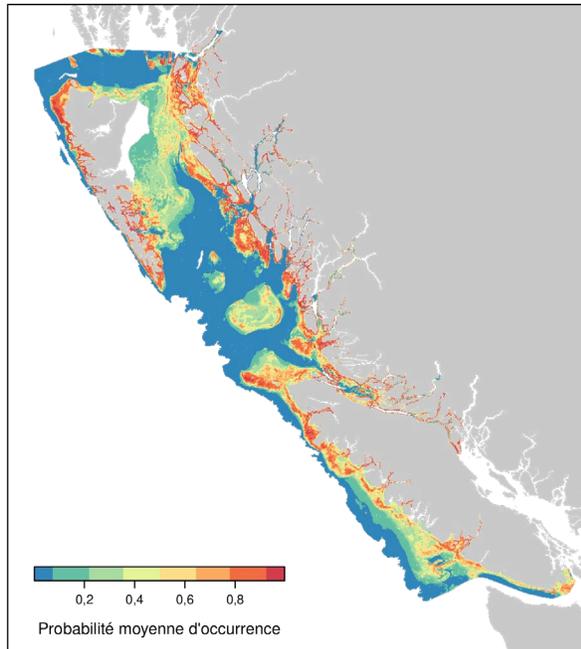




ÉLABORATION D'UN CADRE DE MODÉLISATION DE LA RÉPARTITION DES ESPÈCES ET SON APPLICATION À DOUZE ESPÈCES DE LA CÔTE CANADIENNE DU PACIFIQUE



Exemple de carte de prédiction de la répartition des espèces.

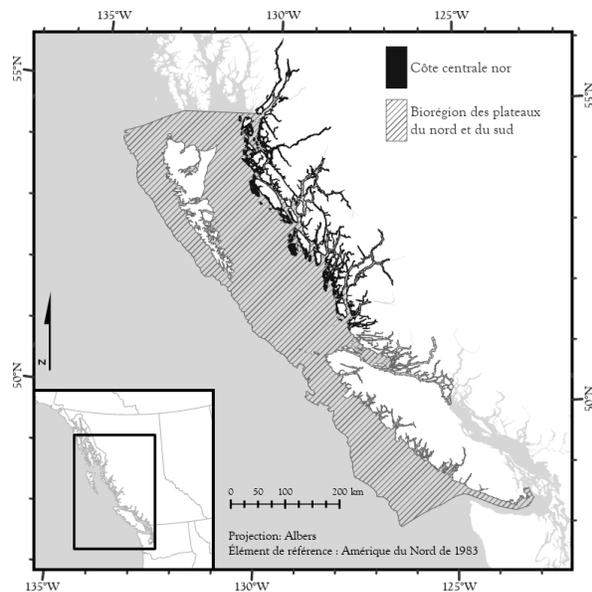


Figure 1. Les deux zones d'étude utilisées dans l'application du cadre. Les eaux côtières des côtes nord et centrale (résolution de 20 m) et les biorégions des plateaux du nord et du sud (résolution de 100 m).

Contexte :

Les modèles de la répartition des espèces (MRE) peuvent servir à prédire la répartition des espèces en établissant un lien entre les observations de l'occurrence des espèces et les données environnementales. La compréhension de la répartition des espèces peut guider diverses activités de gestion, y compris la planification spatiale marine et les interventions d'urgence en cas de pollution. Les MRE peuvent également relever des lacunes dans les connaissances écologiques, ce qui aide à cibler les futurs efforts d'enquête et de recherche. Cependant, les MRE 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 appliquée de façon appropriée (Elith et Leathwick 2009; Hawkins et al. 2003; Roberts et al. 2016).

La Direction des sciences de Pêches et Océans Canada (MPO) a proposé cet examen par les pairs afin de fournir une approche complète et normalisée de l'élaboration de MRE fondés sur des pratiques exemplaires pour assurer une qualité et une rigueur uniformes dans leur utilisation et leur application. L'évaluation et l'avis découlant de cet examen par les pairs régional du Secrétariat canadien de consultation scientifique (SCCS) serviront à élaborer des MRE et à les intégrer, le cas échéant, aux décisions scientifiques et stratégiques relatives à la gestion et à la conservation des espèces marines.

Le présent avis scientifique découle de l'examen régional par les pairs du 11-12 juin 2019 sur la Modélisation de la qualité de l'habitat – Pratiques exemplaires pour l'océan Pacifique du Canada. Toute autre publication découlant de cette réunion sera publiée, lorsqu'elle sera disponible, sur le [calendrier des avis scientifiques de Pêches et Océans Canada \(MPO\)](#).

SOMMAIRE

- Les modèles de répartition des espèces (MRE) sont pertinents pour plusieurs objectifs nationaux liés à la planification spatiale marine, aux évaluations de la vulnérabilité, aux interventions d'urgence et aux évaluations des stocks. L'application efficace des méthodes de MRE présente plusieurs avantages pour Pêches et Océans Canada (MPO), notamment la prise en compte explicite des contextes écologiques et de gestion, la préparation uniforme des données disponibles et l'application de méthodes analytiques appropriées pour construire et évaluer les modèles.
- La Direction des sciences de la région du Pacifique du MPO a proposé un cadre qui comprend un ensemble de lignes directrices sur l'élaboration de MRE uniformes, interprétables et défendables. Ce cadre peut être mis en œuvre à l'aide de scripts conçus spécialement (<https://gitlab.com/dfo-msea/sdm>) et écrits dans le langage de programmation statistique R (R Core Team 2018). Le cadre vise à aider les praticiens en suivant les pratiques exemplaires actuelles et en fournissant des conseils sur les aspects clés de la préparation des données, de l'ajustement et de l'évaluation des modèles, de l'estimation des incertitudes et de l'interprétation des résultats.
- L'application du cadre est illustrée par l'application de trois méthodes de modélisation de plus en plus complexes à douze espèces benthiques, dans deux zones d'étude de la côte canadienne du Pacifique (figure 1). Les prédictions ainsi obtenues sont évaluées à l'aide de paramètres de rendement normalisés, de diagrammes de diagnostic et de cartes, y compris l'importance relative et les effets marginaux des prédicteurs contributifs. L'application démontre l'importance d'un modèle de données cohérent, la mise au point de modèles multiples, dont un modèle fondé sur les connaissances, et l'estimation de l'incertitude.
- Les MRE pour les douze espèces sont destinés à faciliter la planification et la préparation des interventions en cas de déversement d'hydrocarbures. En général, les modèles s'adaptent bien aux données, comme le montrent les paramètres de rendement. L'utilisateur devra évaluer la pertinence de ces prévisions de modèle pour des utilisations autres que l'intervention en cas de déversement d'hydrocarbures, en fonction d'objectifs du modèle clairement énoncés et définis par leur application particulière.
- Les sources d'incertitude découlent des choix méthodologiques, ainsi que de la disponibilité et de la qualité des données. L'incertitude des prévisions du modèle est évaluée au moyen de quatre méthodes différentes qui réduisent ou mesurent l'incertitude dans l'espace sur toute la zone d'étude. D'autres sources d'incertitude sont prises en compte en suivant des pratiques exemplaires pour la sélection des données, dans la mesure du possible.
- Le cadre est une méthode scientifiquement défendable, transparente et reproductible pour produire des MRE. Il peut servir à produire des modèles pour d'autres espèces ou être adapté à d'autres degrés de modélisation.
- Le cadre a donné lieu à plusieurs recommandations pour l'élaboration et l'application réussies des MRE.

RENSEIGNEMENTS DE BASE

Aperçu de la modélisation de la répartition des espèces

Il est de plus en plus important de comprendre la répartition possible des espèces marines et de leurs habitats pour prendre des décisions éclairées sur leur gestion et leur conservation. Comme peu d'espèces marines sont inventoriées, les MRE sont devenus une approche courante pour estimer la répartition des espèces valorisées. Cependant, l'élaboration des MRE s'accompagne de plusieurs défis, notamment l'accessibilité croissante de méthodes statistiques et données environnementales sophistiquées, la variété des données sur les occurrences des espèces et la diversité des méthodes d'échantillonnage. De nombreuses considérations liées aux objectifs du modèle, à la préparation des données, à la sélection des variables et aux méthodes analytiques influent également sur le choix des méthodes appropriées et sur l'interprétation du modèle.

La pratique exemplaire comprend la définition d'un modèle écologique décrivant le contexte de l'analyse, un modèle de données décrivant les données sur les occurrences et les prédicteurs, et un modèle statistique reliant les deux (figure 2). Puisque la gestion d'une espèce ou de son habitat est inséparable de son écologie, le modèle écologique décrit à la fois le contexte écologique et de gestion. Sa définition commence par la pratique exemplaire qui consiste à articuler clairement les objectifs du modèle et inclut une description explicite de ce que l'on sait de l'espèce et de son écologie. Cela permet de déterminer l'étendue et la résolution nécessaires du modèle, ainsi que de construire le modèle de données. Le modèle de données est composé des observations des espèces et des prédicteurs environnementaux présumés caractériser leurs habitats. Les observations (c.-à-d. les données sur les événements) s'accompagnent d'un contexte complet qui décrit comment, quand, pourquoi et où elles ont été recueillies. De même, le contexte des variables prédictives (les données indépendantes) indique leur résolution, si elles sont statiques (p. ex. l'altitude) ou supposées statiques (moyennes à long terme des observations), et toutes les interactions pertinentes entre elles. Le modèle statistique sert à relier les observations d'une espèce aux données prédictives et à prévoir la répartition de l'espèce dans la zone d'étude. La complexité des méthodes statistiques est très variable. Les principaux critères de sélection des méthodes statistiques sont la capacité à atteindre les objectifs et la qualité de l'utilisation des données disponibles.

Facteurs liés à la sélection des données

Toutes les données sont biaisées et tous les modèles formulent des hypothèses. Par exemple, les observations d'une espèce peuvent n'avoir lieu que dans une partie de son aire de répartition, tandis que les hypothèses sur les déplacements d'une espèce détermineront si les résultats du modèle sont pertinents toute l'année ou pendant une saison particulière. La pratique exemplaire pour les MRE exige de prendre en compte ces biais et ces hypothèses lors de la sélection des espèces et des données environnementales appropriées pour la modélisation.

Les données sur les observations des espèces se présentent sous différentes formes et sont recueillies de nombreuses façons. Des plans de relevés ciblés ou non aléatoires peuvent mener à des prédictions de tendances dans la collecte de données plutôt que dans la répartition des espèces, surtout si l'échantillonnage est structuré sur le plan spatial ou environnemental (Araújo *et al.* 2019). Les observations qui couvrent autant que possible l'étendue spatiale et temporelle de la répartition de l'espèce sont préférables. La précision spatiale, temporelle et taxonomique réduite des données sur l'occurrence limitera les options de modélisation. Il faut envisager de

combiner des observations provenant de différentes sources de données pour augmenter la taille et l'étendue de l'échantillon, mais en prenant soin de ne pas masquer les biais d'échantillonnage, car cela pourrait donner des résultats trompeurs et accroître l'incertitude du modèle.

Diverses sources issues de technologies de télédétection (p. ex. imagerie satellitaire, drones, acoustique) et de modèles océaniques (altitude, type de fond et dynamique des océans) proposent des prédicteurs décrivant l'environnement marin. Il faut également tenir compte des variables spatiales, comme la distance par rapport à des caractéristiques physiques ou biotiques importantes (p. ex. le rivage). De plus, l'effort de pêche peut être un prédicteur important de la répartition des espèces dans les écosystèmes touchés (e.g., Foster *et al.* 2015; Tien *et al.* 2017). La résolution des prédicteurs marins peut couvrir des ordres de grandeur allant de l'échelle du kilomètre pour les prédicteurs de modèles océanographiques à l'échelle inférieure au mètre pour les prédicteurs topographiques d'origine acoustique. La pratique exemplaire exige que la résolution des prédicteurs corresponde à l'échelle à laquelle on présume qu'elle influence la répartition de l'espèce (e.g., Wiens 1989). Du fait de l'incohérence inévitable de la résolution entre les observations et les prédicteurs (et souvent entre les prédicteurs), certaines variables prédictives sont probablement bien représentatives d'un processus écologique, mais d'autres serviront de variables de remplacement pour les prédicteurs fonctionnant à des résolutions différentes. L'importance de ce décalage dépendra des objectifs du modèle, surtout si le modèle doit être transférable (c.-à-d. s'il doit fournir des prévisions fiables dans des endroits ou parfois à l'extérieur de la zone d'étude) ou s'il doit expliquer uniquement la tendance observée dans la zone d'étude.

Élaboration d'un cadre de modélisation de la répartition des espèces et son application à douze espèces de la côte canadienne du Pacifique

Région du Pacifique

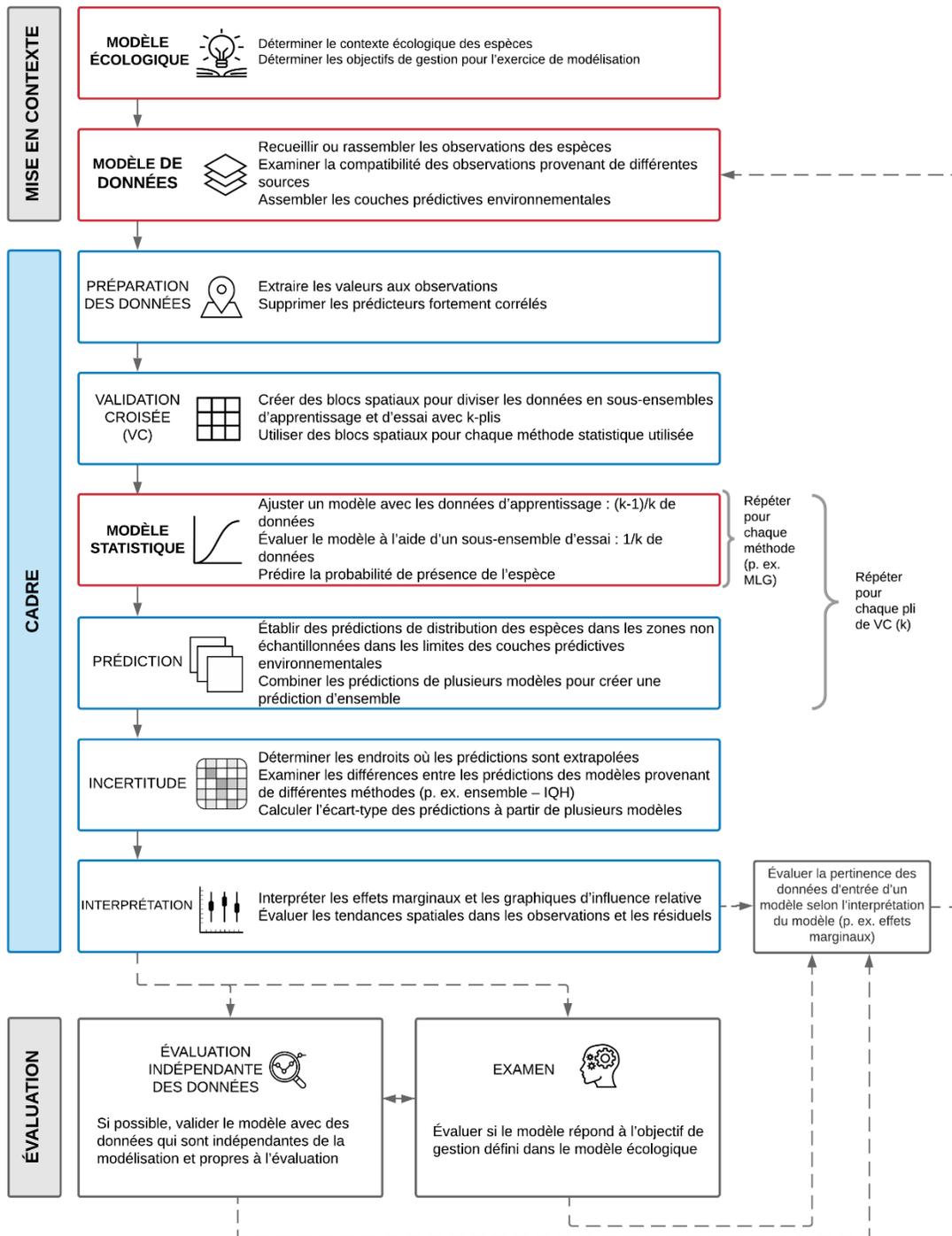


Figure 2. Aperçu du processus complet de modélisation, de l'élaboration du contexte du modèle (contextualisation) à l'évaluation des prédictions du modèle. Le cadre élaboré ici pour la modélisation de la répartition des espèces comprend une série d'étapes prescrites qui automatisent les pratiques exemplaires. Les modèles linéaires généralisés (MLG) sont un exemple de méthode de modélisation axée sur les données, tandis que les modèles d'indice de qualité de l'habitat (IQH) sont un exemple d'approche de modélisation de l'enveloppe fondée sur les connaissances.

ÉVALUATION

Aperçu du cadre de la MRE

Le cadre de la MRE a été élaboré en gardant à l'esprit les pratiques exemplaires en matière d'élaboration de modèles et a été peaufiné grâce à l'expérience acquise au cours de son application. Pour le mettre en œuvre le cadre, des scripts conçus spécialement (<https://gitlab.com/dfo-msea/sdm>) ont été écrits dans le langage de programmation statistique R (R Core Team 2018). Le cadre comprend six étapes (figure 2) : préparation des données, validation croisée, ajustement du modèle, évaluation du modèle, prévision et interprétation. Les modèles sont évalués selon une approche de validation croisée de blocs spatiaux afin d'accroître l'indépendance des observations utilisées pour ajuster (données de formation) et évaluer (données de test) les modèles. L'ajustement des modèles se fait de façon répétée à l'aide d'une variété de méthodes pour créer de multiples modèles qui sont combinés dans une prédiction d'ensemble. Les prévisions d'ensemble sont plus robustes (Oppel *et al.* 2011) et permettent de mieux comprendre l'incertitude des modèles. L'incertitude est réduite en limitant l'extrapolation. Elle est évaluée en mesurant la variation entre les prédictions du modèle qui composent l'ensemble, en comparant une prédiction du modèle fondée sur les connaissances à une prédiction du modèle fondée sur les données, et en mettant en correspondance les résiduels du modèle. On interprète les modèles en examinant l'influence relative des variables prédictives dans le modèle et les effets marginaux de chaque variable prédictive sur les observations.

Le cadre met l'accent sur les MRE axés sur les données (p. ex. modèles linéaires généralisés). Il examine également l'utilité des modèles d'enveloppe fondés sur les connaissances, qui fournissent un moyen de délimiter la répartition des espèces à l'aide de contraintes environnementales hypothétiques. Ces modèles sont utilisés dans le cadre pour décrire la compréhension écologique actuelle de la façon dont les espèces sont liées à leur environnement et pour évaluer dans quelle mesure cette compréhension écologique se reflète dans les MRE axés sur les données.

Données et méthodes utilisées pour l'application

Douze espèces benthiques ont été sélectionnées pour la modélisation afin d'illustrer l'application du cadre et de guider la planification des interventions d'urgence en cas de déversement d'hydrocarbures dans le cadre du plan régional d'intervention. Le cadre a été appliqué aux eaux du plateau continental de la côte canadienne du Pacifique, à l'exclusion du détroit de Georgie, à 2 résolutions : 100 m pour les biorégions des plateaux du nord et du sud et 20 m pour la région des eaux côtières des côtes nord et centrale (figure 1). Les couches de prédicteurs environnementaux utilisées dans les modèles (tableau 1) comprenaient la bathymétrie et ses dérivés, l'océanographie physique, chimique et biologique, le fetch et le substrat. Les prédicteurs étaient un mélange de variables mesurées interpolées, de prédictions tirées de modèles océanographiques et de modèles topographiques de complexité variable.

Les douze espèces benthiques choisies sont des espèces prioritaires pour la conservation (Gale *et al.* 2019) ou sont très vulnérables au pétrole (Hannah *et al.* 2017); elles ont été sélectionnées pour représenter une diversité de caractéristiques du cycle biologique, d'habitats et de communautés écologiques, avec différents niveaux de disponibilité et de qualité des données, afin d'évaluer comment le cadre a fonctionné avec différentes combinaisons de modèles écologiques et de données. Les données sur l'occurrence pour les douze espèces (tableau 2) proviennent de relevés systématiques ciblés. Pour certaines espèces, les

observations de l'occurrence proviennent de plusieurs relevés. Cela était nécessaire lorsque les échantillons étaient de petite taille (p. ex. 100 observations) ou lorsque la couverture spatiale des observations dans la zone d'étude était limitée dans l'espace géographique (p. ex. aucune observation dans la partie sud) ou dans l'espace prédicteur (p. ex. toutes les observations avaient eu lieu sur un habitat rocheux). Afin de réduire le biais d'échantillonnage, les observations des différents relevés n'ont été combinées que si les types d'engins d'échantillonnage étaient les mêmes. Pour quatre des douze espèces sélectionnées (l'étoile ocrée, la moule bleue, la palourde du Pacifique et la pennatule orangée), nous ne disposions pas de données adéquates sur l'occurrence pour la modélisation fondée sur les données dans nos zones d'étude.

L'application a utilisé trois méthodes d'ajustement du modèle qui couvrent toute la gamme de complexité du modèle : les modèles d'indice de qualité de l'habitat (IQH), les modèles linéaires généralisés (MLG) et les arbres à régression stimulée (ARS). Les IQH sont des modèles d'enveloppe fondés sur les connaissances qui dépendent de l'information publiée et de la consultation d'experts pour délimiter les répartitions des espèces en fonction des conditions environnementales (Brooks 1997; US Fish and Wildlife Services [USFWS] 1981). Les MLG et les ARS sont des méthodes statistiques qui relient les données sur l'occurrence aux prédicteurs environnementaux (Elith et Graham 2009). Selon les méthodes du cadre, les prédictions des modèles axés sur les données (p. ex. MLG et ARS) ont été combinées en une prédiction d'ensemble à l'aide d'une moyenne pondérée selon le rendement.

Les données sur l'occurrence utilisées pour l'ajustement et l'évaluation des modèles ont été divisées en cinq ensembles de données spatialement distincts, ou blocs, pour effectuer la validation croisée. Cinq modèles de MLG, d'ARS et d'ensemble ont été construits et évalués. La moyenne et l'écart-type d'un ensemble des paramètres d'évaluation et des surfaces de prévision ont été calculés pour les cinq modèles. Le rendement du modèle a été mesuré à l'aide du paramètre de la surface sous la courbe caractéristique de l'opérateur récepteur, qui varie de 0 à 1. Les scores inférieurs à 0,5 indiquent que les modèles sont pires que les modèles aléatoires et les scores de 1 indiquent que le modèle prédit parfaitement les données sur l'occurrence (Freeman et Moisen 2008; Merckx *et al.* 2011). Pour chaque espèce, tous les modèles ont été évalués avec les mêmes données afin de faciliter la comparaison des modèles. Toutefois, l'élaboration des modèles IQH différait de celle des modèles axés sur les données en ce sens qu'un seul modèle IQH a été construit pour chaque espèce. De plus, nous n'avons pu construire que des modèles IQH pour les espèces jugées déficientes en données (tableau 2).

**Élaboration d'un cadre de modélisation de la
répartition des espèces et son application à douze
espèces de la côte canadienne du Pacifique**

Région du Pacifique

Tableau 2. Sommaire des données sur l'occurrence choisies pour modéliser les douze espèces dans cette étude. Une régression linéaire généralisée et des arbres à régression stimulée ont été utilisés pour construire des modèles pour les huit espèces pour lesquelles on n'a pas jugé que les données étaient insuffisantes (DI). Les relevés de données sources (ABL = ormeau; RSU = oursin rouge; BHM = cartographie de l'habitat benthique; Cuke = holothurie; HBLL = palangre sur fond dur; IPHC = Commission internationale du flétan du Pacifique) représentent les évaluations ou la surveillance des stocks menées par la région du Pacifique du MPO et ses partenaires industriels. Pour cette application, les données tirées de ces relevés étaient de type présence-absence (P-A) ou absence seulement (Abs). La zone d'étude indique les eaux côtières (N) ou le plateau (S); Obs/n indique la prévalence des observations pour chaque espèce.

Espèce	Zone d'étude	Taille de l'échantillon (n)	Obs/n (%)	Années	Source (P-A)	Source (Abs)	Type de données	Précision des données spatiales
Ormeau nordique (<i>Haliotis kamtschatkana</i>)	N	2 293	22	2011-2016	ABL, RSU, BHM	-	Points	Emplacement du site (ABL) ou emplacement estimé le long du transect (RSU, BHM)
Panope du Pacifique (<i>Panopea generosa</i>)	N	9 350	58	2010-2017	GDK, BHM	RSU, Cuke	Points	Emplacement estimé le long du transect
Varech Pterygophora (<i>Pterygophora californica</i>)	N	6 607 ¹	3	2010-2017	ABL, Cuke, RSU, BHM ²	-	Points	Emplacement estimé le long du transect
Oursin rouge (<i>Mesocentrotus franciscanus</i>)	N	3 300	26	2010-2016	RSU, BHM	GDK, Cuke	Points	Emplacement estimé le long du transect
Zostère marine (<i>Zostera</i> spp.)	N	12 567 ³	4	2010-2017	GDK, Cuke, RSU, BHM ⁴	-	Points	Emplacement estimé le long du transect
Crabe dormeur (<i>Metacarcinus magister</i>)	S	391 ⁵	49 ⁶	1982-2009 ⁷	Enquêtes sur le crabe dormeur ⁸	-	Points	Position de début du déploiement des engins ⁹
Sébaste à dos épineux (<i>Sebastes maliger</i>)	S	4 937 ¹⁰	41 ¹¹	2003-2018 ¹²	HBLL, IPHC	-	Lignes	Position de début et de fin de la palangre
Sébaste aux yeux jaunes	S	4 937 ¹⁰	51	2003-2018 ¹²	HBLL, IPHC	-	Lignes	Position de début et de fin de la palangre

¹ Erratum : Octobre 2020 - 6 486 se lit maintenant 6 607

² Erratum : Octobre 2020 - ABL, Cuke, BHM se lit maintenant ABL, Cuke, RSU, BHM

³ Erratum : Octobre 2020 - 12 650 se lit maintenant 12 567

⁴ Erratum : Octobre 2020 - ABL, Cuke, BHM se lit maintenant GDK, Cuke, RSU, BHM

⁵ Erratum : Octobre 2020 - 6 702 se lit maintenant 391

⁶ Erratum : Octobre 2020 - 10 se lit maintenant 49

⁷ Erratum : Octobre 2020 - 1975-2018 se lit maintenant 1982-2009

⁸ Erratum : Octobre 2020 - Crabe, chalut à crevettes, BHM se lit maintenant Enquêtes sur le crabe dormeur

⁹ Erratum : Octobre 2020 - Début du transect (crabe, chalut à crevettes) ou position estimée sur le transect (BHM) se lit maintenant Position de début du déploiement des engins

¹⁰ Erratum : Octobre 2020 - 4,758 se lit maintenant 4,937

¹¹ Erratum : Octobre 2020 - 40 se lit maintenant 41

¹² Erratum : Octobre 2020 - 2003-2017 se lit maintenant 2003-2018

**Élaboration d'un cadre de modélisation de la
répartition des espèces et son application à douze
espèces de la côte canadienne du Pacifique**

Région du Pacifique

Espèce	Zone d'étude	Taille de l'échantillon (n)	Obs/n (%)	Années	Source (P-A)	Source (Abs)	Type de données	Précision des données spatiales
<i>(Sebastes ruberrimus)</i>								
Pennatule orangée <i>(Ptilosarcus gurneyi)</i>	S	DI	-	-	-	-	-	-
Palourde du Pacifique <i>(Leukoma staminea)</i>	N	DI	-	-	-	-	-	-
Étoile ocrée <i>(Pisaster ochraceus)</i>	N	DI	-	-	-	-	-	-
Habitat de la moule bleue <i>(Mytilus edulis, M. trossulus, M. galloprovincialis)</i>	N	DI	-	-	-	-	-	-

Résultats de l'application

Des cartes des observations de l'occurrence, des prévisions des modèles et des mesures de l'incertitude, ainsi que l'influence relative et les effets marginaux des variables prédictives environnementales ont été produits pour toutes les espèces. Les résultats pour le sébaste à dos épineux sont inclus à titre d'exemple (annexe A).

Les modèles d'ensemble les plus performants étaient ceux du sébaste à dos épineux et de l'ormeau (figure 3). Les modèles du panope, du varech, de l'oursin et de la zostère marine ont tous eu un rendement modérément bon. Les modèles les moins performants étaient ceux du crabe dormeur et du sébaste aux yeux jaunes, mais même si leur performance était relativement médiocre, les prédictions des modèles ont tout de même donné de meilleurs résultats que les solutions aléatoires.

Collectivement, les modèles IQH ont donné de piètres résultats, la majorité ayant produit des résultats inférieurs aux pires modèles axés sur les données. Le modèle IQH le moins performant était celui de la zostère marine, qui n'était pas mieux que les solutions aléatoires. Les différences de rendement des modèles sont probablement attribuables aux différences des modèles écologiques, ainsi qu'à la distribution et à la qualité des données disponibles sur l'occurrence, plutôt qu'à la simple taille de l'échantillon des observations.

Pour les espèces pour lesquelles la performance du modèle IQH était particulièrement faible, les comparaisons des tracés des effets marginaux et de l'influence relative (p. ex. figure A3) indiquent que les déterminants pourraient être mal compris ou mal mis à l'échelle. La détermination des problèmes propres à chaque espèce dans les modèles écologiques ou les modèles de données démontre l'utilité diagnostique des paramètres de rendement générés par le cadre. Par exemple, l'inégalité de l'influence relative des prédicteurs dans les modèles à validation croisée implique une mauvaise spécification du modèle, tandis que les différences dans les prévisions du modèle peuvent dénoter une incertitude dans la structure et la résolution du modèle. L'utilisation de plusieurs méthodes de modélisation permet donc d'évaluer la confiance dans la structure du modèle et la pertinence des prédicteurs.

Pour les huit espèces pour lesquelles les observations de l'occurrence sont adéquates, les paramètres de rendement du modèle et les mesures de l'incertitude spatialement explicites fournissent aux gestionnaires des outils pour évaluer la confiance dans les prédictions du modèle dans un domaine d'intérêt donné (voir la figure A2). Pour les quatre espèces pour

Élaboration d'un cadre de modélisation de la répartition des espèces et son application à douze espèces de la côte canadienne du Pacifique

Région du Pacifique

lesquelles les données sont insuffisantes, les modèles d'enveloppe préliminaires fournissent une base de référence à partir de laquelle on pourra élaborer des modèles de répartition fondés sur les données une fois que les données d'observation seront disponibles.

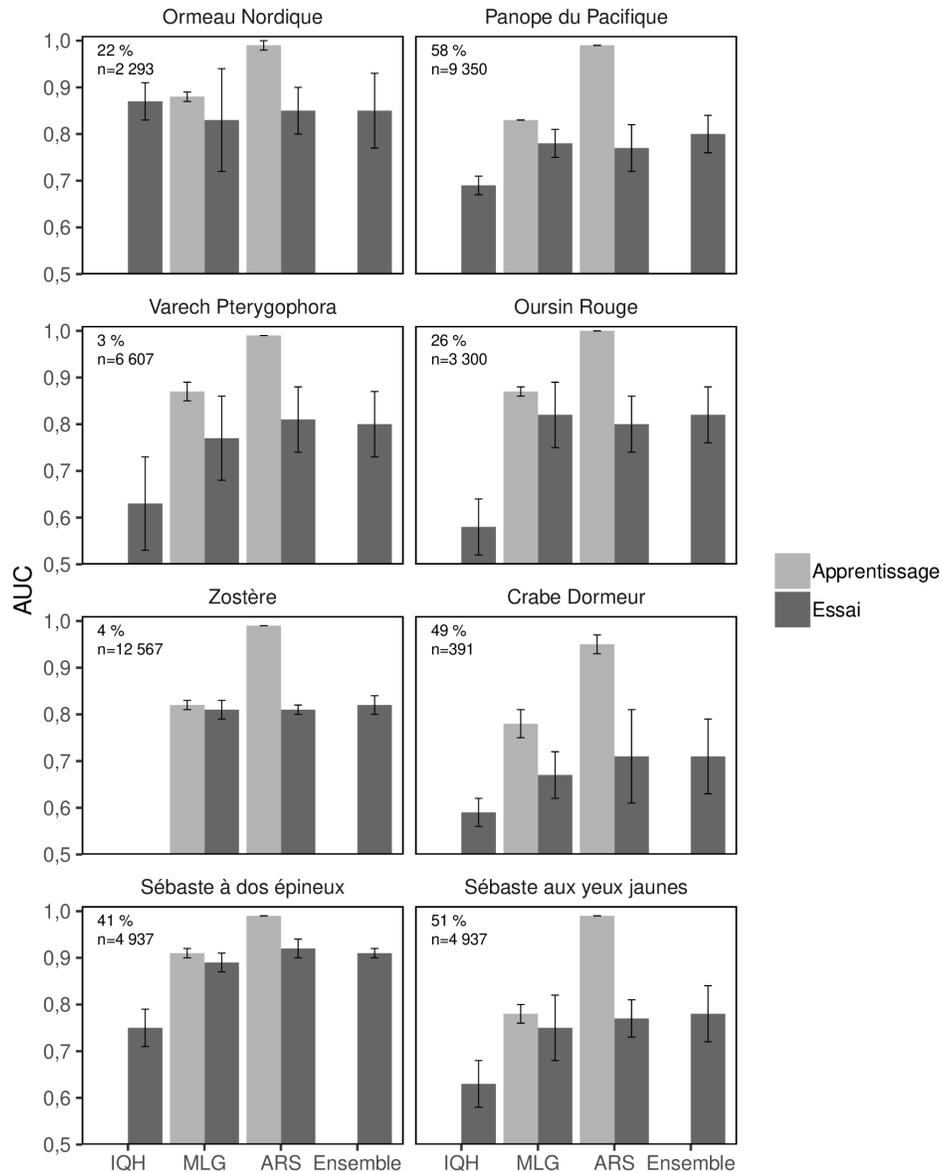


Figure 3. Rendement des huit modèles de répartition pour les espèces pour lesquelles les données sur l'occurrence sont suffisantes. La surface moyenne sous la courbe caractéristique de l'opérateur récepteur (AUC) est basée sur des essais de validation croisée de blocs spatiaux à quintuple épreuve et est indiquée pour chacun des quatre types de modèles : indice de qualité de l'habitat (IQH), modèle linéaire généralisé (MLG), arbre à régression stimulée (ARS), et ensemble. Les barres d'erreur représentent un écart-type. La prévalence et la taille de l'échantillon sont indiquées dans l'angle supérieur gauche. Les barres non représentées ont une valeur d'AUC inférieure à 0,5. Les modèles IQH et d'ensemble n'ont pas été évalués à l'aide de données sur la formation.

Sources d'incertitude

- L'incertitude découle de la disponibilité et de la qualité des données. Le biais d'échantillonnage dans les données sur l'occurrence est une source principale d'incertitude et peut être faussé si les données sur l'occurrence sont combinées à partir de sources multiples. Cela soulève la question de savoir quand il convient de combiner des données provenant de différentes sources. L'incertitude est réduite au minimum en suivant les pratiques exemplaires pour la sélection des données, dans la mesure du possible.
- L'interpolation à l'intérieur de la zone d'étude est généralement appropriée pour les modèles qui donnent de bons résultats. Si l'extrapolation à d'autres moments et lieux est jugée appropriée, elle doit être faite avec prudence.
- Il est possible d'obtenir des estimations de la confiance dans les prévisions des modèles ainsi qu'un aperçu des aspects de la qualité des données et du biais d'échantillonnage en comparant les prévisions des modèles fondés sur les connaissances et celles des modèles axés sur les données.
- Une approche d'ensemble de plusieurs modèles facilite l'estimation de l'incertitude des prévisions en permettant de comparer des modèles construits selon différentes méthodes.
- La variation entre les prédictions des modèles construits à partir de différents sous-ensembles d'observations (à l'aide de la validation croisée de blocs spatiaux) permet de mieux comprendre le biais d'échantillonnage spatial et de mettre en évidence les zones moins certaines. Les endroits où les relations modélisées sont moins certaines peuvent être une indication d'une non-stationnarité locale.
- L'incertitude découlant des erreurs dans les couches de prédicteurs, de la précision spatiale des observations, de l'absence de prédicteurs ou de l'inadéquation d'échelle entre les prédicteurs et les observations est prise en compte pendant la phase de préparation des données avant l'établissement du cadre, mais non dans les estimations numériques des incertitudes. De plus, l'incertitude des paramètres du modèle n'est pas prise en compte dans les évaluations de l'incertitude.

CONCLUSIONS ET AVIS

Le document de recherche présenté ainsi que la réunion officielle d'examen par les pairs et le présent avis scientifique constituent un examen approprié des documents de référence ainsi que des défis et des pratiques exemplaires pour l'élaboration de modèles de répartition des espèces.

Le cadre est une méthode scientifiquement défendable, transparente et reproductible pour produire des MRE. Il convient pour la production de modèles d'espèces supplémentaires ou l'adaptation à d'autres zones d'étude. Le cadre facilite également les mises à jour lorsque de nouvelles données sur l'occurrence ou de meilleurs prédicteurs sont disponibles. Cette approche itérative visant à réévaluer les données d'entrée du modèle pour déterminer si elles conviennent pendant la phase d'interprétation du modèle et à perfectionner par la suite les modèles (voir la figure 2) a permis d'améliorer le rendement du modèle dans de nombreux cas.

Une approche à plusieurs modèles facilite les estimations de l'incertitude en permettant de comparer différents types de modèles. Dans l'application présentée, l'élaboration de modèles d'enveloppe fondés sur les connaissances a aidé à définir la compréhension écologique actuelle des douze espèces et a servi de point de référence auquel les prévisions du modèle

d'ensemble axé sur les données ont pu être comparées. L'élaboration des prévisions du modèle d'ensemble a permis d'estimer l'incertitude en examinant la variation entre les modèles de composantes. L'influence relative et les effets marginaux des prédicteurs ont également contribué à guider les évaluations de la fiabilité des modèles des espèces individuelles.

Le cadre a été appliqué à douze espèces pour lesquelles la qualité et la quantité des données étaient variables. En général, les modèles s'adaptent bien aux données, comme le montrent les paramètres de rendement. Les douze MRE serviront à étayer les interventions d'urgence en cas de déversement d'hydrocarbures en mer. L'utilisateur devra évaluer la pertinence de ces prévisions de modèle pour des utilisations autres que l'intervention en cas de déversement d'hydrocarbures, en fonction de leur application particulière.

L'application a donné des résultats particulièrement bons pour des espèces d'intérêt commercial ou de conservation, pour lesquelles on disposait d'un vaste échantillon d'observations appropriées, même si les données d'observation pour certaines espèces (p. ex. le crabe dormeur) étaient spatialement biaisées vers un habitat plus approprié. Pour ces espèces, des observations supplémentaires obtenues grâce à un programme d'échantillonnage bien conçu permettraient d'améliorer les prévisions du modèle. Pour les espèces pour lesquelles on dispose de données d'observation de grande qualité, la modélisation de la densité des espèces constituerait une prochaine étape précieuse pour fournir des renseignements supplémentaires à une intervention en cas de déversement d'hydrocarbures et d'autres besoins de gestion.

Le cadre a donné lieu à plusieurs recommandations pour l'élaboration et l'application réussies des MRE.

- Définir clairement les objectifs du modèle, car ils guident l'élaboration des données et des modèles statistiques et sont essentiels à l'interprétation des résultats du modèle.
- Être prudent dans le choix des espèces et des données environnementales, car la modélisation avec des données biaisées peut donner des résultats médiocres ou trompeurs.
- Intégrer les connaissances écologiques locales dans la mesure du possible; ces connaissances sont particulièrement précieuses dans les situations où l'on manque de données.
- Élaborer des modèles d'enveloppe fondés sur les connaissances, car ils clarifient les connaissances écologiques actuelles, aident à déterminer d'importants prédicteurs environnementaux et fournissent un moyen d'estimer l'incertitude des modèles.
- Élaborer des prédictions d'ensemble à partir de plusieurs modèles, car l'approche d'ensemble produit des prédictions robustes et fournit un moyen d'estimer l'incertitude des modèles.
- Utiliser des données recueillies de façon indépendante pour une validation supplémentaire du modèle dans la mesure du possible.

Le document de recherche (Nephtin *et al.* 2019) qui accompagne ce processus contient des discussions techniques beaucoup plus détaillées et explicites sur les considérations relatives à la mise en œuvre de la modélisation de la répartition des espèces et constitue une ressource précieuse pour les praticiens.

LISTE DES PARTICIPANTS DE LA RÉUNION

Nom	Prénom	Organisme d'appartenance
Anderson	Sean	Secteur des sciences du MPO, Évaluation des stocks
Beazley	Lindsay	Secteur des sciences du MPO, Maritimes
Benoy	Nicholas	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, Maritimes
Goulet	Pierre	Secteur des sciences du MPO, Terre-Neuve
Gregr	Edward	Université de la Colombie-Britannique
Gullage	Lauren	Secteur des sciences du MPO, Terre-Neuve
Herborg	Matthias	Secteur des sciences du MPO, Sciences des écosystèmes
Hubley	Brad	Secteur des sciences du MPO, Maritimes
Kenchington	Ellen	Secteur des sciences du MPO, 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, Maritimes
Nepkin	Jessica	Secteur des sciences du MPO, Sciences des écosystèmes
Novaczek	Émilie	Secteur des sciences du MPO, Terre-Neuve
O	Miriam	Secteur des sciences du MPO, Sciences des écosystèmes
Pretty	Christina	Secteur des sciences du MPO, Terre-Neuve
Poehlke	Travis	Gestion des écosystèmes du MPO, Océans
Robinson	Cliff	Secteur des sciences du MPO, Sciences des écosystèmes
Rooper	Chris	Secteur des sciences du MPO, Évaluation des stocks
Rubidge	Emily	Secteur des sciences du MPO, Sciences des écosystèmes
Sameoto	Jessica	Secteur des sciences du MPO, Maritimes
Schut	Steven	Secteur des sciences du MPO, Évaluation des stocks
St. Germain	Candice	Secteur des sciences du MPO, Division des sciences océanologiques

Nom	Prénom	Organisme d'appartenance
Warren	Margaret	Secteur des sciences du MPO, Terre-Neuve
Wells	Nadine	Secteur des sciences du MPO, Terre-Neuve
Yakgujaanas	Jaasaljuus	Conseil de la Nation Haïda

SOURCES DE RENSEIGNEMENTS

Le présent avis scientifique découle de l'examen régional par les pairs du 11-12 juin 2019 sur la Modélisation de la qualité de l'habitat – Pratiques exemplaires pour l'océan Pacifique du Canada. Toute autre publication découlant de cette réunion sera publiée, lorsqu'elle sera disponible, sur le [calendrier des avis scientifiques de Pêches et Océans Canada \(MPO\)](#).

Araújo, M.B., Anderson, R.P., Barbosa, A.M., Beale, C.M., Dormann, C.F., Early, R., Garcia, R.A., Guisan, A., Maiorano, L., and Naimi, B. 2019. Standards for distribution models in biodiversity assessments. *Science Advances* 5(1): eaat4858.

Brooks, R.P. 1997. Improving Habitat Suitability Index Models. *Wildlife Society Bulletin (1973-2006)* 25(1): 163-167.

Carignan, K., Eakins, B., Love, M., Sutherland, M., and McLean, S. 2013. Bathymetric Digital Elevation Model of British Columbia, Canada: Procedures, Data Sources, and Analysis. NOAA National Geophysical Data Center (NGDC).

Davies, S.C., Gregr, E.J., Lessard, J., Bartier, P., and Wills, P. 2019. Bathymetric elevation models for ecological analyses in Pacific Canadian coastal waters. *Can. Tech. Rep. Fish. Aquat. Sci.* 3321: vi + 38p.

Elith, J., and Leathwick, J.R. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual review of ecology, evolution, and systematics* 40: 677-697.

Elith, J., and Graham, C.H. 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography* 32(1): 66-77.

Foster, S.D., Dunstan, P.K., Althaus, F., and Williams, A. 2015. The cumulative effect of trawl fishing on a multispecies fish assemblage in south-eastern Australia. *J. Appl. Ecol.* 52(1): 129-139.

Freeman, E.A., and Moisen, G.G. 2008. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecol. Model.* 217(1-2): 48-58.

Gale, K.S.P., Frid, A., Lee, L., McCarthy, J., Robb, C., Rubidge, E., Steele, J., et Curtis, J.M.R. 2019. Cadre d'identification des priorités en matière de conservation écologique pour la planification d'un réseau d'aires marines protégées et son application dans la biorégion du plateau nord. *Secr. can. de consult. sci. du MPO. Doc. de rech.* 2018/055. viii + 204 p.

Gregr, E.J. 2012. BC_EEZ_100m: A 100 m raster of the Canadian Pacific Exclusive Economic Zone. SciTech Environmental Consulting, Vancouver BC.

Gregr, E.J. 2014. Fetch Geometry Calculator Version 1.0 – User Guide. SciTech Environmental Consulting, Vancouver, BC.

Région du Pacifique

- Gregr, E.J., and Haggarty, D. 2017. Background Substrate and the integration of nearshore and deep water classifications (Draft). SciTech Environmental Consulting, Vancouver, BC.
- Hannah, L., St. Germain, C., Jeffery, S., Patton, S., and O, M. 2017. Application of a framework to assess vulnerability of biological components to ship-source oil spills in the marine environment in the Pacific Region. DFO Can. Sci. Advis. Sec. Res. Doc. 2017/057. ix + 145 p.
- Hawkins, D.M., Basak, S.C., and Mills, D. 2003. Assessing model fit by cross-validation. *Journal of Chemical Information and Computer Sciences* 43: 579-586. doi:10.1021/ci025626i.
- Masson, D., and Fine, I. 2012. Modeling seasonal to interannual ocean variability of coastal British Columbia. *Journal of Geophysical Research: Oceans* 117(C10): C10019:10011-10014. doi:10.1029/2012jc008151.
- Merckx, B., Steyaert, M., Vanreusel, A., Vincx, M., and Vanaverbeke, J. 2011. Null models reveal preferential sampling, spatial autocorrelation and overfitting in habitat suitability modelling. *Ecol. Model.* 222(3): 588-597.
- Nephtin, J., Gregr, E.J., St. Germain, C., Fields, C., and Finney, J.L. 2019. Development of a species distribution modelling framework and its application to twelve species on Canada's Pacific coast. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/nnn. Sous presse.
- Oppel, S., Gardner, B., O'Connell, A.F., Louzao, M., Miller, P.I., Meirinho, A., and Ramírez, I. 2011. Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. *Biol. Conserv.* 156: 94-104. doi:10.1016/j.biocon.2011.11.013.
- R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Roberts, D.R., Bahn, V., Ciuti, S., Wintle, B.A., Guisera-Arroita, G., Elith, J., Warton, D.I., Hartig, F., Dormann, C.F., Lahoz-Monfort, J.J., Hauenstein, S., Thuiller, W., Schröder, B., and Boyce, M.S. 2016. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* 40: 913-929. doi:10.1111/ecog.02881.
- Tien, N., Craeymeersch, J., Van Damme, C., Couperus, A., Adema, J., and 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.
- USFWS. 1981. Standards for the development of habitat suitability index models. U.S. Department of Interior, Fish and Wildlife Service, Division of Ecological Services ESM 103.
- Wiens, J.A. 1989. Spatial scaling in ecology. *Funct. Ecol.* 3: 385-397.

ANNEXE A. RÉSULTATS DU MODÈLE DE RÉPARTITION DES ESPÈCES POUR LE SÉBASTE À DOS ÉPINEUX

Cette annexe présente un exemple des résultats de la modélisation découlant de l'application du cadre de la MRE sur la côte canadienne du Pacifique, pour le sébaste à dos épineux. Voir les résultats de la modélisation pour toutes les autres espèces dans Nephin *et al.* (2019). La figure A1 illustre la distribution des observations de la présence et de l'absence utilisées pour la modélisation, la figure A2 montre les prévisions de la répartition du sébaste à dos épineux dans la zone d'étude et l'incertitude correspondante, et la figure A3 représente l'influence relative et les effets marginaux des variables prédictives environnementales.

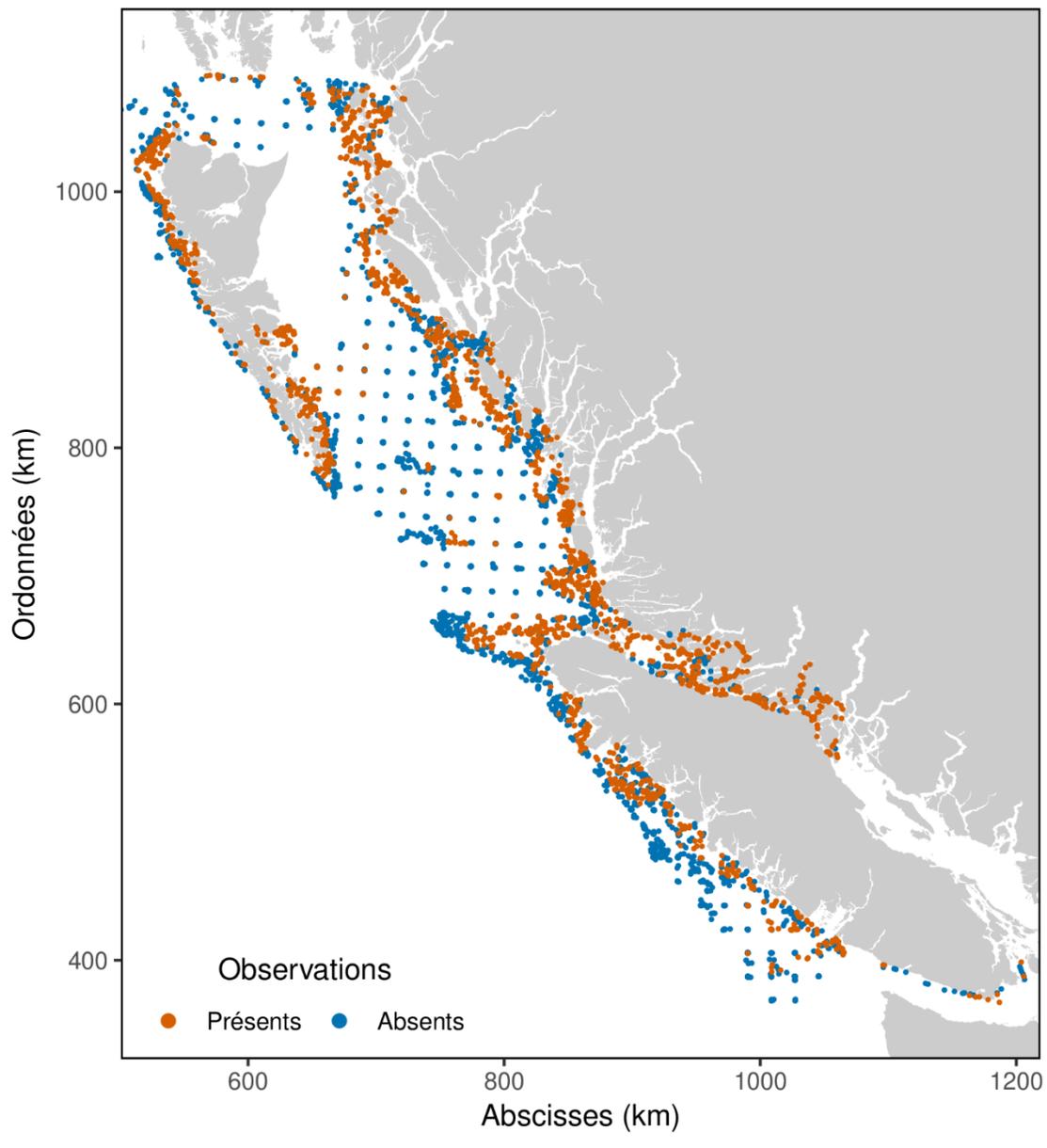


Figure A.1 Observations de la présence et de l'absence du sébaste à dos épineux dans la zone d'étude du plateau.

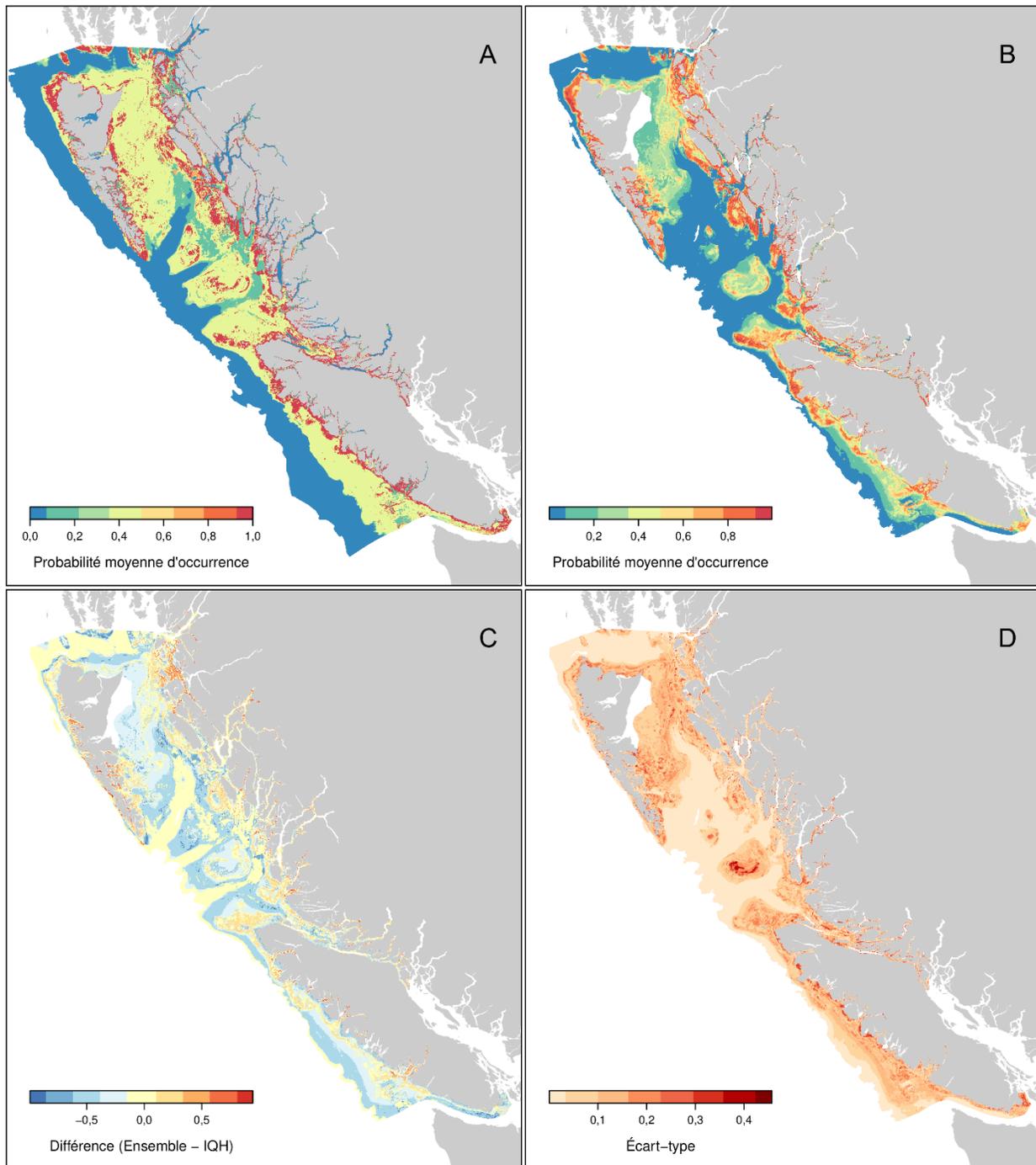


Figure A.2. Prévisions de la répartition du sébaste à dos épineux et incertitude correspondante. Prédictions de la probabilité d'occurrence à partir A) du modèle de l'indice de la qualité de l'habitat (IQH) et B) du modèle d'ensemble basé sur des modèles linéaires généralisés et d'arbres à régression stimulée. L'incertitude du modèle est représentée par C) la différence entre les prédictions du modèle IQH et celles du modèle d'ensemble et D) l'écart-type des prédictions de plusieurs modèles d'ensemble.

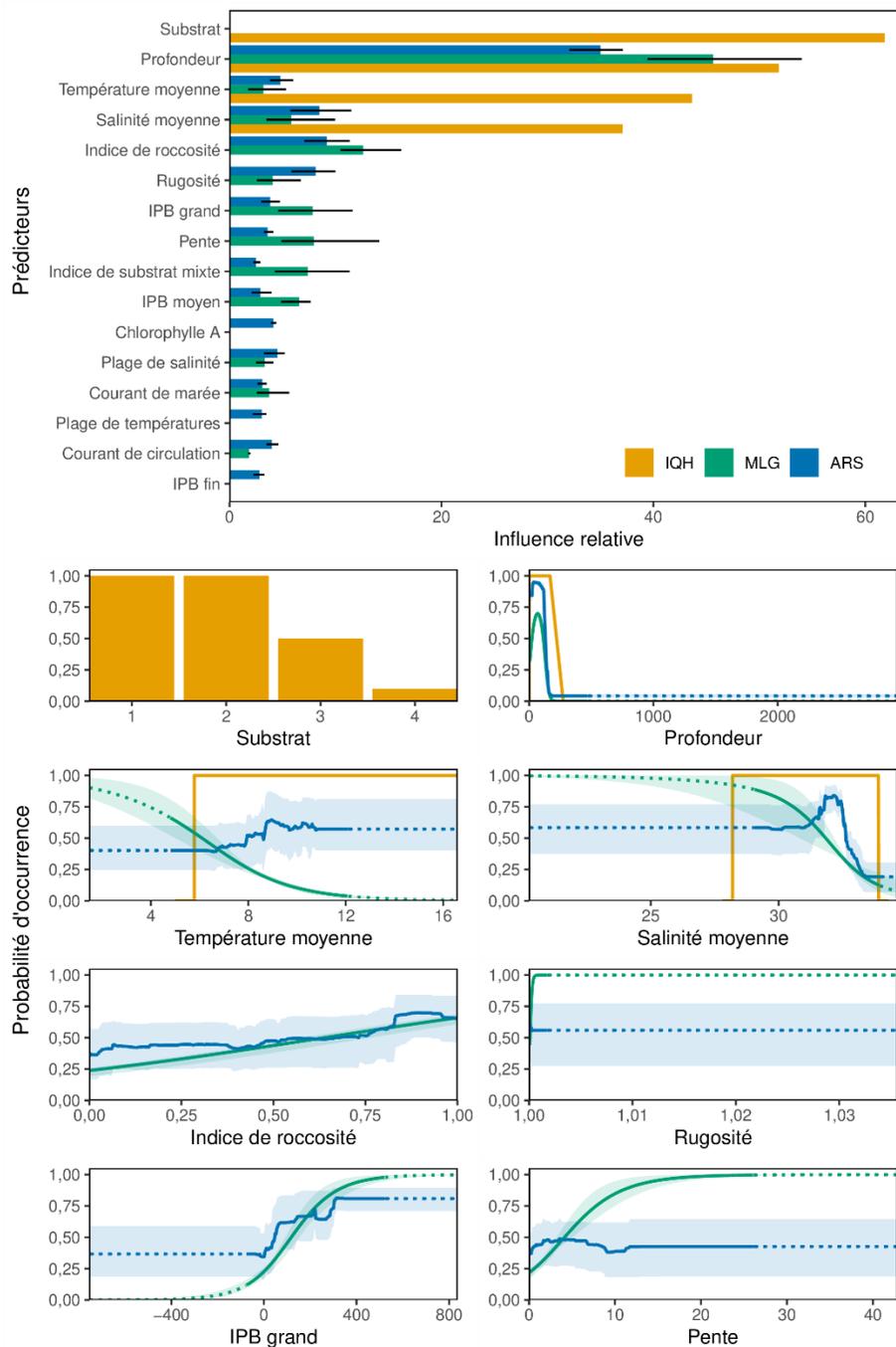


Figure A.3. Influence relative des prédicteurs (en haut) et des effets marginaux (en bas, plusieurs panneaux) pour les huit prédicteurs environnementaux les plus influents de chacun des modèles IQH, MLG et ARS du sébaste à dos épineux. Pour les modèles MLG et ARS, les barres des diagrammes de l'influence relative représentent la moyenne et les lignes montrent l'influence relative minimale et maximale des modèles à validation croisée à quintuple épreuve. Dans les diagrammes des effets marginaux, les lignes pleines représentent les effets marginaux moyens par méthode, et les zones ombrées représentent les effets marginaux minimaux et maximaux des modèles à validation croisée à quintuple épreuve. Le substrat a été représenté comme une variable catégorielle pour le modèle IQH et comme un indice de roccosité continu pour les modèles MLG et ARS.

CE RAPPORT EST DISPONIBLE AUPRÈS DU :

Centre des avis scientifiques (CAS)
Région du Pacifique
Pêches et Océans Canada
3190, chemin Hammond Bay
Nanaimo (Colombie-Britannique) V9T 6N7

Téléphone : (250) 756-7208

Courriel : csap@dfo-mpo.gc.ca

Adresse Internet : www.dfo-mpo.gc.ca/csas-sccs/

ISSN 1919-5117

© Sa Majesté la Reine du chef du Canada, 2020



La présente publication doit être citée comme suit :

MPO. 2020. Élaboration d'un cadre de modélisation de la répartition des espèces et son application à douze espèces de la côte canadienne du Pacifique. Secr. can. de consult. sci. du MPO, Avis sci. 2020/004. (Errata : Octobre 2020)

Also available in English:

DFO. 2020. *Development of a Species Distribution Modelling Framework and its Application to Twelve Species on Canada's Pacific Coast.* DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2020/004. (Errata: October 2020)