

Informations de base

Cette section contient des informations de base sur l'ensemble de données, adaptées à une entrée de métadonnées minimale.

Titre : Modèle de substrat profond (100 m) du plateau canadien du Pacifique

ID de l'ensemble de données : substrate100m-data

Statut : Complété

Contrôle de qualité : Complété

Résumé : Ce modèle de type de fond à substrat en eau profonde a été créé pour faciliter la modélisation de l'habitat et pour compléter les parcelles de fond du littoral. Il a été créé à partir d'une combinaison de couches dérivées de la bathymétrie en plus des observations du type de fond. À l'aide d'une classification par forêt d'arbres décisionnels, la relation entre les substrats observés et les dérivés bathymétriques a été estimée dans l'ensemble du site d'intérêt.

La trame est classée en : 1 - roche 2 - mélange 3 - sable 4 - vase

E-mail de responsable : Joanne.Lessard@dfo-mpo.gc.ca

Citez ces données comme : Dr. Dana Haggarty, Edward Gregr, Joanne Lessard, Cole Fields and Sarah Davies. 2018. Deep substrate (100 m) for the Pacific Canadian shelf. Published Sept 16, 2018. Data Distributor: J. Lessard, Fisheries and Oceans Canada, Nanaimo, BC.

Date de début : 1984-01-01

Date de fin : 2018-03-30

Informations de contact

Cette section contient les coordonnées du créateur des données et du responsable du programme.

Créateur de données :

Nom : Dr. Dana Haggarty

Email : Dana.Haggarty@dfo-mpo.gc.ca

Position : Chef de programme

Organisation : Gouvernement du Canada; Pêches et Océans Canada; STAR

Adresse : 3190, chemin Hammond Bay, Nanaimo, Colombie-Britannique, V9T 6N7, Canada

Phone : 250-756-7386

Co-créateurs : Cole Fields, Edward Gregr, Joanne Lessard, Sarah Davies, Jessica Finney, Jessica Nephin

Gestionnaire de programme :

Nom : Joanne Lessard

Email : Joanne.Lessard@dfo-mpo.gc.ca

Position : Chef de programme

Organisation : Gouvernement du Canada; Pêches et Océans Canada; MSEA Section

Adresse : 3190, chemin Hammond Bay, Nanaimo, Colombie-Britannique, V9T 6N7, Canada

Phone : 250-729-8364

Général

Métadonnées générales compatibles avec la norme de métadonnées Canada Open Data.

Catégorie de sujet : Océans

Date complet : 2018-03-30

Date publiée : 2018-09-17

Fréquence de mise à jour : Irrégulière

Niveau de l'ensemble de données : Base de données

Mots clés (Thésaurus des sujets de base du gouvernement du Canada) : plancher océanique, fond marin

Science

Cette section contient des métadonnées propres à la Direction des sciences du MPO.

Mots-clés scientifiques : rugosité, sédiments, substrat

Thème : Cartographie de base

Méthodes : Les observations du substrat ont été rassemblées à partir des sources suivantes : Service hydrographique du Canada (SHC), données de relevés en plongée, données du SHC sur les marais, Ressources naturelles Canada (RNC) et relevés par véhicule sous-marin téléguidé (VTG). Ces observations ont été mises en correspondance avec les classes de type de fond communes : roche, mixte, sable ou boue. À l'aide d'ArcGIS, les points ont été classés en données d'apprentissage (2/3) ou de test (1/3). Le sous-ensemble d'apprentissage est totalement exclu du processus de création du modèle et est utilisé pour évaluer le rendement de chaque modèle. Le partitionnement aléatoire des données en sous-ensembles d'apprentissage et de test ne permet pas de résoudre le problème de l'autocorrélation spatiale entre les observations.

Le code pour créer le modèle suit les étapes suivantes :

Une série matricielle de prédicteurs environnementaux est créée. Les observations ponctuelles sont superposées à la série matricielle, et les valeurs des prédicteurs sont extraites vers les points. Les enregistrements avec des valeurs BType4 non valides ou avec des valeurs S.O. des prédicteurs sont supprimés. Les données d'apprentissage sont pondérées en fonction de leur prévalence et sont utilisées pour ajuster un modèle de forêt d'arbres décisionnels à l'aide du progiciel Ranger (qui prend en charge la pondération des cas). Le modèle ajusté est utilisé pour prédire l'étendue de la série matricielle environnementale d'entrée, en classant la zone entière dans les catégories roche, mixte, sable ou boue. La surface prédite est exportée sous forme de fichier de données tramées GeoTIFF avec la même résolution que les prédicteurs de la série matricielle (dans ce cas, une résolution de 100 mètres) et avec la même projection. Les statistiques d'évaluation, notamment kappa et la précision par classe prédite, sont générées à l'aide des données de test retenues.

Étapes manuelles après la génération du modèle :

1) Effectuer une nouvelle projection de la couche vers EPSG : 3005 (le progiciel R ne prend pas en charge l'écriture du code EPSG de haut niveau dans les données du système de référence des coordonnées).

- 2) Créer un tableau d'attributs matriciels avec un champ SUBSTRATE (substrat) pour les catégories avec texte.
- 3) Ajouter un champ pour la prévalence dans le tableau des attributs matriciels.

Couches de prédicteur : À l'aide de la GDAL, les couches bathymétriques sources (NOAA et BCMCA) ont été rééchantillonnées et mosaïquées afin de produire une couche matricielle de bathymétrie de 100 mètres. Lorsque les trames de bathymétrie se chevauchaient dans l'espace (une grande partie de la région, à l'exception des zones de l'entrée Dixon, du passage Chatham et du canal de Portland), la couche de bathymétrie de la NOAA a écrasé la couche du BCMCA. Les dérivés de la bathymétrie ont été créés à l'aide du module Arcpy en langage Python (voir la section Scripts dans les métadonnées). La couche de données bathymétriques a été lissée en utilisant une moyenne focale avant de générer les dérivés suivants : pente -> écart-type de la pente, et courbure. Cette étape était nécessaire pour réduire les artefacts trouvés dans les dérivés. Au cours de l'élaboration, ces artefacts provenant des dérivés bathymétriques non lissés se sont répercutés sur les modèles de substrat prédits. Les données bathymétriques non lissées ont été utilisées pour générer les couches de l'Indice de position bathymétrique (IPB) normalisées, car ce traitement implique déjà un traitement avec un voisinage de cellules. Une couche catégorique pour la rugosité a été utilisée à partir du B.C. Marine Conservation Atlas (BCMCA). Elle a été convertie en une grille matricielle avec une résolution de 100 m pour s'aligner sur les autres couches de prédicteurs. Il a été décidé d'utiliser cette couche parce qu'elle avait été dérivée de la bathymétrie et ensuite modifiée manuellement. Voir les méthodes décrites dans le lien des sources de données. Les données bathymétriques originales non lissées ont été utilisées en tant que données d'entrée du modèle de forêt d'arbres décisionnels. Les couches de données sur l'énergie océanique ont également été incluses – les courants océaniques moyens de fond (Masson et Fine 2012), et la vitesse moyenne des marées sur le fond (Masson et Fine 2012). Les données originales du système de modélisation océanique régionale (ROMS) ont une résolution de grille de 3 km sur 3. Ces données ont été interpolées à l'aide de l'outil Spline avec interruptions (ESRI) et rééchantillonnées en trames avec une résolution de 100 m. Voir la section sur les scripts pour obtenir un lien vers le traitement des données du modèle ROMS. La bathymétrie provient principalement de la NOAA, avec une petite partie mosaïquée au nord-ouest à partir de la bathymétrie SciTech (voir Sources de données pour obtenir les liens). Les couches de l'Indice de position bathymétrique (IPB) ont été créées à l'aide de l'outil de modélisation du terrain benthique et ont été normalisées après avoir été calculées. Il est important de noter que lors du calcul de l'IPB, l'outil s'attend à ce que les données bathymétriques aient des valeurs négatives associées aux profondeurs plutôt que des valeurs positives.

Couches de prédicteur :

- 1 : Bathymétrie
- 2 : Pente (dérivé bathymétrique) – degrés
- 3 : Écart-type de la pente (dérivé bathymétrique)
- 4 : Indice de position bathymétrique vaste (dérivé bathymétrique) – Rayon intérieur : 25 – Rayon extérieur : 250
- 5 : Indice de position bathymétrique moyen (dérivé bathymétrique) – Rayon intérieur : 10 – Rayon extérieur : 100
- 6 : Indice de position bathymétrique fin (dérivé bathymétrique) – Rayon intérieur : 3 – Rayon extérieur : 25
- 7 : Courbure (dérivé bathymétrique; pente de la pente)
- 8 : Rugosité (BCMCA)
- 9 : Circulation
- 10 : Marée

Les sources de données :

Source : Bathymétrie :

<https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ngdc.mgg.dem:4956> (3 arc-second resolution)

Source : Littoral :
https://www.gis-hub.ca/dataset/chs_hwl
Source : Rugosité :
https://bcmca.ca/datafeatures/eco_physical_highrugosity/
Source : Océanographique :
<https://www.gis-hub.ca/dataset/bc-roms-bottomwaters> (3 km resolution)

Scripts des routines logicielles :

Préparation/création de couches de données : <https://gitlab.com/dfo-msea/environmental-layers/bathy-derivatives>
Modèle de forêt d'arbres décisionnels utilisant le progiciel Ranger : <https://gitlab.com/dfo-msea/environmental-layers/random-forest-substrate/-/releases/v1.1>
Traitement des ROMS : <https://gitlab.com/dfo-msea/environmental-layers/bc-roms-climatologies>
Interpolation à l'aide de l'outil Spline avec interruptions : <https://gitlab.com/dfo-msea/environmental-layers/spline-barriers-resample>
Alignement des couches : <https://gitlab.com/dfo-msea/environmental-layers/SnapExtractMask>

Qualité des données spatiales : Les points d'échantillonnage du substrat de fond sont considérablement plus denses près du rivage que dans les zones plus profondes, en raison des priorités de collecte de données et de la facilité de collecte. Ainsi, le modèle peut être biaisé en faveur des observations dans les zones moins profondes. De plus, les échantillons du SHC utilisés sont susceptibles d'être biaisés par rapport aux observations liées au substrat « dur », puisque le mandat de navigation sécuritaire du SHC exige que l'on se concentre sur les roches et les récifs. La prévalence des échantillons de substrat dur dans l'analyse introduit probablement un certain biais, mais l'effet sur les résultats du modèle est inconnu.

Précision de la position : La couche a une précision horizontale nominale de 100 m.

Précision de l'attribut : Les données ont été produites à partir d'un ensemble robuste de programmes d'échantillonnage lié au type de fond. Grâce à la composante de validation (modélisation de la forêt naturelle), nous pouvons être assez confiants concernant l'exactitude des attributs. La précision locale sera la plus élevée dans les régions côtières, là où l'échantillonnage du type de fond est le plus élevé, et plus incertaine dans les zones de faible densité d'échantillonnage. La précision est plus généralement compromise par la variabilité inhérente, souvent élevée, du fond de l'océan qui diminue avec la profondeur.

Cohérence logique : Les prévisions sont fondées sur la relation entre les caractéristiques géophysiques et les observations du type de fond provenant de diverses sources. Les méthodes ont été appliquées de manière uniforme dans toute la zone d'étude.

Exhaustivité : Ces données sont complètes pour la zone d'étude.

Données sur l'absence : Aucun. La couverture prévue est complète.

Incertitudes : En tant que sortie du modèle, la concordance entre le substrat prédit et observé variera dans la zone d'étude de manière imprévisible. Les expériences ponctuelles réalisées pendant l'établissement du modèle indiquent que le modèle peut être assez sensible aux données d'entrée. Les zones présentant une grande variabilité spatiale (c'est-à-dire des « mouchetures ») peuvent présenter une plus grande incertitude que les zones plus homogènes.

Restrictions d'utilisation : Les données sont limitées au secteur des Sciences du MPO pour l'analyse spatiale seulement.

Date de changement	Description du changement
2020-07-21	Mise à jour des champs d'attributs
2020-07-21	Mise à jour des champs d'attributs
2020-09-16	Mise à jour des champs d'attributs
2020-09-16	Mise à jour des champs d'attributs
2020-09-16	Mise à jour des champs d'attributs
2020-09-16	Mise à jour des champs d'attributs
2020-09-16	Mise à jour des champs d'attributs
2020-09-16	Mise à jour des champs d'attributs
2020-09-16	Test de mise à jour de l'ordre des résultats 2
2020-09-23	Mise à jour des champs d'attributs
2020-09-23	Mise à jour des champs d'attributs
2020-09-23	Mise à jour des champs d'attributs
2020-09-23	Mise à jour des champs d'attributs
2020-09-23	Mise à jour des champs d'attributs
2020-09-23	Mise à jour des champs d'attributs
2020-09-23	Mise à jour des champs d'attributs
2020-10-15	Suppression du fichier de forme actifs des données du point source concernant les observations
2020-10-15	Inclusion de la version du code en GitLab https://gitlab.com/dfo-msea/environmental-layers/random-forest-substrate/-/releases/v1.0
2020-12-21	Sous-données en eau profonde : MISE À JOUR DU PROGICIEL RANGER : https://github.com/ejgregr/substrate_model
2021-04-27	Sous-données en eau profonde : Mise à jour du nom GeoTIFF pour inclure Ranger dans le nom.

Les données des espèces :

Code et Nom	Les données sur l'âge	Type d'observation
Références :		
Référence : DuPreez, C. 2015. A new arc–chord ratio (ACR) rugosity index for quantifying three-dimensional landscape structural complexity. <i>Landscape Ecology</i> 30:181-192. [en anglais seulement]		
Référence : Masson, D., and I. Fine (2012), Modeling seasonal to interannual ocean variability of coastal British Columbia, <i>J. Geophys. Res.</i> , 117, C10019, doi:10.1029/2012JC008151 [en anglais seulement]		
Référence : S.; Slocum, N.; Pobuda, M.; Wright, D.J. Unified Geomorphological Analysis Workflows with Benthic Terrain Modeler. <i>Geosciences</i> 2018, 8, 94. doi:10.3390/geosciences8030094 [en anglais seulement]		
Référence : https://cran.r-project.org/web/packages/ranger/ranger.pdf		

Collaboration : SciTech Environmental Consulting

Les Autres Informations : Les données ont été validées dans le cadre d'un effort de 2018 sur les métadonnées par le groupe MSEA. Les aspects suivants des données ont été examinés : 1) référence spatiale (EPSG:3005), 2) comparaison visuelle pour valider que les emplacements sont cartographiés correctement, 3) évaluation des lignes de côte pour déterminer le chevauchement des rivages, 4) vérification des lacunes dans les données, 5) évaluation des valeurs nulles. Les origines des données régionales ont également été validées pour s'assurer de l'absence de chevauchement. Ces données n'ont pas encore été communiquées. La confidentialité et les autorisations seront décidées avant la diffusion.

Confidentialité : Pas protégé