



Pêches et Océans
Canada

Fisheries and Oceans
Canada

Sciences des écosystèmes
et des océans

Ecosystems and
Oceans Science

Secrétariat canadien de consultation scientifique (SCCS)

Compte rendu 2020/001

Région du Pacifique

**Compte rendu de l'examen par les pairs de la région du Pacifique sur la
Modélisation de la qualité de l'habitat – Pratiques exemplaires pour l'océan
Pacifique du Canada**

**Les 11 et 12 juin 2019
Nanaimo (Colombie-Britannique)**

**Président : Steven Schut
Rapporteurs : Steven Schut et Jill Campbell**

Station biologique du Pacifique
Pêches et Océans Canada
3190, chemin Hammond Bay
Nanaimo (Colombie-Britannique) V9T 6N7

Avant-propos

Le présent compte rendu a pour but de consigner les principales activités et discussions qui ont eu lieu au cours de la réunion. Il peut contenir des recommandations sur les recherches à effectuer, des incertitudes et les justifications des décisions prises pendant la réunion. Le compte rendu peut aussi faire état de données, d'analyses ou d'interprétations passées en revue et rejetées pour des raisons scientifiques, en donnant la raison du rejet. Bien que les interprétations et les opinions contenues dans le présent rapport puissent être inexactes ou propres à induire en erreur, elles sont quand même reproduites aussi fidèlement que possible afin de refléter les échanges tenus au cours de la réunion. Ainsi, aucune partie de ce rapport ne doit être considérée en tant que reflet des conclusions de la réunion, à moins d'une indication précise en ce sens. De plus, un examen ultérieur de la question pourrait entraîner des changements aux conclusions, notamment si des renseignements supplémentaires pertinents, non disponibles au moment de la réunion, sont fournis par la suite. Finalement, dans les rares cas où des opinions divergentes sont exprimées officiellement, celles-ci sont également consignées dans les annexes du compte rendu.

Publié par :

Pêches et Océans Canada
Secrétariat canadien de consultation scientifique
200, rue Kent
Ottawa (Ontario) K1A 0E6

[http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca](http://www.dfo-mpo.gc.ca/csas-sccs/csas-sccs@dfo-mpo.gc.ca)



© Sa Majesté la Reine du chef du Canada, 2020
ISSN 2292-4264

La présente publication doit être citée comme suit :

MPO. 2020. Compte rendu de l'examen par les pairs de la région du Pacifique sur la Modélisation de la qualité de l'habitat – Pratiques exemplaires pour l'océan Pacifique du Canada; 11 et 12 juin 2019. Secr. can. de consult. sci. du MPO, Compte rendu 2020/001.

Also available in English:

DFO. 2020. *Proceedings of the Pacific Regional Peer Review on Habitat Suitability Modelling Best Practices for Canada's Pacific Ocean; June 11-12, 2019. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2020/001.*

TABLE DES MATIÈRES

SOMMAIRE	v
INTRODUCTION	1
PRÉSENTATION DU DOCUMENT DE TRAVAIL	2
RÉSUMÉ DU DOCUMENT DE RECHERCHE	2
CADRE	2
PRÉSENTATION DES EXAMENS ÉCRITS DU CADRE	2
Chris Rooper, Secteur des sciences du MPO, Sciences des écosystèmes	2
Javier Murillo-Perez, Secteur des sciences du MPO, Maritimes	3
Anders Knudby, Université d'Ottawa	3
DISCUSSION GÉNÉRALE SUR LE CADRE	4
Terminologie	4
Composition du modèle d'ensemble	4
Choix des modèles	5
Utilisation des modèles d'application	5
Incertitude du modèle	5
Divers	6
APPLICATION	6
PRÉSENTATION DES EXAMENS ÉCRITS DE L'APPLICATION	6
Chris Rooper, Secteur des sciences du MPO, Sciences des écosystèmes	6
Javier Murillo-Perez, Secteur des sciences du MPO, Maritimes	7
Anders Knudby, Université d'Ottawa	7
DISCUSSION GÉNÉRALE SUR L'APPLICATION	8
Agrégation des données	8
Inclusion des ensembles de données dans les modèles	8
Termes d'interaction du MLG	9
Prédicteurs environnementaux supplémentaires	10
Seuil d'inclusion de la valeur AUC dans le modèle d'ensemble	10
Relations entre les paramètres et figures illustrant les sorties des modèles	10
Divers	10
CONCLUSIONS	10
RÉVISIONS DU DOCUMENT DE TRAVAIL	11
CADRE	11
APPLICATION	12
RECOMMANDATIONS	13
REMERCIEMENTS	13
ANNEXE A : CADRE DE RÉFÉRENCE	14
ANNEXE B ORDRE DU JOUR	17
ANNEXE C : LISTE DES PARTICIPANTS	19
ANNEXE D : EXAMENS DU DOCUMENT DE TRAVAIL	20

EXAMINATEUR : CHRIS ROOPER, SECTEUR DES SCIENCES DU MPO, SCIENCES DES ÉCOSYSTÈMES.....	20
EXAMINATEUR : JAVIER MURILLO-PEREZ, SECTEUR DES SCIENCES DU MPO, RÉGION DES MARITIMES.....	23
EXAMINATEUR : ANDERS KNUDBY, UNIVERSITÉ D'OTTAWA.....	28

SOMMAIRE

Le présent compte rendu résume les discussions pertinentes et les principales conclusions de la réunion régionale d'examen par les pairs du Secrétariat canadien de consultation scientifique (SCCS) de Pêches et Océans Canada (MPO), tenue les 11 et 12 juin 2019 à la Station biologique du Pacifique, à Nanaimo (Colombie-Britannique). Un document de travail établissant le cadre pour les modèles de qualité de l'habitat et les modèles de répartition des espèces, ainsi qu'un exemple d'application de ce cadre, a été déposé aux fins d'examen par les pairs.

Les principaux sujets examinés portaient sur les points suivants : les différences entre les modèles de qualité de l'habitat et les modèles de répartition des espèces et les prévisions produites par les modèles générés dans le cadre; l'inclusion du modèle d'indice de qualité de l'habitat dans le modèle d'ensemble; l'inclusion d'autres prédicteurs environnementaux; la manière d'agréger plusieurs points de données dans une seule cellule matricielle; et s'il est possible ou obligatoire de combiner les données obtenues à partir de différents engins de relevé.

Les participants, en personne et sur le Web, représentaient les Secteurs des sciences et de la gestion des écosystèmes de Pêches et Océans Canada (MPO), ainsi que des organisations des Premières Nations et le milieu universitaire.

Les conclusions et l'avis découlant de cet examen seront présentés sous la forme d'un avis scientifique fournissant des conseils au Secteur des sciences du MPO pour guider la modélisation de la qualité de l'habitat et de la répartition des espèces.

L'avis scientifique et le document de recherche à l'appui seront rendus publics sur le site Web du [Secrétariat canadien de consultation scientifique](#) (SCCS).

INTRODUCTION

Une réunion régionale d'examen par les pairs a eu lieu les 11 et 12 juin 2019 à la Station biologique du Pacifique à Nanaimo afin d'évaluer le cadre des pratiques exemplaires pour l'élaboration de modèles de qualité de l'habitat et de répartition des espèces établis dans le document de travail.

Le cadre de référence de l'examen scientifique (annexe A) a été élaboré en réponse à une demande d'avis scientifique du Secteur des sciences de Pêches et Océans Canada. Les avis d'examen scientifique et les conditions de participation ont été envoyés aux représentants possédant l'expertise pertinente de Pêches et Océans Canada, des Premières Nations, d'organisations non gouvernementales et du milieu universitaire.

Le document de travail suivant a été rédigé et remis aux participants avant la réunion :

Nephin, J., Gregr, E.J., St. Germain, C., Fields, C. et Finney, J.L. Modélisation de la qualité de l'habitat – Pratiques exemplaires pour l'océan Pacifique du Canada. Document de travail du CASP 2018SCI01.

Le président de la réunion, Steven Schut, souhaite la bienvenue aux participants, passe en revue le rôle du SCCS dans la fourniture des avis évalués par les pairs et donne un aperçu général du processus du SCCS. Le président discute du rôle des participants, de l'objet des diverses publications de la réunion régionale d'examen par les pairs (avis scientifique, compte rendu et document de recherche), ainsi que de la définition et du processus à suivre pour parvenir à des décisions et à des avis consensuels. Chaque personne est invitée à participer pleinement à la discussion et à faire part de ses connaissances pendant le processus, dans le but de formuler des conclusions et des avis défendables sur le plan scientifique. Les participants confirment qu'ils ont tous reçu des copies du cadre de référence, du document de travail et de l'ébauche de l'avis scientifique (AS).

Le président passe en revue l'ordre du jour (annexe B) et le cadre de référence de la réunion, en soulignant les objectifs et en désignant Jill Campbell comme rapporteur pour l'examen. Il décrit ensuite les règles de base et le processus d'échange durant la réunion, en rappelant aux participants que la réunion sert d'examen scientifique et non de consultation. La salle est équipée de microphones pour permettre la participation à distance par conférence Web, et on rappelle aux participants en personne de répondre aux commentaires et aux questions de façon à ce que les participants en ligne les entendent.

On rappelle aux membres que tous les participants à la réunion sont sur un pied d'égalité en tant que participants et qu'ils sont censés apporter leur contribution au processus d'examen s'ils ont des renseignements ou des questions concernant le document de travail faisant l'objet des discussions. Au total, 39 personnes ont participé à l'examen régional par les pairs (annexe C).

On informe les participants que Chris Rooper, Javier Murillo-Perez et Anders Knudby ont été invités avant la réunion à fournir un examen écrit détaillé du document de travail afin d'aider tous les participants à la réunion d'examen par les pairs. Les participants ont reçu des copies des examens écrits.

Les conclusions et l'avis découlant de cet examen seront présentés sous la forme d'un avis scientifique au Secteur des sciences du MPO pour guider la modélisation de la qualité de l'habitat et de la répartition des espèces. L'avis scientifique et le document de recherche à l'appui seront rendus publics sur le site Web du [Secrétariat canadien de consultation scientifique](#) (SCCS).

PRÉSENTATION DU DOCUMENT DE TRAVAIL

Tous les auteurs sont présents : J. Nephin, E.J. Gregr, C. St. Germain, C. Fields et J.L. Finney. Jessica Nephin, Jessica Finney et Candice St. Germain présentent un exposé oral récapitulant le document de travail décrit dans le résumé suivant.

RÉSUMÉ DU DOCUMENT DE RECHERCHE

Les modèles décrivant les habitats convenables pour une espèce sont essentiels pour prendre des décisions éclairées sur la gestion et la conservation des espèces et des lieux marins. Des modèles de qualité de l'habitat (MQH) sont utiles pour plusieurs objectifs nationaux liés à la planification spatiale marine, aux évaluations de la vulnérabilité, aux interventions d'urgence et à l'évaluation des stocks. En tant qu'organisme responsable de l'élaboration des données scientifiques qui sous-tendent ces objectifs stratégiques, Pêches et Océans Canada (MPO) a besoin que les méthodes de MQH soient appliquées de manière uniforme. L'application efficace des MQH dépend des contextes écologiques et de gestion pertinents, de la préparation uniforme des données disponibles et de l'application de méthodes d'analyse appropriées pour construire des modèles et les évaluer.

Ce cadre, mis en œuvre avec des scripts conçus à cette fin et rédigés dans le langage de programmation statistique R, a été préparé comme un ensemble de lignes directrices sur l'élaboration de MQH uniformes, interprétables et défendables afin d'appuyer la contribution du MPO aux politiques canadiennes sur les océans. Le cadre vise à aider les analystes qui souhaitent élaborer des MQH et les autres personnes qui s'intéressent aux pratiques exemplaires liées à la compréhension, à l'intégration et à la communication des résultats de ces modèles à un public plus vaste. Pour le MPO en particulier, les défis liés à la préparation de données d'observation et de prédicteurs marins peuvent être particulièrement intéressants.

Dans la mesure du possible, le cadre met en œuvre des pratiques exemplaires reconnues pour les aspects clés de la préparation des données, de l'adaptation des modèles et de l'évaluation. Ces méthodes sont illustrées à l'aide de trois approches de construction de modèles de complexité croissante, qui sont combinées en un modèle d'ensemble final. En fournissant une collection de résultats statistiques, graphiques et cartographiés, le cadre offre une plateforme pour l'élaboration efficace et normalisée d'une série de MQH utiles pour guider la gestion des ressources marines vivantes.

L'application du cadre est démontrée par l'utilisation d'une série des meilleurs prédicteurs environnementaux disponibles pour prévoir l'habitat potentiel de douze espèces valorisées sur la côte canadienne du Pacifique. Ces espèces comprenaient des invertébrés, des plantes formant un habitat et des poissons. Une modélisation fondée sur des hypothèses a été appliquée à toutes les espèces, en mettant l'accent sur celles pour lesquelles on manquait de données. L'ensemble de prédicteurs, de méthodes et d'outils d'évaluation communs réunis ici facilitera grandement la future élaboration de MQH pour d'autres espèces présentes dans la partie canadienne de l'océan Pacifique.

CADRE

PRÉSENTATION DES EXAMENS ÉCRITS DU CADRE

Chris Rooper, Secteur des sciences du MPO, Sciences des écosystèmes

Veuillez consulter l'examen écrit complet à l'annexe D. Voici les principaux commentaires présentés par Chris Rooper à la réunion de l'examen par les pairs régional :

-
- D'autres types de modélisation pourraient être envisagés, y compris l'entropie maximale, qui est de plus en plus utilisée.
 - Il faudrait mettre à l'essai les modèles du cadre au moyen de données recueillies de façon indépendante ou collectées à l'aide d'un engin de relevé différent de celui utilisé dans l'élaboration des modèles.
 - Il faudrait inclure une définition claire des MQH et des modèles de répartition des espèces (MRW), car ces définitions ne sont claires ni dans le document de travail, ni dans la documentation de façon plus générale. De plus, il serait avantageux de discuter plus à fond des niches réalisées et fondamentales dans la mesure où elles s'appliquent aux MQH et aux MRE.

Les auteurs n'ont pas besoin d'autres précisions sur cette présentation.

Javier Murillo-Perez, Secteur des sciences du MPO, Maritimes

Veillez consulter l'examen écrit complet à l'annexe D. Voici les principaux commentaires présentés par Javier Murillo-Perez à la réunion de l'examen par les pairs régional :

- On pourrait envisager d'autres types de modélisation que les trois présentés ici, notamment des modèles additifs généralisés (MAG) ou des modèles de forêts d'arbres décisionnels. L'utilisation de MAG pourrait fournir d'autres techniques pour définir les prédicteurs spatiaux et l'autocorrélation spatiale (ACS).
- Il faudrait mettre à l'essai les modèles du cadre au moyen de données collectées à l'aide d'un engin de relevé différent de celui utilisé dans l'élaboration des modèles.
- Des modèles semblables élaborés dans les Maritimes ont utilisé des couches superposées pour démontrer l'incertitude des prédicteurs, et on pourrait reprendre des méthodes semblables ici.

Les auteurs n'ont pas besoin d'autres précisions sur cette présentation.

Anders Knudby, Université d'Ottawa

Veillez consulter l'examen écrit complet à l'annexe D. Voici les principaux commentaires présentés par Anders Knudby à la réunion de l'examen par les pairs régional :

- Il faudrait inclure une définition claire des MQH et des modèles de répartition des espèces (MRE), car ces définitions ne sont claires ni dans le document de travail, ni dans la documentation de façon plus générale. De plus, il serait avantageux de discuter plus à fond des niches réalisées et fondamentales dans la mesure où elles s'appliquent aux MQH et aux MRE. La distinction entre niche réalisée et fondamentale peut avoir des répercussions sur les objectifs du modèle final.
- Un participant demande comment les modèles se comparent avec ou sans l'utilisation de l'ACS. Des préoccupations sont soulevées au sujet de la figure 3.13 (*Map of spatial auto-covariate predictor that represents the distance weighted sum of nearby observations*, Carte du prédicteur spatial auto-covariable qui représente la somme pondérée de la distance des observations à proximité) et de la façon dont les cellules sans résiduel calculé ont été traitées à l'étape de l'autocorrélation spatiale.
- Le rendement relatif de chaque modèle devrait être affiché pour chaque espèce puisque l'indice de qualité de l'habitat (IQH) est différent du modèle linéaire généralisé (MLG) et de

l'arbre de régression stimulée (ARS). Une étape pourrait être ajoutée au cadre pour décider s'il faut ajouter l'IQH à l'ensemble.

Les auteurs précisent que l'auto-covariable (ACV) n'a pas été interpolée et n'a été utilisée que pour l'ajustement du modèle et l'évaluation des données d'essai du modèle. Le rôle de l'IQH dans l'ensemble est défini par sa valeur AUC (zone sous la courbe de la caractéristique de fonctionnement du récepteur).

DISCUSSION GÉNÉRALE SUR LE CADRE

Terminologie

Les examinateurs, et par la suite les participants, demandent que les modèles de qualité de l'habitat (MQH) soient définis plus clairement dans le document. Les modèles de l'indice de qualité de l'habitat (IQH), fondés sur des conseils d'experts, cartographient de façon plus réaliste les niches fondamentales et devraient donc être décrits comme des MQH, alors que les modèles MLG et ARS (modèles corrélatifs), fondés sur des observations de présence/absence, cartographient de façon plus réaliste les niches/répartitions réalisées et devraient donc être décrits comme des modèles de répartition des espèces (MRE). Toutefois, un participant fait remarquer que, selon les questions d'entrevue posées aux experts, les modèles IQH peuvent fournir de l'information sur les niches fondamentales ou réalisées. Cependant, il est décidé que les modèles IQH ne prédisent pas les mêmes choses que les modèles corrélatifs, car les conseils d'experts ne tiennent pas compte des facteurs de probabilité d'occurrence comme l'effort ou l'engin utilisé pour le relevé. En outre, un participant indique que l'on observera vraisemblablement la présence d'espèces sans conseils spécialisés, car les experts sont plus susceptibles de pécher par excès de prudence (niches réalisées). De plus, les modèles IQH ont généralement sous-prédict l'habitat en raison des méthodes de paramètres limitatifs. Il est recommandé d'élaborer un modèle d'ensemble de tous les modèles IQH à partir d'entrevues individuelles avec des experts. Un participant précise qu'il existe un spectre allant de la qualité de l'habitat à la répartition des espèces et que les modèles générés ici se situent quelque part dans ce spectre.

Résultats : Des définitions claires de MQH, MRE et IQH seront ajoutées au glossaire du document et une discussion élargie sur les niches fondamentales et réalisées sera incluse dans le corps du document. Le nom du document sera modifié pour indiquer que ce cadre permet de créer des MRE (plutôt que des MQH).

Composition du modèle d'ensemble

Étant donné qu'il a été déterminé que les modèles IQH modélisent la qualité de l'habitat des espèces et que les modèles corrélatifs modélisent la répartition des espèces, les participants s'inquiètent de l'inclusion du modèle IQH dans le modèle d'ensemble. Les auteurs démontrent que lorsque le modèle IQH n'est pas inclus dans le modèle d'ensemble, il y a très peu de différence dans les extrants. Ils élaborent également une figure pour montrer où les deux types de modèles diffèrent sur le plan spatial, et les participants notent que le modèle IQH semble sous-estimer la prévision d'un habitat convenable. Un participant pense que les zones où la différence est plus marquée dans la qualité de l'habitat entre les types de modèles seraient de bons endroits à échantillonner les années suivantes pour mettre les modèles à l'essai.

Résultats : Recommander aux utilisateurs finaux d'examiner les conséquences de l'inclusion de l'IQH dans leurs modèles d'ensemble. L'IQH peut être plus utile avec les espèces pour lesquelles on dispose de peu de données. Recommander aux utilisateurs finaux d'élaborer des cartes de chaque type de modèle individuellement et une carte du modèle d'ensemble, y

compris des modèles corrélatifs seulement. Recommander aux utilisateurs finaux d'élaborer des cartes appariant chaque ensemble de modèles pour déterminer où se situent les différences de prédiction de la modélisation. Ajouter des emplacements d'échantillonnage sur les cartes pour déterminer la corrélation spatiale ou le biais d'échantillonnage. Recommander aux utilisateurs finaux d'avoir des objectifs de modélisation clairs pour les aider à prendre ces décisions.

Choix des modèles

Les participants indiquent que d'autres types de modèles qui pourraient être utilisés dans ce cadre n'ont pas été mentionnés (MAG, modèles de forêts d'arbres décisionnels, Maxent). Les auteurs font remarquer que les autres modèles n'ont pas été exclus délibérément, mais plutôt que les modèles utilisés ont été choisis pour couvrir un éventail de complexités de la modélisation. Il est décidé que le cadre ne vise pas à recommander à l'utilisateur final d'utiliser des modèles précis.

Résultats : Une série de considérations a été élaborée pour aider les utilisateurs finaux à choisir le modèle. Les auteurs préciseront qu'il existe un corpus de travaux sur les types de modèles.

Utilisation des modèles d'application

Les modèles des espèces ont été élaborés pour la planification des interventions en cas de déversement d'hydrocarbures et pour les essais du cadre. Toutefois, cette réunion du SCCS ne cherche pas à évaluer l'acceptation de ces douze modèles d'espèces. Il est décidé que le processus de cette réunion n'est pas suffisant pour accepter les modèles d'espèces tels quels et que le présent document ne formule pas de recommandations pour reprendre ces modèles sans autre développement. Toutefois, un participant fait remarquer que ces modèles constituent les meilleures données disponibles; ils seront probablement utilisés jusqu'à ce que des modèles plus complets puissent être élaborés.

Résultats : Il est décidé d'ajouter au document un texte indiquant que ces modèles doivent être utilisés avec prudence et d'ajouter également un texte semblable à toutes les légendes des figures des modèles d'espèces. Les modèles d'application sont des études de cas plutôt que des modèles reconnus de répartition des espèces.

Incertitude du modèle

Les participants discutent de la différence entre l'incertitude de la validation croisée et l'incertitude du modèle et de la meilleure façon de communiquer ces différences. Un participant souligne que l'incertitude de la validation croisée a plus d'effet sur l'incertitude du modèle que l'incertitude des paramètres. En outre, il faudrait clarifier les différences entre l'incertitude du modèle et l'incertitude des facteurs. Les participants déterminent que la pratique exemplaire actuelle consiste à examiner les résiduels par rapport aux valeurs prévues et aux paramètres du modèle. Toutefois, comme les données sont binaires (présence/absence), il peut être difficile d'interpréter les résiduels et une analyse plus poussée peut être nécessaire (classification par fenêtres, utilisation de résiduels quantiles randomisés ou recherche d'une taille de bloc spatial optimale). L'observation de la carte des résiduels met en évidence les schémas spatiaux, auquel cas il peut être nécessaire d'utiliser un terme auto-covariable.

Résultats : Il est décidé de fournir des lignes directrices à l'utilisateur final au sujet des zones d'incertitude et de la meilleure façon de les déterminer, avec une formulation précise élaborée pour refléter l'incertitude du cadre et des données.

Divers

Un participant est préoccupé par le fait que les CTE (connaissances traditionnelles écologiques) ne sont pas incluses dans les modèles IQH. Ces connaissances devraient être incluses dans les modèles IQH et pourraient servir d'essai du modèle pour les modèles IQH existants. Les CTE pourraient également être utilisées pour examiner les cartes de l'habitat des espèces.

Un examinateur mentionne qu'il faudrait utiliser des données indépendantes pour mettre les modèles à l'essai. Le groupe détermine qu'il s'agit d'une pratique exemplaire actuelle. Les essais par rapport à un autre type d'engin de relevé ne seraient pas suffisants car ils pourraient mettre en évidence un biais, la répartition totale des espèces ou le type d'habitat complet n'étant pas nécessairement couverts.

Un examinateur suggère de définir un modèle idéal, mais il est décidé qu'il n'est pas possible d'en créer un puisque les objectifs du modèle varieront selon l'utilisateur final et qu'un modèle exemplaire n'est pas réalisable car il pourrait décourager l'utilisation de ce cadre.

L'organigramme du cadre devait être mis à jour pour refléter les essais du modèle à l'aide de données indépendantes et pour le rendre plus général plutôt que propre aux méthodes utilisées dans cette application.

Un participant fait remarquer qu'il existe plusieurs façons d'inclure la structure spatiale, comme les champs aléatoires gaussiens. Il est décidé d'ajouter le texte au document de travail à cet effet.

Un examinateur propose d'ajouter une superposition d'extrapolation sur les cartes de sortie du modèle. On conclut que cette question ne relève pas de la portée de ce cadre.

Un participant demande des précisions sur la façon dont le terme d'autocorrélation spatiale (ACS) a été utilisé. Il n'est pas certain que ce terme est inclus dans les prévisions finales. Les auteurs ajouteront du texte dans le document pour préciser que l'ACS a été utilisée uniquement pour déterminer si elle avait une incidence.

On demande des précisions sur la façon dont le modèle ARS a été réglé. Les auteurs préciseront que les points de données des paramètres sont supprimés au hasard et que les ensembles de données de réglage et d'essai sont spatiaux.

Des participants suggèrent aux auteurs d'inclure un texte recommandant à l'utilisateur final de tenir compte de la couverture spatiale de ses données et des différentes conditions le long de la côte lorsqu'il désigne le domaine du modèle.

APPLICATION

PRÉSENTATION DES EXAMENS ÉCRITS DE L'APPLICATION

Chris Rooper, Secteur des sciences du MPO, Sciences des écosystèmes

Veuillez consulter l'examen écrit complet à l'annexe D. Voici les principaux commentaires présentés par Chris Rooper à la réunion de l'examen par les pairs régional :

- De nombreuses espèces n'ont pas été échantillonnées dans l'ensemble de leur habitat. Les sources de données utilisées ne sont pas optimales aux fins de l'élaboration de ce modèle. Il faut justifier le choix des ensembles de données inclus dans les modèles, car il semble y avoir des incohérences.
- Le seuil d'inclusion du modèle dans le modèle d'ensemble (AUC <0,5) pourrait être plus élevé.

-
- La section des conclusions devrait présenter des conclusions plus générales.
 - Les modèles IQH et corrélatifs prédisent différentes choses et doivent être traités différemment.
 - Ajouter des figures pour montrer toutes les cartes du modèle ensemble et représenter toutes les relations de données sur une seule figure.
 - Il semble que la valeur AUC moyenne des cinq modèles de validation croisée soit supérieure à celle des modèles individuels, ce qui n'est pas logique.

Les auteurs ont démontré que la valeur AUC du modèle d'ensemble n'est pas plus élevée que celle des modèles individuels. De même, si un ou deux modèles de validation croisée donnent de mauvais résultats, le modèle IQH pourrait compenser lorsqu'il est inclus dans le modèle d'ensemble.

Javier Murillo-Perez, Secteur des sciences du MPO, Maritimes

Veillez consulter l'examen écrit complet à l'annexe D. Voici les principaux commentaires présentés par Javier Murillo-Perez à la réunion de l'examen par les pairs régional :

- D'autres variables prédictives pourraient être utilisées. Elles pourraient également être ajoutées pour tenir compte des incohérences attribuables au type d'engin de relevé.
- La création d'une couche de substrat secondaire pourrait être utile en ce qui concerne la perte de données lors de la conversion de données linéaires en données ponctuelles. Ou convertir le substrat en points de données proportionnels.
- Les prédicteurs environnementaux semblent porter sur le printemps ou l'été seulement.
- Il serait utile d'inclure une analyse complète des quatre espèces qui n'ont pas été entièrement modélisées pour discuter de l'influence de la qualité des données sur la sortie du modèle.
- Recommander d'inclure les points de données sur la présence ou l'absence sur les cartes.

Les auteurs expliquent que les données des prédicteurs environnementaux du printemps et de l'été sont plus fortement corrélées avec la répartition des espèces. Cette précision sera ajoutée dans le document.

Anders Knudby, Université d'Ottawa

Veillez consulter l'examen écrit complet à l'annexe D. Voici les principaux commentaires présentés par Anders Knudby à la réunion de l'examen par les pairs régional :

- Il faut préciser ce qu'est le modèle écologique et comment il est appliqué.
- Il faut poursuivre la discussion sur la façon dont les termes d'interaction des MLG ont été dérivés.
- Préoccupations entourant la façon d'inclure le modèle IQH avec les modèles corrélatifs dans l'ensemble.

Les auteurs décrivent la façon dont le modèle écologique a été inclus pour déterminer si les prédicteurs ont un effet quadratique ou d'interaction dans le MLG et dans le processus de sélection des prédicteurs, car certains prédicteurs n'ont pas été utilisés si une relation n'était pas indiquée dans le modèle écologique.

DISCUSSION GÉNÉRALE SUR L'APPLICATION

Agrégation des données

La préparation des données a consisté à regrouper les données sur la présence ou l'absence au fil du temps dans une cellule matricielle. Des participants craignent que cette méthode ne soit pas une bonne façon d'éclaircir les données et que des observations soient perdues. Dans les données d'entrée du modèle, il serait plus précis de remplacer la valeur binaire présence/absence par une probabilité d'occurrence dans cette cellule matricielle. L'agrégation des données aura une incidence sur la pondération des cellules comportant de nombreuses observations comparativement aux cellules comportant peu d'observations. Cela serait particulièrement fréquent pendant les relevés linéaires sur plusieurs cellules, car ils auraient tendance à surprédire la présence. L'échantillonnage aléatoire de points de données est une autre solution envisagée, mais cette méthode est également considérée comme entraînant une perte d'observations, et pourrait être influencée par des couches de prédicteurs environnementaux qui varient selon l'année et sont fortement corrélées à la présence des espèces. Un autre participant propose de calculer la moyenne des données pour l'ensemble des chaluts et de créer des données proportionnelles plutôt que des données binaires, mais on ne sait pas trop comment cela pourrait être intégré au cadre du modèle et les participants rejettent cette méthode. En fin de compte, la méthode utilisée dépendra des objectifs de l'utilisateur final.

Résultats : Les auteurs préciseront dans le texte la façon dont leurs données ont été agrégées, les implications correspondantes et le fait que d'autres méthodes sont décrites dans la documentation.

Inclusion des ensembles de données dans les modèles

Des préoccupations sont soulevées au sujet de la justification de l'inclusion des données de relevé dans les modèles du sébaste et du crabe dormeur. Les données tirées du relevé du sébaste au chalut n'ont pas été utilisées, car les auteurs ne voulaient pas combiner les différentes données obtenues à l'aide de différents types d'engins de relevé (comme cela a été fait pour les modèles des autres espèces), mais le modèle du crabe dormeur a été élaboré avec les données tirées des relevés au casier, au chalut et en plongée. Les participants notent que le type d'engin modifie la détectabilité, la prévisibilité des prises et le type d'habitat. Les auteurs ont préparé une présentation sur le modèle du crabe dormeur et demandent aux participants des conseils sur la meilleure façon de combiner différents ensembles de données. Les participants craignent que les données de plongée prédisent l'habitat dans une zone plus profonde qu'elles ne le devraient, peut-être en raison d'erreurs de géoréférencement dans les données. Chaque ensemble de données produit des cartes de répartition des espèces différentes, mais lorsque l'on combine les ensembles de données en un seul modèle, la carte permet des prédictions sur une échelle spatiale plus grande. Certains participants s'inquiètent du fait que la valeur AUC des modèles d'entraînement est élevée ($> 0,7$) et que leurs courbes ROC sont très variables, ce qui montre que les statistiques utilisées peuvent être trompeuses. Un participant suggère d'élaborer un facteur de mise à l'échelle ou d'étalonnage entre les types d'engins pour aplanir les différences de capturabilité. Les participants discutent de l'ensemble de données à utiliser, qu'il s'agisse de votre « meilleur » ensemble de données ou de celui qui cible le plus directement les espèces modélisées (dans le cas du crabe dormeur, il s'agirait des données sur les relevés au casier) puisqu'il s'agit d'un habitat convenable connu. Une pratique exemplaire actuelle consiste à exécuter chaque ensemble de données dans le modèle ainsi que l'ensemble de données combiné et à comparer les cartes, comme les auteurs l'ont fait dans leur présentation, ce qui permettrait aux utilisateurs finaux de voir comment leurs ensembles de

données se comparent les uns aux autres. Le document présentera des lignes directrices sur la façon de choisir l'ensemble de données le plus approprié à utiliser. L'utilisateur final doit garder à l'esprit ses objectifs de modélisation lorsqu'il choisit l'ensemble de données, car différents types de données prédisent différentes choses (c.-à-d. que les données tirées du relevé au chalut concernent uniquement l'habitat des substrats meubles et indiquent les endroits où il sera sans doute possible de pêcher des crabes, mais ne comprennent pas tous les habitats convenables). De plus, il doit être conscient des lacunes dans les données et des fausses absences (c.-à-d. que les estuaires ou les deltas sont rarement échantillonnés, alors que c'est là que les crabes femelles et juvéniles sont plus susceptibles de se trouver et que ces emplacements ne sont pas indiqués dans les modèles). Un participant propose de pouvoir combiner chaque ensemble de données modélisé en un modèle d'ensemble en ajoutant un facteur de pondération aux valeurs AUC; on note cependant que les données des relevés en plongée ont une valeur AUC élevée, mais ne prédisent pas bien l'habitat. Un participant demande si les modèles de validation croisée avec de faibles valeurs AUC devraient être inclus dans les modèles. Les auteurs expliquent que les cinq modèles de validation croisée étaient inclus et qu'ils entraînent un écart-type important. Un participant indique que ces ensembles de données sont biaisés sur le plan spatial, tout en extrapolant sur une zone étendue, et remet en question la justification de cette démarche. Les auteurs ont choisi de le faire parce qu'ils n'estimaient pas qu'il y avait une grande différence dans la façon dont les données étaient recueillies dans les biorégions de la plate-forme Nord et de la plate-forme Sud. Les auteurs notent également que cette question a été abordée dans leur choix des échelles spatiales des régions côtières et des plates-formes.

Résultats : Les auteurs fourniront d'autres précisions dans leur document pour justifier leur décision concernant les différences de sélection des données entre le sébaste et le crabe dormeur. Ils recommanderont aux utilisateurs finaux de créer des cartes à l'aide de chacun de leurs ensembles de données pour explorer la façon dont leurs données peuvent être biaisées. De plus, pour répondre aux préoccupations entourant les répercussions de l'extrapolation sur une échelle spatiale plus grande que l'ensemble de données initial, il est décidé qu'il est nécessaire d'ajouter des lignes directrices dans le document pour informer les utilisateurs finaux de cette préoccupation et les encourager à tenir compte de l'échelle de leurs ensembles de données, des prédicteurs environnementaux qu'ils choisissent, et à avoir des objectifs de modèle clairs. Des lignes directrices seront également ajoutées pour que les utilisateurs finaux envisagent d'utiliser le modèle IQH plutôt que les modèles corrélatifs si les ensembles de données sur les engins de relevé ne donnent pas des cartes exactes de la répartition des espèces.

Termes d'interaction du MLG

Un participant fait valoir que des interactions autres que celles figurant dans le MLG pourraient se produire. Seules les interactions connues ont été incluses dans le modèle, pas toutes les interactions possibles. Certains participants ne comprennent pas bien si des termes d'interaction ont été omis pour des raisons de calcul ou des raisons écologiques. Les auteurs déclarent que les termes d'interaction ont été inclus sur les conseils d'experts et que compte tenu de la forte corrélation de certains prédicteurs environnementaux, l'inclusion de ces termes d'interaction pourrait entraîner une surcorrélacion.

Résultats : Il est décidé de préciser dans le texte que des raisons écologiques limitent le nombre de termes d'interaction utilisés et que les interactions peuvent changer à la suite de recherches ultérieures ou sous l'effet des changements climatiques.

Prédicteurs environnementaux supplémentaires

Les participants discutent de la création de prédicteurs environnementaux supplémentaires, peut-être fondés sur l'écologie des espèces individuelles, sur des échelles spatiales plus petites ou sur les conditions associées aux événements de mortalité. Cependant, le but de ce document n'est pas d'élaborer les meilleurs modèles pour chaque espèce, mais un cadre.

Résultats : Un texte sera ajouté au document pour informer les utilisateurs finaux de la sélection des prédicteurs environnementaux propres aux espèces et pour indiquer qu'il existe beaucoup plus de prédicteurs environnementaux que ceux utilisés dans cette application.

Seuil d'inclusion de la valeur AUC dans le modèle d'ensemble

Les participants discutent de la possibilité de relever le seuil de la valeur AUC pour l'inclusion des modèles de validation croisée dans le modèle d'ensemble. Dans le document de travail, c'est la valeur de 0,5 qui a été retenue car toute valeur supérieure est mieux qu'aléatoire et tout autre seuil peut être interprété comme étant arbitraire. De plus, un seuil trop élevé de la valeur AUC pourrait empêcher de créer des modèles.

Résultats : La valeur restera fixée 0,5 et les auteurs fourniront aux utilisateurs finaux des lignes directrices sur la façon d'examiner l'entente sur le modèle, l'écart-type entre les modèles de validation croisée et la qualité des données afin de leur permettre de prendre leurs propres décisions au sujet des seuils de la valeur AUC lorsqu'ils créeront leurs modèles.

Relations entre les paramètres et figures illustrant les sorties des modèles

Les participants demandent des figures supplémentaires, notamment une figure illustrant toutes les relations entre les paramètres et une figure en six graphiques avec une carte des sorties de chaque modèle. Ces figures aideront les utilisateurs finaux à analyser pleinement les sorties de leur modèle. Afin d'accroître la résolution des modèles des régions côtières, on ajoutera une superposition en zoom avant du modèle sur la zone terrestre sur les cartes plus vastes de la biorégion de la plate-forme.

Résultats : Afin de réduire la quantité de travail pour les auteurs, il est convenu de produire une série complète de figures pour un exemple d'espèce pour laquelle on dispose de données de grande qualité et un exemple d'espèce pour laquelle on dispose de données de faible qualité.

Divers

Les participants remarquent que les espèces n'ont pas été échantillonnées dans l'ensemble de leur habitat et que cela pourrait fausser les modèles. Un texte sera ajouté pour avertir les utilisateurs finaux de tenir compte de l'étendue spatiale de leurs données lorsqu'ils élaboreront leurs propres modèles.

Il sera précisé dans le texte que la corrélation de Spearman s'ajoute au facteur d'inflation de la variance.

Certains participants prennent part à la modélisation d'archétypes et se demandent comment elle pourrait s'appliquer au cadre. Il est décidé que la modélisation d'archétypes sort de la portée du document, mais qu'elle sera mentionnée comme une possibilité future.

CONCLUSIONS

Le document de travail est accepté avec des modifications. Les participants conviennent que le cadre de référence a été respecté. Les conclusions suivantes figurent également dans l'avis scientifique.

-
- Le document de recherche présente un examen approprié des documents de référence, ainsi que des défis et des pratiques exemplaires en matière d'élaboration de MQH.
 - L'application du cadre, mis en œuvre au moyen de scripts R, est une méthode scientifiquement défendable, transparente et reproductible pour la production de MRE. Le cadre peut être utilisé pour produire des modèles pour d'autres espèces ou adapté à l'étendue d'autres exercices de modélisation.
 - Une approche à plusieurs modèles facilite les estimations de l'incertitude en permettant de comparer les différents types de modèles.
 - Le cadre a été appliqué à douze espèces en utilisant des données de qualité et quantité variables sur la présence et l'absence comme études de cas. Les modèles correspondent généralement bien aux données, comme le montrent les paramètres de rendement. L'application démontre l'importance de préciser les composantes clés, notamment l'exploration des données, l'interprétation du modèle, etc.
 - Il incombera à l'utilisateur d'évaluer la pertinence des prévisions du modèle découlant de l'application du cadre.

RÉVISIONS DU DOCUMENT DE TRAVAIL

Cadre

- Mentionner d'autres modèles (p. ex. Maxent).
- Mentionner que la collecte de nouvelles données indépendantes pour mettre le modèle à l'essai est un modèle exemplaire et l'ajouter aux recommandations.
- Définition claire d'un MQH, d'un MRE et d'une niche réalisée par rapport à une niche fondamentale dans le document et le glossaire.
- Préciser que les prévisions de l'IQH par rapport au MLG et au modèle ARS sont différentes et que si l'IQH est inclus dans l'ensemble, on ne sait pas bien sur quoi portent les prédictions.
- Discuter des modèles MLG et MAG et justifier pourquoi ils n'ont pas été inclus.
- Mentionner que l'inclusion d'un plus grand nombre de méthodes dans un ensemble pourrait être une approche valable.
- Indiquer d'autres façons d'inclure la structure spatiale (p. ex. champs aléatoires gaussiens).
- Inclure une discussion sur la façon dont les connaissances traditionnelles écologiques et les connaissances écologiques locales pourraient être incluses dans l'analyse du cadre.
- Préciser que l'ACS a été utilisée seulement pour voir si elle avait une incidence et qu'elle n'a pas été incluse dans les prévisions finales.
- Indiquer que les points de données des paramètres sont supprimés de façon aléatoire dans le réglage du modèle ARS et que les ensembles de données de réglage et d'essai sont spatiaux.
- Préciser que l'incertitude de l'application est principalement une incertitude de validation croisée, pas vraiment une incertitude des paramètres.
- Clarifier les différences entre l'incertitude du modèle et l'incertitude des facteurs.

-
- Chercher une autre approche pour analyser les résiduels au moyen de la classification par fenêtres ou de résiduels quantiles randomisés, et fournir une explication plus complète des méthodes de création et d'interprétation des résiduels.
 - Fournir des lignes directrices sur les cas où il est approprié d'inclure l'IQH dans les modèles d'ensemble.
 - Indiquer que les modèles des espèces ne sont pas des modèles définitifs et qu'il peut être nécessaire de les perfectionner pour pouvoir les utiliser dans des applications de gestion.
 - Inclure un paragraphe sur les « signaux d'alarme » du cadre dont les utilisateurs finaux devraient tenir compte avant/pendant l'élaboration d'un modèle.
 - Généraliser l'organigramme du cadre.
 - Ajouter une référence pour le seuil de la valeur AUC pour les modèles de validation croisée à inclure dans le modèle d'ensemble.
 - Préciser pourquoi et comment les observations ponctuelles ont été regroupées, et les répercussions de ces décisions. Expliquer que d'autres options sont possibles pour agréger les données.
 - Fournir des lignes directrices sur la portée du domaine du modèle en ce qui concerne la couverture spatiale de vos données et les différentes conditions le long de la côte. Inclure une discussion sur la stationnarité.

Application

- Il est à noter que toutes les espèces n'ont pas été échantillonnées dans l'ensemble de leur habitat.
- Justifier les incohérences entre les recommandations pour la sélection des données pour le crabe dormeur et le sébaste à dos épineux.
- Recommander des travaux futurs (p. ex. validation, essais de simulation).
- Inclure une figure à six graphiques de tous les modèles, ajouter des points de données aux cartes, ajouter une figure de l'écart-type de l'ensemble, inclure des cartes des zones littorales en médaillon.
- Inclure une figure illustrant toutes les relations entre les paramètres.
- Inclure dans la section Prochaines étapes une discussion de l'inclusion d'autres variables prédictives (p. ex. couche mixte, aragonite, effort de pêche, autres espèces (comme les coraux et les éponges), température de pointe, baisses de salinité, événements létaux).
- Indiquer que la corrélation de Spearman s'ajoute au facteur d'inflation de la variance.
- Préciser pourquoi les autres saisons n'ont pas été incluses dans les prédicteurs environnementaux.
- Mentionner la modélisation d'archétypes comme une possibilité future.
- Indiquer si tous les termes d'interaction des prédicteurs n'ont pas été inclus dans les modèles IQH pour des raisons écologiques ou de calcul.
- Formuler des commentaires sur ce qui constituerait le meilleur modèle.
- Recommander de créer plusieurs modèles IQH fondés sur divers experts des espèces.

-
- Ajouter un texte pour recommander de modéliser individuellement les ensembles de données lorsqu'ils sont différents et que l'utilisateur final décide s'il est approprié de combiner tous les ensembles de données ou d'en utiliser un seul. Les utilisateurs finaux devraient prendre cette décision en fonction de leurs objectifs finaux, des limitations de leurs données et de leur modèle et de leur biais. Le modèle d'espèce pour le crabe dormeur servira d'exemple.

RECOMMANDATIONS

- Pour bien appliquer le cadre en vue d'aborder les questions de gestion, il faut faire correspondre les objectifs de gestion aux intrants et extrants du modèle.
- Il incombera à l'utilisateur d'évaluer la pertinence des prévisions de l'étude de cas découlant de l'application du cadre.
- Pour appliquer le cadre :
 - Définir clairement les objectifs du modèle.
 - Examiner la cohérence spatiale et temporelle de la résolution entre le prédicteur et les données sur l'espèce.
 - Diagnostiquer les erreurs possibles en recherchant des relations inattendues dans les effets marginaux et l'influence relative des prédicteurs.
 - Tenir compte de la façon dont la qualité des données sur l'occurrence de l'espèce influe sur le rendement du modèle.
 - Envisager de réduire la taille de l'échantillon si une partie des données sur l'occurrence réduit la qualité globale des données.
 - Examiner les graphiques de validation et les résiduels (y compris les schémas spatiaux dans les résiduels) en plus des statistiques individuelles pour évaluer le rendement du modèle.
 - Utiliser les statistiques AUC et TSS (statistiques sur la qualité vraie) sur le kappa et la précision pour évaluer le rendement du modèle.
 - Utiliser le CV des blocs spatiaux pour construire des modèles et évaluer l'ajustement et la transférabilité des modèles.
 - Envisager d'effectuer la modélisation avec un prédicteur spatial pour faciliter l'interprétation du modèle.
 - Fournir une mesure spatialement explicite de l'incertitude en examinant la variation entre plusieurs modèles.
- Utiliser des données recueillies de façon indépendante pour l'évaluation du modèle, dans la mesure du possible.

REMERCIEMENTS

Nous apprécions le temps que tous les participants ont consacré au processus d'examen régional par les pairs. Nous remercions en particulier les examinateurs, Chris Rooper, Javier Murillo-Perez et Anders Knudby, pour leur temps et leur expertise.

ANNEXE A : CADRE DE RÉFÉRENCE

Modélisation de la qualité de l'habitat – Pratiques exemplaires pour l'océan Pacifique du Canada

Réunion d'examen par les pairs régionale : Région du Pacifique

Du 11 au 12 juin 2019

Nanaimo (Colombie-Britannique)

Président : Steven Schut

Contexte

Les modèles de la qualité de l'habitat (MQH) sont un outil pouvant servir à prédire la répartition de l'habitat d'une espèce, en établissant un lien entre les observations de l'occurrence d'une espèce et les données environnementales. Cette catégorie de modèles étant diversifiée, les MQH servent à diverses applications, comme la prévision de l'habitat des espèces vulnérables (Rengstorf *et al.* 2013, Anderson *et al.* 2016, Rowden *et al.* 2017), la détermination des zones importantes du cycle biologique des poissons (Le Pape *et al.* 2014, Rooper *et al.* 2019), la détermination des zones candidates en vue d'une protection (Embling *et al.* 2009), et la prévision des changements pour l'habitat des espèces aquatiques dus aux changements climatiques (Cheung *et al.* 2010). Les MQH peuvent également aider à optimiser l'utilité des efforts de recherche et du financement à venir, en éclairant la planification des levés et en orientant la recherche sur les lacunes existantes en matière de connaissances.

Les MQH peuvent, par conséquent, aider à répondre aux besoins en matière de conservation et de gestion, en comblant certains de ces renseignements manquants et en améliorant notre compréhension de la répartition des espèces pour de nombreuses initiatives de planification spatiale marine, comme la configuration du réseau d'aires marines protégées (Abecasis *et al.* 2014), l'intervention en cas de déversement d'hydrocarbures, et la désignation de zones d'importance écologique et biologique (Beazley *et al.* 2016). Cependant, les MQH peuvent aussi être mal utilisés, si les données sur les espèces ou l'environnement ne sont pas correctement examinées et préparées, et si la validation des modèles n'est pas effectuée de façon appropriée (Roberts *et al.* 2017, Hawkins *et al.* 2009, Elith and Leathwick 2009).

Cet examen par les pairs vise à produire un cadre exhaustif de pratiques exemplaires en fonction de 12 espèces présentes dans la région du Pacifique, afin d'orienter l'élaboration et l'utilisation futures des MQH. Même si le cadre sera élaboré à l'aide d'espèces du Pacifique, il est possible de l'utiliser pour d'autres espèces et régions. Des discussions et des conseils seront fournis sur l'utilisation appropriée, le prétraitement et la préparation des données, les méthodes de modélisation, l'élaboration et la validation des modèles, l'interprétation des résultats, ainsi que la façon de communiquer les incertitudes sous-jacentes.

La Direction des sciences de la Région du Pacifique de Pêches et Océans Canada propose d'élaborer une approche normalisée pour la création de MQH en fonction des pratiques exemplaires, afin de s'assurer qu'ils soient tous utilisés avec la même qualité et la même rigueur. L'évaluation et les conseils découlant de cet examen régional par les pairs du Secrétariat canadien de consultation scientifique serviront à élaborer des MQH et à les intégrer aux décisions scientifiques et stratégiques relatives à la gestion et à la conservation des espèces marines.

Objectifs

Le document de travail suivant sera examiné et servira de fondement aux discussions et aux avis sur les différents objectifs énumérés ci-après.

Nephin, J., E.J. Gregr, C. St. Germain, C. Fields, J.L. Finney. Modélisation de la qualité de l'habitat – Pratiques exemplaires pour l'océan Pacifique du Canada. Document de travail du CASP 2018SCI01

Les objectifs précis de cet examen sont les suivants :

1. Fournir le contexte, les méthodes, les défis et les pratiques exemplaires actuelles pour l'élaboration des MQH.
2. Évaluer les défis associés à la recherche de données appropriées sur les espèces et l'environnement, et fournir des conseils sur les pratiques exemplaires en vue de préparer les données pour les MQH.
3. Élaborer un cadre qui applique les pratiques exemplaires à l'élaboration des MQH, y compris la façon de valider les modèles et de présenter les hypothèses et l'incertitude relatives aux modèles.
4. Démontrer la façon d'utiliser le cadre des MQH à l'aide d'une étude de cas dans la région du Pacifique.
5. Examiner les constatations pertinentes à l'élaboration des MQH et formuler des recommandations pour les utilisations futures.

Publications prévues

- Avis scientifique
- Document de recherche
- Compte rendu

Participation prévue

- Pêches et Océans Canada (Secteur des sciences, des écosystèmes et des océans, Gestion des pêches, Direction de la gestion des écosystèmes)
- Milieu universitaire
- Groupes autochtones
- Organisations non gouvernementales

Références

- Abecasis, D., Afonso, P. and Erzini, K., 2014. Combining multispecies home range and distribution models aids assessment of MPA effectiveness. *Marine Ecology Progress Series*, 513, pp.155-169.
- Anderson, O. F., Guinotte, J. M., Rowden, A. A., Tracey, D. M., Mackay, K. A., & Clark, M. R., 2016. Habitat suitability models for predicting the occurrence of vulnerable marine ecosystems in the seas around New Zealand. *Deep-Sea Research Part I: Oceanographic Research Papers*, 115, 265–292. <https://doi.org/10.1016/j.dsr.2016.07.006>
- Beazley, L., Kenchington, E., Murillo, F.J., Lirette, C., Guijarro, J., McMillan, A. and Knudby, A., 2016. Species distribution modelling of corals and sponges in the maritimes region for use in the identification of Significant Benthic Areas. *Canadian Technical Report of Fisheries and Aquatic Sciences* 3171.
- Cheung, W.W., Lam, V.W., Sarmiento, J.L., Kearney, K., Watson, R.E.G., Zeller, D. and Pauly, D., 2010. Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Global Change Biology*, 16(1), pp.24-35.

-
- Elith, J., & Leathwick, J., 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Embling, C. B., Gillibrand, P. A., Gordon, J., Shrimpton, J., Stevick, P. T., & Hammond, P. S., 2009. Using habitat models to identify suitable sites for marine protected areas for harbour porpoises (*Phocoena phocoena*). *Biological Conservation*, 143(2), 267–279. <https://doi.org/10.1016/j.biocon.2009.09.005>
- Hawkins, D. M., Basak, S. C., & Mills, D., 2003. Assessing model fit by cross-validation. *Journal of Chemical Information and Computer Sciences*, 43(2), 579–586. <https://doi.org/10.1021/ci025626i>
- Le Pape, O., Delavenne, J., & Vaz, S., 2014. Quantitative mapping of fish habitat: A useful tool to design spatialised management measures and marine protected area with fishery objectives. *Ocean and Coastal Management*, 87, 8–19. <https://doi.org/10.1016/j.ocecoaman.2013.10.018>
- Rengstorf, A. M., Yesson, C., Brown, C., & Grehan, A. J., 2013. High-resolution habitat suitability modelling can improve conservation of vulnerable marine ecosystems in the deep sea. *Journal of Biogeography*, 40(9), 1702–1714. <https://doi.org/10.1111/jbi.12123>
- Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, Hauenstein S, Lahoz-Monfort JJ, Schröder B, Thuiller W. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* 40:913-929.
- Rowden, A. A., Anderson, O. F., Georgian, S. E., Bowden, D. A., Clark, M. R., Pallentin, A., & Miller, A., 2017. High-Resolution Habitat Suitability Models for the Conservation and Management of Vulnerable Marine Ecosystems on the Louisville Seamount Chain, South Pacific Ocean. *Frontiers in Marine Science*, 4(October). <https://doi.org/10.3389/fmars.2017.00335>
- Rooper, C.N., Hoff, G.R., Stevenson, D.E., Orr, J.W. and Spies, I.B., 2019. Skate egg nursery habitat in the eastern Bering Sea: a predictive model. *Marine Ecology Progress Series*, 609, pp.163-178.

ANNEXE B ORDRE DU JOUR

Modélisation de la qualité de l'habitat – Pratiques exemplaires pour l'océan Pacifique du Canada

Les 11 et 12 juin 2019

Nanaimo (Colombie-Britannique)

Président : Steve Schut

JOUR 1 – Mardi 11 juin

Heure	Sujet	Présentateur
9 h	Présentations Examen de l'ordre du jour et gestion interne Aperçu et procédure du SCCS	Président
9 h 15	Examen du cadre de référence	Président
9 h 30	Présentation du document de travail – Cadre des MQH	Auteurs
10 h 15	Pause	
10 h 45	Aperçu des examens écrits – Cadre des MQH	Président + Examineurs et auteurs
11 h 30	Définition des enjeux clés aux fins de discussion en groupe - Cadre des MQH	Groupe
12 h	Pause repas	
13 h	Présentation du document de travail – Application du cadre	Auteurs
13 h 30	Aperçu des examens écrits – Application du cadre	Président + Examineurs et auteurs
14 h 30	Pause	
14 h 45	Définition des enjeux clés aux fins de discussion en groupe - Application du cadre	Participants à l'examen régional par les pairs
15 h 15	Discussion et résolution des résultats et conclusions	Participants à l'examen régional par les pairs
17 h	Levée de la réunion pour la journée	

JOUR 2 - Mercredi 12 juin

Heure	Sujet	Présentateur
9 h	Présentations Examen de l'ordre du jour et gestion interne Récapitulation de la première journée (<i>au besoin</i>)	Président
9 h 15	Reprise des questions en suspens du jour 1	Participants à l'examen régional par les pairs
10 h 30	Pause	
10 h 45	<i>Avis scientifique (AS)</i> Établir un consensus sur les éléments suivants en vue de leur inclusion : <ul style="list-style-type: none">• Points saillants• Sources d'incertitude• Résultats et conclusions• Figures et tableaux• Avis supplémentaire pour la direction (<i>au besoin</i>)	Participants à l'examen régional par les pairs
12 h	Pause repas	
13 h	<i>Avis scientifique (suite)</i>	Participants à l'examen régional par les pairs
14 h 45	Pause	
15 h	Prochaines étapes – Examen par le président <ul style="list-style-type: none">• Processus d'examen et d'approbation de l'avis scientifique et échéanciers• Échéanciers relatifs au document de recherche et au compte rendu• Autres mesures de suivi ou engagements (<i>au besoin</i>)	Président
15 h 45	Autres questions découlant de l'examen	Président et participants
16 h	Levée de la réunion	

ANNEXE C : LISTE DES PARTICIPANTS

Nom	Prénom	Organisme d'appartenance
Anderson	Sean	Secteur des sciences du MPO, Évaluation des stocks
Beazley	Lindsay	Secteur des sciences du MPO, Région des Maritimes
Benoy	Nicholas	Secteur de la gestion des écosystèmes du MPO, Océans
Campbell	Jill	Secteur des sciences du MPO
Candy	John	Secteur des sciences du MPO, Centre des avis scientifiques
Chiang	Eric	Secteur des océans du MPO
Christensen	Lisa	Secteur des sciences du MPO, Centre des avis scientifiques
Curtis	Janelle	Secteur des sciences du MPO, Sciences des écosystèmes
Davies	Sarah	Secteur des sciences du MPO, Sciences des écosystèmes
Dudas	Sarah	Secteur des sciences du MPO, Sciences des écosystèmes
English	Philina	Secteur des sciences du MPO, Évaluation des stocks
Ferguson	Kiyomi	Secteur des sciences du MPO
Fields	Cole	Secteur des sciences du MPO, Sciences des écosystèmes
Finney	Jessica	Secteur des sciences du MPO, Sciences des écosystèmes
Gale	Katie	Secteur des sciences du MPO, Sciences des écosystèmes
Gomez	Catalina	Secteur des sciences du MPO, Terre-Neuve
Goulet	Pierre	Secteur des sciences du MPO, Région de Terre-Neuve
Gregr	Ed	Université de la Colombie-Britannique
Gullage	Lauren	Secteur des sciences du MPO, Région de Terre-Neuve
Herborg	Matthias	Secteur des sciences du MPO, Sciences des écosystèmes
Hubleby	Brad	Secteur des sciences du MPO, Région des Maritimes
Kenchington	Ellen	Secteur des sciences du MPO, Région des Maritimes
Knudby	Anders	Université d'Ottawa
Lessard	Joanne	Secteur des sciences du MPO, Sciences des écosystèmes
Leus	Dan	Secteur des sciences du MPO, Sciences des écosystèmes
Murillo-Perez	Javier	Secteur des sciences du MPO, Région des Maritimes
Nephin	Jessica	Secteur des sciences du MPO, Sciences des écosystèmes
Novaczek	Emilie	Secteur des sciences du MPO, Région de Terre-Neuve
O	Miriam	Secteur des sciences du MPO, Sciences des écosystèmes
Pretty	Christina	Secteur des sciences du MPO, Terre-Neuve
Robinson	Cliff	Secteur des sciences du MPO, Sciences des écosystèmes
Rooper	Chris	Secteur des sciences du MPO, Sciences des écosystèmes
Rubidge	Emily	Secteur des sciences du MPO, Sciences des écosystèmes
Sameoto	Jessica	Secteur des sciences du MPO, Région des Maritimes
Schut	Steve	Secteur des sciences du MPO, Section des données sur les pêches et l'évaluation
St. Germain	Candice	Secteur des sciences du MPO, Division des sciences océaniques
Warren	Margaret	Secteur des sciences du MPO, Région de Terre-Neuve
Wells	Nadine	Secteur des sciences du MPO, Terre-Neuve
Yakgujaanas	Jaasaljuus	Conseil de la Nation haïda

ANNEXE D : EXAMENS DU DOCUMENT DE TRAVAIL

EXAMINATEUR : CHRIS ROOPER, SECTEUR DES SCIENCES DU MPO, SCIENCES DES ÉCOSYSTÈMES

Je vous remercie de me donner l'occasion de participer à cet examen. J'admire et j'apprécie vraiment le travail réalisé pour le document du SCCS. Les auteurs méritent des félicitations pour les efforts incroyables qu'ils ont déployés pour élaborer ce cadre.

Vous trouverez ci-après des commentaires qui portent sur les questions précises de cet examen, suivis de certains commentaires précis qui devraient être traités ou justifiés. J'ai également inclus des commentaires plus précis et quelques suggestions grammaticales mineures dans le texte à l'aide du suivi des modifications.

1. *Fournir le contexte, les méthodes, les défis et les pratiques exemplaires actuelles pour l'élaboration des MQH.*

Le domaine de la modélisation de la répartition de l'habitat et des espèces est vaste et a fait l'objet d'un bon nombre d'examens et de discussions détaillées sur les méthodologies, les défis et les pratiques exemplaires. Beaucoup de ces exercices, comme les statistiques d'essai à utiliser pour évaluer les modèles, sont en cours depuis de nombreuses années et ne sont toujours pas résolus. Les auteurs ont fait un excellent travail de mise en contexte sur ces sujets et ont fait référence aux documents pertinents pour ceux qui s'intéressent aux détails de ces types de discussions. En général, les auteurs ont bien cerné les points sur lesquels ils ont fait des choix et les ont justifiés en présentant des documents ou des exemples pertinents. Le traitement des types de modélisation autres que ceux pris en compte dans ce cadre était relativement léger, mais compréhensible compte tenu de l'espace nécessaire pour rendre justice à ce sujet et du fait qu'il sortait un peu du sujet de ce document.

2. *Évaluer les défis associés à la recherche de données appropriées sur les espèces et l'environnement, et fournir des conseils sur les pratiques exemplaires en vue de préparer les données pour les MQH.*

Les auteurs donnent un bon aperçu des défis liés aux données et des exemples utilisant les espèces examinées dans ce cadre. J'ai souligné quelques questions pour obtenir plus d'explications dans les commentaires. Dans l'ensemble, je pense que les auteurs tracent une voie qui peut être suivie pour la compilation de données, surtout en ce qui concerne la côte de la Colombie-Britannique.

3. *Élaborer un cadre qui applique les pratiques exemplaires à l'élaboration des MQH, y compris la façon de valider les modèles et de présenter les hypothèses et l'incertitude relatives aux modèles.*

J'ai vraiment aimé les organigrammes qui illustrent l'élaboration des MQH. Ils ont été très utiles et ont fourni un cadre facile à suivre. Il était facile de voir les endroits où de nouvelles données ou méthodes pourraient être insérées dans l'organigramme. J'aime aussi beaucoup les figures utilisées pour montrer leurs analyses de l'évaluation des modèles à l'aide de données indépendantes et dépendantes des seuils. Elles étaient très utiles et très claires. La méthode du CV pour la validation du modèle a été très bien présentée, y compris l'explication de la méthode du CV des blocs. Lorsque des choix ont été faits (comme l'utilisation de l'écart type au lieu du coefficient de variation pour montrer l'incertitude), ces décisions ont été bien expliquées et bien réfléchies. Il existe certainement d'autres façons d'aborder certains des exercices de validation, mais souligner la nécessité d'avoir des données indépendantes pour valider les modèles est conforme aux pratiques exemplaires.

4. Un commentaire est que même si la méthode du CV des blocs est utile pour réduire l'indépendance, elle ne l'élimine pas. Par exemple, le modèle de l'ormeau nordique était fondé en grande partie sur des sites témoins qui avaient été choisis au départ comme zones d'habitat convenable. Il y a donc une certaine dépendance ou un certain biais dans le plan du relevé. La meilleure façon de valider les modèles est de mettre leurs prédictions à l'essai de façon indépendante en recueillant des données supplémentaires. Cela se fait rarement, mais c'est le modèle exemplaire. Dans le cas de l'ormeau en particulier, les auteurs disposaient des données provenant d'un certain nombre de types de relevés et il aurait été relativement facile d'utiliser les données d'un relevé pour paramétrer le modèle, puis de les valider avec un autre relevé. L'indépendance aurait été encore plus réduite dans ce processus.

5. *Démontrer la façon d'utiliser le cadre des MQH à l'aide d'une étude de cas dans la région du Pacifique.*

L'application du cadre aux espèces de la région du Pacifique a été très bien faite. J'ai quelques commentaires à faire sur un autre exposé et d'autres parties du document Word.

6. *Examiner les constatations pertinentes à l'élaboration des MQH et formuler des recommandations pour les utilisations futures.*

La section Conclusions a besoin d'être un peu étoffée (voir le commentaire ci-après), mais elle était généralement bonne. Les recommandations suivaient très bien les données et les analyses présentées. Il y avait quelques incohérences entre les recommandations et les choix qui ont été faits dans les analyses (en pensant en particulier aux compromis des données sur le crabe dormeur, où quelques sources différentes plutôt difficiles à comparer ont été utilisées, et au sébaste, pour lequel deux sources de données raisonnablement bonnes ont été présentées, mais une seule a été utilisée). Il serait utile de resserrer cette norme avec des explications dans la section Crabe dormeur.

Il n'y avait pas de recommandation explicite sur les types de modèles à inclure dans les ensembles. Les auteurs ont utilisé n'importe quel modèle dont la valeur AUC était supérieure à 0,5 (n'est-ce pas?). Est-ce ce qu'ils recommandent ici dans le cadre?

Ce qui manquait, c'était une recommandation sur la validation indépendante des modèles par la collecte de nouvelles données. J'aimerais qu'elle soit incluse, même si elle n'était pas nécessairement appliquée ici.

Principaux commentaires

Section 1.

Dans la section 1.1, l'aperçu de la modélisation de la qualité de l'habitat prête à confusion lorsqu'il est lu dans le contexte de ce cadre. Un MQH est défini comme la détermination de la niche fondamentale de l'espèce en l'absence d'obstacles et d'interactions biotiques, par opposition aux modèles de répartition des espèces qui intègrent ces processus. Il semble que cette discussion s'applique surtout aux méthodes de l'IQH, mais pas aux méthodes du MLG et du modèle ARS. Toutefois, les données et les relations utilisées pour générer tous les modèles sont fondées soit sur des relations établies principalement à partir d'observations sur le terrain (c.- à- d. opinions d'experts), soit sur les données elles-mêmes. Ces deux types de modèles sont au moins en partie des modèles de la niche réalisée. Il serait probablement utile de définir exactement ce qu'est un MQH dans le contexte de ce cadre, puis d'expliquer comment il définit une niche fondamentale. Sinon (probablement ma solution préférée), il serait possible de discuter de la théorie des niches, de définir ce que signifie MQH pour ce cadre et de supprimer la discussion sur la façon dont le MQH, tel que vous l'avez défini, s'inscrit dans la théorie des

niches. La partie importante que vous essayez de communiquer se trouve dans les deux paragraphes suivants.

Si vous choisissez de combiner les types d'engins de pêche et de relevés et que l'étendue spatiale ou la capturabilité diffère d'un type d'engin à l'autre, il faut le prendre en compte (ou le mettre à l'essai) dans le modèle. Un biais spatial peut se produire dans le premier cas et un biais non aléatoire se produira dans l'autre. Il faut en tenir compte dans la modélisation ou en mettant à l'échelle les prédictions de la prévalence. Ou, comme vous l'avez fait pour le sébaste, rejeter un ensemble de données.

Le MQH et la répartition des espèces sont utilisés de façon interchangeable dans le premier paragraphe de l'introduction, ce qui pourrait être une partie du problème. Il faut une bonne définition claire de ce dont vous parlez lorsque vous faites référence à un MQH. En me fondant sur une partie de la discussion sur le cadre dans les méthodes de mise à l'échelle des prédicteurs temporels, etc., je pourrais être tenté de dire explicitement et d'emblée que votre intérêt pour ce cadre est de produire des cartes statiques de l'habitat.

À ce sujet, pour les modèles MLG/ARS, le manuscrit utilise un MQH d'une manière très précise, la probabilité qu'un organisme ait été trouvé à un endroit au moins une fois. Puisque vos présences peuvent être des combinaisons de 0 et de 1 et que vos absences n'ont jamais été trouvées ici, c'est un peu différent de la modélisation typique présence/absence ou même de l'IQH comme vous l'utilisez. Il faudrait l'indiquer quelque part, et aussi discuter des conséquences de cette définition. Les données d'observation provenant de l'IQH sont également différentes (il y a beaucoup de mélange de densité dans les définitions de ces fonctions). Je veux dire que puisque les modèles prédisent différentes choses, cela nous amène à nous demander s'ils devraient être regroupés ensemble.

Section 1. J'envisagerais de réorganiser et de repenser cette introduction. Je pense que les principales recommandations que vous formulez ici sont les suivantes : 1) examiner le contexte de gestion du modèle, les questions auxquelles il faut répondre; 2) examiner l'objectif des modèles, le produit dont vous avez besoin pour répondre aux questions; 3) examiner les caractéristiques de l'espèce et ses relations écologiques; 4) examiner les types de données dont vous avez besoin pour traiter ces relations; et 5) examiner le type de modèle adapté à la qualité/quantité des données dans le contexte du produit recherché. Ces concepts sont un peu mélangés dans la structure actuelle et, bien que certains points importants soient soulevés (comme le lien entre l'écologie et la gestion), il serait préférable, à mon avis, d'avoir une séquence qui mène le lecteur aux prochaines sections où vous discutez plus en détail afin que le cadre soit organisé selon le même flux.

Section 3.1.

Pourquoi avez-vous choisi d'utiliser n'importe quelle présence comme indication de présence, plutôt que tout l'ensemble des données? D'un point de vue statistique, je ne vois pas immédiatement le raisonnement, sauf peut-être que vos densités ponctuelles étaient trop élevées (biais spatial) pour certains relevés? Dans ce cas, il y a d'autres méthodes d'éclaircissement qui peuvent être utilisées. De plus, combien de fois cela s'est-il produit (quel était le n pour l'ensemble de données original et n pour les ensembles de données réduits)?

Section 4.4.

Cette section des conclusions porte en grande partie sur les méthodes de l'IQH et les raisons pour lesquelles elles devraient être conservées. Il devrait y avoir d'autres conclusions importantes qui sont ajoutées ici.

Section 5.

La validation par des données indépendantes n'est pas abordée.

Avez-vous mis à l'essai le rendement du modèle d'ensemble sans inclure les modèles IQH? L'une des conclusions était que l'augmentation de la valeur AUC à partir du modèle d'ensemble était en partie le résultat de l'inclusion du modèle IQH, qui a introduit un certain réalisme biologique dans les modèles statistiques. C'est peut-être vrai, mais ce n'était pas indiqué dans les comparaisons. À l'appui de ce point, il serait utile de calculer l'ensemble avec seulement le MLG et le modèle ARS et de le comparer à la version à trois modèles.

À cet égard, en l'absence de bonnes données, l'IQH est très utile, mais lorsqu'il y a des données raisonnables, il semble préférable de laisser les données guider les relations et de les examiner dans le contexte de l'IQH/de la documentation/de l'opinion d'experts plutôt que de recommander d'inclure un modèle dont le rendement est médiocre pour l'ensemble. C'est peut-être une opinion impopulaire, mais si vous recommandez de toujours utiliser l'IQH dans ce cadre, vous devriez démontrer que le rendement des ensembles est aussi bon ou meilleur lorsqu'on inclut ce modèle que lorsqu'il est éliminé. L'utilisation de l'IQH dans des situations où l'on manque de données (comme pour les moules) semble plus appropriée que pour les espèces bien documentées (comme les sébastes).

Annexe B

À l'annexe B, il serait bon d'avoir des figures en six graphiques des cartes avec un graphique par modèle, un graphique des données sur une carte, les cartes de l'ensemble et les cartes de l'écart-type de l'ensemble. À l'heure actuelle, la seule sortie du modèle est tirée de l'ensemble et de l'IQH, et nous ne voyons jamais les données réelles. Les pratiques exemplaires exigent probablement de montrer ces choses. Les cartes avec les facteurs limitatifs pour l'IQH peuvent probablement être supprimées.

Pour ce qui est de l'espace économisé, j'aimerais voir les formes des divers modèles tracées l'une à côté de l'autre. L'un des points du document est que des modèles robustes devraient présenter les mêmes formes par rapport aux types de modèles, mais il serait bon de le voir en annexe pour que le lecteur puisse juger les ajustements du modèle en fonction de leurs mérites et de leur réalisme écologique.

EXAMINATEUR : JAVIER MURILLO-PEREZ, SECTEUR DES SCIENCES DU MPO, RÉGION DES MARITIMES

Un document de travail a été préparé pour examen dans le cadre de la réunion du SCCS sur la Modélisation de la qualité de l'habitat – Pratiques exemplaires pour l'océan Pacifique du Canada. Ce document présente un cadre pour élaborer et appliquer des modèles de qualité de l'habitat sur la côte canadienne du Pacifique. La préparation des données, l'élaboration de modèles, l'évaluation des modèles et l'interprétation des résultats et des incertitudes sont abordées afin de normaliser et de faciliter le processus de création de modèles. Ce cadre a été mis en œuvre dans le langage de programmation statistique R.

Il convient de féliciter les auteurs pour la qualité et la quantité des analyses présentées dans le document. L'approche comprend la plupart des recommandations récentes sur la modélisation de la qualité de l'habitat et chaque étape du processus de modélisation, de la préparation des données à l'interprétation des résultats et des incertitudes, est examinée. Les données et les méthodes sont bien expliquées et appuient les principales conclusions.

La plupart des commentaires et des questions que je présente ici sont mineurs et reflètent des orientations potentielles ou des modèles de rechange qui pourraient être utilisés dans les révisions futures de la méthodologie.

Points importants à prendre en considération :

1. SÉLECTION DU TYPE DE MODÈLE.

Les auteurs ont choisi trois types de modèles, d'un degré de complexité différent, qui sont utilisés dans la prédiction finale de l'ensemble : les modèles d'indice de qualité de l'habitat (IQH) dérivés d'opinions d'experts, le modèle linéaire généralisé (MLG) et les arbres de régression stimulée (ARS). L'utilisation de l'IQH est discutée et justifiée dans les conclusions, et le modèle ARS est couramment utilisé dans les applications récentes. Cependant, bien que l'on dise que le cadre convient également aux modèles additifs généralisés (MAG), il n'y a pas de discussion expliquant pourquoi on préfère les MLG aux MAG. L'utilisation de MAG, qui peuvent inclure une relation non linéaire entre la réponse et les prédicteurs, pourrait régler certains des problèmes relevés dans le MLG, comme la relation de la température sur le crabe dormeur.

2. DONNÉES PRÉDICTIVES (page 6).

La profondeur maximale de la couche mixte est un autre prédicteur estimé par les modèles de circulation océanique qui peut avoir une influence importante sur la production primaire dans les eaux de surface (Polovina *et al.* 1995, Carstensen *et al.* 2002). En outre, l'effort de pêche peut également être un prédicteur important pour comprendre une partie de la répartition actuelle sur les écosystèmes touchés (p. ex. Foster *et al.* 2015, Tien *et al.* 2017). Et dans certains cas, les variables biotiques peuvent aussi renseigner sur la répartition d'autres organismes lorsqu'il existe une forte relation entre eux. On pourrait aussi discuter de ces prédicteurs.

3. PRÉDICTEURS SPATIAUX (page 8).

Ces dernières années, l'approche des approximations de Laplace imbriquées et intégrées (INLA) (Rue *et al.* 2009) a permis de réaliser des progrès importants en modélisation spatiale. L'utilisation d'une approche INLA permet d'inclure la structure spatiale dans les estimations et les prévisions des paramètres (p. ex. Boudreau *et al.* 2017, Penino *et al.* 2019). Pour un examen de la question, voir Bakka *et al.* (2018).

4. PRÉPARATION DES DONNÉES (page 12).

Pour les observations représentées comme des lignes, on dit que les valeurs moyennes sont calculées à partir des cellules matricielles qui croisent les segments de la ligne. Alors que pour les variables prédictives catégoriques, on utilise la catégorie la plus dominante se produisant dans les cellules matricielles qui s'entrecroisent. Toutefois, dans certains cas, il peut être plus utile de calculer le pourcentage de chevauchement entre les catégories et les lignes. Par exemple, si un transect recouvre principalement des sédiments sablonneux, mais traverse une parcelle rocheuse convenant à un organisme donné, l'utilisation de l'information dominante ne permettrait pas de relever cette caractéristique, alors que l'utilisation d'un pourcentage de chevauchement (p. ex. 80 % de sable, 20 % de roche) ajouterait certains renseignements qui pourraient être utiles pour le modèle.

L'utilisation du coefficient de corrélation du rang de Spearman (ou Pearson) en plus du facteur d'inflation de la variance (FIV) pourrait ajouter de l'information à la corrélation entre les prédicteurs. De plus, on pourrait ajouter un second seuil basé sur le coefficient de Spearman (> 0,8 ou 0,9) au seuil du FIV pour faciliter la décision de supprimer les prédicteurs. Cela pourrait aussi ajouter des renseignements sur le degré de corrélation entre la température et la salinité (page 35).

5. VALIDATION CROISÉE (page 12).

La validation du modèle utilise habituellement le même type de données que celui utilisé pour construire le modèle. Toutefois, lorsque l'on utilise de nouvelles données obtenues à partir d'un engin différent pour valider les modèles, les résultats peuvent être très différents (p. ex. Rooper *et al.* 2019). Cela ajoute de l'incertitude quant aux prévisions, qui pourrait être discutée dans le rapport et on pourrait peut-être recommander d'utiliser différents types de données (lorsqu'ils sont disponibles) pour valider les modèles.

6. ÉVALUATION DU MODÈLE (page 15).

D'autres valeurs AUC et TSS pourraient être ajoutées pour montrer l'exactitude de la prédiction du modèle entre 0,5 et 1. Selon Pearce et Ferrier (2000) et Jones *et al.* (2010), les valeurs AUC > 0,9 sont considérées comme bonnes, comme modérément bonnes entre 0,7 et 0,9 et comme médiocres < 0,7, alors que les valeurs TSS > 0,6 sont considérées comme bonnes, comme passables à modérément bonnes entre 0,2 et 0,6 et comme médiocres < 0,2 (Landis et Koch 1977, Jones *et al.* 2010).

7. PRÉDICTIONS ET INCERTITUDES (figure 3.7, page 18).

La Région des Maritimes a utilisé une autre option de délimitation des zones d'extrapolation qui consiste à ajouter une couche indiquant ces zones, mais avec une certaine transparence pour toujours voir les valeurs prévues en dessous. Cela permet de voir les zones extrapolées dans les surfaces de la prédiction, mais ajoute un degré élevé d'incertitude (voir un exemple à la figure 1). Dans ce cas, les zones extrapolées ne sont pas très informatives parce qu'un modèle de forêts d'arbres décisionnels a été utilisé. Mais dans le cas d'un MLG, d'un MAG ou d'un modèle d'ensemble, les zones extrapolées pourraient fournir certains renseignements et donner des idées de recherches futures.

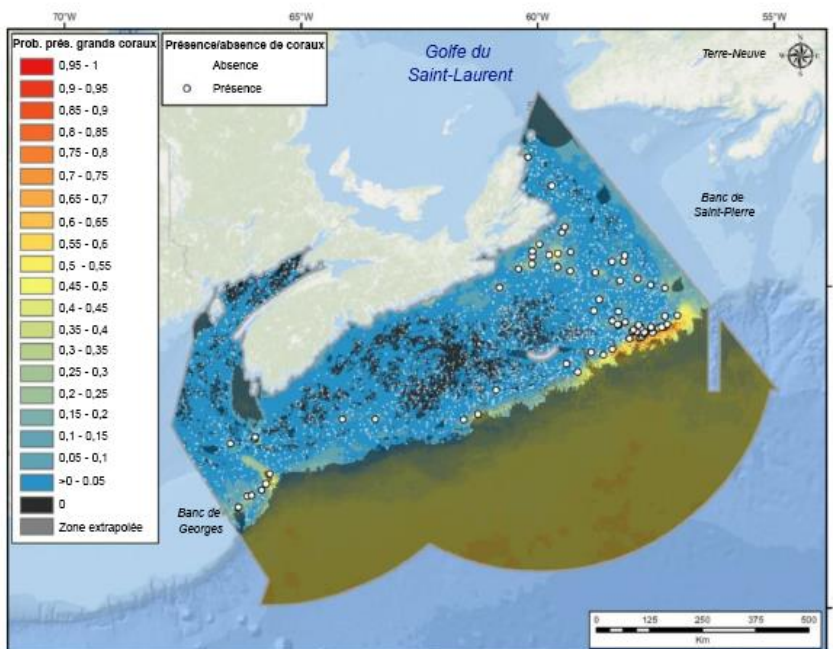


Figure 1. Zones d'extrapolation du modèle de forêts d'arbres décisionnels sur les données non équilibrées sur la présence et l'absence et les prises de grandes gorgones recueillies dans la Région des Maritimes entre 2002 et 2015. Les observations de la présence et de l'absence de grandes gorgones, ainsi que les prédictions de la probabilité de présence (Beazley *et al.* 2017) sont également illustrées.

8. PRÉDICTEURS ENVIRONNEMENTAUX. Couches océanographiques (page 29).

Les conditions océanographiques peuvent varier considérablement entre les saisons dans cette région, surtout pour les eaux côtières. Il semble que seules les données de la saison printemps-été ont été utilisées pour calculer la température, la salinité et la vitesse du courant. Y a-t-il une raison pour laquelle les données de l'automne et de l'hiver n'ont pas été incluses? Des précisions ou des explications seraient utiles.

9. OBSERVATIONS DES ESPÈCES (page 33).

Quatre des douze espèces sélectionnées n'ont pas été modélisées parce qu'on ne disposait pas de données adéquates sur leur occurrence en raison de la petite taille des échantillons, de la faible précision de l'emplacement spatial des observations ou de la faible résolution taxonomique. Toutefois, dans certains cas, il serait intéressant de voir les résultats de ces modèles et de déterminer si la baisse de rendement prévue est attribuable à l'une ou l'autre de ces raisons. Par exemple, les modèles de l'est du Canada pour les éponges ont en général un rendement inférieur à celui des autres groupes benthiques, probablement en raison de la faible résolution taxonomique. Mais il se pourrait que certains taxons à résolution taxonomique élevée (ordre des Pennatulacea) aient des affinités environnementales similaires. Cela pourrait ajouter de l'information à l'étape suivante mentionnée dans les conclusions (page 49) pour évaluer la taille de l'échantillon à laquelle le rendement de ces modèles commence à se dégrader. De plus, cela pourrait donner un exemple de la façon dont la qualité des données sur l'occurrence des espèces influe sur les statistiques du rendement des modèles (4^e recommandation).

10. ÉVALUATION DU MODÈLE (page 15) ET COMPARAISON DES STATISTIQUES DU RENDEMENT (page 48)

Ce cadre estime un seuil optimal qui dépendra du but de la modélisation. Quelle est l'optimisation du seuil utilisée dans le rapport? Un seuil qui maximise la TSS? La prévalence? Il serait utile d'avoir des précisions.

Considérations futures :

Données sur l'abondance

Ce cadre est axé sur les données sur la présence et l'absence, et les estimations de l'abondance ne sont pas prises en compte. Toutefois, si l'on veut utiliser les modèles pour l'évaluation des stocks ou les plans de récolte, il serait bon d'avoir un cadre tenant compte des données sur l'abondance, comme il est dit dans les conclusions (page 49). Bien qu'un cadre semblable puisse être appliqué aux données sur l'abondance, il faudrait ajouter certains renseignements supplémentaires, principalement liés aux statistiques d'évaluation du modèle, afin d'utiliser ce cadre pour les données sur l'abondance à l'avenir.

Modèles au niveau de la communauté

Douze espèces ont été sélectionnées parmi un plus grand groupe. Toutefois, à l'avenir, il pourrait être intéressant de modéliser toutes ces espèces en essayant d'inclure les interactions entre elles. Les progrès récents des techniques de modélisation des communautés permettent de modéliser les données sur les communautés afin de rechercher des groupes d'espèces présentant des réponses communes à l'environnement, comme la modélisation des archétypes d'espèces (Dunstan *et al.* 2011, Murillo *et al.* 2018) ou l'utilisation de modèles conjoints de répartition des espèces (p. ex. Warton *et al.* 2015; Ovaskainen *et al.* 2017) qui peuvent tenir compte des interactions entre les espèces et permettre d'intégrer des caractéristiques des espèces dans un cadre de modélisation utile pour déterminer les caractéristiques-réponses qui peuvent fournir des perspectives fonctionnelles, mécanistes et prédictives sur les processus qui

façonnent l'assemblage et la dynamique des communautés écologiques. Ces modèles sont probablement hors de la portée de ce cadre, mais il y aura lieu d'en tenir compte à l'avenir lorsqu'on inclura plus d'espèces.

Commentaires généraux :

L'inclusion de points de données d'entrée (présence et absence) dans les surfaces de la prédiction pourrait ajouter de l'information sur la distribution des données. Elle peut être plus compliquée à visualiser pour la zone d'étude des plates-formes, mais il faut en tenir compte.

Une fois que les résultats seront définitifs et que le cadre sera approuvé, il serait utile de rendre les scripts R disponibles.

D'autres commentaires d'ordre rédactionnel ont été fournis aux auteurs séparément.

Références

- Bakka, H., Rue, H., Fuglstad, G.-A., Riebler, A., Bolin, D., Illian, J., Krainski, E., Simpson, D., Lindgran, F. 2018. Spatial modelling with R-INLA: A review. *WIREs Computational Statistics* 10, <https://doi.org/10.1002/wics.1443>.
- Beazley, L., Kenchington, E., Murillo, F.J., Lirette, C., Gujjarro, J., McMillan, A., and Knudby, A. 2016. Species Distribution Modelling of Corals and Sponges in the Maritimes Region for Use in the Identification of Significant Benthic Areas. *Can. Tech. Rep. Fish. Aquat. Sci.* 3172: vi + 189p.
- Boudreau, S. A., Shackell, N. L., Carson, S., and den Heyer, C. E. 2017. Connectivity, persistence, and loss of high abundance areas of a recovering marine fish population in the northwest Atlantic ocean. *Ecology and evolution*, 7(22):9739-9749.
- Carstensen, J., Conley, D.J., Lophaven, S., Danielsson, A., Rahm, L., and Toompuu, A. 2002. Statistical analysis and modelling of phytoplankton dynamics: exploitation of data in the Nordic and Baltic Monitoring Programs, Nordic Council of Ministers.
- Dunstan PK, Foster SD, Darnell R (2011) Model based grouping of species across environmental gradients. *Ecol Modell* 222:955–963
- Foster, S.D., Dunstan, P.K., Althaus, F., Williams, A. 2015. The cumulative effect of trawl fishing on a multispecies fish assemblage in south-eastern Australia. *J. Appl. Ecol.*, 52:129-139.
- Jones, C.C., Acker, S.A., Halpern, C.B. 2010. Combining local- and large-scale models to predict the distributions of invasive plant species. *Ecol. Appl.*, 20: 311-326.
- Landis, J.R., and Koch, G.G. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33: 159- 174.
- Murillo, F.J., Kenchington, E., Tompkins, G., Beazley, L., Baker, E., Knudby, A., Walkusz, W. 2018. Sponge assemblages and predicted archetypes in the eastern Canadian Arctic. *Mar. Ecol. Prog. Ser.*, 597: 115-135.
- Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., Abrego, N. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecol. Lett.*, 20:561–576.
- Pearce, J., and Ferrier, S. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecol. Model.*, 133: 225-245.

-
- Pennino, M.G., Guijarro-Garcia, E., Vilela, R. del Rio, J.L., Bellido, J.M. 2019. Modeling the distribution of thorny skate (*Amblyraja radiata*) in the Southern Grand Banks (Newfoundland, Canada). *Can. J. Fish. Aquat. Sci.*, <https://doi.org/10.1139/cjfas-2018-0302>
- Polovina, J.J., Mitchum, G.T., and Evans, G.T. 1995. Decadal and basin-scale variation in mixed layer depth and the impact on biological production in the Central and North Pacific, 1960-88. *Deep-Sea Res.* 42: 1701-1716.
- Rooper, C., Wilborn, R., Goddard, P. 2019. Determining the best methods for model validation. In: Use of species distribution modeling in the deep sea (Kenchington et al., eds.). *Can. Tech. Rep. Fish. Aquat. Sci.* 3296. pp.24-26.
- Rue, H., Martino, S., Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. B*, 71: 319-392.
- Tien, N.S.H., Craeymeersch, J., van Damme, C., Couperus, A.S., Adema, J., Tulp, I. 2017. Burrow distribution of three sandeel species relates to beam trawl fishing, sediment composition and water velocity, in Dutch coastal waters. *J. Sea Res.*, 127: 194-202.
- Warton, D.I., Blanchet, F.G., O'Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C., Hui, F.K.C. 2015. So many variables: joint modeling in community ecology. *Trends Ecol. Evol.*, 30:766–779.

EXAMINATEUR : ANDERS KNUDBY, UNIVERSITÉ D'OTTAWA

Pour cet examen, je me suis laissé guider par les questions suggérées et j'ai fourni des réponses à chacune en caractères gras ci-après.

1. Le but du document de travail est-il clairement énoncé?
Oui, le sujet est clairement défini, et l'objet du document est clairement énoncé (et significatif).
2. Les données et les méthodes sont-elles adéquates pour étayer les conclusions?
Oui, mais j'ai quelques suggestions d'éléments qui pourraient ou devraient être mis à l'essai. Elles sont décrites dans mes commentaires ci-après.
3. Les données et les méthodes sont-elles expliquées de façon suffisamment détaillée pour évaluer correctement les conclusions?
Oui, le document est très détaillé et clair.
4. Si le document présente des conseils aux décideurs, les recommandations sont-elles présentées sous une forme utilisable, et les conseils reflètent-ils l'incertitude des données, de l'analyse ou du processus?
Je ne pense pas que le document fournisse des conseils aux décideurs en tant que tels, plutôt des conseils aux scientifiques qui, à leur tour, pourraient conseiller les décideurs.
5. Pouvez-vous suggérer d'autres domaines de recherche qui sont nécessaires pour améliorer nos capacités d'évaluation?

J'ai quelques suggestions d'éléments à prendre en considération – aucune n'est cruciale, mais certaines pourraient améliorer le document.

- D'après la figure de la page 13, il semble que la méthode des blocs ne réduise pas au maximum la dissémination des facteurs du CV d'entraînement et d'essai. La principale faiblesse réside dans l'attribution aléatoire des blocs aux facteurs, qui semble s'éloigner fort

peu de l'attribution aléatoire des points aux facteurs. Je suggérerais au moins de mettre à l'essai une méthode de mise en grappes.

- À la page 46 et à la page 51, vous dites que cette approche « élimine la dépendance spatiale entre les ensembles de données d'entraînement et d'essai », mais ce n'est pas exact. Elle la réduit un peu, mais il y a encore des blocs de données d'essai entourés de tous les côtés par des blocs de données d'entraînement.
- Pour une prévalence donnée, il y aura toujours un lien linéaire entre la précision et le kappa, alors il n'est pas nécessaire de les utiliser tous les deux. Je pense qu'il y a une certaine variation entre les deux parce que votre prévalence varie d'un facteur de CV à l'autre (pages 15-16).
- Comme Jessica le sait déjà, je travaille à l'élaboration d'un code pour évaluer si un paramètre d'incertitude donné (comme celui décrit aux pages 17 et 18) est effectivement lié à l'exactitude des prédictions. Le manque de preuves à cet égard, malgré ce que laisse entendre l'intuition, mérite peut-être d'être souligné.
- Pages 21-22 Lorsque vous avez parlé de l'importance des variables et des méthodes de calcul, vous devriez au moins mentionner :
 - L'approche fondée sur la permutation.
 - Que la méthode de l'omission d'une variable (« drop-one ») produira une très faible importance pour les prédicteurs fortement corrélés avec d'autres (même si cela aura été réglé, dans une certaine mesure, par le retrait de la variable plus tôt dans le processus).
- Est-ce qu'une interpolation par krigeage exécutée sur le prédicteur auto-covariable ne serait pas un prolongement naturel de l'approche de l'ACS, afin de permettre de l'étendre à l'ensemble du modèle? On peut aussi faire la même chose, mais pour les résiduels du modèle initial (une méthode communément appelée krigeage de régression).
- Je pense que l'une des faiblesses du document est que vous i) décrivez dans le cadre comment, p. ex., le modèle écologique constitue une base importante pour le modèle de données, mais que, ensuite, vous ii) construisez un modèle de données unique pour des espèces dont l'écologie est vraisemblablement très variable. Cela n'est pas du tout inhabituel parce que la plupart des MQH sont créés sans qu'on ait examiné attentivement les caractéristiques de l'espèce, les prédicteurs et les échelles des relations écologiques imaginées, mais cela semble contraire à l'objectif que vous avez énoncé de suivre les pratiques exemplaires. Exemples :
 - Vous dites que le substrat a été cartographié par voisinage plutôt que par cellule parce qu'on suppose que c'est préférable pour les espèces mobiles. Mais vous utilisez également cette couche de substrat pour les espèces non mobiles.
 - Vous ne modifiez pas la résolution spatiale des prédicteurs pour refléter l'échelle à laquelle chaque espèce peut « interagir avec » un prédicteur donné.
 - Vous n'adaptez aucun prédicteur pour qu'il soit spécifiquement pertinent pour la période en question (p. ex. les couches océanographiques sont quantifiées pour la saison printemps/été de façon identique pour toutes les espèces).

Je pense en fait que ce que vous faites est correct, parce que, comme vous le mentionnez à la page 49, « L'utilisation d'une seule série de données prédictives (c.-à-d. un modèle de données cohérent) était essentielle à la capacité du cadre de gérer efficacement plusieurs modèles d'habitat et espèces. » C'est simplement que vous aviez mentionné précédemment la façon dont ces caractéristiques propres à chaque espèce devraient être prises en compte dans

l'élaboration des données et des modèles. Mais vous écrivez aussi (page 50) : « La création d'un modèle IQH, quelles que soient les données d'observation disponibles, est utile au processus de modélisation sur deux plans. Premièrement, elle contribue au modèle écologique en clarifiant la compréhension actuelle de l'écologie de l'espèce. » Mais ensuite, vous ne continuez pas vraiment à utiliser ce « modèle écologique » pour quoi que ce soit. En fait, je ne suis même pas tout à fait certain de bien comprendre ce qu'est ce « modèle écologique ».

- Pour ce qui est de la chlorophylle, il ne faut pas oublier que les algorithmes d'estimation de la chlorophylle donnent eux-mêmes de très mauvais résultats à proximité de la terre; il serait donc probablement préférable d'exclure simplement les données sur la chlorophylle pour les pixels qui contiennent une quantité de terre.
- La distribution de fréquence des directions du vent et des vitesses du vent a-t-elle été incluse dans le calcul du fetch? Il semble pertinent de l'inclure, mais cela n'a pas été mentionné dans le document.
- Je suis intrigué par le prédicteur du substrat qui est lui-même prédit à partir d'autres prédicteurs. Cela semble inutile, car ce genre d'approche devrait être inutile, du moins dans les modèles ARS très bien adaptés (bien qu'elle puisse mener à des modèles plus interprétables).
- Vous devriez indiquer qui a élaboré les modèles d'IQH et la connaissance qu'avaient ces personnes des données sur l'occurrence au moment où ils l'ont fait. Il est pratiquement impossible d'élaborer des modèles IQH sans que les données introduisent un biais, si l'on sait à quoi ressemble la distribution des données.
- Page 42 : Vous avez des profondeurs négatives à la figure 4.4.
- Page 43 : Je ne sais pas si j'ai mal compris quelque chose, mais avez-vous tenu compte de toutes les données sur l'occurrence dans l'élaboration du prédicteur auto-covariable?
 - Dans l'affirmative, il est quelque peu injuste de comparer des modèles avec et sans ce terme, parce que les données « d'essai » ont ensuite été utilisées pour élaborer un des prédicteurs pour les modèles, y compris le terme auto-covariable.
 - Dans la négative, comment le terme a-t-il contribué aux prédictions pour les données « d'essai »?
- Vous mentionnez que l'IQH a un effet modérateur parce qu'il n'est pas biaisé par les données. Je comprends la logique ici, mais avez-vous vérifié si un ensemble composé uniquement de vos MLG et modèles ARS avait un rendement supérieur ou inférieur à celui de votre ensemble IQH+MLG+ARS lorsque vous l'avez évalué par rapport aux données d'essai? Je crois fermement que c'est au fruit qu'on juge l'arbre, alors à moins que l'IQH ne contribue à améliorer certaines statistiques d'essai, je ne vois pas sa valeur.