_eEstimator*

N_eEstimator (Do et al., 2014a) is a tool for estimating contemporary effective population size (Ne) using multi-locus diploid genotypes from population samples. Data can be microsatellites of SNPs. Four methods are available to calculate ; three single-sample (point estimation) methods and one two-sample (temporal) method. The single sample methods available is the Linkage-Disequilibrium method, the Heterozygosity Excess method and the Molecular co-ancestry method. The user needs to provide genotypic data in one of the accepted formats (FSTAT, GENEPOP).

The last version available is the 2.01 on `http://www.molecularfisherieslaboratory.com.au/neestimator-software/`.

▷ *POPGENE*

POPGENE (Yeh and Boyle, 1997) is a user-friendly computer freeware for the analysis of genetic variation among and within populations using co-dominant and dominant markers. It computes both comprehensive genetic statistics (e.g., allele frequency, gene diversity, genetic distance, G-statistics, F-statistics) and complex genetic statistics (e.g., gene flow, neutrality tests, Linkage-Disequilibrium, multi-locus structure).

The current version is the 1.32 available at `https://www.ualberta.ca/~fyeh/popgene_download.html`.

▷ *simuPOP*

simuPOP (Peng and Kimmel, 2005) is a general-purpose individual-based forward-time population genetics simulation environment. It models individuals with genotypes and simulates the transmission of individual genotype when a population evolves generation by generation. Although the basic evolutionary scenario follows a discrete non-overlapping generation model, aged structured populations can be mimicked using special non-random mating schemes. simuPOP evolves populations forward in time, subject to arbitrary number of genetic and environmental forces such as mutation, recombination, migration and population/subpopulation size changes. Statistics of populations can be calculated and visualized dynamically. simuPOP is provided as a number of Python modules, which provide of a large number of Python objects and functions, including

population, mating schemes, operators (objects that manipulate populations) and simulators to coordinate the evolutionary processes. Users have to write a Python script to link these modules and run simulation.

The procedure to install simuPOP is available at the following address : http://simupop.sourceforge.net/Main/Download.

▷ *STRUCTURE*

STRUCTURE (Pritchard et al., 2000) is a software package for using multi-locus genotype data to investigate population structure. Its uses include inferring the presence of distinct populations, assigning individuals to populations, studying hybrid zones, identifying migrants and admixed individuals, and estimating population allele frequencies in situations where many individuals are migrants or admixed. It can be applied to most of the commonly-used genetic markers, including SNPS, microsatellites, RFLPs and AFLPs.

The last version available is the 2.3.4 at http://pritchardlab.stanford.edu/structure.html.

5.2.2 Use of population genetic software for estimating contemporary N_e

Softwares used in a few published studies to estimating contemporary Ne are listed in table 6. The table compiles a few examples and is in no way an exhaustive review.

Case studies and publication	*N_eEstimator*	*OneSAMP*	*COLONY*	Other software
Cetaceans populations Waples (1991)				PP
Human populations Tenesa et al. (2007)				PP
Crocodile populations Bishop et al. (2009)	X X			TM3.1
Ruminant populations Cervantes et al. (2011)	X			Molkin
Gecko populations Hoehn et al. (2012)	X X	X		
Mackerel populations Macbeth et al. (2013)	X			
Salmon populations Van Doornik et al. (2013)	X			SALMONNb
Sturgeon populations Wilson et al. (2014)	X X X			
Shark populations Dudgeon and Ovenden (2015)	X			
Fly population Francuski and Milankov (2015)	Methods not precised			
Simulation only Gilbert and Whitlock (2015)	X X X X			
Cougar populations Juarez et al. (2015)		X	X	*CoN_e*; *MLN_e*
Snake populations Monzón-Argüello et al. (2015)	X X X			
Salmon populations Perrier et al. (2015)	X			
Endemic fishes populations Pilger et al. (2015)	X		X	

Table 6: Software used for calculating the contemporary effective population size and the associated method : Linkage-Disequilibrium, Heterozygosity Excess, Co-ancestry method, Temporal estimation method. Personnal Programm : PP, *N_eEstimator*: Do et al. (2014a), *ONeSAMP*: Tallmon et al. (2008), *COLONY*: Jones and Wang (2010), *SN_eP*: Barbato et al. (2015), *TM*3.1: Berthier et al. (2002), *SALMONNb*: Waples et al. (2006), *CoN_e*: Anderson (2005), *MLN_e*: Wang and Whitlock (2003), *MolKin*: (Gutierrez et al., 2005).

5.3 R-packages for studying population genetics

As for software, several R package exist for studying population genetics but only a few for estimating effective population size. Here, choice has been made to only present packages which are still available on the CRAN (https://cran.r-project.org/) or which can be downloaded directly from the programmer website. The following package descriptions are those provided in the CRAN R Project web page of the package or the programmer's description on his own website.

Only two packages are available for calculating the population effective size. The first one uses a two samples method and the second one calculates the population effective size of simulated populations with known demographic data. No R-package which calculate the population effective size of natural population with single sample methods was found when this review was compiled (2015). However, several packages can estimates Linkage-Disequilibrium or Heterozygosity Excess.

▷ *adegenet*: tool-set for the exploration of genetic and genomic data.

Adegenet provides formal classes for storing and handling various genetic data, including genetic markers with varying ploidy and hierarchical population structure, alleles counts by populations, and genome-wide SNP data. It also implements original multivariate methods (DAPC, sPCA), graphics, statistical tests, simulation tools, distance and similarity measures, and several spatial methods. A range of both empirical and simulated data sets is also provided to illustrate various methods (Jombart, 2008; Jombart and Ahmed, 2011).

User manual: https://cran.r-project.org/web/packages/adegenet/adegenet.pdf.

▷ *ape*: Analyses of Phylogenetics and Evolution

This package provides functions for reading, writing, plotting, and manipulating phylogenetic trees, analyses of comparative data in a phylogenetic framework, ancestral character analyses, analyses of diversification and macro-evolution, computing distances from allelic and nucleotide data (Paradis et al., 2004).

User manual: https://cran.r-project.org/web/packages/ape/ape.pdf.

▷ *diveRsity*: A comprehensive, general purpose population genetics analysis package

This package allows the calculation of both genetic diversity partition statistics, genetic differentiation statistics, and locus informativeness for ancestry assignment. It also provides users with various option to calculate bootstrapped 95% confidence intervals both across loci, for pairwise population comparisons, and to plot these results interactively. Parallel computing capabilities and pairwise results without bootstrapping are provided. Weir and Cockerham's 1984 F-statistics are also calculated. Various plotting features are also provided, as well as Chi-square tests of genetic heterogeneity are also provided. Functionality for the calculation of various diversity parameters

is possible for RAD-seq derived SNP data sets containing thousands of marker loci. A shiny application for the development of microsatellite multiplexes is also available.

User manual: https://cran.r-project.org/web/packages/diveRsity/diveRsity.pdf.

▷ *gap*: genetic analysis package

This package is designed as an integrated package for genetic data analysis of both population and family data. Currently, it contains functions for sample size calculations of both population-based and family-based designs, probability of familial disease aggregation, kinship calculation, statistics in linkage analysis, and association analysis involving genetic markers including haplotype analysis with or without environmental covariates (Zhao, 2015).

User manual: https://cran.r-project.org/web/packages/gap/gap.pdf.

▷ *Geneland*: detection of structure from multilocus genetic data

This package provides tools for stochastic simulations and MCMC inferences of structure from genetic data.

User manual: https://cran.r-project.org/web/packages/Geneland/Geneland.pdf.

▷ *genetics*: population genetics

The package *genetics* provides classes and methods for handling genetic data. It includes classes to represent genotypes and haplotypes at single markers up to multiple markers on multiple chromosomes. Functions include allele frequencies, flagging homo/heterozygotes, flagging carriers of certain alleles, estimating and testing for Hardy-Weinberg disequilibrium, estimating and testing for Linkage-Disequilibrium, ...

User manual: https://cran.r-project.org/web/packages/genetics/genetics.pdf.

▷ *hapassoc*: inference of trait associations with SNP haplotypes and other attributes using the EM algorithm

The package is used for inference of trait associations with haplotypes and other covariates in generalized linear models. The functions are developed primarily for data collected in cohort or cross-sectional studies. They can accommodate uncertain haplotype phase and handle missing genotypes at some SNPs (Burkett et al., 2006).

User manual: https://cran.r-project.org/web/packages/hapassoc/hapassoc.pdf.

▷ *haplo.stat*: statistical analysis of haplotypes with traits and covariates when linkage phase is ambiguous

A suite of R routines for the analysis of indirectly measured haplotypes. The statistical methods assume that all subjects are unrelated and that haplotypes are ambiguous (due to unknown

linkage phase of the genetic markers).

User manual: https://cran.r-project.org/web/packages/haplo.stats/haplo.stats.pdf.

▷ *hwde*: models and tests for departure from Hardy-Weinberg equilibrium and independence between loci

Fits models for genotypic disequilibrium.

User manual: https://cran.r-project.org/web/packages/hwde/hwde.pdf.

▷ *LDcorSV*: Linkage-Disequilibrium corrected by the structure and the relatedness

The package provides a set of functions which aim is to propose four measures of Linkage-Disequilibrium: the usual r^2 measure, the r^2_S measure (r^2 corrected by the structure sample), the r^2_V (r^2 corrected by the relatedness of genotyped individuals), the $r^2_V S$ measure (r^2 corrected by both the relatedness of genotyped individuals and the structure of the sample).

User manual: https://cran.r-project.org/web/packages/LDcorSV/LDcorSV.pdf.

▷ *NB*: Maximum Likelihood method in estimating effective population size from genetic data

The allele frequencies in a closed population change over time, which is known as genetic drift. The magnitude of change is directly related to N_e. This package aims to estimate this quantity from genetic samples collected over multiple time points.

User manual: https://cran.r-project.org/web/packages/NB/NB.pdf.

▷ *NEff*: Calculating effective sizes based on known demographic parameters of a population

This package estimates effective population size with data obtained within less than a generation but considering demographic parameters is possible. This individual based model uses demographic parameters of a population to calculate annual effective sizes and effective population sizes (per generation). A defined number of alleles and loci will be used to simulate the genotypes of the individuals. Step-wise mutation rates can be included. Variations in life history parameters (sex ratio, sex-specific survival, recruitment rate, reproductive skew) are possible. These results will help managers to define whether existing populations as viable or not.

Demographic parameter of a population are used as input values to calculate annual effective sizes and effective population sizes (per generation). The demographic input are the age of sexual maturity, the number of offspring per female per year, the sex ratio, the female and male variances in reproductive success and the female, male and juvenile survival rates. Heterozygosity over time is observed and used to calculate effective sizes. The population can be adapted to every life history parameter combination or genetic variation.

```
> population(max.repeat=3, max.time=50, Na=200, n.recruit.fem=2, surv.ad.fem=0.7,
+ N.loci=190, N.allele=2)
I've done the repeat number 1
I've done the repeat number 2
I've done the repeat number 3


Summary table of an optimal behaviour of your population:
$parameters
                                    parameter    value 95% lower CI 95% upper CI
1       slope of heterozygosity loss per year  -0.0011      -0.0013       -9e-04
2 slope of heterozygosity loss per generation  -0.0033      -0.0036      -0.0029
3                       mean generation length      3.2        2.694        3.706
4                              Ny[simulation] 457.4719     421.0097      493.934
5                                   Ny[calc]  402.224         <NA>         <NA>
6                              Ne[simulation] 153.7963     132.7237     174.8688
7                                   Ne[calc]  125.695         <NA>         <NA>
```

Figure 9: Example of application of the packkage *NEff*. A population of 200 individuals is simulated over 50 years. The survival rate of the female is 0.7 and the number of offspring per female per year is 2. Genetic data used are simulated as 490 bi-allelic SNPs. The simulated effective population size is estimated at 154 (IC 95%: [133 ; 175]) and the calculated population effective size is estimated at 125.7.

User manual: https://cran.r-project.org/web/packages/NEff/NEff.pdf.

▷ *NeON*: R-package to estimate human effective population size and divergence time from patterns of Linkage-Disequilibrium between SNPs

The *NeON* R package has been designed to explore population's LD patterns in order to reconstruct two key parameters of human evolution: the effective population size and the divergence time between populations. *NeON* starts with binary or pairwise-LD PLINK files, and allows (a) to assign a genetic map position using HapMap (NCBI release 36 or 37) (b) to calculate the effective population size over time exploiting the relationship between N_e and the average squared correlation coefficient of LD ($r2LD$) within predefined recombination distance categories, and (c) to calculate the confidence interval about N_e based on the observed variation of the estimator across chromosomes. This package also allows to estimating the divergence time between populations given the Ne values calculated from the within-population LD data and a matrix of between-populations F_{ST}. These routines can easily be adapted to any species whenever genetic map positions are available.

Package and user manual available at http://www.unife.it/dipartimento/biologia-evoluzione/ricerca/evoluzione-e-genetica/software.

▷ *pegas*: population and evolutionary genetics analysis system

This package provides functions for reading, writing, plotting, analysing, and manipulating allelic and haplotypic data, and for the analysis of population nucleotide sequences and microsatellites including coalescence analyses (Paradis, 2010).

User manual: https://cran.r-project.org/web/packages/pegas/pegas.pdf.

▷ *Popgen*: statistical and population Genetics

This is a package that implements a variety of statistical and population genetic methodology, for example, LD measures from genotypes or haplotypes, clustering of SNPs, inferences of population structure, etc...

User manual: https://cran.r-project.org/web/packages/popgen/popgen.pdf.

▷ *poppr*: Population genetic analyses for hierarchical analysis of partially clonal populations built upon the architecture of the 'adegenet' package.

This package provides tools for population genetic analysis, which include genotypic diversity measures, genetic distances with bootstrap support, native organization and handling of population hierarchies, and clone correction (Kamvar et al., 2014).

User manual: https://cran.r-project.org/web/packages/poppr/poppr.pdf.

Glossary

Allele Variant or alternative form of the DNA sequence at given locus. 5

Allele exclusion criteria Exclusion criteria of rare allele. Commonly fixed at 0.05, 0.02 or 0.01. 33

Ascertainment bias A possible form of bias that occurs when genetic loci are assumed to represent population (or species) wide genetic variation but are actually a feature of a small subset of the population (or species). 14

Bottleneck Sudden reduction in population size including a loss in genetic variation. It increases allele frequencies sampling error and has a disproportionate impact on the effective population size in later generations even if census sizes increases. 22

Census size The total number of individuals in a population including immatures. 19

Codominant genetic marker A genetic marker in which both alleles are expressed, thus heterozygous individuals can be distinguished from either homozygous state. 41

Dioecious Species having the sexual organs (male and female) upon distinct individuals. 30

Diploid A nucleus or individual having two copies of each chromosome. 12

Dispersal Leaving an area of birth or activity for another area. 8, 14

Effective population size The number of individuals in an ideal Wright-Fisher population reproducting and contributing to the alleles present in the next generation. This population experiences as much genetic drift as an actual population regardless of census size. 4, 14

Evolution Change in gene frequency through time resulting from natural selection and producing cumulative changes in characteristics of a population. 3

F-statistics A set of statistics used to estimate deviations from the Hardy-Weinberg model in populations and to estimate the degree to which a group of populations is genetically subdivided. 15

Fixation The loss of alleles from a polymorphic population until only one remeains, i.e., becomes monomorphic. 7, 10

Fixation index The proportion by which heterozygosity is reduced or increased relative to the heterozygosity in a randomly mating population with the same allele frequencies. 15

Gene Specific nucleotide sequence of DNA that codes for a particular protein, tRNA or rNA, usually means an exon or series of exons. 5

Gene conversion Nonreciprocal transfer of genetic information between homologous chromosome. 32

Gene flow Exchange of genetic material between populations. 8, 14, 27

Genetic drift Change in allele frequecies within a population over time due to the sampling effect of small population size. 3, 10, 27

Genetic hitchhiking Occurs when an allele changes frequency not because it itself is under natural selection, but because it is near another gene on the same chromosome that is undergoing a selective sweep. 31

Genetic marker A sequence of DNA or protein than can be screened to reveal key attributes of its state or composition and thus used to reveal genetic variation. Example : SNP, microsatellites,.... 11

Genotype State for a particular genetic locus of an organism. 4

Haplotype Genetic data from a single chromosome. 5

Hardy-Weinberg equilibrium The proposition that genotypic ratios resulting from random mating remained unchanged from one generation to another, provided natural selection, genetic drift and mutation are absent. 10

Heterozygosity Can be seen as the probability of a gene to be heterozygous. 15

Heterozygosity Excess excess of heterozygotes in the progeny relative to the HW expected proportions due to differences in allele frequencies in males and females caused by binomial sampling error. 26

Inbreeding Reproduction between closely related individuals; include autogamy. 15

Linkage-Disequilibrium Two alleles from different loci on the same chromosome co-occuring at a significantly greater frequency than expected by a random association. 26

Locus A specific region or position on the genome or chromosome. 5

Mating system Behavioural mechanism involved in the acquisition of a mate, including the number of mates acquired, the manner in which they are acquired, the nature of the pair bond and provision of parental care. 3

Microsatellite Short tandem repeats of a short sequence of (typically two to four) nucleotides randomly distributed throughout the genome. 11

Migration Intentional, directional, usually seasonal movements of animals between two regions or habitats; involving departure and return of the same individual; *i.e.* a round trip movement. 3, 8, 14

Minimum Viable Population The minimum effective population required to persist despite genetic drift, demographic and environmental stochasticity. 22

Mutation Transmissible change in structure of a gene or chromosome. 3, 27

Mutation rate The frequency at which a particular mutation occurs in a genome. 11

Non overlapping generations Populations where individuals of each generation die before the birth of individuals from the next generation. 10

Population (Genetics) All the individuals connected by gene flow, i.e., the gene pool. 3

Population genetic Studies of populations' reproduction and not individuals' reproduction, *i.e.* the studies of the distribution and the evolution of the alleles and genotypes frequencies in populations. 4

Population structure Heterogeneity in allele frequencies across a population caused by limited gene flow. 3, 14

Recombination Exchange of genetic material between two homologous chromosomes to break up linkage groups and yield allelic combinaisons not occuring in parental generations. 3, 27

Selection The influence of the environnement in determining which individuals will breed and pass their genes on to the next generation and which will not breed. 3, 27

Sex ratio The relative number of males and females in a population. 10

Siblings Individuals having the same mother and the same father. 20

Single Nucleotide Polymorphism (SNP) Single base pair substitution distributed throughout the nuclear genome. The majority of SNPs are diallelic with low chances of homoplasy. 11

Subpopulation Spatially distinct unit of a population. 14

Wright-Fisher model A simplified version of the biological life cycle where all sampling to found the next generation occurs from an infinite pool of gametes built from equal contributions of all individuals. This approximation is commonly employed to model genetic drift. 10

Bibliography

Alter, S. E., Rynes, E., and Palumbi, S. R. (2007). DNA evidence for historic population size and past ecosystem impacts of gray whales. *Proceedings of the National Academy of Sciences*, 104(38):15162–15167.

Anderson, E. C. (2005). An Efficient Monte Carlo Method for Estimating Ne From Temporally Spaced Samples Using a Coalescent-Based Likelihood. *Genetics*, 170(2):955–967.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., and Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, 3(10):e3376.

Balloux, F. (2001). EASYPOP (Version 1.7): A Computer Program for Population Genetics Simulations. *Journal of Heredity*, 92(3):301–302.

Balloux, F. (2004). Heterozygote excess in small populations and the heterozygote excess effective population size. *Evolution*, 58(9):1891–1900.

Barbato, M., Orozco-terWengel, P., Tapio, M., and Bruford, M. W. (2015). SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Frontiers in Genetics*, 6.

Beaumont, A. R., Boudry, P., and Hoare, K. (2010). *Biotechnology and genetics in fisheries and aquaculture*. Blackwell, Chichester ; Ames, Iowa, 2nd ed edition.

Beaumont, M. A. (1999). Detecting population expansion and decline using microsatellites. *Genetics*, (153):2013–2029.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F. (1996). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations.

Berthier, P., Beaumont, M. A., Cornuet, J.-M., and Luikart, G. (2002). Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: a genealogical approach. *Genetic Society of America*, 160(2):741–751.

Bishop, J. M., Leslie, A. J., Bourquin, S. L., and O'Ryan, C. (2009). Reduced effective population size in an overexploited population of the Nile crocodile (*Crocodylus niloticus*). *Biological Conservation*, 142(10):2335–2341.

Brumfield, R. T., Beerli, P., Nickerson, D. A., and Edwards, S. V. (2003). The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology & Evolution*, 18(5):249–256.

Burkett, K., Graham, J., and McNeney, B. (2006). hapassoc: software for likelihood inference of trait associations with SNP haplotypes and other attributes. *J Stat Soft*, 16(2):1–19.

Castro, A. L. F., Stewart, B. S., Wilson, S. G., Hueter, R. E., Meekan, M. G., Motta, P. J., Bowen, B. W., and Karl, S. A. (2007). Population genetic structure of Earth's largest fish, the whale shark (*Rhincodon typus*). *Molecular Ecology*, 16(24):5183–5192.

Caughley, G. (1994). Directions in conservation biology. *The Journal of Animal Ecology*, 63(2):215.

Cervantes, I., Pastor, J., Gutiérrez, J., Goyache, F., and Molina, A. (2011). Computing effective population size from molecular data: The case of three rare Spanish ruminant populations. *Livestock Science*, 138(1-3):202–206.

Chevolot, M. (2006). *Assessing genetic structure of thornback ray, Raja clavata : A thorny situation?* PhD thesis, University of Groningen.

Cockerham, C. C. and Weir, B. S. (1977). Digenic descent measures for finite populations. *Genetical Research*, 30(02):121.

Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M., and Estoup, A. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, 30(8):1187–1189.

Darwin, C. (1859). *The origin of species.* Collins classic. Harper Press, 2011 edition.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., and Ovenden, J. R. (2014a). NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size Ne from genetic data. *Molecular Ecology Resources*, 14(1):209–214.

Do, K.-T., Lee, J.-H., Lee, H.-K., Kim, J., and Park, K.-D. (2014b). Estimation of effective population size using single-nucleotide polymorphism (SNP) data in Jeju horse. *Journal of Animal Science and Technology*, 56(1):28.

Dudgeon, C. L., Blower, D. C., Broderick, D., Giles, J. L., Holmes, B. J., Kashiwagi, T., Krück, N. C., Morgan, J. A. T., Tillett, B. J., and Ovenden, J. R. (2012). A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *Journal of Fish Biology*, 80(5):1789–1843.

Dudgeon, C. L. and Ovenden, J. R. (2015). The relationship between abundance and genetic effective population size in elasmobranchs: an example from the globally threatened zebra shark *Stegostoma fasciatum* within its protected range. *Conservation Genetics*, 16(6):1443–1454.

Excoffier, L., Laval, G., and Schneider, S. (2005). Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, (1):47–50.

Excoffier, L. and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3):564–567.

Excoffier, L., Smouse, P., and Quattro, J. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131(2):479–491.

Francuski, L. and Milankov, V. (2015). Assessing spatial population structure and heterogeneity in the dronefly: spatial population structure in the dronefly. *Journal of Zoology*, pages n/a–n/a.

Franklin, I. (1980). Evolutionary change in small populations. In *Conservation biology: an evolutionary-ecological perspective*, pages 135–149. Sinauer Associates, Sunderland, Mass.

Gilbert, K. J. and Whitlock, M. C. (2015). Evaluating methods for estimating local effective population size with and without migration: estimating N_e in the presence of migration. *Evolution*, 69(8):2154–2166.

Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3).

Grant, W. and Waples, R. S. (2000). Scales of temporal and spatial genetic variability in marine fishes: implications for fisheries oceanography. In *Fisheries Oceanography: Fish Biology and Aquatic Resources*, pages 61–93. Blackwell science edition.

Gutierrez, J. P., Royo, L. J., Alvarez, I., and Goyache, F. (2005). MolKin v2.0: A Computer Program for Genetic Analysis of Populations Using Molecular Coancestry Information. *Journal of Heredity*, 96(6):718–721.

Hamilton, M. B. (2009). *Population genetics*. Wiley-Blackwell, Chichester, UK ; Hoboken, NJ.

Hardy, G. (1908). Mendelian proportions in a mixed population. *Science*, (28):49–50.

Hare, M. P., Nunney, L., Schwartz, M. K., Ruzzante, D. E., Burford, M., Waples, R. S., Ruegg, K., and Palstra, F. (2011). Understanding and estimating effective population size for practical application in marine species management: applying effective population size estimates to marine species management. *Conservation Biology*, 25(3):438–449.

Hartl, D. L. (1994). *Génétique des populations*. Sciences et histoire. Médecine Sciences Flammarion.

Hartl, D. L. and Jones, E. W. (1998). *Genetics: principles and analysis*. Jones and Bartlett Publishers, Sudbury, Mass, 4th ed edition.

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., Cariani, A., Maes, G. E., Diopere, E., Carvalho, G. R., and Nielsen, E. E. (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges: analytical approaches. *Molecular Ecology Resources*, 11:123–136.

Hill, W. (1974). Estimation of Linkage-Disequilibrium in randomly mating populations. *Heredity*, 33(2):229–239.

Hill, W. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetics research*, 38(03):209–216.

Hill, W. and Robertson, A. (1968). Linkage Disequilibrium in finite populations. *Theoretical and applied genetics*, 38(226-231).

Hoehn, M., Gruber, B., Sarre, S. D., Lange, R., and Henle, K. (2012). Can Genetic Estimators Provide Robust Estimates of the Effective Number of Breeders in Small Populations? *PLoS ONE*, 7(11):e48464.

Holsinger, K. (2001). Hardy–Weinberg Law. In *Encyclopedia of Genetics*, pages 912–914. Elsevier.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11):1403–1405.

Jombart, T. and Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27(21):3070–3071.

Jones, O. R. and Wang, J. (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, 10(3):551–555.

Jorde, P. and Ryman, N. (1995). Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics*, 139(2):1077–1090.

Juarez, R. L., Schwartz, M. K., Pilgrim, K. L., Thompson, D. J., Tucker, S. A., Smith, J. B., and Jenks, J. A. (2015). Assessing temporal genetic variation in a cougar population: influence of harvest and neighboring populations. *Conservation Genetics*.

Kalinowski, S. T., Taper, M. L., and Marshall, T. C. (2007). Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment: CERVUS likelihood model. *Molecular Ecology*, 16(5):1099–1106.

Kamvar, Z. N., Tabima, J. F., and Grünwald, N. J. (2014). *Poppr* : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2:e281.

Kimura, M. (1990). *The neutral theory of molecular evolution*. Cambridge Univ. Press, Cambridge.

Kuhner, M. K. (2006). LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics*, 22(6):768–770.

Kuo, C.-H. and Janzen, F. J. (2003). BottleSim: a bottleneck simulation program for long-lived species with overlapping generations. *Molecular Ecology Notes*, 3(4):669–673.

Lamarck, J. (1809). *Philosophie géologique*. Dentu, Paris.

Lande, R. (1995). Mutation and conservation. *Conservation Biology*, 9(4):782–791.

Launey, S., Barre, M., Gerard, A., and Naciri-Graven, Y. (2001). Population bottleneck and effective size in Bonamia ostreae-resistant populations of Ostrea edulis as inferred by microsatellite markers. *Genetics Research*, 78(03):259.

Leberg, P. (2005). Genetic approaches for estimating the effective size of populations. *Journal of Wildlife Management*, 69(4):1385–1399.

Lewontin, R. (1964). The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics*, 49:49–67.

Librado, P. and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25(11):1451–1452.

Lowe, A., Harris, S., and Ashton, P. (2004). *Ecological genetics: design, analysis, and application*. Blackwell Pub, Malden, MA, USA.

Luikart, G. and Cornuet, J.-M. (1999). Estimating the effective number of breeders from heterozygote excess in progeny. *Genetics*, 151(3):1211–1216.

Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K., and Allendorf, F. W. (2010). Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics*, 11(2):355–373.

Macbeth, G. M., Broderick, D., Buckworth, R. C., and Ovenden, J. R. (2013). Linkage Disequilibrium Estimation of Effective Population Size with Immigrants from Divergent Populations: A Case Study on Spanish Mackerel (Scomberomorus commerson). *G3: Genes/Genomes/Genetics*, 3(4):709–717.

Meffe, G. K. and Carroll, C. R. (1997). *Principles of conservation biology*. Sinauer, Sunderland, Ma, 2nd ed edition.

Mendel, G. (1866). Versuche über Plflanzen-hybriden. (Bateson translation). *Verhandlungen des naturforschenden Ver-eines in Brünn*, (4):3–47.

Monzón-Argüello, C., Patiño-Martínez, C., Christiansen, F., Gallo-Barneto, R., Cabrera-Pérez, M. n., Peña-Estévez, M. n., López-Jurado, L. F., and Lee, P. L. M. (2015). Snakes on an island: independent introductions have different potentials for invasion. *Conservation Genetics*, 16(5):1225–1241.

Morin, P. A., Martien, K. K., and Taylor, B. L. (2009). Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources*, 9(1):66–73.

Murray, B. W., Wang, J. Y., Yang, S.-C., Stevens, J. D., Fisk, A., and Svavarsson, J. (2008). Mitochondrial cytochrome b variation in sleeper sharks (Squaliformes: Somniosidae). *Marine Biology*, 153(6):1015–1022.

Nance, H. A., Klimley, P., Galván-Magaña, F., Martínez-Ortíz, J., and Marko, P. B. (2011). Demographic processes underlying subtle patterns of population structure in the scalloped hammerhead shark, *Sphyrna lewini*. *PLoS ONE*, 6(7):e21459.

Nei, M. (1973). Analysis of Gene Diversity in Subdivided Populations. *Proc. Nat. Acad. Sci. USA*, 70(12):3321–3323.

Nielsen, R. (2000). Estimation of population parameters and recombination rates from single nucleotide polymorphisms. *Genetics*, 154(2):931–942.

Nikolic, N. (2009). *Diversité génétique et taille efficace chez les populations de poissons sauvages : le cas du Saumon atlantique un poisson migrateur amphihalin menacé.* PhD thesis, Université Toulouse III – Paul Sabatier, Toulouse.

Nomura, T. (2008). Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evolutionary Applications*, 1(3):462–474.

Nomura, T. (2009). Interval Estimation of the Effective Population Size from Heterozygote-Excess in SNP Markers. *Biometrical Journal*, 51(6):996–1016.

Paradis, E. (2010). pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*, 26(3):419–420.

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, 20(2):289–290.

Peng, B. and Kimmel, M. (2005). simuPOP: a forward-time population genetics simulation environment. *Bioinformatics*, 21(18):3686–3687.

Perrier, C., April, J., Cote, G., Bernatchez, L., and Dionne, M. (2015). Effective number of breeders in relation to census size as management tools for Atlantic salmon conservation in a context of stocked populations. *Conservation Genetics*.

Pigliucci, M. (2009). An Extended Synthesis for Evolutionary Biology. *Annals of the New York Academy of Sciences*, 1168(1):218–228.

Pilger, T. J., Gido, K. B., Propst, D. L., Whitney, J. E., and Turner, T. F. (2015). Comparative conservation genetics of protected endemic fishes in an arid-land riverscape. *Conservation Genetics*, 16(4):875–888.

Piry, S. (1999). Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity*, 90(4):502–503.

Portnoy, D. S. and Heist, E. J. (2012). Molecular markers: progress and prospects for understanding reproductive ecology in elasmobranchs. *Journal of Fish Biology*, 80(5):1120–1140.

Pritchard, J., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2):945–959.

Pudovkin, A., Zaykin, D., and Hedgecock, D. (1996). On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics*, 144(1):383–387.

Pudovkin, A. I., Zhdanova, O. L., and Hedgecock, D. (2010). Sampling properties of the heterozygote-excess estimator of the effective number of breeders. *Conservation Genetics*, 11(3):759–771.

Rai, U. K. (2006). Minimum sizes for viable population and conservation biology. *Our Nature*, 1(1).

Rasmussen, D. I. (1979). Sibling Clusters and Genotypic Frequencies. *The American Naturalist*, 113(6):948.

Reed, D. H. and Bryant, E. H. (2000). Experimental tests of minimum viable population size. *Animal Conservation*, 3(1):7–14.

Robertson, A. (1965). The interpretation of genotypic ratios in domestic animal populations. *Animal Production*, 7(03):319–324.

Robinson, J. D. and Moyer, G. R. (2013). Linkage disequilibrium and effective population size when generations overlap. *Evolutionary Applications*, 6(2):290–302.

Rogers, A. R. and Huff, C. (2009). Linkage Disequilibrium Between Loci With Unknown Phase. *Genetics*, 182(3):839–844.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, 8(1):103–106.

Russell, J. C. and Fewster, R. M. (2009). Evaluation of the Linkage Disequilibrium method for estimating effective population size. In Thomson, D. L., Cooch, E. G., and Conroy, M. J., editors, *Modeling Demographic Processes In Marked Populations*, pages 291–320. Springer US, Boston, MA.

Ryman, N. and Utter, F., editors (1987). *Population genetics & fishery management*. Washington Sea Grant Program : Distributed by University of Washington Press, Seattle.

Schaid, D. J. (2004). Linkage disequilibrium testing when linkage phase is unknown. *Genetics*, 166:505–512.

Schleif, R. F. (1993). *Genetics and molecular biology*. Johns Hopkins University Press, Baltimore, 2nd ed edition.

Schultz, J. K., Feldheim, K. A., Gruber, S. H., Ashley, M. V., Mcgovern, T. M., and Bowen, B. W. (2008). Global phylogeography and seascape genetics of the lemon sharks (genus *Negaprion*). *Molecular Ecology*, 17(24):5336–5348.

Schwartz, M., Luikart, G., and Waples, R. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology & Evolution*, 22(1):25–33.

Selander, R. (1970). Behaviour and genetic variation in natural populations. *American Zoologist*, 10:53–66.

Shaffer, M. L. (1981). Minimum population sizes for species conservation. *BioScience*, 31(2):131–134.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, 236(4803):787–792.

Smith, R. L. and Smith, T. M. (2001). *Ecology & field biology*. Benjamin Cummings, San Francisco, 6th ed edition.

Sober, E., editor (1994). *Conceptual issues in evolutionary biology*. MIT Press, Cambridge, Mass, 2nd ed edition.

Soulé, M. E. and Wilcox, B. A., editors (1980). *Conservation biology: an evolutionary-ecological perspective*. Sinauer Associates, Sunderland, Mass.

Tallmon, D. A., Koyuk, A., Luikart, G., and Beaumont, M. A. (2008). Computer programs: onesamp: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources*, 8(2):299–301.

Tenesa, A., Navarro, P., Hayes, B. J., Duffy, D. L., Clarke, G. M., Goddard, M. E., and Visscher, P. M. (2007). Recent human effective population size estimated from linkage disequilibrium. *Genome Research*, 17(4):520–526.

Van Doornik, D. M., Eddy, D. L., Waples, R. S., Boe, S. J., Hoffnagle, T. L., Berntson, E. A., and Moran, P. (2013). Genetic Monitoring of Threatened Chinook Salmon Populations: Estimating Introgression of Nonnative Hatchery Stocks and Temporal Genetic Changes. *North American Journal of Fisheries Management*, 33(4):693–706.

Vignal, A., Milan, D., SanCristobal, M., and Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, 34(3):275–305.

Wang, J. (1996). Deviation from Hardy–Weinberg proportions in finite populations. *Genetical Research*, 68(03):249.

Wang, J. (2005). Estimation of effective population sizes from data on genetic markers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1459):1395–1409.

Wang, J. and Whitlock, M. C. (2003). Estimating effective population size and migration rates from genetic samples over space and time. *Genetics*, 163(1):429–446.

Waples, R., K., Larson, W., A., and Waples, R. (2016). Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. *Heredity*, 117(4):233–240.

Waples, R. S. (1991). Genetic method for estimating the effective size of Cetacean populations. *Rep. int. Whal. Commn*, 13:279–300.

Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci*. *Conservation Genetics*, 7(2):167–184.

Waples, R. S. (2015). Testing for Hardy-Weinberg Proportions: Have We Lost the Plot? *Journal of Heredity*, 106(1):1–19.

Waples, R. S. (2016). Tiny estimates of the N_e/N ratio in marine fishes: Are they real? *Journal of Fish Biology*, 89(6):2479–2504.

Waples, R. S., Antao, T., and Luikart, G. (2014). Effects of overlapping generations on Linkage Disequilibrium estimates of effective population size. *Genetics*, 197(2):769–780.

Waples, R. S. and Do, C. (2008). ldne : a program for estimating effective population size from data on linkage disequilibrium: COMPUTER PROGRAMS. *Molecular Ecology Resources*, 8(4):753–756.

Waples, R. S. and Do, C. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3(3):244–262.

Waples, R. S. and England, P. R. (2011). Estimating contemporary effective population size on the basis of Linkage Disequilibrium in the face of migration. *Genetics*, 189(2):633–644.

Waples, R. S., Masuda, M., and Pella, J. (2006). salmonnb: a program for computing cohort-specific effective population sizes (Nb) in Pacific salmon and other semelparous species using the temporal method: PROGRAM NOTE. *Molecular Ecology Notes*, 7(1):21–24.

Waples, R. S. and Yokota, M. (2006). Temporal estimates of effective population size in species with overlapping generations. *Genetics*, 175(1):219–233.

Weinberg, W. (1908). Über den Nachweis der Vererbung beim Menschen (English translations in Boyer 1963 and Jameson 1977. *Jahresh. Ver. Vaterl. Naturkd. Württemb.*, (64):369–382.

Weir, B. and Cockerham, C. (1984). Estimating f-statistics for the analysis of population structure. *Evolution*, 38(06):1358–1370.

Weir, B. S. (1979). Inferences about Linkage-Disequilibrium. *Biometrics*, 35:235–254.

Weir, B. S. (1996). *Genetic data analysis II: methods for discrete population genetic data*. Sinauer Associates, Sunderland, Mass.

Wilson, C. C., McDermid, J. L., Wozney, K. M., Kjartanson, S., and Haxton, T. J. (2014). Genetic estimation of evolutionary and contemporary effective population size in lake sturgeon (*Acipenser fulvescens* Rafinesque, 1817) populations. *Journal of Applied Ichthyology*, 30(6):1290–1299.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, (16):97–159.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress on Genetics*, pages 355–366.

Wright, S. (1949). The genetical structure of populations. *Annals of Eugenics*, 15(1):323–354.

Wright, S. (1950). Genetical Structure of Populations. *Nature*, 166(4215):247–249.

Wright, S. (1965). The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating. *Evolution*, 19(3):395.

Wright, S. (1984). *Variability within and among natural populations*. Number Sewall Wright ; Vol. 4 in Evolution and the genetics of populations. Univ. of Chicago Press, Chicago, Ill., paperback ed edition.

Yeh, F. and Boyle, T. (1997). Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany*, 129(157).

Zhao, J. (2015). gap: Genetic Analysis Package. R package version 1.1-16.

Zhdanova, O. L. and Pudovkin, A. I. (2008). Nb_hetex: A Program to Estimate the Effective Number of Breeders. *Journal of Heredity*, 99(6):694–695.

Zhivotovsky, L. A., Yurchenko, A. A., Nikitin, V. D., Safronov, S. N., Shitova, M. V., Zolotukhin, S. F., Makeev, S. S., Weiss, S., Rand, P. S., and Semenchenko, A. Y. (2015). Eco-geographic units, population hierarchy, and a two-level conservation strategy with reference to a critically endangered salmonid, Sakhalin taimen Parahucho perryi. *Conservation Genetics*, 16(2):431–441.

ANNEXE B

Principales raies du golfe de Gascogne : fiches synthétiques des traits d'histoires de vie et pêche

RAIE BOUCLÉE

Dénominations

Français : raie bouclée
Latin : *Raja clavata*
Anglais : thornback ray

Photo : F. Marandel

Quelques données disponibles

Taille maximum observée

91,3 cm (Serra Pereira et al. 2008)
104 cm (Gallagher et al. 2005)
89 cm (Serra Pereira et al. 2005)
94.5 cm (Whittamore and McCarthy 2005)
107 cm (Holden 1972)

Age de maturité théorique

7,3 (Wiegand et al. 2011)
6,13 (Gallagher et al. 2005)
4,6 (Whittamore and McCarthy 2005)

Croissance (Von Bertalanffy)

Linf = 118 ; K=0,155 ; t0=-0,655
(Wiegand et al. 2011)
Linf = 130,5 ; K=0,153; t0=-0,654
(Serra Pereira et al. 2008)

Relation Taille-Poids

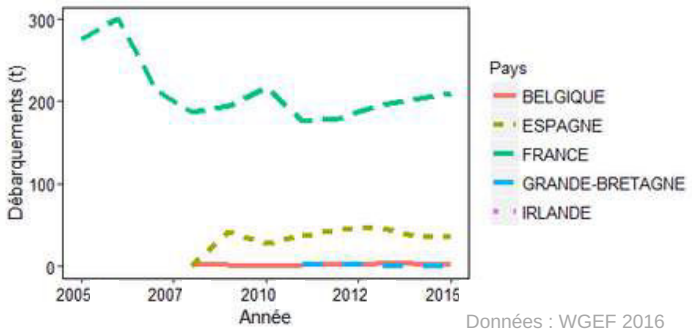
$$Poids_t = a * Taille_t{}^b$$

a=0,0039 ; b=3,1034
(McCully et al. 2012)
a=00035 ; b=3,1807
(Dorel et al. 1986)

Statut IUCN

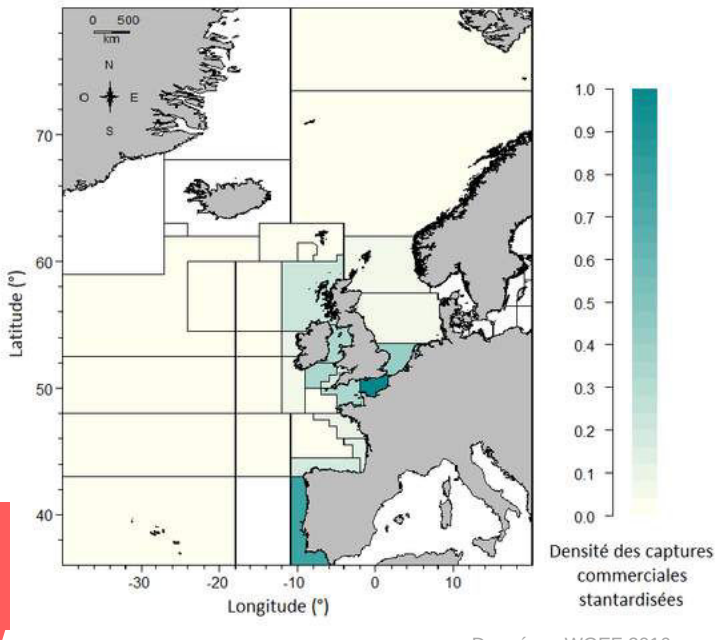
Quasi menacée

Débarquements (Golfe de Gascogne)



Données : WGEF 2016

Occurrence géographique



Données : WGEF 2016

Florianne Marandel

RAIE CHARDON

Dénominations

Français : raie chardon
Latin : *Leucoraja fullonica*
Anglais : shagreen ray

Photo : Ifremer

Quelques données disponibles

Taille maximum observée

90 cm (McCully et al. Unpublished)
96 cm (Mc Cully et al. 2012)
103 cm (Du Buit 1974)

Relation Taille-Poids

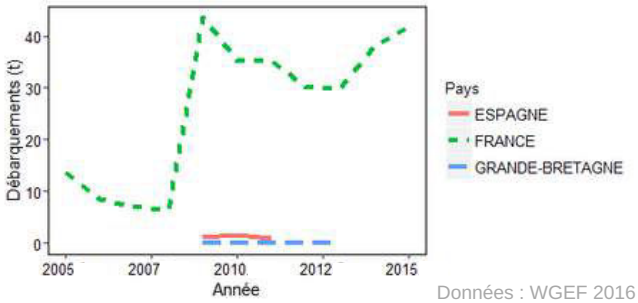
$$Poids_t = a * Taille_t{}^b$$

a=0,0025 ; b=3,196
(McCully et al. 2012)

Statut IUCN

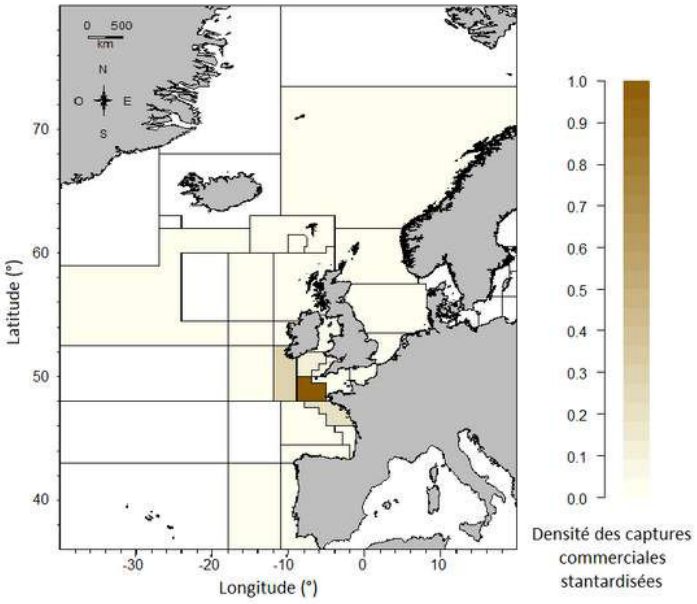
Vulnérable

Débarquements (Golfe de Gascogne)



Données : WGEF 2016

Occurrence géographique



Données : WGEF 2016

Florianne Marandel

RAIE CIRCULAIRE

Dénominations

Français : raie circulaire
Latin : *Leucoraja circularis*
Anglais : sandy ray

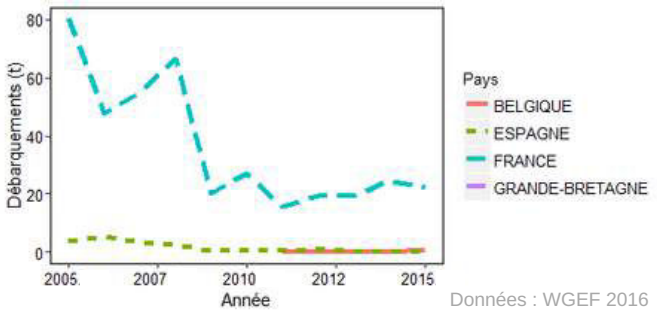
Photo : Quéro et Vayne 1997

Quelques données disponibles

Taille maximum observée

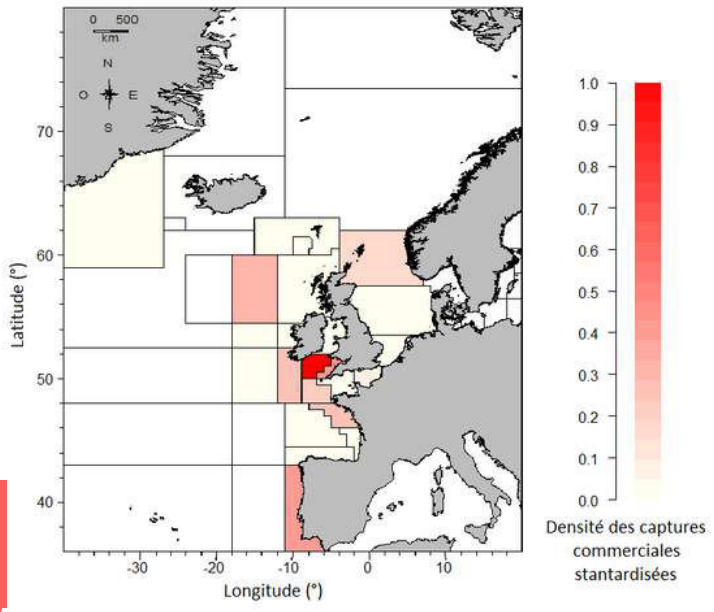
80 cm (Ebert and Stehmann 2013)
120 cm (Stehmann 1990)

Débarquements (Golfe de Gascogne)



Données : WGEF 2016

Occurrence géographique



Données : WGEF 2016

Statut IUCN

En voie d'extinction

Florianne Marandel

RAIE DOUCE

Dénominations

Français : raie douce
Latin : *Raja montagui*
Anglais :spotted ray

Photo : E. Farrell

Quelques données disponibles

Taille maximum observée

77 cm (Gallagher et al. 2005)
74 cm (Walker 1999)
70 cm (Holden 1972)

Age de maturité théorique

3,77 (Gallagher et al. 2005)
6,51 (Walker 1999)

Croissance (Von Bertalanffy)

Linf = 75,4 ; K=0,277 ; t0=-0,952
(Gallagher et al. 2005)
Linf = 77,1 ; K=0,201; t0=-1,471
(Walker 1999)
Linf = 70,7 ; K=0,185; t0=-0,464
(Holden 1972)

Relation Taille-Poids

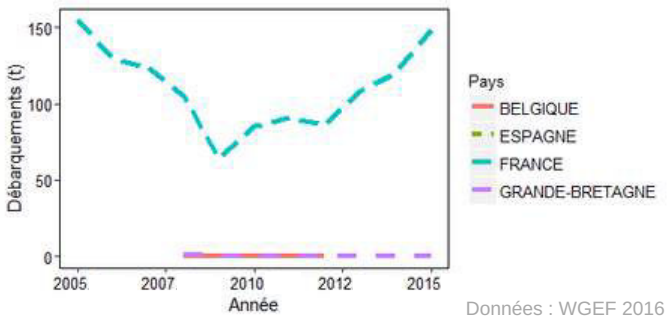
$$Poids_t = a * Taille_t{}^b$$

a=0,0037 ; b=3,147
(McCully et al. 2012)
a=00020 ; b=3,312
(Dorel et al. 1986)

Statut IUCN

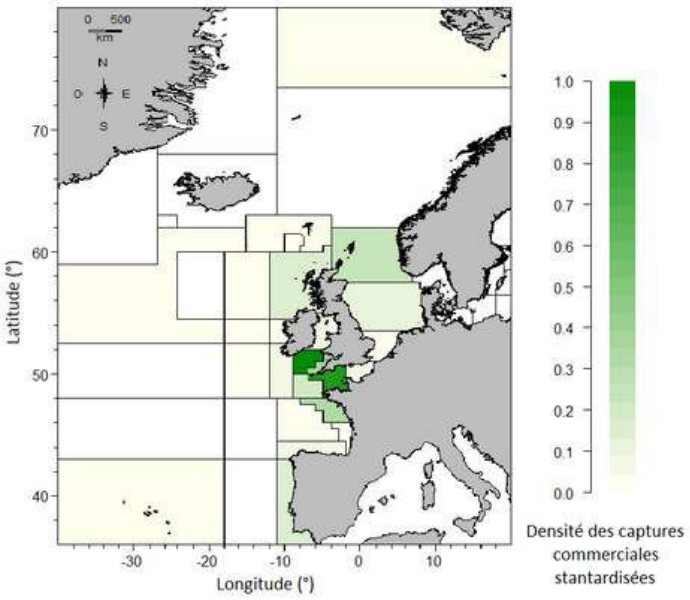
Préoccupation mineure

Débarquements (Golfe de Gascogne)



Données : WGEF 2016

Occurrence géographique



Données : WGEF 2016

Florianne Marandel

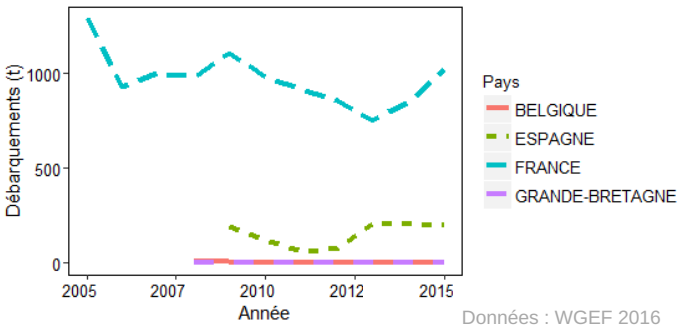
RAIE FLEURIE

Dénominations

Français : raie fleurie
Latin : *Leucoraja naevus*
Anglais : cuckoo ray

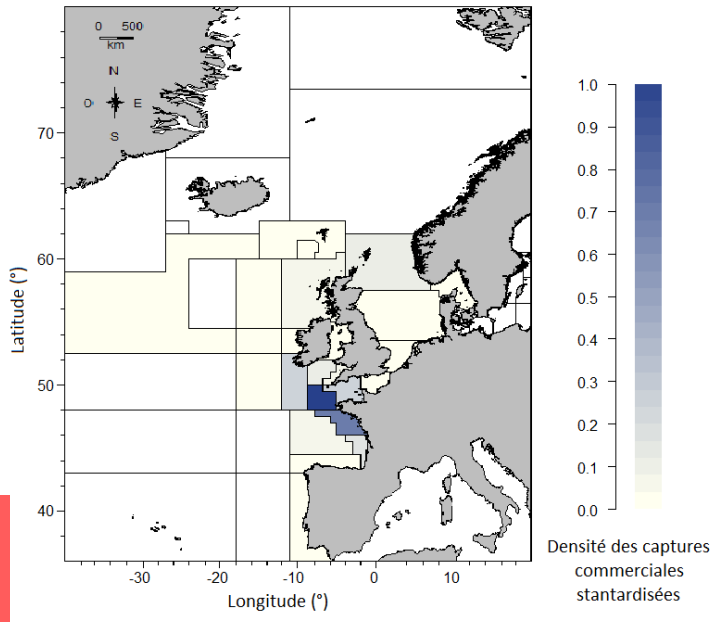
Photo : Ifremer

Débarquements (Golfe de Gascogne)



Données : WGEF 2016

Occurrence géographique



Données : WGEF 2016

Quelques données disponibles

Taille maximum observée

72 cm (Maia et al. 2012)
71 cm (Gallagher et al. 2005)
66 cm (Walker 1999)
72 cm (Du Buit 1972)

Age de maturité théorique

4,21 (Gallagher et al. 2005)
7,1 (Walker 1999)

Croissance (Von Bertalanffy)

Linf = 79 ; K=0,246 ; t0=-0,574
(Gallagher et al. 2005)
Linf = 70,4 ; K=0,217; t0=-1,016
(Walker 1999)

Relation Taille-Poids

$$Poids_t = a * Taille_t{}^b$$

a=0,009 ; b=3,105
(McCully et al. 2012)
a=0,005 ; b=3,1864
(Dorel et al. 1986)

Statut IUCN

Préoccupation mineure

Florianne Marandel

RAIE LISSE

Dénominations

Français : raie lisse
Latin : *Raja brachyura*
Anglais : blonde ray

Photo : Fishbase

Quelques données disponibles

Taille maximum observée

109 cm (Gallagher et al. 2005)
115 cm (Holden 1972)

Age de maturité théorique

5,07 (Gallagher et al. 2005)

Croissance (Von Bertalanffy)

Linf = 150,1 ; K=0,137 ; t0=-0,885
(Gallagher et al. 2005)
Linf = 116,7 ; K=0,190; t0=-0,485
(Holden 1972)

Relation Taille-Poids

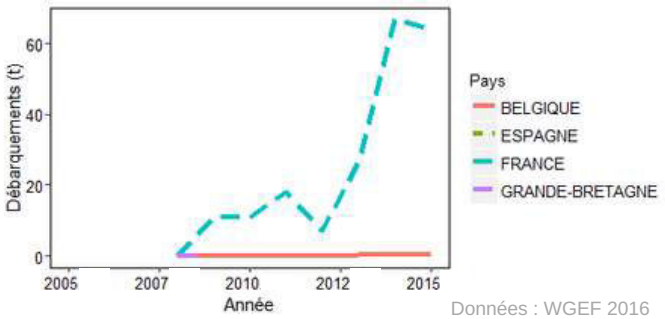
$$Poids_t = a * Taille_t^{\,b}$$

a=0,0027 ; b=3,2635
(McCully et al. 2012)
a=0,0028 ; b=3,2334
(Dorel et al. 1986)

Statut IUCN

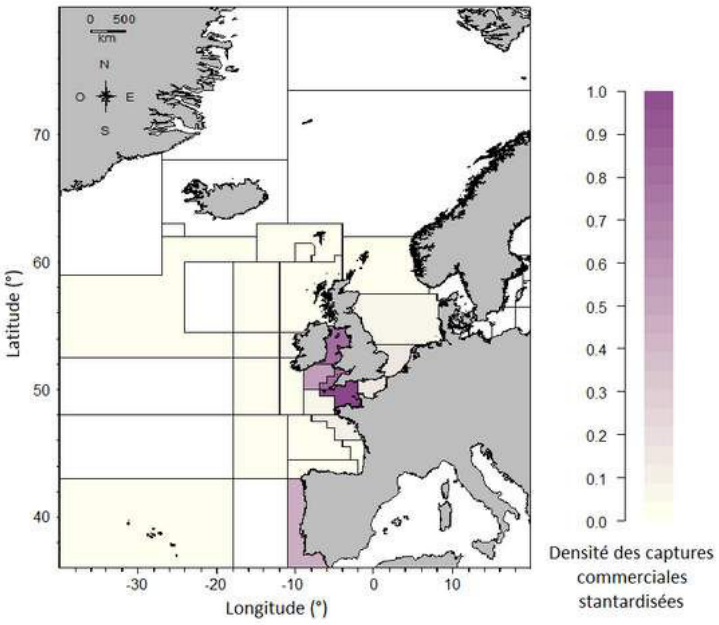
Quasi menacée

Débarquements (Golfe de Gascogne)



Données : WGEF 2016

Occurrence géographique



Données : WGEF 2016

Florianne Marandel

ANNEXE C

Evaluation de l'impact des contaminants chimiques sur deux Rajidae méconnues

NOAA

Marine Pollution Bulletin 123 (2017) 387–394

Contents lists available at ScienceDirect

Marine Pollution Bulletin

journal homepage: www.elsevier.com/locate/marpolbul

Baseline

Concentrations of mercury and other trace elements in two offshore skates: sandy ray *Leucoraja circularis* and shagreen ray *L. fullonica*

E.E. Manuel Nicolaus[a,*], Jon Barry[a], Thi P.C. Bolam[a], Pascal Lorance[b], Florianne Marandel[b], Sophy R. McCully Phillips[a], Suzanna Neville[a], Jim R. Ellis[a]

[a] Cefas, Pakefield Road, Lowestoft, Suffolk NR33 0HT, United Kingdom
[b] IFREMER, Unité Ecologie et Modèles pour l'Halieutique, B.P. 21105, 44311 Nantes Cedex 03, France

ARTICLE INFO

Keywords:
Trace metals
Sandy ray
Shagreen ray
Bay of Biscay
Celtic Sea

ABSTRACT

Trace metal concentrations in muscle and liver tissues from two offshore species of skate were examined. Concentrations of mercury in muscle of *Leucoraja circularis* (n = 20; 23–110.5 cm total length, 157–490 m water depth) and L. *fullonica* (n = 24; 28.5–100 cm total length, 130–426 m water depth) were 0.02–1.8 and 0.04–0.61 mg kg^{-1}, respectively. Concentrations of both As and Hg increased with total length. Only the largest specimen had a concentration of Hg in muscle > 1.0 mg kg^{-1}. Data were limited for specimens > 90 cm long, and further studies on contaminants in larger-bodied skates could usefully be undertaken.

Skates (Rajformes) are demersal elasmobranchs that are widespread in shelf seas and deep-water habitats. This speciose order includes approximately 290 species (Last et al., 2016), that range up to ca. 250 cm total length. In terms of their trophic position, the order contains many species that are benthic predators, whilst some are more piscivorous, including some species that predate on other elasmobranchs. The estimated trophic levels of skates range from 3.48–4.22 (Ebert and Bizzarro, 2007). As some species are of low market value, skates are an important group of commercial fish in many parts of the world, including Europe. Given their demersal habitat and potential longevity, they may bioaccumulate various contaminants, such as mercury and persistent organic contaminants (Nicolaus et al., 2016a; Lyons and Adams, 2017). Trace elements reach the marine environment via anthropogenic and natural inputs (Nicolaus et al., 2016b) and ultimately bind to sediments due to their strong affiliation with particulate matter (Zhang et al., 2007). Consequently, demersal fish that may forage and bury in upper surficial sediments, such as skates, may be exposed to trace metals in sediments.

Various studies have examined the contaminants of skates from the inner continental shelf of European seas (Dixon and Jones, 1994; De Gieter et al., 2002; Storelli and Barone, 2013) and elsewhere, but data are limited for those skates living in deeper water (Mormede and Davies, 2001), despite there being some evidence that Hg concentrations in marine fish can increase with water depth (Choy et al., 2009).

Sandy ray *Leucoraja circularis* and shagreen ray L. *fullonica* are two offshore skates that are widespread along the edge of the continental shelf from Iceland and northern Norway to north-west Africa, including the Mediterranean Sea (Ebert and Stehmann, 2013). Despite their broad distribution range, the offshore nature of these two species means there are very few published biological investigations (Du Buit, 1972; Consalvo et al., 2009; Mnsari et al., 2009). Currently, the life histories of these two species are poorly known, they are both relatively large-bodied species, reportedly attaining maximum lengths of 120 cm. Both species predate on crustaceans and fish (Du Buit, 1972), with L. *fullonica* (trophic level = 4.6) also predating on other elasmobranchs (Ebert and Bizzarro, 2007).

The United Kingdom's Clean Seas Environmental Monitoring Programme (CSEMP) samples flatfish species, mostly at sites within 22 km (12 nautical mile) of shore, to assess spatial and temporal concentrations of contaminants in fish liver and muscle (Nicolaus et al., 2016b). The dab *Limanda limanda* is a useful indicator species for inshore waters, but it does not occur along the edge of the continental shelf. To better understand the concentrations of trace elements in offshore fish, samples of muscle and liver from L. *circularis* and L. *fullonica* caught in the Bay of Biscay and Celtic Sea were analysed.

Specimens of L. *circularis* (n = 20) and L. *fullonica* (n = 24) were caught during annual trawl surveys of the Bay of Biscay and Celtic Sea (EVHOE survey: ÉValuation Halieutique de l'Ouest de l'Europe; see Mahé and Poulard, 2005 and ICES, 2010) in 2014–2016. These specimens were collected from the deeper parts of the survey area (Fig. 1) in waters of 130–490 m depth. Specimens were frozen whole for subsequent collection of biological data (total length (L_T), disc width, total

* Corresponding author.
E-mail address: manuel.nicolaus@cefas.co.uk (E.E.M. Nicolaus).

http://dx.doi.org/10.1016/j.marpolbul.2017.08.054
Received 27 June 2017; Received in revised form 9 August 2017; Accepted 23 August 2017
Available online 30 August 2017

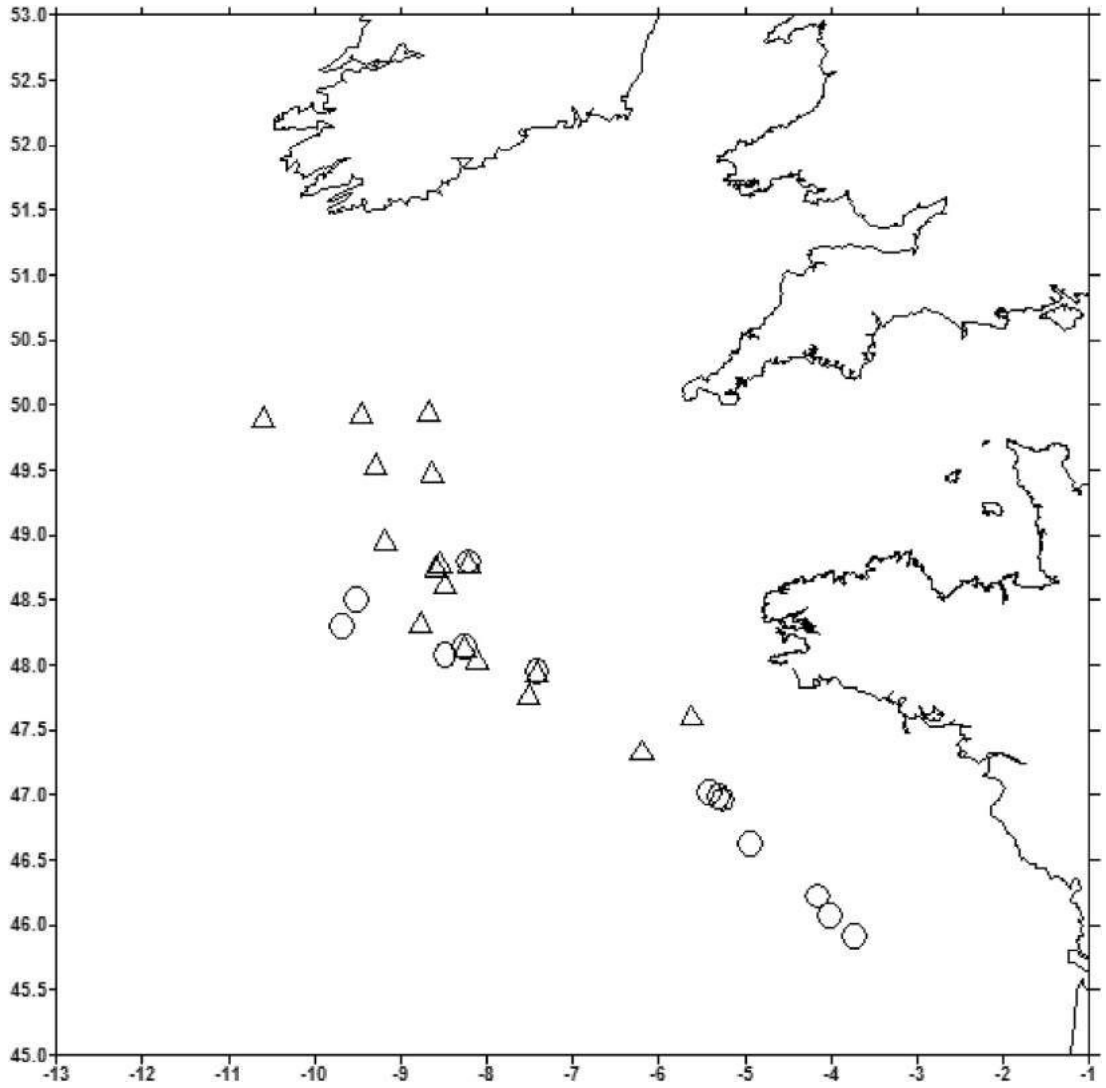


Fig. 1. Sampling locations of *Leucoraja circularis* (open circles) and L. *fullonica* (open triangles) in the Bay of Biscay and Celtic Sea.

weight, sex and maturity) and tissue sampling. Sections of dorsal muscle, excluding skin and ceratotrichia, were excised, and either the whole liver (small specimens) or sub-samples from the three main lobes taken. These samples were then re-frozen until analysed for trace elements.

Trace metal analysis followed standard procedures (Jones and Laslett, 1994). Tissue samples underwent an acid digestion using an enclosed vessel microwave (Multiwave 3000, Anton Paar, Hertford, UK). Typically, approximately 1 g of homogenised sample was weighed out and pre-digested overnight in 6 mL of nitric acid (Aristar grade 69%, VWR, Leicestershire, UK). The digestion was performed using a temperature-controlled microwave programme specific for the sample matrix. The digest was diluted further prior to analysis by inductively coupled plasma-mass spectrometry (ICP-MS) using an Agilent 7500ce (Agilent Technologies, Waldbronn, Germany). Quantification of the trace elements was performed by external calibration and deploying eight levels (0, 0.5, 1, 5, 10, 20, 100 and 500 $\mu g\,L^{-1}$) of working standard solutions which were prepared from a customised mixed metal standard solution of 100 $mg\,L^{-1}$ (SPEX Certiprep Ltd., Middlesex, UK).

To ensure a high level of quality assurance, a reagent blank and a certified reference material (CRM TORT-2-Lobster hepatopancreas, National Research Council Canada, Halifax, Nova Scotia, Canada) was analysed within-batch to monitor method performance on a day-to-day basis. Concentration data derived from the analysis of the CRM were then added to existing quality control Shewhart charts (using North West Analytical Quality Analyst™, Northwest Analytical Inc., USA) for the assessment of the on-going method performance from the batch analysis of real samples. The validity of results was established using the warning and control limits of the Shewhart chart, which are defined as 2σ and 3σ of the mean, respectively.

In addition to internal quality control, the analytical laboratory biannually participates in the proficiency testing scheme Quasimeme (Quality Assurance of Information for Marine Environmental Monitoring in Europe) as external quality assurance. A summary of the accuracy of the analytical methods is provided in the Supplementary Material (Table S1).

To carry out an effect-based assessment, the approach of Nicolaus et al. (2015) and Nicolaus et al. (2017) was used, which compares the measured environmental concentrations (MEC) for set determinants to derived assessment criteria (Table 1). A risk characterisation ratio

Table 1
Assessment criteria used to analyse the contaminant status of *Leucoraja circularis* and L. *fullonica* for concentrations (mg kg^{-1}) of Cd, Pb and Hg in muscle (based on European regulations on the maximum levels in foodstuffs (CEC, 2006)) and Cd and Pb in liver (based on the preliminary OSPAR, 2009 indicators of environmental quality, whereby it was suggested using the statutory dietary limits of Cd and Pb in bivalves as proxy thresholds for concentrations in fish liver). Note: CEC (2006) lists Hg limits for "rays (*Raja* species)" as 1.0, and this is the assumed level used in the current analysis, as both species were formerly in the genus *Raja*.

Metal	Threshold concentration (mg kg^{-1})		Percentage of *Leucoraja circularis* exceeding limits		Percentage of L. *fullonica* exceeding limits	
	Muscle	Liver	Muscle	Liver	Muscle	Liver
Cd	0.05	1	0%	5%	0%	0%
Pb	0.3	1.5	0%	0%	0%	0%
Hg	1.0	–	5%	–	0%	–

(RCR) was then calculated by dividing the MEC by either the limits defined in European Commission Regulations (proxy Environmental Assessment Criteria - EAC) that cite the safe maximum levels of contaminants in seafood (above which the concentration may be harmful to human health (CEC, 2006)) or the thresholds suggested by OSPAR (2009) for indicators of environmental quality.

It was also assessed whether there were differences in contaminant concentrations at length (L_T) between the two species for three contaminants: As, Se and Hg. The models used for As and Se were linear models (contaminant concentration = $a + b\ L_T$ + error), with a non-linear model used for Hg (contaminant concentration = $a\ {L_T}^b$ + error), where a and b are parameters estimated by least squares and the error is assumed to be Normally distributed with mean 0 and constant variance. To compare the concentration for the species across all lengths, the fit to the data was compared (as defined by the residual sum of squares RSS1) when a model was fitted to each species separately against the fit when a single model was fitted to all data (RSS0). An F-test was used to evaluate a p-value based on the differences between RSS0 and RSS1 (Mead and Curnow, 1984).

Concentrations of 11 trace elements (As, Cd, Cr, Cu, Fe, Hg, Mn, Pb, Ni, Se and Zn) were analysed in the muscle and liver of L. *circularis* (n = 20; 23–110.5 cm total length, 157–490 m water depth) and L. *fullonica* (n = 24; 28.5–100 cm total length, 130–426 m water depth). Data (Table 2, and Tables S2–S3 for raw data) indicate that both Pb and Cr occurred in low concentrations in the muscle, with 52.3% and 54.5% of samples (species combined) below the detection limits for Pb and Cr, respectively. For those specimens with detectable limits, the mean concentrations of Cr and Pb were 0.010 and 0.011 mg kg^{-1}, respectively (species combined).

Concentrations of As in the muscle ranged from 4.9–95 and 22–141 mg kg^{-1} in L. *circularis* and L. *fullonica*, respectively, and there was a significant difference in the concentration of As between the two species with length ($F = 161.8$, df = 4,40, $p < 0.001$). Whilst concentrations of As increased significantly with length (Fig. 2), data were limited for skates > 90 cm L_T.

Concentrations of Hg in muscle were 0.02–1.8 and 0.04–0.61 mg kg^{-1} in L. *circularis* and L. *fullonica*, respectively (Table 2), and increased with L_T (Fig. 2). Only one specimen (a 110.5 cm L_T L. *circularis*) had a concentration of Hg in muscle > 1.0 mg kg^{-1} (and had a Hg concentration of 4.1 mg kg^{-1} in the liver). There is a need to examine more samples of skates > 90 cm L_T to better understand the proportion of large skate specimens that might exceed EC regulations on safe maximum levels of contaminants. For Hg, there was also a significant difference between the two species at different lengths ($F = 452.7$, df = 4,40, $p < 0.001$; Fig. 2).

Concentrations of Se in muscle ranged from 0.36–0.64 and 0.29–0.4 mg kg^{-1} in L. *circularis* and L. *fullonica*, respectively, and concentrations were significantly different between the two species at different lengths ($F = 348.2$, df = 4,40, $p < 0.001$; Fig. 2).

Whilst sample sizes were relatively limited, there was the indication that concentrations (at length) of both Hg and Se were higher in L. *circularis*, and As higher in L. *fullonica*, although the reasons for this are unclear.

Comparing the results to the available assessment criteria, one specimen of L. *circularis* failed the proxy EAC for Cd in liver, no samples failed the proxy EAC for Hg or Pb in liver. For Hg in the muscle tissue, only one specimen of L. *circularis* failed the proxy EAC (1.0 mg kg^{-1}).

By analysing the Hg:Se ratio, it became apparent that only one sample for *Leucoraja circularis* showed a ratio above one, indicating that the organism itself could suffer from physiological impacts (Nicolaus et al., 2016a). If the ratio is above one, it means that there are more Hg

Table 2
Mean concentrations (mg kg^{-1}; ± SD and range) of 11 trace elements in muscle and liver tissues of *Leucoraja circularis* (n = 20) and L. *fullonica* (n = 24) [samples below detection limit concentrations were omitted from the summary data, which explains the lower sample sizes for some trace elements].

Species	Trace element	Liver			Muscle		
		Mean ± SD	Range	N	Mean ± SD	Range	N
L. circularis	As	16.77 ± 6.41	6.4–34	20	46.65 ± 19.38	4.9–95	20
	Cd	0.39 ± 0.56	0.09–2.7	20	0.01 ± 0.00	0.01–0.01	7
	Cr	0.03 ± 0.04	0.01–0.17	20	0.06 ± 0.05	0.01–0.19	20
	Cu	7.34 ± 4.15	2.3–20	20	0.26 ± 0.11	0.13–0.57	20
	Fe	89.15 ± 60.24	39–301	20	2.01 ± 0.99	0.96–5.4	20
	Pb	0.02 ± 0.01	0.01–0.05	9	0.01 ± 0.00	0.01–0.02	10
	Mn	0.44 ± 0.17	0.26–1	20	0.29 ± 0.24	0.08–1.0	20
	Hg	0.33 ± 0.89	0.01–4.1	20	0.43 ± 0.38	0.02–1.8	20
	Ni	0.03 ± 0.03	0.01–0.1	18	0.04 ± 0.03	0.01–0.14	17
	Se	1.13 ± 0.63	0.65–3.7	20	0.45 ± 0.08	0.36–0.64	20
	Zn	10.65 ± 2.60	7.9–19	20	4.73 ± 1.17	3.4–7.8	20
L. fullonica	As	19.33 ± 5.77	4.9–35	24	47.88 ± 24.15	22–141	24
	Cd	0.17 ± 0.05	0.08–0.26	24	0.01 ± 0.00	0.01–0.01	13
	Cr	0.04 ± 0.03	0.01–0.16	22	0.11 ± 0.08	0.01–0.37	24
	Cu	3.00 ± 1.72	1.4–8.4	24	0.32 ± 0.09	0.17–0.48	24
	Fe	40.63 ± 24.46	15–113	24	2.38 ± 0.96	1.2–5	24
	Pb	0.01 ± 0.00	0.01–0.02	8	0.01 ± 0.00	0.01–0.02	11
	Mn	0.59 ± 0.13	0.35–0.84	24	0.30 ± 0.16	0.11–0.61	24
	Hg	0.04 ± 0.03	0.01–0.12	23	0.13 ± 0.14	0.04–0.61	24
	Ni	0.04 ± 0.03	0.01–0.16	22	0.08 ± 0.07	0.02–0.36	23
	Se	0.96 ± 0.25	0.16–1.3	24	0.35 ± 0.03	0.29–0.4	24
	Zn	10.71 ± 5.33	4.5–34	24	4.91 ± 1.06	3.4–6.7	24

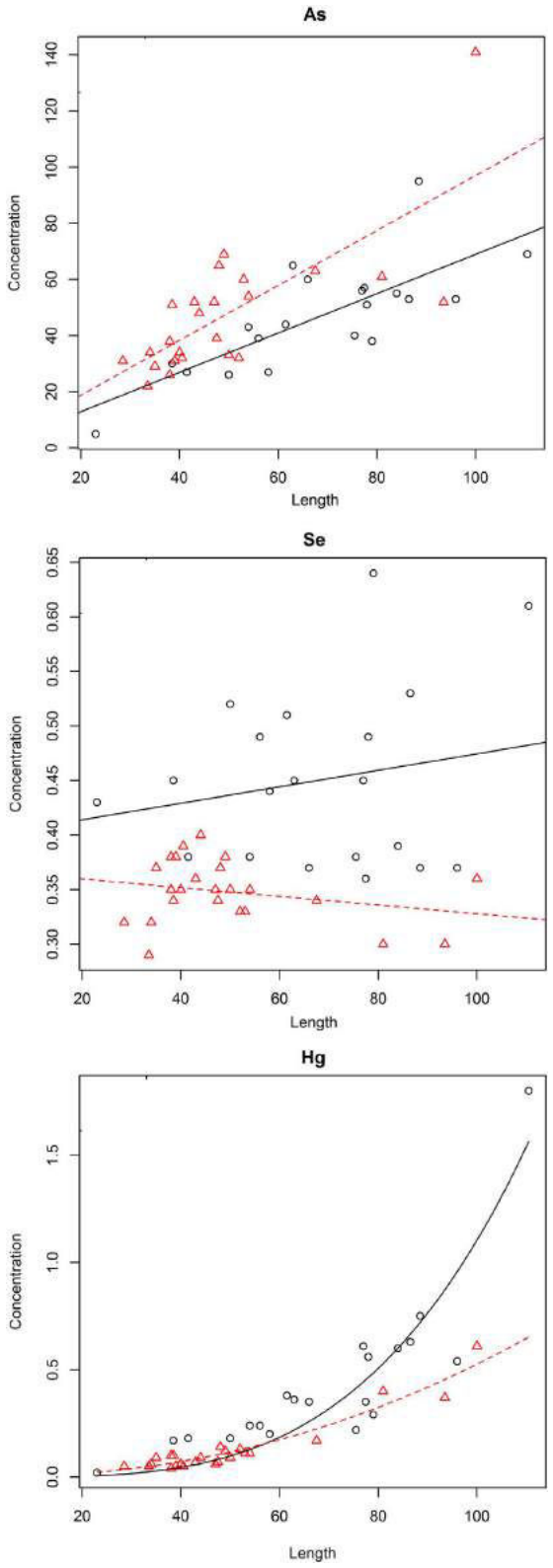


Fig. 2. Concentrations of As, Hg and Se (mg kg^{-1}) in muscle in relation to total length in *Leucoraja circularis* (SAR, circles) and L. *fullonica* (SHR, triangles) from the Bay of Biscay and Celtic Sea. The relationship between concentrations and length in L. *circularis* were As = $-1.167 + 0.701L_T$, Se = $0.399 + 0.000755L_T$, and Hg = $1.156 * 10^{-7}L_T^{3.49}$. The relationship between concentrations and length in L. *fullonica* were As = $-0.846 + 0.979L_T$, Se = 0.386–$0.000396L_T$, and Hg = $2.462 * 10^{-5}L_T^{2.165}$.

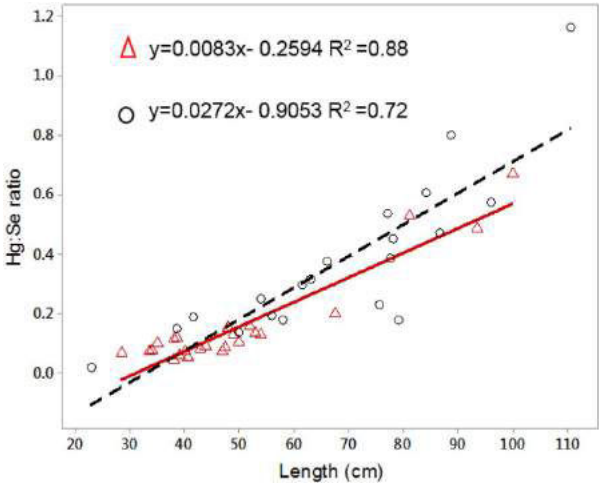


Fig. 3. Relationship between the Hg:Se ratio and total length in the muscle of *Leucoraja circularis* (circles; dashed trendline) and L. *fullonica* (triangles; solid trendline).

moles than Se moles present in the individual, which suggests that methylmercury exposure could limit selenoprotein synthesis, which is part in the routine functioning of enzymes in multiple biological processes (EERC, 2011). There was a positive and similar relationship between the Hg:Se ratio in both species (Fig. 3).

Mean concentrations of other trace elements (Ni, Cu, Zn, Mn and Fe) were low (Table 2). The concentrations of Cu in muscle were all at detectable limits, the highest concentrations were only 0.57 and 0.48 mg kg^{-1} for L. *circularis* and L. *fullonica*, respectively. Furthermore, the mean Cu concentrations in the muscle were only 0.26 (*L. circularis*) and 0.32 (*L. fullonica*) mg kg^{-1}, which are generally lower than the concentrations reported in shallower-water fish species (Collings et al., 1996; Mormede and Davies, 2001).

Monitoring programmes examining metals and other contaminants in fish and shellfish around the UK provided data for a range of skate species during the 1970s (Murray, 1979, 1981; Murray and Portmann, 1984), with thornback ray *Raja clavata* also examined regularly into the 1990s (Franklin, 1987; Franklin and Jones, 1995). Since then, however, monitoring efforts have focused on a smaller groups of core species, with dab presently the main indicator species studied (Nicolaus et al., 2016b).

Other previous studies on contaminants in skates from European seas have generally focused on the more coastal and more widespread *R. clavata* (Dixon and Jones, 1994; Chouvelon et al., 2012; Türkmen et al., 2013; Torres et al., 2016), a species that has also been used in experimental studies of metal accumulation (Pentreath, 1973, 1976, 1977a, 1977b). In contrast, the levels of contaminants in other skate species have been subject to more limited study (Table 3).

This is the first published study on the trace element concentrations in L. *fullonica*, and in L. *circularis* from Atlantic waters, and results indicated that concentrations of Hg could exceed 1.0 mg kg^{-1} in larger specimens. Storelli et al. (1998) reported a mean Hg concentration of 1.47 mg kg^{-1} in muscle of L. *circularis* from the southern Adriatic Sea, with this based on two pooled samples from 10 specimens. The case-study species are both members of the outer shelf and upper slope fish assemblage, and neither species are frequent in shallow waters.

Table 3
Published studies on mercury in the muscle of skates (Rajiformes). Scientific names updated as per Last et al. (2016) and trophic levels (TL) as reported by Ebert and Bizzarro (2007). Fish size refers to total length, unless specified otherwise (D = disc width) with values in parentheses indicating mean size. Hg concentrations (mg kg^{-1}) refer to total mercury concentrations by wet weight, unless specified otherwise. Notes: [1] Concentration of Hg related to dry weight; [2] Concentration of Hg refers to methylmercury; [3] Concentration of Hg refers to whole animal, not specific tissues; [4] Concentration of Hg reported to increase with length; [5] Study provided data for other tissues; [6] Study provided data for other trace elements or contaminants; [7] Species identification might be questionable, given location of capture and/or maximum size; [8] Concentrations refer to mean and 95% CI.

Scientific name	TL	Geographic area	Year	Depth range (m)	Sample size (number of pools)	Size range (cm)	Hg concentration mean ± SD (range)	Notes	Source
Amblyraja radiata	3.82	Nova Scotia	1970	–	10	–	0.03–0.24	[2]	Zitko et al. (1971)
		Halifax	1972	–	10	–	0.12–0.41	–	Freeman et al. (1974)
		Georges Bank	1971	27–113	– (2)	60–93	0.24 (0.21–0.26)	[5]	Greig et al. (1975)
		North Sea	1975	–	10	–	0.09 (0.06–0.17)	[6]	Murray (1981)
		Barents Sea	1994	–	5	–	0.72 ± 0.70	[1,5,6,8]	Zauke et al. (1999)
Dentiraja cerva	–	Tasmania	1976–1977	–	32	42.2–59.7 (D)	0.31 ± 0.04 (0.03–0.85)	–	Thomson (1985)
Dipturus batis	4.06	Irish Sea	1992–1993	–	8	(72.8)	1.2 ± 1.1 (0.27–3.14)	[5,6,7]	Collings et al. (1996)
Dipturus kwangtungensis	–	East China Sea	2001–2002	131–133	3	(35)	< 0.05	[1,3,6]	Asante et al. (2008)
Dipturus oxyrinchus	–	Southern Adriatic Sea	1995	–	12 (3)	–	1.56 ± 0.95 (1.00–2.65)	–	Storelli et al. (1998)
Leucoraja circularis	–	Southern Adriatic Sea	1995	–	10 (3)	–	1.47 ± 1.12 0.68–2.27	–	Storelli et al. (1998)
		Bay of Biscay and Celtic Sea	2014–2016	157–490	20	23–110.5	0.43 ± 0.38 (0.02–1.8)	[4,5,6]	This study
Leucoraja erinacea	3.70	George's Bank	1971	24–113	–	45–50	0.15 (0.13–0.16)	[5]	Greig et al. (1975)
		Southern New England	2009–2012	–	173	21.5–29.5 (D)	0.4 ± 0.3	[1,4]	Taylor et al. (2014)
Leucoraja fullonica	4.06	Bay of Biscay and Celtic Sea	2014–2016	130–426	24	28.5–100	0.13 ± 0.14 (0.04–0.61)	[4,5,6]	This study
Leucoraja naevus	3.91	North Sea (Aberdeen)		–	10	–	0.054 ± 0.029 (0.02–0.11)	[6]	Davies (1981)
		Irish Sea	1976	–	2	–	0.35 (0.28–0.41)	[6]	Murray and Portmann (1984)
		Bay of Biscay	2001–2010	120–199	10	60.4 ± 2.8	0.569 ± 0.239 (0.396–1.205)	[1]	Chouvelon et al. (2012)
Leucoraja ocellata	4.04	George's Bank	1971	18–62	(1)	–	0.15	[5]	Greig et al. (1975)
		Southern New England	2009–2012	–	148	20.5–54.7 (D)	0.3 ± 0.2	[1,4]	Taylor et al. (2014)
Okamejei kenojei	–	East China Sea	2001–2002	147–149	2	(39)	0.17	[1,3,6]	Asante et al. (2008)
Raja asterias	–	Southern Adriatic	1995	–	100 (20)	–	0.73 ± 0.47 (0.05–1.50)	–	Storelli et al. (1998)
Raja brachyura	3.82	Irish Sea	1976	–	7	–	0.18 (0.07–0.66)	[6]	Murray and Portmann (1984)
		Bristol Channel	1976	–	5	–	0.32 (0.17–0.71)	[6]	Murray and Portmann (1984)
Raja clavata	3.69	Sagres (Portugal)	1972	–	2	–	0.53	[1,5,6]	Stenner and Nickless (1975)
		North Sea	1975	–	7	–	0.16 (0.06–0.28)	[6]	Murray (1981)
		North Sea (coastal)	1975	–	10	–	0.16 (0.07–0.24)	[6]	Murray (1981)
		North Sea	1976	–	50	–	(0.03–2.0)	[5, 6]	Murray and Portmann (1984)
		Bristol Channel	1976	–	10	–	(0.06–0.38)	[5, 6]	Murray and Portmann (1984)
		Irish Sea	1976	–	18	–	(0.14–1.0)	[5, 6]	Murray and Portmann (1984)
		English Channel	1977	–	10	(41)	0.11 (0.09–0.13)	[6]	Franklin (1987)
		Morecombe Bay	1979	–	9	(55)	0.22 (0.12–0.37)	[6]	Franklin (1987)
		Liverpool Bay	1981	–	10	(38)	0.29	[6]	Franklin (1987)
		Bristol Channel	1983	–	9	(68)	0.39	[6]	Franklin (1987)
		Swansea Bay	1983	–	24	(53)	0.21	[6]	Franklin (1987)
		Southern North Sea	1983	–	14	(48)	0.15	[6]	Franklin (1987)
		English Channel	1990	–	18	–	0.07–0.10	[6]	Franklin and Jones (1995)
		Southern North Sea	1990–1992	–	8	–	0.05–0.11	[6]	Franklin and Jones (1995)
		North Sea	1992	–	22	41–50.6	0.097 (0.007–0.270)	[4]	Dixon and Jones (1994)
		Southern Adriatic	1995	–	84 (14)	–	1.24 ± 0.23 (0.86–1.60)	–	Storelli et al. (1998)

(continued on next page)

Table 3 *(continued)*

Scientific name	TL	Geographic area	Year	Depth range (m)	Sample size (number of pools)	Size range (cm)	Hg concentration mean ± SD (range)	Notes	Source
		North Sea and eastern Channel	1997–1999	–	19	–	0.039 ± 0.021	[5]	Baeyens et al. (2003)
							0.037 ± 0.019	[2,5]	
		Bay of Biscay	2001–2010	120–199	11	73.5 ± 11.1		1.02-1 ± 0.-816 (0.524–3.147)	[1]
		Chouvelon et al. (2012)							
Azores	20-13	–	30	40.5–79.5	0.14–1.47*	[1,4,5,6]	Torres et al. (2016)		
Raja eglanteria	3.68	S. Carolina-Florida	–	(inshore)	1	–	1.3	[1,5,6]	Windom et al. (1973)
		Delaware Bay	1975		3	25–27 (D)	0.214 ± 0.098 (0.119–0.321)		Gerhart (1977)
Raja microocellata	3.88	Bristol Channel	1976	–	5	–	0.15 (0.06–0.20)	[6]	Murray and Portmann (1984)
					5		0.07 (0.04–0.10)		
		Irish Sea	1976	–	3	–	0.37 (0.23–0.49)	[6]	Murray and Portmann (1984)
		Bay of Biscay	2001–2010	< 30	5	69.4 ± 9.9	0.169 ± 0.04 (0.128–0.217)	[1]	Chouvelon et al. (2012)
Raja miraletus	3.67	Southern Adriatic Sea	1995	–	40 (10)	–	1.10 ± 0.38 (0.60–1.78)	–	Storelli et al. (1998)
		Adriatic Sea	2010	–	127 (10 pooled samples)	33.5–48.5	1.09 ± 0.39 (0.40–1.78)	[4,6]	Storelli and Barone (2013)
		Adriatic Sea	2011	–	35 (pooled into 11 samples)	35.0–85.7	0.98 ± 0.33 (0.28–1.54)	[6,7]	Storelli et al. (2013)
Raja montagui	3.59	English Channel	1975	–	17	–	0.12 (0.06–0.19)	[6]	Murray (1981)
		Bristol Channel	1976	–	2	–	0.06–0.07	[6]	Murray and Portmann (1984)
		Irish Sea	1992–1993	–	6	(39.3)	0.3 ± 0.1	[5,6]	Collings et al. (1996)
Rajella fyllae	3.78	Rockall Trough	1998	850–950		38–54	0.044–0.410	[5,6]	Mormede and Davies (2001)
		Barents Sea	1994	–	5	–	1.35 ± 0.19	[1,5,6,8]	Zauke et al. (1999)
Rostroraja velezi	–	Mexico (Pacific)	2012	–	1	83 (D)	1.127	[1,5]	Ruelas-Inzunza et al. (2013)
		Costa Rica (Pacific)	2010–2011	< 100	19	17.5–55.4 (D)	0.25 ± 0.16 (0.01–0.50)	[1,4]	Sandoval-Herrera et al. (2016)
Sympterygia bonapartei	–	Bahia Blanca, Argentina	1985–1986	–	–	29–50.5	0.18 ± 0.06	[6]	Marcovecchio et al. (1988)
Zearaja chilensis	–	Chile	2009–2012	–	102	–	0.088 ± 0.05	[6]	Lopez et al. (2014)
Bathyraja spinicauda	4.02	Barents Sea	1994	–	3	–	0.58–1.21	[1,5,6,8]	Zauke et al. (1999)
Beringraja binoculata	–	Gulf of Alaska	2012–2013	–	20	92–175	0.09 ± 0.06 (up to 0.61)	[4,5,6]	Farrugia et al. (2015)
Beringraja rhina	–	Gulf of Alaska	2013	–	20	109–133	0.34 ± 0.18 (up to 0.61)	[4,5,6]	Farrugia et al. (2015)

Consequently, the concentrations of contaminants are not expected to be related to recent anthropogenic inputs. Observed contaminants levels, in particular Hg, are therefore likely to reflect ocean basin scale burden (from natural processes and longer-term, historical inputs over a continental scale), rather than local/regional input.

Most published studies on metals in skate muscle have indicated that Hg concentrations are often $< 1.0\ mg\ kg^{-1}$ (Table 3), Hg concentrations $> 1.0\ mg\ kg^{-1}$ wet weight have been observed in a range of skate species, including *R. clavata* (Storelli et al., 1998), long-nosed skate *Dipturus oxyrinchus* (Storelli et al., 1998), common skate *Dipturus batis* (Collings et al., 1996; although the timing and location of sampling infers a degree of doubt to the species, and it may refer to *R. clavata*), and brown ray *Raja miraletus* (Storelli et al., 1998, 2013). The latter study reported that 35% of samples of *R. miraletus* from the Adriatic Sea exceeded $1.0\ mg\ g^{-1}$ (wet weight). Other studies have reported Hg concentrations $> 1.0\ mg\ kg^{-1}$ dry weight, including for round skate *Rajella fyllae* and spinytail skate *Bathyraja spinicauda* (Zauke et al., 1999) and *R. clavata* (Chouvelon et al., 2012). Skates can also have high concentrations of other toxic contaminants, such as As (Collings et al., 1996; De Gieter et al., 2002). Consequently, further studies could usefully ascertain levels of contaminants of skates, focusing on areas of known contamination, and also on longer-lived (or larger-bodied) species.

Whilst some studies (e.g. Storelli and Barone, 2013; Lopez et al., 2014; Taylor et al., 2014) have used robust sample sizes, many previous studies examining the concentrations of metals in skates have been based on low sample sizes (see Table 3). Studies indicating high concentrations of contaminants but based on small sample sizes, should be used with a degree of caution, given the potential variability in metal concentrations.

Within Europe, EC Regulation 1881/2006 established a maximum level of $1.0\ mg\ kg^{-1}$ of Hg in *Raja* spp., presumably with the intent to refer to all members of the family, with the maximum concentrations of Pb and Cd being 0.30 and $0.05\ mg\ kg^{-1}$, respectively (CEC, 2006). Whilst the concentrations of Pb and Cd observed in the present study were all below these limits, concentrations of Hg can exceed maximum levels, and further studies of skates in either area of known historical contamination and of larger-bodied species could usefully be undertaken to better ascertain spatial and ontogenetic influences on Hg concentrations.

Skates in European seas are not aged routinely by national fisheries laboratories, and age estimates for the species studied here are not currently available. The increasing concentrations of contaminants in larger specimens may be related to both biomagnification (e.g. through eating higher trophic level prey) and/or bioaccumulation (Lyons and Adams, 2017). Older fish typically exhibit reduced growth rates, and so there can be more variability in the length at age for older fish. Given this, and that bioaccumulation of contaminants can be related more closely to age than length (Braune, 1987), there is a need to have appropriate sample sizes of fish in larger size classes.

Skates are large-bodied demersal species, often of commercial value, and the order Rajiformes has both a broad bathymetric range and an extensive biogeographic range, especially in temperate and subtropical seas. Hence, they could potentially be a useful taxon for examining broader scale patterns in contaminant levels. It would be useful if published studies on contaminants in skates also included depth information, clarification of whether the 'size' reported is disc width or total length, and covered as broad a size (age) range as possible.

Acknowledgements

We thank the scientists and crew of RV *Thalassa* for collecting samples of the two species, and David James and Lee Warford for assisting with laboratory analyses of metals. The collection of biological material was supported by Defra contracts MB5201 and MF6001, and the analysis of trace elements by SLA22. We also thank Robin Law for his editorial advice given during the compilation of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.marpolbul.2017.08.054.

References

Asante, K.A., Agusa, T., Mochizuki, H., Ramu, K., Inoue, S., Kubodera, T., Takahashi, S., Subramanian, A., Tanabe, S., 2008. Trace elements and stable isotopes ($\delta^{13}C$ and $\delta^{15}N$) in shallow and deep-water organisms from the East China Sea. Environ. Pollut. 156, 862–873.

Baeyens, W., Leermakers, M., Papina, T., Saprykin, A., Brion, N., Noyen, J., De Gieter, M., Elskens, M., Goeyens, L., 2003. Bioconcentration and biomagnification of mercury and methylmercury in North Sea and Scheldt Estuary fish. Arch. Environ. Contam. Toxicol. 45, 498–508.

Braune, B.M., 1987. Mercury accumulation in relation to size and age of Atlantic herring (*Clupea harengus harengus*) from the southwestern Bay of Fundy, Canada. Arch. Environ. Contam. Toxicol. 16, 311–320.

CEC, 2006. Commission regulation (EC) no 1881/2006 of 19 December 2006. Setting maximum levels for certain contaminants in foodstuffs. Off. J. Eur. Union L 364, 5–24.

Chouvelon, T., Spitz, J., Caurant, F., Mendez-Fernandez, P., Autier, J., Lassus-Débat, A., Chappuis, A., Bustamante, P., 2012. Enhanced bioaccumulation of mercury in deep-sea fauna from the Bay of Biscay (north-east Atlantic) in relation to trophic positions identified by analysis of carbon and nitrogen stable isotopes. Deep Sea Res. I: Ocean. Res. Pap. 65, 113–124.

Choy, C.A., Popp, B.N., Kaneko, J.J., Drazen, J.C., 2009. The influence of depth on mercury levels in pelagic fishes and their prey. Proc. Natl. Acad. Sci. 106, 13865–13869.

Collings, S.E., Johnson, M.S., Leah, R.T., 1996. Metal contamination of angler-caught fish from the Mersey Estuary. Mar. Environ. Res. 41, 281–297.

Consalvo, I., Psomadakis, P.N., Bottaro, M., Vacchi, M., 2009. First documented record of *Leucoraja circularis* (Rajidae) in the central Tyrrhenian Sea. Mar. Biodiv. Rec. 2, e24.

Davies, I.M., 1981. Survey of Trace Elements in Fish and Shellfish Landed at Scottish Ports, 1975–76. Department of Agriculture and Fisheries for Scotland (28 pp).

De Gieter, M., Leermakers, M., Van Ryssen, R., Noyen, J., Goeyens, L., Baeyens, W., 2002. Total and toxic arsenic levels in North Sea fish. Arch. Environ. Contam. Toxicol. 43, 406–417.

Dixon, R., Jones, B., 1994. Mercury concentrations in stomach contents and muscle of five fish species from the north east coast of England. Mar. Pollut. Bull. 28, 741–745.

Du Buit, M.J., 1972. The role of geographical and seasonal factors in the diet of *R. naevus* and *R. fullonica*. Trav. Lab. Biol. Halieut. Rennes 6, 35–50.

Ebert, D.A., Bizzarro, J.J., 2007. Standardized diet compositions and trophic levels of skates (Chondrichthyes: Rajiformes: Rajoidei). Environ. Biol. Fish 80, 221–237.

Ebert, D.A., Stehmann, M.F.W., 2013. Sharks, batoids, and chimaeras of the North Atlantic. In: FAO Species Catalogue for Fishery Purposes. No. 7. FAO, Rome (523 pp).

EERC, 2011. Selenium and Mercury: Fishing for Answers. (https://www.icmj.com/userfiles/files/Selenium-Mercury.pdf last visited on 01/06/2017).

Farrugia, T.J., Oliveira, A.C., Knue, J.F., Seitz, A.C., 2015. Nutritional content, mercury, and trace element analyses of two skate (Rajidae) species in the Gulf of Alaska. J. Food Compos. Anal. 42, 152–163.

Franklin, A., 1987. The concentration of metals, organochlorine pesticides and PCB residues in marine fish and shellfish: Results from MAFF fish and shellfish monitoring programmes 1977–1984. In: Aquatic Environment Monitoring Report. Vol. 16 MAFF Directorate of Fisheries Research, Lowestoft (38 pp).

Franklin, A., Jones, J., 1995. Monitoring and surveillance of non-radioactive contaminants in the aquatic environment and activities regulating the disposal of wastes at sea, 1993. In: Aquatic Environment Monitoring Report. Vol. 44 MAFF Directorate of Fisheries Research, Lowestoft (68 pp).

Freeman, H.C., Horne, D.A., McTague, B., McMenemy, M., 1974. Mercury in some Canadian Atlantic coast fish and shellfish. J. Fish. Res. Bd. Can. 31, 369–372.

Gerhart, E.H., 1977. Concentrations of total mercury in several fishes from Delaware Bay, 1975. Pest. Monit. J. 11, 132–133.

Greig, R.A., Wenzloff, D., Shelpuk, C., 1975. Mercury concentrations in fish, North Atlantic offshore waters-1971. Pestic. Monit. J. 9, 15–20.

ICES, 2010. Manual for the International Bottom Trawl Surveys in the Western and Southern Areas Revision III. (64 pp. Available at: http://www.ices.dk/marine-data/Documents/DATRAS%20Manuals/Addendum_2_Manual_IBTS_Western_and_Southern_Areas_Revision_III.pdf).

Jones, B.R., Laslett, R.E., 1994. Methods for trace metals in marine and other samples. In: Aquatic Environment Protection: Analytical Methods. MAFF Directorate Fisheries Research, Lowestoft, pp. 11 (29 pp).

Last, P.R., White, W.T., de Carvalho, M.R., Séret, B., Stehmann, M.F.W., Naylor, G.J.P., 2016. Rays of the World. CSIRO Publishing/Comstock Publishing Associates (832 pp).

Lopez, S.A., Abarca, N., Concha, F., Meléndez, R., 2014. Heavy metal concentrations in two important fishes caught in artisanal fisheries of southeastern pacific waters. International J. Agric. Pol. Res. 2, 414–420.

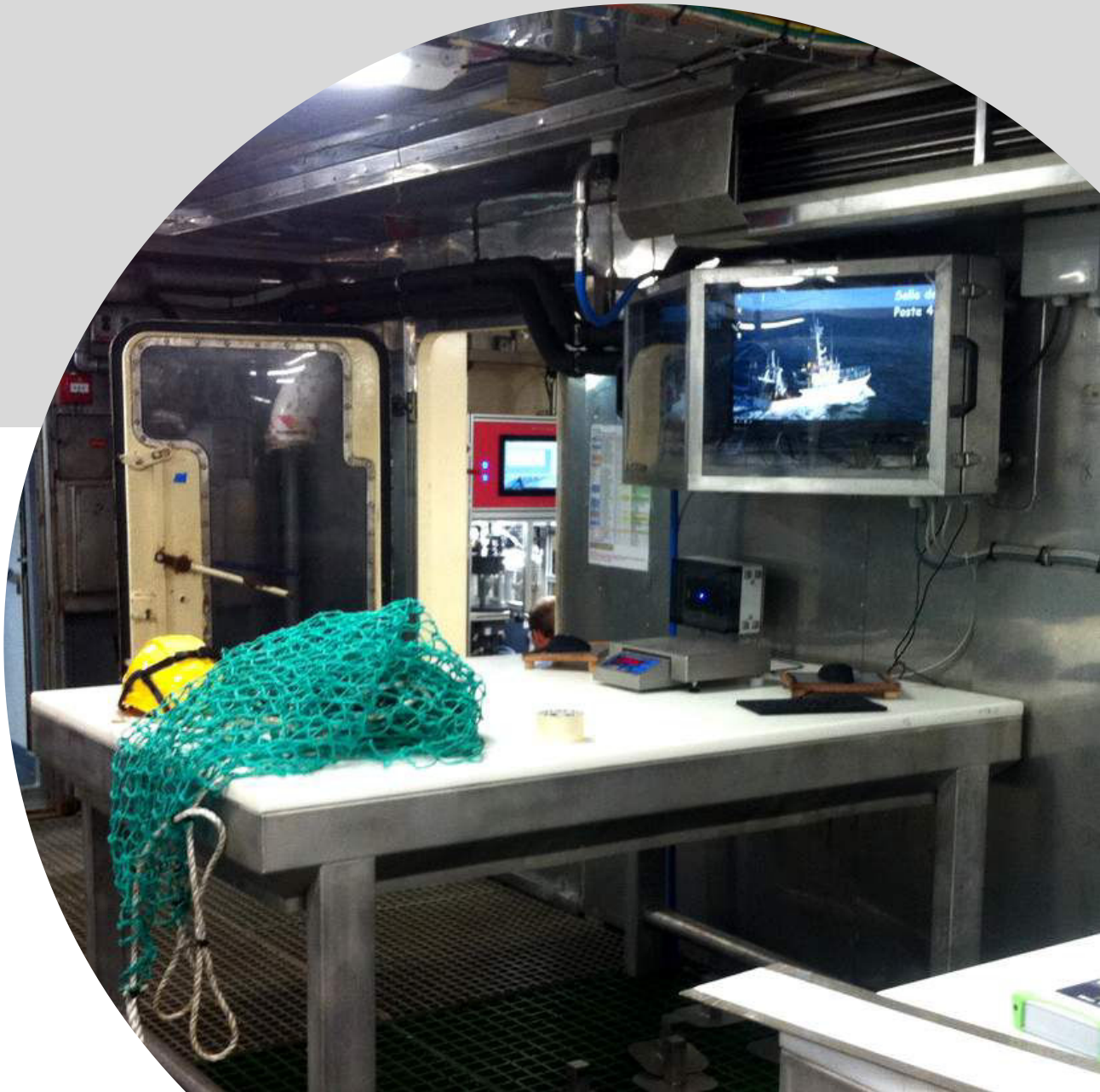
Lyons, K., Adams, D.H., 2017. First evidence of persistent organic contaminants as potential anthropogenic stressors in the barndoor skate *Dipturus laevis*. Mar. Pollut. Bull. 116, 534–537.

Mahé, J.C., Poulard, J.C., 2005. Spatial distribution and abundance trends of main elasmobranch species in the Bay of Biscay and Celtic Sea from bottom trawl surveys. In: ICES CM 2005/N:04, (27 pp).

Marcovecchio, J.E., Moreno, V.J., Perez, A., 1988. Determination of heavy metal concentrations in biota of Bahía Blanca, Argentina. Sci. Total Environ. 75, 181–190.
Mead, R., Curnow, R., 1984. Statistical Methods in Agriculture and Experimental Biology. Chapman and Hall (488 pp).
Mnsari, N., Boumaïza, M., Capapé, C., 2009. Morphological data, biological observations and occurrence of a rare skate, *Leucoraja circularis* (Chondrichthyes: Rajidae), off the northern coast of Tunisia (central Mediterranean). Pan-Am. J. Aquat. Sci. 4, 70–78.
Mormede, S., Davies, I.M., 2001. Heavy metal concentrations in commercial deep-sea fish from the Rockall Trough. Cont. Shelf Res. 21, 899–916.
Murray, A.J., 1979. Metals, organochlorine pesticides and PCB residue levels in fish and shellfish landed in England and Wales during 1974. In: Aquatic Environment Monitoring Report. Vol. 2 MAFF Directorate Fisheries Research, Lowestoft (11 pp + tables).
Murray, A.J., 1981. Metals, organochlorine pesticides and PCB residue levels in fish and shellfish landed in England and Wales during 1975. In: Aquatic Environment Monitoring Report. Vol. 5 MAFF Directorate Fisheries Research, Lowestoft (40 pp).
Murray, A.J., Portmann, J.E., 1984. Metals and organochlorine pesticide and PCB residues in fish and shellfish in England and Wales in 1976 and trends since 1970. In: Aquatic Environment Monitoring Report. Vol. 10 MAFF Directorate of Fisheries Research, Lowestoft (79 pp).
Nicolaus, E.E.M., Law, R.J., Wright, S.R., Lyons, B.P., 2015. Spatial and temporal analysis of the risks posed by polycyclic aromatic hydrocarbon, polychlorinated biphenyl and metal contaminants in sediments in UK estuaries and coastal waters. Mar. Pollut. Bull. 95, 469–479.
Nicolaus, E.E.M., Bendall, V.A., Bolam, T.P.C., Maes, T., Ellis, J.R., 2016a. Concentrations of mercury and other trace elements in porbeagle shark. Mar. Pollut. Bull. 112, 406–410.
Nicolaus, E.E.M., Wright, S.R., Bolam, T.P.C., Barber, J.L., Bignell, J.P., Lyons, B.P., 2016b. Spatial and temporal analysis of the risks posed by polychlorinated biphenyl and metal contaminants in dab (*Limanda limanda*) collected from waters around England and Wales. Mar. Pollut. Bull. 112, 399–405.
Nicolaus, E.E.M., Wright, S.R., Barry, J., Bolam, T.P.C., Ghareeb, K., Ghaloom, M., Al-Kanderi, N., Harley, B.F.M., Le Quesne, W.J.F., Devlin, M.J., Lyons, B.P., 2017. Spatial and temporal analysis of the risks posed by total petroleum hydrocarbon and trace element contaminants in coastal waters of Kuwait. Mar. Pollut. Bull. http://dx.doi.org/10.1016/j.marpolbul.2017.04.031.
OSPAR, 2009. Background Document on CEMP Assessment Criteria for QSR 2010. (Publication Number: 461/2009. 23 pp).
Pentreath, R.J., 1973. The accumulation from sea water of 65-Zn, 54-Mn, 58-Co and 59-Fe by the thornback ray, *Raja clavata* L. J. Exp. Mar. Biol. Ecol. 12, 327–334.
Pentreath, R.J., 1976. The accumulation of mercury by the thornback ray (*Raja clavata*). J. Exp. Mar. Biol. Ecol. 25, 131–140.
Pentreath, R.J., 1977a. The accumulation of 110Ag by the plaice, *Pleuronectes platessa* L. and the thornback ray, *Raja clavata* L. J. Exp. Mar. Biol. Ecol. 29, 315–325.
Pentreath, R.J., 1977b. The accumulation of cadmium by the plaice, *Pleuronectes platessa* L. and the thornback ray, *Raja clavata* L. J. Exp. Mar. Biol. Ecol. 30, 223–232.
Ruelas-Inzunza, J., Escobar-Sánchez, O., Patrón-Gómez, J., Moreno-Sanchez, X.G., Murillo-Olmeda, A., Spanopoulos-Hernández, M., Corro-Espinosa, D., 2013. Mercury in muscle and liver of ten ray species from Northwest Mexico. Mar. Pollut. Bull. 77, 434–436.
Sandoval-Herrera, N.I., Vargas-Soto, J.S., Espinoza, M., Clarke, T.M., Fisk, A.T., Wehrtmann, I.S., 2016. Mercury levels in muscle tissue of four common elasmobranch species from the Pacific coast of Costa Rica, Central America. Reg. Stud. Mar. Sci. 3, 254–261.
Stenner, R.D., Nickless, G., 1975. Heavy metals in organisms of the Atlantic coast of SW Spain and Portugal. Mar. Pollut. Bull. 6, 89–92.
Storelli, M.M., Barone, G., 2013. Toxic metals (Hg, Pb, and Cd) in commercially important demersal fish from Mediterranean Sea: contamination levels and dietary exposure assessment. J. Food Sci. 78, T362–T366.
Storelli, M.M., Stuffler, R.G., Marcotrigiano, G.O., 1998. Total mercury in muscle of benthic and pelagic fish from the South Adriatic Sea (Italy). Food Addit. Contam. 15, 876–883.
Storelli, M.M., Barone, G., Perrone, V.G., Storelli, A., 2013. Risk characterization for polycyclic aromatic hydrocarbons and toxic metals associated with fish consumption. J. Food Compos. Anal. 31, 115–119.
Taylor, D.L., Kutil, N.J., Malek, A.J., Collie, J.S., 2014. Mercury bioaccumulation in cartilaginous fishes from Southern New England coastal waters: contamination from a trophic ecology and human health perspective. Mar. Environ. Res. 99, 20–33.
Thomson, J.D., 1985. Mercury concentrations of the axial muscle tissues of some marine fishes of the continental shelf adjacent to Tasmania. Aust. J. Mar. Freshw. Res. 36, 509–517.
Torres, P., da Cunha, R.T., Micaelo, C., dos Santos Rodrigues, A., 2016. Bioaccumulation of metals and PCBs in *Raja clavata*. Sci. Total Environ. 573, 1021–1030.
Türkmen, M., Tepe, Y., Türkmen, A., Sangün, M.K., Ateş, A., Genç, E., 2013. Assessment of heavy metal contamination in various tissues of six ray species from İskenderun Bay, Northeastern Mediterranean Sea. Bull. Environ. Contam. Toxicol. 90, 702–707.
Windom, H., Stickney, R., Smith, R., White, D., Taylor, F., 1973. Arsenic, cadmium, copper, mercury, and zinc in some species of North Atlantic finfish. J. Fish. Res. Bd. Can. 30, 75–279.
Zauke, G.P., Savinov, V.M., Ritterhoff, J., Savinova, T., 1999. Heavy metals in fish from the Barents Sea (summer 1994). Sci. Total Environ. 227, 161–173.
Zhang, L., Ye, X., Feng, H., Jing, Y., Ouyang, T., Yu, X., Liang, R., Gao, C., Chen, W., 2007. Heavy metal contamination in western Xiamen Bay sediments and its vicinity, China. Mar. Pollut. Bull. 54, 974–982.
Zitko, V., Finlayson, B.J., Wildish, D.J., Anderson, J.M., Kohler, A.C., 1971. Methylmercury in freshwater and marine fishes in New Brunswick, in the Bay of Fundy and on the Nova Scotia Banks. J. Fish. Res. Bd. Can. 28, 1285–1291.

ANNEXE D

De la difficulté à obtenir des paramètres biologiques : étude de deux Rajidae méconnues



Working Document to the ICES Working Group on Elasmobranch Fishes, Lisbon, 15–24 June 2016

Biological studies of *Leucoraja fullonica* and *Leucoraja circularis* in the Northeast Atlantic

S. R. McCully Phillips[1], P. Lorance[2], F. Marandel[2] and J. R. Ellis[1]

[1]Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk, NR33 0HT, UK

[2]IFREMER, Unité Ecologie et Modèles pour l'Halieutique B.P. 21105, 44311 Nantes Cedex 03, France

Summary

Biological data were collected from 25 specimens of *Leucoraja fullonica* and 14 specimens of *Leucoraja circularis* collected from the Northeast Atlantic. Conversion factors are presented along with data on hepato- and gonado-somatic indices, maturity information, nidamental gland width and clasper length data. Preliminary information on diet composition is also given.

Introduction

To date, there are very limited published investigations on the life history of either of these relatively large-bodied skates. Shagreen ray *Leucoraja fullonica* reaches a maximum size of between 100–120 cm total length (Bauchot, 1987; Muus and Nielsen, 1999). Its reputation in the literature has been largely restricted to date, on presence in trawl surveys, and distributional range (Ellis *et al*., 2015). Very little is known regarding its biology and reproductive cycle, other than that it is oviparous, and produces egg cases that measure about 80 mm by 50 mm (Stehmann and Bürkel 1984). McCully *et al*. (2012) reported on a limited number of specimens from trawl surveys of the Celtic Sea (1992–2011), with total length (L_T) ranging from 21 to 96 cm and 24 to 70 cm in males and females, respectively. All female specimens were immature, while only two males at 75 and 96 cm L_T were mature; the largest immature male caught was 82 cm L_T. It is an offshore species, usually occurring on the outer parts of the continental shelf. In the Northeast Atlantic, it is suspected that there have been continuing population declines of 30–50% over three generations, and is classified as '*Vulnerable*' by the IUCN (McCully and Walls, 2015).

Sandy ray *Leucoraja circularis* is even more data-limited than its congener. The maximum recorded size is 120 cm L_T (Stehmann, 1990), but most individuals caught are between 70 and 80 cm L_T (Serena, 2005, Ebert and Stehmann, 2013). Very little is known regarding its biology and reproductive cycle, other than that it is oviparous, and produces egg cases that measure 88-90 by 50-60 mm (Stehmann and Bürkel 1984; Mnsari *et al*. 2009). Age at maturity, longevity, size at birth, reproductive age, gestation time, reproductive periodicity, fecundity, rate of population increase and natural mortality are all unknown (McCully *et al*., 2015). It is also an offshore species, occurring in deeper shelf and slope waters, down to depths of up to 800 m. Given that this species was often reported in historical accounts as occurring in shelf seas, but is now restricted to deeper waters, it is

suspected to have declined in the Northeast Atlantic and Mediterranean Sea by more than 50% in the last three generations, and is classified as '*Endangered*' by the IUCN (McCully *et al*., 2015).

Materials and Methods

Specimens of both species were caught during the Ifremer EVHOE surveys of the Celtic Sea in 2014 and 2015. Some additional specimens were included from Cefas' observer programme and fisheries-independent surveys. Specimens were initially frozen prior to detailed examination in the laboratory (see Table 1 for measurements collected). Some specimens that were subjected to prolonged freezing were dehydrated and therefore excluded from length-weight analyses. Maturity for males was assigned based on gross external examination of the claspers and internal inspection of the testes, while for females, it was assigned following internal examination of the ovaries, oocytes and development of nidamental gland. Specimens were classified as immature (A), maturing (B), mature (C), or active (D), according to the maturity key given in Appendix I. Only fish at Stages C and D are considered to be mature (i.e. capable of reproducing).

Table 1: Parameters collected from the *Leucoraja* cadavers.

All specimens (where possible)*:	Males only:
• Sex • Total length (cm) • Disc width (mm) • Total weight (g) • Liver weight (0.1 g) • Gonad weight (0.1 g, including epigonal organ) • Weight of stomach contents (0.1 g) • Gutted weight (g) • Maturity stage • Stomach 'fullness' score • Identification of stomach contents	• Outer clasper length (mm) • Inner clasper length (mm) Females only: • Nidamental gland width (mm) • Number of mature follicles • Maximum follicle diameter (mm)

*not all parameters (e.g. stomach contents) were possible for all specimens, due to freezer damage

Results

To date, 14 *Leucoraja circularis* (7 female, 7 male) and 25 *Leucoraja fullonica* (12 females, 13 male) samples (Figure 1) have been obtained and fully dissected for scientific study.

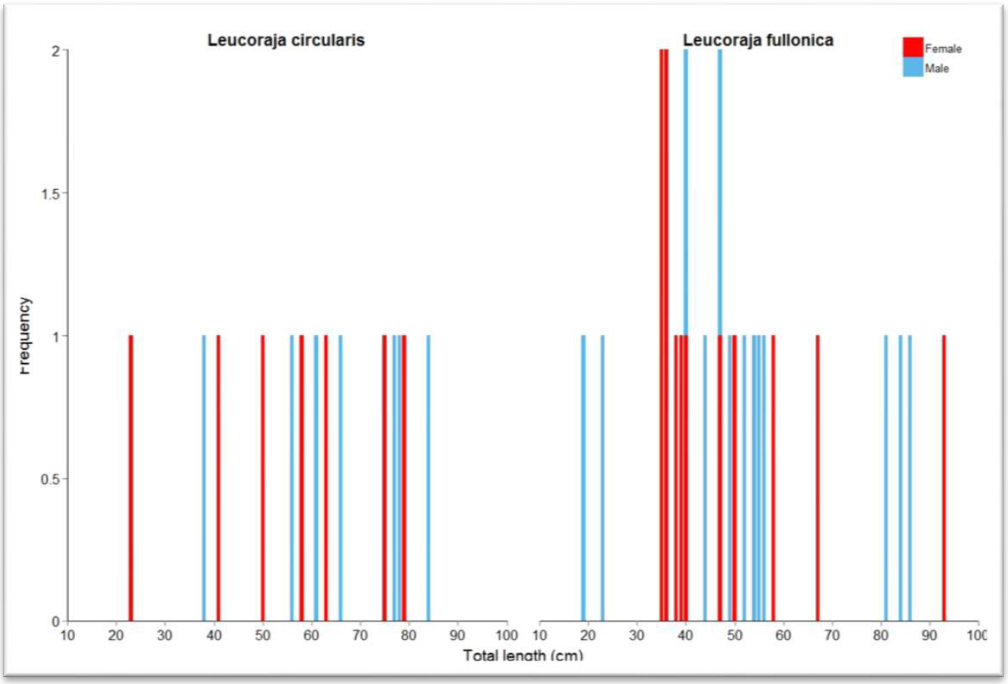


Figure 1: Number of *Leucoraja* biologically sampled by length and sex

Length-weight conversion factors

In fisheries science, conversion factors between total length and weight are important parameters. Assessment of species status is either length- or age-based, and given that elasmobranchs are not routinely aged, length information is critical to studies of population dynamics. Figure 2 shows the relationship between total length and weight in the specimens sampled by sex. The trend is very similar between the species, especially in the smaller specimens. In the larger skate, *L. circularis* were marginally heavier for a given L_T than *L. fullonica*, although data for larger *L. circularis* were lacking, and the low sample size of larger specimens may skew the estimated length-weight relationship. Skates are traditionally landed for the market gutted. Therefore, a conversion factor of eviscerated weight to length is also a useful parameter to determine (Figure 3) to augment data collected during market sampling programmes. Again, the relationship was very similar between the species, with large overlap within the 95% confidence limits throughout most of the size range. These relationships were not examined by sex, given the small sample size.

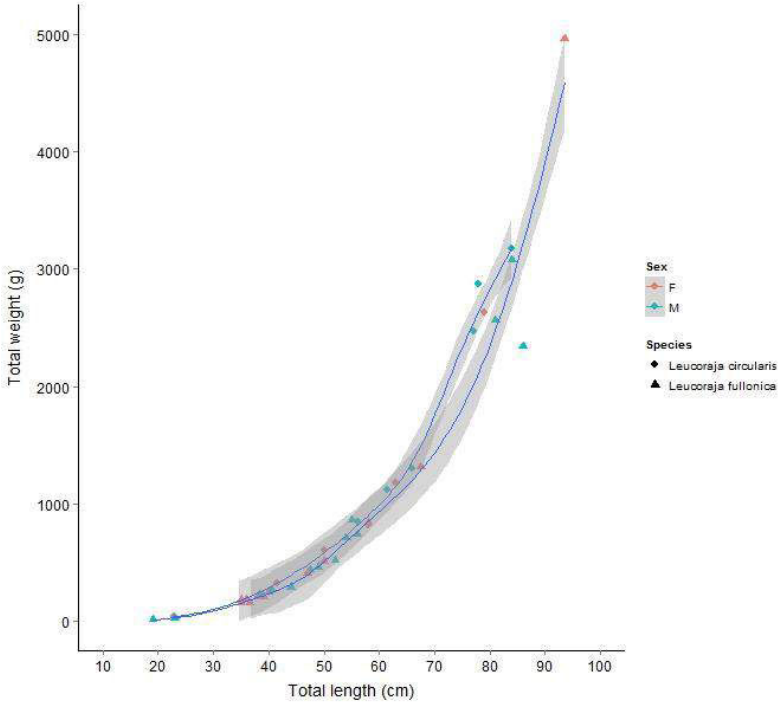


Figure 2: Relationship between total weight and total length by sex, represented by the equation: $y=7.3x^{0.3}$ (*L. circularis* $r^2 = 0.996$) and $y=8.3x^{0.29}$ (*L. fullonica* $r^2 = 0.990$)

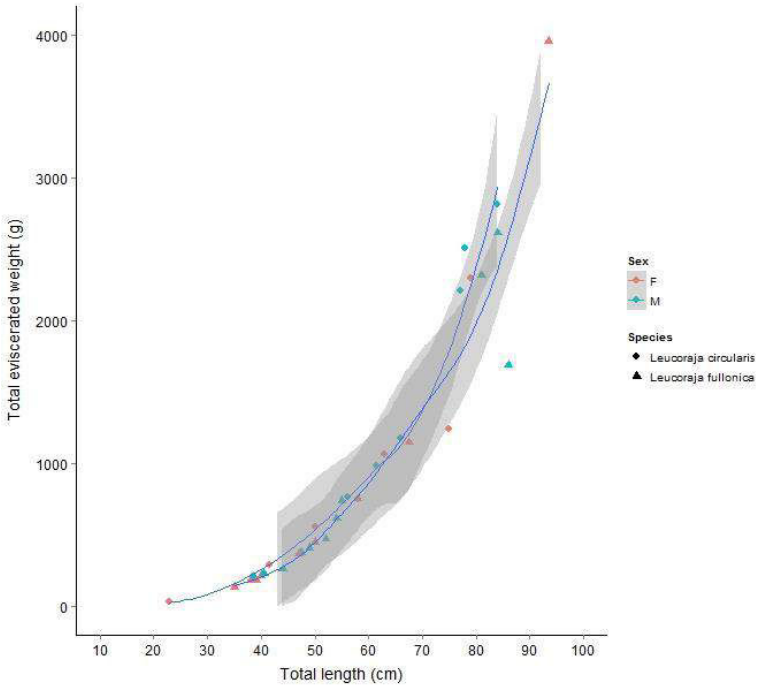


Figure 3: Total length to gutted weight relationship, represented by the equation: $y=8.2x^{0.29}$ (*L. circularis* $r^2 = 0.985$) and $y=7.8x^{0.3}$ (*L. fullonica* $r^2 = 0.983$)

Hepatosomatic index

Livers were removed and weighed for each specimen, as relating the relative weight of this organ to fish condition and maturity stage is critical to understanding the reproductive cycle of elasmobranch fish. The relationship between liver weight and total length was examined (Figure 4) and, although positively correlated, it is likely to be dependent upon several factors, including sex, maturity stage and season. The liver weight can be expressed as a proportion of and body weight (the hepatosomatic index, HSI; Table 2). It is a good indicator of the energy reserve in an animal, thus the lowest values are usually seen in females nearing the end of the reproductive cycle (McCully Phillips & Ellis, 2015), however, this is not evident here due to the lack of mature and active females. Changes in this index can indicate spawning seasons and environmental quality. The average HSI across all samples was 4.94, with the smallest (2.78) exhibited by the smallest sandy ray in the samples (L_T = 23cm), with two smaller shagreen ray specimens having larger HSI, indicating that the smaller bodied species (sandy ray) is likely to have a reduced HSI to the larger bodied species (shagreen ray). The largest index (10.28) from a mature shagreen male (L_T = 86cm), with the largest shagreen female (L_T = 93.5) having a lower HSI (8.96) due to the presence of large mature follicles, which would have reduced the available energy reserve.

Table 2: Hepatosomatic index (HSI) of *Leucoraja* sampled by sex and maturity stage

Maturity Stage	Mean HSI (females)		Mean HSI (males)	
	L. circularis (n)	*L. fullonica* (n)	*L. circularis* (n)	*L. fullonica* (n)
A	4.02 (7)	4.77 (11)	3.91 (3)	4.67 (9)
B	-	-	4.76 (2)	-
C	-	8.96 (1)	5.02 (2)	7.64 (3)

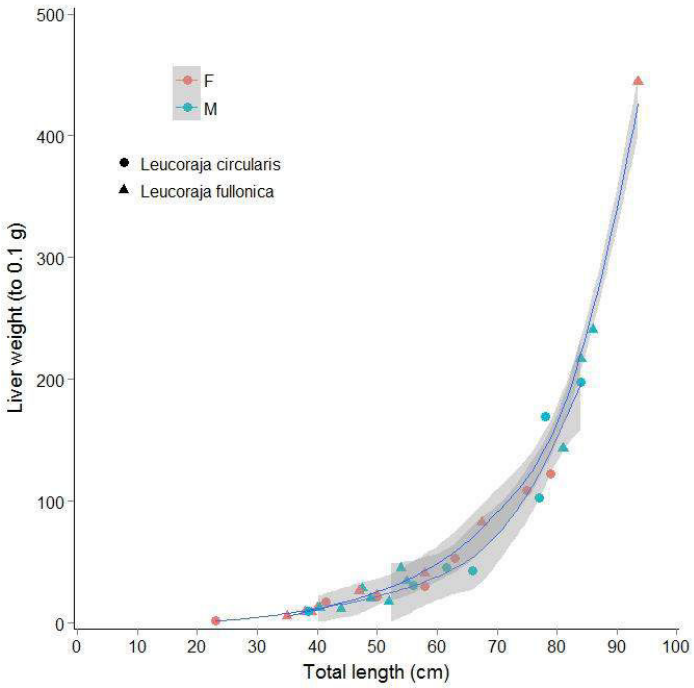


Figure 4: Relationship between total length and liver weight represented by the equation: $y=22x^{0.26}$ (*L. circularis* r^2 = 0.973) and $y=23x^{0.24}$ (*L. fullonica* r^2 = 0.970)

Gonadosomatic index

The basic relationship between total length and gonad weight was explored (Figure 5). The pattern seen across this length range showed a clear increase with size. The association between gonad weight and total body length was expressed as the gonadosomatic index (GSI), and the average GSI by sex and maturity stage is given in Table 3. As expected, this increases throughout life to the 'mature' stage.

Table 3: Mean gonad weight and gonadosomatic index (GSI) by sex and maturity

Sex	**Female**				**Male**			
Maturity Stage	**Mean gonad weight (g)**		**Mean GSI (%)**		**Mean gonad weight (g)**		**Mean GSI (%)**	
	L. circularis (n)	*L. fullonica (n)*	*L. circularis*	*L. fullonica*	*L. circularis (n)*	*L. fullonica (n)*	*L. circularis*	*L. fullonica*
A	3.36 (7)	1.6 (11)	0.32	0.31	2.37 (3)	1.46 (9)	0.31	0.25
B	-		-	-	10.9 (2)	-	0.48	-
C	-	36.3 (1)	-	0.73	18.25 (2)	21.5 (3)	0.69	0.81

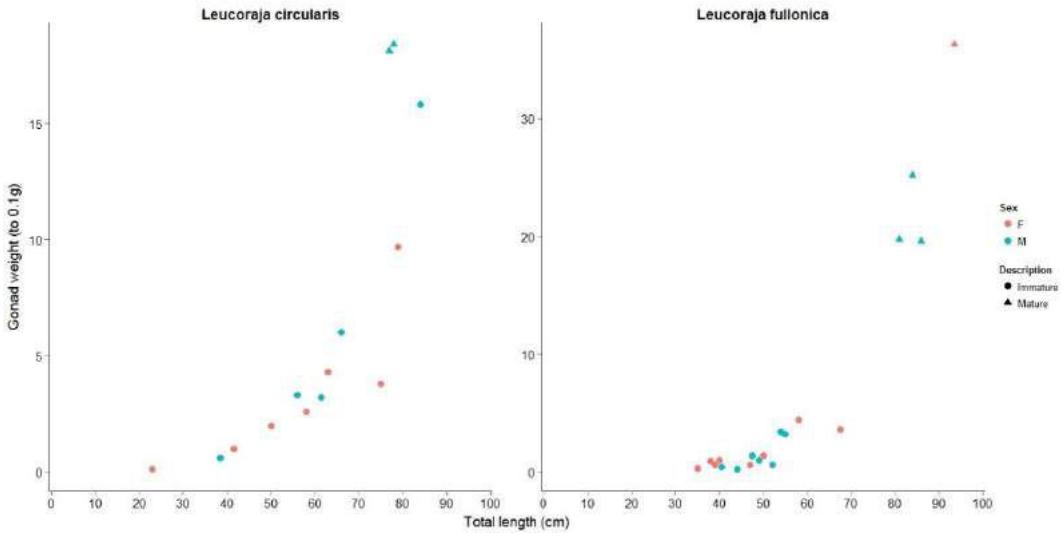


Figure 5: Relationship between total length and gonad weight

Maturity

The size at which fish mature is an important factor to determine in fisheries management. Often in elasmobranchs, it may be better to 'protect' the larger female proportion of the population, as they are the most fecund and may contribute the most to recruitment. Indeed, maximum landing lengths have been used as measures to protect the largest individuals of both spurdog and porbeagle in EU waters. For other populations, skate in particular, minimum landings sizes are introduced to protect the juveniles until adulthood. Length at 50% maturity ogives and estimates cannot be given due to

the lack of samples, especially of mature females. What these data do show is that maturity is at a large size in comparison to other rajidae, and is conceivably similar to that of *Raja brachyura* (L_{50} of 78.0 and 83.4 for males and females respectively; McCully *et al*., 2012).

Table 4: Maturity estimates for the samples to date (number of samples available in brackets)

	Female		**Male**	
	L. circularis	***L. fullonica***	***L. circularis***	***L. fullonica***
Smallest mature	NA	93.5 (1)	77 (2)	81 (3)
Largest immature	79 (7)	67.5 (11)	84 (5)	56 (10)

The nidamental gland is where the oocytes become fertilised before passing down into the uterus where they form the egg-cases. The width of this gland is closely associated with the onset of maturity. This can be an important parameter to measure – especially where different biological studies employ different maturity scales. Figure 6 shows this relationship for both immature and mature females, which separate between approximately 10–35 mm width mark, although unfortunately a more accurate estimate is hampered by the lack of mature females.

Similar to the nidamental gland width in females, clasper length of males can also provide a quantitative measure of maturity to augment the qualitative assignment of maturity scales. The outer clasper length to total length relationship for the males sampled to date, by maturity stage is shown in Figure 7, with the maturation of males occurring approximately between 90–100 mm outer clasper length.

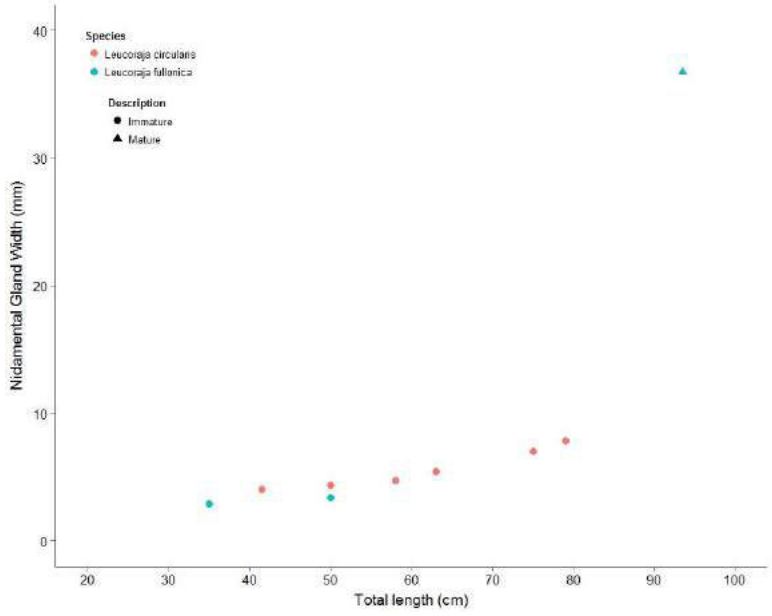


Figure 6: Relationship between nidamental gland width and total length in females.

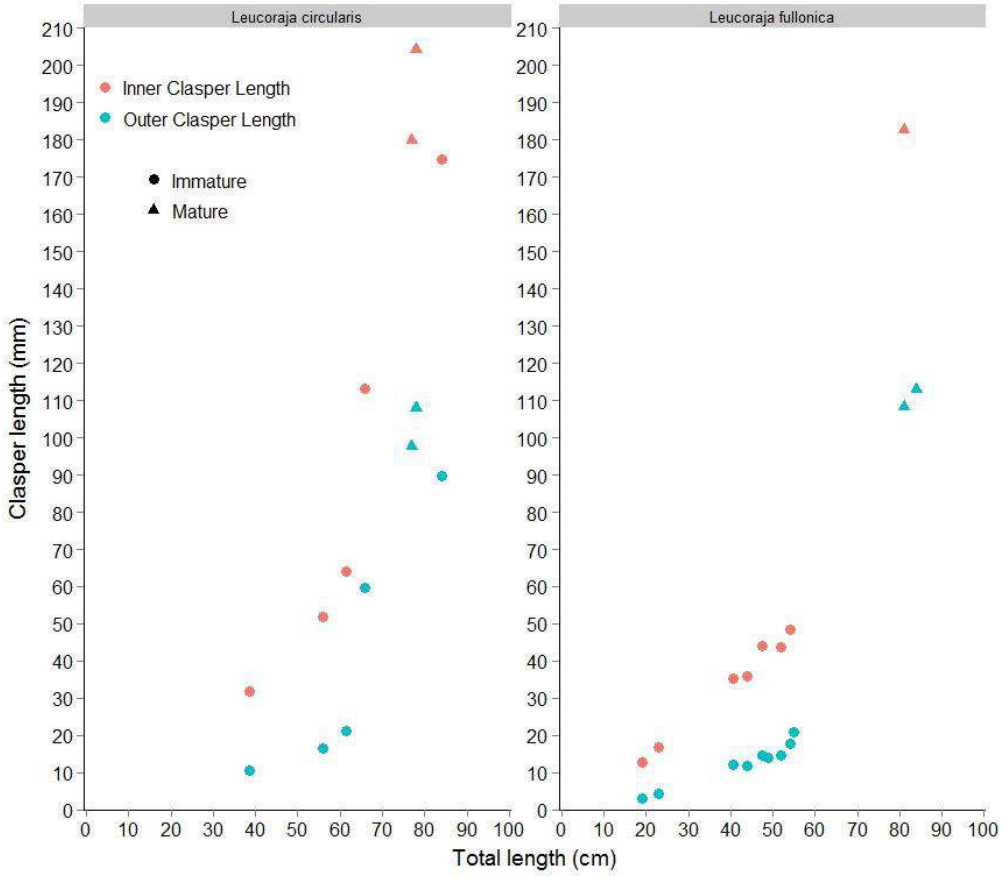


Figure 7: Relationship between outer clasper length and total length in males.

Diet composition and stomach fullness

The stomachs of each specimen were extracted, qualitatively given a 'fullness' score, before being emptied and the contents weighed, then analysed for prey species identification. The weights on the stomach contents (to nearest 0.1 g) ranged from 0–261.5g, with an average of 19.4g (n = 38). The species identified are given in Table 5. Primarily, the most common contents were digested, crustacean and fish remains, but of those remains which could be identified, shrimps and *Capros aper* (Figure 8) dominated the diet of *L. circularis*. The diet of *L. fullonica*, was generally more piscivorous, with *Scyliorhinus canicula*, euphausiids, *Leucoraja naevus* and *Processa* spp. dominating. One 86 cm L_T specimen had a 37cm L_T *L. naevus* in its stomach (possibly consumed in the net, Figure 9), while two specimens contained *S. canicula*, one of which had consumed 13 individuals up to 22 cm L_T.

Figure 8: Stomach contents of *L. circularis* showing *Capros aper* (left) and an unidentified Polychelidae (right)

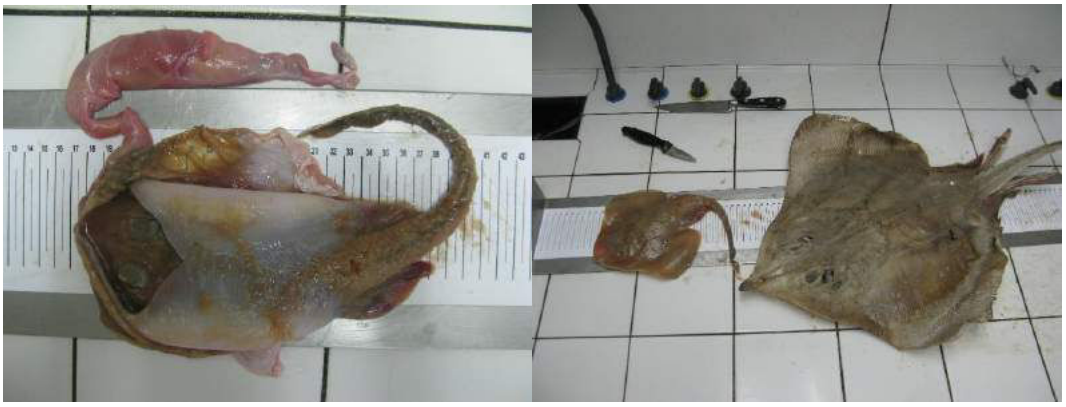


Figure 9: Stomach contents of *L. fullonica* showing a freshly consumed 37cm L_T *L. naevus*

Table 5: Species present in the stomach contents of *Leucoraja*.

Amphipoda
Capros aper
Crab (brachyura indet.)
Crustacean remains
Digested remains
Euphausiids
Fish remains
Polychelidae
Leucoraja naevus
Polychaete
Processa spp.
Scyliorhinus canicula
Shrimps indet.
Solenocera membranacea

References:

Bauchot, M.L. 1987. Raies at autres batoidés. In: M. Fisher, M. Schneider and M.-L. Bauchot (eds), *Fiches FAO d'Identification des Espècs pour les Besoins de la Peche. Méditerranée et Mer Noire. Zone de Peche 37. Revision 1. II*, pp. 847-885. FAO, Rome.

Ebert, D.A. and Stehmann, M.F.W. 2013. *Sharks, batoids, and chimaeras of the North Atlantic. FAO Species Catalogue for Fishery Purposes No. 7. Food and Agricultural Organization of the United Nations (FAO)*. FAO, Rome.

Ellis, J. R., Heessen, H. J. L. and McCully Phillips, S. R. 2015. Skates (Rajidae). In: *Fish atlas of the Celtic Sea, North Sea, and Baltic Sea* (Heessen, H. J. L., Daan, N. and Ellis, J. R., Eds.). Wageningen Academic Publishers / KNNV Publishing, 96–124.

McCully, S. and Walls, R. 2015. *Leucoraja fullonica*. The IUCN Red List of Threatened Species 2015: e.T161461A48938639. http://dx.doi.org/10.2305/IUCN.UK.2015-1.RLTS.T161461A48938639.en. Downloaded on **26 May 2016**.**16**.

McCully, S. R., Scott, F., and Ellis, J. R. 2012. Lengths at maturity and conversion factors for skates (Rajidae) around the British Isles, with an analysis of data in the literature. *ICES Journal of Marine Science: Journal du Conseil*, *69*(10), 1812-1822.

McCully, S., Ellis, J., Walls, R. and Fordham, S. 2015. *Leucoraja circularis*. The IUCN Red List of Threatened Species 2015: e.T161464A48938919. http://dx.doi.org/10.2305/IUCN.UK.2015-1.RLTS.T161464A48938919.en. Downloaded on **26 May 20**

Mnsari, N., Boumaïza, M. and Capapé, C. 2009. Morphological data, biological observations and occurrence of a rare skate, *Leucoraja circularis* (Chondrichthyes: Rajidae), off the northern coast of Tunisia (central Mediterranean). *Pan-American Journal of Aquatic Sciences* 4(70-78).

Muus, B.J. and J.G. Nielsen, 1999. Sea fish. Scandinavian Fishing Year Book, Hedehusene, Denmark. 340 p.

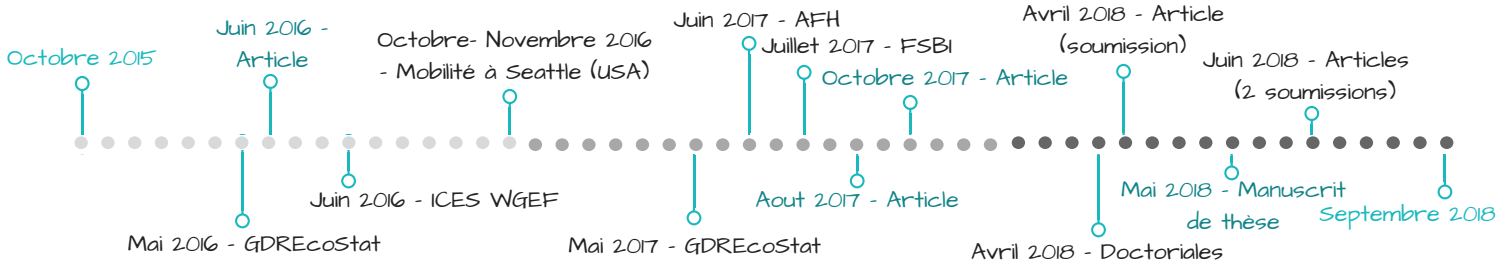
Serena, F. 2005. *Field identification guide to the sharks and rays of the Mediterranean and Black Sea*.

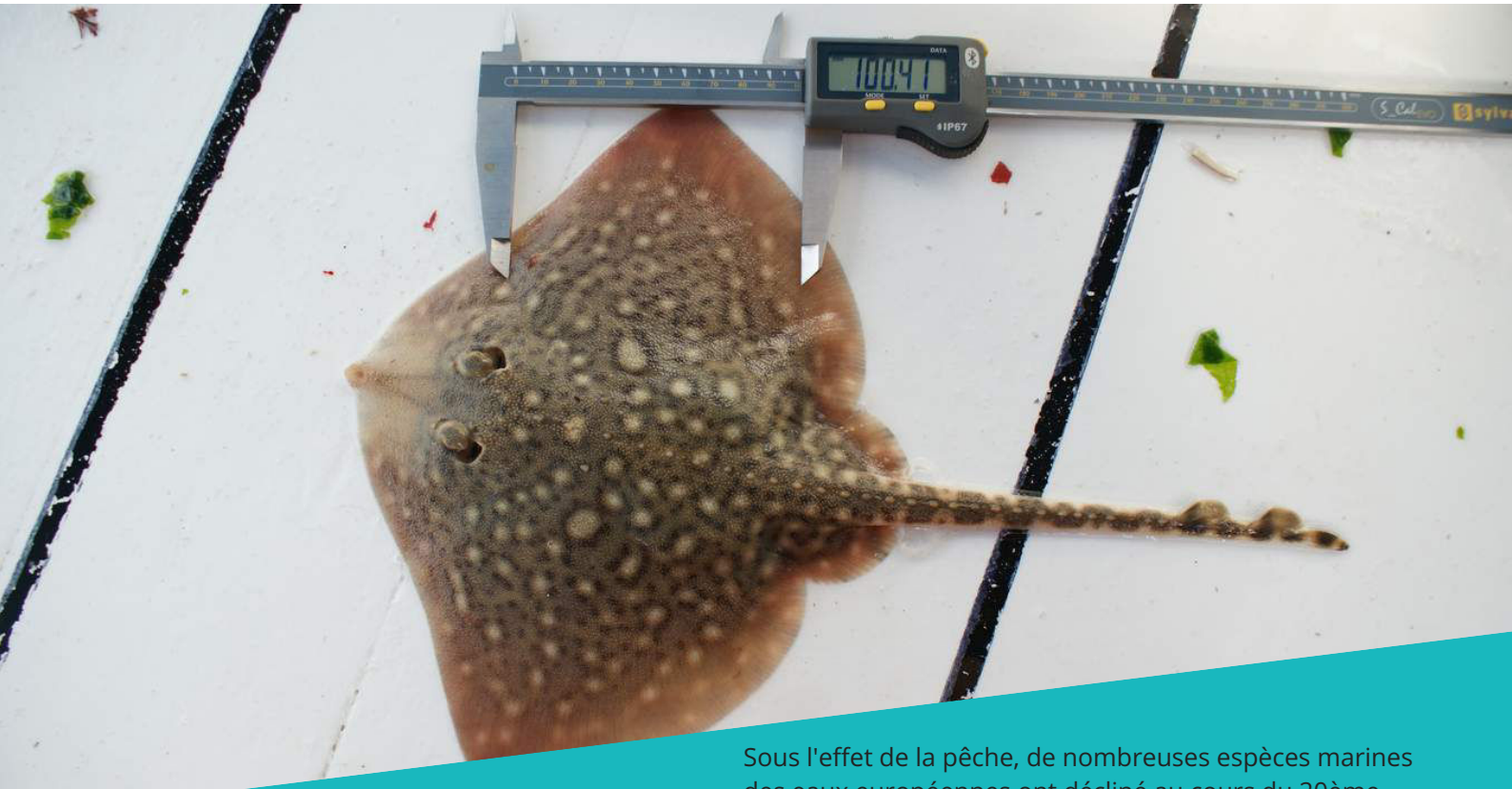
Stehmann, M. and Burkel, D.L. 1984. Rajidae. In: P.J.P. Whitehead, M.-L. Bauchot, J.-C. Hureau, J. Nielsen and E. Tortonese (eds), *Fishes of the North-eastern Atlantic and Mediterranean*, pp. 163–196. UNESCO, Paris.

Stehmann, M., 1990. Rajidae. p. 29-50. In J.C. Quero, J.C. Hureau, C. Karrer, A. Post and L. Saldanha (eds.) Check-list of the fishes of the eastern tropical Atlantic. Junta Nacional de Investigaçao Cientifica e Tecnológica, Lisbon, Portugal. Vol. 1.

Appendix I: Maturity scale used in the present study

Maturity stage	Males	Females
A (Immature)	Claspers undeveloped, shorter than extreme tips of posterior margin of pelvic fin Testes small and thread-shaped	Ovaries small, gelatinous, or granulated, but with no differentiated oocytes visible Oviducts small and thread-shaped, width of shell gland not much greater than the width of oviduct
B (Maturing)	Claspers longer than posterior margin of pelvic fin, their tips more structured, but claspers soft and flexible and cartilaginous elements not hardened Testes enlarged, sperm ducts beginning to meander	Ovaries enlarged and with more transparent walls. Oocytes differentiated in various small sizes (<5 mm). Oviducts small and thread-shaped, width of shell gland greater than width of the oviduct, but not hardened
C (Mature)	Claspers longer than posterior margin of pelvic fin, cartilaginous elements hardened, and claspers stiff Testes enlarged, sperm ducts meandering and tightly filled with sperm	Ovaries large with enlarged oocytes (>5 mm), with some very large, yolk-filled oocytes (ca. 10 mm) also present Uteri enlarged and wide, shell gland fully formed and hard
D (Active)	Clasper reddish and swollen, sperm present in clasper groove, or flowing if pressure exerted on cloaca	Egg capsules beginning to form in shell gland, partially visible in uteri, or egg capsules fully formed and hardened and in oviducts/uteri, or egg case being exuded from cloaca





ÉVALUATION DE L'ÉTAT DES POPULATIONS DE RAIE BOUCLÉE

Sous l'effet de la pêche, de nombreuses espèces marines des eaux européennes ont décliné au cours du 20ème siècle. Les raies et requins, plus sensibles à la pêche que les poissons osseux, ont été particulièrement touchés. La conservation de ces taxons constitue donc un objectif majeur de la gestion des ressources marines. Toutefois, l'estimation de l'état actuel des populations par les approches halieutiques usuelles est souvent impossible pour ces espèces peu abondantes, en raison notamment du nombre insuffisant d'observations. Il est donc primordial de disposer de techniques alternatives pour estimer l'état de ces populations et améliorer la connaissance de leur biologie en lien avec leur gestion et conservation.
Cette thèse présente une comparaison d'outils démographiques et génétiques afin d'évaluer l'état des populations d'élasmobranches en prenant la raie bouclée comme cas d'étude. La thèse est ainsi organisé en trois parties : délimitation des populations, évaluation génétique et démographique de leur état.

CONNECTIVITÉ

Étude de la définition d'unités de population appropriées pour la gestion de la raie bouclée par méthodes génétique et démographique.

MODÉLISATION

Étude de l'estimation de l'abondance de raie bouclée du Golfe de Gascogne à partir de données de pêches professionnelles et scientifiques.

POPULATION EFFICACE

Étude de l'estimation de la taille de population efficace de la raie bouclée à partir de données génétiques empiriques et simulées.

Titre : Evaluation de l'état des populations de raie bouclée

Mots clés : Modélisation, génétique, taille de population efficace, connectivité, conservation, élasmobranches

Résumé : Sous l'effet de la pêche, de nombreuses espèces de raies des eaux européennes ont décliné au cours du 20ème siècle. La conservation de ces espèces est un objectif majeur quant à la gestion des ressources marines. La raie bouclée (*Raja clavata*) est l'espèce de raie la plus répandue d'Atlantique Nord-Est. Sa gestion, basée sur un quota non spécifique, repose principalement sur les observations scientifiques et professionnelles et non sur des méthodes d'évaluations d'abondance. Les objectifs de cette thèse consistent ainsi à comparer les méthodes d'évaluations d'abondances disponibles pour cette espèce et à les appliquer aux données disponibles. Deux grands axes sont creusés : l'utilisation de méthodes basées sur la démographie de l'espèce et de méthodes basées sur la génétique. Cette thèse s'articule autour de 3 grandes parties : 1) délimitation des populations de raie bouclée pour leur gestion ; 2) évaluation de l'abondance des populations de raie bouclée ; 3) comparaison de l'état de la raie bouclée avec les autres espèces de raies de l'Atlantique Nord-Est.
Les résultats mettent en avant l'utilité des outils démographiques pour l'exploitation de l'espèce. En effet, les modèles halieutiques développés permettent d'estimer l'état de la biomasse de la raie bouclée grâce à des tendances. Le stock du golfe de Gascogne a notamment été estimé dans un bon état écologique. Les outils génétiques présentent une grande utilité dans la conservation des espèces marines car ils renseignent principalement sur la santé génétique au long terme.

Title : Assessment of the state of thornback ray populations

Keywords : Modelling, genetics, effective population size, connectivity, conservation, elasmobranchs

Abstract : During the 20th century, several skates and rays species in European waters declined because of fishing. Conservation of these species is a major objective of the management of marine resources. The thornback ray (*Raja clavata*) is the most widespread species of the North-Est Atlantic. Its management is based on a nonspecific quota and lay on observations only as no stock assessment is available. Thus this thesis aims to compare the available stock assessments methods for this species and to apply them to empiric data. Two types of methods are investigated: methods based on population demography and methods based on population genetics. The thesis is separated in three parts: 1) assessment of management units; 2) assessment of the state of the thornback ray populations; 3) comparison of the thornback ray population's states with the other rays species of the Northeast Atlantic.
Results highlight the usefulness of the demographics methods for the harvest of the species. Indeed, the developed halieutics models are able to estimate the biomass state of the thornback ray by using tendencies. The Bay of Biscay stock was estimated in a good ecological state. Genetics methods are very useful in conservation as they inform on the genetic health as a long timescale.