

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 peu profond (20 m) de la côte canadienne du Pacifique

ID de l'ensemble de données : substrate20m

Statut : En cours

Contrôle de qualité : Complété

Résumé : Le modèle de type de fond à substrat peu profond a été élaboré à l'appui de la modélisation des habitats littoraux. Les sources de données comprennent à la fois les observations disponibles sur le type de fond et les couches prédictives environnementales, notamment les couches océanographiques, le fetch, la bathymétrie et ses dérivés. En utilisant la classification par forêt d'arbres décisionnels pondérée du logiciel Ranger R, la relation entre le type de fond observé et les couches prédictives peut être déterminée, ce qui permet de classer le type de fond dans l'ensemble des zones d'étude. Les fichiers de données tramées prédits sont classés comme suit : 1) roche, 2) mélange, 3) sable, 4) vase.

Les domaines du modèle de substrat catégoriel sont limités à l'étendue des couches bathymétriques d'entrée (voir les sources de données), qui est de 5 km à partir de la courbe de profondeur de 50 m.

E-mail de responsable : Dana.Haggarty@dfo-mpo.gc.ca

Citez ces données comme : Dr. Dana Haggarty and Edward Gregr, Joanne Lessard, Cole Fields, Sarah Davies, 2020. Shallow substrate model (20m) of the Pacific Canadian coast. Published Jan 31 2020. Data distributor: Dr. Dana Haggarty, Marine Spatial Ecology Section, Fisheries and Oceans Canada, Nanaimo, BC.

Date de début : 1984-01-01

Date de fin : 2019-12-31

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 sébaste côtier et morue-lingue

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

Gestionnaire de programme :

Nom : Dr. Dana Haggarty

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

Position : Chef de programme sébaste côtier et morue-lingue
Organisation : Gouvernement du Canada; Pêches et Océans Canada; STAR
Adresse : 200 - 401 Rue Burrard, Vancouver, Colombie-Britannique, V6C 3S4, Canada
Phone : 250-756-7386

Général

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

Catégorie de sujet : Océans

Date complet : 2015-01-01

Date publiée : 2020-01-31

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 de l'océan, fond marin

Science

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

Mots-clés scientifiques : rugosité, sédiments, substrat

Thème : Classification

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.

La bathymétrie et ses dérivés 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, courbure et rugosité. 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. 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 et la vitesse moyenne des marées sur le fond. La région du DG utilise le modèle NEMO pour la mer des Salish (Allen) comme données de base pour les couches de données sur l'énergie océanique. Les quatre autres régions sont issues du modèle ROMS de la Colombie-Britannique (Masson). 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 20 m. Voir la section sur les scripts pour obtenir un lien vers le traitement des données du modèle ROMS. Les données de base du modèle NEMO avaient une meilleure résolution dans le détroit de Georgia et il a donc été décidé que ce modèle serait utilisé pour la région du DG. Elles ont d'abord été interpolées à une résolution de cellule de 40 m (en raison des limitations de calcul) et ensuite rééchantillonnées à une résolution de 20 m. Les couches de données sur les marées (Tidal) et la circulation (Circulation) du modèle NEMO ont été lissées en utilisant l'outil Statistiques focales d'ArcGIS avec un voisinage de 13 cellules. Le fetch a également été inclus en tant que prédicteur. Les points de fetch ont été interpolés en fonction de la série matricielle (les méthodes précises et les données sources exactes ne sont pas claires). Les couches de données bathymétriques à une résolution de 20 m du MPO (qui comprenaient des données d'élévation terrestre) ont été utilisées comme source bathymétrique. 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 (dérivé bathymétrique; rugosité arc-corde)
- 9 : Circulation
- 10 : Marée
- 11 : Fetch

Les sources de données :

Source : <https://www.gis-hub.ca/dataset/coastline-fetch-bc>
Source : https://www.gis-hub.ca/dataset/bathy_20m
Source : <https://www.gis-hub.ca/dataset/substrate-preds-qcs-20m>
Source : <https://www.gis-hub.ca/dataset/substrate-preds-hg-20m>
Source : <https://www.gis-hub.ca/dataset/substrate-preds-sog-20m>
Source : <https://www.gis-hub.ca/dataset/substrate-preds-ncc-20m>
Source : <https://www.gis-hub.ca/dataset/substrate-preds-wcvi-20m>

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 : <https://gitlab.com/dfo-msea/environmental-layers/random-forest-substrate/-/releases/v1.1>
Traitement des données ROMS : <https://gitlab.com/dfo-msea/environmental-layers/bc-roms-climatologies>
Traitement des données NEMO : <https://gitlab.com/dfo-msea/environmental-layers/salishsea-nemo>

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 20 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 meilleure 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 les zones d'étude décrites.

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.

Restrictions d'utilisation : Data is restricted to DFO Science for spatial analysis only

L'historique des modifications :

| Date de changement | Description du changement |
|--------------------|---|
| 2020-12-21 | Substrat – ensemble de données à une résolution de 20 m : PROGICIEL R : remplacé par progiciel Ranger et dépôt de code (https://github.com/ejgregr/substrate_model). Le progiciel Ranger prend en charge la pondération des cas. |

Les données des espèces :

| Code et Nom | Les données sur l'âge | Type d'observation |
|--|-----------------------|--------------------|
| Références : | | |
| Référence : Du Preez, 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 : Soontiens, N., Allen, S., Latornell, D., Le Souef, K., Machuca, I., Paquin, J.-P., Lu, Y., Thompson, K., Korabel, V., 2016. Storm surges in the Strait of Georgia simulated with a regional model. <i>Atmosphere-Ocean</i> 54 1-21. https://dx.doi.org/10.1080/07055900.2015.1108899 [en anglais seulement] | | |

Collaboration : SciTech Environmental Consulting

Confidentialité : Pas protégé