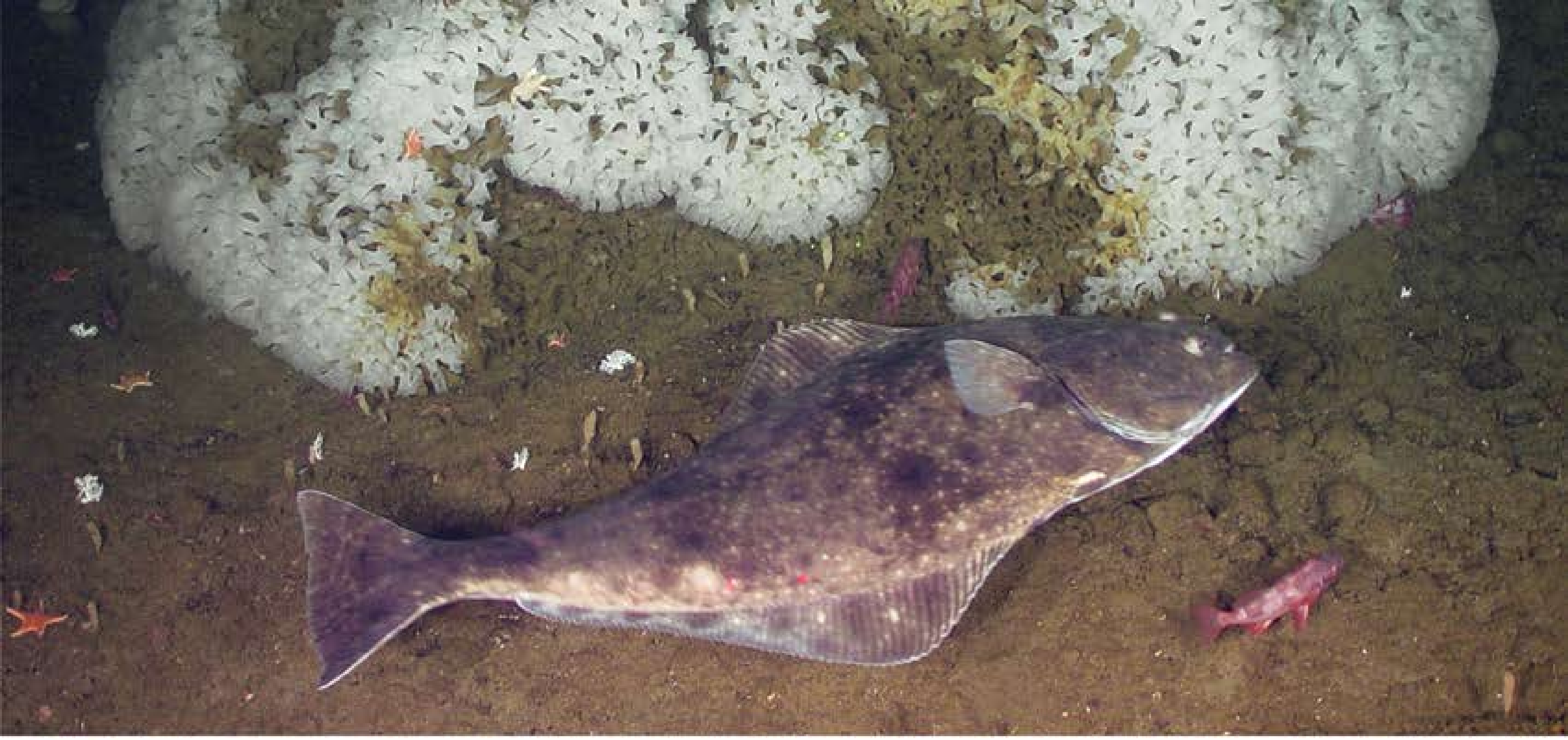
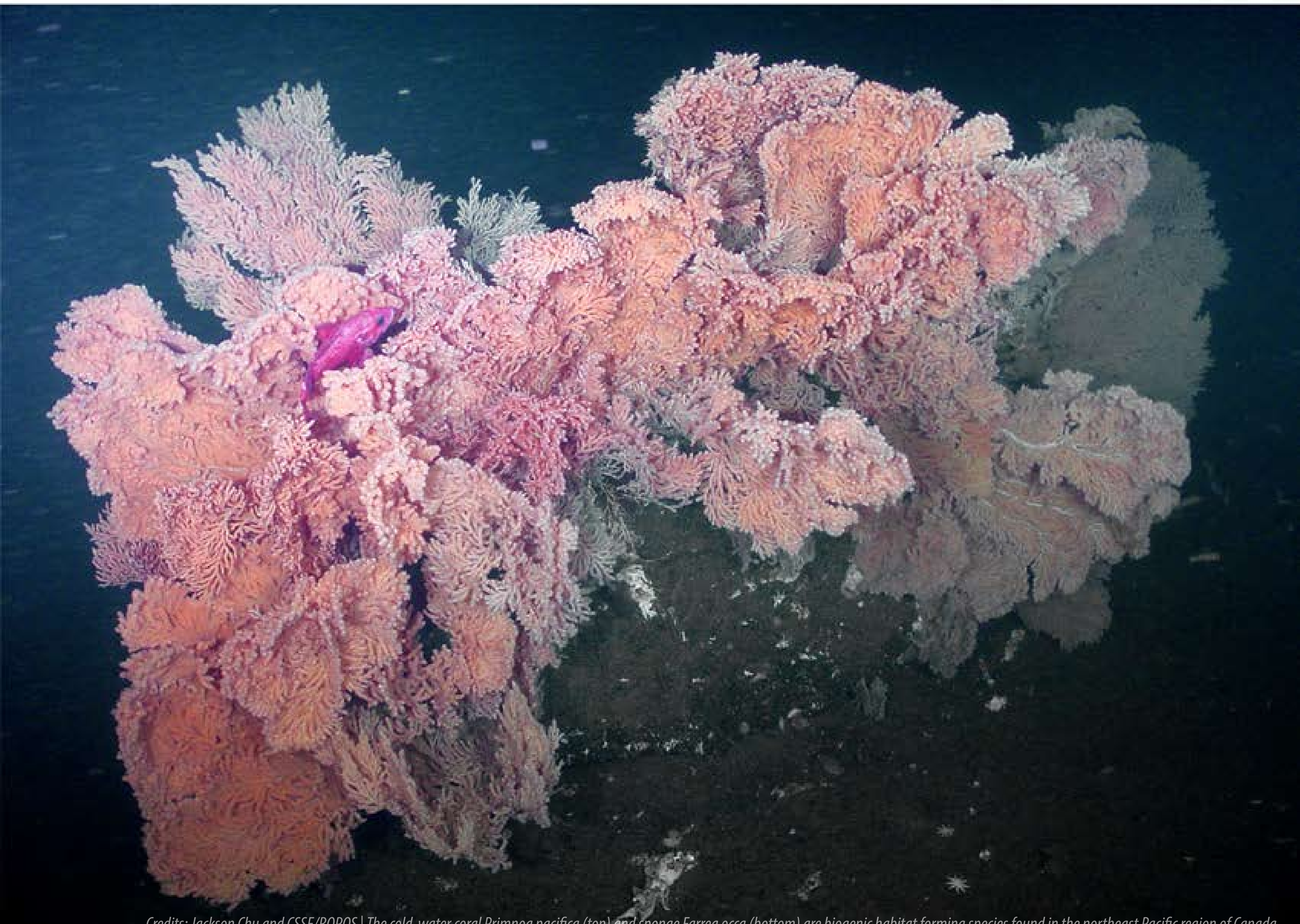


Report of Working Group 32 on Biodiversity of Biogenic Habitats

Edited by Janelle M.R. Curtis and Masashi Kiyota



Credits: Jackson Chu and CSSF/ROPOS | The cold-water coral *Primnoa pacifica* (top) and sponge *Farrea occa* (bottom) are biogenic habitat forming species found in the northeast Pacific region of Canada.

Report of Working Group 32 on Biodiversity of Biogenic Habitats

Edited by Janelle M.R. Curtis and Masashi Kiyota

Acknowledgements:

We thank all members of WG32 for their insightful contributions to its activities, for their rich body of research on the biodiversity of biogenic habitats, and for diligently ensuring that WG32 met its Terms of Reference. We are grateful to Anya Dunham (Canada) for making perceptive research contributions and for stepping in to serve as the interim co-chair from 2016-2018 while Janelle Curtis (Canada) was recovering from a severe accident in the mountains. We appreciate thoughtful guidance from PICES' BIO Committee over the years. We also thank members of the PICES Secretariat for their support of WG32, especially during Janelle Curtis' recovery. WG32's contributions support the work of the North Pacific Fisheries Commission (NPFC), which now works closely with PICES through the joint NPFC – PICES Framework for Enhanced Scientific Collaboration in the North Pacific Ocean.

Executive Summary

PICES' Convention came into force in 1992 while the North Pacific Fisheries Commission's (NPFC) Convention on the Conservation and Management of High Seas Fisheries Resources in the North Pacific Ocean came into force in 2015. In 2019, the two organizations agreed on a joint NPFC – PICES Framework for Enhanced Scientific Collaboration in the North Pacific Ocean. This framework identified three broad areas of joint interest to PICES and NPFC including research on stock assessment for NPFC's priority species, an ecosystem approach to fisheries management, and vulnerable marine ecosystems (VMEs) in the North Pacific Ocean. The Food and Agriculture Organizations (FAO) defined VMEs as "groups of species, communities, or habitats that may be vulnerable to impacts from fishing activities." NPFC identified four orders of corals – Alcyonacea, Antipatharia, Gorgonacea, and Scleractinia – as indicators of potential VMEs, but recognized that sponges and other invertebrates may also be identified as VME indicator taxa in the future. Although NPFC was formed a few years before the onset of this joint NPFC – PICES framework, WG32's activities on the biodiversity of biogenic habitats support joint research activities related to VMEs.

One of the many motivations for developing a working group to focus on the biodiversity of biogenic habitat (WG32) are the threats of bottom fishing and climate change to the

distribution and abundance of corals, sponges, and other organisms that provide habitat for marine organisms, including commercially valuable species. Biogenic corals and sponges provide habitat for early life history stages of fishes, and perform important functions in nutrient cycling and carbon sequestration. Catches of rockfishes and other commercially important species were higher where corals and sponges occurred as bycatch. The goals of PICES WG32 centered on collating data on the distribution of corals, sponges, and associated fauna, using models to predict the distribution of these biogenic habitats, proposing indicators for assessing them, and reviewing their associations with commercially important species.

The distribution of fauna is more difficult to monitor in marine environments, particularly in the deep sea. Yet, assessing the distributions of biogenic habitat is a precursor to understanding their role in the greater dynamics of the entire ecosystem (e.g. biodiversity, ecosystem functioning, fisheries, etc.). Therefore, predictive modeling methods are often used to estimate the distribution of marine fauna using available data. Knowledge of the spatial distribution of species is valuable for understanding ecosystem structure and function. Predictions of species' ranges from distribution models are often used to inform marine management and conservation efforts, but few studies justify the model selected or quantify the uncertainty of the model predictions in a spatial manner.

One of the key activities of WG32 was to review modeling approaches to predict the potential distributions of species and habitat suitability for biogenic habitats within national waters. Due to the now numerous species distribution modeling methods, there is some difficulty in selecting an appropriate algorithm. Habitat modelling approaches that were reviewed by WG32 members included MaxEnt, generalized additive models (GAMs), and Boosted Regression Trees. During the 2016 PICES annual meeting, WG32 convened a workshop on modeling approaches for corals and sponges and provided recommendations for data and modeling approaches that should be considered for those biogenic habitats. Technical aspects of the species distribution modeling, including the best-practices for generating input data, creating models and evaluating the results, a data-driven approach to defining bioregions, a multi-scale assessment of species distribution models, and an assessment of model transferability were also examined during the workshop.

Members recognized the value of producing multiple models for multiple areas because oceanic environments are subject to different currents and water masses and different levels of terrestrial influence can vary drastically in their biogeochemistry. They also recognized the importance of selecting an appropriate grid cell size and spatially examining uncertainty in model predictions and how it varies over space.

Across biogenic habitat types, assessment and monitoring efforts shared the following common themes: defining study objectives, assembling preliminary data, determining scale of interest, selecting indicators, determining study methods and sampling design, and full protocol review. Monitoring the status of biogenic habitats is difficult and costly, so WG32 focused on developing effective indicators for assessing and monitoring diversity and reviewing associations between biogenic habitats and commercial species. Indicators included trends in bycatch in commercial fisheries. WG32 members proposed a structured, iterative approach to designing monitoring programs for marine biogenic habitats that allows for rigorous data collection to inform management strategies, even when data and resource are limited.

WG32 members developed species distribution models (SDMs) for several major groupings of biogenic habitat-forming corals and sponges, in part to assess the primary drivers of suitable habitat for these taxa and to identify potential areas of high diversity of biogenic habitats. Some members and their collaborators introduced a new method for assessing the validity of VME indicator taxa (Gorgonians, Alcyonacea, Antipatharia and Scleractinia) and applied association analysis for identifying VME indicator taxa on the basis of sea-floor visual imagery. Others reconstructed long-term climate-driven range shifts in biogenic habitats and associated fishes in the western North Pacific Ocean from tropical to subarctic zones.

A key outcome of WG32's activities included the identification of large scale environmental and ecological predictors for the distribution and biodiversity of coral, sponge and associated taxa. Members put together an exhaustive group of measured and derived predictor variables for the North Pacific Ocean on a 1 km^2 grid These predictor variables were made available to all WG32 members through a shared drive, which allowed them to use the layers in their own modeling efforts.

Overall, WG32 members illustrated the value of SDMs for assessing potentially important environmental variables that could influence the distributions of biogenic habitat in areas with historically few observations (e.g. offshore deep waters). By combining multi-model outputs into a single composite index for corals and sponges, potential areas of suitable habitat for multiple biogenic habitats were identified. Members showed how multi-models can also be used to empirically assess areas that have been identified as ecologically and biologically significant marine areas (EBSAs).

Data availability will likely remain poor in offshore areas. SDMs are one tool that can extrapolate modelled species-environment relationships into areas where species records are rare and provide an empirical foundation for hypothesis development. Members also showed how applying a multi-model and multi-area approach can improve the interpretation of the modeled species-environment relationships and how to show which areas are predicted by the model to have high or low levels of uncertainty.

WG32 members also developed a biogeographical scheme for the Upper Bathyal zone (200-1000m) in the Pacific Ocean using octocoral distributions. The main driver for twelve proposed biogeographical units seems to be temperature, which is a defining feature of water masses. Members also investigated how environmental variables influence the distribution of corals, including those changing as seas warm. Climate change is causing coral declines in southern Japan due to bleaching but range expansion in northern Japan.

Contents

Activities of Working Group Members

MaxEnt modelling of biogenic habitat-forming cold-water corals and sponges in the Northeast Pacific Region of Canada

Jackson W.F. Chu[1], Jessica Nephin[2], Samuel Georgian[3], Anders Knudby[4], Chris Rooper[5], Katie S.P. Gale[2], Janelle Curtis[5]

[1]*Department of Ocean Sciences, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada*

[2]*Institute of Ocean Sciences, Fisheries and Oceans Canada, Sidney, BC, Canada*

[3]*Marine Conservation Institute, Seattle, WA, USA*

[4]*Department of Geography, Environment and Geomatics, University of Ottawa, Ottawa, ON, Canada*

[5]*Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC, Canada*

Background and Overview

This section summarizes the recent Canadian efforts in predictive modelling of the distributions of biogenic habitat-forming, cold-water corals and sponges (CWCS) in the northeast Pacific region of Canada (NEPC). Biogenic habitat-forming CWCS are often the focal taxa when applying empirical frameworks designed to identify sensitive benthic areas or those that are vulnerable to significant adverse impacts as a result of fishing activities (Ardron *et al.*, 2014; Dunn *et al.*, 2014). These frameworks focus on criteria that can almost entirely be met by the presence of dense CWCS communities. Key ecosystem roles that CWCS have include habitat provisioning for early life history stages of rockfish and sharks (Baillon *et al.*, 2012; Henry *et al.*, 2013), nutrient cycling, and carbon sequestration (Henry & Roberts, 2007; Oevelen *et al.*, 2009; Chu & Leys, 2010; Chu *et al.*, 2011; Kahn *et al.*, 2015). Biological traits such as slow growth, low reproductive output and dispersal rates, and long life spans (Roark *et al.*, 2009; Jochum *et al.*, 2012), make CWCS especially vulnerable to destructive bottom contact fisheries. Thus, activities that remove or destroy CWCS directly results in the net loss of the aforementioned ecosystem functions and services.

Our NEPC case study area is the Pacific exclusive economic zone (EEZ) of Canada. This region is notable for several globally unique biogenic habitat-forming CWCS communities (e.g. glass sponge reefs, Krautter *et al.*, 2001) and is where the majority of Canadian seamounts are located (>80%, DFO, 2019). Seamount ecosystems are of noted interest to the scientific community because of their global biogeography patterns and high productivity. Dense populations of commercially important fish and communities of biogenic-habitat forming CWCS are frequently found at seamounts (Clark *et al.*, 2011; Guinotte & Davies, 2014). Historically, the majority of empirical knowledge on the distribution of habitat-forming CWCS in the NEPC has focused on the samples collected on the continental shelf and slope with only a sparse number of records coming from the offshore, deep waters in the NEPC.

Published knowledge of biogenic habitat-forming CWCS on NEPC seamounts has come from research mostly done at two outliers with summits occurring in <30 m depths (Bowie and Cobb seamounts); most seamounts in the NEPC do not have summits that extend into the epipelagic zone.

Species distribution models (SDMs) were developed for several major groupings of biogenic habitat-forming CWCS. The goals of these analyses were to assess the primary drivers of suitable habitat for CWCS in the NEPC, identify potential areas of high CWCS diversity (i.e., areas that were suitable habitat for multiple CWCS groups), and determine the extent of CWCS suitable habitat at seamounts in this region. To assess areas as suitable habitat for multiple biogenic habitat-forming CWCS, a 'CWCS composite index' was created from combining spatial predictions from the individual models developed for each CWCS group. Several environmental data layers (n=32) were generated for SDM development; methods on how they were generated are described in detail. Additional details describe how historical CWCS occurrence records were queried and compiled from several regional databases and filtered before being used in SDM development. The CWCS-composite index was applied to empirically assess seamounts that have been provisionally identified as ecologically significant habitats within the Canadian EEZ (Ban *et al.*, 2016).

Data compilation

Environmental data layers

SDMs require environmental data that cover the extent of the area of interest and are selected for their potential to influence the distribution of modelled taxa. The North Pacific Marine Science Organization Working Group 32 on Biodiversity of Biogenic habitats (PICES WG32) created an expansive set of 32, coarse-resolution (1 km^2), environmental data layers for use in CWCS SDM development for the PICES convention area. These data layers were created following the methodology first described by

Davies & Guinotte (2011) and since expanded upon by others (Guinotte & Davies, 2014; Rowden *et al.*, 2017; Georgian *et al.*, 2019). Variables were obtained from a variety of sources (Table 1) and cover a range of bathymetry-derived variables, physico-chemical variables, and water column properties that are useful in predicting the potential distributions of benthic species in major ocean basins elsewhere. This case study focused on the area that covers the exclusive economic zone of the NEPC.

Table 1. Environmental data layers generated by PICES WG32 with associated data or methods reference. Shortened variable names are in square parentheses.

Variable name	Units	Native Resolution	Reference
Bathymetry-derived variables			
Bottom depth	metres	0.0083°	Becker *et al.*, 2009; Sandwell *et al.*, 2014
Aspect – east-facing [*eastness*]		0.0083°	Jenness, 2013a
Aspect – north-facing [*northness*]		0.0083°	Jenness, 2013a
Curvature – General [*gencurve*]		0.0083°	Jenness, 2013a
Curvature – Cross-Sectional [*crosscurve*]		0.0083°	Jenness, 2013a
Curvature – Longitudinal [*longcurve*]		0.0083°	Jenness, 2013a
Slope	degrees	0.0083°	Jenness, 2013a
Roughness [*VRM*]		0.0083°	Sappington *et al.*, 2007
Bathymetric position index [*bpi*] (1000m, 5000m, 10000m 20000m)		0.0083°	Jenness, 2013b
Seamounts polygon [*seamounts*]			Yesson *et al.*, 2011
Chemical variables			
Alkalinity	µmol l^{-1}	3.6 x 0.8-1.8°	Steinacher *et al.*, 2009
Dissolved inorganic carbon [*DIC*]	µmol l^{-1}	3.6 x 0.8-1.8°	Steinacher *et al.*, 2009
Omega - aragonite (Ω_{ARAG}) [*arag*]		3.6 x 0.8-1.8°	Steinacher *et al.*, 2009
Omega - calcite (Ω_{CALC}) [*calc*]		3.6 x 0.8-1.8°	Steinacher *et al.*, 2009
Dissolved oxygen [*oxygen*]	ml l^{-1}	1°	Garcia *et al.*, 2014a
Phosphate	µmol l^{-1}	1°	Garcia *et al.*, 2014b
Silicic acid [*dSi*]	µmol l^{-1}	1°	Garcia *et al.*, 2014b
Nitrate	µmol l^{-1}	1°	Garcia *et al.*, 2014b
Particulate organic carbon [*POC*]	g C m^{-2} yr^{-1}	0.05°	Lutz *et al.*, 2007
Physical variables			
Temperature	°C	0.25°	Locarnini *et al.*, 2013
Salinity	pss	0.25°	Zweng *et al.*, 2013
Current velocity – regional [*regfl*]	m s^{-1}	0.5°	Carton & Giese, 2008
Current velocity – vertical [*vertfl*]	m s^{-1}	0.5°	Carton & Giese, 2008
Current direction [*curdir*]	degrees	0.5°	Carton & Giese, 2008
Current direction – relative to aspect [*curapsect*]	degrees	0.5°	Rooper *et al.*, 2014
3D current-surface angle [*curang*]	degrees	0.5°	This study
Surface-layer properties			
Chlorophyll-a [*chl-a*]	mg m^{-3}	4 km	Aqua Modis (NOAA)
Photosynthetically active radiation [*PAR*]	W m^{-2}	4 km	Aqua Modis (NOAA)
Sea Surface Temperature [*SST*]	°C	4 km	Aqua Modis (NOAA)

Bathymetry

Bathymetry data and their derivatives are ubiquitous in benthic SDM studies. For this study, bathymetric data (https://topex.ucsd.edu/WWW_html/srtm30_plus.html) were obtained from the SRTM30+ layer at a native resolution of 0.0083° (~1 km)(Becker *et al.*, 2009; Sandwell *et al.*, 2014). The SRTM30+ layer (hereafter bathymetry) is derived from Sandwell et al. (2014), the Lamont-Doherty Earth Observatory Multibeam Synthesis Project, the JAMSTEC Data Site for Research Cruises, the National Center for Environmental Information (formerly the National Geophysical Data Center) Coastal Relief Model, and the International Bathymetric Chart of the Oceans.

Several derivative variables were calculated from the bathymetry layer. Slope, aspect, and curvature were calculated using the toolkit 'DEM Surface Tools v.2' (Jenness, 2004, 2013a) for ArcGIS (v.10.4, ESRI). The slope of each grid cell (in degrees) was calculated using the four-cell method (Horn, 1981; Jones, 1998). Aspect, or the maximum slope direction in degrees, is a circular variable (i.e. the difference between 0° and 359° is one unit) and thus was converted into two components: north-facing aspect (sin(aspect)) and east-facing aspect (cos(aspect)). Curvature generally describes the shape of the seafloor as a proxy to how the water column can interact with the substratum. Three types of curvature were calculated: general curvature, cross-sectional curvature, and longitudinal curvature. For general curvature, convex features have more positive values and concave features are more negative. For cross-sectional curvature, positive values are indicative of local features that may induce water divergence, and negative values are indicative of features that induce water convergence. Longitudinal curvature assigns positive values to features where water velocity is expected to decrease and negative values to features where velocity is expected to increase.

Roughness is a measure of topographical complexity and was calculated using the vector ruggedness measure (VRM) index method (Sappington *et al.*, 2007). VRM generates a dimensionless index that incorporates the bathymetry, aspect and slope layers. This process uncouples the slope from the resulting roughness index calculated for a raster cell by measuring the dispersion of vectors orthogonal to the terrain surface for a user-defined neighbourhood of cells. Here, the VRM index layer was generated using a neighbourhood of the adjacent eight cells from the one km^2 bathymetry layer.

Bathymetric Position Index (BPI) quantifies the relative elevation of a feature relative to the surrounding seafloor, with positive values indicating features that are elevated and negative values indicating features that are depressed. BPI values close to zero indicate relatively flat surfaces or areas with constant slopes. As biological processes are scale-dependent, and because BPI is calculated at a user-defined scale, a range of BPI layers was generated: 1,000 m (the fine-scale limit of the method based on the bathymetry layer), 5,000 m, 10,000 m, and 20,000 m. BPI layers were generated using the Land Facet Corridor Designer using the toolkit ‘Land Face Corridor Designer v1.2’ (Jenness, 2013b).

Oceanographic properties

Data layers for temperature, salinity, dissolved oxygen, and several dissolved nutrients were generated from data obtained from the World Ocean Atlas (WOA, v.2 2013). Carbonate chemistry (dissolved inorganic carbon, total alkalinity, $\Omega_{ARAGONITE}$, $\Omega_{CALCITE}$) were obtained from Steinacher et al. (2009). Chlorophyll-*a* (chl-a), sea surface temperature (SST), and photosynthetically active radiation (PAR) data were generated using mission composites (average of 2002-2016 data) from the MODIS/Aqua NOAA program at a resolution of 4 km and resampled to match the extent and resolution of the bathymetry layer without interpolation.

Several layers that characterize current and flow patterns were generated because of the strong influence of water movement on sessile species distributions (Genin *et al.*, 1986; Leys *et al.*, 2011). A bottom current velocity layer was generated using data from the Simple Ocean Data Assimilation model (v.3.4.1, Carton & Giese 2008) averaged as the composite of the years 1990-2007. Current velocities were calculated in both the horizontal and vertical dimensions, and current direction for each grid cell was calculated from zonal (U) and meridional (V) velocities according to the formula:

$$\text{Direction} = \frac{180}{\pi} \times \text{atan2}([u], [v])$$

Current flows to the south when values are close to +180° and -180°, flows to the east at +90°, flows to the west at -90°, and flows to the north at 0°.

Two additional current layers were created to capture flow patterns relative to bathymetry. The first layer generated was a two-dimensional (2D) current layer that quantified current-flow direction relative to seafloor aspect. In this 2D current layer, values of 0° indicate current flow is in the same direction as the direction of the steepest slope and values of 180° indicate current flows in the opposite direction of the steepest slope (*sensu* Rooper *et al.* 2014). The second layer generated was a three-dimensional (3D) current layer that quantified the current direction relative to the seafloor plane. For this 3D current layer, the slope and aspect layers were used to define the 3D orientation of the seafloor for each raster cell, from which the direction normal to the plane was defined in Cartesian coordinates. The direction of the water current was then defined at the seafloor based on the velocities relative to the east-flowing (x), north-flowing (y), and vertical-flowing (z) directions, and the angle between the two vectors was calculated. 3D current layer values near 90° indicate the current is flowing near-parallel to the seafloor and values less than 90° indicate current is flowing into the seafloor (e.g. northward flowing current into a south-facing seafloor slope).

WOA, carbonate chemistry, and current data layers were transformed to match the extent and resolution of the bathymetry layer using a variable up-scaling approach that approximates conditions at the seafloor (Davies & Guinotte, 2011). Each layer was first interpolated to a slightly higher resolution (0.5°) than its native resolution using inverse distance weighting, resampled to match the extent and resolution of the bathymetry data, and draped over the bathymetry data within its depth range. WOA data were available as 102 depth-binned layers from depths of 0-5500 m. Vertical resolution of WOA depth layers were 5 m (from 0-100 m), 25 m (100-500 m), 50 m (500-2000 m), and 100 m (2000-5500 m). Carbonate chemistry data (Steinacher *et al.*, 2009) were available in 33 depth-binned layers (6, 19, 38, 62, 93, 133, 183, 245, 322, 415, 527, 661, 818, 1001, 1211, 1449, 1717, 2014, 2340, 2693, 3072, 3473, 3894, 4329, 4775 m). Simple Ocean Data Assimilation (SODA) current data were available in 50 depth bins (5.0, 15.1, 25.2, 35.4, 45.6, 55.9, 66.3, 76.8, 87.6, 98.6, 110.1, 122.1, 134.9, 148.7, 164.1, 181.3, 201.3, 224.8, 253.1, 287.6, 330.0, 382.4, 446.7, 525.0, 618.7, 728.7, 855.0, 996.7, 1152.4, 1320.0, 1497.6, 1683.1, 1874.8, 2071.3, 2271.3, 2474.0, 2678.8, 2884.9, 3092.1, 3300.1, 3508.6, 3717.6, 3926.8, 4136.3, 4345.9, 4555.566, 4765.4, 4975.2, 5185.1, and 5395.0 m). This up-scaling approach has repeatedly been shown to be effective for many global and regional scale variables (Davies & Guinotte, 2011; Yesson *et al.*, 2012).

All data layers were projected in the world equidistant conic PICES azimuthal equidistance projection (-180 central meridian, 1.0 km linear unit). The edges of data layers do not completely extend into the coastal fjords habitats in the NEPC thus this study could not capture this habitat type which is known to harbour dense populations of CWCS taxa (Leys *et al.*, 2004; Gasbarro *et al.*, 2018).

Species records

While CWCS communities have been studied extensively in the NEPC, a comprehensive dataset of georeferenced occurrence records had not been compiled prior to this study. The majority of the CWCS records came from Fisheries and Oceans Canada (DFO) research and commercial catch databases. CWCS are recorded as incidental catch in the commercial databases. DFO research records spanned 1963-2017 and include targeted surveys for stock assessments of commercial invertebrate and groundfish populations and synoptic research bottom trawl surveys used for monitoring biogeographic areas within the Canadian EEZ. DFO commercial groundfish catch records include fisher and observer logbooks and dockside validation data from trawl, trap, and longline fisheries throughout this region; only records from 1996-2017 were used in this study because of recording reliability (M. Surry, DFO pers comm.). Queries of the DFO databases were done using internal, three-digit DFO codes (n=581) that uniquely identify CWCS taxa to varying taxonomic levels (e.g. 2A0 = Porifera, 3S6 = *Paragorgia arborea*). Additional records up to 2014 were compiled from the Royal British Columbia Museum (RBCM) archives which are now available in open access (Wheeler, 2018).

Although over 17,900 individual CWCS records were compiled from the various data sources, additional data management and quality control and assurance steps were required before species data could be meaningfully used in SDMs. CWCS records spanned several decades and several levels of taxonomic resolution (e.g. identified down to only phylum level or down to species level). Up-to-date taxonomic names and a complete taxonomic hierarchy were manually appended to records and verified in the World Register of Marine Species (Worms Editorial Board, 2018). *A priori*, expert-knowledge guided parsing of the records in order to filter out records that were inappropriate for use in SDMs focused on biogenic habitat-forming marine CWCS. For example, fresh-water sponges (e.g. *Spongila lacustris*) present in the museum records were filtered out based on expert-knowledge. Records of Calcarea class of sponges were excluded because no biogenic habitat-forming calcareous sponges occur in the NEPC. Because of the varying degrees of taxonomic resolution, only records with taxonomic resolution down to at least the class-level for sponges and order-level for corals were used in models. Records were pooled into six general biogenic habitat-forming CWCS groups. Four of these groups were orders of cold-water corals: Alcyonacea (soft corals), Scleractinia (stony corals), Antipatharia (black corals), and Pennatulacea (sea pens). We also modeled the distributions of two classes of sponges: Hexactinellida (glass sponges) and; Demospongiae (demosponges) (Figure 1).

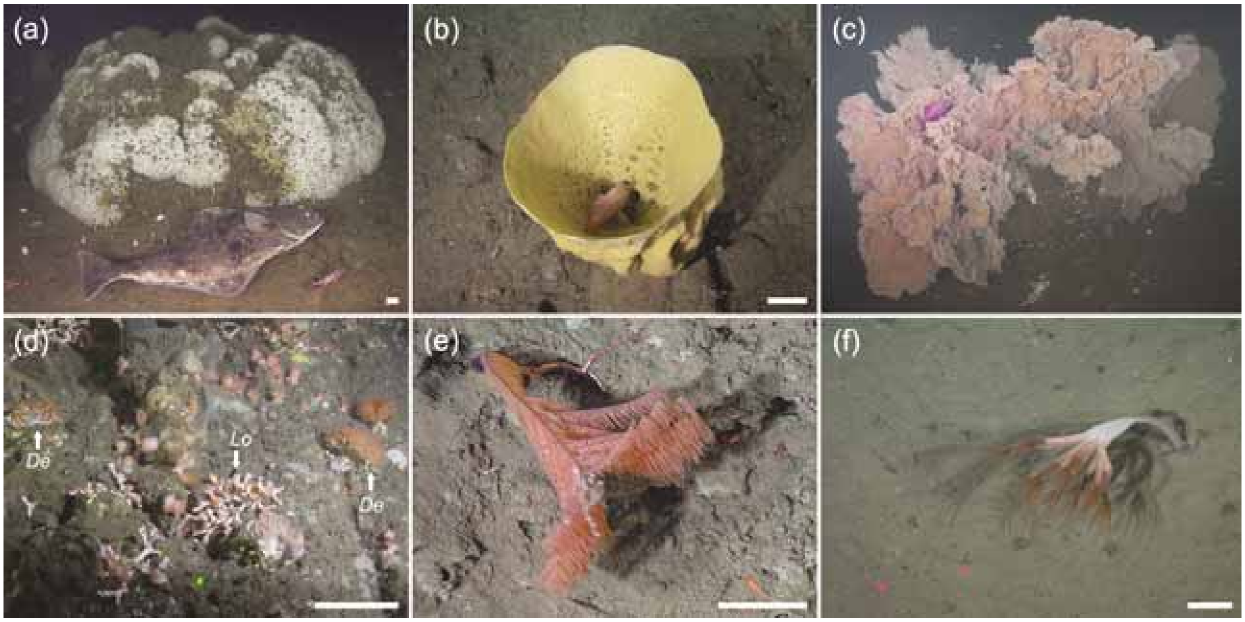


Figure 1. *In situ* examples of the major, biogenic habitat-forming cold-water coral and sponges from the Northeast Pacific Region of Canada. (a) Glass sponge *Farrea occa* (class Hexactinellida), (b) demosponge *Mycale loveni* (class Demospongiae), (c) soft coral *Primnoa pacifica* (order Alcyonacea); also a gorgonian coral. (d) Stony corals *Desmophyllum* sp. (De) and *Lophelia pertusa* (Lo) (order Scleractinia), (e) black coral *Bathypathes* sp. (order Antipatharia), (f) sea pen *Umbellula* sp. (order Pennatulacea). Scale bars: (a,b,d) 10 cm, (e-f) 5 cm. No scale bar was available for (c).

Historically, several families of biogenic habitat-forming corals were associated with the now-defunct order Gorgonacea. However, since it is still common to refer to these corals (now in the order Alcyonacea) as 'gorgonians', we created a gorgonian model by using a subset of records (n=428) in the 'soft coral' group that had at least a family-level of identification from the coral families: Anthothelidae, Paragorgiidae, Corallidae, Keroeididae, Acanthogorgiidae, Plexauridae, Gorgoniidae, Chrysogorgiidae, Primnoidae, and Isididae following the gorgonian grouping of Miyamoto *et al.* (2017). Sponge diversity and plasticity is problematic for species-level identification.

While the majority of glass sponges occurring in the NEPC are biogenic, habitat-forming types, demosponge diversity occupy niches that range from intertidal to deep-sea with many species in this region being encrusting, non-habitat forming morphotypes. To prevent intertidal species from adding uncertainty to models developed for biogenic habitat-forming types, only demosponge records occurring in >100 m depths were used in the models which improved the performance of the final models (Chu *et al.*, 2019).

Following Davies & Guinotte (2011), coral and sponge data for each group were gridded to the 1 km^2 resolution of the environmental data layers and redundant records in each cell were removed (i.e., only one presence record per cell per group). Table 2 summarizes the final number of records in each of the CWCS groups used in this study.

Table 2. The number of records used in group-specific cold-water coral and sponge (CWCS) models. Trawl-absences were generated from synoptic trawl surveys. The shallow extent of demosponge records was truncated at 100 m to remove non-biogenic habitat-forming types (e.g. encrusting intertidal spp.). Depth-range is the observed depth of the presence records. Note that the gorgonian coral presence records are a subset of the soft coral records (see Methods).

CWCS group	Presence	Trawl-absences	Depth-range (m)
Glass sponges	1494	3248	28-3368
Demosponges	570	3465	100-2660
Soft corals	1960	4030	18-3624
Stony corals	717	4330	24-1388
Sea pens	3050	3678	8-2458
Black corals	51	4351	81-1985
Gorgonian corals	428	4447	41-3624

Species distribution modelling

Generating targeted absence data and depth masking

Individual MaxEnt species distribution models were developed for the seven CWCS groups. MaxEnt is often used in study areas where data are limited to only species presence data. However, the predictive power of SDMs is increased with the knowledge of absence locations (Phillips *et al.*, 2009).

Therefore, targeted absence data were generated by taking advantage of the DFO synoptic research trawl surveys that occur in the study area (Nottingham *et al.*, 2018) rather than use the random background sampling method that is the default setting in MaxEnt. Since 2003, DFO has followed a depth-stratified, random sampling trawl design over the area spanning the extent of the presence records in this study (Figure 2 a,b). Trawl surveys identify all species caught in a trawl thus those without the occurrence of any of the model-specific CWCS species codes were considered a targeted absence record (Table 2). DFO synoptic research trawl surveys have been fairly reliable when used to generate absence data in SDMs that require presence-absence species data (e.g. Random Forest, Beazley et al., 2018).

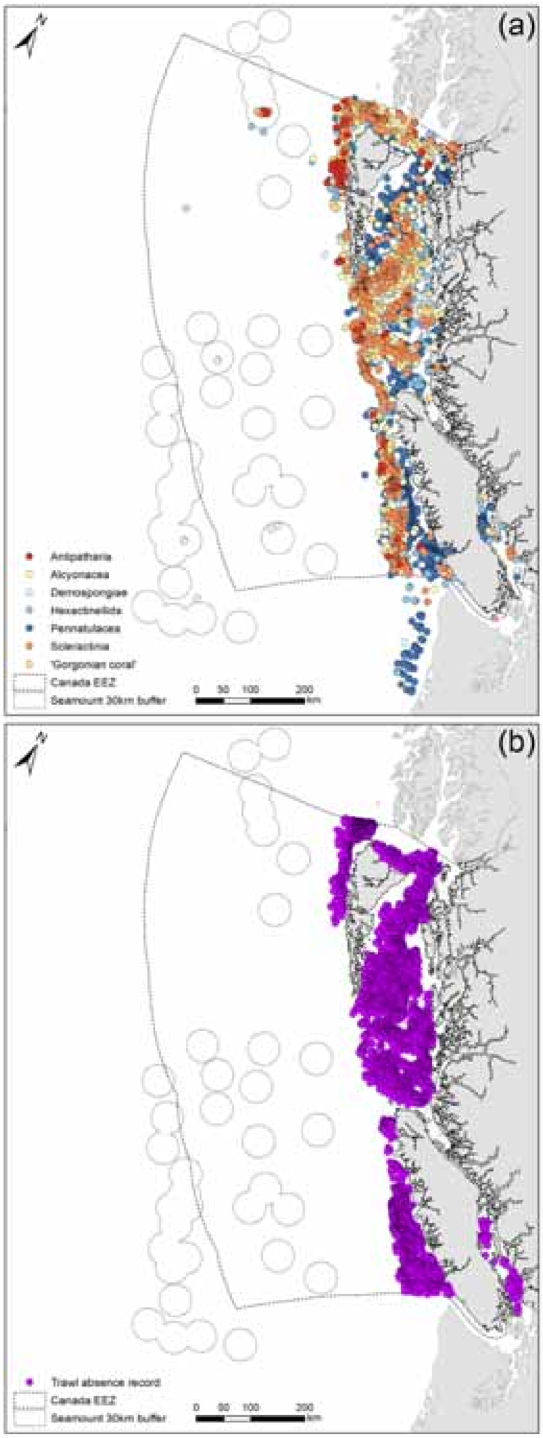


Figure 2. Biogenic habitat-forming cold-water coral and sponge (CWCS) records used in this study. (a) Presence records for six CWCS groups and the gorgonian coral subgroup. (b) Trawl absence records used in this study. Location of named seamounts are delineated by a 30 km buffer zone around their respective summits.

The combined CWCS presence records occupied a depth range of 8-3600 m which reflects the depths of potential fishing activity in the NEPC. Seamount species are characteristically distributed within a preferred depth range (Clark *et al.*, 2010). Therefore, the maximum modelled depth was limited to the depth range occupied by the presence records of each CWCS group (Table 2). MaxEnt (v.3.3.3e, Phillips et al. 2010) models were run with the default parameters for the convergent threshold (1025), maximum iteration value (500), and regularization multiplier (1) using the R package 'dismo' (Hijmans *et al.*, 2017)

Environmental variable selection

Although MaxEnt is reasonably robust to multicollinearity among environmental variables used as predictors of species distributions (Elith *et al.*, 2011), best SDM practices should always reduce the degree of correlation among model predictors. Using the R package 'usdm' (Naimi *et al.*, 2014), variance inflation factors (VIFs) were examined among environmental variables. Values of the environmental variables were extracted at the location of the species data and VIFs were calculated starting with the complete set of environmental data layers. Variables with the highest VIF were iteratively removed until the final set of variables all had VIF values < 10. Ecophysiologically relevant variables were preferentially retained (e.g. silicic acid for sponge groups, carbonate chemistry variables for corals) rather than variables with no direct influence on organism physiology (e.g. depth). This process of variable reduction was done separately using the species records belonging to each CWCS group. The final set of environmental variables used in each CWCS model is summarized in Table 3. All models shared 18 variables: east-facing aspect, north-facing aspect, cross-sectional curvature, longitudinal curvature, slope, roughness, dissolved oxygen, regional current velocity, vertical current velocity, current direction, current direction relative to aspect, 3D current-surface angle, chlorophyll-a, photosynthetically active radiation, sea surface temperature, and topographic position index at three scales (1000 m, 5000 m, 20000 m). The normalized relative importance for each variable used in their respective final CWCS models was assessed using a jack-knife procedure that compared models with and without the variable and the corresponding decrease in the area under the receiver operating characteristic curve (AUC, Phillips, 2005).

Table 3. Final sets of environmental data layers used in models. Full name and units for each variable are summarized in Table 1.

Variable	Glass sponges	Demo-sponges	Soft corals	Stony corals	Black corals	Sea pens	Gorgonian corals
eastness	X	X	X	X	X	X	X
northness	X	X	X	X	X	X	X
crosscurve	X	X	X	X	X	X	X
longcurve	X	X	X	X	X	X	X
slope	X	X	X	X	X	X	X
VRM	X	X	X	X	X	X	X
oxygen	X	X	X	X	X	X	X
regfl	X	X	X	X	X	X	X
vertfl	X	X	X	X	X	X	X
curdir	X	X	X	X	X	X	X
curaspect	X	X	X	X	X	X	X
curang	X	X	X	X	X	X	X
chl-a	X	X	X	X	X	X	X
PAR	X	X	X	X	X	X	X
SST	X	X	X	X	X	X	X
BPI1000	X	X	X	X	X	X	X
BPI5000	X	X	X	X	X	X	X
BPI20000	X	X	X	X	X	X	X
BPI10000	X	X	X	X	X	X	
dSi	X	X		X	X		
arag		X		X	X		
calc			X			X	X

Model validation

Model performance was assessed using five-fold cross-validation where occurrence data (presence and absences) were randomly sampled to create five equal data partitions that follow the same data ratio of presence-absence. Models were trained on four folds and tested with the remaining fold. Each iteration of this procedure (n=5) rotated through the partitions always using a unique partition of records as the testing data. Mean and standard deviation of AUC, percentage correctly classified (PCC), correctly predicted presence (sensitivity), correctly predicted absence (specificity), and kappa was calculated to assess general model performance. Model thresholds were calculated by maximizing the sum and sensitivity using the R package 'PresenceAbsence' (Freeman & Moisen, 2008).

Model predictions of CWCS habitat suitability

A bootstrap resampling procedure (n=200 iterations) was applied to each CWCS model to generate predictions of habitat suitability and spatially explicit measurements of uncertainty associated with the predictions at each raster cell (following Anderson et al., 2016; Rowden et al., 2017). Occurrence data and associated environmental predictors at those locations were randomly sampled with replacement to match the data ratio of presence and trawl-absences in Table 2. MaxEnt models were fit to each iteration, and logistic predictions of habitat suitability (0-1) were generated with values close to one indicating more suitable habitat. Mean and standard deviation (SD) was calculated from the 200 predictions; we use SD to quantify the uncertainty of the model predictions (i.e. predictions are more variable in areas with high SD).

In addition to generating predictions for each CWCS group, a 'CWCS composite index' was created that combined mean predictions among the four coral orders and two sponge class models. The gorgonian model outputs were excluded from the CWCS composite index because the species data were a subset of the Alyconacea model. Rasters of predicted habitat suitability for each CWCS group were reclassified into binary presence-absence layers using model-specific threshold (average of the five-folds) and then combined into a single, composite index (CI) layer where CI values of six would indicate suitable habitat for all six biogenic habitat-forming CWCS groups.

Results and Application

Based on the multiple evaluation metrics, CWCS MaxEnt models performed reasonably well (e.g. AUC values ranged from 0.78-0.91, Table 4). *A priori* use of expert knowledge to parse the records was particularly effective as the Demospongiae model performance improved without shallow records (<100 m) when compared to a model that included shallow records (e.g. AUC increased by ~0.6 and sensitivity increased by 0.11). Water column properties were generally the most important predictors for CWCS occurring in the study area (Table 5). All CWCS models shared dissolved oxygen ($[O_2]$) as a top-3 ranked predictor based on relative importance (10-47%). Examination of the marginal response curves for $[O_2]$ indicates the probability of CWCS occurrence is inversely related to $[O_2]$ levels in the study area with maximum probability occurring at the lower $[O_2]$ distribution, or ~0.3 ml L^{-1}. For both sponge groups, the top-ranked predictor was silicic acid (24-29%) with maximum probability occurring at the higher end of the silicic acid distribution or > 100 µmol L^{-1}. All CWCS models predict areas of high habitat suitability (HSI mean ~1) to occur in the NEPC although to varying extents (Figure 3). Relative extent and locations of prediction uncertainty also varied but was generally low in predicted areas of high habitat suitability (Figure 4).

Table 4. Model performance statistics for cold-water coral and sponge MaxEnt models developed for this study. Values are mean (±1 SD) calculated from five cross-validation folds.

CWCS group	AUC	Threshold	PCC	Sensitivity	Specificity	Kappa
Glass sponges	0.83 (0.01)	0.43 (0.02)	0.76 (0.02)	0.71 (0.05)	0.78 (0.05)	0.46 (0.03)
Demosponges	0.81 (0.02)	0.38 (0.06)	0.71 (0.06)	0.78 (0.11)	0.70 (0.09)	0.29 (0.05)
Soft corals	0.86 (0.01)	0.41 (0.04)	0.79 (0.02)	0.77 (0.06)	0.80 (0.05)	0.54 (0.02)
Stony corals	0.79 (0.01)	0.38 (0.05)	0.70 (0.05)	0.77 (0.08)	0.68 (0.07)	0.27 (0.04)
Black corals	0.91 (0.09)	0.31 (0.12)	0.97 (0.01)	0.84 (0.11)	0.97 (0.01)	0.39 (0.08)
Sea pens	0.78 (0.01)	0.57 (0.02)	0.71 (0.01)	0.70 (0.06)	0.71 (0.05)	0.41 (0.02)
Gorgonian corals	0.85 (0.01)	0.32 (0.06)	0.69 (0.03)	0.89 (0.04)	0.67 (0.04)	0.22 (0.03)

The CWCS composite index resolved extensive areas of suitable habitat on the continental shelf and slope, between 500 and 1400 m bottom depths, for all six biogenic habitat-forming CWCS groups (index score = 6, Figure 5). Smaller isolated patches with a composite index = 6 also occur at five seamounts (Bowie, Hodgkins, Oshawa, Dellwood, and Union) within the Canadian EEZ boundaries. If we combine the total area with a composite index value of ≥ 5, 95% of the areas of potentially diverse biogenic habitat in the NEPC occurs on the continental shelf and slope of the study area (19,568 km^2) with the remainder occurring on seamount and seamount-like features occurring in offshore waters (1,084 km^2).

Table 5. Ranked environmental predictors based on normalized, permuted importance for variables used in cold-water coral and sponge MaxEnt models. Full variable names are in Table 1.

Rank	**Glass sponges**		**Demosponges**		**Soft corals**		**Stony corals**		**Black corals**		**Sea pens**		**Gorgonian**
	Variable	%	Variable	%	Variable	%	Variable	%	Variable	%	Variable	%	Variable
1	*dSi*	28.0	*dSi*	29.9	*oxygen*	36.0	*oxygen*	20.6	*oxygen*	39.0	*oxygen*	29.0	*oxygen*
2	*oxygen*	17.4	*chl-a*	14.2	*calcite*	16.4	*SST*	16.2	*chl-a*	14.1	*chl-a*	13.4	*PAR*
3	*PAR*	12.8	*oxygen*	10.2	*SST*	15.7	*aragonite*	15.3	*VRM*	11.5	*SST*	11.9	*SST*
4	*regfl*	11.6	*northness*	6.6	*PAR*	11.8	*PAR*	13.4	*calcite*	8.7	*regfl*	10.0	*bpi20000*
5	*slope*	4.9	*slope*	6.5	*chl-a*	9.6	*dSi*	9.2	*SST*	6.7	*bpi20000*	8.3	*eastness*
6	*northness*	4.8	*curdir*	6.2	*VRM*	3.1	*vertfl*	3.3	*eastness*	3.4	*slope*	5.8	*curdir*
7	*chl-a*	4.7	*curaspect*	5.6	*bpi20000*	2.1	*bpi5000*	3.2	*northness*	3.3	*calcite*	5.1	*regfl*
8	*VRM*	3.9	*PAR*	4.3	*eastness*	1.3	*curdir*	3.0	*dSi*	3.0	*PAR*	3.7	*northness*
9	*curdir*	3.7	*SST*	3.0	*northness*	1.1	*northness*	2.5	*slope*	2.4	*curaspect*	2.6	*chl-a*
10	*vertfl*	2.0	*bpi20000*	2.9	*regfl*	1.0	*slope*	2.3	*Curapsect*	1.8	*northness*	2.5	*VRM*
11	*curaspect*	1.6	*eastness*	2.6	*slope*	0.6	*regfl*	2.1	*bpi20000*	1.5	*vertfl*	2.4	*calcite*
12	*eastness*	1.3	*regfl*	2.3	*curaspect*	0.5	*eastness*	1.8	*curdir*	1.5	*eastness*	2.3	*slope*
13	*SST*	1.3	*bpi10000*	1.9	*vertfl*	0.3	*bpi10000*	1.6	*longcurve*	1.0	*curdir*	2.2	*vertfl*
14	*bpi20000*	1.2	*VRM*	1.3	*curdir*	0.3	*curaspect*	1.5	*regfl*	1.0	*bpi10000*	0.4	*curaspect*
15	*bpi10000*	0.7	*vertfl*	0.9	*longcurve*	0.04	*chl-a*	1.2	*salinity*	0.4	*bpi5000*	0.1	*bpi5000*
16	*bpi5000*	0.2	*arag*	0.7	*bpi10000*	0.03	*bpi20000*	1.0	*bpi10000*	0.3	*curang*	0.08	*bpi1000*
17	*crosscurve*	0.04	*bpi5000*	0.4	*crosscurve*	0.01	*VRM*	0.9	*PAR*	0.2	*VRM*	0	*crosscurve*
18	*curang*	0.02	*longcurve*	0.4	*bpi5000*	0.01	*longcurve*	0.5	*crosscurve*	0.1	*bpi1000*	0	*curang*
19	*longcurve*	0	*curang*	0.1	*curang*	0.001	*curang*	0.4	*vertfl*	0.1	*longcurve*	0	*longcurve*
20	*bpi1000*	0	*crosscurve*	0.0	*bpi1000*	0	*bpi1000*	0.2	*curang*	0.0	*crosscurve*	0	

				5						2			
21			*bpi1000*	0.04			*crosscurve*	0.07	*bpi5000*	0.02			
22									*bpi1000*	0			

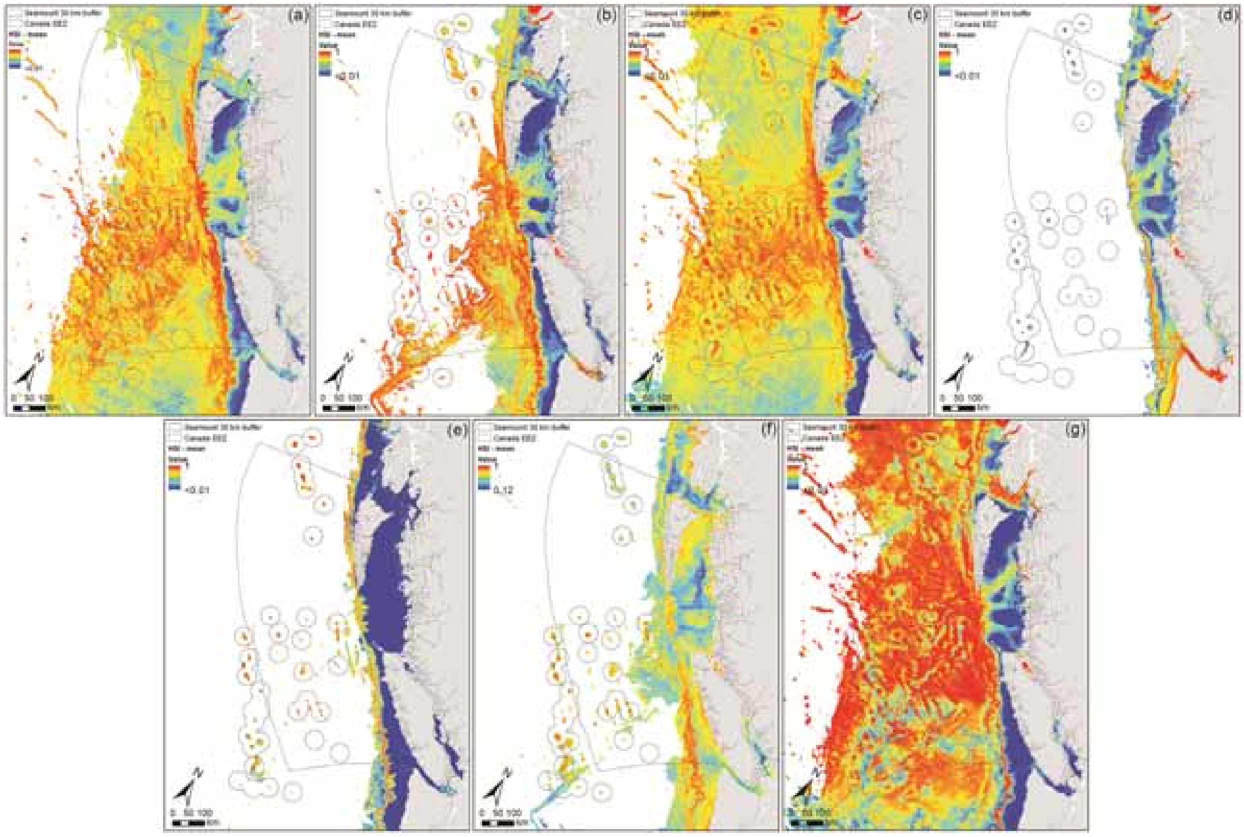


Figure 3. Means of predictions of habitat suitability index (HSI) for (a) Hexactinellida sponges, (b) Demospongiae sponges, (c) Alcyonacea corals, (d) Scleractinia corals, (e) Antipatharia corals, (f) Pennatulacea corals, and (g) Gorgonian corals.

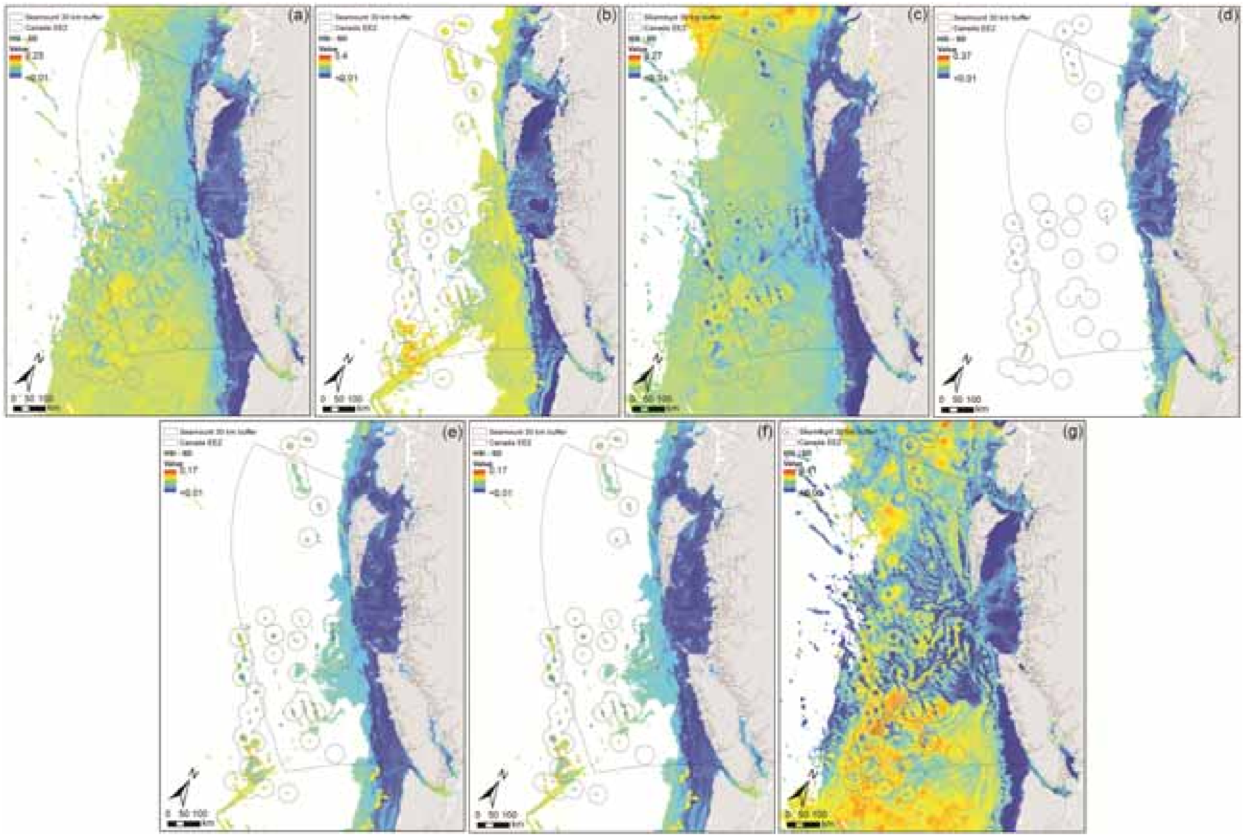


Figure 4. Standard deviation of predictions of habitat suitability index (HSI) for (a) Hexactinellida sponges, (b) Demospongiae sponges, (c) Alcyonacea corals, (d) Scleractinia corals, (e) Antipatharia corals, (f) Pennatulacea corals, and (g) Gorgonian corals.

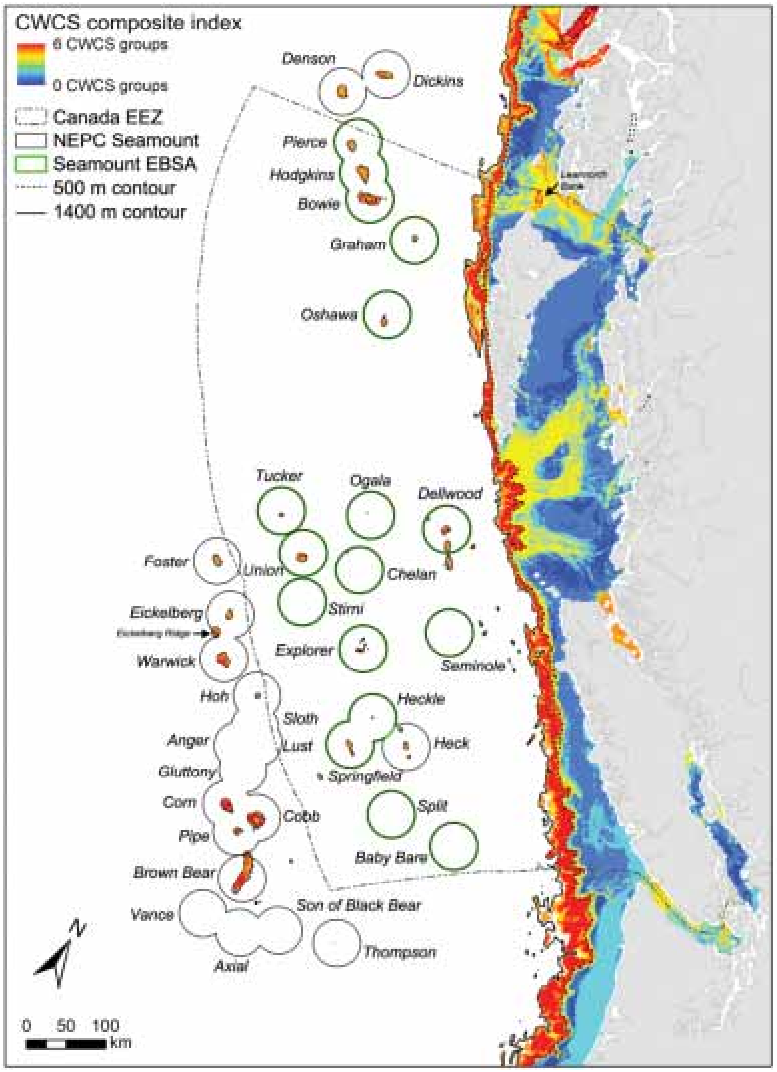


Figure 5. Composite index of areas of suitable habitat for multiple groups of biogenic habitat-forming cold-water corals and sponges. Colour indicates areas of suitable habitat for multiple CWCS groups modelled in this study. Circles outline a 30 km radius buffer around the summits of named seamounts in the NEPC. Green highlighted circles are seamounts identified by Ban et al. (2016) as ecologically and biologically significant areas. Contours delineating 500 m and 1400 m bottom depth which covers the majority of the area that is highly suitable habitat for a diversity of habitat-forming CWCS.

Discussion and conclusions

This NEPC case study illustrates the value of species distribution models for assessing potentially important environmental variables that could influence the distributions of biogenic habitat-forming

CWCS and generating habitat suitability predictions in areas with historically few observations (e.g. offshore deep waters). Although there is variation in the level of importance among individual predictor variables included in each CWCS model, severely low $[O_2]$ ($[O_2] < 0.5$ ml l^{-1}) is an important predictor of habitat suitability for all CWCS groups in the NEPC (Chu *et al.*, 2019). Preliminary assessment of model predictions has validated the model prediction of severely low $[O_2]$ being a strong predictor of biogenic habitat-forming CWCS in the NEPC. Using underwater vehicles to perform post-hoc, visual surveys guided by the model predictions; dense communities of CWCS were discovered at Union and Dellwood seamounts in $[O_2] < 0.2$ ml L^{-1} (Chu et al. 2019). By combining multi-model outputs into a single CWCS composite index to represent biogenic habitat diversity, areas of suitable habitat for multiple CWCS can also be used to empirically assess areas that have been provisionally identified as candidates of conservation interest (e.g. ecologically and biologically significant areas).

In addition to the importance of validating SDMs using post-hoc data collection, it is important to outline considerations when interpreting SDM outputs and how to apply them to facilitate future research directions. While oceanographic characteristics appeared to be the most important predictors in assessing CWCS habitat suitability, interpretations of results need to be constrained to scale, extent, and focal taxa of the study. Information on substratum type (e.g. most CWCS require hard substratum) can influence CWCS recruitment and are often important predictors of CWCS distributions (Krigsman *et al.*, 2012; Masuda & Stone, 2015). However, these data were unavailable at the resolution and extent of the study area. Because most of the historical CWCS records in the NEPC have low-taxonomic resolution, broad-scale models developed for this region require the available species data to be pooled into higher taxonomic groups. Although the regional models developed in this study still performed reasonably well, higher taxonomic resolution should generally improve model performance as this would account for species-specific niche differences which may spatially manifest in studies that focus on smaller scales.

In addition to improving overall data quality, applying different modelling approaches could also improve the degree of confidence ascribed to outputs and decrease the uncertainty associated with predictions generated using SDMs. Several different SDMs approaches exist; aspects that differentiate models in their applied use include species input data requirements (e.g. presence-only, presence-absence, abundance) and model-specific assumptions and uncertainty. Therefore, ensemble modelling could be a future step that assesses model-specific uncertainty (Araújo & New, 2007) by 'averaging' uncertainty among models similar to the forecasting approaches used in climatology (Rooper *et al.*, 2017). An additional modelling consideration could be to incorporate interspecific relationships into model assumptions (e.g. Joint SDMs, Harris, 2015), thus modelling community-level habitat patterns while accounting for interactions among co-occurring species.

Because extensive sampling plans in logistically challenging environments are expensive, data will continue to be sparse in the immediate future. Assessing the distributions of biogenic habitat-forming CWCS is a precursor to understanding their role in the greater dynamics of the entire ecosystem (e.g. biodiversity, ecosystem functioning, fisheries, etc.). In offshore areas where data are scant, data availability will likely remain poor given the remote setting and the complexities of international stakeholder dynamics. SDMs are one tool that can extrapolate modelled species-environment relationships into areas where species records are rare, thus providing an empirical foundation that can promote hypothesis development which can, in turn, concentrate limited science resources into targeted data collection in logistically challenging environments (e.g. Chu *et al.*, 2019).

References

Anderson OF, Guinotte JM, Rowden AA, Tracey DM, Mackay KA, Clark MR (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.

Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.

Ardron JA, Clark MR, Penney AJ et al. (2014) A systematic approach towards the identification and protection of vulnerable marine ecosystems. *Marine Policy*, **49**, 146–154.

Baillon S, Hamel JF, Wareham VE, Mercier A (2012) Deep cold-water corals as nurseries for fish larvae. *Frontiers in Ecology and the Environment*, **10**, 351–356.

Ban S, Curtis JMR, St. Germain C, Perry RI, Therriault TW (2016) Identification of Ecologically and Biologically Significant Areas (EBSAs) in Canada' s Offshore Pacific Bioregion. *Canadian Science Advisory Secretariat*, **2016/34**, 1–152.

Beazley L, Wang Z, Kenchington E et al. (2018) Predicted distribution of the glass sponge Vazella pourtalesi on the Scotian Shelf and its persistence in the face of climatic variability. *PLoS ONE*, **13**, e0205505.

Becker J, Sandwell D, Smith W et al. (2009) Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution: SRTM30. *Marine Geodesy*, **32**, 355–371.

Carton A, Giese B (2008) A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Monthly Weather Review*, **136**, 2999–3017.

Chu JWF, Leys SP (2010) High resolution mapping of community structure in three glass sponge reefs (Porifera, Hexactinellida). *Marine Ecology Progress Series*, **417**, 97–113.

Chu JWF, Maldonado M, Yahel G, Leys SP (2011) Glass sponge reefs as a silicon sink. *Marine Ecology Progress Series*, **441**, 1–14.

Chu JWF, Nephin J, Georgian S, Knudby A, Rooper C, Gale KSP (2019) Modelling the environmental niche space and distributions of cold-water corals and sponges in the Canadian northeast Pacific Ocean. *Deep-Sea Research Part I*.

Clark MR, Rowden AA, Schlacher T et al. (2010) The Ecology of Seamounts: Structure, Function, and Human Impacts. *Annual Review of Marine Science*, **2**, 253–278.

Clark MR, Watling L, Rowden AA, Guinotte JM, Smith CR (2011) A global seamount classification to aid the scientific design of marine protected area networks. *Ocean and Coastal Management*, **54**, 19–36.

Davies AJ, Guinotte JM (2011) Global habitat suitability for framework-forming cold-water corals. *PLoS ONE*, **6**, e18483.

DFO (2019) Biophysical and Ecological Overview of the Offshore Pacific Area of Interest (AOI). *Canadian Science Advisory Secretariat*, **2019.011**.

Dunn DC, Ardron J, Bax N et al. (2014) The Convention on Biological Diversity's Ecologically or Biologically Significant Areas: Origins, development, and current status. *Marine Policy*, **49**, 137–145.

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Freeman EA, Moisen G (2008) PresenceAbsence: An R package for presence absence analysis. *Journal of Statistical Software*, **23**, 1–31.

Garcia HLR, Boyer T, Antonov J, Baranova O, Zweng M, Reagan J, Johnson D (2014a) World Ocean Atlas 2013, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. In: *NOAA*

Atlas NESDIS 75 (eds Levitus S, Mishonov A), pp. 1–25.

Garcia H, Locarnini R, Boyer T et al. (2014b) World Ocean Atlas 2013, Volume 4: Dissolved Inorganic Nutrients (phosphate, nitrate, silicate). In: *NOAA Atlas NESDIS 76* (eds Levitus S, Mishonov A), p. 25.

Gasbarro R, Wan D, Tunnicliffe V (2018) Composition and functional diversity of macrofaunal assemblages on vertical walls of a deep northeast Pacific fjord. *Marine Ecology Progress Series*, **597**, 47–64.

Genin A, Dayton PK, Lonsdale PF, Spiess FN (1986) Corals on seamount peaks provide evidence of current acceleration over deep-sea topography. *Nature*, **322**, 59–61.

Georgian SE, Anderson OF, Rowden AA (2019) Ensemble habitat suitability modeling of vulnerable marine ecosystem indicator taxa to inform deep-sea fisheries management in the South Pacific Ocean. *Fisheries Research*, **211**, 256–274.

Guinotte JM, Davies AJ (2014) Predicted deep-sea coral habitat suitability for the U.S. West Coast. *PLoS ONE*, **9**, e93918.

Harris DJ (2015) Generating realistic assemblages with a joint species distribution model. *Methods in Ecology and Evolution*, **6**, 465–473.

Henry LA, Roberts JM (2007) Biodiversity and ecological composition of macrobenthos on cold-water coral mounds and adjacent off-mound habitat in the bathyal Porcupine Seabight, NE Atlantic. *Deep-Sea Research Part I: Oceanographic Research Papers*, **54**, 654–672.

Henry L, Moreno J, Hennige SJ, Wicks LC, Vad J, Roberts JM (2013) Cold-water coral reef habitats benefit recreationally valuable sharks. *Biological Conservation*, **161**, 67–70.

Hijmans RJ, Phillips S, Leathwick J, Elith J, Hijmans MRJ (2017) Package “dismo.” *Circles*, **9**.

Horn B (1981) Hill shading and the reflectance map. *Proceedings of the IEEE*, **69**, 14–47.

Jenness JS (2004) Calculating landscape surface area from digital elevation models. *Wildlife Society Bulletin*, **32**, 829–839.

Jenness JS (2013a) DEM Surface Tools for ArcGIS. *DEM Surface Tools*.

Jenness JS (2013b) Land Facet Corridor Designer.

Jochum KP, Wang X, Vennemann TW, Sinha B, Müller WEG (2012) Siliceous deep-sea sponge Monorhaphis chuni: A potential paleoclimate archive in ancient animals. *Chemical Geology*, **300**, 143–151.

Jones KH (1998) A comparison of algorithms used to compute hill slope as a property of the DEM. *Computers & Geosciences*, **24**, 315–323.

Kahn AS, Yahel G, Chu JWF, Tunnicliffe V, Leys SP (2015) Benthic grazing and carbon sequestration by deep-water glass sponge reefs. *Limnology and Oceanography*, **60**, 78–88.

Krautter M, Conway KW, Barrie JV, Neuweiler M (2001) Discovery of a “Living Dinosaur”: Globally unique modern hexactinellid sponge reefs off British Columbia, Canada. *Facies*, **44**, 265–282.

Krigsman LM, Yoklavich MM, Dick EJ, Cochrane GR (2012) Models and maps: predicting the distribution of corals and other benthic macro-invertebrates in shelf habitats. *Ecosphere*, **3**, 1–16.

Leys SP, Wilson K, Holeton C, Reiswig HM, Austin WC, Tunnicliffe V (2004) Patterns of glass sponge(Porifera, Hexactinellida) distribution in coastal waters of British Columbia. *Marine Ecology Progress Series*, **283**, 133–149.

Leys SP, Yahel G, Reidenbach MA, Tunnicliffe V, Shavit U, Henry M, Reiswig HM (2011) PLoS ONE The sponge pump : the role of current induced flow in the design of the sponge body plan. *PLoS ONE*, **6**, e27787.

Locarnini RA, Mishonov AV, Antonov JI et al. (2013) World Ocean Atlas 2013, Volume 1: Temperature. In: *NOAA Atlas NESDIS 73*, pp. 1–40.

Lutz MJ, Caldeira K, Dunbar RB, Behrenfeld MJ (2007) Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. *Journal of Geophysical Research: Oceans*, **112**, C10.

Masuda MM, Stone RP (2015) Bayesian logistic mixed-effects modelling of transect data: relating red tree coral presence to habitat characteristics. *ICES Journal of Marine Science*, **72**, 2674–2683.

Miyamoto M, Kiyota M, Hayashibara T, Nonaka M (2017) Megafaunal composition of cold-water corals and other deep-sea benthos in the southern Emperor Seamounts area , North Pacific Ocean. *Galaxea, Journal of Coral Reef Studies*, **19**, 19–30.

Naimi B, Hamm N, Groen T, Skidmore A, Toxopeus A (2014) Where is positional uncertainty a problem for species distribution modelling. *Ecography*, **37**, 191–203.

NASA Goddard Space Flight Center Ocean Ecology Laboratory OBPG (2014) *Aqua Modis*. Greenbelt, MD, USA.

Nottingham MK, Williams DC, Wyeth MR, Olsen N (2018) Summary of the West Coast Haida Gwaii Synoptic Bottom Trawl Survey, August 24 - September 19, 2012. *Canadian Manuscript Report of Fisheries and Aquatic Sciences*, **3133**, 1–55.

Oevelen D van, Duineveld G, Lavaleye M, Mienis F, Soetaert K, Heip CHR (2009) The cold-water coral community as hotspot of carbon cycling on continental margins: A food-web analysis from Rockall Bank (northeast Atlantic). *Limnology and Oceanography*, **54**, 1829–1844.

Phillips SJ (2005) A brief tutorial on Maxent. *AT&T Research*.

Phillips SJ, Dudik M, Elith J et al. (2009) Sample Selection Bias and Presence-Only Distribution Models : Implications for Background and Pseudo-Absence Data. *Ecological Applications*, **19**, 181–197.

Roark EB, Guilderson TP, Dunbar RB, Fallon SJ, Mucciarone DA (2009) Extreme longevity in proteinaceous deep-sea corals. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 5204–5208.

Rooper CN, Zimmermann M, Prescott MM, Hermann AJ (2014) Predictive models of coral and sponge distribution, abundance and diversity in bottom trawl surveys of the Aleutian Islands, Alaska. *Marine Ecology Progress Series*, **503**, 157–176.

Rooper CN, Zimmermann M, Prescott MM (2017) Deep-Sea Research Part I Comparison of modeling methods to predict the spatial distribution of deep- sea coral and sponge in the Gulf of Alaska. *Deep-Sea Research Part I*, **126**, 148–161.

Rowden AA, Anderson OF, Georgian SE, Bowden DA, Clark MR, 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**.

Sandwell DT, Müller RD, Smith HF, Garcia E, Francis R (2014) New global marine gravity model from CryoSat-2 and Jason-1 reveals buried tectonic structure. *Science*, **346**, 65–67.

Sappington JM, Longshore KM, Thompson DB (2007) Quantifying Landscape Ruggedness for Animal Habitat Analysis: A Case Study Using Bighorn Sheep in the Mojave Desert. *The Journal of Wildlife Management*, **71**, 1419–1426.

Steinacher M, Joos F, Frolicher TL, Plattner CK, Doney SC (2009) Imminent ocean acidification in the Arctic projected with the NCAR global coupled carbon cycle-climate model. *Biogeosciences*, **6**, 515–533.

Wheeler E (2018) Royal BC Museum - Invertebrates Collection. v1.1.

Worms Editorial Board (2018) World Register of Marine Species.

Yesson C, Clark MR, Taylor ML, Rogers AD (2011) The global distribution of seamounts based on 30 arc seconds bathymetry data. *Deep-Sea Research Part I: Oceanographic Research Papers*, **58**, 442–453.

Yesson C, Taylor ML, Tittensor DP et al. (2012) Global habitat suitability of cold-water octocorals. *Journal of Biogeography*, **39**, 1278–1292.

Zweng MM, Reagan JR, Antonov JI et al. (2013) World Ocean Atlas 2013, Volume 2: Salinity. In: *NOAA Atlas NESDIS* (eds Levitus S, Mishonov A), pp. 1–39.

Predicting Glass Sponge (Porifera, Hexactinellida) Distributions in the North Pacific Ocean and Spatially Quantifying Model Uncertainty

Editor's note: This contribution to WG32's final report was reproduced from: Davidson, Fiona (2019) Predicting Glass Sponge (Porifera, Hexactinellida) Distributions in the North Pacific Ocean and Spatially Quantifying Model Uncertainty (unpublished Master of Science in Geography, University of Ottawa.

Fiona Davidson

Department of Geography, Environment and Geomatics, University of Ottawa, Ottawa, ON,

Canada

Abstract

Predictions of species' ranges from distribution modeling are often used to inform marine management and conservation efforts, but few studies justify the model selected or quantify the uncertainty of the model predictions in a spatial manner. This thesis employs a multi-model, multi-area species distribution modeling (SDM) analysis to develop a higher certainty in the predictions where similarities exist across models and areas. Partial dependence plots and variable importance rankings were shown to be useful in producing further certainty in the results. The modeling indicated that glass sponges (*Hexactinellida*) are most likely to exist within the North Pacific Ocean where alkalinity is greater than 2.2 µmol l^{-1} and dissolved oxygen is lower than 2 ml l^{-1}. Silicate was also found to be an important environmental predictor. All areas, except Hecate Strait, indicated that high glass sponge probability of presence coincided with silicate values of 150 µmol l^{-1} and over, although lower values in Hecate Strait confirmed that sponges can exist in areas with silicate values of as low as 40 µmol l^{-1}. Three methods of showing spatial uncertainty of model predictions were presented: the standard error (SE) of a binomial GLM, the standard deviation (SD) of predictions made from 200 bootstrapped Generalized linear models (GLMs), and the SD of eight commonly used SDM algorithms. Certain areas with few input data or extreme ranges of predictor variables were highlighted by these methods as having high uncertainty. Such areas should be treated cautiously regardless of the overall accuracy of the model as indicated by accuracy metrics (Area Under the receiver operator characteristic Curve (AUC), a threshold-independent model evaluation indicator and True Skill Statistic (TSS), a threshold-dependent measure of model accuracy), and such areas could be targeted for future data collection. The uncertainty metrics produced by the multi-model SE varied from the GLM SE and the bootstrapped GLM. The uncertainty was lowest where models predicted low probability of presence and highest where the models predicted high probability of presence and these predictions differed slightly, indicating high confidence in where the models predicted the sponges would not exist.

Introduction

Biogeographical patterns for benthic marine taxa are poorly understood due to a lack of accessible geospatial information. Knowledge of the spatial distribution of species is a crucial prerequisite for the understanding of ecosystem functioning and processes as well as conservation management (Reiss et al., 2011). Fauna is more difficult to access and monitor in marine environments than in terrestrial environments. Due to the resulting limitation of data on marine taxa, predictive modeling methods are often employed in order to estimate their full distribution from the available data (Guisan et al., 2000; Guisan et al., 2005). Species distribution models (SDMs) used for this purpose, also referred to as habitat suitability models or ecological niche models, estimate the relationship between species' coordinate data and the environment within which they exist (Franklin, 2009; Elith et al., 2011). While terrestrial SDM work is a fairly robust field, marine applications of SDM have been more recent in their developments (Reiss et al., 2011; Robinson et al., 2011). The field of marine SDM has been stimulated by increasingly available large-scale environmental data on ocean biogeochemistry and the need for prediction methods to quantify and estimate changes in species distribution in response to climatic changes (Reiss et al., 2011). However, a systematic review of 236 published papers on marine-based SDMs (Robinson et al., 2017) noted some shortcomings typical in the field. Only 9% of the reviewed studies tested their model results against independent data, which is generally accepted as an unbiased method of assessing model performance, and 94% of the reviewed studies failed to report the amount of uncertainty derived from data deficiencies and model parameters. When model predictions are evaluated against independent data, there is usually no spatial component to the summary statistics or visualization of patterns of uncertainty such as spatial clustering or links with specific predictor variables (Elith et al., 2002b). A popular method of interpreting and calculating prediction uncertainty involves measuring similarities between distribution predictions, where multiple taxa, models, areas and/or spatial resolution are tested and the resultant predictions are compared (Monk et al., 2012; Pennino et al., 2016; Svensson et al., 2013). Besides similarity calculations, measuring and quantifying prediction uncertainty is an underdeveloped aspect of marine SDM work, and, along with testing several SDMs, will be explored in this contribution.

This contribution will focus on hexactinellid sponges in the North Pacific Ocean. The class Hexactinellida (kingdom Animalia, phylum Porifera), consists of between 400-500 species in two subclasses which contain five orders, 17 families, and 118 genera (Reiswig et al., 1983). They are often referred to as glass sponges because their skeletons are composed of spicules of silica. Hexactinellid sponges are sessile, relying on filter-feeding to obtain the macroscopic detritus material they subsist on (Atwater et al., 2001). After hatching, sponge larvae drift in the water column for a limited time before settling on the seafloor as juvenile sessile sponges (Maldonado, 2006). Observations of planktonic larval life in laboratories indicate that most larvae are anchiplanic, which means they generally remain in the water column for minutes to a few days, usually less than two weeks (Maldonado, 2006). Larvae are known to disperse under the influence of hydrodynamic processes that operate at a spatial scale of tens of meters to kilometers, and are not thought to be affected by active substratum selection, which operates at a smaller scale of centimeters to meters (Maldonado, 2006). Little is known about dispersion in hexactinellid sponges, and limited information about species within the hexactinellid class is available in the dataset used for this contribution. Therefore, it must be mentioned that this contribution operates under the assumption that glass sponges of different species react similarly to their environment, due to a lack of more specific data.

When the sponges die, their siliceous skeletons remain and future sponges grow directly on them, forming reefs. Their skeletons have left a fossil record as far back as the Cambrian/Pre-Cambrian, which would make them possibly the earliest living metazoans on earth (Leys, 2003). While they are found in every ocean in the world, they remain a rare taxon with a seemingly specific set of environmental conditions required to thrive. Research on deep-sea reef-forming benthic taxa is crucial as they are important indicators of the health of benthic ecosystems and often increase biodiversity where they are found (Knudby et al., 2013; Beazley et al., 2013). Their high diversity, large biomass, complex physiology and chemistry, and long evolutionary history lend sponges (and their endosymbionts) to play a key role in diverse ecological processes, including but not limited to predation, habitat provision, nutrient cycling, food chains, and bioerosion (Rützler, 2004). Unfortunately, benthic marine taxa are vulnerable to climate change as well as human activities such as fisheries, specifically deep-sea trawling (Rooper et al., 2017). Glass sponge reefs in the waters off the coast of British Columbia have recently been permanently protected from trawling and other human activities through the establishment of marine protected areas (MPAs). Since little is known about the distribution of this taxon outside British Columbia coastal waters, applying SDMs to glass sponges throughout the North Pacific Ocean will shed light on this otherwise difficult-to-research taxon.

The primary goals for this research were threefold:

1) To map hexactinellid (glass) sponge distribution for the entire North Pacific Ocean, as well as several smaller areas, by testing commonly used species distribution modeling methods

2) To assess the model outputs both in terms of the relative importance of different environmental variables in making predictions about glass sponge presence/absence, as well as the specific dependence of glass sponge presence probability on these environmental variables

3) To compare existing methods for mapping prediction uncertainty.

Species Distribution Modeling in Marine Environments

Species distribution modeling allows for the understanding of processes that create habitat distribution patterns and has become increasingly important in the face of threats such as habitat destruction, species invasions, pollution and climate change (Robinson et al., 2011). SDM algorithms require high-quality species presence/absence records as well as high-quality environmental information to infer the macroecological preferences of species (Tyberghein et al., 2011). By transferring SDMs from terrestrial to marine environments, the validity of the model and its predictive performance will be affected by the unique physical properties of marine habitats (Robinson et al., 2011). This is largely due to the fact that marine ecosystems have significantly less permanence than terrestrial ecosystems; for example, a treeline or grasslands may remain stable during a timeline of decades, while ecological and physical conditions in the water are in continual flux (Longhurst, 2007).

Existing Guidance on Model Selection

Model complexity has increased greatly over time from environmental matching (e.g. BIOCLIM, DOMAIN) to more complex non-linear relationships between species and their environment (e.g. generalized additive models (GAMs), MaxEnt) (Elith et al., 2009). BIOCLIM is an early SDM package which relates the bioclimatic environment species exist within to a number of environmental predictor variables, such as temperature or elevation (Booth et al., 2013). Due to the now numerous SDM methods, there is some difficulty in selecting an appropriate algorithm. The advice that would assist making an informed choice of method is currently scattered throughout literature (Elith et al., 2009). It remains difficult to know which model is 'best' for the given data before comparing multiple models. This contribution therefore focuses on several commonly used SDMs.

Input data required for SDM work involves biological data: information about the species (single or multiple species) distribution, and environmental data: usually raster data describing the landscape the species is found within (Pearson, 2010). Biological data can be obtained in numerous ways: from surveys, museum collections, or personal collection and may be *presence-only* (PO, coordinates of where the species has been observed), or *presence/absence* (PA, coordinates of where the species has and has not been observed). Generally models are thought to have more ecological validity when fit with PA data as opposed to PO data, however the quality of absence data is often questioned due to possibility of 'false absences', which refers to instances when a species was present but not detected, or the environment was suitable but the species was absent (Pearson, 2010). Environmental data refers to predictor variables depicting climate, topography, land cover and vegetation, substrate, and other physical and chemical attributes of the area being modeled (Franklin et al., 2010). Spatial scale is often considered when collecting data and has two components: extent and resolution. Spatial extent refers to the size of the area being modeled and spatial resolution refers to the size of grid cells of the data. It is often common for datasets with large extents to have coarse resolution, and small extents to have high resolution (Pearson, 2010). As with other deep-sea species modeling efforts, due to the lack of information available concerning the niche environmental preferences of the relevant taxa it is difficult to ascertain the importance of individual environmental variables prior to modeling. When working with taxa for which there are limited data, environmental input layers are by necessity often selected primarily based on their availability and presumed relevance, and less important variables can be identified and removed during the modeling process.

Statistical Models

The Linear Model

Linear multiple regression models predict the response variable (Y) from a vector of multiple predictor variables, X = (X_1, X_2,…,X_p):

$$\hat{Y} = \hat{\beta}_0 + \sum_{j=1}^{p} X_j \hat{\beta}_j + \varepsilon \qquad \text{(Eq. 1)}$$

where $\hat{\beta}$ is the vector of estimated coefficients and $\hat{\beta}_0$ is an estimated constant known as the intercept (Franklin et al., 2010). The error term, ε , is normally distributed with zero mean and constant variance, and the variance of Y is constant across observations (Franklin et al., 2010).

Generalized Linear Models (GLMs)

While Franklin & Miller (2010) note that ecological data often violate the assumptions of the linear model, GLMs are often used in modeling and can be described as extensions of the linear model that can cope with non-normal distributions of the response variable (Venables et al., 1994). Distributions that are often used to characterize response variables in ecology include Gaussian, Poisson, binomial, negative binomial, and gamma (Franklin et al., 2010).

The linear model can be generalized using a link function that describes how the mean of Y depends on linear predictors, and a variance function that describes how the variance of Y depends on its mean (Chambers et al., 1992). The equation for the GLM can be seen in Equation 2:

$$\delta\big(E(Y)\big) = LP = \hat{\beta}_0 + \sum_{j=1}^{p} X_j \beta_j + \varepsilon \qquad \text{(Eq.2)}$$

where the predictor variables (far right side of the equation) are combined to produce a linear predictor, LP, and the expected value of Y, E(Y), is related to the LP through the link function, δ () (Franklin et al., 2010). Formulating a GLM for SDM involves selecting the response distribution and the link function (collectively known as the family of the GLM), the variance function, and the predictors (Franklin et al., 2010). The link function describes how the mean of Y depends on the linear predictor. For a binary response variable, a binomial distribution and logit link function are used.

Generalized Additive Models (GAMs)

Generalized additive models (GAMs) differ from GLMs in their ability to identify and describe a non-linear relationship between response and predictor variables; they are non-parametric extensions of GLMs (Franklin et al., 2010).

$$\delta\left(E(Y)\right) = LP = \hat{\beta}_0 + \sum_{j=1}^{p} X_j f_j + \varepsilon \qquad \text{(Eq. 3)}$$

where the coefficients of the GLM are replaced by a smoothing function, f (Franklin et al., 2010). The fit of a GAM model is generally evaluated by testing the non-linearity of a predictor versus the non-parametric fit (Franklin et al., 2010). GAMs are used for characterizing non-linear response curves of species because they can suggest the shape of the parametric response curve and are thus more flexible than GLMs (Franklin et al., 2010). GAMs are popular in SDM work because they tend to have high prediction accuracy, have been subjected to comparisons with other models and have proven to be useful (Franklin et al., 2010).

Machine Learning Models:

Maximum Entropy (MaxEnt) Distribution Modeling

The MaxEnt model was created in order to make predictions and inferences from incomplete data (Phillips et al., 2006), for example presence-only data. MaxEnt is one of the most common forms of SDM and “has been described as especially efficient to handle complex interactions between response and predictor variables” (Fourcade et al., 2014). MaxEnt is an acronym created for the concept of maximum entropy modeling (Guinotte et al., 2014), which extrapolates the likelihood a species has of existing in any specific geographic space. This can also be defined as a measure of dispersiveness. The underlying principle is that one should assume uniform distributions are preferred, given certain constraints (Nigam et al., 1999). Since becoming available in 2004, MaxEnt has been used to publish diverse projects including finding correlates of species occurrences, mapping current distributions, and other related tasks in ecological, evolutionary, conservation and biosecurity applications (Elith et al., 2011).

MaxEnt has often been explained as estimating a distribution across geographic space (Phillips et al., 2006). Elith et al. (2011) give an alternative view: a characterization that focuses on comparing probability densities in covariate space. Their research examines how MaxEnt can be understood by looking at Bayes’ rule:

$$\Pr(y = 1|z) = f_1(z)\Pr(y = 1)/f(z) \qquad \text{(Eq. 4)}$$

where y = 1 indicates presence, y = 0 indicates absence, and z indicates a vector of environmental covariates. It must be assumed that all environmental variables z are available landscape-wide, and L is the extent of the landscape (not included in Eq. 4). $f(z)$ can be defined as the probability density of covariates across L, $f_1(z)$ can be defined as the probability density of covariates across locations within L where the species is present, and $f_0(z)$ can be defined as where the species is absent (Elith et al., 2011). The quantity to be estimated is the probability of presence of the species, conditioned on the environment: $\Pr(y = 1|z)$.

Equation 4 can theoretically be explained by the following: that if the conditional density of the covariates at presence sites is known, $\mathbf{f_1(z)}$, and if the unconditional density of covariates across the study area is known, $\mathbf{f(z)}$, the prevalence $\mathbf{Pr(y = 1)}$ is the only remaining value necessary to calculate the probability of occurrence (Ward, 2007; Elith et al., 2011). First, MaxEnt's core output involves estimating the ratio $\mathbf{f_1(z)}$/ $\mathbf{f(z)}$. This gives insight about which features are important and how suitable one place is compared to another, which is the core of the MaxEnt model output. This explanation of MaxEnts' structure by ecologists rather than statisticians can be helpful in understanding the complicated processes that the data undergo.

Boosted Regression Trees (BRT)

Boosted regression trees (BRT) is an ensemble method for fitting statistical models that differs from conventional techniques to fit a single parsimonious model; BRTs combine the strength of two algorithms: regression trees and boosting (Elith et al., 2008). Regression trees are models that relate a response to their predictors by recursive binary splits, and boosting is an adaptive method which combines simple models to give improved prediction performance (Elith et al., 2008).

The decision trees in BRT are tree-based models which partition the predictor space into rectangles, doing this using a series of rules to identify regions having homogeneous responses to predictors (Elith et al., 2008). Then, a constant is fitted to each region, with regression trees fitting the mean response for observations in that region. Fitting a single decision tree is often done by growing a large tree and afterwards pruning it by collapsing the weakest links (identified through cross-validation) (Elith et al., 2008). Decision trees are popular because they allow for information to be represented in an intuitive manner that is easy to visualize. Trees are insensitive to outliers and are able to accommodate missing data in predictor variables by using surrogates (Breiman et al., 1984).

Study Area

The Pacific Ocean, the largest and deepest of the earth's oceans, is about 15 times the size of the United States, and is almost equal in area to the total land area of the world. The ocean can be divided by the equator into two separate areas: the north and south. The study area for this project is contained by the boundaries of the North Pacific Ocean: bordered by the Arctic Ocean in the north, Asia in the west, the Americas in the east, and the equator in the south. It provides habitat for thousands of species, including cold-water sponges and corals. Due to the size of this study area, five sub-areas were delineated within the North Pacific Ocean to account for the likely varying physical and chemical environments across an area as large as the North Pacific Ocean. These sub-regions within the North Pacific Ocean include a) The US Exclusive Economic Zone (EEZ) around Alaska, b) The Canadian EEZ around British Columbia, and c) The US EEZ around the Washington-Oregon-California coast, as well as d) two smaller areas within the Canadian EEZ which were manually delineated but roughly correspond to i) Hecate Strait and ii) the shelf waters west of Vancouver Island (Figs. 1-4). It is likely that there are varying environments within an area as large as the North Pacific Ocean, as well as varying groups of sponges. Using a multi-area analysis ensures a more comprehensive attempt at capturing these potentially different species-environment relations.

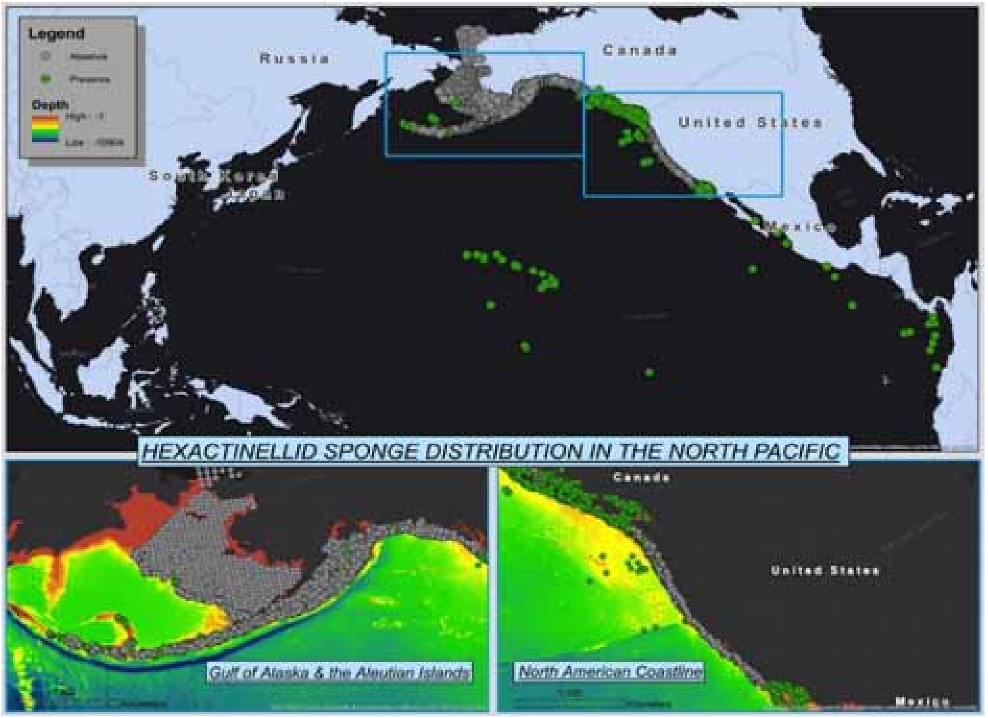


Figure 1. Hexactinellid Sponge Distribution in the North Pacific Ocean with insets of the Gulf of Alaska and British Columbia coastline.

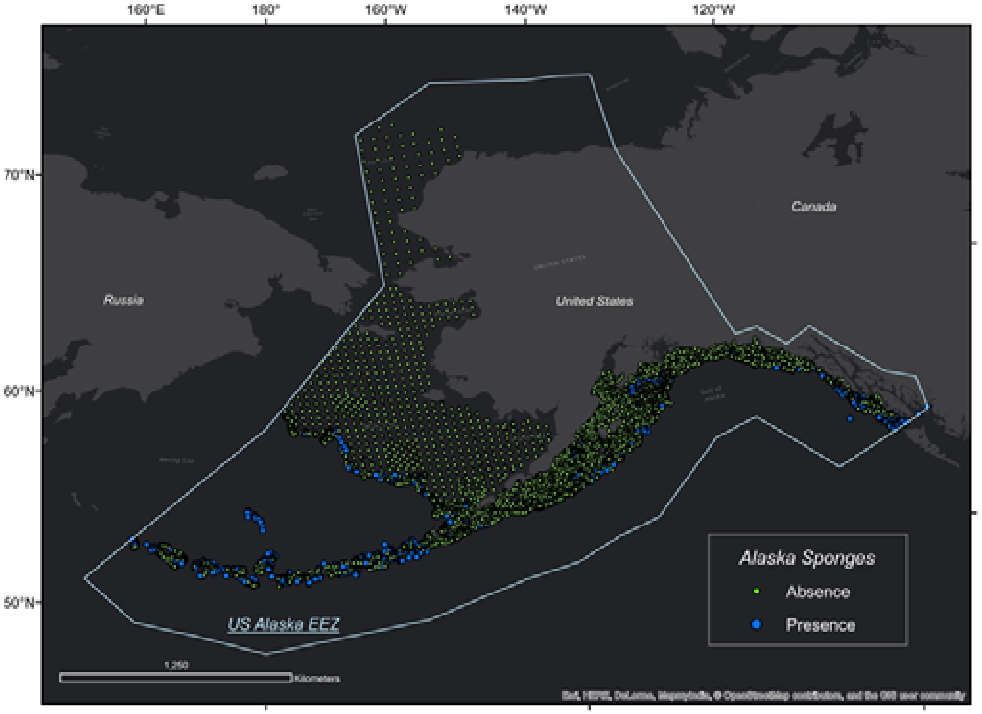


Figure 2. Alaska sub-area with contained sponge presence-absence (PA) data.

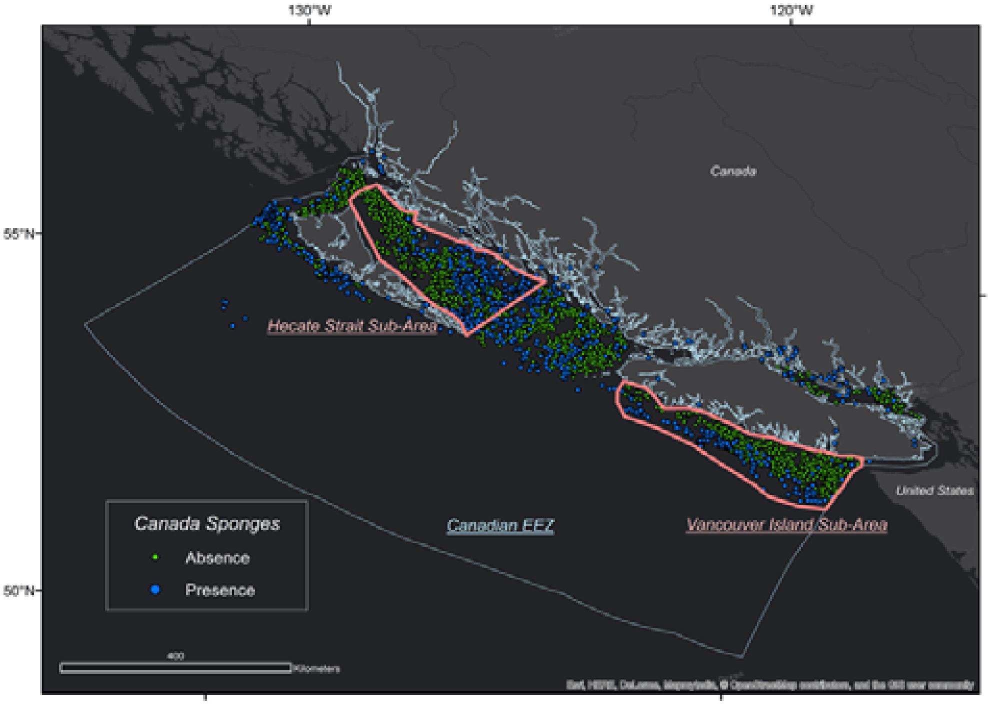


Figure 3. British Columbia, Hecate Strait and Vancouver Island sub-areas with contained sponge PA data.

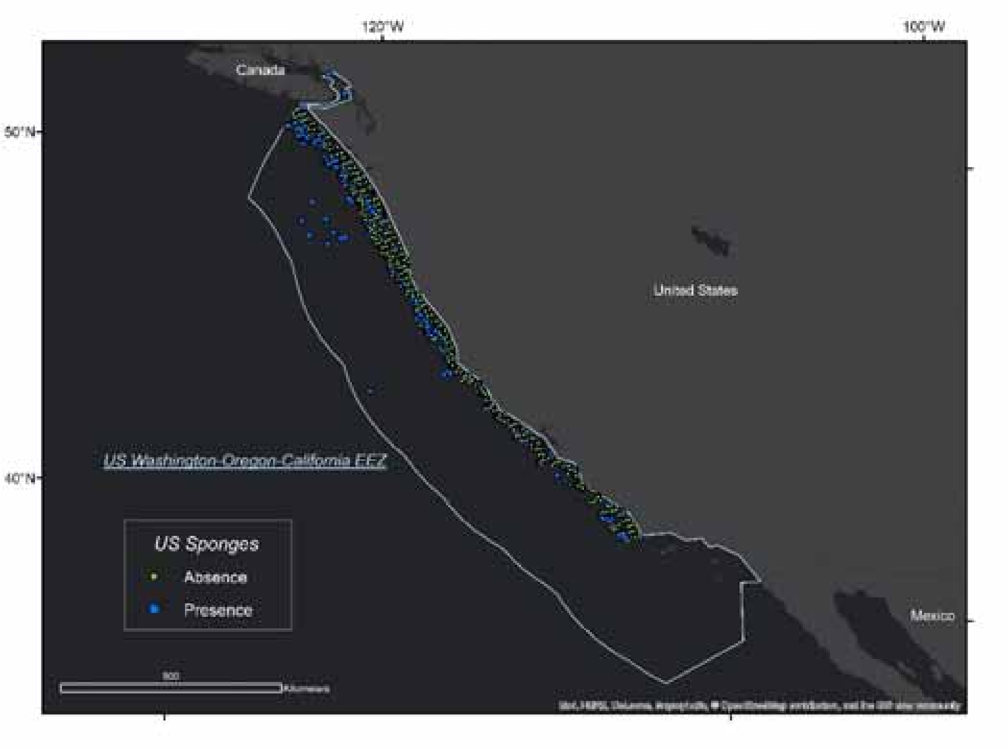


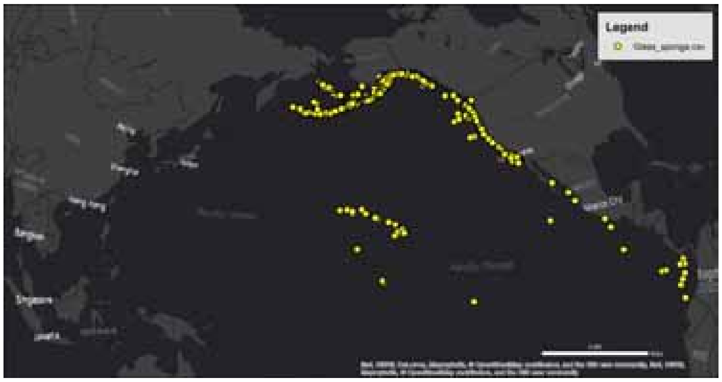
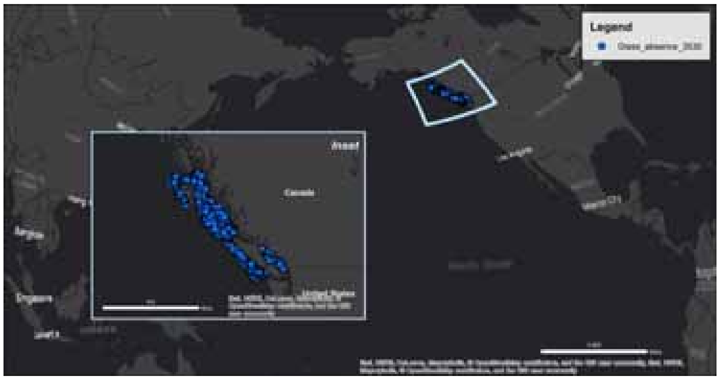
Figure 4. United States Washington-Oregon-California sub-area with contained sponge PA data.

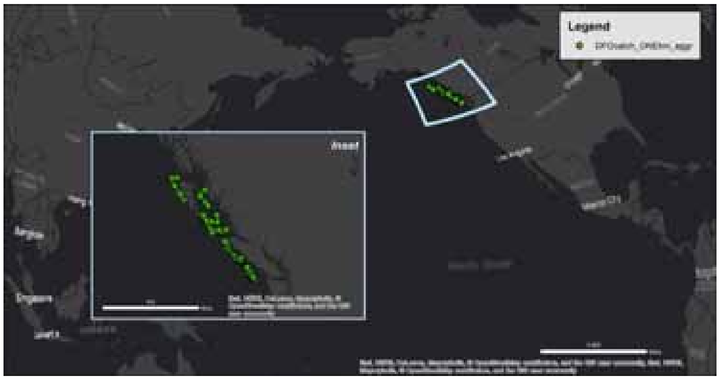
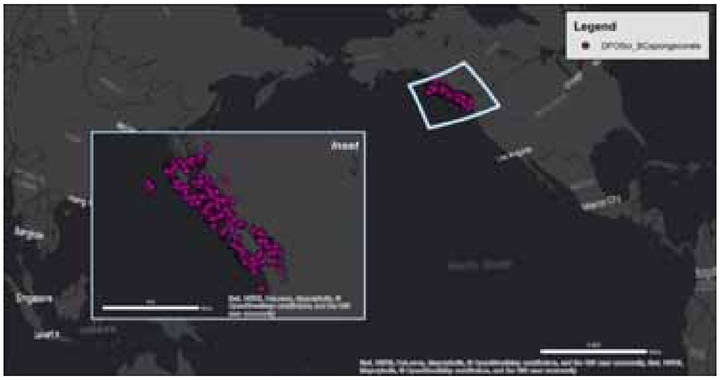
Biological Data

Presence and absence glass sponge data were obtained from trawl surveys conducted by Fisheries and Oceans Canada (DFO) and the US Government. Data from several surveys were collected and merged to create a dataset containing both presence and absence information for 42,113 coordinate locations sampled between 1996 and 2016. The dataset contains 16,148 presence points and 25,684 absence points. As can be seen in Figures 1–4, and Table 1, the species presence/absence points are located largely in coastal waters along the coast of North America, and out along the Aleutian Islands, with a few data points from Hawaii. No data from the western North Pacific Ocean were used.

In an attempt to decrease sample bias, the original dataset of 42,113 coordinate points was thinned based on environmental variation (See Methods section for more detail). After the data was thinned based on local environmental variation of the predictor variables, the resultant dataset had 12,467 sponge PA points.

Table 1. Hexactinellid sponge data: location of data points and source.

Data	Geographic Extent	Number of Presences	Number of Absences
US bottom trawl surveys from Alaska (1996-2016) and US West Coast (1996-2004)		1008	22,322
Presence data from North Pacific Marine Science Organization (PICES) Working Group 32		14,134	0
DFO commercial bycatch logs		0	3530

DFO commercial catch records aggregated to 1km grid		251	0
DFO research databases and museum records		868	0
Total		42,113	

Environmental Data

Environmental variables were selected based on availability and presumed likelihood of being relevant to the distribution of glass sponges. Potential environmental variables relevant to the distribution of glass sponges have been compiled from various sources through the North Pacific Marine Science Organization (PICES) Working Group 32 (Table 2). The environmental data layers are in a raster format, with a cell size of 1000 m by 1000 m, using an azimuthal equidistant projection with a central meridian of -180. The values reflect the near-sea floor part of the water column.

Table 2. Environmental variables, units and reference.

Variable Name	Units	Reference
Alkalinity	µmol l^{-1}	Steinacher et al. (2009)
Aragonite saturation state	Ω_{ARAG}	Steinacher et al. (2009)
Aspect	degrees	Becker et al. (2009)
Calcite saturation state	Ω_{ARAG}	Steinacher et al. (2009)
Depth	m	Becker et al. (2009)
Dissolved inorganic carbon	µmol l^{-1}	Garcia et al. 2014a
Eastness	degrees	Wilson et al. (2007)
Nitrate	µmol l^{-1}	Garcia et al. 2014b
Northness	degrees	Wilson et al. (2007)
Oxygen	ml l^{-1}	Garcia et al. 2014a
Phosphate	µmol l^{-1}	Garcia et al. 2014b
Roughness	unitless	Wilson et al. (2007)
Rugosity	unitless	Becker et al. (2009)
Salinity	pss	Zweng et al. 2013
Silicate	µmol l^{-1}	Garcia et al. 2014b
Slope	degrees	Becker et al. (2009)
Temperature	°C	Locarnini et al. 2013
TPI (Topographic Position Index)	unitless	Wilson et al. (2007)
TRI (Terrain Ruggedness Index)	unitless	Wilson et al. (2007)

Methods

Many of the choices in the following methodology section were made in an effort to maximize reproducibility of this study, however had other methods been selected, the results could have differed. Within the SDM field, reproducibility is a common problem. Advanced modeling techniques, data selection and processing require many choices to be made which decreases the replicability yet is nonetheless common and relatively unavoidable in SDM.

Data Pre-Processing

Spatial sampling bias, a common problem in marine and terrestrial SDM, decreases the accuracy and interpretability of SDM outputs. Spatial filtering is a common method of removing spatial bias as a data pre-processing step. For example, Boria et al. (2014) filtered clustered data to discard any data point within 10 miles of another point, and Varela et al. (2014) applied an environmental filter which discards presence points that are too clustered in environmental space. The following steps were taken to spatially thin the data in a manner which takes into account the differences in areas with high environmental variation and areas with low environmental variation. The principle at the basis of this method is that areas with low environmental variation across space require less geographically dense data to cover environmental variability than do areas with high environmental variation, and they can therefore be thinned more than areas with high variation in an effort to reduce bias in the dataset.

1. The local SD of each predictor variable was calculated for a 9 x 9 km window centered on each cell, and then normalized to a 0-1 scale. The 19 normalized values were then added together to produce a single raster with a theoretical value range of 0-19, quantifying local environmental variation across the study area.
2. A histogram was plotted to view the frequency distribution of this local environmental variation (Fig. 5). If this histogram had been multimodal, spatial areas corresponding to each local maximum, i.e. clusters of low or high local environmental variation, could have been identified. However, the histogram was unimodal, so instead quintiles were calculated to separate the study area into five regions ranging from lowest to highest local environmental variation. The maximum value for each quintile can be seen in Table 3.

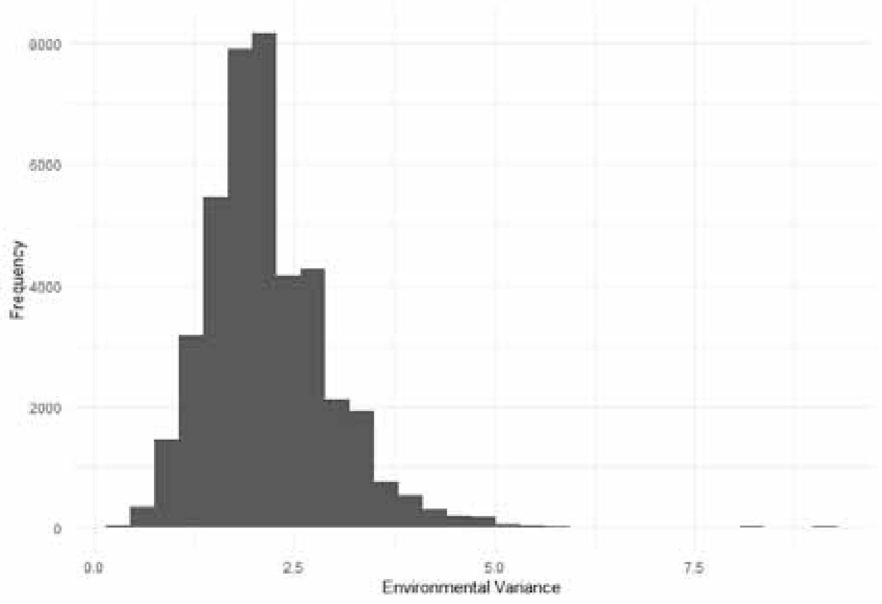


Figure 5. Histogram showing environmental variance data distribution.

Table 3. Quintile breaks for the variance data (environmental variance values from Figure 5).

20%	40%	60%	80%	100%
1.576018	1.893480	2.200223	2.730262	9.180244

3. Five subsets of the presence/absence data were then generated, one for each quintile, and semi-variograms were generated based on the bathymetric values from each subset. Depth was chosen to be the predictor for which to produce semi-variograms for several reasons: 1) depth often is one of the most important variables in any SDM for hexactinellid sponges, and 2) it can be used as a proxy for many other variables in this study.
4. The semi-variograms all used 1000 m bins and a cut-off of 25,000 m to ensure standardization (Fig. 6). For each plot, the distance at which semi-variance increased to more than 500 was noted. For the first quintile (representing the area with lowest local environmental variation) this distance was ~15,000m, for the second quintile it was ~7500m, and for the third, fourth and fifth quintile it was ~2500m. The semi-variance

value of 500 was selected visually to provide a range of reasonable distances that were considered suitable to inform the scale of spatial thinning.

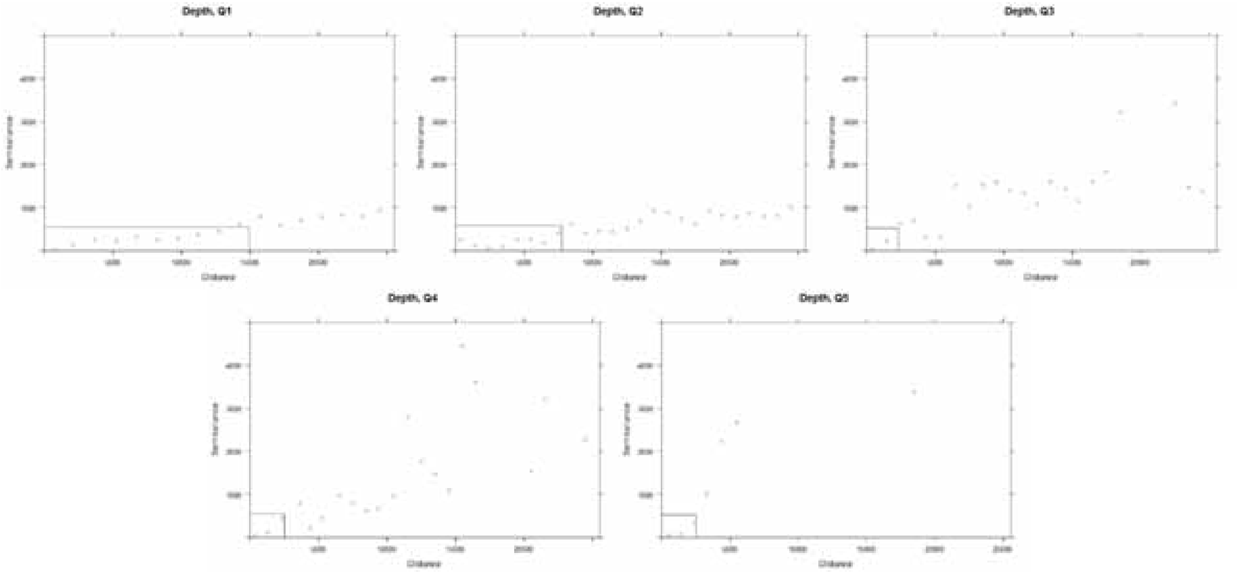


Figure 6. Semi-variogram plots showing five quintiles, with semi-variance value of 500 indicated to show approximate calculation of range value.

5. Based on these semi-variograms, block-based thinning was applied to the presence/absence data:
    i. A grid with 15,000 m cell sizes was overlaid on the study area, and the mean value of the local environmental variation raster was calculated for each cell. For those cells falling in the first quintile, i.e. with mean local environmental variation <1.576 (see Table 3), the presence/absence data were aggregated according to the following rules:
        a) If no presence/absence observations were found in the cell, the output would be empty.

        b) If there were at least as many presence as absence observations in the cell, the output would be a presence point, located in the center of the cell.
        c) If there were more absence than presence observations in the cell, the output would be an absence point, located in the center of the cell.
    ii. This process was repeated with 7500 m cells applied to points in the second quintile, and with 2500 m cells applied to points in the third, fourth and fifth quintiles.
    iii. The outputs were combined to produce a single set of spatially thinned presence/absence observations. Out of the original 42,113 presence and absence points, 12,467 remained after thinning. A sample section from the Aleutian Arc can be seen in Figure 7, showing the difference between the original points and the thinned points.

Species Distribution Modeling

For each of the predefined areas, the following SDM types were tested using the 'Biomod2' (Thuiller et al., 2016) package in R (R Core Team, 2013): GAM, BRT and MaxEnt (resulting in 18 model-area combinations). The following parameters for modeling were used:

- 'Number of Evaluation Runs: 3'; Running three evaluations means the calibration and evaluation is run 3 separate times independently, which allows for a more robust test of the models when independent data is not available.

- 'Data Split: 80%'; this sets 80% of the data aside for calibration of models, with the remaining 20% used for validation.
- Model accuracy measures: KAPPA, TSS, AUC; KAPPA refers to Cohen's Kappa Coefficient, and TSS to True Skill Statistic (Zhang et al., 2015). Both Kappa and TSS are threshold-dependent measures of model accuracy. They range from −1 to +1, where +1 indicates perfect agreement between predictions and observations and values of 0 or less indicate agreement no better than random classification (Landis et al., 1977). The Area Under the receiver operator characteristic Curve (AUC) is an effective, threshold-independent model evaluation indicator and is also independent of prevalence (i.e. the frequency of occurrence) of the target species (Zhang et al., 2015). Ranges used to interpret accuracy metrics from these statistics can be found in Table 4.

Table 4. Model accuracy ranges for AUC, Kappa and TSS measures (Zhang et al., 2015).

	Poor	Moderate	Good/Excellent
AUC	< 0.7	0.7 - 0.9	> 0.9
Kappa and TSS	< 0.4	0.4 – 0.8	> 0.8

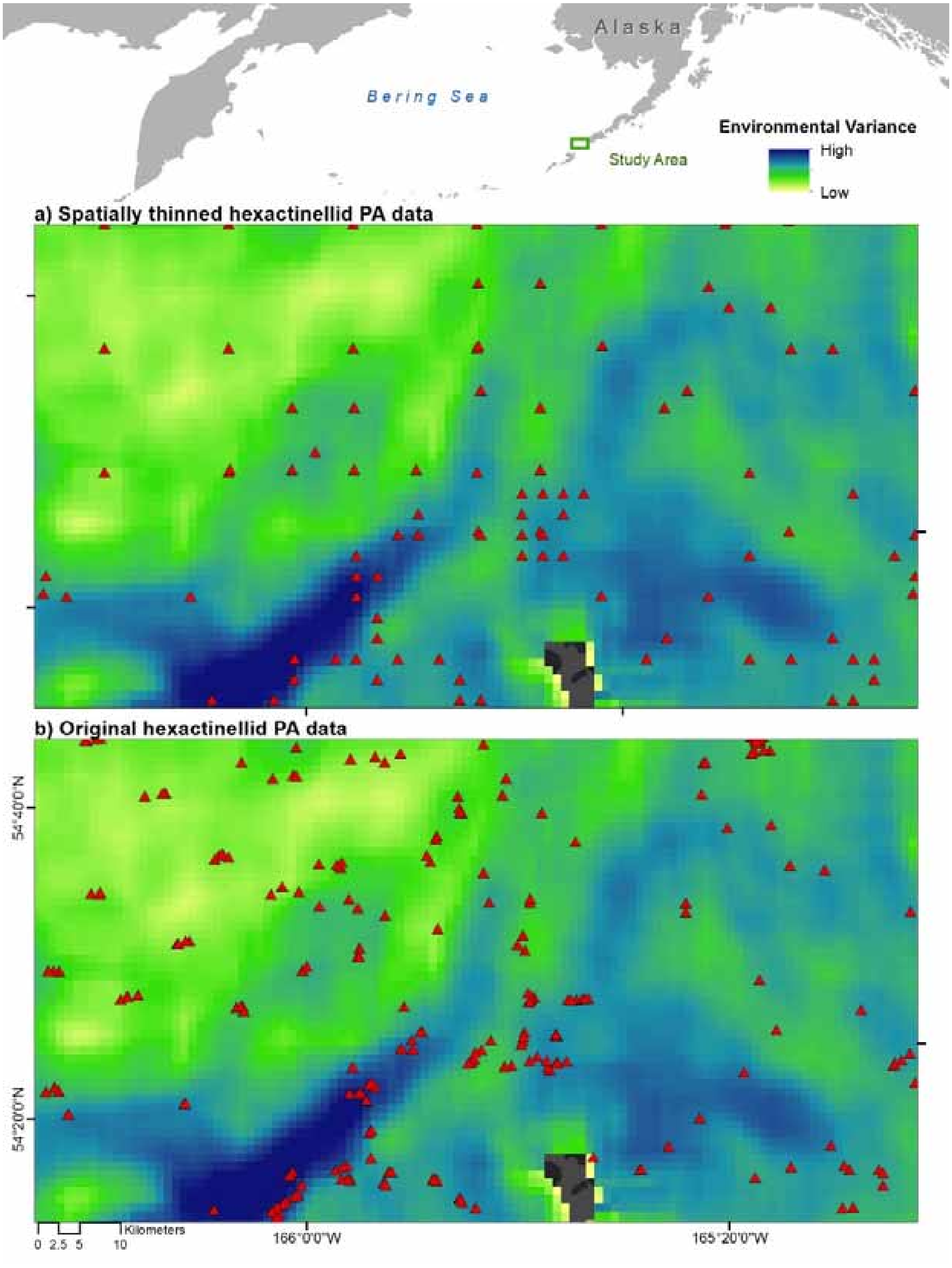


Figure 7. Sample area showing spatially thinned data and original data on the Aleutian Arc.

Interpretation of Model Results

Model output can be classified into two types: aspatial and spatial. Aspatial outputs from models consist of variable importances and partial dependence plots depicting fitted functions relating probability of occurrence to each selected predictor, while spatial outputs consist of GIS layers depicting the probability of presence across the study area (Ferrier et al., 2002).

Variable importance values were calculated for every model-area combination. These values quantify to what extent each predictor variable contributes to the predictions made by the model. The variable importance value is calculated as the result of one minus the correlation between the original prediction and the prediction with only the individual variable of interest. While the individual values are dependent on the algorithm used, they can still be used to provide relative information on predictor importances within the model.

Partial dependence plots provide a graphical representation of how likely the species is to be present, given a gradient of the specific environmental predictor. As with variable importance, these plots are calculated when the model is built, by averaging every other predictor variable except the one chosen predictor, and the change in model response is measured in relation to changes in the one variable. Partial dependence plots showing results from multiple algorithms can be used to visually compare species responses to environmental variable values. Partial dependence plots were generated for predictor variables that ranked in the top 25% in variable importance in several models and areas. The predictor variables chosen for closer analysis were alkalinity, oxygen, silicate, and phosphate. Additionally, frequency distribution plots were produced for alkalinity, oxygen, silicate, and phosphate in each area, showing the percentage of presence (as opposed to absence) data points across the range of environmental variable values. This information can shed light on which value ranges the species most commonly exists within. Partial dependence plots can be compared for similarities. Variable importance was also considered in an effort to find strong trends in how predictors contribute to different models/areas. It can be posited that if the response curves of a particular predictor variable are similar across multiple models/areas, the variable importance is likely to be high as well. When the response curves vary significantly, it is more likely those variables ranked toward the lower range of variable importance for the model.

Mapping Prediction Uncertainty

For the purpose of testing and comparing uncertainty metrics spatially, a binomial GLM was fit to the Hecate Strait boundary (Fig. 3), a subset of the original, North-Pacific-wide dataset. 1,255 presence/absence points were included within the area and three environmental predictors with high variable importance were selected and clipped to the same extent; alkalinity, oxygen and silicate. A GLM was selected because it can provide a model-based uncertainty measure that can be mapped in addition to the actual model predictions. The logit-link function was selected for the binomial GLM because it is appropriate for binary data and ensures the predicted values will be between 0 and 1 (Kindt et al., 2005). Three spatially explicit uncertainty metrics were compared using this GLM model:

GLM Prediction SE

Bootstrapped GLM SD

SD of Multiple Model Predictions

The first uncertainty metric involved producing partial dependence plots from the GLM outputs and adding confidence intervals to the partial dependence plots. In order to obtain the predicted values from the estimates of the coefficients, the inverse link function needs to be calculated. Using the inverse link function, the confidence interval was calculated as the fitted value plus/minus two times the SE on the link scale. Confidence intervals (Cis) were calculated for alkalinity, oxygen and silicate. Adding CIs can provide information on why certain areas would have predictions with high or low confidence. Next (and separate from the CIs), the SE of the prediction was calculated. SE provides the absolute measure of the typical distance between the data points and the regression line, in the units of the dependent variable. The SE of the prediction was then written to a raster and thus the uncertainty of the model can be seen spatially.

The second uncertainty metric – Bootstrapped GLM SD – was obtained by bootstrapping the GLM. Bootstrapping is an approach to statistical inference based on building a sampling distribution for a statistic by resampling repeatedly from the data. 200 bootstrap samples were created from the data. GLMs were then calibrated on the bootstrap samples, still using alkalinity, oxygen and silicate as predictors for the models. The calibrated models were then used to make predictions, and the SD was calculated for the predictions.

The third uncertainty metric – SD of Multiple Model Predictions – aims to test if standard deviations are geographically comparable among a variety of models. By running eight models available in the Biomod2 package (GLM, BRT, GAM, FDA, MARS, RF, MAXENT.Phillips and MAXENT.Tsuruoka) on the Hecate Strait subset, the SD of all the predictions can be calculated, as for the GLM bootstrapping above. This provided a spatial view of where the models produced similar results and where they differed.

Results

Model Performance

Model fit statistics and variable importance values from the GAM, BRT, and MaxEnt models run on the North Pacific basin-wide data, as well as the five sub-areas are presented in Table 5. Figures 8-11 present the outputs of these models in the form of partial dependence plots and show the data distribution for alkalinity, oxygen, phosphate and silicate for each model/area. Only these four variables were selected because they had consistently high variable importance values.

As can be seen in Table 5, AUC values for the majority of the models were between 0.7-0.9. These values are interpreted to indicate these models performed moderately well (See Table 4 for value ranges associated with model accuracy) (Zhang et al., 2015). Two MaxEnt models for the BC and Alaska areas performed poorly, with AUC values of 0.655 and 0.428 respectively, and the GAM and BRT models for the Vancouver sub-area performed especially well, with AUC values of 0.946 and 0.978, the highest of all the models and areas. The Kappa and TSS values reported similar results in terms of models in the North Pacific, BC, and Alaska generally performing poorly, and models in the Vancouver and Hecate Strait sub-areas performing well (more detailed results for all models and areas can be found in Table 5).

Variable Importance

For each area and model type, the model assigns a variable importance value to each of the 19 environmental predictors which were used as input to the model. The variables are arranged by importance to the model on a scale of 0-1. Individual variables ranked among the top 25% within each specific model are highlighted in Table 5. Alkalinity is ranked within the top 25% of variables in 13 out of 18 models. Oxygen is the next variable of highest importance being ranked within the top 25% of variables in 9 out of 18 models. Variables which are ranked within the top 25% in at least 4 of the 18 models include phosphate, silicate, temperature, nitrate, depth, omega aragonite, and omega calcite. The remaining variables are ranked within the top 25% for less than four models.

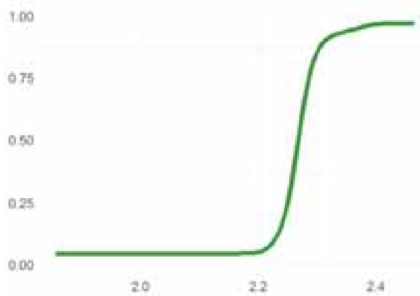
Table 5. Model Results, Fit Statistics and Variable Importance Values (top 25% of variables in each model are bolded in blue).

		North Pacific			***BC EEZ***			***ALASKA EEZ***			***US WOC EEZ***			***BC Hecate Strait***			***BC Vancouver Island***		
Model		GAM	BRT	MaxEnt	GAM	BRT	MaxEnt	GAM	BRT	MaxEnt	GAM	BRT	MaxEnt	GAM	BRT	MaxEnt	GAM	BRT	MaxEnt
Kappa		0.371	0.368	0.343	0.321	0.398	0.250	0.330	0.359	0.033	0.620	0.595	0.516	0.449	0.514	0.453	0.818	0.824	0.787
TSS		0.447	0.449	0.409	0.378	0.440	0.254	0.491	0.506	0.022	0.618	0.598	0.516	0.425	0.526	0.486	0.825	0.864	0.7
AUC Value		0.799	0.880	0.776	0.750	0.797	0.655	0.804	0.816	0.428	0.890	0.878	0.836	0.777	0.850	0.811	0.946	0.978	0.85
Variable Importance (Scale of 0-1)	Alkalinity	***1***	***0.410***	***0.071***	***1***	0.01	0.001	***1***	***0.163***	0.024	***1***	***0.22***	0.01	***1***	***0.136***	***0.137***	***1***	***0.013***	0.002
	Omega Aragonite	***0.973***	0.003	0.005	***0.703***	0.008	0	***0.837***	0.009	***0.05***	***0.679***	0.008	0.376	0.832	***0.048***	0.069	0.68	0.003	0
	Aspect	0.011	0.001	0	0.016	0.005	0	0.001	0.002	0.002	0	0	0.118	0.425	0.102	0.075	0.045	0.002	0
	Omega Calcite	***0.885***	0.009	0.015	***1***	0.021	0	***0.953***	0.014	0.023	***0.976***	0.011	0.219	***1***	0.005	0.046	***1***	0.001	0
	Depth	0.303	***0.025***	0.027	0.509	0.021	0.017	0.256	0.026	0.007	0.043	0.016	***0.7***	0.622	0.009	0.001	0.718	***0.101***	***0.166***
	Dissolved Inorganic Carbon	0.389	0.009	0	0.638	0.002	0.069	0.641	0.008	0.01	0.658	0.005	***0.629***	0.642	0.013	0.059	0.656	0.003	0.005
	Eastness	0.029	0.005	0.011	0.01	0.003	0.066	0.008	0.001	0	0.017	0.003	0.035	0.166	0.006	0	0.059	0	0
	Nitrate	0.367	***0.012***	***0.028***	0.559	***0.148***	***0.249***	0.35	0.021	***0.117***	0.2	0.027	0.322	0.099	***0.05***	***0.076***	0.609	0.003	***0.314***
	Northness	0.006	0.001	0	0.014	0.003	0	0.001	0	0	0.001	0	0	0.002	0.021	0.012	0.156	***0.016***	0.003
	Oxygen	***0.393***	***0.020***	***0.167***	0.48	0.007	0.111	0.113	***0.108***	***0.134***	0.405	***0.127***	0.239	***1***	0.011	0.074	***1***	0	***0.013***
	Phosphate	0.156	0.004	0.014	***1***	***0.121***	***0.13***	0.333	0.004	0.003	0.266	0.011	0.018	0.456	***0.048***	0.011	0.557	0	0
	Roughness	0.02	0.001	0.012	0.056	0.002	0.113	0.032	0.001	0	0.016	0.001	0.029	0.009	0.003	0	0.641	0.001	0.001
	Rugosity	0.003	0	0.004	0.05	***0.03***	***0.114***	0.025	0.001	***0.94***	0.015	0.003	0	0.029	0.006	0	0.443	0.001	0.001
	Salinity	0.021	0.008	0.014	0.633	0.018	***0.263***	0.116	***0.056***	0.021	0.017	***0.028***	***0.877***	0.589	0.011	0	0.562	0.002	0
	Silicate	0.241	0.002	***0.028***	0.482	0.002	0.071	***0.643***	***0.052***	0.003	0.504	***0.038***	0.001	***1***	0.008	0.028	0.821	0.001	0.009
	Slope	0.002	0.001	0.002	0.051	0.012	0.049	0.041	0.003	0.026	0.036	0.008	0	0.03	0.012	0	0.382	***0.053***	0
	Temp	0.257	0.002	0.001	0.211	***0.034***	0.059	0.144	0.011	0.041	***0.818***	0.019	***0.411***	0.599	0.002	0.034	***0.897***	0.001	0
	TPI	0.024	0.008	0.001	0.012	0.01	0.046	0.028	0.027	0.002	0.031	0.007	0.064	0.003	0.006	0.066	0.31	0.006	***0.031***
	TRI	0.003	0	0	0.004	0.001	0.037	0.032	0.001	0	0.257	0.001	0.021	0.009	0.025	***0.087***	0.449	0.004	0.002

Figure 8. Partial dependence plots for alkalinity in all areas and models with ranked variable importance.

GAM	BRT	MaxEnt	Data Distribution For Alkalinity
1 (1st)	0.41 (1st)	0.071 (2nd)	Percentage of Presence Values

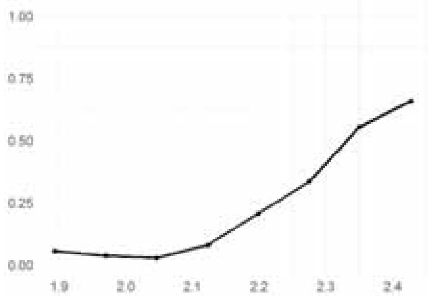
North Pacific



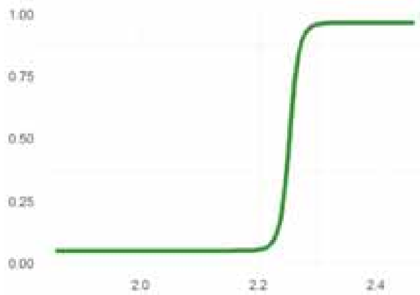
1 (Tied for 1st)

0.01 (Tied for 8th)

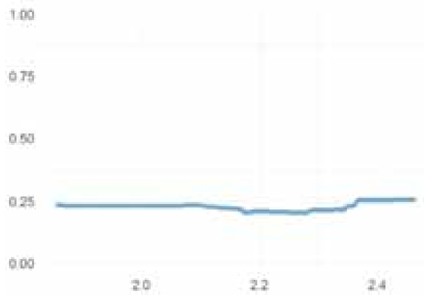
0.001 (14th)



BC

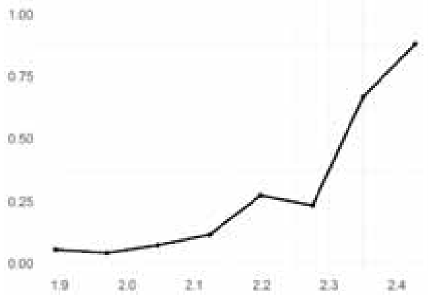


1 (1st)

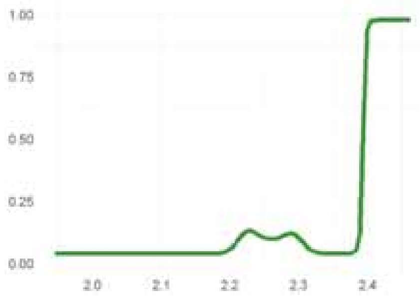


0.163 (1st)

0.024 (7th)

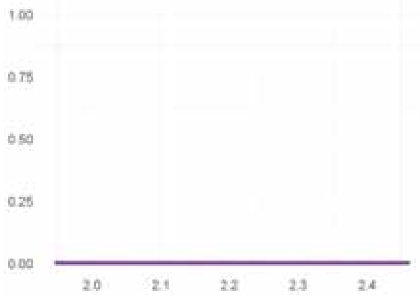


Alaska

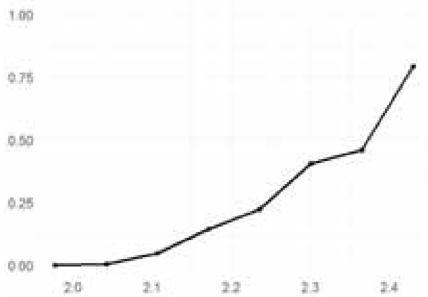


1 (1st)

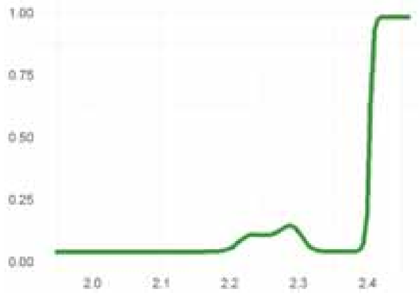
0.22 (1st)



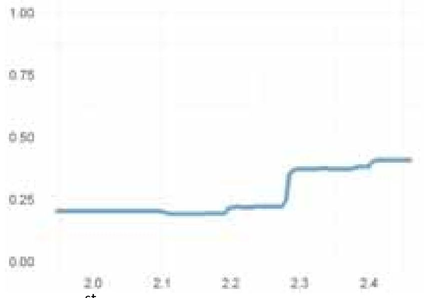
0.01 (15th)



US

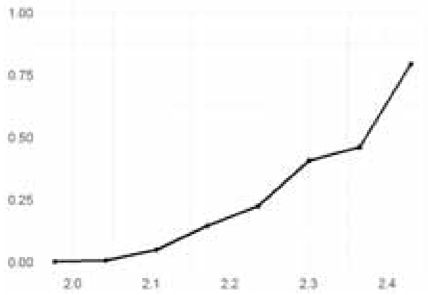


1 (Tied for 1st)



0.136 (1st)

0.137 (1st)



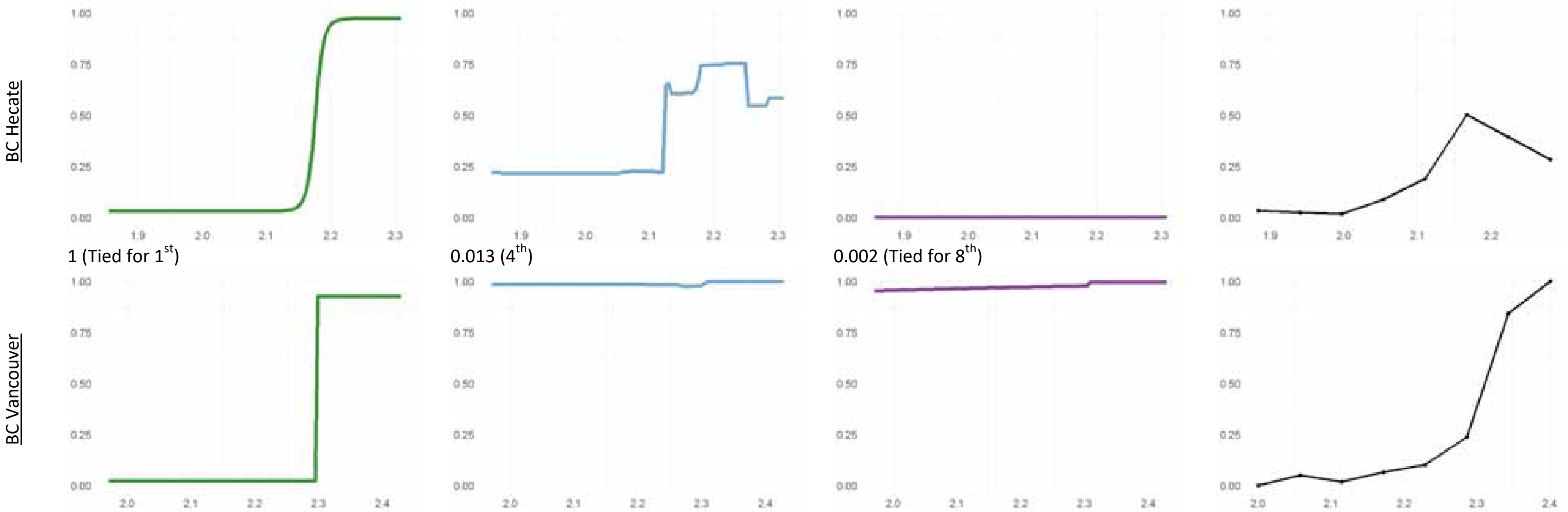


Figure 9. Partial dependence plots for oxygen in all areas and models with ranked variable importance.

GAM

0.393 (4th)

BRT

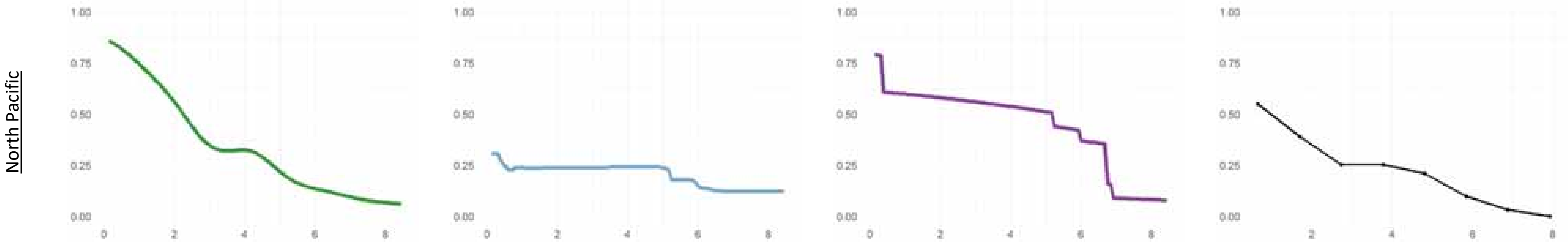
0.02 (3rd)

MaxEnt

0.167 (1st)

Data Distribution for Oxygen

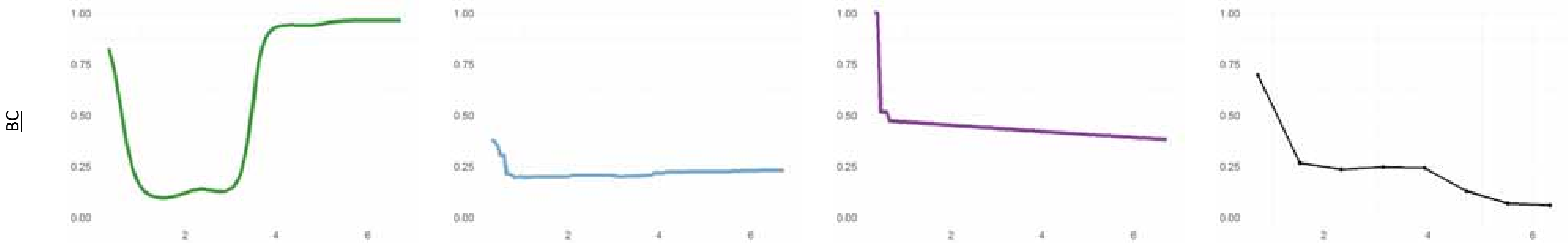
Percentage of Presence Values



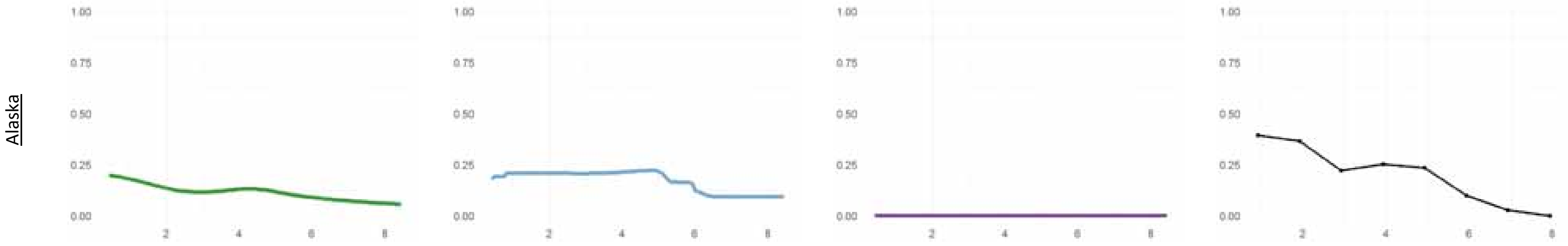
0.48 (10th)

0.007 (11th)

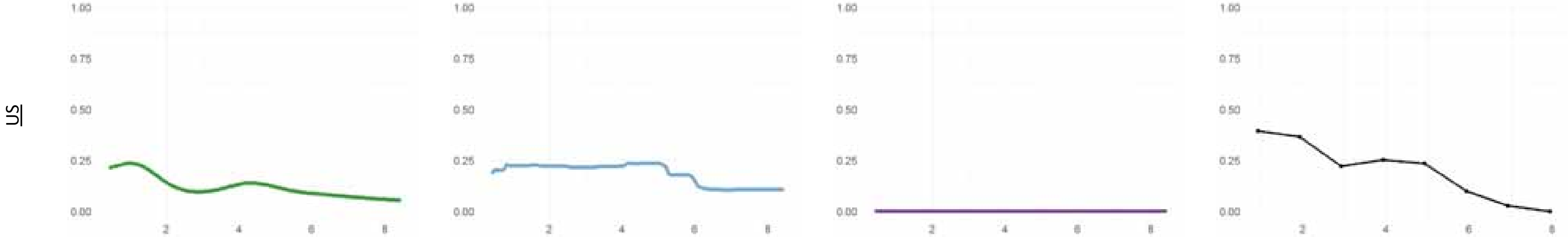
0.111 (6th)



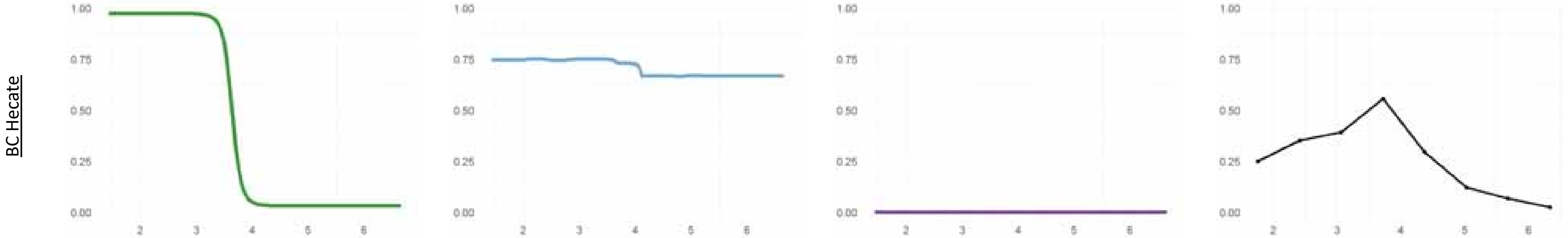
0.113 (11th) 0.108 (2nd) 0.134 (2nd)



0.405 (7th) 0.127 (2nd) 0.239 (7th)



1 (Tied for 1st) 0.011 (Tied for 10th) 0.074 (5th)



1 (Tied for 1st) 0 (Tied for 19th) 0.013 (4th)

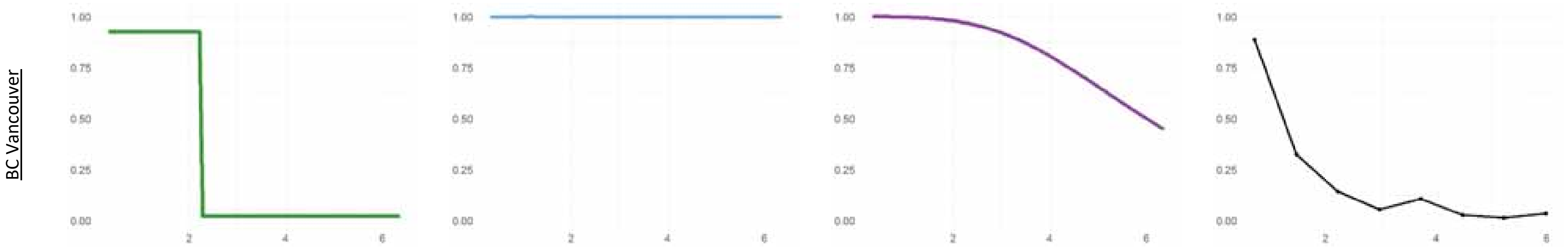
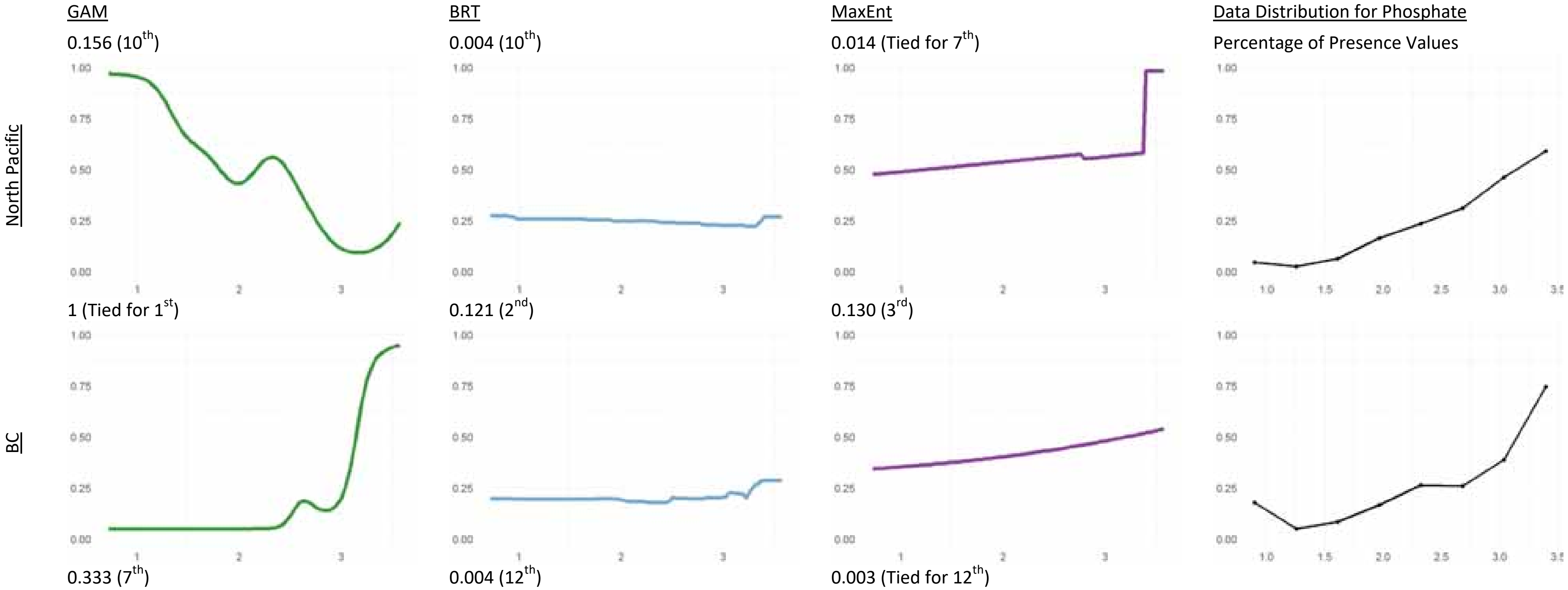


Figure 10. Partial dependence plots for phosphate in all areas and models with ranked variable importance.



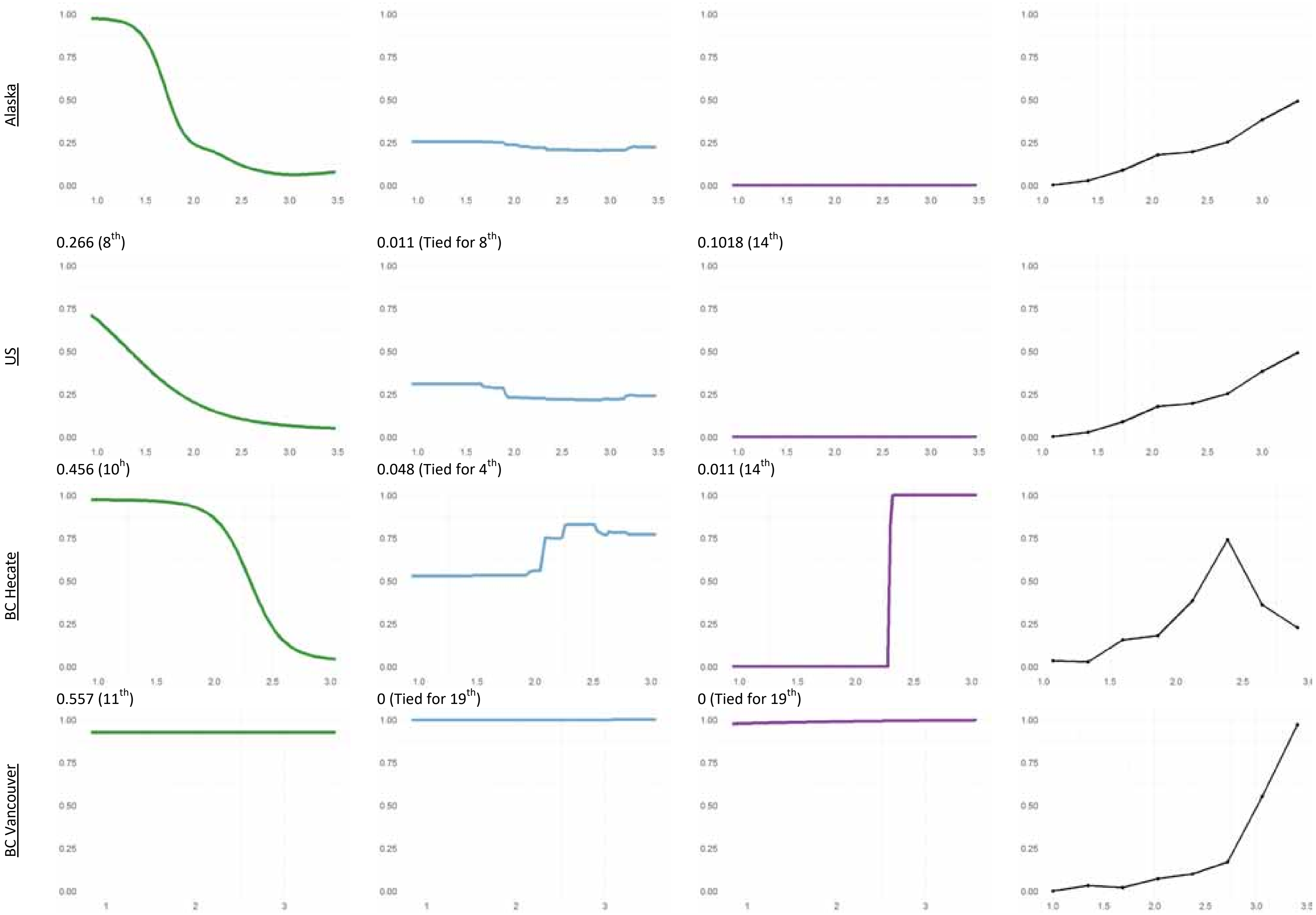
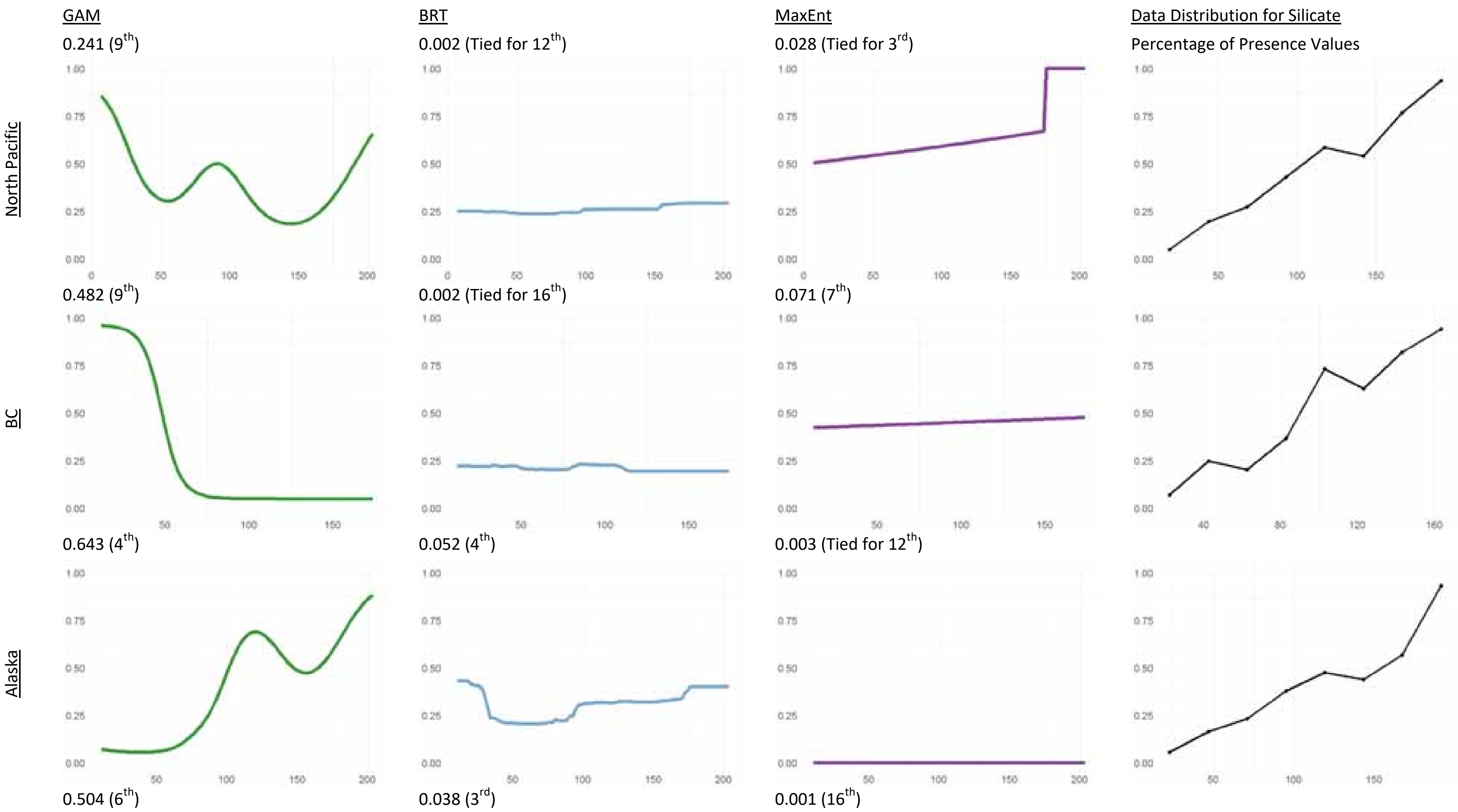

Alaska
US
BC Hecate
BC Vancouver
0.266 (8th)
0.011 (Tied for 8th)
0.1018 (14th)
0.456 (10h)
0.048 (Tied for 4th)
0.011 (14th)
0.557 (11th)
0 (Tied for 19th)
0 (Tied for 19th)
0.00
0.25
0.50
0.75
1.00
1.0
1.5
2.0
2.5
3.0
3.5
1
2
3

Figure 11. Partial dependence plots for silicate in all areas and models with ranked variable importance.



GAM

0.241 (9th)

BRT

0.002 (Tied for 12th)

MaxEnt

0.028 (Tied for 3rd)

Data Distribution for Silicate

Percentage of Presence Values

North Pacific

0.482 (9th)

0.002 (Tied for 16th)

0.071 (7th)

BC

0.643 (4th)

0.052 (4th)

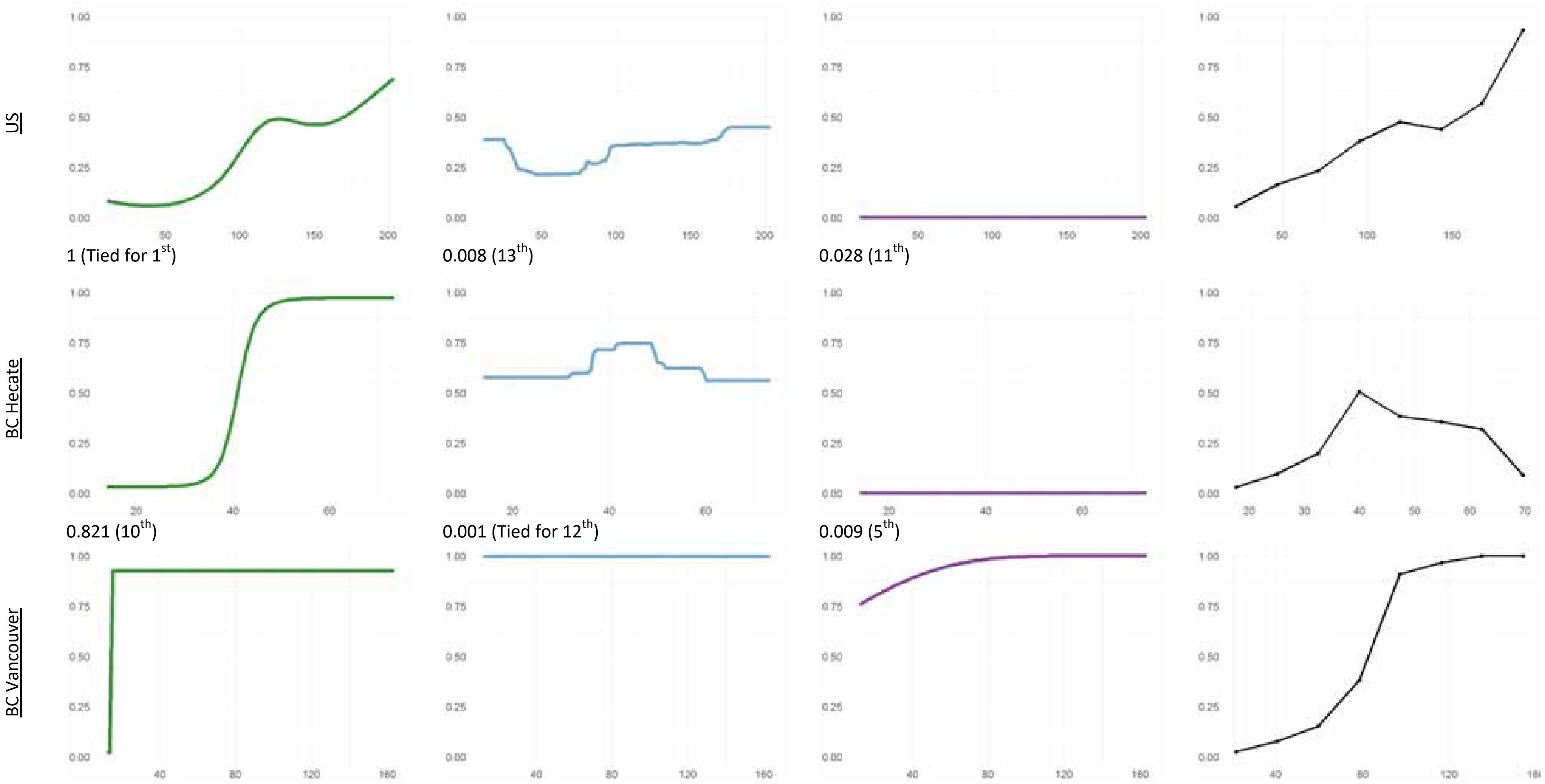
0.003 (Tied for 12th)

Alaska

0.504 (6th)

0.038 (3rd)

0.001 (16th)


US
BC Hecate
BC Vancouver
1 (Tied for 1st)
0.008 (13th)
0.028 (11th)
0.821 (10th)
0.001 (Tied for 12th)
0.009 (5th)
0.00
0.25
0.50
0.75
1.00
50
100
150
200
20
30
40
50
60
70
80
120
160

Partial Dependence Plots

Partial dependence plots generated for alkalinity, oxygen, phosphate and silicate can be seen in Figures 8-11. Table 5 presents depth as a frequently important variable in this analysis of glass sponges. Partial dependence plots from multiple models and areas show that as the taxon encounters depths shallower than 1000 m, the probability of presence decreases, confirming they are more likely to be found in deep waters (Fig. 12). Figure 13 shows the taxon data in the BC Vancouver sub area. It can be easily noted here that the majority of the presence values are in the deeper waters. It is important to note that glass sponges also exist in shallow waters, as evidenced by the glass sponge reefs of coastal British Columbia (Fig. 15).

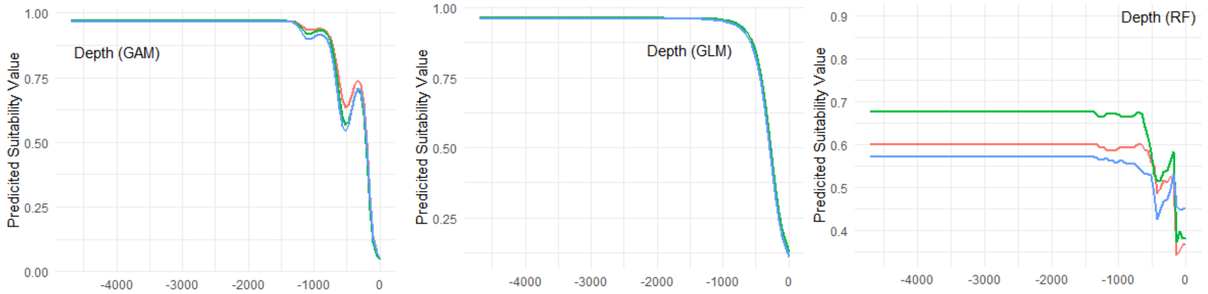
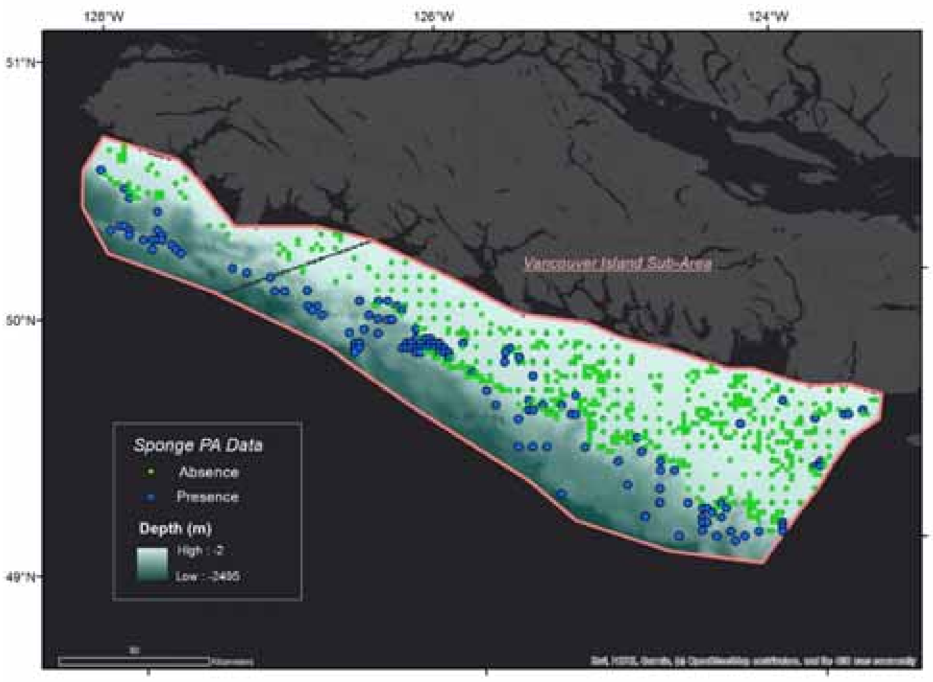


Figure 12. Partial dependence plots for depth from GAM, GLM, and random forest (RF) models.



Each plot has three lines for each time the evaluation was run.
Figure 13. Bathymetry and species data points within BC Vancouver Island sub-area (grey line is missing data).

Alkalinity

Figure 8 shows partial dependence plots and ranked variable importance for alkalinity, for each of the 18 model-area combinations. Within the North Pacific area, all three models suggest a high probability of glass sponge presence within highly alkaline waters. In the British Columbia area, the GAM model suggests an increase in glass sponge presence probability with alkalinity values higher than 2.2 µmol l^{-1}, a trend that is also present in each of the sub-areas. The BC Hecate area, in particular, suggests an increased probability of presence for glass sponges in alkalinity values of 2.2 µmol l^{-1} and higher. All three models produced comparable partial dependence plots for this area, where variable importance values for alkalinity rank 1st out of 19 variables for all three models. Finally, the data distribution plot for alkalinity values in the BC Hecate area shows a high percentage of presence values in alkalinity ranges of between 2.15 and 2.25 µmol l^{-1}. The BC Vancouver area has a comparable GAM response curve as BC Hecate Strait, however the BRT and MaxEnt models for BC Vancouver have lower variable importance ranks and do not show an increase in probability of presence with increased alkalinity values. The BC Hecate Strait area has a range of alkalinity values of 1.855 – 2.308 µmol l^{-1}, while most other areas have a maximum of closer to 2.4 or 2.5 µmol l^{-1}.

Oxygen

Partial dependence plots for oxygen (Fig. 9) are interestingly varied as well; the general trend seen in GAM models from several of the areas suggests an increased probability of presence with lower oxygen values, except that the GAM model for the BC area suggests the opposite. Oxygen was 10th in variable importance in the GAM model for the BC area, which means there is less indication in the model that glass sponges are strongly influenced by oxygen. The GAM plot for the BC area is interesting because it presents a pattern opposite to the oxygen plots for other areas, opposite to the plots for other models, and opposite to the data distribution itself. The variable importance values for oxygen in these models are not as consistently high as for alkalinity. The data distribution plots for all of the areas show a greater proportion of presence values in area with relatively low oxygen concentrations.

Phosphate

Response curves for phosphate can be seen in Figure 10 and present a wide variety of possible responses of probability of sponge presence in relation to phosphate content. Due to the lower variable importance values, it is more difficult to find strong environmental trends in the data. Phosphate was within the top three variables influencing the GAM, BRT and MaxEnt models within the BC area. For the remaining areas, the variable importances range from 4th to last (19th). The plots from the BC area indicate that probability of sponge presence increases with phosphate levels of roughly 3 µmol l^{-1} and higher. The next highest variable importance values are a result of the BRT and MaxEnt models for the BC Hecate Strait area. Phosphate was

4th in variable importance for these two models and both indicate a slight increase in probability of presence between 2.0 and 2.5 µmol l^{-1}.

Silicate

Models which indicated that higher silicate content is more suitable for sponges included MaxEnt (North Pacific Ocean, BC Hecate Strait and BC Vancouver), and GAM (Alaska, US, BC Hecate Strait and BC Vancouver) (Fig. 11). The data distribution plots largely indicate a greater proportion of presence values with increasing silicate value. The two partial dependence plots with the highest variable importance values (3rd) are the MaxEnt model in the North Pacific area and the BRT model in the US area. While both these plots indicate a general increase in probability of presence in relation to an increase in silicate levels, the BRT model in the US also indicates an increased probability of presence with very low silicate contents. Generally, because silica plots have lower variable importance values than alkalinity or oxygen plots, less weight can be placed on their accuracy. The silicate GAM plot for the BC area produced an opposite result to the remaining plots.

Spatial Predictions

Figure 14 shows the predicted probability of glass sponge presence from the BRT, GAM and MaxEnt models for Alaska, in the form of a raster prediction probability of presence.

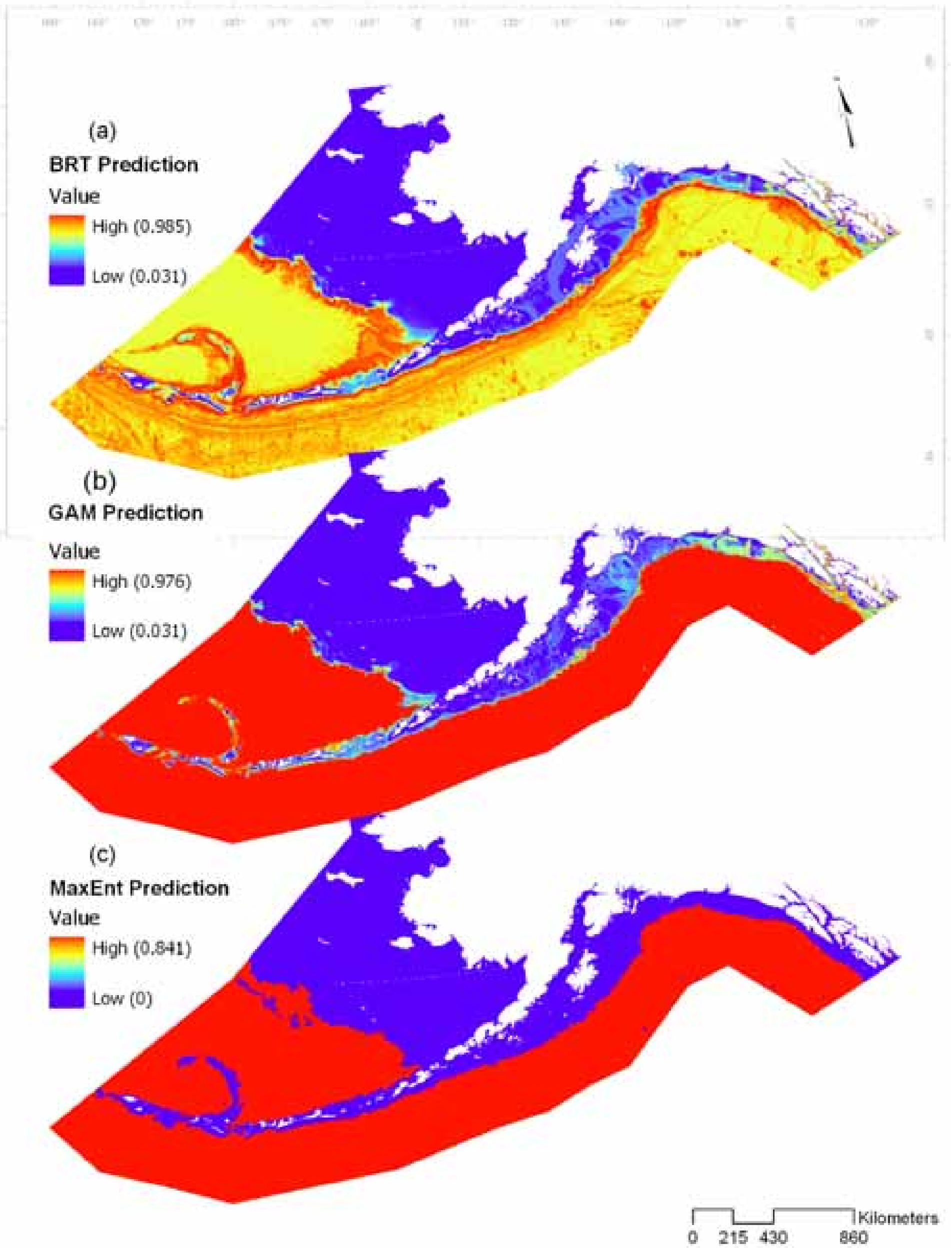


Figure 14. Model predictions from BRT, GAM, and MaxEnt models for the Alaska area.

Alaska was selected for this section because it is a smaller area than the North Pacific (which is a large area, thus making predictions difficult to see in detail) but larger than the BC and US areas (which are quite small and have less variation in predictions of probability of presence). The AUC values for these three models respectively are 0.816, 0.804, and 0.428, meaning the BRT and GAM models performed very well and the MaxEnt model performed poorly. The MaxEnt model for the Alaska area is the model that performed least well across all models and areas, and as can be seen in Figure 14c, the area is divided into red (high probability of presence) and blue (low probability of presence) without much variation between those two predictions. Figure 14a and b show much more variation of probability of presence across the prediction.

Uncertainty Metrics: BC Hecate Strait

Uncertainty refers to a lack of sureness or confidence about something (Elith et al., 2002b). Most outputs of SDM work are presented with confidence, with no indication of uncertainties, but it has been proposed that maps of uncertainty would help in the interpretation of these predictions (Elith et al., 2002b). The Hecate Strait sub-area in BC was used for the uncertainty metric analysis because of its high environmental variation and interesting patterns of alkalinity, oxygen and silicate distributions. The prediction from a GLM run on this area can be seen in Figure 15, along with the outlines of the Hecate Strait/Queen Charlotte Sound Glass Sponge Reefs Marine Protected Area (MPA).

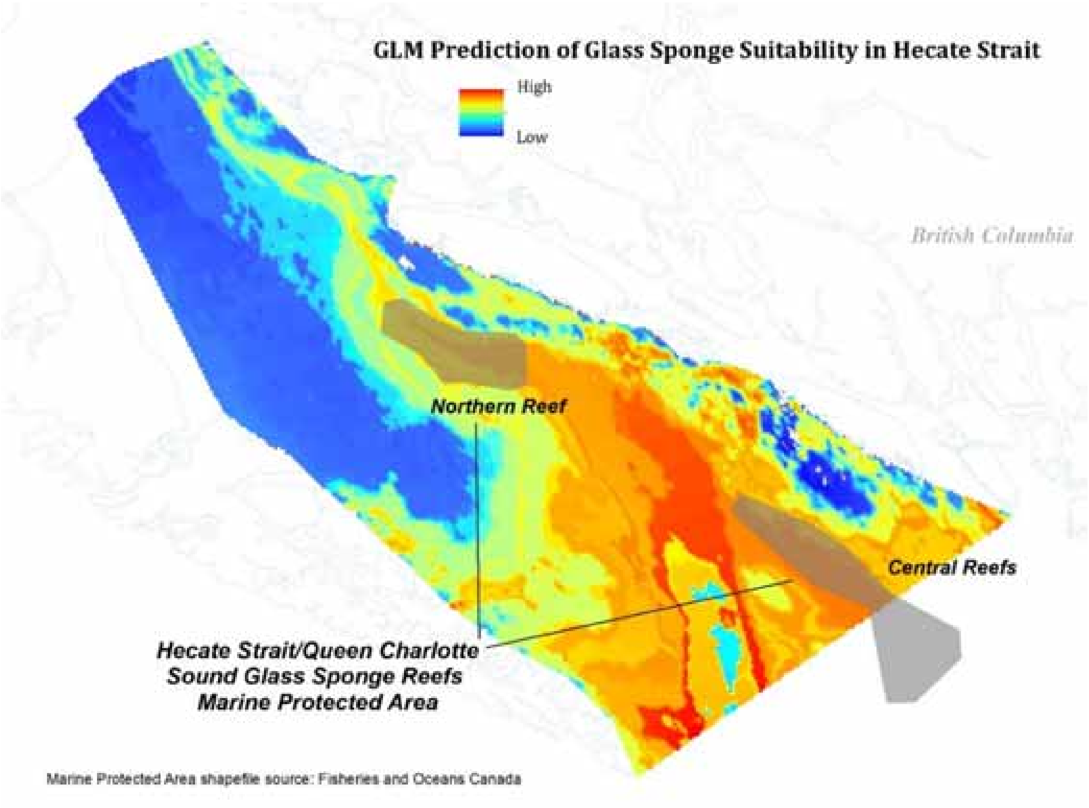


Figure 15. GLM prediction of glass sponge probability of presence in Hecate Strait with MPA boundaries.

The Northern Reef and part of the Central Reef are contained within the Hecate Strait boundaries employed for this study. The MPA boundaries overlap with moderately high suitability for glass sponges; providing some confidence in the model predictions and their real-world accuracy, despite the MPA area boundaries not falling within the highest probability of presence areas (red areas).

Standard Error of GLM Prediction

Figure 16 shows the partial dependence plots for alkalinity, oxygen and silicate for the GLM model of the BC Hecate Strait area, with model-based confidence intervals added. Generally, the confidence intervals are narrow, corresponding to a low expected error, for predictor value ranges with many data points, shown in the figure as a high density of red/blue lines. Value ranges with wider intervals have fewer data points. If the areas with wide confidence intervals overlap spatially, the relevant areas are likely to produce less certain predictions.

The SE of the GLM fit was written to a raster and can be seen in Figure 17a. Area 2 in Figure 17a has high uncertainty.

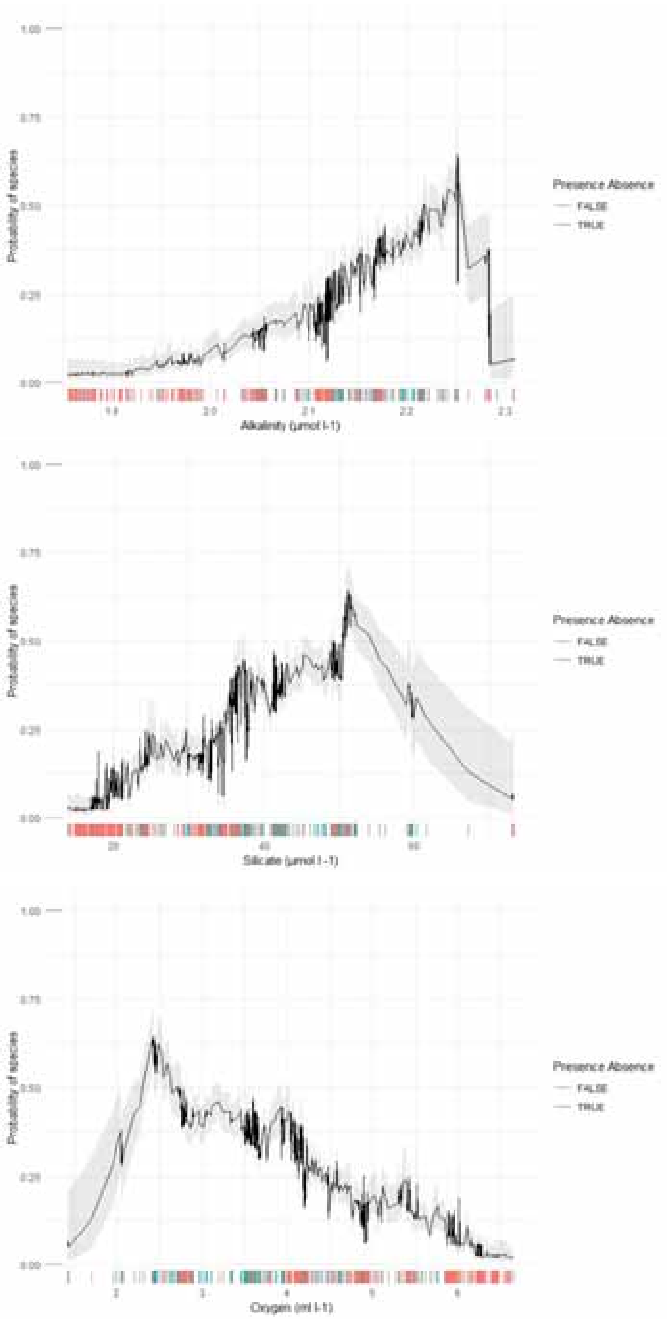


Figure 16. Partial dependence plots for alkalinity, silicate, and oxygen with estimated confidence intervals.

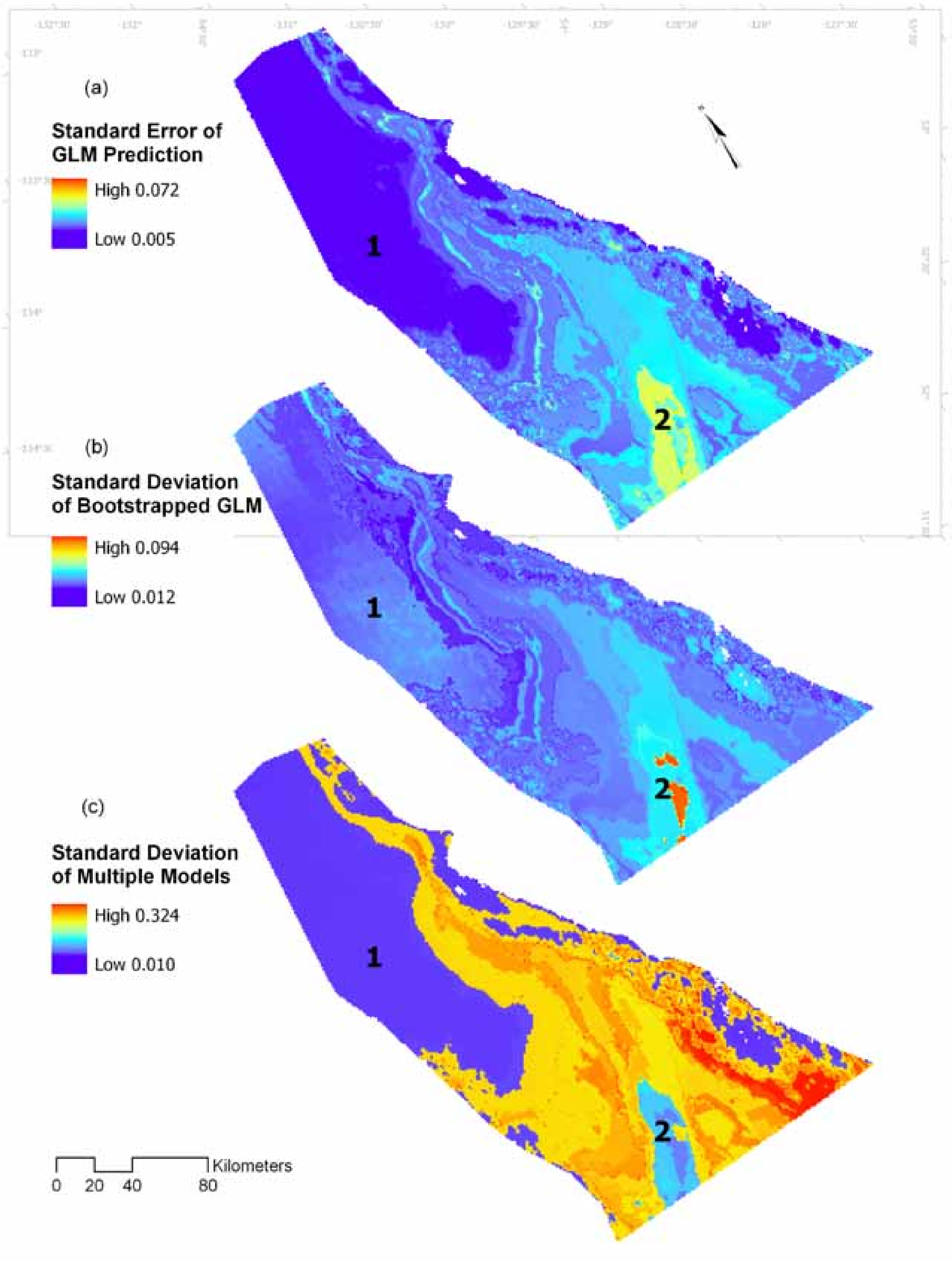


Figure 17. Uncertainty metrics mapped to Hecate Strait: a) SE of GLM prediction, b) SD of bootstrapped GLM, and c) SD of multiple models.

There are very clear environmental gradients throughout Hecate Strait which can be visually confirmed to have an influence on the uncertainty metrics. Area 1 in Figure 17a has low alkalinity and silicate levels with high oxygen levels (Figs. 19-21). This combination of environmental values generally coincides with absence data for glass sponges, which the models interpret as unsuitable habitat. The SE in Fig. 17a is low, indicating high certainty in the prediction of low probability of sponge presence. Area 2 in Figure 17a has high uncertainty values. Area 2 corresponds with opposite niche environmental characteristics to Area 1; very high alkalinity levels, very high silica levels, and very low oxygen levels, all of which are value ranges that are poorly represented in the data. As can be seen in the partial dependence plots (Fig. 16), these value ranges are associated with low data density and high CIs. The GLM is forced to make predictions for these areas based on a combination of few data points with similar values and extrapolation from more data-dense value ranges, which leads to extreme predictions and higher uncertainty.

Standard Deviation of Bootstrapped GLMs

Figure 17b presents the result of bootstrapping the binomial GLM 200 times, calculating the SD for each cell and then writing this result to a raster. This method of spatially showing prediction uncertainty yields similar results to the initial method of calculating the SE from the GLM. It shares an area of high uncertainty with the first method (Area 2), which was noted as having extreme values of all three input predictors. This method, as well as the first method, does not show Area 1 to have high SE, indicating consistency across methods.

Standard Deviation of Multiple SDMs

Finally, Figure 17c presents the result of running multiple SDMs and mapping the SD of the model predictions. The models used and their individual predictions can be seen in Figure 18. This method shows the highest uncertainty to exist in the lower right corner of Hecate Strait, which corresponds to medium uncertainty in the first two methods. While taking a different approach from the first two methods, this final method is equally as important for determining spatial uncertainty from predictive models and yields interesting results that could aid policy makers in making informed decisions based on SDMs. Area 1 in Figure 17c has low SD because the majority of the SDMs produced a similar probability of presence for this area (Fig. 18). This is a consistent result from all three methods, indicating low uncertainty in environments considered unsuitable for the taxon in question. The areas that are yellow and orange yield higher error values because the models produced different results, despite having been calibrated with the same data. Each of the specific model outputs showed the highest probability of species to be in the area which has high standard deviation in Figure 17c. The resultant high SD is a result of this being the area of the model output that changes the most with different SDMs.

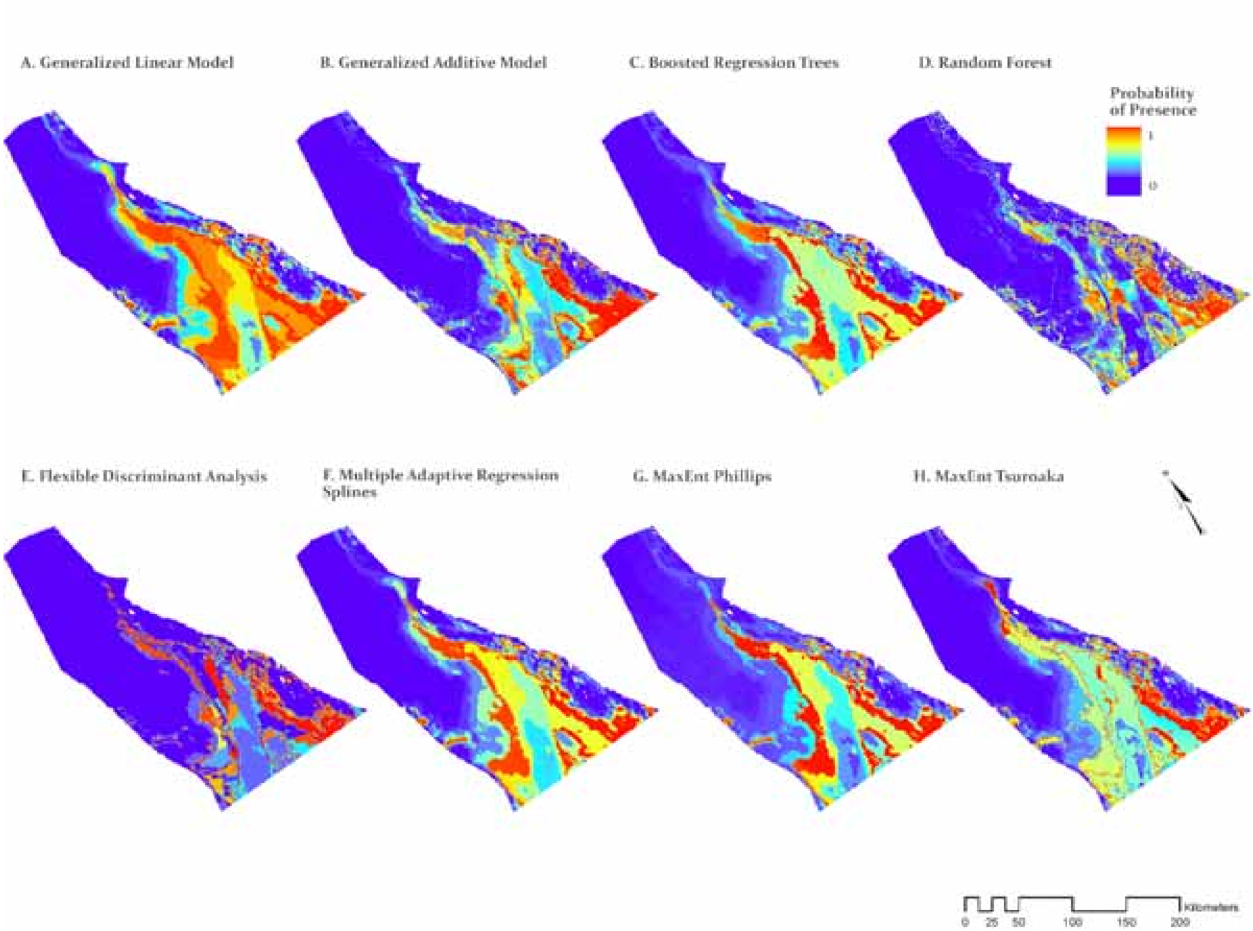


Figure 18. Predictions of glass sponge probability of presence in Hecate Strait: a) GLM, b) GAM, c) BRT, d) RF, e) FDA, f) MARS, g) MaxEnt Phillips, and h) MaxEnt Tsuroaka.

Discussion

Aspatial Model Predictions: Partial Dependence Plots and Variable Importance

As a method of assessing the validity of models, partial dependence plots and variable importance rankings were presented and analyzed. The results presenting partial dependence plots and variable importance values from the GAM, BRT and MaxEnt models can provide information about 1) the ability of the model to describe the environment-species relationships, and therefore 2) the potential for using the model to make inferences about the ecology of glass sponges and characteristics of their habitats.

In addition to standard model performance metrics such as AUC, TSS and Kappa, ranked variable importance and partial dependence plots can inform how certain or uncertain an SDM-based prediction is. If a multi-model and multi-area approach has been used for the modeling, and a given environmental variable has high importance values across multiple model types and areas, a higher confidence can be placed in that variable having a non-spurious effect on the distribution of the response variable, e.g. glass sponges in the present case. It is likely that the use of more models, and more environmentally distinct areas would serve to strengthen the multi-model, multi-area approach even more. The expected result from performing a multi-area, multi-model approach was that strong habitat preferences will be reflected similarly in partial dependence plots from different models and areas, while weak habitat preferences will not.

Using this approach, the partial dependence plots for alkalinity strongly suggest a causal relationship between alkalinity and glass sponge presence. According to the model results, glass sponge probability of presence increases in conjunction with higher alkalinity values, specifically at concentrations of 2.1 µmol l^{-1} or 2.2 µmol l^{-1} and higher. While several of the areas modeled produced this trend, the BC Hecate area displays arguably the most consistent result based on the fact that all of the three models for this area ranked alkalinity as first out of nineteen other variables. Less confidence can be placed in certain areas and models where alkalinity is ranked lower and the associated response curves are inconsistent with those of high variable importance. For example, MaxEnt models in BC and US areas have alkalinity importance rankings of 14th and 15th respectively, and neither of the corresponding partial dependence plots provide any useful ecological information about how sponges respond to alkalinity levels. Although the models are too complex to provide a definitive explanation, it is likely that the BC and US areas have other environmental variables which are more influential for the model, and therefore the relationship between sponge presence and alkalinity is more difficult for the model to identify. Every highly ranked (top 25 percentile) partial dependence plot of alkalinity shows an increased probability of presence associated with high alkalinity values, generally above 2.2 µmol l^{-1}. Areas of Hecate Strait which meet these alkalinity values are relatively few and can be seen in Figure 19.

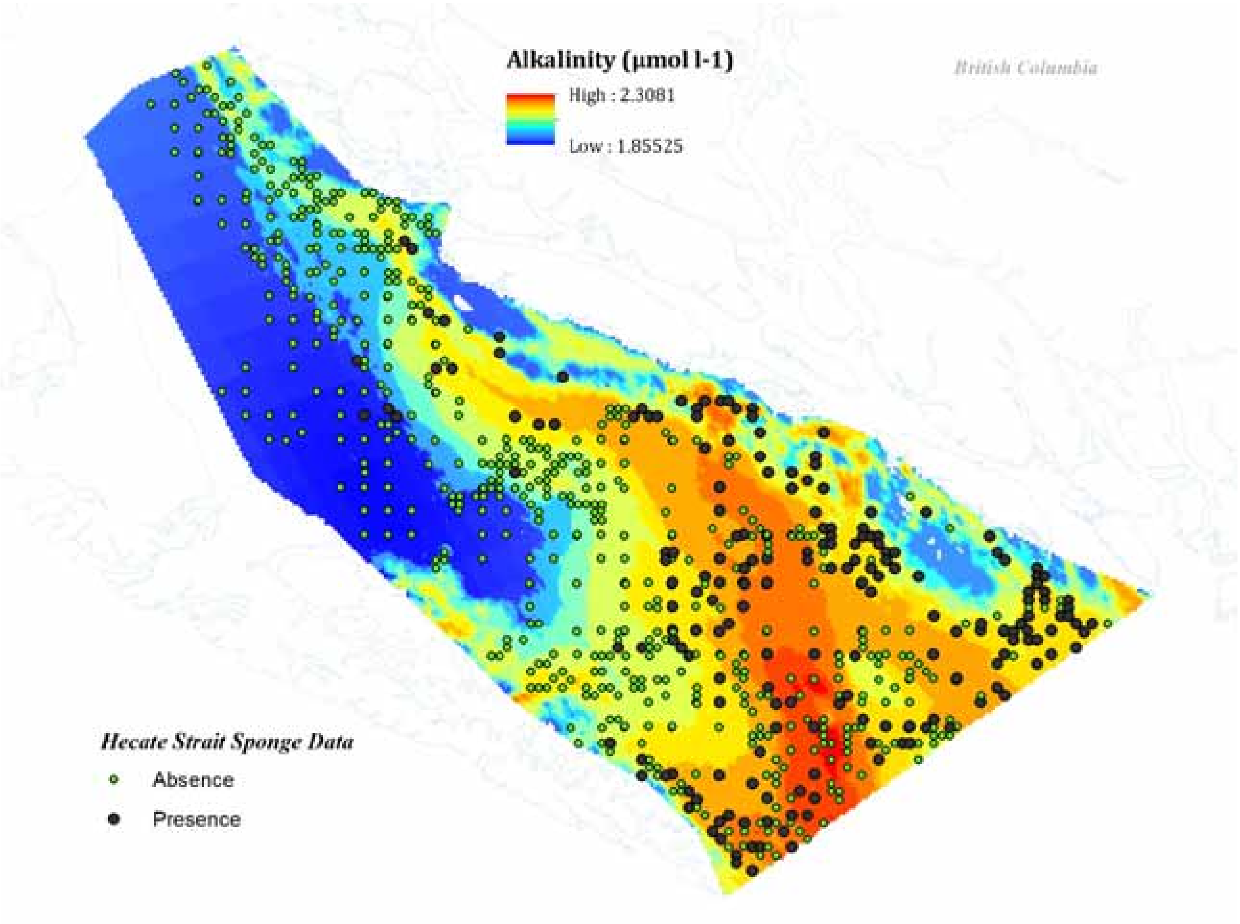


Figure 19. Alkalinity values across Hecate Strait.

According to the partial dependence plots with high variable importance rankings, a high probability of glass sponge presence is associated with low oxygen values, generally below 2 ml l^{-1}. Deep-water sponges, and many other invertebrates, use little oxygen and have adapted to live in low-oxygen environments, for example during low tide or in benthic sediments (Leys et al., 2018). Leys and Kahn (2018) note that glass sponges tolerate long-term hypoxic conditions by reducing their filtration rate and feeding activity. Filtration, they concluded, is costly to glass sponges and attempting to slow their filtration has driven innovations in their morphology and physiology (Leys et al., 2018). Chu et al. (2019) also found that dissolved oxygen was a highly ranked positive predictor of habitat for cold-water coral and sponge grounds in the Canadian northeast Pacific Ocean. As a result of this finding, Chu et al. (2019) predicted that cold-water coral and sponge taxa would have lower oxygen requirements in comparison to highly mobile taxa such as fish. It was also found, in an attempt to validate the models predicting that cold-water corals and sponges are likely to occur in severely low oxygen environments, that these taxa exist in oxygen levels as low as 0.2 ml l^{-1} at the Union and Dellwood seamounts (both are southwest of the southern point of Haida-Gwaii). Figure 20 shows the distribution of oxygen content across Hecate Strait. It is important to note how related oxygen and alkalinity are to each other in this area. With many sponges existing in high-alkalinity, low-oxygen waters, it is difficult to know whether this is because of the high alkalinity, the low oxygen, some combination of the two, or a third variable that is also correlated with both alkalinity and oxygen. This illustrates the benefit of using a multi-area approach, because these two variables may be less related in other areas which have also been modeled.

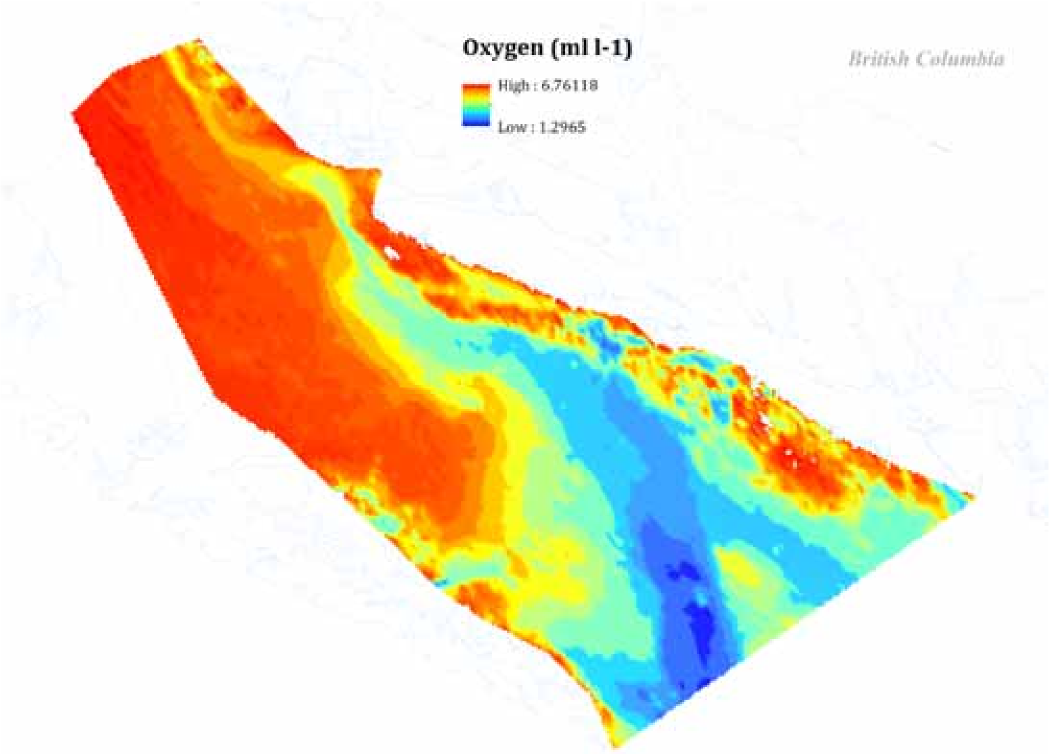
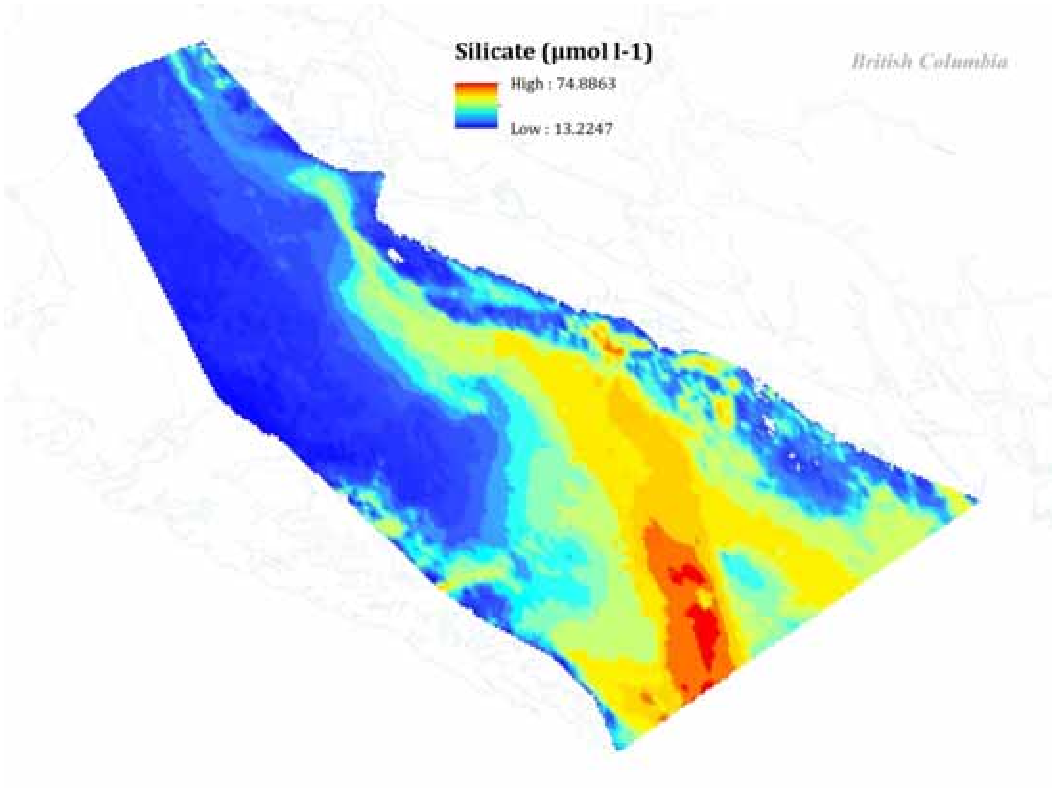


Figure 20. Oxygen values across Hecate Strait.

Silicate was not often returned from the models as one of the top 25% of variables, however certain models and areas did produce silicate as the most important variable. These included MaxEnt for the North Pacific area, GAM and BRT for the Alaska area, BRT for the US area, and GAM for the BC Hecate Strait area. Out of these, most show an increased probability of glass sponge presence with high levels of silicate. High levels of silicate often overlap with high levels of alkalinity within Hecate Strait (compare Fig. 19 and Fig. 21).

Figure 21. Silicate values across Hecate Strait.

In published literature, it has been indicated that glass sponges need high levels of dissolved silica (Leys et al., 2004; Austin., 1984; Chu et al., 2019). Chu et al. (2019) found that silicic acid was a top predictor for sponge groups because biogenic silica (biogenic silica occurs when dissolved



silicate transforms to particulate skeletal matter (Treguer et al., 1995)) can constitute over 90% of the biomass of cold-water sponges. Silicate levels are high in both the Antarctic Ocean as well as the coastal northeast Pacific Ocean, which are both regions of high glass sponge abundance (Leys et al., 2004; Treguer et al., 1995). Interestingly, the Hecate Strait area shows an increased probability of glass sponge presence at much lower levels of silicate than the remaining areas tested within this contribution. Many areas indicate high probability of glass sponge presence in areas with silicate values of 150 µmol l^{-1}, however Hecate Strait indicates high probability of presence beginning where silicate values reach over 40 µmol l^{-1}. The highest level of silicate within the Hecate Strait subarea is 74 µmol l^{-1}. Whitney et al. (2004) identified silicate levels of over 40 µmol l^{-1} around sponge reefs in Hecate Strait, thus confirming this result.

Evaluating partial dependence plots in addition to model accuracy metrics as an additional way to assess SDM outputs is a descriptive and largely qualitative exercise. The challenge is that there can be valuable information concerning ecological relationships, but also nonsensical and spurious relationships, presented in these plots. But there is value in producing multiple models for multiple areas because oceanic environments subject to different currents and water masses and different levels of terrestrial influence can vary drastically in their biogeochemistry. This contribution presented one approach to disentangling the two; by looking for species-environment relationships that are strong (as indicated by high variable importance), consistent between model types and consistent between different areas, it is possible to extract only those relationships most likely to be caused by ecological processes. The two anomalous plots mentioned in the results section, the oxygen and silicate plots from the GAM model for the BC area, presented the opposite relationships of what the remaining models presented. These two plots are an excellent example of why it is important to not make inferences about ecological relationships based on single-model and single-area partial dependencies.

Using ranked variable importance values and selecting a threshold for a confidence cut-off can provide a quantitative measure of accuracy. Providing a measure of probable accuracy alongside model outputs can be helpful for environmental managers and stakeholders who require numerical models to estimate species distribution to design effective spatial management measures for conservation and protection.

Spatial Model Uncertainty Predictions

Spatially examining the uncertainty of model predictions is not commonly done in SDM studies, however it is important that potential users of SDM products have an understanding of the predictive accuracy of models and how this may vary across geographic space (Elith et al., 2005). Most evaluation metrics of predictive performance use a comparison of predictions against observations at a particular set of sites (Fielding et al., 1997). As also done in this contribution, statistics such as kappa and AUC values are widely used to assess whether predictions are suitably accurate for their intended use, however these statistics are somewhat restricted because they do not assess the predictions in geographic space and do not allow for exploration of spatial errors (Elith et al., 2005; Elith et al., 2002a; Fielding et al., 1997).

Confidence Intervals and Standard Error

It is suggested in SDM literature that plotting CIs around model predictions could be crucial to the interpretation of the models' performance, particularly mapping CIs of these predictions (Elith et al., 2002b; Elith et al., 2005). CIs around plotted responses (such as partial dependence plots) help show where species-predictor variable relationships are most uncertain (Ferrier et al., 2002). Adding these error metrics is instrumental for producing models which can be understood as ecological realities (Elith et al., 2005). Figure 16 shows CIs added to partial dependence plots of alkalinity, oxygen and silicate. The largest CI on each plot correlates with the lowest density of data points in each variable. Sponge data where silicate values are greater than 55 µmol l^{-1} have the highest uncertainty, as there are fewer data points for silicate values over this threshold. As mentioned earlier, silicate levels in Hecate Strait have been documented to be lower than surrounding areas, yet glass sponges remain in great abundance in Hecate Strait.

According to Figure 15, the highest probability of glass sponge presence within Hecate Strait coincides with areas comprised of high alkalinity, very low oxygen, and medium-high silicate levels. Hecate Strait is a shallow asymmetric channel between Haida Gwaii and the northern mainland of British Columbia (Perry et al., 1994). It is a unique area due to its shape; it is roughly 140km wide at its southern end and narrows to 48km in the north, covering around 23,000km^2 with depth values reaching down to 494m. The shallowest part is the northwest area, which has low alkalinity levels, high oxygen levels and low silicate levels. This is also the area for which the lowest probability of presence for glass sponges was predicted (Fig. 15). This northwest area (labelled as Area 1 in Fig. 17) shows consistently low uncertainty with all three methods. All models used predicted low probability of sponge presence in this area, due to its physical characteristics mentioned previously, and all methods of quantifying uncertainty show low uncertainty in this area, indicating it is highly probably this area is unsuitable for sponges.

Using CIs as the only means of quantifying the uncertainty of SDM predictions is not a complete method, according to Elith et al. (2002b), who mention that uncertainty in model outputs is not explicitly accounted for in the CIs of GLMs. They suggest that bootstrapped CIs can better account for different sources of uncertainty rather than simply applying CIs to GLM predictions. This is an interesting avenue for further work on spatially quantifying model uncertainty, as only the SE of the GLM fit and SD of the bootstrapped GLMs were calculated in the work that underlies this contribution. Area 2 was identified as an area of highly uncertain predictions by both the SE of the GLM fit and the SD of the bootstrapped GLMs. The bootstrapped runs of the GLM produce very consistent predictions in the northwest part of Hecate Strait (Area 1). The third method of measuring uncertainty involved running eight SDMs on the same data used for the prior methods and then calculating SE of all eight predictions. Figure 17c presents high SE values around the southeast corner of Hecate Strait. This area has medium uncertainty in the first two methods, indicating slight differences in model predictions. The eight models produced consistent predictions for the northwest area of Hecate Strait, suggesting with a low level of uncertainty that sponge probability of presence in this area is low.

One method of comparing these three metrics of estimating uncertainty is by looking at the original biological input data. Area 1 mostly contains absence values and almost no presence values (Fig. 19), therefore it seems that when every model shows low probability of presence, one can assume with some confidence it is likely correct. Additionally, Area 1 is shallow, has high levels of oxygen and low levels of both silicate and alkalinity, which are environmental conditions that are the opposite of what models generally predict as suitable habitat for sponges. Figure 17a, b and c present Area 1 as having low uncertainty, indicating all assign low uncertainty to the prediction of low probability of presence in this area.

These results allow for the conclusion that if the model predicts low probability of presence, it has higher certainty in this prediction than in predictions of high probability of presence. The first two methods differ from the last method in what they show, but the first two methods show medium uncertainty in the same areas which are highly uncertain in the last method (the areas where the models all predict relatively high probability of presence, but of varying values and slightly different geographic spreads of this high probability of presence). This leads to the conclusion that uncertainty is generally lowest where the models predict the species not to be, and highest where the models predict the highest probability of species presence to be.

Both bootstrapping a model and running multiple SDMs are useful methods of calculating prediction uncertainty, and both these methods could be extremely useful for providing planners with information to consider when employing the predictions in conservation planning and decision making (Ferrier et al., 2002).

Limitations

Numerous limitations exist within SDM work; not all models are transposable to distinct environments, they are strongly dependent on the considered scale, they are difficult to implement in a management context, many models are not easily interpretable, and software is not always available to practitioners (Guisan et al., 2005). Alongside these limitations, a consistent limitation is the fact that any model will rely heavily on the quality of the input data. This study used a spatial data thinning method based on local environmental variation to eliminate the spatial sampling bias that was present in the original data set. Spatial bias is a common limitation in the SDM field because it may cause biased model results and it is difficult to tell if the species-environment relationships in the model are representative of the real world or if they are a function of how the data was sampled. When using biological data from another organization, such is the case here, it can be challenging to find sufficient details about the data to ensure its quality. There are also limitations in the interpretation of regression-based models and machine learning models. The two methods produce different results. For example, calculating uncertainty metrics is more easily done from regression models as opposed to machine learning models. This contribution compared a regression-based model with two machine learning models. A final limitation often overlooked in SDM studies is the spatial dependency of accuracy of the model outputs. Presenting the error spatially is an important aspect of SDM moving forward, as it will be easily understood by those in environmental management who are unfamiliar with the modeling methods. Spatially presented error metrics add value to the already used aspatial error metrics.

Conclusion

By analyzing a variety of commonly used SDMs and examining different spatial and aspatial metrics to quantify model accuracy and uncertainty, this contribution has shown how applying a multi-model and multi-area approach can improve the interpretation of the modeled species-environment relationships. It has also shown how different methods of uncertainty mapping can provide increased insight as to which areas are predicted by the model to have high/low levels of uncertainty.

Running three models on six areas showed that partial dependence plots can differ substantially between model types and adjacent geographical areas. It is therefore necessary to not overstate the ecological results presented in individual plots, and to be careful while interpreting them ecologically. One way to assess the ecological interpretability of partial dependence plots is to perform a multi-model, multi-area study, and compare plots across models and areas prior to drawing ecological inferences.

Based on the results presented in this contribution, it appears that glass sponges are most likely to be found in areas with alkalinity values greater than 2.2 µmol l^{-1} and oxygen values lower than 2 ml l^{-1}. While silicate was also an important environmental predictor, the results for the probability of sponge presence in relation to silicate are more variable. Every area except Hecate Strait indicated that glass sponges are more likely to exist in areas with silicate values of 150 µmol l^{-1} and over, however lower values in Hecate Strait confirm sponges can exist in areas with silicate values of 40 µmol l^{-1} and over.

While model accuracy metrics like AUC and TSS contain important information about the ability of a model to produce good predictions, spatial uncertainty metrics can outline areas where predictions are more or less likely to be correct. There is a small area in the south of Hecate Strait (Area 2) that both the GLM and the bootstrapped GLM indicate as being subject to highly uncertain predictions. Such areas should be treated cautiously regardless of the overall accuracy of the model as indicated by the accuracy metrics, and such areas could be targeted for future data collection.

Finally, it was shown that different approaches to estimating prediction uncertainty can yield different but important results. This can be seen in Hecate Strait: predictions for the shallow, low-alkalinity area in the northwest part of Hecate Strait (Area 1) are consistent between models (all models give low probability of sponge presence), consistent between bootstrapped runs of the GLM models and the SE of the GLM model (which also all predict low probability of sponge presence), and are also consistent with the data points from that area. This indicates we are very confident in the model prediction of low probability of sponge presence in Area 1. The highest uncertainty corresponds to areas where models have presented high probability of presence, since these areas do not all overlap neatly, the uncertainty arises from these varying predictions.

References

Aarts, G., J. Fieberg, and J. Matthiopoulos. (2012). Comparative interpretation of count, presence–absence and point methods for species distribution models. *Methods in Ecology and Evolution, 3*(1), 177-187.

Allouche, O., Tsoar, A., and Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology, 43*, 1223-1232.

Araújo, M. B., Pearson, R. G., Thuiller, W., & Erhard, M. (2005). Validation of species-climate impact models under climate change. *Global Change Biology, 11*(9), 1504-1513.

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology and Evolution, 22*(1), 42-47.

Atwater, D., & Fautin, D. G. (2001). "Hexactinellida". University of Michigan, Museum of Zoology. On-line resource: https://animaldiversity.org/accounts/Hexactinellida/

Austin, W. C. (1984). Underwater birdwatching. *Can Tech Rep Hydrogr Ocean Sci 38*, 83–89.

Austin, W. C. (1999). The relationship of silicate levels to the shallow water distribution of hexactinellids in British Columbia. *MEMOIRS-QUEENSLAND MUSEUM*, 44, 44-49.

Beazley, L. I., Kenchington, E. L., Murillo, F. J., & Sacau, M. D. M. (2013). Deep-sea sponge grounds enhance diversity and abundance of epibenthic megafauna in the Northwest Atlantic. *ICES Journal of Marine Science, 70*(7), 1471–1490.

Becker, J. J., Sandwell, D. T., Smith, W. H. F., Braud, J., Binder, B., Depner, J., Fabre, D., Factor, J., Ingalls, S., Kim, S. H., Ladner, R., Marks, K., Nelson, S., Pharaoh, A., Trimmer, R., von. Rosenberg, J., Wallace, G., & Weatherall, P. (2009). Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution: SRTM30_PLUS. *Marine Geodesy, 32*(4), 355–371.

Booth, T. H., Nix, H. A., Busby, J. R., & Hutchinson, M. F. (2014). Bioclim: The first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. *Diversity and Distributions, 20*(1), 1-9.

Boria, R. A., Olson, L. E., Goodman, S. M., & Anderson, R. P. (2014). Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, *275*, 73–77.

Breiman, L. (2001). Random forests. Machine Learning, 45, 5-32.

Cañadas, A., Sagarminaga, R., De Stephanis, R., Urquiola, E., & Hammond, P. S. (2005). Habitat preference modelling as a conservation tool: Proposals for marine protected areas for cetaceans in southern Spanish waters. *Aquatic Conservation: Marine and Freshwater Ecosystems, 15*(5), 495–521.

Chefaoui, R. M., & Lobo, J. M. (2008). Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling, 210*(4), 478-486.

Chu, J. W. F., Nephin, J., Georgian, S., Knudby, A., Rooper, C., & Gale, K. S. P. (2019). Modelling the environmental niche space and distributions of cold-water corals and sponges in the Canadian northeast Pacific Ocean. *Deep-Sea Research Part I: Oceanographic Research Papers, 151*.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46.

Conway, K. W., Barrie, J. V., & Krautter, M. (2005). Geomorphology of unique reefs on the western Canadian shelf: sponge reefs mapped by multibeam bathymetry. *Geo-Marine Letters, 25*(4), 205-213.

Dempsey, S.J., E.M. Gese, G.M. Kluever, R.C. Lonsinger, and L.P. Waits. (2015). Evaluation of Scat Deposition Transects versus Radio Telemetry for Developing a Species Distribution Model for a Rare Desert Carnivore, the Kit Fox. *PLOS One 10*(10):e0138995.

Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F. T., D'Amen, M., Randin, C., Engler, R., Pottier, J., Pio, D., Dubius, A., Pellissier, L., Mateo, R. G., Hordijk, W., Salamin, N., & Guisan, A. (2017). ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography, 40*(6), 774-787.

Džeroski, S., & D. Drumm. (2003). Using regression trees to identify the habitat preference of the sea cucumber (Holothuria leucospilota) on Rarotonga, Cook Islands. *Ecological modeling, 170*, 219–226.

Elith, J., & Burgman, M.A. (2002a). Predictions and their validation: rare plants in the Central Highlands, Victoria, Australia. In: *Predicting Species Occurrences: Issues of Accuracy and Scale,* Editors: Scott, J.M., Heglund, P. J., Morrison, M. L., Raphael. M. G., Wall, W. A., & Samson, F. B. Island Press, Covelo, CA, 303-314.

Elith, J., Burgman, M. A., & Regan, H. M. (2002b). Mapping epistemic uncertainties and vague concepts in predictions of species distribution. *Ecological Modelling, 157*, 313–329.

Elith, J., Ferrier, S., Huettmann, F., & Leathwick, J. (2005). The evaluation strip: A new and robust method for plotting predicted responses from species distribution models. *Ecological Modelling, 186*, 280–289.

Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M. M., Peterson, A. T., Phillips, S. J., Richardson, K., Scachetti-Pereira, R. , Schapire, R. E., Soberón, J., Williams, S., Wisz, M. S., Zimmermann, N. E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography, 29*(2), 129–151.

Elith, J., Leathwick, J, R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology, 77*, 802-813.

Elith, J., & 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.

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions, 17*(1), 43-57.

Ferrier, S., Watson, G., Pearce, J., & Drielsma, M. (2002). Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. *Biodiversity and Conservation, 11*, 2275–2307.

Ferrier, S., & Guisan, A. (2006). Spatial modelling of biodiversity at the community level. Journal of *Applied Ecology, 43*(3), 393-404.

Fieldling, A., & Bell, J. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation, 24*(1), 38-49.

Findley, J. (2010). *Presence-only modeling with Maxent* (Rep.). Syracuse University.

Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. *PLoS ONE, 9*(5).

Franklin, J., & Miller, J. A. (2010). Mapping species distributions: spatial inference and prediction. Cambridge, UK: Cambridge University Press. 338 pp.

Friedman, J. H. Multivariate adaptive regression splines. (1991). *Ann. Stat 19,* 1–141.

Garcia, H. E., R. A. Locarnini, T. P. Boyer, J. I. Antonov, A. V. Mishonov, O. K. Baranova, M. M. Zweng, J. R. Reagan, D. R. Johnson, (2013a). World Ocean Atlas 2013. Vol. 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. S. Levitus, Ed.; A. Mishonov, Technical Ed. NOAA Atlas NESDIS 75, 27 pp

Garcia, H. E., R. A. Locarnini, T. P. Boyer, J. I. Antonov, O. K. Baranova, M. M. Zweng, J.R. Reagan, D. R. Johnson, (2013b). World Ocean Atlas 2013. Vol. 4: Dissolved Inorganic Nutrients (phosphate, nitrate, silicate). S. Levitus, Ed.; A. Mishonov, Technical Ed. NOAA Atlas NESDIS 76, 25 pp.

Guinotte, J. M., & Davies, A. J. (2014). Predicted Deep-Sea Coral Habitat Suitability for the U.S. West Coast. *PLoS ONE, 9*(4).

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling, 135*, 147–186.

Guisan, A., T.C. Edwards, Jr., and T. Hastie. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling 157*(2-3), 89–100.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters, 8,* 993-1009.

Guisan, A., & Rahbek, C. (2011). SESAM - a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *Journal of Biogeography, 38,* 1433-1444.

Hijmans, R. J., & Elith, J. (2017). Species distribution modeling with R. R Package Vignette.

Jiménez-Valverde, A., Lobo, J. M., & Hortal, J. (2008). Not as good as they seem: The importance of concepts in species distribution modelling. *Diversity and Distributions, 14*, 885–890.

Kent, R., & Carmel, Y. (2011). Presence-only versus presence-absence data in species composition determinant analyses. *Diversity and Distributions, 17*(3), 474-479.

Kestrup, Åsa M., Smith, D.L. and Therriault, T.W. (Eds.) (2015). Report of Working Group 21 on Non-indigenous Aquatic Species. PICES Sci. Rep. No. 48, 176 pp.

Knudby, A., Ledrew, E., & Brenning, A. (2010). Predictive mapping of reef fish species richness, diversity and biomass in Zanzibar using IKONOS imagery and machine learning techniques. *Remote Sensing of Environment, 114*(6), 1230-1241.

Knudby, A., Kenchington, E., & Murillo, F. J. (2013). Modeling the distribution of geodia sponges and sponge grounds in the northwest Atlantic. *PLoS ONE, 8*(12).

Landis, J. R., Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

Levesque, C., & Jamieson, G.S. (2015). Identification of Ecologically and Biologically Significant Areas in the Strait of Georgia and off the West Coast of Vancouver Island: Phase I - Identification of Important Areas. DFO Can. Sci. Advis. Sec. Res. Doc. 2014/100. viii + 68 p.

Leys, S. P. (2003). The significance of syncytial tissues for the position of the hexactinellida in the Metazoa. *Integrative and Comparative Biology, 43*(1), 19–27.

Leys, S., Wilson, K., Holeton, C., Reiswig, H., Austin, W., & Tunnicliffe, V. (2004). Patterns of glass sponge (Porifera, Hexactinellida) distribution in coastal waters of British Columbia, Canada. *Marine Ecology Progress Series, 283*, 133-149.

Leys, S. P., Yahel, G., Reidenbach, M. A., Tunnicliffe, V., Shavit, U., & Reiswig, H. M. (2011). The Sponge Pump: The Role of Current Induced Flow in the Design of the Sponge Body Plan. *PLoS ONE, 6*(12).

Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M., Paver, C. R., Reagan, J. R., Johnson, D. R., Hamilton, M., & Seidov, D. (2013). World Ocean Atlas 2013, Volume 1: Temperature. S. Levitus, Ed.; A. Mishonov, Technical Ed.; NOAA Atlas NESDIS 73, 40 pp.

Leys, S. P., & Kahn, A. S. (2018). Oxygen and the energetic requirements of the first multicellular animals. *Integrative and Comparative Biology, 58*(4), 666–676.

MacKenziem, D. I., (2005). Was it there? Dealing with imperfect detection for species presence/absence data. *Australian & New Zealand Journal of Statistics, 47*(1), 65-74.

Maldonado, M. (2006). The ecology of the sponge larva. *Canadian Journal of Zoology, 84*, 175–194.

Mi, C., Huettmann, F., Guo, Y., Han, X., & Wen, L. (2017). Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *Peer J, 5*:e2849

Monk, J., Ierodiaconou, D., Harvey, E., Rattray, A., & Versace, V. L. (2012). Are we predicting the actual or apparent distribution of temperate marine fishes? *PLoS ONE, 7*(4).

Monserud, R. A., Leemans, R. (1992). Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling, 62*, 275–293

Nedjah, N. & Luiza de Macedo, M. (2005). Fuzzy Systems Engineering: Theory and Practice; Springer: New York, NY, USA.

Nigam, K., Lafferty, J., & Mccallum, A. (1999). Using Maximum Entropy for Text Classification. IJCAI-99 Workshop on Machine Learning for Information Filtering, 61–67.

Pagel, J., & Schurr, F. M. (2012). Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography, 21*(2), 293–304.

Pearce, J., & Ferrier, S. (2000). An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling, 128*, 127–147.

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global Ecology and Biogeography, 12*(5), 361–371.

Pearson, R. G. (2010). Species' distribution modeling for conservation educators and practitioners. Network of Conservation Educators and Practitioners, Center for Biodiversity and Conservation, American Museum of Natural History. *Lesson in Conservation*, *3*, 54–89.

Pennino, M. G., Conesa, D., Lo´pez-Quı´lez, A., Mun˜oz, F., Ferna´ndez, A., & Bellido, J. M. (2016). Fishery-dependent and -independent data lead to consistent estimations of essential habitats. *ICES Journal of Marine Science, 73*(9), 2302–2310.

Perry, R. I., Stocker, M., & Fargo, J. (1994). Environmental effects on the distributions of groundfish in Hecate Strait, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, *51*, 1401–1409.

Phillips, S. (n.d.). A Brief Tutorial on Maxent (pp. 1-38, Rep.). AT&T Research.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modeling, 190*(3-4), 231-259.

Phillips, S.J., Dudik, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications, 19*, 181–197.

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems, 9*(2), 181-199

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Redding, D. W., Lucas, T. C. D., Blackburn, T. M., & Jones, K. E. (2017). Evaluating Bayesian spatial methods for modelling species distributions with clumped and restricted occurrence data. *PLoS ONE 12*(11): e0187602.

Reiss, H., Cunze, H., König, K., Neumann, K., & Kröncke, I. (2011). Species distribution modelling of marine benthos: A North Sea case study. *Marine Ecology Progress Series, 442*, 71–86.

Reiswig, H. M., & Mackie, G. O. (1983). Studies on Hexactinellid Sponges. III. The Taxonomic Status of Hexactinellida within the Porifera. *Philosophical Transactions of the Royal Society B: Biological Sciences, 301*, 419-428.

Rivera, O. R., & Lopez-Quilez, A. (2017). Development and comparison of species distribution models for forest inventories. *ISPRS International Journal of Geo-Information, 6*(6), 176.

Roberts, J. M., Long, D., Wilson, J. B., Mortensen, P. B., & Gage, J. D. (2003). The cold-water coral Lophelia pertusa (Scleractinia) and enigmatic seabed mounds along the north-east Atlantic margin: are they related? *Marine Pollution Bulletin 46*, 7–20.

Robinson, L. M., Elith, J., Hobday, A. J., Pearson, R. G., Kendall, B. E., Possingham, H. P., & Richardson, A. J. (2011). Pushing the limits in marine species distribution modelling: Lessons from the land present challenges and opportunities. *Global Ecology and Biogeography, 20*(6), 789-802.

Robinson, N. M., Nelson, W. A., Costello, M. J., Sutherland, J. E., & Lundquist, C. J. (2017). A Systematic Review of Marine-Based Species Distribution Models (SDMs) with Recommendations for Best Practice. *Frontiers in Marine Science, 4*(421).

Rooper, C. N., Zimmermann, M., & Prescott, M. M. (2017). Comparison of modeling methods to predict the spatial distribution of deep-sea coral and sponge in the Gulf of Alaska. *Deep Sea Research Part I: Oceanographic Research Papers*, *126*, 148–161.

Rowden, 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*:335.

Rudnick, D., Beier, P., Cushman, S., Dieffenbach, F., Epps, C.W., Gerber, L., Hartter, J., Jenness, J., Kintsch, J., Merenlender, A.M., Perkle, R.M., Preziosi, D.V., Ryan, S.J., and S. C. Trombulak. (2012). The Role of Landscape Connectivity in Planning and Implementing Conservation and Restoration Priorities. *Issues in Ecology.* Report No. 16. Ecological Society of America. Washington, DC.

Shabani, F., Kumar, L., & Ahmadi, M. (2016). A comparison of absolute performance of different correlative and mechanistic species distribution models in an independent area. *Ecology and Evolution, 6*(16), 5973–5986.

Steinacher, M., Joos, F., Frölicher, T. L., Plattner, G. K., & Doney, S. C. (2009). Imminent ocean acidification in the Arctic projected with the NCAR global coupled carbon cycle-climate model. *Biogeosciences*, *6*, 515–533.

Svensson, J.R., Jonsson, L., & Lindegarth, M. (2013). Excessive spatial resolution decreases performance of quantitative models, contrary to expectations from error analyses. *Marine Ecology Progress Series, 485*, 57-73.

Swets, K. A. (1988). Measuring the accuracy of diagnostic systems. *Science, 240*, 1285-1293.

Syfert, M. M., Smith, M. J., & Coomes, D. A. (2013). The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. *PLoS ONE, 8*(2).

Thuiller, W. (2003). BIOMOD – optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology, 9*, 1353–1362.

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography, 32*(3), 369-373.

Thuiller, W., Georges, D., Engler, R., & Breiner, F. (2016). biomod2: Ensemble Platform for Species Distribution Modeling. R package version 3.3-7. https://CRAN.R-project.org/package=biomod2

Thurber, A. R., Sweetman, A. K., Narayanaswamy, B. E., Jones, D. O., Ingels, J., & Hansman, R. L. (2014). Ecosystem function and services provided by the deep sea. *Biogeosciences, 11*(14), 3941-3963.

Treguer, P., Nelson, D, M., Van Bennekom, A. J., DeMaster, D. J., Leynaert, A., & Queguiner, B. (1995). The silica balance in the world ocean: a reestimate. *Science, 268*, 375–379.

Tsoar, A., Allouche, O., Steinitz, O., Rotem, D., & Kadmon, R. (2007). A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions, 13*, 397–405.

US Department of Commerce, & National Oceanic and Atmospheric Administration. (2018). Are glass sponges made of glass? Retrieved from https://oceanexplorer.noaa.gov/facts/glass-sponges.html

Václavík, T., J.A. Kupfer, and R.K. Meentemeyer. (2012). Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography 39*(1), 42-55.

Varela, S., Anderson, R. P., García-Valdés, R., & Fernández-González, F. (2014). Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography*, *37*(11).

Vaughan, I. P., & Ormerod, S. J. (2005). The continuing challenges of testing species distribution models. *Journal of Applied Ecology, 42*, 720-730.

Vieilledent, G., Merow, C., Guelat, J., Latimer, A., Kery, M., Gelfand, A., Wilson, A., Mortier, F., & Silander, Jr. J. (2014) Hierarchical Bayesian species distribution models with the hSDM R Package. URL http://hsdm.sourceforge.net/wp-content/uploads/2014/07/hSDM-vignette.pdf

Wang, Y., Naumann, U., Wright, S.T. & Warton, D.I. (2012). mvabund - an R package for model-based analysis of multivariate abundance data. *Methods Ecol. Evol.,* 3, 471–474.

Ward, G. (2007). Statistics in ecological modeling: presence-only data and boosted mars. Stanford University, Palo Alto.

Ward, G., Hastie, T., Barry, S., Elith, J., & Leathwick, J. R. (2009). Presence-only data and the EM algorithm. *Biometrics, 65*(2), 554–563.

Whitney, F., Conway, K., Thomson, R., Barrie, V., Krautter, M., & Mungov, G. (2004) Oceanographic habitat of sponge reefs on the Western Canadian continental shelf. *Continental Shelf Research.*

Wilson, M. F. J., O'Connell, B., Brown, C., Guinan, J. C., & Grehan, A. J. (2007). Multiscale Terrain Analysis of Multibeam Bathymetry Data for Habitat Mapping on the Continental Slope. *Marine Geodesy, 30*, 3-35.

Wintle, B. A., McCarthy, M. A., Parris, K. M., & Burgman, M. A. (2004). Precision and bias of methods for estimating point survey detection probabilities. *Ecological Applications, 14*(3), 703-712.

Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dormann, C. F., Forchhammer, M. C., Grytnes, J. A., Guisan, A., Heikkinen, R. K., Høye, T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M. C., Normand, S., Öckinger, E., Schmidt, N. M., Termansen, M., Timmermann, A., Wardle, D. A., Aastrup, P., & Svenning, J. C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews, 88*, 15–30.

Zhang, L., Liu, S., Sun, P., Wang, T., Wang, G., Zhang, X., & Wang, L. (2015). Consensus forecasting of species distributions: The effects of niche model performance and niche properties. *PLoS ONE, 10*(3).

Mapping biogenic habitats: distribution of glass sponge reefs and key variables likely to influence their condition

Anya Dunham

Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC, Canada

Background

Glass sponge bioherms, or reefs, are unique biogenic habitats found along the coasts of western Canada and the United States. Sponge reefs were known only from fossil records until analogous live reefs were discovered in Hecate Strait and Queen Charlotte Sound in the 1980s (Conway et al., 1991). More recently, the reefs were found in Southeast Alaska (Stone et al., 2014) and in Chatham Sound near the border between Canada and Alaska (Shaw et al., 2018). A number of smaller reefs have also been discovered in the Strait of Georgia and Howe Sound (Conway et al., 2005, 2007; Cook et al., 2008; Chu and Leys, 2010; Clayton and Dennison, 2017; DFO, 2018; Dunham et al., 2018), hereinafter referred to as the Salish Sea.

The reefs in the Salish Sea are built by the Dictyonine glass sponges *Aphrocallistes vastus* and *Heterochone calyx*. These sponges possess rigid three-dimensional skeletal frameworks that remain intact after the sponges' death (Leys et al., 2007). The reefs are formed when larval sponges settle onto exposed skeletons of dead sponges and fine sediments entrained in bottom currents are baffled and trapped by the dead reef matrix, solidifying reef structure (Leys et al., 2004; Krautter et al., 2006). The bulk of the reef thus consists of dead sponge matrix cemented by sediments, with only the most recent generation growing 1–2m above the reef surface (Conway et al., 2001).

Several studies have provided insight into the ecosystem function of the glass sponge reefs. The reefs contribute to the productivity of benthic ecosystems by forming habitat for diverse communities of invertebrates and fish (Chu and Leys, 2010; Cook et al., 2008; Dunham et al., 2015; Marliave et al., 2009), act as regionally important silica sinks (Chu et al., 2011; Tréguer and De La Rocha, 2013), and, being one of the densest known communities of deep-water filter feeders, link benthic and pelagic environments through carbon and nitrogen processing (Kahn et al., 2015). However, understanding of the reefs' overall role in the Salish Sea ecosystem was limited, as most of the empirical work to date has been constrained to a few well-studied areas.

Research summary

Members of PICES Working Group 32 and collaborators undertook research to map glass sponge reefs in the Salish Sea, quantify their condition and ecosystem function, describe biodiversity associated with the reefs, and identify drivers behind live sponge cover variation (Dunham et al., 2018a,b; DFO, 2018; Conway et al., 2019). Standardized visual surveys using a Remotely Operated Vehicle (ROV) were undertaken and quantitative assessments of all known reef complexes in the Strait of Georgia (Dunham et al., 2018a), as well as the newly discovered and mapped reefs in Howe Sound (DFO, 2018) were completed. An information-theoretic approach was used to examine six seabed terrain characteristics (rugosity, slope, curvature, broad and fine bathymetric position indices [BBPI and FBPI, respectively], and minimum depth surveyed), bottom current estimates, and two measures of potential human impact (density of anthropogenic objects and past fishing pressure) as predictor variables for live reef-building sponge percent cover.

The reefs varied widely in their estimated sponge cover: mean live reef-building sponge cover ranged from 0.2% to 17.5% (Fig. 1A) and dead sponge cover ranged from 0.1% to 42% (Fig. 1C). The frequency of occurrence of habitat categories also varied between reef complexes (Fig. 1D). Sponge rubble was observed in all reefs; rubble cover ranged from 0.1% to 14% (Fig. 1B).

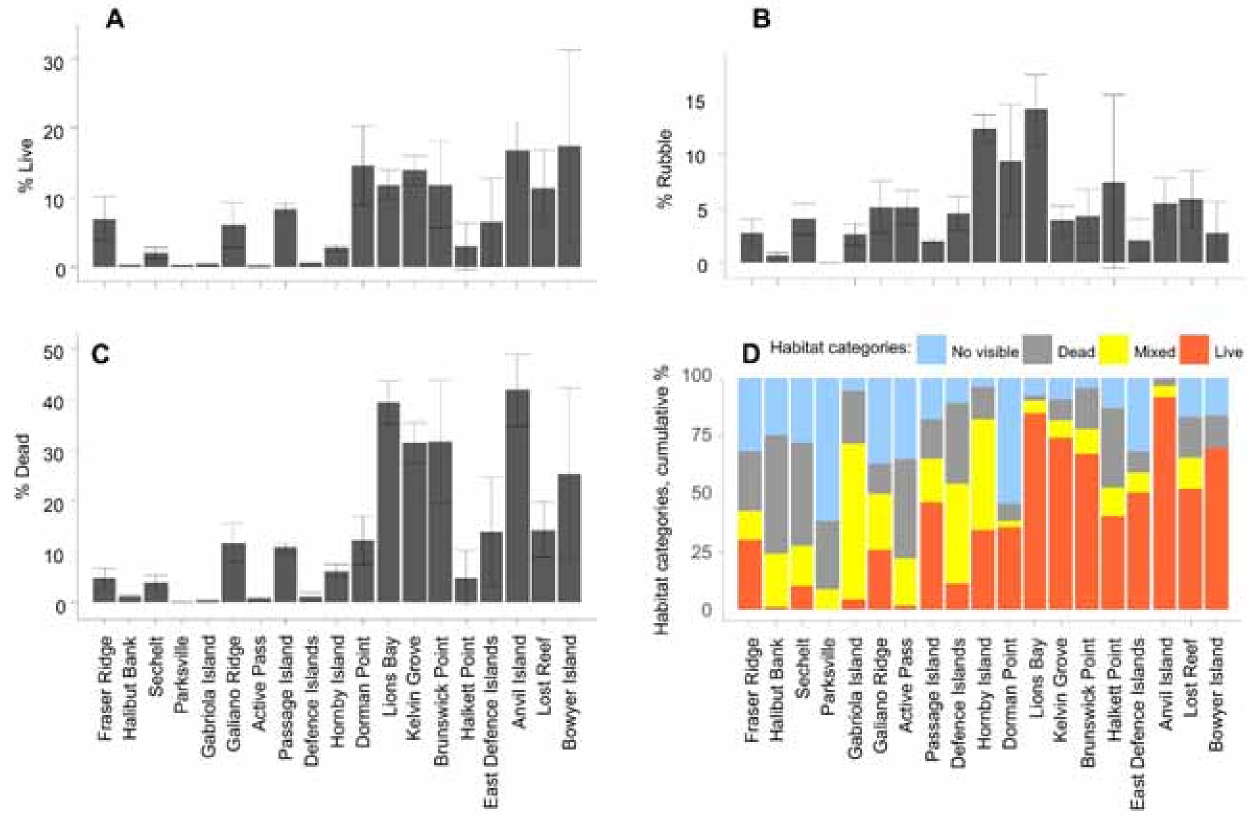


Figure 1. Reef-building sponge cover and condition across 19 reefs: (A) live reef-building sponge percent cover per image, mean ± 95% confidence Interval, (B) sponge rubble percent cover per image, mean ± 95% confidence Interval (C) dead reef-building sponge percent cover per image, mean ± 95% confidence Interval, and (D) frequencies of occurrence of habitat categories per reef. Reproduced from Dunham et al. (2018a).

The reefs were found to support diverse and abundant communities of invertebrates and fish, with 115 unique taxonomic groups observed (Dunham et al., 2018a, online supplement S7). Expected species richness differed between reef complexes and ranged between 18 and 69.

For live reef-building sponge percent cover, the best fit model that explained 75% of the variation included the following explanatory variables: range of rugosity, curvature, minimum depth, and the interactions of range of rugosity with curvature and depth. In general, reefs with a broader range of rugosity had higher live sponge cover. However, for reefs located on strongly concave slopes (i.e. negative profile curvature values), increased rugosity range led to lower live sponge cover. The deeper the reef was located the more concave the slope had to be before an increase in rugosity range switched from increasing to decreasing live sponge cover (Fig. 2).

Overall, differences in live cover appeared to be largely driven by seabed terrain characteristics.

In general, reefs with a broader range of rugosity exhibited higher live sponge cover. This may be explained by the more rugose seafloor generating local turbulence that may deliver nutrient-rich water to filter-feeding sponges. In addition, more rugose areas may promote infilling of the dead reef matrix with sediment which is crucial for supporting the reef mass and preventing sponge skeletons from dissolving in ambient seawater over time (Krautter et al., 2006), thus enabling reef existence and growth. However, for reefs located on strongly concave slopes, and especially those in relatively shallow waters, increased rugosity led to lower live sponge cover. These areas may be experiencing turbulence and suspended sediment concentrations that are too high to support reef-building glass sponges. Our results support the conclusion that glass sponge reefs require a delicate balance of turbidity and suspended sediment concentration.

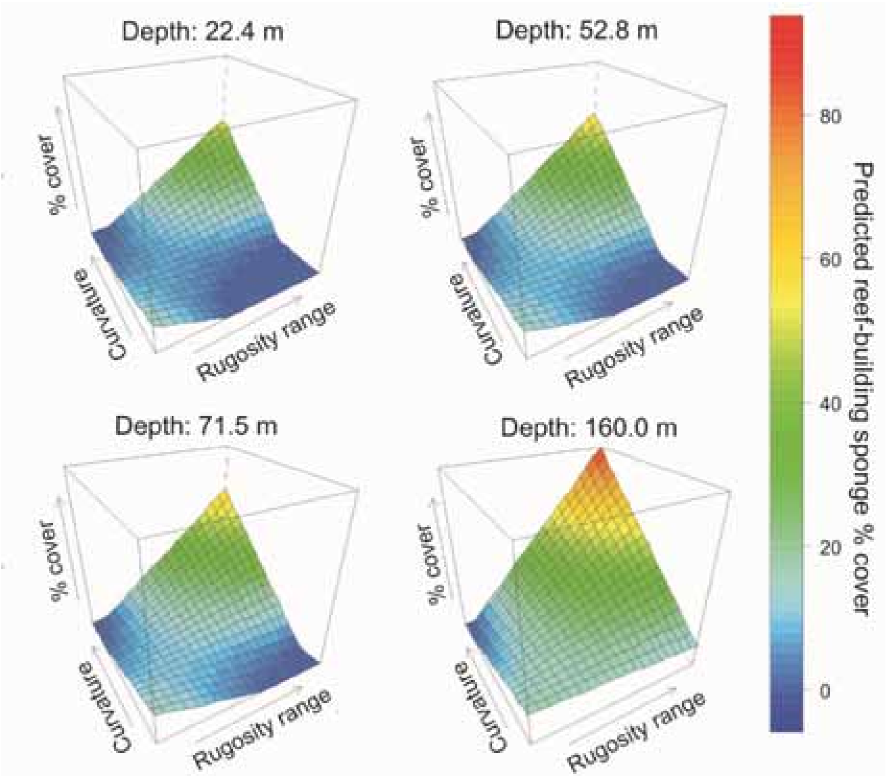


Figure 2. Predicted live reef-building sponge cover from the model with seabed terrain variables that explained 75% of variation between reefs: % cover * rugosity-range + curvature + depth+ rugosity-range * curvature + rugosity-range* depth. Predicted values less than 0 are displayed as 0% cover for clarity. Curvature gradient is from concave to convex. Reproduced from Dunham et al. (2018a).

Members of PICES Working Group 32 and collaborators also reviewed glass sponge reef geological expressions (geomorphologies) across all known locations in the Pacific Ocean and combined these observations with available ecological datasets on sponge cover (Conway et al., 2019). The reef morphologies that develop over time are remarkably variable. We found a significant association between reef morphologies and the patch size of live reef habitat and live reef patch isolation. Ridge and bioherm reef morphologies support larger patches of live reef habitat surrounded by distinct areas of dead or dead and buried reef. Conversely, the small wave and thin biostrome reef morphologies typically have many small patches of live reef habitat surrounded by mixed live and dead reef. This supports the hypothesis that there is a link between the geologic reef morphology and the distribution of suitable habitat for live sponges within the reef. However, regardless of the reef morphology, live reef occurred more commonly on the slopes up to the promontories of the reef form. Overall, the development of the diverse sponge reef morphologies and associated reef habitats appear to be driven by both physical and biological factors.

Reef-forming glass sponges are long-lived, but slow growing, exceptionally fragile, and thus slow to recover, especially from impacts that damage the reef's skeletal framework. Most reefs have been impacted by bottom-contact fishing activities. Using recent visual surveys conducted in Hecate Strait, Chatham Sound, and the Georgia Basin as reference baseline (2010-2017) and aggregating surveyed habitats at the geomorphic feature level (i.e. all BC sponge reefs), glass sponge reefs can be assigned a condition score of 4 (poor) following classification of Ward (2011). It is important to note, however, that condition varies widely between reefs, and that the current level of glass sponge reef ecology and ecosystem function knowledge is not sufficient to confidently and comprehensively define and assess reef health. Therefore, this condition score should be interpreted with caution. The reef habitat trend over the last 5 years can be best characterized as stable. Because temporal data to support a trend in reef status requires a time frame of over 5 years (Dunham et al., 2018b), this estimate is based on expert judgement and associated confidence level is low. Integrated, comprehensive monitoring that employs relevant metrics of reef health at appropriate spatial and temporal scales and provides well-resolved time series is necessary to further understand sponge reef ecosystems, to assess the effectiveness of recently enacted protection measures (MPA, fishing closures), and to ensure adaptive management responsive to the state of the protected areas (Conway et al., 2019).

References

Chu J.W.F., Leys S.P. 2010. High resolution mapping of community structure in three glass sponge reefs (Porifera, Hexactinellida). *Marine Ecology Progress Series* 417: 97–113

Chu J.W.F., Maldonado M., Yahel G., Leys S.P. 2011. Glass sponge reefs as a silicon sink. *Marine Ecology Progress Series* 441: 1–14

Clayton L., Dennison G. 2017. Inexpensive video drop-camera for surveying sensitive benthic habitats: Applications from glass sponge (Hexactinellida) reefs in Howe Sound, British Columbia. *Canadian Field-Naturalist* 131(1): 46–54

Conway K., Barrie J., Austin W, Luternauer J. 1991. Holocene sponge bioherms on the western Canadian continental shelf. *Continental Shelf Research* 11: 771–790

Conway K.W., Krautter M., Barrie J.V., Neuweiler, M. 2001. Hexactinellid sponge reefs on the Canadian continental shelf: A unique" living fossil". *Geoscience Canada* 28(2): 71–78

Conway K.W., Barrie J.V., Krautter M. 2005. Geomorphology of unique reefs on the western Canadian shelf: sponge reefs mapped by multibeam bathymetry. *Geo-Marine Letters* 25:205-213

Conway K.W., Barrie J.V., Hill P.R., Austin W.C., Picard K. 2007. Mapping sensitive benthic habitats in the Strait of Georgia, coastal British Columbia: deep-water sponge and coral reefs. *Geological Survey of Canada Current Research* 2007-A2: 1–6

Conway K, Dunham A, Archer S, Burke L, Kung R, and Shaw J. 2019. Pacific Margin Sponge Reefs: Geomorphology and Ecology. In: Seafloor Geomorphology as Benthic Habitat: GeoHab Atlas of seafloor geomorphic features and benthic habitats (2nd edition). Eds: P. Harris, E. Baker. Elsevier Science, USA. 1076 pp.

Cook S.E., Conway K.W., Burd B. 2008. Status of the glass sponge reefs in the Georgia Basin. *Marine Environmental Research 66* (Suppl 1): S80–S86

DFO. 2018. Glass sponge aggregations in Howe Sound: locations, reef status, and ecological significance assessment. DFO Can. Sci. Advis. Sec. Sci. Resp. 2018/032.

Dunham A., Pegg J., Carolsfeld W., Davies S., Murfitt I., Boutillier J. 2015. Effects of submarine power transmission cables on a glass sponge reef and associated megafaunal community. *Marine Environmental Research* 107: 50–60

Dunham A, Archer SK, Davies S, Burke L, Mossman J, Pegg J. 2018a. Assessing ecological role of deep-water biogenic habitats: Glass sponge reefs in the Salish Sea. *Marine Environmental Research* 141: 88-99.

Dunham A., Mossman J., Archer S., Davies S., Pegg J., Archer E. 2018b. Glass sponge reefs in the Strait of Georgia and Howe Sound: Status assessment and ecological monitoring advice. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/010.

Kahn A.S., Yahel G., Chu J.W., Tunnicliffe V., Leys S.P. 2015. Benthic grazing and carbon sequestration by deep-water glass sponge reefs. *Limnology and Oceanography* 60: 78–88

Krautter M., Conway K.W., Barrie J.V. 2006. Recent hexactinosidan sponge reefs (silicate mounds) off British Columbia, Canada: frame-building processes. *Journal of Paleontology* 80: 38–48.

Leys S.P., Mackie G.O., Reiswig H.M. 2007. The biology of glass sponges. *Advances in Marine Biology* 52: 1–145

Marliave J.B., Conway K.W., Gibbs D.M., Lamb A., Gibbs C. 2009. Biodiversity and rockfish recruitment in sponge gardens and bioherms of southern British Columbia, Canada. *Marine Biology* 156: 2247–2254

Shaw J., Conway K., Wu Y., Kung R. 2018. Distribution of hexactinellid sponge reefs in the Chatham Sound region, British Columbia Geological Survey of Canada. 10.4095/306310.

Stone R.P., Conway K., Csepp D., Barrie J. 2014. The boundary reefs: Glass sponge (Porifera: Hexactinellidae) reefs on the international border between Canada and the United States. NOAA Technical Memorandum NMFS-AFSC 264.

Tréguer P.J., De La Rocha C.L. 2013. The world ocean silica cycle. *Annual Review of Marine Science* 5: 477–501

Ward T.J., 2011. SOE 2011 National marine condition assessment – decision model and workshops. Report prepared for the Department of Sustainability, Environment, Water, Population and Communities on behalf of the State of the Environment 2011 Committee. Department of Sustainability, Environment, Water, Population and Communities Canberra, ACT, Australia, 22 p.

Potential indicators for assessing and monitoring diversity of biogenic habitats

Anya Dunham

Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC, Canada

Background

Understanding the ecology and functioning of various biogenic habitats has shed light on their vital importance for the continued conservation of natural resources and associated ecosystem services. As resource management moves toward holistic, habitat-based approaches such as ecosystem-based fisheries management (Pikitch et al., 2004), there is a growing need for management actions that promote healthy biogenic habitats. These management actions require ecological monitoring with clear research questions, appropriate indicators, and a well-designed data collection process to produce robust data and useful outcomes (Underwood & Chapman, 2013). A recently developed framework for biological monitoring (Reynolds, Knutson, Newman, Silverman, & Thompson, 2016) offered an overarching view of the steps required for successful monitoring programs and emphasized the importance of linkages among various planning decisions. However, for data-limited biogenic habitats, designing monitoring programs can be challenging. The structure and functioning of such habitats and underlying ecosystem-level processes (e.g., spatial extent, magnitude of natural variability in abundance and distribution of foundation species, species-habitat associations) are not well understood, often due to these habitats being remote and/or deep, limiting accessibility, and increasing monitoring costs. These limitations may jeopardize effective management and conservation of these habitats, many of which are threatened by human activities (Rossi, Bramanti, Gori, & Orejas Saco del Valle, 2017). A clear road map for designing robust, efficient monitoring programs in the face of data and resource limitations was required.

Research summary

Members of PICES Working Group 32 and collaborators reviewed recent publications (2012-2017) to obtain an overview of benthic assessment and monitoring approaches, methods, and indicators across a range of relatively well-studied marine biogenic habitats. Common themes relevant for all habitat types were identified and, drawing upon these themes, a systematic approach for establishing monitoring programs for data-limited biogenic habitats was developed (Loh et al., 2019).

Biogenic habitat monitoring efforts largely focus on the characteristics, distribution, and ecological function of foundation species, but may target other habitat-forming organisms, especially when community shifts are observed or expected, as well as proxies of habitat status, such as indicator species. Broad-scale methods cover large spatial areas and are typically used to examine the spatial configuration of habitats, whereas fine-scale methods tend to be laborious and thus restricted to small survey areas but provide high-resolution data. Recent, emerging methods enhance the capabilities of surveying large areas at high spatial resolution and improve data processing efficiency, bridging the gap between broad- and fine-scale methods. Although sampling design selection may be limited by habitat characteristics and available resources, it is critically important to ensure appropriate matching of ecological, observational, and analytical scales.

Across biogenic habitat types, assessment and monitoring efforts share the following common themes: defining study objectives, assembling preliminary data, determining scale of interest, selecting indicators, determining study methods and sampling design, and full protocol review. Drawing on these common themes, we propose a structured, iterative approach to designing monitoring programs for marine biogenic habitats that allows for rigorous data collection to inform management strategies, even when data and resource limitations are present (Fig. 3).

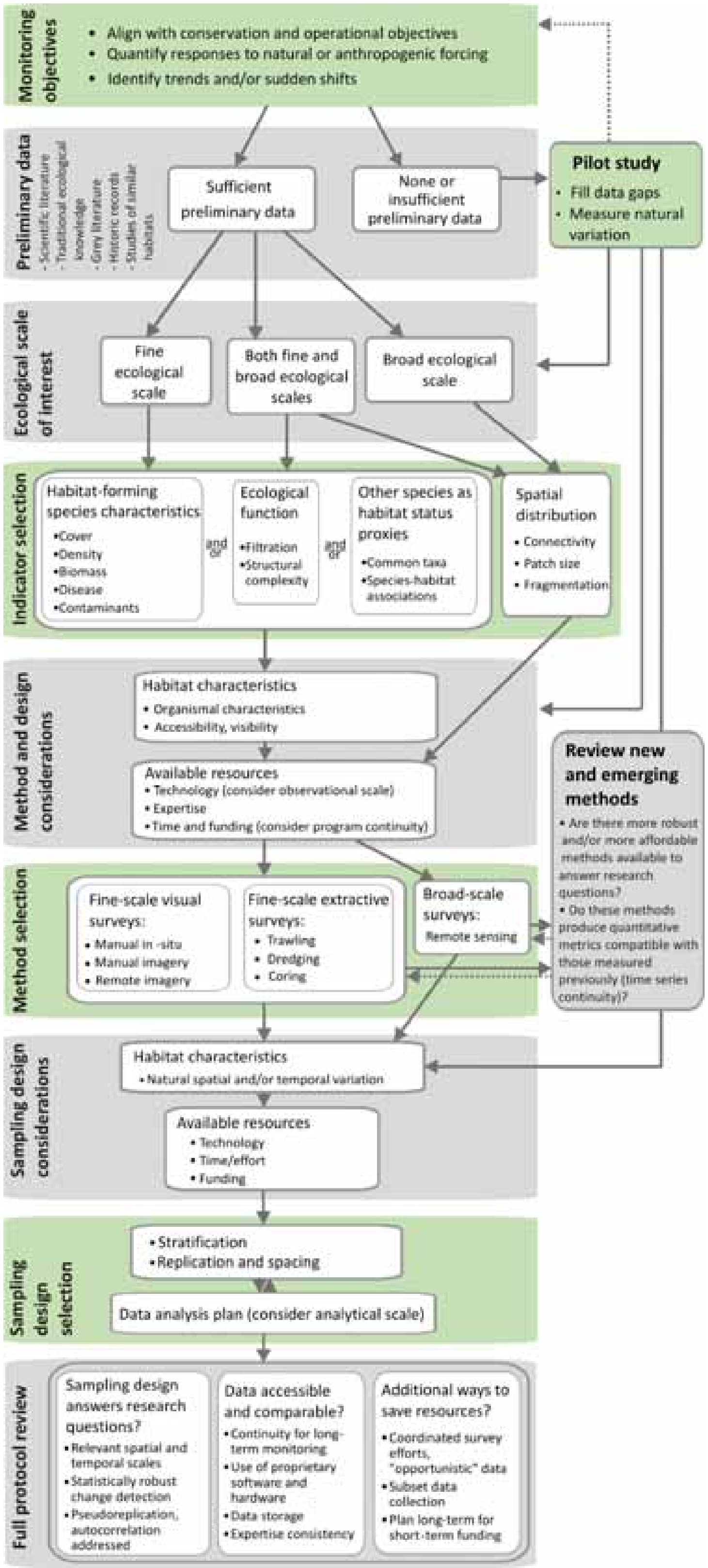


Figure 3. Flowchart illustrating the development of a monitoring protocol for marine biogenic habitats. Gray boxes contain considerations (to facilitate evaluation of available data, methods,

and protocols), and green boxes denote decisions. Key times for iteration back through earlier steps are denoted by the return dashed arrows. Reproduced from Loh et al. (2019).

As part of this project, members of PICES Working Group 32 and collaborators also produced (1) an overview of the applicability, advantages, and disadvantages of broad- and fine-scale survey methods for biogenic habitats; and (2) a summary of attributes, representative indicators, and associated metrics used for biogenic habitat assessments, with corresponding survey methods (Loh et al. 2019). These summaries can be used for developing monitoring programs for any type of biogenic habitat in the North Pacific and elsewhere in the world.

In oceans impacted by human pressures, biogenic habitat assessment and monitoring are crucial for attributing causes of decline and for providing solutions to mitigate habitat damage from anthropogenic impacts and monitoring environmental change (Downs et al., 2005). Systematic monitoring approaches, as laid out in Loh et al. (2019), are urgently required to implement science-based management, evaluate the success of protective measures, and guide adaptive management strategies for data-limited marine biogenic habitats.

References

Downs C.A., Woodley C.M., Richmond R.H., Lanning L.L., Owen R. 2005. Shifting the paradigm of coral-reef “health” assessment. *Marine Pollution Bulletin* 51: 486–94.

Loh T-L., Archer S.K., Dunham A. 2019. Monitoring program design for data-limited marine biogenic habitats: A structured approach. *Ecol. Evol*. 9: 7346–7359.

Pikitch E.K., Santora C., Babcock E.A., Bakun A., Bonfil R., Conover D.O., … Sainsbury K.J. 2004. Ecosystem-based fishery management. *Science* 305: 346–347.

Reynolds J.H., Knutson M.G., Newman K.B., Silverman E.D., Thompson W.L. 2016. A road map for designing and implementing a biological monitoring program. *Environmental Monitoring and Assessment* 188: 399.

Rossi S., Bramanti L., Gori A., Orejas Saco del Valle C. (eds.) 2017. Marine Animal Forests: The Ecology of Benthic Biodiversity Hotspots. Springer International Publishing. Online ISBN: 978-3-319-17001-5.

Underwood A.J., Chapman M.G. 2013. Design and analysis in benthic surveys in environmental sampling. In A. Eleftheriou (Ed.), *Methods for the Study of Marine Benthos* (4th Ed., pp. 1–45). Oxford, UK: John Wiley & Sons, Ltd.

Climate change and the distribution of habitat-forming shallow-water corals

Hiroya Yamano[1] and Naoki H. Kumagai

[1]National Institute for Environmental Studies, Japan.

Background

Tropical and subtropical islands are associated with coral reefs, which provide ecosystem services, including fisheries, tourism and coastal protection. This is especially true to reef islands that are composed fully of reef-derived materials. Global-scale environmental changes, including climate change, have been causing significant change on corals. Japan provides an ideal setting to examine these changes, because it covers a wide latitudinal range, stretching from subtropical to temperate areas, and latitudinal limits of coral reefs and coral distributions are observed around the Japanese islands.

Seas around Japan showed significant sea surface temperature (SST) rises (0.8°C–1.3°C/100 years in annual mean values) (http://www.data.jma.go.jp/kaiyou/data/shindan/a_1/japan_warm/japan_warm.html), and the rising SSTs could have caused two consequences on Japanese corals: decline in the south due to anomalously high SSTs in summer that caused coral bleaching, and range expansion in the north due to rising winter SSTs that allowed survival of warm-water corals in winter. The latter may be associated with decline of macroalgae.

Research Summary

Members of PICES Working Group 32 and collaborators have worked on the detection and projection of these issues as follows:

Bleaching in the south (Kumagai et al., 2018b)

Excessive SSTs can cause coral bleaching, resulting in coral death and decreases in coral cover. A SST threshold of 1˚C over the climatological maximum is widely used to predict coral bleaching. In this study, we refined thermal indices predicting coral bleaching at high-spatial resolution (1 km) by statistically optimizing thermal thresholds, as well as considering other environmental influences on bleaching such as ultraviolet (UV) radiation, water turbidity, and cooling effects. We used a coral bleaching dataset derived from the web-based monitoring system Sango Map Project, at scales appropriate for the local and regional conservation of Japanese coral reefs. We recorded coral bleaching events in the years 2004–2016 in Japan. We revealed the influence of multiple factors on the ability to predict coral bleaching, including selection of thermal indices, statistical optimization of thermal thresholds, quantification of multiple environmental influences, and use of multiple modeling methods (generalized linear models and random forests). After optimization, differences in predictive ability among thermal indices were negligible. Thermal index, UV radiation, water turbidity, and cooling effects were important predictors of the occurrence of coral bleaching. Predictions based on the best model revealed that coral reefs in Japan have experienced recent and widespread bleaching.

Range shift in the north (Kumagai et al., 2018a)

Coral and macroalgal communities are threatened by global stressors. However, recently reported community shifts from temperate macroalgae to tropical corals offer conservation potential for corals at the expense of macroalgae under climate warming. Although such community shifts are expanding geographically, our understanding of the driving processes is still limited. Here, we reconstruct long-term climate-driven range shifts in 45 species of macroalgae, corals, and herbivorous fishes from over 60 years of records (mainly 1950–2015), stretching across 3,000 km of the Japanese archipelago from tropical to subarctic zones. Based on a revised coastal version of climate velocity trajectories, we found that prediction models combining the effects of climate and ocean currents consistently explained observed community shifts significantly better than those relying on climate alone. Corals and herbivorous fishes performed better at exploiting opportunities offered by this interaction. The contrasting range dynamics for these taxa suggest that ocean warming is promoting macroalgal-to-coral shifts both directly by increased competition from the expansion of tropical corals into the contracting temperate macroalgae, and indirectly via deforestation by the expansion of tropical herbivorous fish. Beyond individual species' effects, our results provide evidence on the important role that the interaction between climate warming and external forces conditioning the dispersal of organisms, such as ocean currents, can have in shaping community level responses, with concomitant changes to ecosystem structure and functioning. Furthermore, we found that community shifts from macroalgae to corals might accelerate with future climate warming, highlighting the complexity of managing these evolving communities under future climate change.

Literature cited

Kumagai, N.H., García Molinos, J., Yamano, H., Takao, S., Fujii, M., and Yamanaka, Y. (2018a) Ocean currents and herbivory drive macroalgal-coral community shift under climate warming. *Proceedings of the National Academy of Sciences of the United States of America*, 115, 8990-8995.

Kumagai, N.H., Yamano, H., and Sango Map Project Committee (2018b) High-resolution modeling of thermal thresholds and environmental influences on coral bleaching for local and regional reef management. *PeerJ*, 6: e4382; DOI 10.7717/peerj.4382

Selection of the proper spatial resolution for habitat modeling of cold-water corals

Mai Miyamoto[1], Masashi Kiyota[2], Hiroto Murase[3], Takeshi Nakamura[4], Takeshi Hayashibara[5]

[1]Environmental Consulting Department
Japan NUS Co. Ltd.
Nishi-Shinjuku Prime Square 5F,
7-5-25 Nishi-Shinjuku, Shinjuku-Ku, Tokyo 160-0023, Japan

[2]Graduate School of Fisheries and Environmental Sciences Nagasaki University
1-14 Bunkyo-machi, Nagasaki 852-8521, Japan

[3]School of Marine Resources and Environment Tokyo University of Marine Science and Technology
4-5-7 Konan Minato-ku, Tokyo 108-8477, Japan

[4]Department of Fisheries Distribution and Management National Fisheries University, Japan Fisheries Research and Education Agency
2-7-1 Nagata-Honmachi, Shimonoseki 759-6595, Japan.

[5]Research Center for Subtropical Fisheries Seikai National Fisheries Research Institute, Japan Fisheries Research and Education Agency 148, Fukai-Ohta, Ishigaki-shi, Okinawa, 907-0451, Japan

Background

Cold-water corals are benthic cnidarians that generally inhabit deep-sea floors of the world ocean. Some deep-sea corals form complex, reef-like structures and provide habitats for other animals. Due to their slow growth, long life span, and slow recovery from physical damage, cold-water corals are claimed as important components of vulnerable marine ecosystems (VMEs). The assessment and management of impacts on VMEs are urgent tasks for deep sea bottom fisheries to fulfill the global requests for ecosystem-based fishery management. Habitat suitability modeling of cold-water corals has been used to estimate the priority areas of conservation interests. Habitat suitability modeling of cold-water corals are conducted at various spatial resolutions partially due to the limited availability of data from the deep-sea environment. Whereas coarse spatial resolutions, such as 1-degree grid-cells, are helpful for broader strategic consideration of the selection of priority areas for conservation on regional and global scales, fine spatial resolutions are required for tactical management at local scale. Members of PICES Working Group 32 and collaborators examined the effects of spatial resolution of bathymetric data on habitat suitability modeling of cold-water corals at a local scale, namely on seamounts. Large cold-water gorgonian corals on the southern Emperor Seamounts in the high seas of the central North Pacific Ocean were used as an example for planning the spatial management of the deep-sea habitat. The grid-cell size of the bottom bathymetry raster generated from the multi-beam echo sounder data was manipulated to investigate the effects of the spatial resolution of bathymetric data on terrain attributes and habitat suitability modeling of these corals.

Research Summary

The WG32 members and collaborators have worked on the consideration of spatial resolutions as follows (Miyamoto et al., 2017). Species occurrence data and high-resolution multi-beam bathymetry data were collected by ship-borne surveys in the Emperor Seamounts area in 2009 - 2013. Depth and terrain parameters were generated at six different grid cell sizes from 25 x 25 m to 800 x 800 m and used as environmental variables for habitat analysis (Fig 1). The values of terrain parameters showed different patterns at smaller (≤ 100 m) and larger (> 100 m) grid cell sizes (Fig 2). Accordingly, the topographic structures expressed by the raster maps changed with the grid cell sizes. Maxent habitat models showed higher prediction accuracy at smaller grid cell sizes, and predicted high habitat suitability at such locations as ridges on upper slopes and terrace edges and surface undulation on seamount tops, suggesting the importance of sloped and/or irregular sea floor as habitat of large gorgonian corals. Within the available data, the model of 25 m grid-cell resolution showed the best performance in habitat modeling of cold-water corals on the Emperor Seamounts. The sea floor structures predicted to be suitable for large gorgonians were consistent with the biological characteristics of large gorgonians. It was confirmed that MaxEnt models gave satisfactory performance at smaller grid-cell sizes (≦ 200 m). The variation in terrain attribute values also differed between smaller (≦ 100 m) and larger (≧ 200 m) grid-cell sizes. These results demonstrate that it is desirable to obtain bathymetric grid data at resolutions of 100 m or less for the purpose of predicting the distributions of corals at a local scale (e.g. within a seamount). It is important to decide the optimum spatial resolution in consideration of the objective of analysis, data availability and geographical or biological characteristics.

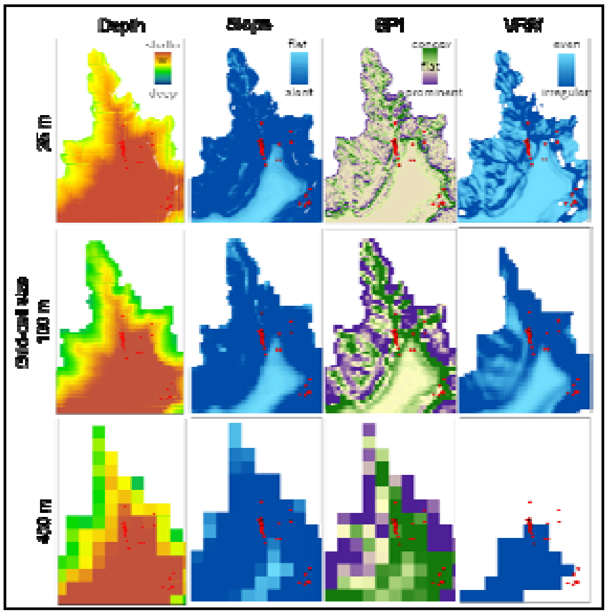


Figure 1. Example of maps of depth and three terrain attributes [slope, bathymetric position index (BPI) and vector ruggedness measure (VRM)] at three grid-cell sizes for the Colahan seamount. Red dots represent location where large gorgonians were observed or collected. Reproduced from Miyamoto et al. (2017).

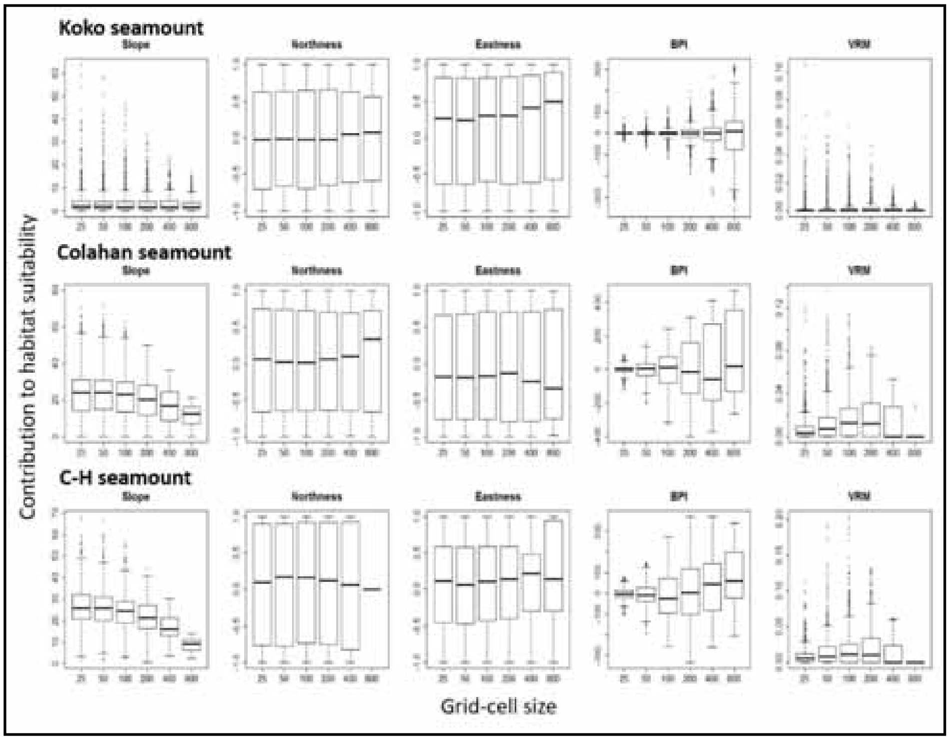


Figure 2. Boxplots of terrain parameter as six grid-cell sizes on the three seamounts studied. Boxplots show median (thick black lines), range (dashed lines), upper and lower fifth and 95th percentiles (boxes), and outliers (points). X-axis indicates grid-cell sizes. BPI, bathymetric position index; VRM, vector ruggedness measure. Reproduced from Miyamoto et al. (2017).

Literature cited

Miyamoto, M., Kiyota, M., Murase, H., Nakamura, T., and Hayashibara, T. (2017) Effects of bathymetric grid-cell sizes on habitat suitability analysis of cold-water gorgonian corals on seamounts. Marine Geodecy, 40, 205-223.

Method for assessing the validity of the VME indicator taxa in the Emperor Seamounts area

Mai Miyamoto[1] and Masashi Kiyota[2]

[1]Environmental Consulting Department
Japan NUS Co. Ltd.
Nishi-Shinjuku Prime Square 5F,
7-5-25 Nishi-Shinjuku, Shinjuku-Ku, Tokyo 160-0023, Japan

[2]Graduate School of Fisheries and Environmental Sciences Nagasaki University
1-14 Bunkyo-machi, Nagasaki 852-8521, Japan

Background

Some species of cold-water corals are known to provide biogenic habitats to other animals on the deep sea-floor in several areas of the world ocean. Due to their slow growth, long life span and slow recovery from physical damage, as well as the habitat forming property, cold-water corals are considered as important components of vulnerable marine ecosystems (VMEs). In the Emperor Seamounts area of the North Pacific Ocean, four orders of cold-water corals; i.e., Gorgonians (Scleraxonia, Holaxonia and Calcaxonia), Alcyonacea (excluding Gorgonians), Antipatharia, and Scleractinia, were selected as VME indicator taxa by the North Pacific Fisheries Commission (NPFC), and the bycatch threshold was set at a tentative value similar to other Regional Fisheries Management Organizations (RFMOs). However, appropriateness of these indicator species has not been examined scientifically. In this study, benthic samples collected through bottom surveys from 2009 to 2014 were identified and listed for delineating the characteristics of benthic animal fauna in the Emperor Seamounts area. The association analysis was applied to examine the validity of candidate VME indicator taxa as the biodiversity indicator through examination of their co-occurrence with other benthic animals.

Research Summary

Members of PICES Working Group 32 and collaborators have introduced a new method for assessing the validity of VME indicator taxa as follows (Miyamoto and Kiyota, 2017). The benthos samples collected by R/V Kaiyo-maru in the past six years were used to evaluate the effectiveness of four orders of cold-water corals (i.e., Gorgonians, Alcyonacea, Antipatharia and Scleractinia) and other benthic animals as VME indicator taxa in the Emperor Seamounts area. The benthos samples were identified and listed, and their occurrence frequencies and total weights were calculated (Fig 1).

Then association analysis, which is often used for discovering hidden relationships among purchased items in market transaction data (Blattberg et al., 2008), was applied to the occurrence data per sampling haul to explore the co-occurring relationships of benthic animals. For example, an association rule which indicates that many customer purchasing item A also purchase item B is expressed in the form of {A} → {B} (Silverstein et al., 1998, Hahsler et al., 2005). In this expression, {A} is called the antecedent part and {B} is called the consequent part. In this study, such association rules were explored that indicate co-occurrence of benthic taxa such as “A habitat where benthos taxon A occurs is also inhabited by taxon B”. The haul-by-haul occurrence data were applied to the association analysis, and those rules that include one of the six candidate VME indicator taxa, i.e., Gorgonians, Alcyonacea (sea fans), Antipatharia (black corals), Scleractinia (stone corals), Stylasterina (hydrocorals) or Porifera (sponges), in the consequent part were extracted. The effectiveness of an association rule is evaluated by the values of Support, Confidence and Lift. Gorgonians and Scleractinia showed high occurrence frequencies and large total wet weight (Fig 2). The occurrence frequencies of Stylasterina, Alcyonacea, Antipatharia and Porifera were low, but the total wet weight of Porifera was relatively high because of the large water content. Many association rules that include Gorgonians or Scleractinia in the consequent part were extracted and showed high confidence and lift values. Only a small number of association rules were extracted for Porifera, and no rules were extracted for Alcyonacea, Antipatharia and Stylasterina. These results demonstrate that Gorgonians and Scleractinia frequently co-exist with other benthic animals and suggest their potential as VME indicator taxa in the Emperor Seamounts area. This study presents a new method to assess characteristics of benthic communities and to screen for potential indicator taxa based on the analysis of co-occurrence tendencies among benthic taxa.

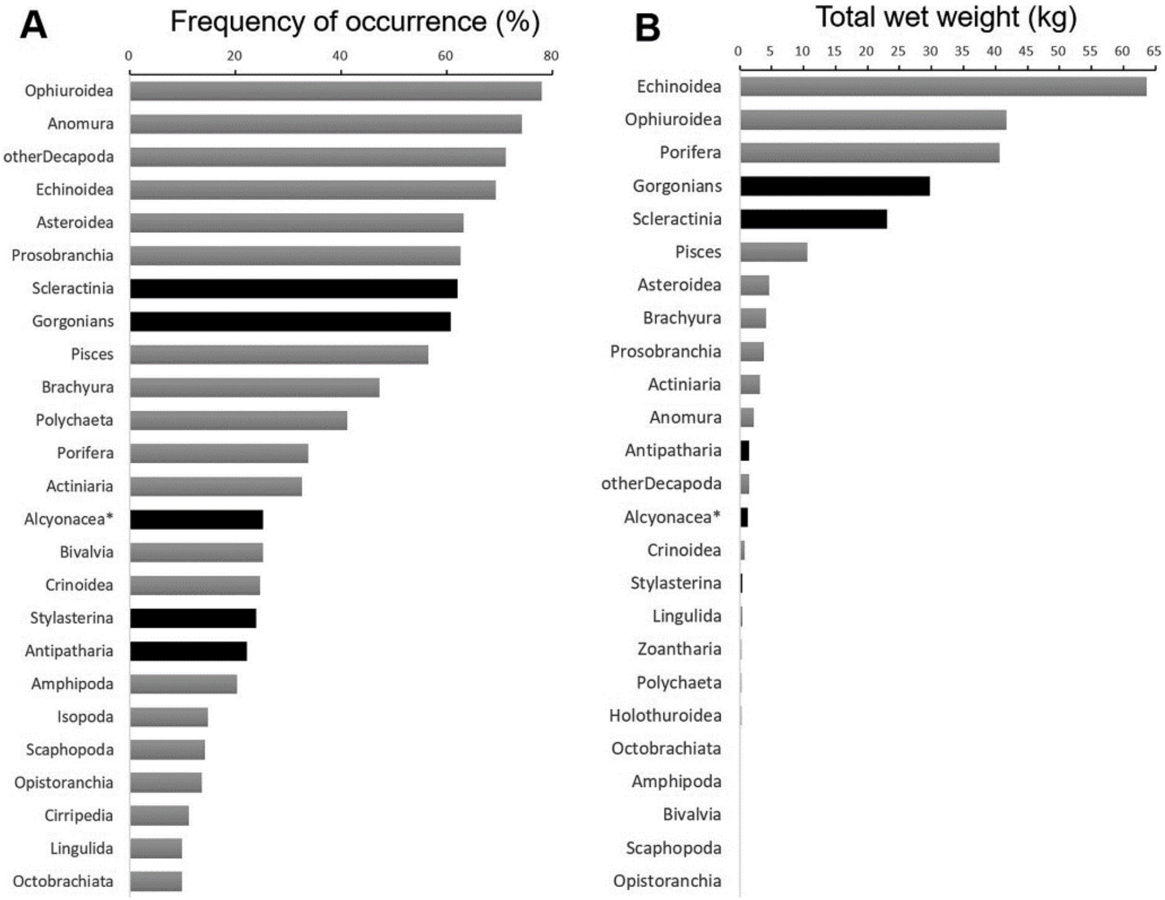
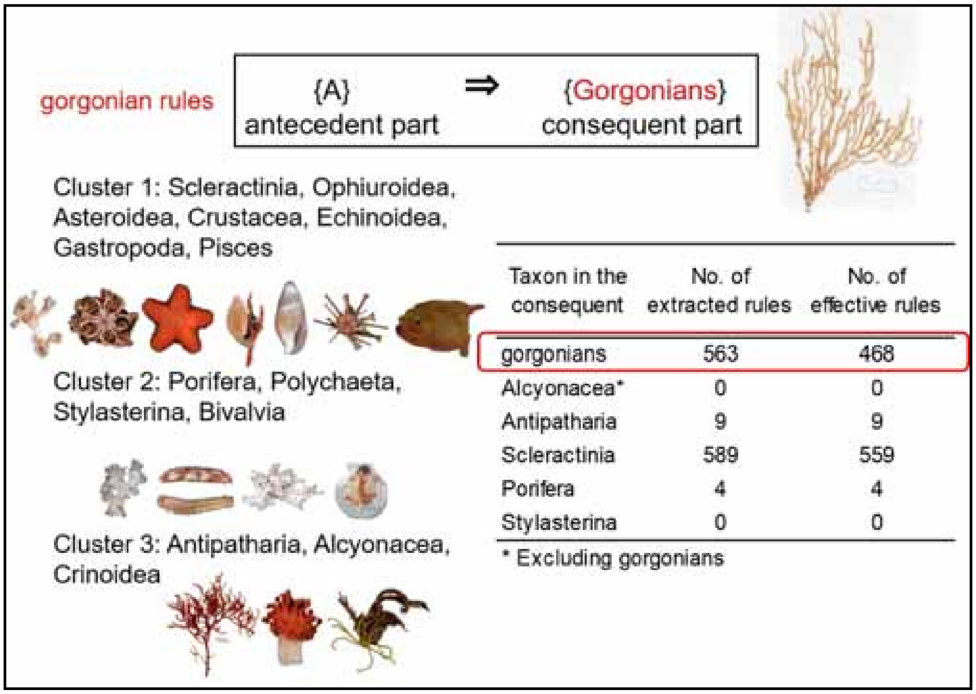


Figure 1. Frequency of occurrence (A) and total wet-weights (B) of benthic megafauna collected by scientific surveys in the southern Emperor Seamounts area. Only top 25 groups are shown in this figure. Alcyonacea* excluding Gorgonians (Scleraxonia, Holaxonia and Calcaxonia).



Taxon in the consequent	No. of extracted rules	No. of effective rules
gorgonians	563	468
Alcyonacea*	0	0
Antipatharia	9	9
Scleractinia	589	559
Porifera	4	4
Stylasterina	0	0

* Excluding gorgonians

Figure 2. Example of results of Gorgonians association analysis. Most effective association rules had Gorgonians in the consequent part and other taxa in the antecedent parts.

References

Blattberg, R.C., Kim, B-Do., and Neslin, S.A. (2008) Market basket analysis. In: Database Marketing. Springer, New York, USA, pp. 339–351.

Hahsler, M., Bettina, G., Hornik, K. (2005) A Computational Environment for Mining Association Rules and Frequent Item Sets. WU Vienna University Research Report Series. No. 15.

Miyamoto, M., and Kiyota, M. (2017). Application of association analysis for identifing indicator taxa of vulnerable marine ecosystems in the Emperor Seamounts area, North Pacific Ocean. Ecological Indicators, 78, 301-310.

Silverstein, G., Brin, S., Motwani, R. (1997) Beyond market baskets: generalizing association rules to dependence rules. Data Min. Knowl. Discov. 2, 39–68.

Application of association analysis for identifying VME indicator taxa on the basis of sea-floor visual images

Mai Miyamoto[1] and Masashi Kiyota[2]

[1]Environmental Consulting Department
Japan NUS Co. Ltd.
Nishi-Shinjuku Prime Square 5F,
7-5-25 Nishi-Shinjuku, Shinjuku-Ku, Tokyo 160-0023, Japan

[2]Graduate School of Fisheries and Environmental Sciences Nagasaki University
1-14 Bunkyo-machi, Nagasaki 852-8521, Japan

Research Summary

Members of PICES Working Group 32 and collaborators further explored another usage of the association analysis for identifying VME indicator taxa on the basis of visual seafloor surveys. The visual images of the sea-floor were collected through the observation surveys using the drop camera system operated by R/V Kaiyo-maru in the Emperor Seamounts area. The benthic animals taken in each sea-floor footage were identified to family or order levels (Porifera, Stylasterina, Pennatulacea, Gorgonians, Alcyonacea (excluding Gorgonians), Antipatharia, Scleractinia, Actiniaria, Corallimorpharia, Zoantharia, jellyfish, Gastropoda, Cephalopoda, Polychaeta, Crustacea, Crinoidea, Asteroidea, Ophiuroidea, Echinoidea and Pisces), and their occurrence frequencies were calculated. Using the density of benthic taxa in sampling locations as multivariate distance data, the sea-floor photographing sites or benthic taxa were classified into clusters according to Ward’s method. The Canonical Correspondence Analysis (CCA) was applied to characterize the benthic community by environmental parameters (depth, flatness, softness, and roughness). Cluster analysis indicated that the survey sites were classified into two large clusters such as hard bottom or soft bottom, and further

classified into 6 clusters relevant to the variations of depth, flatness, softness and roughness. CCA plot demonstrated the influence of sea-floor features on benthos occurrence (Fig 1).

Then association analysis, which is commonly used for discovering hidden relationships among purchased items in market transaction data (Blattberg et al., 2008) and also applied to assess the validity of benthic indicator taxa (Miyamoto and Kiyota, 2015), was applied to the occurrence data per sea-floor photographing site to explore the co-occurring relationships of benthic animals. Association rules that represent strong relationship like A (condition part) →B (conclusion part) were extracted. The Gorgonians showed 27 effective association rules that included 10 taxa as antecedent (Table 1). Only a small number of effective rules were extracted when Scleractinia or Porifera were placed in the consequent. The extracted rules with Scleractinia or Porifera as the consequent included few taxa as antecedents. These results demonstrate that Scleractinia and Porifera are less effective as indicators of the benthic communities in the Emperor Seamounts area than Gorgonians. In summary, the composition of the benthic community varied greatly depending on the bottom substratum, and Gorgonians were dominant benthos on hard bottom that occupied many stations. Association analysis demonstrated co-occurrence of Gorgonians with many other benthic animals.

References

Blattberg, R.C., Kim, B-Do., and Neslin, S.A. (2008) Market basket analysis. In: Database Marketing. Springer, New York, USA, pp. 339–351.

Miyamoto, M., and Kiyota, M. (2017). Application of association analysis for identifing indicator taxa of vulnerable marine ecosystems in the Emperor Seamounts area, North Pacific Ocean. Ecological Indicators, 78, 301-310.

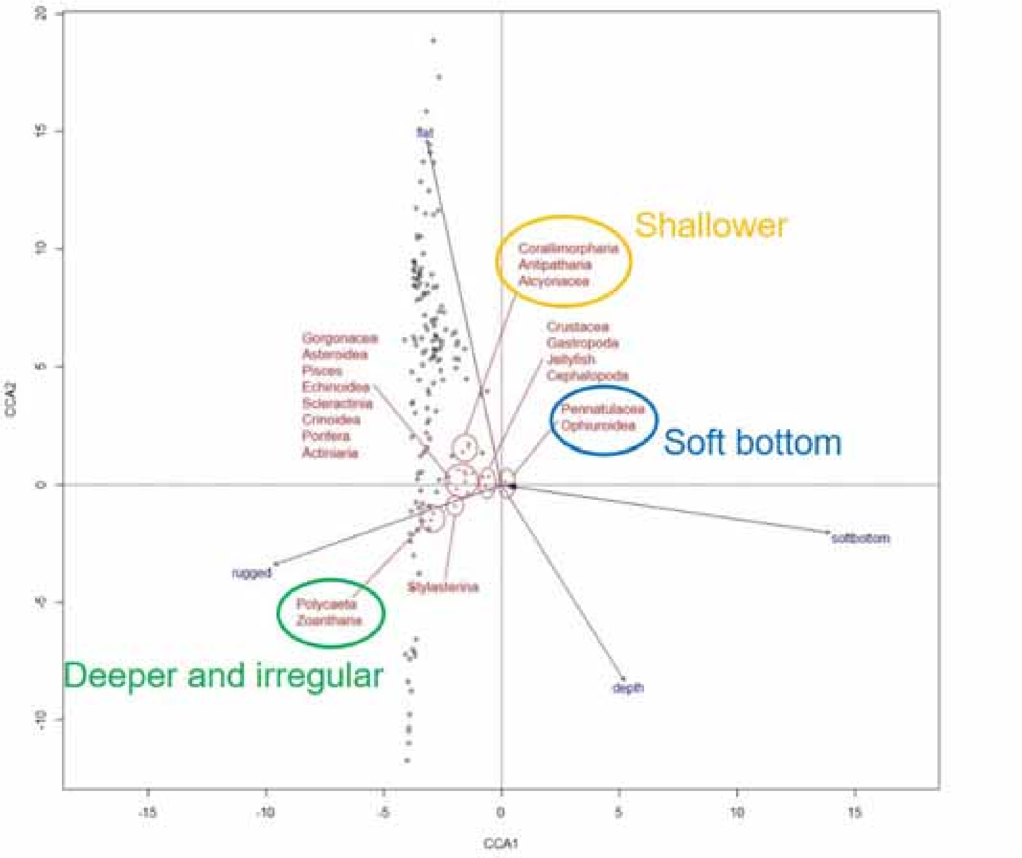


Figure 1. CCA plot characterizes the benthic community by four environmental parameters (i.e. depth, flatness, softness and roughness).

Table 1. Number of effective association rules that include candidate vulnerable marine ecosystem indicator taxa as the consequent part.

Taxon in the consequent	No. of rules	Taxa in the antecedents
Gorgonacea	27	Alcyonacea, Antipatharia, Scleractinia, Porifera, Stylasterina, Crustacea, Crinoidea, Asteroidea, Echinoidea, Pisces
Alcyonacea	0	-
Antipatharia	0	-
Scleractinia	2	Antipatharia, Crustacea, Echinoidea, Pisces
Porifera	1	Gorgonacea, Echinoidea
Stylasterina	0	-

Assessment of species diversity and dominance of shallow water corals using environmental DNA

Go Suzuki[1] and Chuya Shinzato[2]

[1]Yaeyama Laboratory, Seikai National Fisheries Research Institute

[2]Atmosphere and Ocean Research Institute, The University of Tokyo, Japan

Background

DNA that originated from various sources, such as mucus, metabolic waste, and damaged tissues, from multicellular organisms exists in seawater and is called environmental DNA (eDNA). Recently eDNA has begun to be employed in aquatic environmental research, and it is also being used to monitor marine biodiversity. However, we are not aware of any reports that have attempted to use eDNA from seawater to monitor coral reefs. In this study, we performed tank experiments with running seawater as an initial proof of the concept.

Corals release massive amounts of soluble mucus, which transfers large amounts of energy and nutrients to the reef substrate. Due to the symbiont's natural rate of increase, corals steadily release *Symbiodinium* cells into the surrounding environment, suggesting that close to reefs, seawater should contain detectable quantities of DNA from both corals and *Symbiodinium*. Recently whole genome sequences of an *Acropora* coral and *Symbiodinium* have been published, and next-generation sequencing (NGS) technologies have been used to investigate coral reef biodiversity. For easier assessment of the species diversity of reef building corals, we developed a novel method for monitoring of *Acropora* corals from environmental DNA (eDNA) in seawater using NGS.

Research Summary

We performed a tank experiment with running seawater using 19 *Acropora* species; *A. acuminata*, *A. austera*, *A. awi*, *A. cytherea*, *A. carduus*, *A. digitifera*, *A. echinata*, *A. florida, A. grandis*, *A. hyacinthus*, *A. intermedia*, *A.* sp1 aff. *echinata*, *A. microphthalma*, *A. muricata*, *A. nasuta*, *A. selago*, *A. tenuis* (Shinzato et al., 2018). Complete mitochondrial genomes of all the *Acropora* species were assembled to create a database and major types of their *Symbiodinium* symbionts were identified. Then eDNA was isolated by filtering inlet and outlet seawater from the tanks. We detected all of the tested *Acropora* types from eDNA samples. Proportions and numbers of DNA sequences were both positively correlated with masses of corals in the tanks. In this trial, we detected DNA sequences from as little as 0.04 kg of *Acropora* colony, suggesting that existence of at least one adult *Acropora* colony (∼30 cm diameter = 1 kg) per m^2 at depths <10m could be detected using eDNA in the field. Although this is the initial attempt to detect coral and *Symbiodinium* simultaneously from eDNA in seawater, this method may allow us to perform high-frequency, high-density coral reef monitoring of coral species composition and their health conditions without specialized skills to identify coral species using morphological traits.

References

Shinzato, C., Zayasu, Y., Kanda, M., Kawamitsu, M., Satoh, N., Yamashita, H., and Suzuki, G. (2018) Using seawater to document coral-zoothanthella diversity: A new approach to coral reef monitoring using environmental DNA. Front. Mar. Sci., 5, 28.

Studies of the distribution and diversity of biogenic habitat forming taxa in the USA

Rooper Christopher N.[1,2], Georgian Samuel E.[3], Guinotte John M.[4], and Les Watling[5]

[1]Alaska Fisheries Science Center, National Marine Fisheries Service, Seattle WA, USA
[2]Present address: Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC, Canada
[3]Marine Conservation Biology Institute, Seattle WA, USA
[4]US Fish and Wildlife Service, Lakewood CO, USA
[5]Department of Biology, University of Hawaii at Manoa, Honolulu, HI, USA

Background and Overview

This chapter summarizes the recent contributions by the USA and its WG32 members (Sam Georgian, John Guinotte, Chris Rooper and Les Watling) to assess the distribution, abundance and species associations of deep-sea corals and sponges within the EEZ. It should be noted that the USA conducts extensive research on deep-sea coral and sponge within its EEZ and much of that research is summarized in reports by NOAA's Deep Sea Coral Research and Technology Program (https://www.fisheries.noaa.gov/national/habitat-conservation/deep-sea-coral-habitat). The most recent report on the state of deep-sea coral and sponge ecosystems was published in 2017 (Hourigan et al. 2017) and contains an expanded summary of the work on deep-sea corals and sponges in the USA. Here we focus on the research activities that were linked conceptually, temporally or directly to PICES WG32 activities and terms of reference.

Coral and sponge ecosystem data in the USA

Records of deep-sea coral and sponge presence and abundance in the North Pacific Ocean have been historically collected through a number of scientific activities by research organizations in the USA. These include data from the extensive fisheries-independent surveys of ecosystems on the US West Coast and Alaska (both longline and bottom trawl) where the relative abundance and species of corals and sponges are recorded. It also includes a number of other studies using visual survey methods, such as submersibles, remote operated vehicles, autonomous underwater vehicles and towed camera systems. Often these data have records of associated fish and invertebrate assemblages. In North Pacific Ocean waters a number of studies have documented associations of Rockfishes (Sebastes sp.) and other demersal fish species with deep-sea corals and sponges, as well as identifying some key relationships between fish reproduction and these ecosystems.

Beginning in about 2015 these historical data have been housed in a publicly accessible data portal (https://deepseacoraldata.noaa.gov/). From 2014-2019, members of WG32 compiled data from studies in Alaska and the US West Coast and contributed these data (n > 10,000 observations) to the publicly accessible database making it available to the wider PICES community. The database follows the standards of OBIS-USA and is an extension of the international Darwin Core Standard. Key points for these aggregated data are:

- Most studies that collect data on coral and sponge distribution in the USA have records included in a publicly available database that is updated quarterly
- Care in the use of these data is needed, as the quality can vary. A source listed for each data record can be used to check the validity and suitability for a given analysis

Modelling approaches

One of the key activities of WG32 was to review modeling approaches to predict the potential distributions of species and habitat suitability for corals and sponges (e.g., MaxEnt, Boosted Regression Trees, or high resolution bathymetry-based models) within National EEZs. During the 2016 PICES annual meeting a 2-day workshop was held on modeling approaches for deep-sea corals and sponges. It was co-convened by Drs. Kwang-Sik Choi (Korea), Janelle Curtis (Canada), Masashi Kiyota (Japan) and Chris Rooper (USA). The outcomes of the workshop were recommendations for data and modeling approaches that should be considered for deep-sea coral and sponge ecosystems. Technical aspects of the species distribution modeling, including the best-practices for generating input data, creating models and evaluating the results, data driven approach to define bioregions, a multi-scale assessment of species distribution models, and an assessment of the model transferability were also examined during the workshop. The workshop also included a “hands-on” exercise of building some preliminary models of corals and sponges for data limited taxa in the North Pacific Ocean. The goals of the modeling session were to 1) evaluate existing environmental variables/mechanisms effecting basin-wide distribution of coral and sponge, 2) construct preliminary basin-wide habitat models for taxa including glass sponges and corals in the North Pacific Ocean and, 3) provide model-based information for predicting potential changes in distributions of coral and sponge with climate change. An example of the model predictions for Antipatharia in the North Pacific Basin developed during the workshop are shown in Figure 1, with the important variables predicting the distribution shown in Table 1.

In addition to the PICES workshop in 2016, a number of concurrent modeling efforts for Alaska and US West Coast were conducted by WG32 members. A maximum entropy model was used to predict habitat suitability for deep-sea corals on the US West Coast (Guinotte and Davies 2014), a generalized additive modeling approach was used to predict the distribution, abundance and diversity of corals and sponges in the Aleutian Islands and eastern Bering Sea (Rooper et al. 2014, Rooper et al. 2016). In the Gulf of Alaska, a suite of modeling methods (maximum entropy, general linear models, generalized additive models, boosted regression tree and random forest) were evaluated and used in an ensemble to predict the distribution of corals and sponges. In all, these efforts resulted in predictions of the distribution of coral taxa for the entire USA EEZ in the North Pacific. These models continue to be developed and improved (Guinotte et al. 2017) as new data and techniques become available. Key findings from this work are:

- Model validation using independently collected data is important to consider
- Ensemble models can perform better across a range of species than individual models
- The modeling method has less effect on the result than the quality of the underlying predictor and distribution data

Environmental and ecological predictors

A key outcome in modeling the distribution of deep-sea coral and sponge ecosystems in the North Pacific Ocean that was advanced within WG32 was the identification of large scale environmental and ecological predictors for the distribution and biodiversity of coral, sponge and associated taxa. Dr. Samuel Georgian (USA) and colleagues put together an exhaustive group of measured and derived predictor variables for the North Pacific Ocean on a 1 km^2 grid (Table 2; Figure 3). These included bathymetric and terrain variables as well as environmental variables thought to influence the distribution of deep-sea corals and sponges. A number of topographical variables were created, as complex seafloor features generally elevate local currents, which increases larval dispersal, food supply, sediment and waste removal, and dissolved oxygen flux (e.g., Dorschel et al., 2007). Since cold-water corals are filter feeders and therefore reliant on the vertical transfer of surface productivity (Duineveld et al., 2007), the particulate organic carbon (POC) flux (mg C m^{-2} d^{-1}) to the seafloor was also included as a proxy for food availability. The saturation state of the calcite polymorph of calcium carbonate, temperature (°C), salinity (psu), and dissolved oxygen (ml L^{-1}) were included due to their known biological relevance to cold-water octocorals (Mortensen & Buhl-Mortensen, 2004) and importance in previous habitat suitability models (Quattrini et al., 2013; Georgian et al., 2014; Etnoyer et al., 2017). Finally, the distribution of hard bottom substrates was included as the availability of hard substrata is essential for coral recruitment in a number of cold-water corals (e.g., Georgian et al., 2014) including Paramuricea (Mortensen et al., 2007).

Bathymetric data for the North Pacific were obtained from the SRTM30+ layer (Becker et al. 2009; Sandwell et al. 2014) downloaded from: topex.ucsd.edu/WWW_html/srtm30_plus.html. Ocean data in the SRTM30+ layer are derived from Sandwell et al. (2014), the LDEO Ridge Multibeam Synthesis Project, the JAMSTEC Data Site for Research Cruises, the NGDC Coastal

Relief Model, and the International Bathymetric Chart of the Oceans. The native resolution of the SRTM30+ layer is 0.0083° or approximately one kilometer.

A suite of terrain variables was constructed using the SRTM30+ bathymetry layer. Slope, aspect, roughness, and curvature were calculated using the ArcGIS (v.10.4, ESRI) toolkit ‘DEM Surface Tools’ (v.2, Jenness 2004; Jenness 2013a), and the topographic position index were calculated using the toolkit ‘Land Face Corridor Designer (v1.2, Jenness et al. 2013b). The slope of each grid was measured in degrees and calculated using the 4-cell method which has been shown to outperform other methods (Horn 1981, Jones 1998). Aspect represents the direction of the maximum slope and was converted to an index of ‘eastness’ using a sine transformation and an index of ‘northness’ using a cosine transformation. Curvature describes the shape of the seafloor to quantify how water should interact with the terrain. We calculated three types of curvature: general curvature, cross-sectional curvature, and longitudinal curvature. General curvature assigns more positive values to more convex features, and more negative values to more concave features. Cross-sectional curvature assigns positive values to features where water is expected to locally diverge, and negative values to features where water is expected to locally converge. Longitudinal curvature assigns positive values to features where water is expected to decelerate, and negative values to features where water is expected to accelerate. Roughness is a measure of topographical complexity and was calculated as the ratio of surface area to planimetric area. Topographic Position Index (TPI) quantifies the elevation of a feature relative to the surrounding seafloor, with positive values indicating features that are elevated and negative values indicating features that are depressed. Values close to zero may indicate either flat surfaces or areas with constant slopes. As TPI is heavily dependent on the analysis scale, and because benthic organisms may be simultaneously affected by both fine- and broad-scale features, we calculated TPI at multiple scales: 1,000 m (the finest resolution allowed by the bathymetry), 5,000 m, 10,000 m, and 20,000 m. Seamount locations were obtained from Yesson et al. (2011), and include all seafloor features greater than 1000 m in height with a conical shape.

Environmental variables believed a priori to influence the distribution of benthic marine organisms were obtained from a variety of sources (Table 1). Dissolved oxygen, salinity, temperature, and nutrient data were obtained from the World Ocean Atlas (v.2 2013). Carbonate data (Ω_{ARAG}, Ω_{CAL}, dissolved inorganic carbon, and total alkalinity) were obtained from Steinacher et al. (2009). Chlorophyll a, sea surface temperature (SST), and photosynthetically available radiation (PAR) data were obtained as mission composites (average of 2002-2016 data) from NOAA's Aqua MODIS program at a resolution of 4 km, and were resampled to match the extent and resolution of the bathymetry data with no interpolation. Particulate organic carbon (POC) flux at the seafloor was obtained from Lutz et al. (2007). Bottom current velocity data were obtained from the Simple Ocean Data Assimilation (SODA v.3.4.1) model (Carton et al. 2005), with data averaged as the composite of the years 1990-2007. Current velocities were calculated as both horizontal and vertical velocities (m s-1). Current direction for each grid cell was calculated from zonal (U) and meridional (V) velocities according to the formula:

$$\text{Direction}=180/\pi \times \text{atan2}\ 〖([u],[v])〗$$

with values of +180° and -180° indicating that the current flows to the south, +90° to the east, -90° to the west, and 0° to north. A current layer that quantifies the direction of current flow relative to the aspect of the seafloor was also calculated, with values of 0° indicating that the current flows the same direction that the steepest slope is facing, and values of 180° indicating that the current flows opposite to the direction of the steepest slope (sensu Rooper et al. 2014).

Benthic variables (WOA data, carbonate data, and current data) were transformed to match the extent and resolution of the bathymetry layer using a variable up-scaling approach that approximates conditions at the seafloor (Davies and Guinotte 2011). Briefly, each gridded layer was first interpolated to a slightly higher resolution (0.5°) than its native resolution using inverse distance weighting, resampled to match the extent and resolution of the bathymetry data, and draped over the bathymetry data within its depth range. This technique has been demonstrated to work effectively for many global and regional scale variables (Davies and Guinotte 2011; Yesson et al. 2012). WOA data were available as 102 depth-binned layers from depths of 0-5500 m, with a vertical resolution of 5 m (from 0-100 m), 25 m (100-500 m), 50 m (500-2000 m), and 100 m (2000-5500 m). Carbonate data (Steinacher et al. 2009) were available in 33 depth-binned layers: 6, 19, 38, 62, 93, 133, 183, 245, 322, 415, 527, 661, 818, 1001, 1211, 1449, 1717, 2014, 2340, 2693, 3072, 3473, 3894, 4329, 4775 m. SODA current data were available in 50 depth bins: 5.03355, 15.10065, 25.21935, 35.35845, 45.57635, 55.86325, 66.26175, 76.80285, 87.57695, 98.62325, 110.0962, 122.1067, 134.9086, 148.7466, 164.0538, 181.3125, 201.2630, 224.7773, 253.0681, 287.5508, 330.0078, 382.3651, 446.7263, 524.9824, 618.7031, 728.6921, 854.9935, 996.7153, 1152.376, 1319.997, 1497.562, 1683.057, 1874.788, 2071.252, 2271.323, 2474.043, 2678.757, 2884.898, 3092.117, 3300.086, 3508.633, 3717.567, 3926.813, 4136.251, 4345.864, 4555.566, 4765.369, 4975.209, 5185.111, and 5395.023 m.

These predictor variables were made available to all WG32 members through a shared drive and allowed individuals and groups from the PICES community to utilize the layers in their own modeling efforts within their own EEZ's. Key aspects of this predictor data set are:

- A wide variety of variables have been compiled into raster layers for use in modeling the distribution of deep-sea corals and sponges.
- The data captures long-term and large-scale patterns in variables for the North Pacific Ocean

Preliminary North Pacific-Wide Deep-Sea Coral Models

This suite of environmental data was used to construct preliminary habitat suitability models for a number of deep-sea coral taxa across the North Pacific Ocean. Models were constructed using a presence-only Maxent approach. The inclusion of correlated environmental variables may inhibit model performance and interpretation (e.g., Huang et al. 2011). Therefore, highly correlated variables (Figure 2) were removed based on their relationship with other variables and performance in preliminary Maxent models. The final variable set included the saturation state of calcium carbonate (either as calcite or aragonite depending on the biology of each taxon), seafloor roughness, temperature, silicate, topographic position index (TPI; 20,000 m scale), dissolved oxygen, dissolved inorganic carbon, total alkalinity, regional current flow, and vertical current flow.

Georeferenced occurrence data were obtained for each taxon from the NOAA Deep Sea Coral and Sponge Portal. While a common source of error in species distribution models, spatial bias in the sampling of occurrence data considerably weakens the performance and interpretability of models (Phillips et al. 2009), and is often found in presence-only deep-sea species datasets due to the difficulties associated with sampling design in the deep-sea. However, it is possible to reduce the effects of sampling bias by selecting targeted background data that reflect the same bias as the occurrence data (Phillips et al. 2009). Given the relatively high sampling bias observable in our field surveys, we generated a targeted set of background points in addition to a random set of points (see Figure 1). To generate background points preferentially in areas that have been more extensively surveyed, we first created a two-dimensional kernel density estimate of sampling effort based on the presence locations for each taxon. This created a probability grid from which 10,000 background points were sampled according to the probability grid weights. Previous studies using a similar approach to reduce the influence of sampling bias found that model performance was significantly improved (Fitzpatrick et al.

2013). Habitat suitability models were produced using this targeted-background approach in a Maxent environment run with default model parameters for the following taxa: Scleractinia (stony corals), Antipatharia (black corals), Pennatulacea (sea pens), and *Primnoa* (preliminary results in Figure 4).

Proposed Biogeography of the Upper Bathyal Benthos in the Pacific Ocean Based on Octocoral Distributions

One of the topics for discussion during WG32 meetings was the global biogeography of benthic invertebrates and how invertebrates are taxonomically organized in the PICES area. Biogeographical classification schemes such as Briggs' biogeographic provinces (Briggs 1974) and the Marine Ecoregions of the World (MEOW) (Spaulding et al. 2007) have been developed for continental shelf depths, however, the lack of faunal data in the deep sea has led to the development of biogeographical units based on oceanographic characteristics (Watling et al. 2013). During the lifetime of WG32, Dr. Natalie Summers and Dr. Les Watling developed a biogeographical scheme for the Upper Bathyal (200-1000m) in the Pacific Ocean using octocoral distributions.

They retrieved over 200 000 octocoral data records from the Deep Sea Coral Data Portal (DSCDP), Ocean Biogeographic Information System (OBIS), Tropical Deep-Sea Benthos program (French National Museum of Natural History), Queensland Museum from the CIDARIS expeditions, and records retrieved from the Siboga expedition reports. They used cluster analysis to test octocoral distributions against four different biogeographical classification schemes used in the shallower and deeper zones of the ocean as well as one used for subdividing the mesopelagic.

The four classifications used were based on: 1) The MEOW classification in the Pacific (Spalding et al. 2007) from the coast to 200m depth (immediately above the Upper Bathyal), which consisted of 72 ecoregions containing at least one octocoral genus; 2) Watling et al. (2013) Lower Bathyal Provinces from 800 to 3500m in the Pacific; 3) Sutton et al. (2017) 14 Mesopelagic Provinces in the Pacific based on daytime mesopelagic faunal communities as well as environmental proxies derived from the MEOW ecoregions; 4) The three-dimensional Ecological Marine Unit (EMU) classification based on temperature, salinity, dissolved oxygen, nitrate, phosphate, and silicate data obtained from the 2013 World Ocean Atlas, which using cluster analysis resulted in 9 EMUs in the Pacific Ocean. The latter is the only scheme that divides the Upper Bathyal into several depth zones (Sayre et al., 2017).

All classification schemes produced mostly concordant patterns with three major faunal distribution barriers: the North Pacific Current isolates the subarctic units by creating a steep temperature gradient; the Subantarctic Front separates the Subantarctic from the rest of the Pacific Ocean; and the East Pacific Barrier separates the East Pacific Ocean from the Central and West Pacific Ocean. Two other smaller but distinct provinces are the Indo-Pacific where Lower Bathyal genera are found in the Upper Bathyal, and Torres Strait/Coral Sea characterised by mesophotic genera. Dr. Summers and Dr. Watling proposed 12 Upper Bathyal provinces for the

Pacific Ocean based on octocoral distributions (Figure 4). The main driver for these units seems to be temperature, a defining feature of water masses. These units could potentially be subdivided into smaller regions based on habitat. Additionally, the clustering of EMUs provides evidence that the Upper Bathyal should in certain regions be divided vertically into two depth zones based on water masses. Key findings recommended to PICES WG32 members from this work include:

- There was consistency in results across multiple biogeographic classification schemes indicating strong trends in the data on octocorals,
- The biogeography of octocorals in the North Pacific are largely consistent with the oceanographic barriers created by temperature and
- There are four separate biogeographic provinces for octocoral (Figure 5) found in the PICES region.

Indicators of diversity of biogenic habitats

Deep-sea coral and sponge ecosystems in the North Pacific Ocean are influenced by multiple climatological threats, such as rising sea temperature, harmful algal blooms, marine invasive species, hypoxia, and eutrophication. They are also affected by direct anthropogenic activities such as sea bed mining and bottom fishing. These multiple threats can act synergistically, but perhaps differently, from region to region to change ecosystem structure, function and dynamics. A goal of PICES WG32 was to advance the monitoring of deep-sea coral and sponge ecosystems.

In the USA, the major threats to deep-sea corals and sponges are fishing and climate change. Monitoring the status of deep-sea corals and sponges is difficult and costly, so effective indicators of biogenic habitat health are needed. One group of indicators that has been developed in the USA are trends in bycatch (Figure 6) in commercial fisheries. Another indicator is the abundance of deep-sea corals and sponges estimated using multispecies geo-spatial modeling techniques (Thorson et al. 2015) applied to fishery independent trawl survey data. Finally, the spatial extent and trend over time in bottom contacting fishing effort can be used as an indicator of potential fishing impacts on deep-sea corals and sponges. These data and indicators are currently updated and reported annually in the Ecosystem Status Reports of Stock Assessment and Fishery Evaluation documents for Alaska (Zador et al. 2019). The data and trends over time are also available to be downloaded (https://access.afsc.noaa.gov/reem/ecoweb/index.php). These indicators do not necessarily capture the diversity of deep-sea corals and sponges, but they indicate instead what the population status might be and document the potential threats due to fishing activity. Further work is needed on this topic, especially work to integrate the spatially explicit fishing effort with the distribution models for deep-sea corals and sponges. Key recommendations to PICES WG32 members for this work include:

- There is a need to share relevant fisheries and environmental data that addresses trends and threats to biodiversity
- There is a need for the develop marine spatial planning tools and tools to measure marine protected area performance

Associations between commercial species and biogenic habitats

In the US EEZ of the North Pacific Ocean there have been a number of recent studies that have examined the association of commercially important fish and invertebrate species and deep-sea coral and sponge ecosystems. On a larger scale (ecosystem) the studies have demonstrated a strong correlation between rockfishes (Sebastes sp.) and corals and sponges in bottom trawl survey catches (Sigler et al. 2015, Laman et al. 2016, Thorson and Barnett 2017, Laman et al. 2019). In these studies, higher catches of rockfishes (and some other commercially important species) is higher where corals or sponges also occur in the catch. This indicates spatial correlation in density across large areas. The strength of these correlations is variable however, with stronger associations in Alaska than on the US West Coast. However, functional relationships are difficult to resolve using bottom trawl survey data.

Two new studies conducted by members of WG32 have examined the seasonal use of deep-sea coral and sponge habitat and cross-ecosystem associations at differing scales in Alaska. These studies found that the associations documented for rockfishes were consistent across seasons; indicating that rockfishes exhibit the same habitat associations during all times of the year (Conrath et al. 2019). Since most previous work in Alaska has focused on only summertime distributions, this is an important finding.

In a second study, comparisons of habitat associations for rockfishes were compared for the same species group across two different ecosystems (the Aleutian Islands and eastern Bering

Sea) at scales ranging from 1 m to 1000's of km. The study found that rockfishes, in particular, utilized habitat in the same ways in both ecosystems (Rooper et al. 2019). This was in spite of the large differences in quantity and quality of habitat between the two ecosystems. In all, this study found that having more structured habitats (deep-sea coral and sponge) led to a greater abundance of rockfishes, regardless of the larger ecosystem characteristics. The key findings of these studies and others assessed by WG32 members were:

- Deep-sea corals and sponges serve as important habitats for commercially important rockfishes in the Northeastern Pacific Ocean
- The associations between rockfishes and habitat are constant over seasonal and annual time periods and over all spatial scales examined
- A higher abundance of structured habitats, such as deep-sea corals and sponges leads to an increased abundance of rockfishes

Conclusions

Deep-sea coral and sponge ecosystems in the USA EEZ of the North Pacific Ocean are widely distributed. Data on deep-sea corals and sponges have been collated into a global database that is publicly available. The work of WG32 has advanced our knowledge of deep-sea coral and sponge ecosystem distributions by providing a motivation and tools for modeling presence, absence and abundance of deep-sea corals and sponges. Variables developed in workshops sponsored by WG32 have also been crucial to moving the distribution modeling forward on a relevant scale (1 km) for management. Since the working group was formed, published models for all regions of the USA EEZ have been developed and used in management. In addition, some key work in the USA has documented the importance of deep-sea coral and sponge ecosystems to supporting productivity of commercially important fish stocks, particularly rockfishes. Current monitoring efforts relative to deep-sea corals and sponges in the USA EEZ focus primarily on time series of abundance, bycatch and fishing effort. The monitoring could be improved with more focused studies that address the biodiversity of biogenic habitats.

References

Aqua MODIS. 2014. NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group (2014): MODIS-Aqua Ocean Color Data. http://dx/doi.org/10.5067/AQUA/MODIS_OC.2014.0

Becker JJ, Sandwell DT, Smith WHF, Braud J, Binder B, Depner J, Fabre D, Factor J, Ingalls S, Kim SH, Ladner R, Marks K, Nelson S, Pharaoh A, Trimmer R, Von Rosenberg J, Wallace G, Weatherall P. 2009. Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution: SRTM30_PLUS. Marine Geodesy 32(4):355-371.

Briggs, J.C., 1974. Marine Zoogeography. McGraw-Hill, New York.

Carton A, Giese BA. 2008. A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). Monthly Weather Review 136(8):2999-3017.

Chu JWF, Nephin J, Georgian S, Knudby A, Rooper CN, Knudby A, Gale KSP. In press. Modeling the environmental niche space and distributions of cold-water corals and sponges in the Canadian northeast Pacific Ocean. Deep Sea Res I

Conrath CL, Rooper CN, Wilborn RE, Knoth BA, Jones DT. 2019. Seasonal habitat use and community structure of rockfish in the Gulf of Alaska. Fish Res

Dorschel, B., Hebbeln, D., Fouber,t A., White, M. & Wheeler, A.J. (2007). Hydrodynamics and cold-water coral facies distribution related to recent sedimentary processes at Galway Mound west of Ireland. Mar. Geol., 244, 184−195.

Duineveld, G.C., Lavaleye, M.S., Bergman, M.J., De Stigter, H. & Mienis, F. Trophic structure of a cold-water coral mound community (Rockall Bank, NE Atlantic) in relation to the near-bottom particle supply and current regime. Bulletin of Marine Science, 81(3), 449–467.

Etnoyer, P.J., Wagner, D., Fowle, H.A., Poti, M., Kinlan, B., Georgian, S.E. & Cordes, E.E. (2017). Models of habitat suitability, size, and age-class structure for the deep-sea black coral Leiopathes glaberrima in the Gulf of Mexico. Deep Sea Research Part II: Topical Studies in Oceanography doi: 10.1016/j.dsr2.2017.10.008

Fitzpatrick MC, Gotelli NJ, Ellison AM. 2013. MaxEnt versus MaxLike: empirical comparisons with ant species distributions. Ecosphere 4(5):1-15.

Garcia HE, Locarnini RA, Boyer TP, Antonov JI, Baranova OK, Zweng MM, Reagan JR, Johnson DR. 2014a. World Ocean Atlas 2013, Volume 3: Dissolved Oxygen, Apparent Oxygen

Utilization, and Oxygen Saturation. Eds: Levitus S, Mishonov A. NOAA Atlas NESDIS 75. 27 pp.

Garcia HE, Locarnini RA, Boyer TP, Antonov JI, Baranova OK, Zweng MM, Reagan JR, Johnson DR. 2014b. World Ocean Atlas 2013, Volume 4: Dissolved Inorganic Nutrients (phosphate, nitrate, silicate). Eds: Levitus S, Mishonov A. NOAA Atlas NESDIS 76. 25 pp.

Georgian, S.E., Shedd, W. & Cordes, E.E. (2014). High-resolution ecological niche modelling of the cold-water coral Lophelia pertusa in the Gulf of Mexico. Marine Ecology Progress Series, 506, 145–161.

Horn BK. 1981. Hill shading and the reflectance map. Proceedings of the IEEE. 69(1):14-47.

Hourigan TF, Etnoyer PJ, Cairns SD (2017) The State of Deep-Sea Coral and Sponge Ecosystems of the United States. NOAA Technical Memorandum NMFS-OHC-4. Silver Spring, MD. 467 p.

Huang Z, Brooke B, Li J. 2011. Performance of predictive models in marine benthic environments based on predictions of sponge distribution on the Australian continental shelf. Ecological Informatics 6(3-4):205-16.

Jenness JS. 2004. Calculating landscape surface area from digital elevation models. Wild Soc Bull 32:829-839.

Jenness JS. 2013b. Land Facet Corridor Designer. Jenness Enterprises. Available at: www.jennessent.com/arcgis/landfacets.htm

Jenness, JS. 2013a. DEM Surface Tools. Jenness Enterprises. Available at: www.jennessent.com/arcgis/surface_area.htm

Jones KH. 1998. A comparison of algorithms used to compute hill slope as a property of the DEM. Computers & Geosciences. 24(4):315-23.

Laman, N., S. Kotwicki & C.N. Rooper. 2015. Sponge and coral morphology influences the distribution of Pacific ocean perch life stages. Fish Bull. 113:270-289.

Locarnini RA Mishonov AV, Antonov JI, Boyer TP, Garcia HE, Baranova OK, Zweng MM, Paver CR, Reagan JR, Johnson DR, Hamilton M, Seidov D. 2013. World Ocean Atlas 2013, Volume 1: Temperature. Eds: Levitus S, Mishonov A. NOAA Atlas NESDIS 73. 40 pp.

Lutz MJ, Caldeira K, Dunbar RB, Behrenfeld MJ. 2007. Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. Journal of Geophysical Research: Oceans, 112(C10).

MacLean SA, Rooper CN, Sigler MS. 2017. Corals, Canyons, and Conservation: Science based fisheries management decisions in the eastern Bering Sea. Frontiers Marine Science 4: 1-5.

Mortensen, P.B. & Buhl-Mortensen, L. (2004). Distribution of deep-water gorgonian corals in relation to benthic habitat features in the Northeast Channel (Atlantic Canada). Marine Biology, 144(6), 1223–1238.

Mortensen, P.B., Buhl-Mortensen, L., Gebruk, A.V. & Krylova, E.M. (2007). Occurrence of deep-water corals on the Mid-Atlantic Ridge based on MAR-ECO data. Deep Sea Research II, 55, 142–152.

Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological applications 19(1):181-197.

Quattrini, A.M., Georgian, S.E., Byrnes, L., Falco, R., Stevens, A. & Cordes, E.E. (2013). Niche divergence by deep-sea octocorals in the genus Callogorgia across the upper continental slope of the Gulf of Mexico. Mol. Ecol., 22, 4123−4140.

Rooper CN, Goddard P, Wilborn R. 2019. Are fish associations with corals and sponges more than an affinity to structure: evidence across two widely divergent ecosystems. Can J Fish Aquat Sci

Rooper CN, Sigler M, Goddard P, Malecha PW, Towler R, Williams K, Wilborn R, Zimmermann M (2016) Validation and improvement of species distribution models for structure forming invertebrates in the eastern Bering Sea with an independent survey. Mar. Ecol. Prog. Ser. 551:117-130

Rooper CN, Wilborn RE, Goddard P, Williams K, Towler R, Hoff GR. 2018. Validation of deep-sea coral and sponge distribution models in the Aleutian Islands, Alaska. ICES J. Mar. Sci. 75:199-209.

Rooper, C.N., M. Zimmermann, M. Prescott, A. Hermann. 2014. Predictive models of coral and sponge distribution, abundance and diversity in bottom trawl surveys of the Aleutian Islands, Alaska. Mar. Ecol. Prog. Ser. 503:157-176.Sandwell DT, Müller RD, Smith HF, Garcia E, Francis R. 2004. New global marine gravity model from CryoSat-2 and Jason-1 reveals buried tectonic structure. Science 346(6205):65-67.

Rooper, C.N., M. Zimmermann, M. Prescott. 2017. Comparisons of methods for modeling coral and sponge distribution in the Gulf of Alaska. Deep Sea Research II 126:148-161.

Sayre, R., Wright, D., Breyer, S., Butler, K., Van Graafeiland, K., Costello, M., Harris, P., Goodin, K., Guinotte, J., Basher, Z., Kavanaugh, M., Halpin, P., Monaco, M., Cressie, N., Aniello, P., Frye, C., Stephens, D., 2017. A three-dimensional mapping of the ocean based on environmental data. Oceanography 30, 90–103.

Sigler, M.S., Rooper, C.N., Hoff, G.R, Stone, R.P., McConnaughey, R.A. and T. K. Wilderbeur. 2015. Faunal features of submarine canyons on the eastern Bering Sea slope. Mar. Ecol. Prog. Ser. 526:21-40.

Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdaña, Z.A., Finlayson, M.A.X., Halpern, B.S., Jorge, M.A., Lombana, A.L., Lourie, S.A., Martin, K.D., Manus, M.C., Molnar, J., Recchia, C.A., Robertson, J., 2007. Marine ecoregions of the world: A bioregionalization of coastal and shelf areas 57, 573–583.

Steinacher M, Joos F, Frolicher TL, Plattner CK, Doney SC. 2009. Imminent ocean acidification in the Arctic projected with the NCAR global coupled carbon cycle-climate model. Biogeosciences 6:515-533.

Sutton, T.T., Clark, M.R., Dunn, D.C., Halpin, P.N., Rogers, A.D., Guinotte, J., Bograd, S.J., Angel, M. V., Perez, J.A.A., Wishner, K., Haedrich, R.L., Lindsay, D.J., Drazen, J.C., Vereshchaka, A., Piatkowski, U., Morato, T., Błachowiak-Samołyk, K., Robison, B.H., Gjerde, K.M., Pierrot-Bults, A., Bernal, P., Reygondeau, G., Heino, M., 2017. A global biogeographic classification of the mesopelagic zone. Deep. Res. Part I Oceanogr. Res. Pap. 126, 85–102.

Thorson, J.T., Shelton, A.O., Ward, E.J., Skaug, H.J., 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES J. Mar. Sci. J. Cons. 72(5), 1297–1310. doi:10.1093/icesjms/fsu243. URL: http://icesjms.oxfordjournals.org/content/72/5/1297

Watling, L., Guinotte, J., Clark, M.R., Smith, C.R., 2013. A proposed biogeography of the deep ocean floor. Prog. Oceanogr. 111, 91–112.

Wilborn, R.E., Rooper, C.N., Goddard, P., Li, L., Williams, K., Towler, R. 2018. Distribution, abundance, and size of cold-water corals and sponges in the Aleutian Islands, Alaska, using camera surveys. Hydrobiologia 811:251-268.

Yesson C, Clark MR, Taylor ML, Rogers AD. 2011. The global distribution of seamounts based on 30 arc seconds bathymetry data. Deep Sea Research Part I: Oceanographic Research Papers. 58(4):442-53.

Yesson C, Taylor ML, Tittensor DP, Davies AJ, Guinotte J, Baco A, Black J, Hall-Spencer JM, Rogers AD. 2012. Global habitat suitability of cold-water octocorals. Journal of Biogeography 39(7):1278-92.

Zweng MM, Reagan JR, Antonov JI, Locarnini RA, Mishonov AV, Boyer TP, Garcia HE, Baranova OK, Johnson DR, Seidov D, Biddle MM. 2013. World Ocean Atlas 2013, Volume 2: Salinity. Eds: Levitus S, Mishonov A. NOAA Atlas NESDIS 74. 39 pp.

Tables

Table 1. Variables important in modeling the distribution of Antipatharia in the North Pacific Ocean. Two alternative models are shown, one without a bias grid correction for sampling distribution and one corrected for the sampling bias. AUC is the area under the receiver-operator curve (an indication of overall model fit).

Variable	% Contribution without bias grid	% Contribution with bias grid
Calcite	58.6	65.5
Roughness	14.3	3.6
Temperature	7.8	11.2
Silicate	4.2	
TPI 20,000 m	3.6	3.8
Dissolved Oxygen	3.3	--
DIC	3.2	--
Alkalinity	--	2.1
Regional currents	--	2.1
Vertical currents	--	2.0
AUC	0.945 (0.004)	0.925 (0.003)

Table 2. Geophysical and environmental variables available for modeling the distribution of deep-sea corals and sponges in the North Pacific Ocean.

Variable name	Filename	Units	Native Resolution	Reference
Bathymetry	srtm30	meters	0.0083°	Becker et al. 2009 Sandwell et al. 2014
Terrain variables				
Aspect	aspect	degrees	0.0083°	Jenness 2013a
Aspect – Eastness	eastness		0.0083°	Jenness 2013a
Aspect – Northness	northness		0.0083°	Jenness 2013a
Curvature – General	gencurve		0.0083°	Jenness 2013a
Curvature – Cross-Sectional	crosscurve		0.0083°	Jenness 2013a
Curvature – Longitudinal	longcurve		0.0083°	Jenness 2013a
Roughness	roughness		0.0083°	Jenness 2013a
Slope	slope	degrees	0.0083°	Jenness 2013a
Topographic Position Index	tpi		0.0083°	Jenness 2013b
Seamounts	seamounts			Yesson et al. 2011
Environmental variables				
Alkalinity	alk_stein	µmol l^{-1}	3.6x0.8-1.8°	Steinacher et al. (2009)
Dissolved inorganic carbon	dic_stein	µmol l^{-1}	3.6x0.8-1.8°	Steinacher et al. (2009)
Omega aragonite (Ω_{ARAG})	arag_stein		3.6x0.8-1.8°	Steinacher et al. (2009)
Omega calcite (Ω_{CALC})	calc_stein		3.6x0.8-1.8°	Steinacher et al. (2009)
Dissolved oxygen	dissox	ml l^{-1}	1°	Garcia et al. 2014a
Salinity	salinity	pss	0.25°	Zweng et al. 2013
Temperature	temp	°C	0.25°	Locarnini et al. 2013
Phosphate	phosphate	µmol l^{-1}	1°	Garcia et al. 2014b
Silicate	silicate	µmol l^{-1}	1°	Garcia et al. 2014b
Nitrate	nitrate	µmol l^{-1}	1°	Garcia et al. 2014b
Particulate organic carbon	POC	g C m^{-2} yr^{-1}	0.05°	Lutz et al. (2007)
Regional current velocity	regfl	m s^{-1}	0.5°	Carton et al. (2005)
Vertical current velocity	vertfl	m s^{-1}	0.5°	Carton et al. (2005)
Current direction	curdir	degrees	0.5°	Carton et al. (2005)
Current relative to aspect	curaspect	degrees	0.5°	Rooper et al. (2014)
Chlorophyll *a*	chla	mg m^{-3}	4 km	Aqua Modis (NOAA)
hotosynthetically Available Radiation	PAR	W m^{-2}	4 km	Aqua Modis (NOAA)
Sea Surface Temperature	SST	°C	4 km	Aqua Modis (NOAA)

Figures

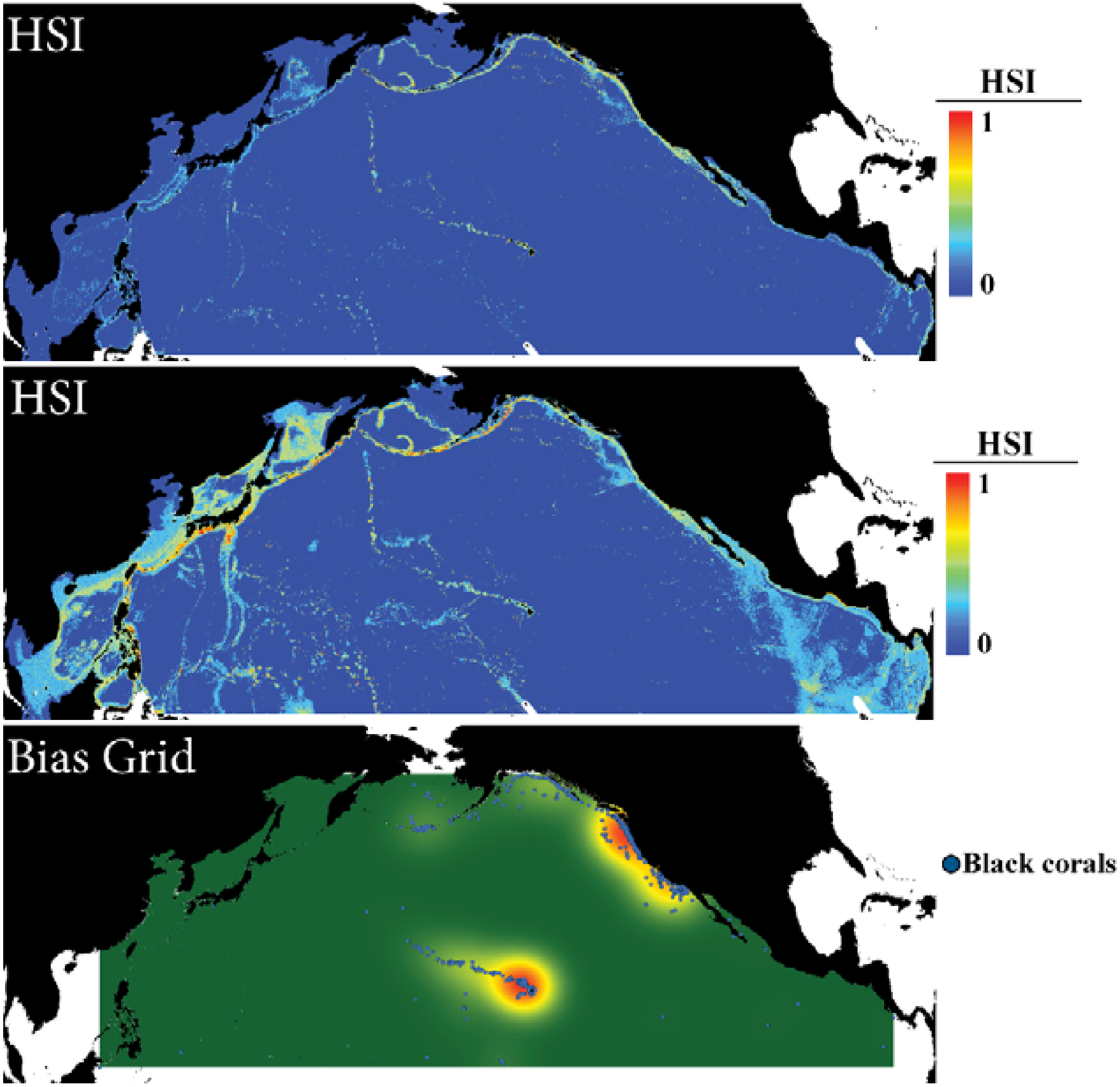


Figure 1. Example models of the distribution for Antipatharia developed by WG32 during the 2016 modeling workshop. The predictions were developed from maximum entropy models without (Panel A) and with (Panel B) a correction for sampling density (Panel C).

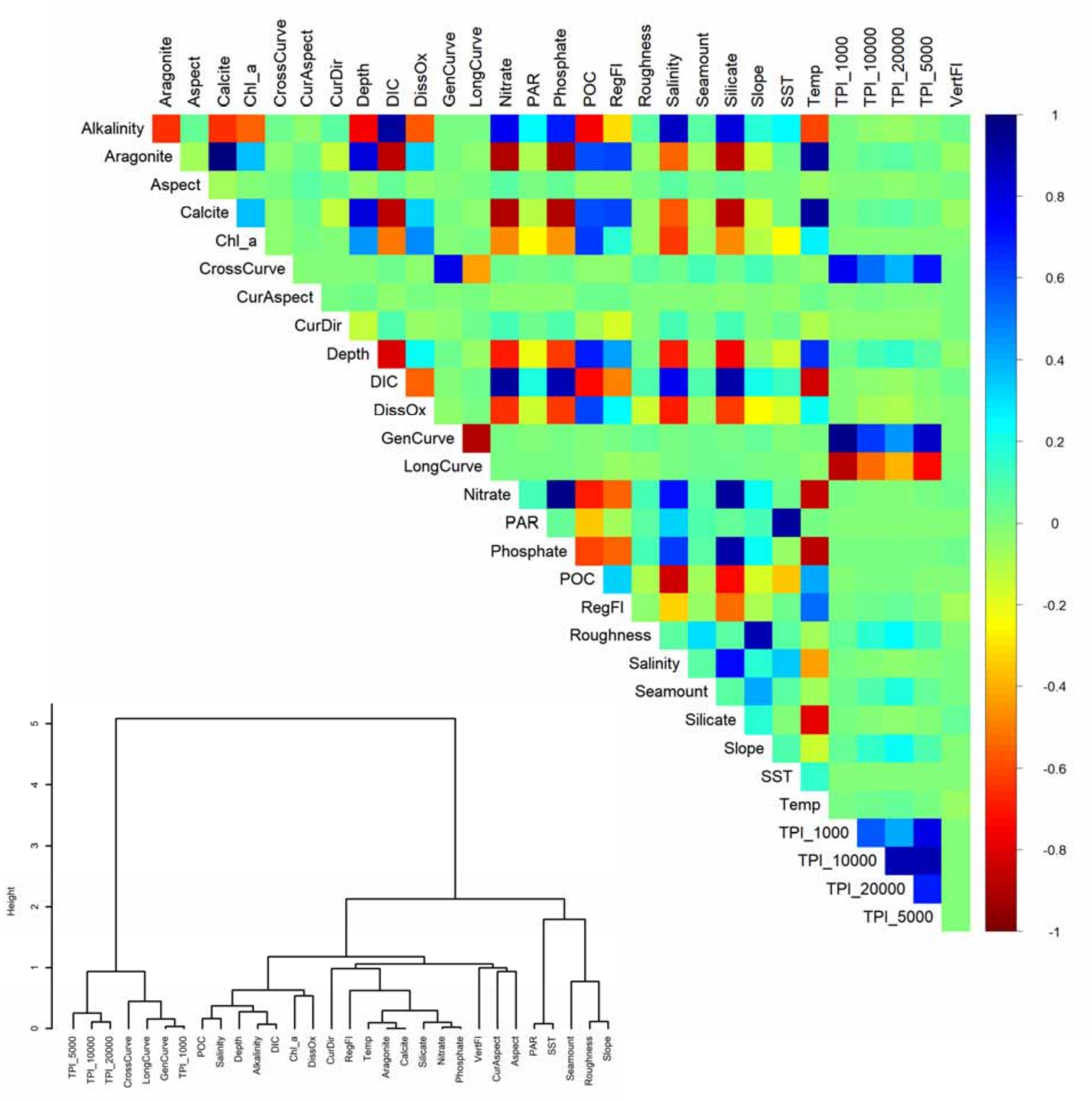
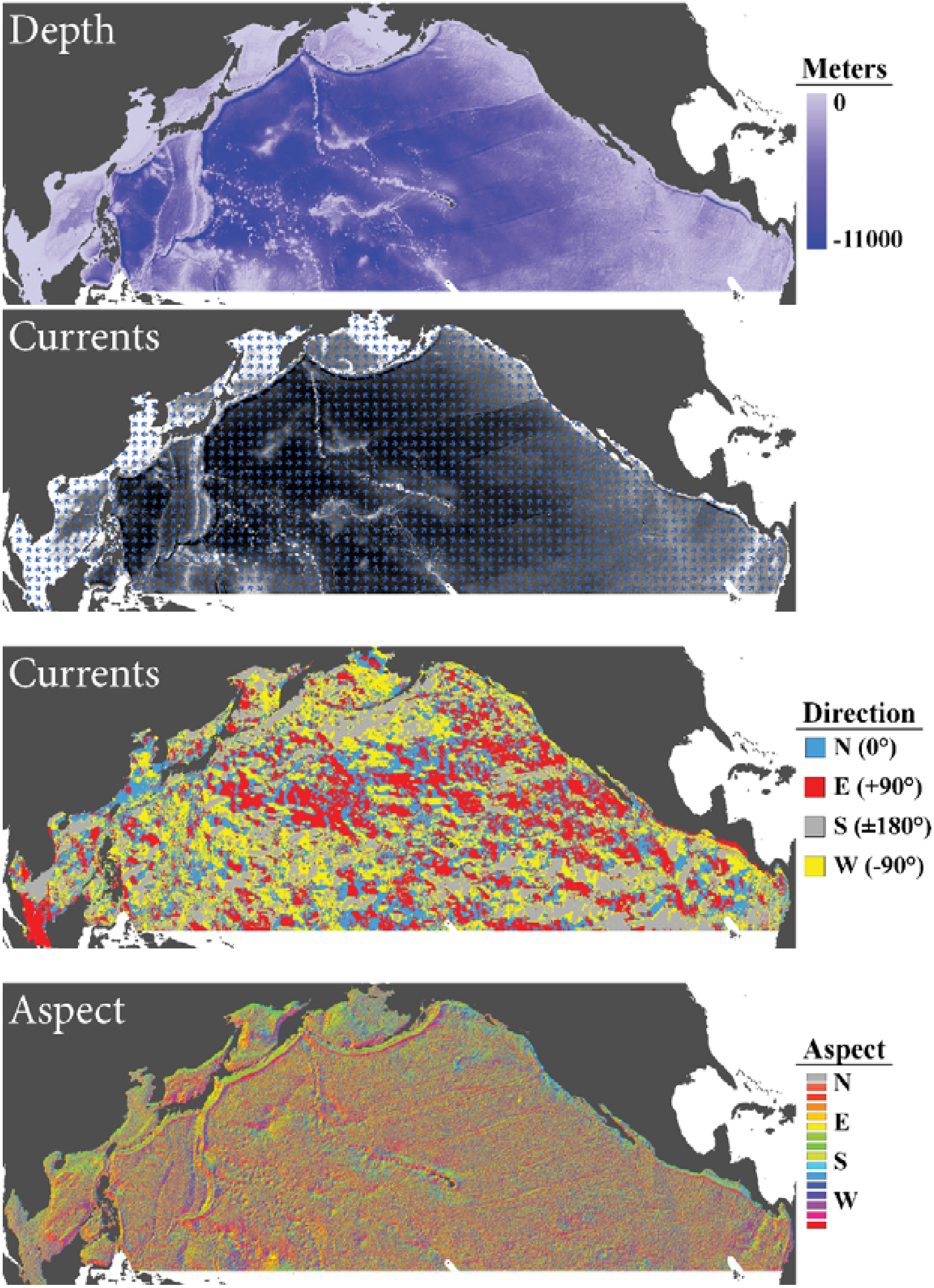
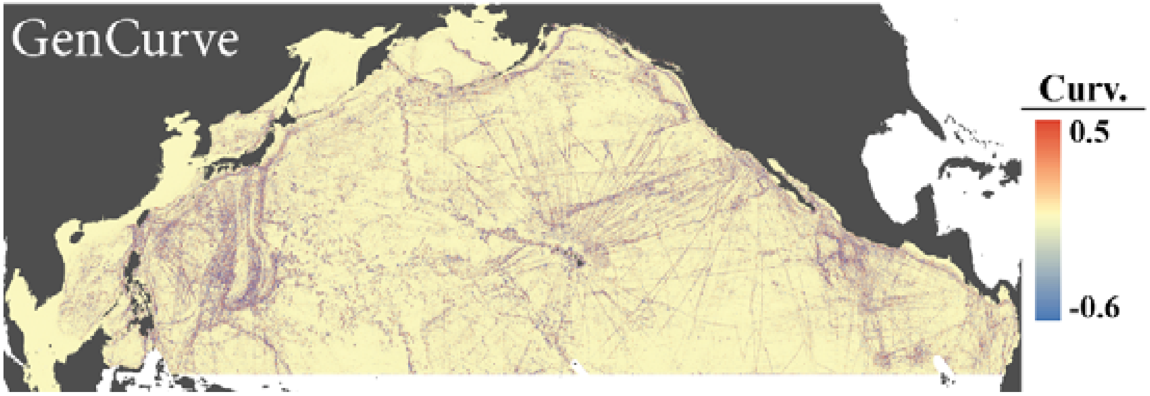
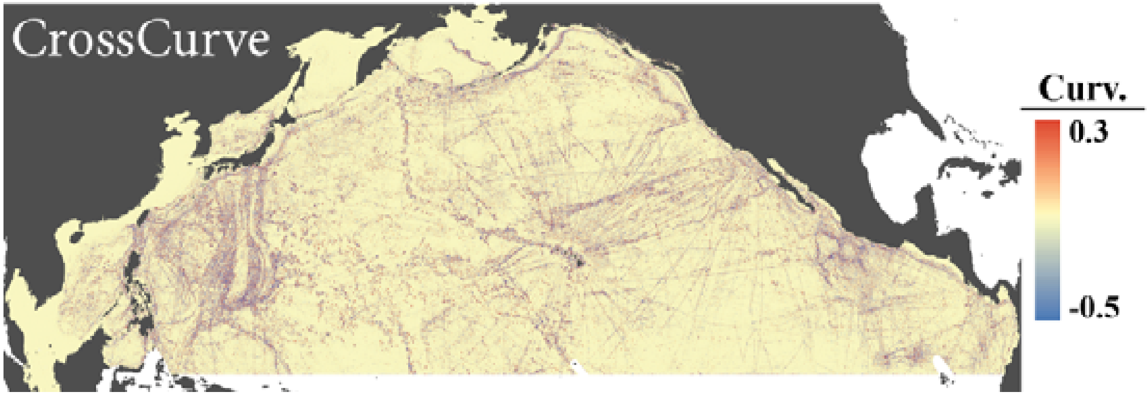


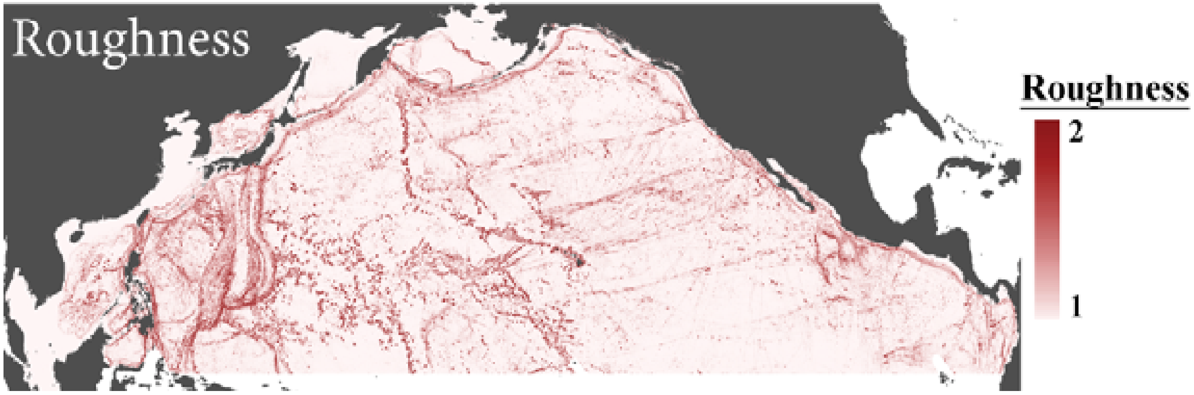
Figure 2. Pearson correlations and cluster analysis of variables available for modeling the distribution of corals and sponges in the Pacific Ocean.

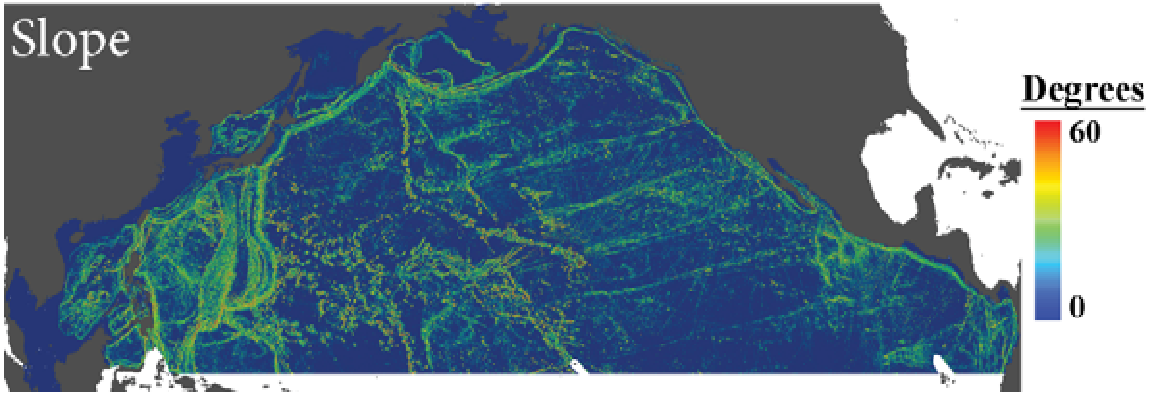

Depth
Meters
0
-11000
Currents
Currents
Direction
N (0°)
E (+90°)
S (±180°)
W (-90°)
Aspect
Aspect
N
E
S
W

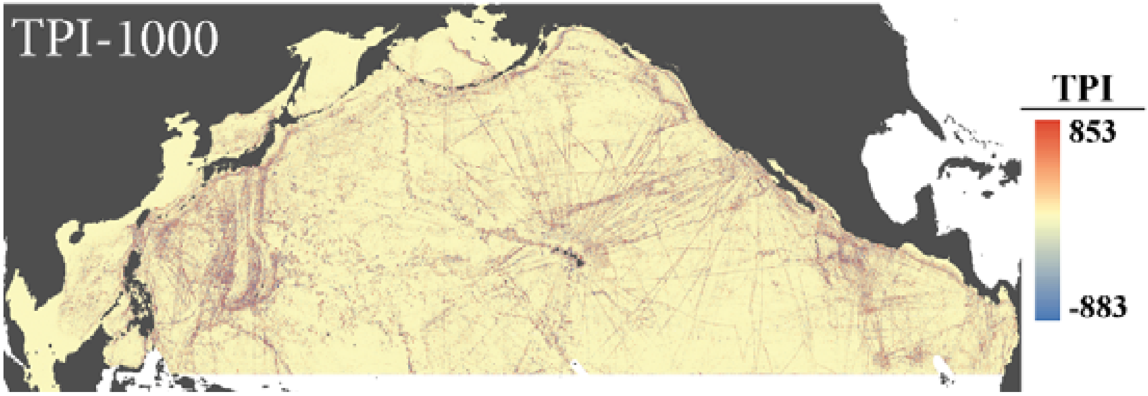

GenCurve
Curv.
0.5
-0.6

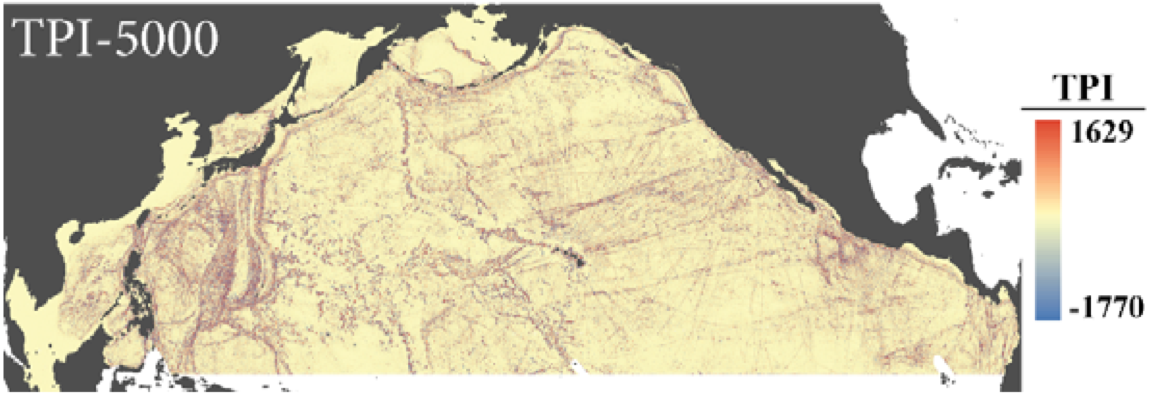

CrossCurve
Curv.
0.3
-0.5

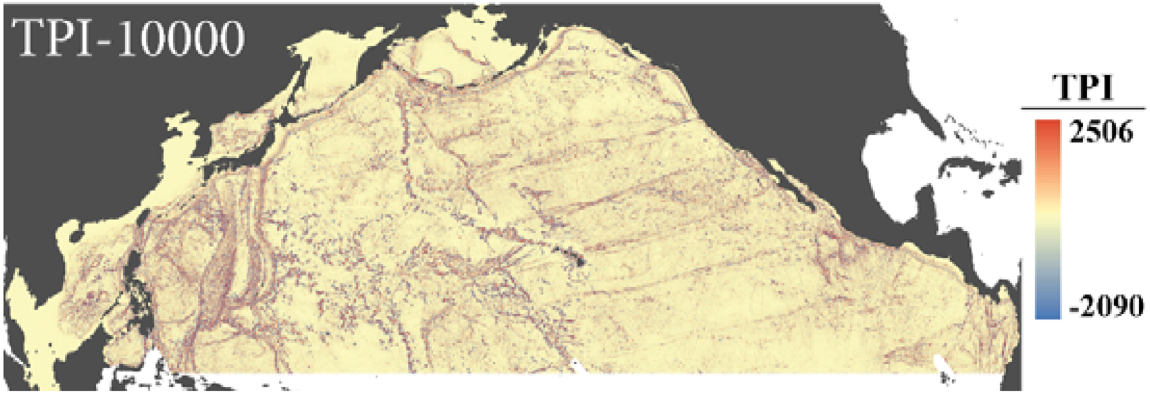

LongCurve
Curv.
0.4
-0.4

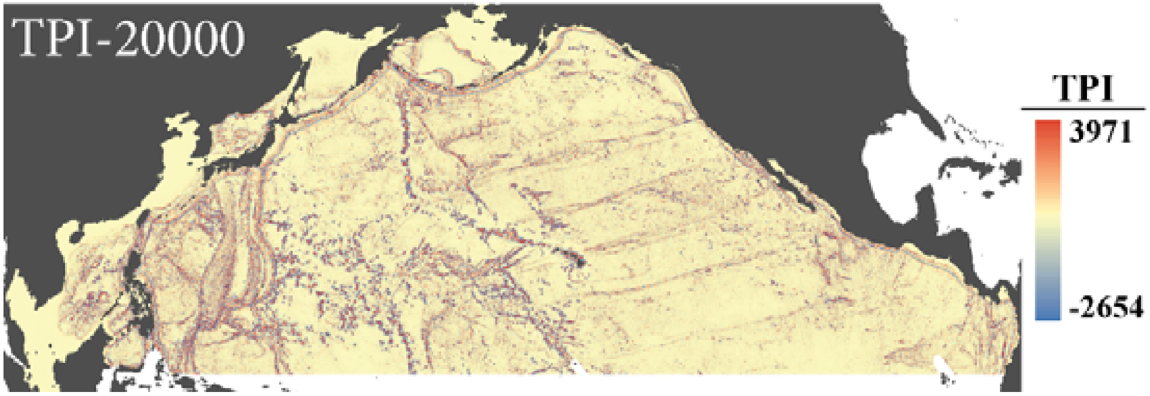

Roughness
Roughness
2
1

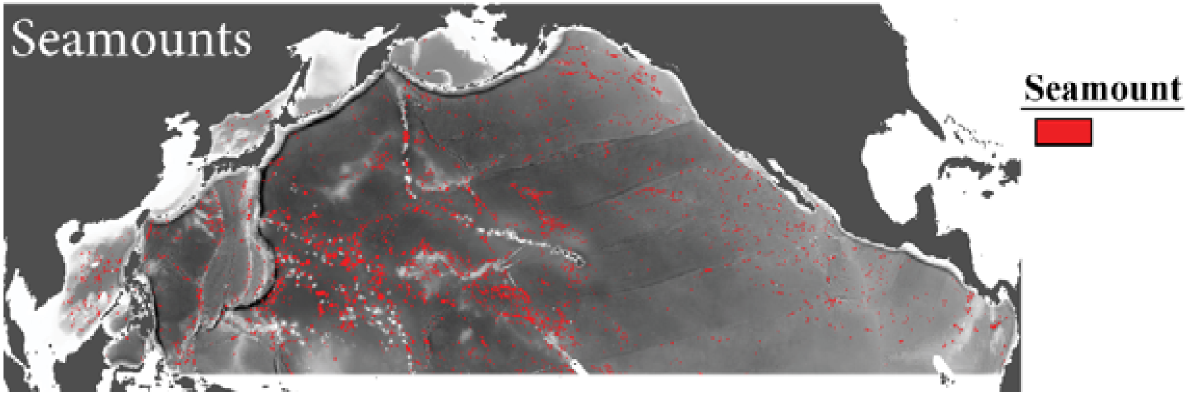

Slope
Degrees
60
0

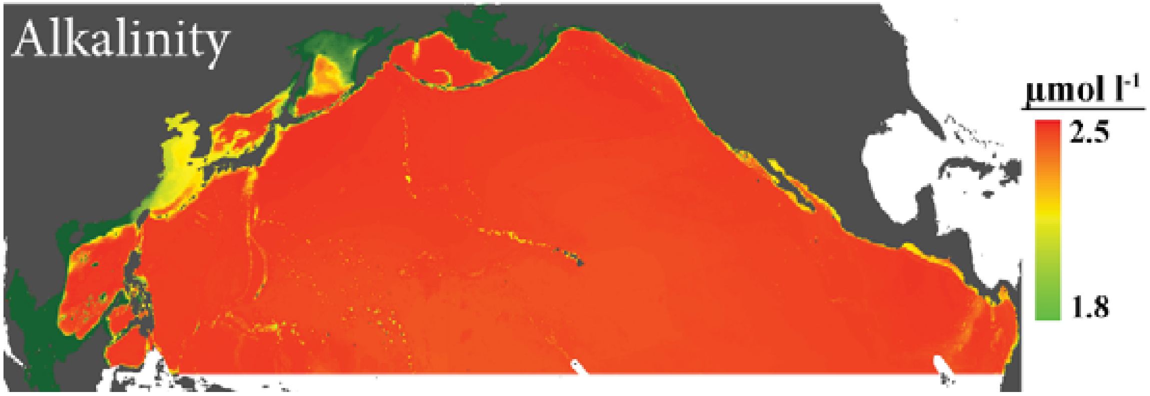

TPI-1000
TPI
853
-883

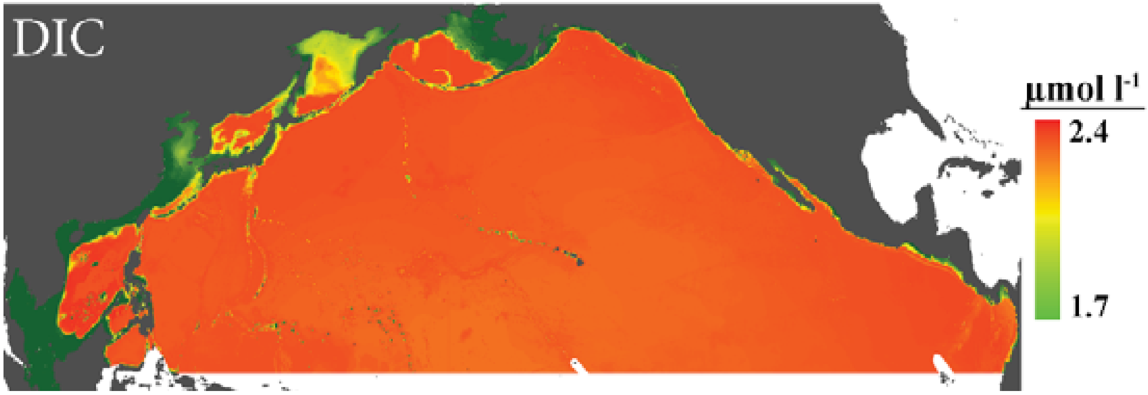

TPI-5000
TPI
1629
-1770

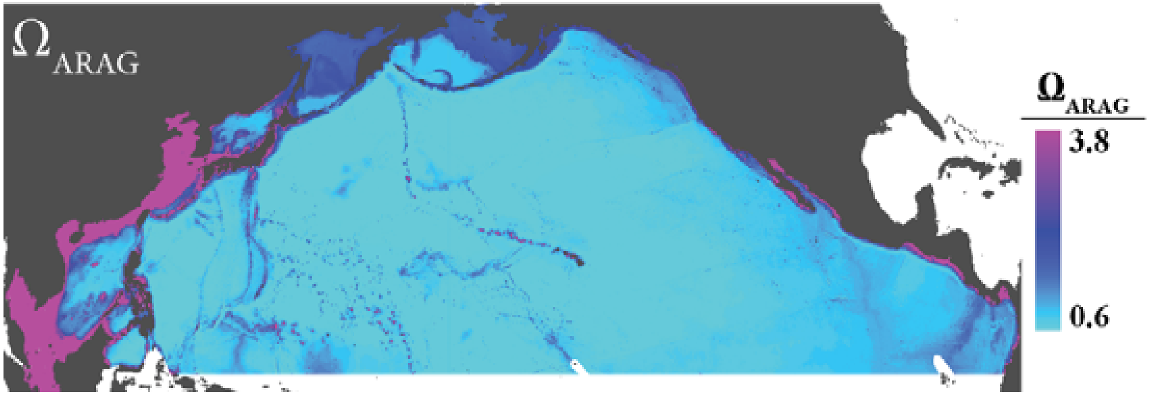

TPI-10000
TPI
2506
-2090

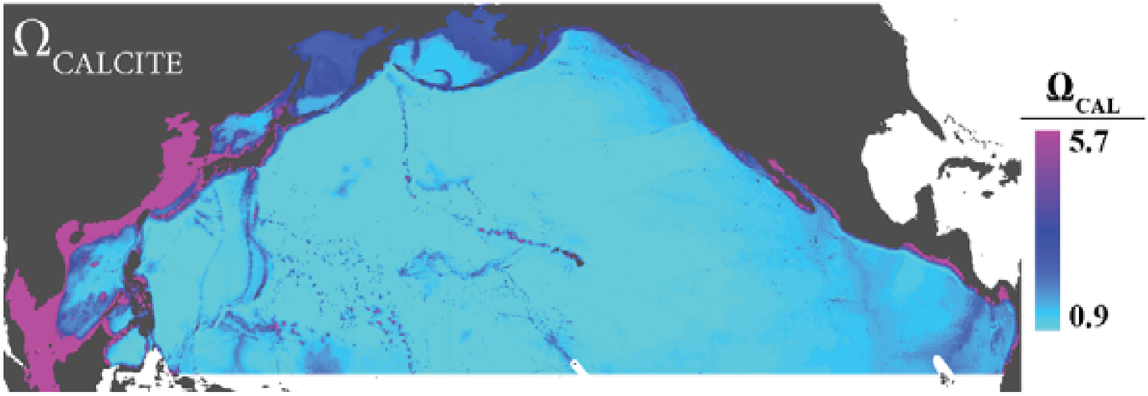

TPI-20000
TPI
3971
-2654

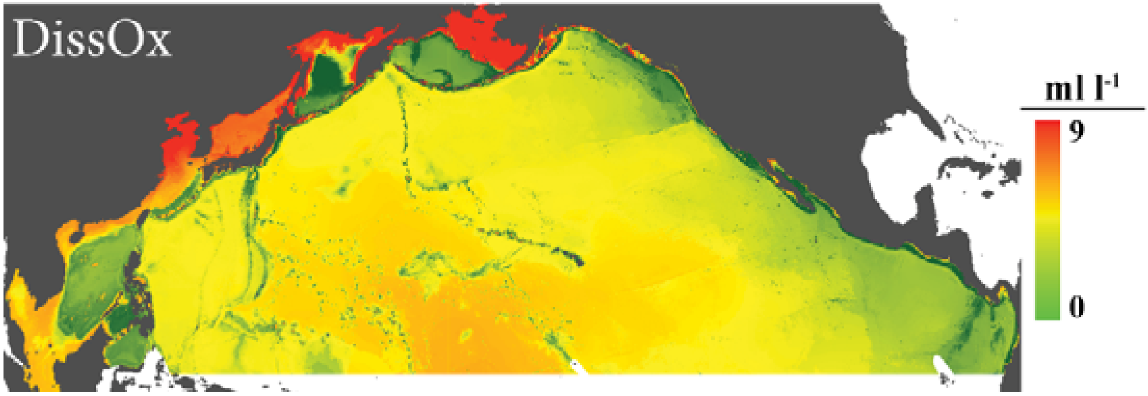

Seamounts
Seamount

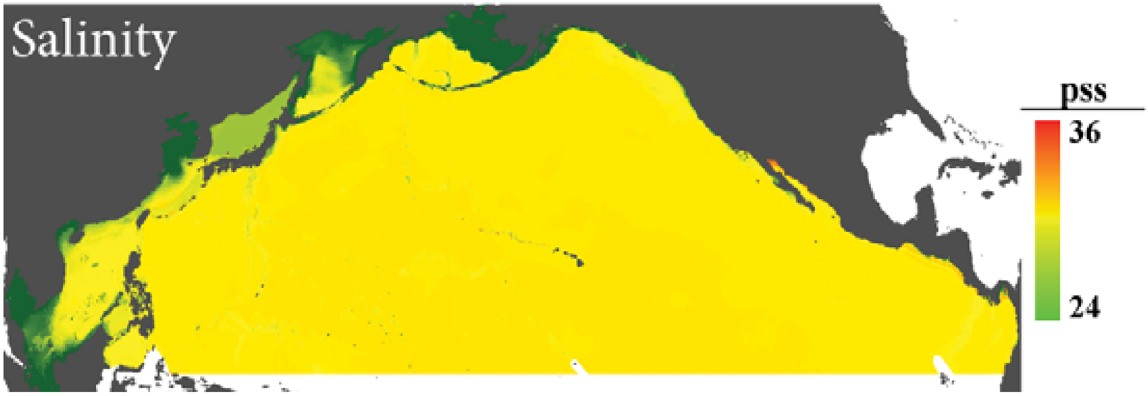

Alkalinity
µmol l-1
2.5
1.8


DIC
µmol l-1
2.4
1.7

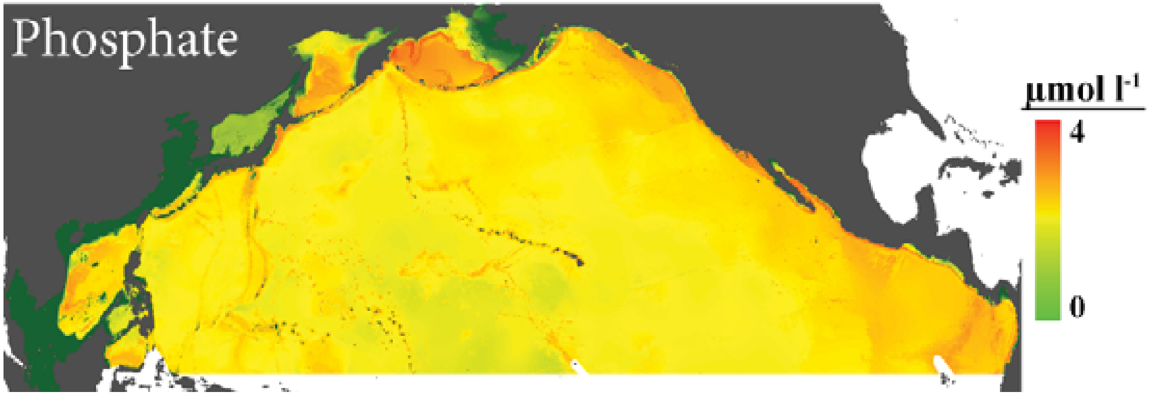

Ω_{ARAG}
Ω_{ARAG}
3.8
0.6

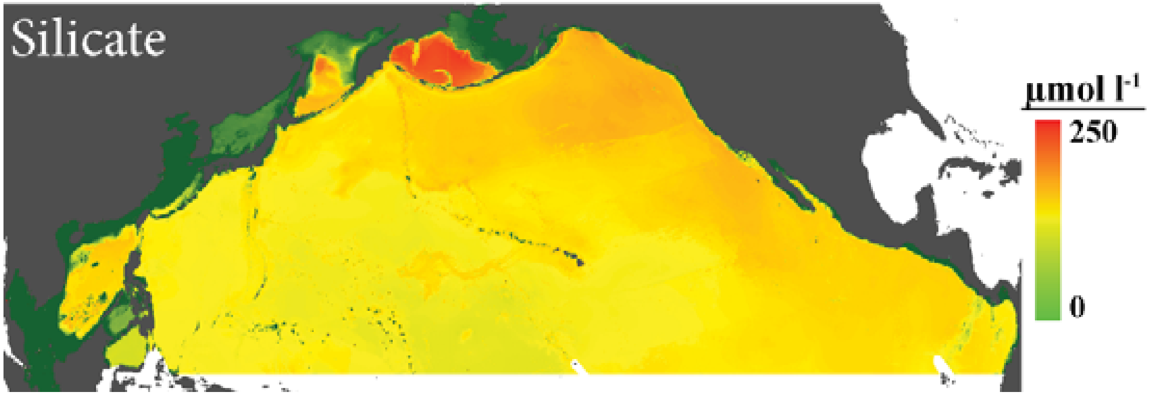

$\Omega_{CALCITE}$
Ω_{CAL}
5.7
0.9

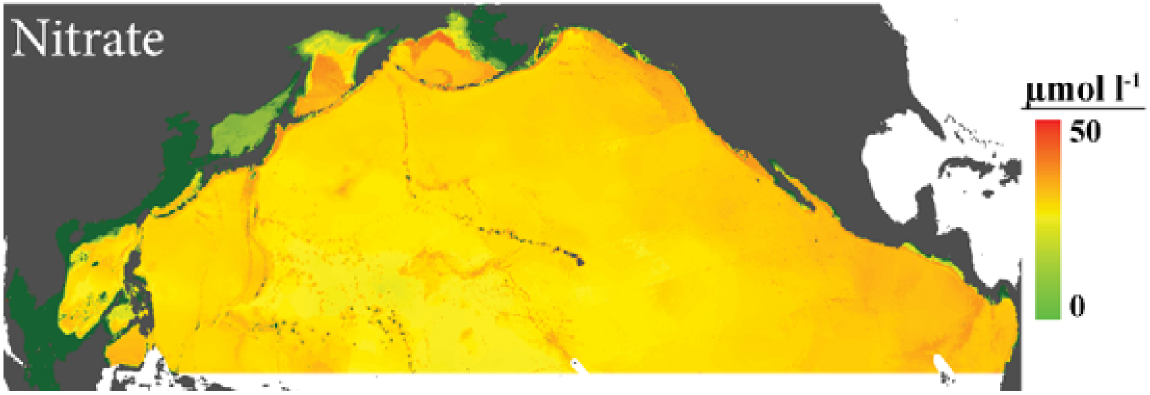

DissOx
ml l^{-1}
9
0

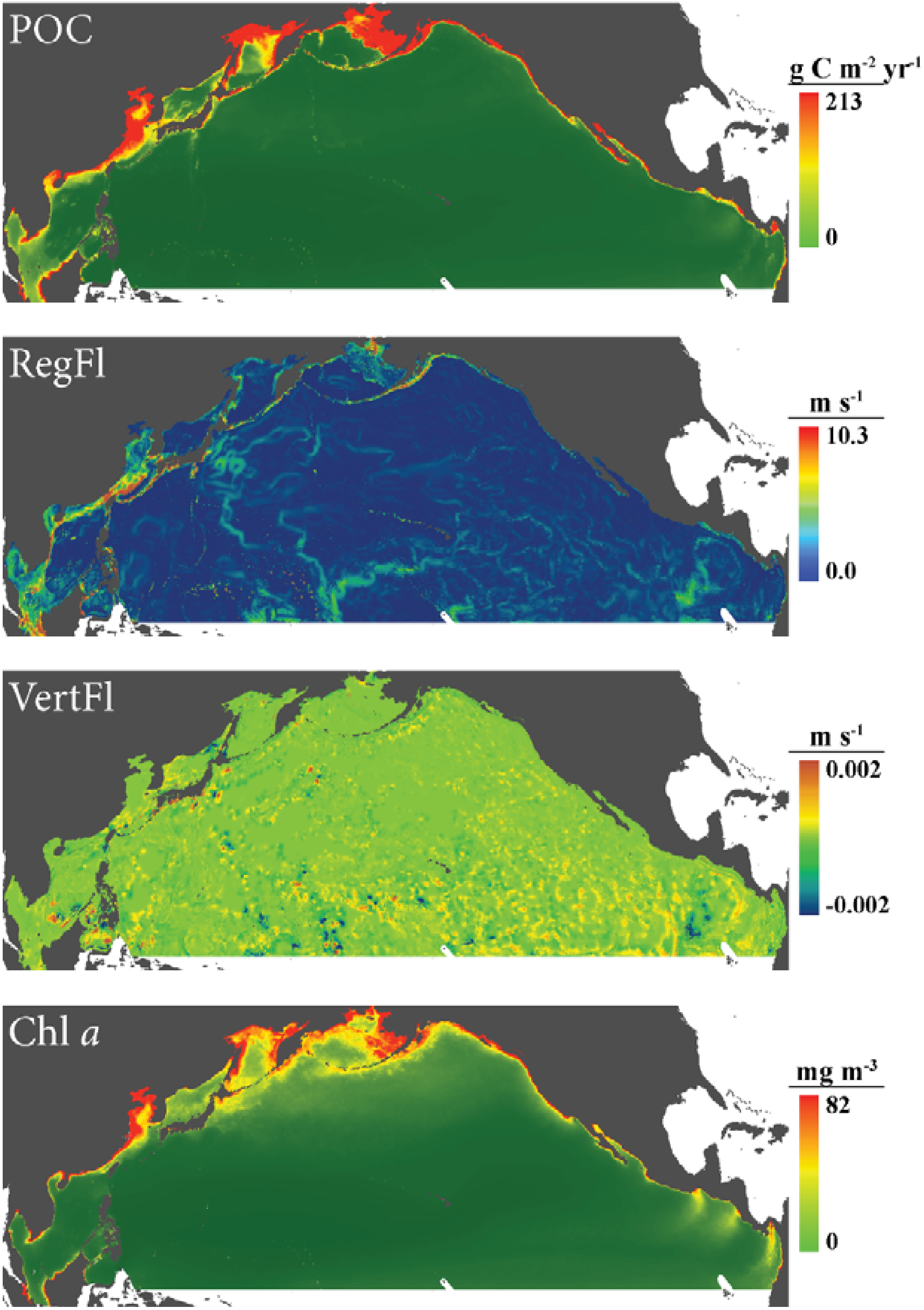

Salinity
pss
36
24


Temp
°C
30
-2


Phosphate
µmol l-1
4
0


Silicate
µmol l-1
250
0


Nitrate
µmol l-1
50
0


POC
g C m-2 yr-1
213
0
RegFl
m s-1
10.3
0.0
VertFl
m s-1
0.002
-0.002
Chl a
mg m-3
82
0

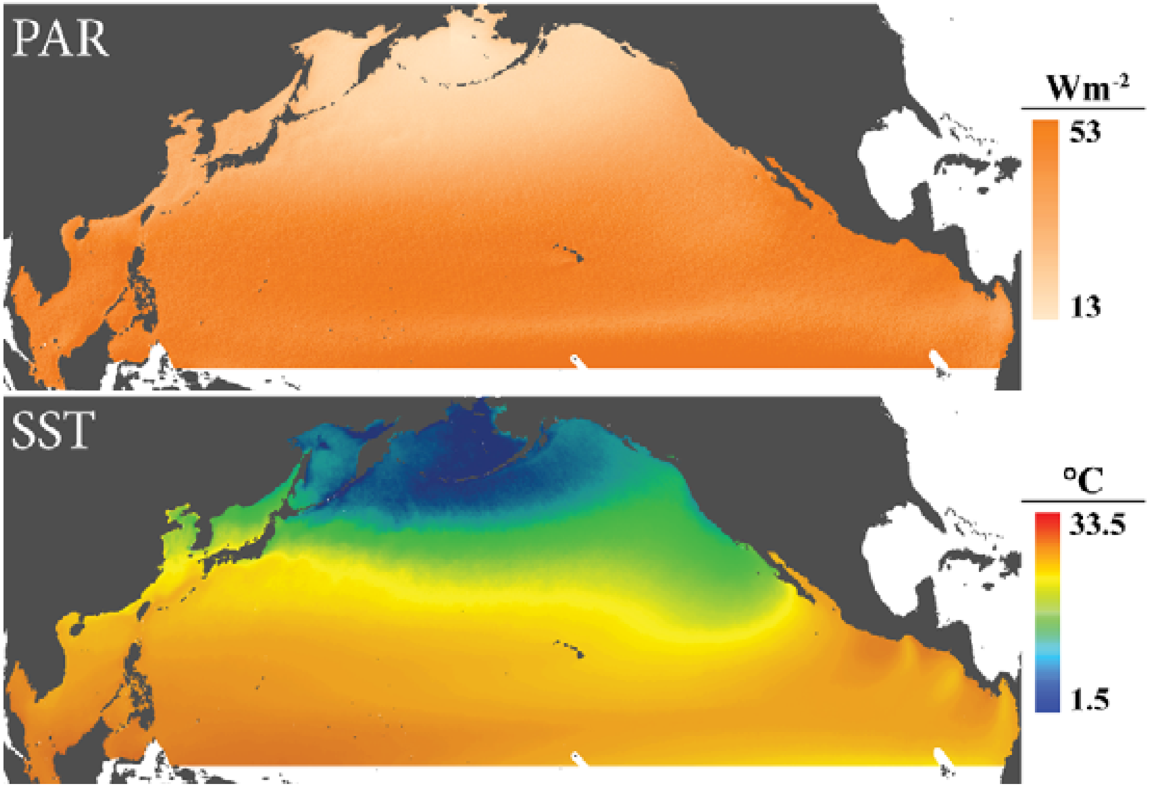
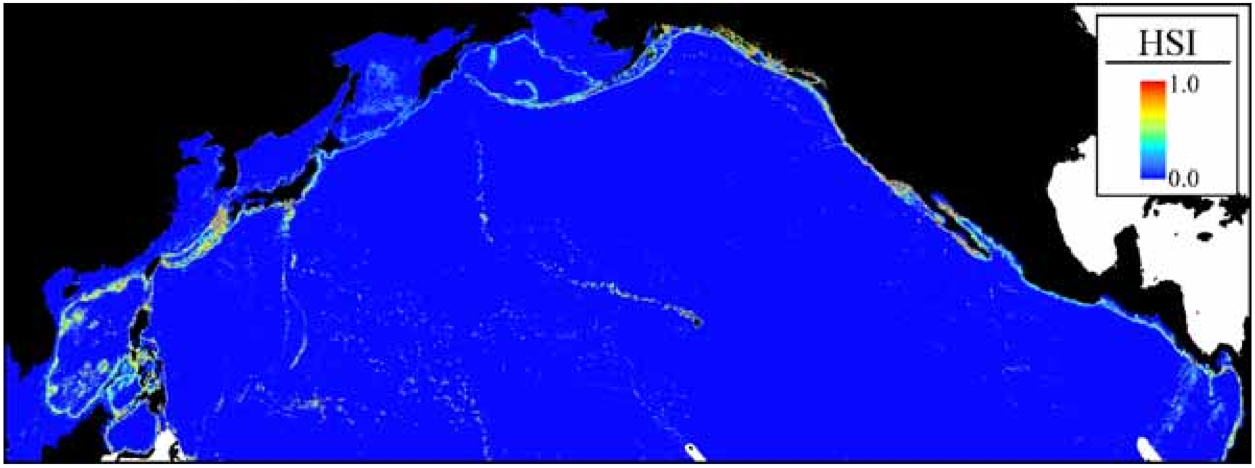
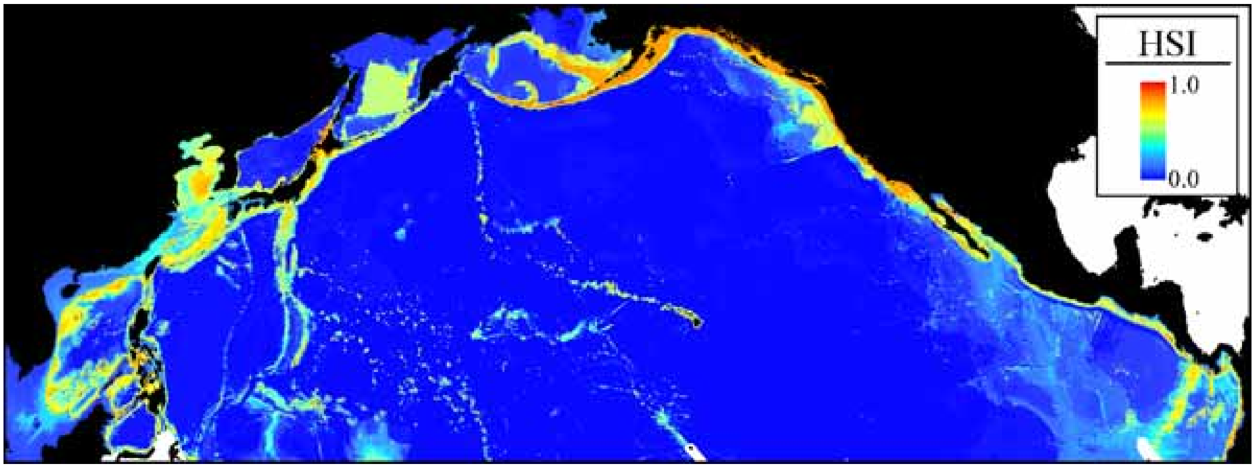


Figure 3. Maps of individual variables available for modeling the distribution of deep-sea corals and sponges in the North Pacific Ocean.

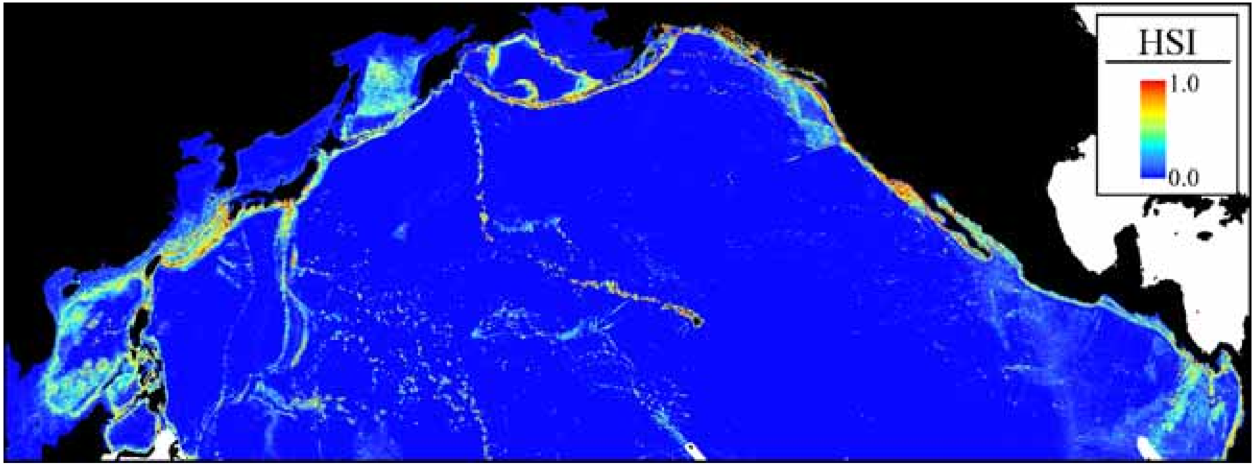
Scleractinia (Stony Corals)



Pennatulacea (Sea Pens)



Antipatharia (Black Corals)



Primnoa

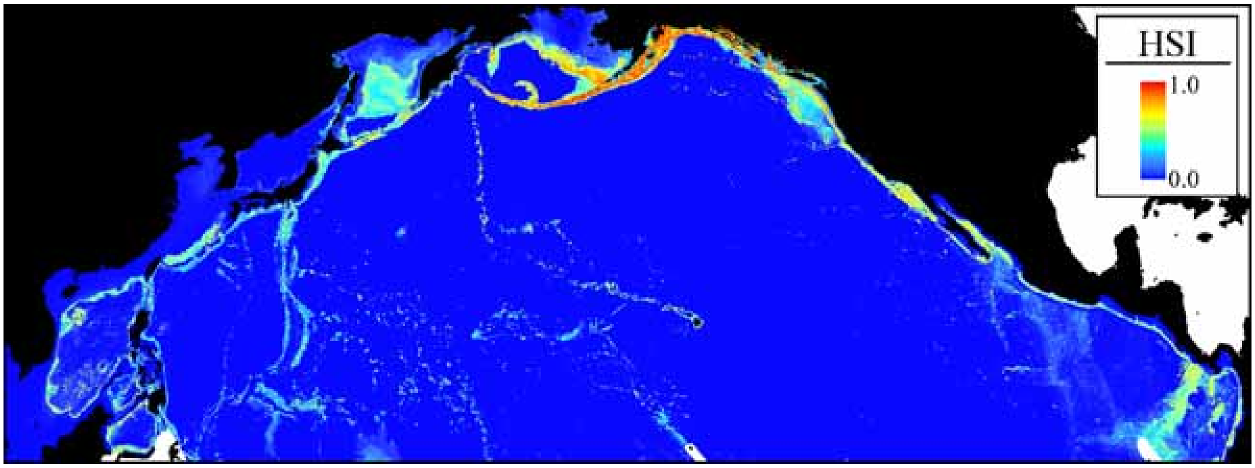


Figure 4. Preliminary habitat suitability models for deep-sea coral taxa in the North Pacific Ocean.

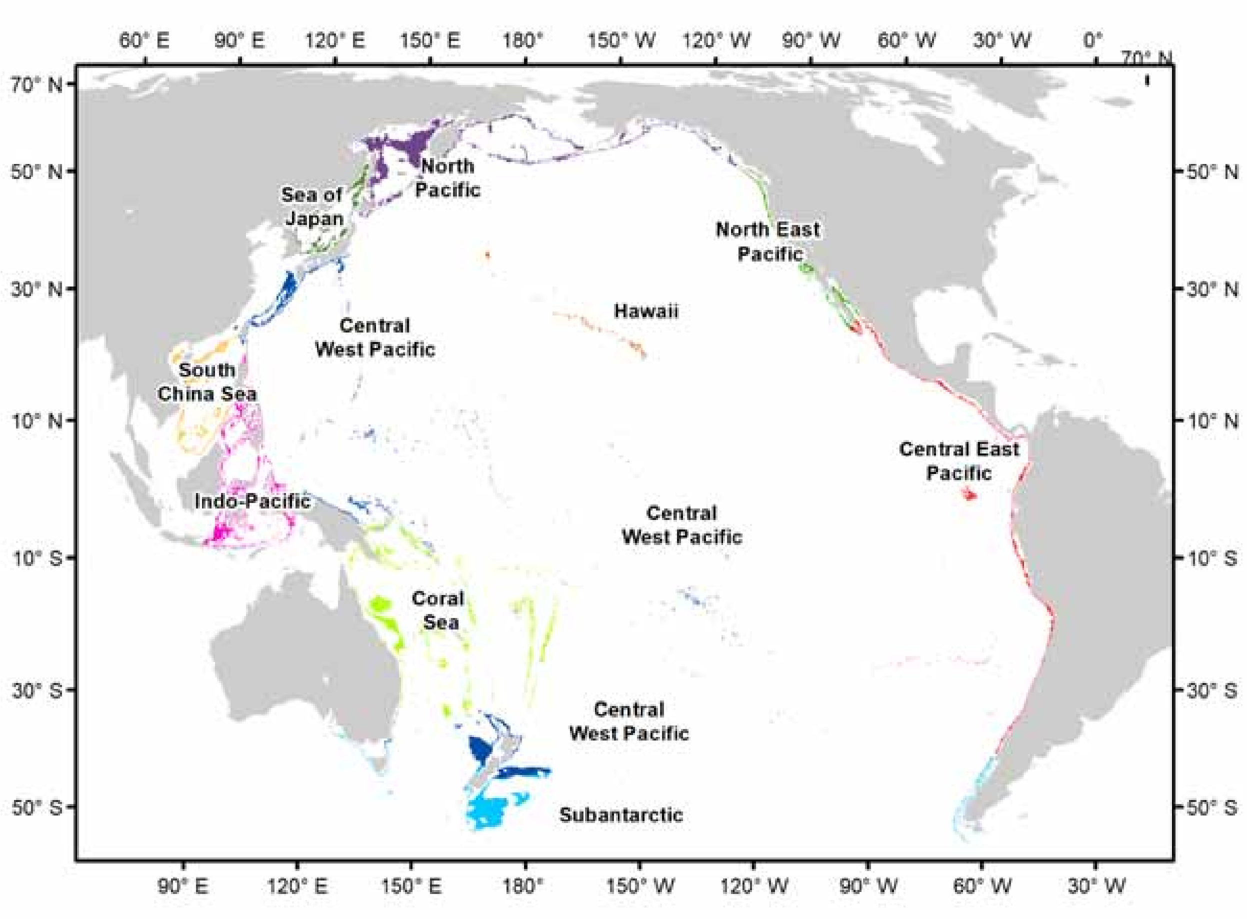
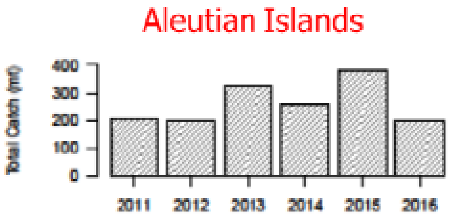
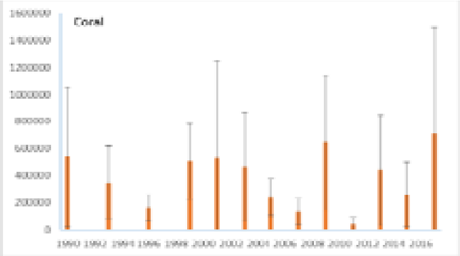
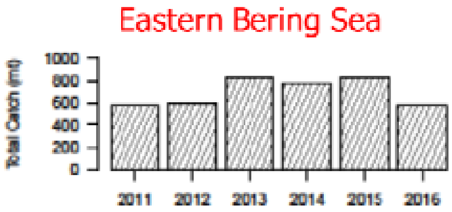
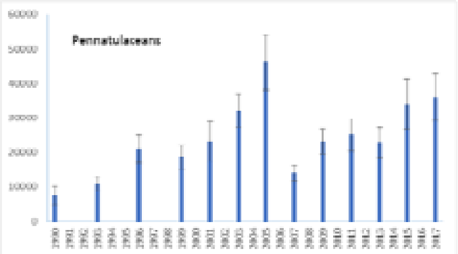


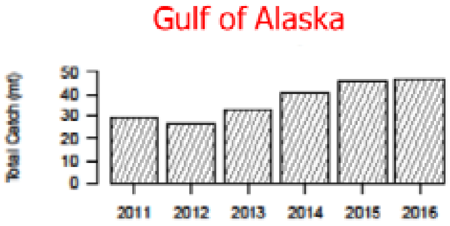
Figure 5. Proposed Upper Bathyal Provinces for the Pacific Ocean based on analysis of 20,000 records of octocorals in the Pacific Ocean.

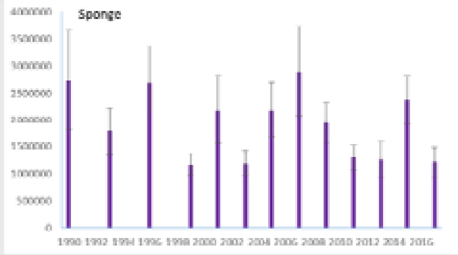

Aleutian Islands
Total Catch (mt)
400
300
200
100
0
2011
2012
2013
2014
2015
2016

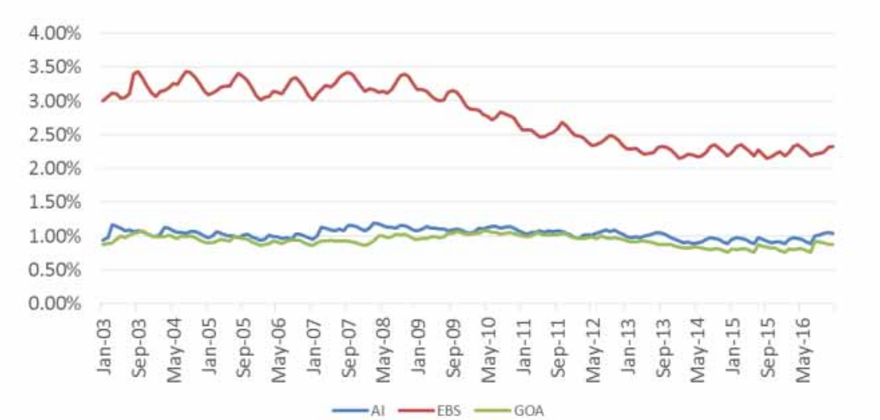

Coral
1600000
1400000
1200000
1000000
800000
600000
400000
200000
0
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016


Eastern Bering Sea
Total Catch (mt)
1000
800
600
400
200
0
2011
2012
2013
2014
2015
2016


Pennatulaceans
60000
50000
40000
30000
20000
10000
0
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017


Gulf of Alaska
Total Catch (mt)
50
40
30
20
10
0
2011
2012
2013
2014
2015
2016


Sponge
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016


4.00%
3.50%
3.00%
2.50%
2.00%
1.50%
1.00%
0.50%
0.00%
Jan-03
Sep-03
May-04
Jan-05
Sep-05
May-06
Jan-07
Sep-07
May-08
Jan-09
Sep-09
May-10
Jan-11
Sep-11
May-12
Jan-13
Sep-13
May-14
Jan-15
Sep-15
May-16
AI
EBS
GOA

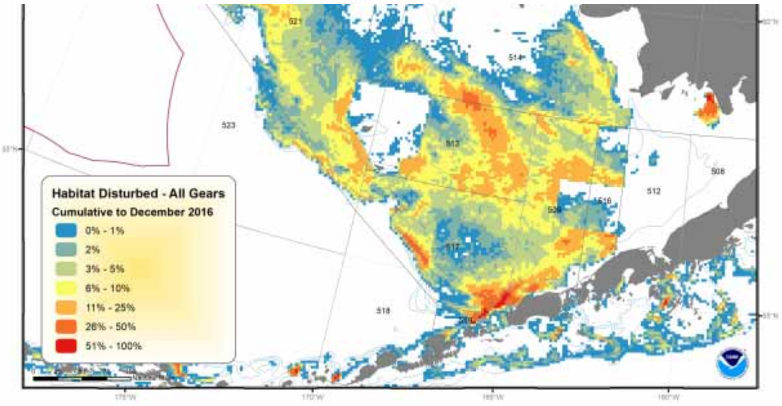

Habitat Disturbed - All Gears
Cumulative to December 2016
0% - 1%
2%
3% - 5%
6% - 10%
11% - 25%
26% - 50%
51% - 100%
521
514
523
513
508
512
509
516
517
518
519

Figure 6. Examples of indices used for monitoring biogenic habitats in Alaska. Total bycatch (panel A) in groundfish fisheries (A. Whitehouse, Alaska Fisheries Science Center, NMFS), time-series of abundance of deep-sea corals and sponges in the Gulf of Alaska (panel B), time series of area impacted by fishing gear in Alaska regions (panel C) and cumulative spatial distribution of fishing effort (panel D) (J. Olson, Alaska Regional Office, NMFS).

Conclusions

WG32 members collaborated effectively with each other and accomplished the overarching vision of contributing a deeper understanding of the diversity and distribution of biogenic habitats in the North Pacific Ocean, and by doing so, enhancing the ability to engage in ecosystem-based fisheries management.

A key outcome of WG32's activities included the identification of large scale environmental and ecological predictors for the distribution and biodiversity of coral, sponge and associated taxa. Predictor variables on a 1 km^2 grid were made available to all WG32 members through a shared drive, which allowed them to use the layers in their own modeling efforts.

Because extensive sampling plans in logistically challenging environments are expensive, data on biogenic habitats will continue to be sparse, particularly in the deep sea, in the immediate future. Hence WG32 reviewed modeling approaches to predict the potential distributions of species and habitat suitability for corals and sponges and identified environmental and ecological predictors of patterns in the distribution and biodiversity of corals, sponges, and associated taxa. WG32 focussed on the technical development of habitat modeling and ecological indicators for data-limited taxa in the deep sea.

MaxEnt habitat models showed higher prediction accuracy at smaller grid cell sizes, and predicted high habitat suitability at such locations as ridges on upper slopes and terrace edges and surface undulation on seamount tops, suggesting the importance of sloped and/or irregular sea floor as habitat of large gorgonian corals. These results demonstrate that it is desirable to obtain bathymetric grid data at resolutions of 100 m or less for the purpose of predicting the distributions of corals at a local scale (e.g. within a seamount). It is important to decide the optimum spatial resolution in consideration of the objectives, data availability and geographical or biological characteristics.

Species distribution models developed by WG32 members and their collaborators suggested that glass sponge reefs require a delicate balance of turbidity and suspended sediment concentration. Another study suggested that the thermal index, UV radiation, and water turbidity were important predictors of the occurrence of coral bleaching, which has affected coral reefs in Japan. That same study found that prediction models combining the effects of climate and ocean currents consistently explained observed community shifts significantly better than those relying on climate alone.

When it comes to applying and interpreting species distribution models in marine ecosystems, WG32 suggested keeping three points in mind:

- Ensemble models can perform better across a range of species than individual models
- The modeling method has less effect on the result than the quality of the underlying predictor and distribution data
- It is important to validate model predictions using independently collected data whenever possible

WG32 members also focused on proposing potential indicators for assessing and monitoring the diversity of biogenic habitats, and documenting associations between commercially important species and biogenic habitats. Systematic monitoring approaches are needed to guide adaptive management strategies for data-limited marine biogenic habitats.

WG32 members demonstrated that habitat-forming species are associated with many commercial fishes and invertebrates. Association analysis demonstrated that Gorgonians frequently co-existed with other benthic animals in the Emperor Seamounts area. One study proposed a new method to assess characteristics of benthic communities and to screen for potential indicator taxa based on the analysis of co-occurrence tendencies among benthic taxa. Other analyses found that areas with more deep-sea corals and sponges had more rockfishes.

The work of WG32 has advanced our knowledge of deep-sea coral and sponge ecosystem distributions by providing tools for modeling presence, absence and abundance of deep-sea corals and sponges. Variables developed in workshops sponsored by WG32 have also been crucial to moving the distribution modeling forward on a relevant scale (i.e. 1 km^2) for management. Since the working group was formed, published models for all regions of the USA EEZ have been developed and used in management. In addition, some key work in the USA has documented the importance of deep-sea coral and sponge ecosystems to supporting productivity of commercially important fish stocks, particularly rockfishes. Current monitoring efforts relative to deep-sea corals and sponges focus primarily on time series of abundance, bycatch and fishing effort. The monitoring could be improved with more focused studies that address the biodiversity of biogenic habitats.

SDMs are one tool that can extrapolate modelled species-environment relationships into areas where species records are rare, thus providing an empirical foundation that can promote hypotheses development which can, in turn, concentrate limited science resources into targeted data collection in logistically challenging environments. Enhancing our community’s ability to better predict where diverse biogenic habitats occurs is an important precursor to understanding how these habitats support other elements of the ecosystem, including commercially valuable species.

Recommendation

WG32 members recommend that PICES engage in further research on biodiversity in the North Pacific Ocean. Specifically, members recommend that PICES convene a new Working Group on *Biodiversity of Seamounts*, with a focus on understanding the distribution of benthic, demersal, and pelagic species that are associated with seamounts. Seamounts are unique habitats and are biodiversity hotspots with high rates of endemism. They can host diverse communities of benthic filter feeders, including corals and sponges. The biodiversity of fishes is also high; almost 800 species of fish have been recorded from seamounts, representing half of the orders of fishes. As such, seamounts are important sources of food. There are approximately 100,000 seamounts worldwide and their abundance is greatest in the North Pacific Ocean. The ecology of only a few seamounts has been studied, in part because of how deep and remote most are. The difficulty in studying the ecology of seamounts means that they are poorly understood habitats in terms of the benthic, demersal, and pelagic species that they support. A Working Group on *Biodiversity of Seamounts* would build on the contributions of WG32 by mapping the distribution of seamount biodiversity and expanding research in some of the unique and abundant ecosystems of the North Pacific Ocean for PICES.

The merits of a new Working Group on *Biodiversity of Seamounts* include (1) the application of concepts developed by WG32, (2) new data to better understand factors that influence the distribution and trends in seamount biodiversity and test key questions about the interactions among taxa that differ in life history (e.g. plankton, filter feeders, fish, mammals), (3) identification of indicators to monitor change, (4) development of hypotheses to forecast responses to multiple stressors, which is aligned with the spirit of FUTURE, (5) maps of the distribution of benthic, demersal, and pelagic biodiversity and its indicators, (5) a new research avenue for PICES with clear linkages to the activities of other PICES activities, including BIO Committee and the NPFC – PICES *Framework for Enhanced Scientific Collaboration in the North Pacific Ocean*.

WG32's focus on biogenic habitat provided a proof of concept on how to undertake collaborative biodiversity research in the North Pacific Ocean. Major applications of the science products developed by a Working Group on *Biodiversity of Seamounts* would be the provision of further technical guidance on the application of SDMs, maps of known and predicted distributions of the benthic, demersal, and pelagic taxa associated with seamounts, and the development of seamount biodiversity indicators.

Annexes

Publications related to WG32 Research

Cairns, Stephen D., Stone, Robert P., Moon, Hye-Won, and Lee, Jong Hee Primnoidae (Octocorallia: Calcaxonia) from the Emperor Seamounts, with Notes on *Callogorgia elegans* (Gray, 1870) Pacific Science, 72(1):125-142. https://doi.org/10.2984/72.1.8

Chu JWF, Nephin J, Georgian S, Knudby A, Rooper C, Gale KSP (2019) Modelling the environmental niche space and distributions of cold-water corals and sponges in the Canadian northeast Pacific Ocean. *Deep-Sea Research Part I*.

Conway K., Dunham A., Archer S.K., Burke L., Kung R., Shaw J. 2019. Pacific Margin Sponge Reefs: Geomorphology and Ecology. In: Seafloor Geomorphology as Benthic Habitat: GeoHab Atlas of seafloor geomorphic features and benthic habitats (2nd edition). Eds: P. Harris, E. Baker. Elsevier Science, USA. 1076 pp.

Dautova T.N., Galkin S.V., Tabachnik K.R., Minin K.V. and Kireev P.A. (2019). The first data on the structure of vulnerable marine ecosystems of the Emperor Chain Seamounts: Indicator taxa, landscapes, and biogeography. Russian Journal of Marine Biology 45(6): 408-417.

Dautova T.N. (2019). Octocorallia as a key taxon in the vulnerable marine ecosystems of the Emperor Chain (Northwest Pacific): diversity, distribution and biogeographical boundary. In: K.A. Lutaenko (ed.). Marine Biodiversity for a Healthy Ocean – Biodiversity, Functional Groups and Ocean Health. Proceedings of the Russia-China Bilateral Workshop, October 10–11, 2019, Vladivostok, Russia. Vladivostok: Publishing House of the Far Eastern Federal University. P. 68–80.

Davidson, Fiona and Knudby, Anders 2019. Mapping Uncertainty of SDM Predictions. In Kenchington, E., Callery, O., Davidson, F., Grehan, A., Morato, T., Appiott, J., Davis, A., Dunstan, P., Du Preez, C., Finney, J., González-Irusta, J.M., Howell, K., Knudby, A., Lacharité, M., Lee, J., Murillo, F.J., Beazley, L., Roberts, J.M., Roberts, M., Rooper, C., Rowden, A., Rubidge, E., Stanley, R., Stirling, D., Tanaka, K.R., Vanhatalo, J., Weigel, B., Woolley, S. and Yesson, C. 2019. Use of Species Distribution Modeling in the Deep Sea. Can. Tech. Rep. Fish. Aquat. Sci. 3296: ix + 76 p.

DFO (2018) Glass sponge aggregations in Howe Sound: locations, reef status, and ecological significance assessment. DFO Can. Sci. Advis. Sec. Sci. Resp. 2018/032

Dunham A, Archer SK, Davies S, Burke L, Mossman J, Pegg J. (2018a) Assessing ecological role of deep-water biogenic habitats: Glass sponge reefs in the Salish Sea. *Marine Environmental Research* 141: 88-99.

Dunham A, Mossman J, Archer S, Pegg J, Davies S, Archer E. (2018b) Glass sponge reefs in the Strait of Georgia and Howe Sound: status assessment and ecological monitoring advice. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/010

Lehnert, Helmut and Stone, Robert P. (2019) Two new species of Geodiidae (Porifera, Demospongiae, Astrophorina) from the Emperor Seamounts, North Pacific Ocean. Zootaxa 4671(3): 381-395

Loh T-L., Archer S.K., Dunham A. 2019. Monitoring program design for data-limited marine biogenic habitats: A structured approach. *Ecology and Evol*ution 9: 7346–7359.

Kumagai, N.H., García Molinos, J., Yamano, H., Takao, S., Fujii, M., and Yamanaka, Y. (2018a) Ocean currents and herbivory drive macroalgal-coral community shift under climate warming. *Proceedings of the National Academy of Sciences of the United States of America*, 115, 8990-8995.

Kumagai, N.H., Yamano, H., and Sango Map Project Committee (2018b) High-resolution modeling of thermal thresholds and environmental influences on coral bleaching for local and regional reef management. *PeerJ*, 6: e4382; DOI 10.7717/peerj.4382

Loh T-L, Archer SK, Dunham A. Monitoring program design for data-limited marine biogenic habitats: A structured approach. *Ecol. Evol*. 2019; 9: 7346–7359.

Miyamoto M., Kiyota M., Murase H., Nakamura T., Hayashibara T. (2017) Effects of Bathymetric Grid-Cell Sizes on Habitat Suitability Analysis of Coldwater Gorgonian Corals on Seamounts, Marine Geodesy 40: 205-223.

Miyamoto M., Kiyota M. (2017) Application of association analysis for identifying indicator taxa of vulnerable marine ecosystems in the Emperor Seamounts area, North Pacific Ocean. Ecological Indicator 78: 301-310.

Miyamoto M, Kiyota M, Hayashibara T, Nonaka M (2017) Megafaunal composition of cold-water corals and other deep-sea benthos in the southern Emperor Seamounts area, North Pacific Ocean 19-30.

Rooper CN, Zimmermann M, Prescott MM (2017) Deep-Sea Research Part I Comparison of modeling methods to predict the spatial distribution of deep-sea coral and sponge in the Gulf of Alaska. *Deep-Sea Research Part I*, **126**, 148–161

Shinzato C, Zayasu Y, Kanda M, Kawamitsu M, Satoh N, Yamashita H and Suzuki G (2018) Using Seawater to Document Coral-Zoothanthella Diversity: A New Approach to Coral Reef Monitoring Using Environmental DNA. Front. Mar. Sci. 5:28. doi: 10.3389/fmars.2018.00028