Serial No. N7663 NAFO SCR Doc. 25/035

SCIENTIFIC COUNCIL MEETING –2025

Vulnerable Marine Ecosystems in the NAFO Regulatory Area: Updated Species Distribution Models of Selected Vulnerable Marine Ecosystem Indicators (Large and Small Gorgonian Corals, Erect Bryozoans and Sea Squirts)

by

F.J. Murillo[1], S. Abalo Morla[2], A.-L. Downie[3], C. Lirette[1], N. Paulin[1], Z. Wang[1], E. Devred[1], S. Clay[1], M. Sacau[2], C. Nozères[1], M. Koen-Alonso[4], L. Gullage[4], V. Hayes[4], M. Caetano[5], P. Gonçalves[5], and E. Kenchington[1]

[1]Department of Fisheries and Oceans, Bedford Institute of Oceanography, Dartmouth, NS, Canada.

[2]Instituto Español de Oceanografía (COV-IEO), CSIC, Vigo, Spain.

[3]Centre for Environment, Fisheries and Aquaculture Science, Lowestoft, UK.

[4] Department of Fisheries and Oceans, Northwest Atlantic Fisheries Centre, St. John's, NL, Canada.

[5] Instituto Português do Mar e da Atmosfera (IPMA), Lisbon, Portugal.

MURILLO, F.J., ABALO MORLA, S., DOWNIE, A.-L., LIRETTE, C., PAULIN, N., WANG, Z., DEVRED, E., CLAY, S., SACAU, M., NOZÈRES, C., KOEN-ALONSO, M., GULLAGE, L.,. HAYES, V., CAETANO, M., GONÇALVES, P., & KENCHINGTON, E. 2025. Vulnerable Marine Ecosystems in the NAFO Regulatory Area: Updated Species Distribution Models of Selected Vulnerable Marine Ecosystem Indicators (Large and Small Gorgonian Corals, Erect Bryozoans and Sea Squirts). *NAFO Scientific Council Research Document*, SCR Doc. 25/035: 1-62.

Abstract

The Northwest Atlantic Fisheries Organization (NAFO) Commission has called for a reassessment of the vulnerable marine ecosystems (VMEs) and impact of bottom fisheries on VMEs for 2027. Species distribution models (SDMs) help to inform on the closed area boundaries and have been used to modify the areas of significant concentrations of Large-Sized Sponges and Large Gorgonian Corals produced through kernel density analyses (KDE) in the previous review completed in 2019. Here we provide Random Forest SDMs for the Large Gorgonian Corals, Small Gorgonian Corals, Erect Bryozoans and Sea Squirts, completing the distribution modeling for the VME Indicator taxa which was presented to the 2024 meeting of the Working Group on Ecosystem Science and Assessment (WG-ESA). As was done for the Large-Sized Sponges, Sea Pens and Black Corals in 2024, we provide maps of uncertainty associated with the areas of predicted presence and absence. For the Small Gorgonian Corals we had sufficient data to model the distributions of subsets of data for each, viz. two species, *Acanella arbuscula* and *Radicipes gracilis*. Additionally, we provided a SDM for Erect Bryozoan VME habitat as well as the Erect Bryozoan Functional Group. To model the VME habitat we used the data above the catch threshold used to delineate the VME polygons in the kernel density analyses performed in 2019, using research vessel catches ≥ 0.2 kg of erect bryozoans. This modeling was performed to focus the large area of suitable habitat for erect bryozoans on the smaller areas where they form VMEs.

Introduction

Species Distribution Modeling

Species distribution models (SDMs) predict the presence, absence, or abundance/biomass of a species or habitat (the response variable) from environmental variables (the predictor variables) thought to influence it. Potential uses of SDMs include 1) explanation, 2) mapping, and 3) transfer (Zurell et al., 2020), with the first focused on identifying the main factors driving the species distributions, the second on producing maps of the distribution, and the third on forecasting or projecting the distributions into a different geographic region or time period. The primary objective of the SDMs presented here is 'Mapping' (Zurell et al., 2020). The maps will also be used to evaluate the area between trawl sets to determine if the full vulnerable marine ecosystem (VME) polygon (derived from kernel density analyses (KDE), which does not consider environmental variables (Kenchington et al., 2019)) is a potential habitat, and to modify the boundaries of the VME polygons if they include areas of predicted species absence. The latter was previously done to modify the VME polygons for Large-Sized Sponges and Large Gorgonian Corals (NAFO, 2019). At the 2024 meeting of the Northwest Atlantic Fisheries Organization (NAFO) Working Group on Ecosystem Science and Assessment (WG-ESA), maps showing the predicted distributions of Large-Sized Sponges, Sea Pens and Black Corals were presented, along with those for a number of subgroups of Sponges (i.e., sub-order Astrophorina, the families Tetillidae and Polymastiidae and sponge grounds (catches above a weight threshold from the kernel density analyses (KDE)) and Sea Pens (i.e., genera *Balticina*, *Funiculina*, *Anthoptilum* and *Pennatula*) (Murillo et al., 2024).

SDMs for Large Gorgonian Corals (Knudby et al., 2013), Erect Bryozoans and Sea Squirts (*Boltenia ovifera*) (Kenchington et al., 2019) have previously been incorporated into the NAFO assessment of VMEs (NAFO, 2019), however, the Small Gorgonian Corals have not previously been assessed. In support of the 2024 NAFO Commission Request#6, b: *Work towards the reassessment of VMEs and impact of bottom fisheries on VMEs for 2027*, SDMs have been created using a common set of environmental predictors and response variables updated to include data through to 2023 (Murillo et al., 2024). We followed the Overview/Conceptualisation, Data, Model fitting, Assessment and Prediction (ODMAP) steps recommended by Zurell et al. (2020) for standard reporting of SDMs, complemented by the recommendations of Sofaer et al. (2019) for the use of SDMs in decision-making.

Here we present SDMs for the VME functional groups Large Gorgonian Corals, Small Gorgonian Corals, Erect Bryozoans and Sea Squirts (mostly a single species, *Boltenia ovifera*). For the Erect Bryozoans, we have modeled the distribution of the Erect Bryozoan VME habitat as determined through use of the kernel density analysis (KDE) threshold established in 2019 (Kenchington et al., 2019). Within the Small Gorgonian Corals we also present models for *Acanella arbuscula* and *Radicipes gracilis*. These additional models will be used to compare the results of the predicted distributions of individual taxa versus that of their functional group. They also will be used to assess whether the current closed areas and fishing impacts differentially impact these species. We follow the methods applied to the VME functional groups Large-Sized Sponges, Sea Pens and Black Coral (Murillo et al., 2024).

Methods

Environmental Data

All layers were displayed using a NAD83 UTM 23N projection and the resolution of the final raster surfaces was 1 km. The spatial extent of the modelled area is bounded by the Canadian Exclusive Economic Zone (EEZ) to the west, and to the north, south and east by the 2500 m depth contour (derived from GEBCO 2024, see below). This area is referred to as the NAFO Regulatory Area (NRA) and includes Flemish Cap and the Nose and Tail of Grand Bank.

Water column variables

Environmental layers representing water column properties were the physical oceanographic variables bottom temperature, bottom salinity, bottom current speed, bottom stress, mixed layer depth, surface temperature, surface salinity, and the biological oceanographic variables chlorophyll *a* and primary productivity (Table 1). Monthly temperature, salinity, current speed, bottom stress and mixed layer depth were extracted from the Bedford Institute of Oceanography North Atlantic Model (BNAM; Wang et al., 2018) for the period 1990-2023. Mean, maximum, minimum and range values derived from BNAM were calculated for all months within a year and averaged across all years. Bottom stress (τ_b) was calculated as outlined in Murillo et al. (2024). For mixed layer depth, only maximum values were calculated as above, averaged across the time period and for seasonal time periods (Winter: January – March; Spring: April - June; Summer: July - September; Fall: October – December). Using ArcGIS Pro's Geostatistical Wizard, BNAM (and BNAM-derived) point data were interpolated using ordinary kriging, and the resulting geostatistical layers were exported to the final raster surfaces. Salinity is considered unitless because it's defined as a ratio of conductivity, rather than a direct mass measurement.

Table 1. Water column variables used in the Random Forest models (Max: Maximum; Min: Minimum; MLD: Mixed Layer Depth; Chl: Chlorophyll; PP: Primary Production; BNAM: Bedford Institute of Oceanography North Atlantic Model (Wang et al., 2018); SOPhyE: Satellite Ocean Colour and Phytoplankton Ecology Group at the Bedford Institute of Oceanography).

Variable	Metric	Unit	Native Resolution	Source
Bottom Salinity	Mean, Max, Min, Range	N/A	1/12º lat/long	BNAM
Bottom Temperature	Mean, Max, Min, Range	°C	1/12º lat/long	BNAM
Bottom Current Speed	Mean, Max, Min, Range	m s^{-1}	1/12º lat/long	BNAM
Bottom Stress	Mean, Max, Min, Range	m s^{-1}	1/12º lat/long	BNAM
Surface Salinity	Mean, Max, Min, Range	N/A	1/12º lat/long	BNAM
Surface Temperature	Mean, Max, Min, Range	°C	1/12º lat/long	BNAM
Surface Current Speed	Mean, Max, Min, Range	m s^{-1}	1/12º lat/long	BNAM
Annual MLD	Max	m	1/12º lat/long	BNAM
Summer MLD	Max	m	1/12º lat/long	BNAM
Fall MLD	Max	m	1/12º lat/long	BNAM
Winter MLD	Max	m	1/12º lat/long	BNAM
Spring MLD	Max	m	1/12º lat/long	BNAM
Annual Chl *a*	Mean, Max, Min, Range	mg m^{-3}	4 km	SOPhyE
Spring Chl *a*	Mean, Max, Min, Range	mg m^{-3}	4 km	SOPhyE
Fall Chl *a*	Mean, Max, Min, Range	mg m^{-3}	4 km	SOPhyE
Winter Chl *a*	Mean, Max, Min, Range	mg m^{-3}	4 km	SOPhyE
Summer Chl *a*	Mean, Max, Min, Range	mg m^{-3}	4 km	SOPhyE
Fall PP	Mean, Max, Min, Range	mg C m^{-2} day^{-1}	4 km	SOPhyE
Winter PP	Mean, Max, Min, Range	mg C m^{-2} day^{-1}	4 km	SOPhyE
Summer PP	Mean, Max, Min, Range	mg C m^{-2} day^{-1}	4 km	SOPhyE
Spring PP	Mean, Max, Min, Range	mg C m^{-2} day^{-1}	4 km	SOPhyE
Annual PP	Mean, Max, Min, Range	mg C m^{-2} day^{-1}	4 km	SOPhyE

Daily photosynthetically active radiation (PAR) and surface chlorophyll *a* concentration data was calculated as detailed in Murillo et al. (2024). Averaged across years and seasonal periods, the mean, maximum, minimum and range values were derived from the 8-day chlorophyll *a* and primary production composites. Seasons were delimited in the following manner: Winter: Jan 01 to Mar 29; Spring: Mar 30 to Jul 03; Summer: Jul 04 to Sep 29; Fall: Sep 30 to Dec 31. As with the BNAM data, the resulting statistical layers were interpolated using ordinary kriging and the geostatistical layers were exported to the final raster surfaces.

Terrain variables

GIS tools from the R package MultiscaleDTM (Ilich et al., 2023) and the System for Automated Geoscientific Analyses (SAGA) (v. 8.4.1; Conrad et al., 2015) accessed with the R package RSAGA (Brenning et al., 2022) were used to calculate terrain variables (Table 2) in the free statistical computing software R (v. 4.3.2; R

Development Core Team, 2023). Terrain variables were derived from a digital elevation model (DEM) produced from the 15 arc-second gridded General Bathymetric Chart of the Oceans (GEBCO) 2024 (GEBCO Compilation Group, 2024) covering the modelled area (NRA). The bathymetric horizontal resolution corresponds to approximately 388 m at the study area's latitude. The GEBCO bathymetry data layer was then projected onto NAD83 UTM23N using the 'terra' R package function with EPSG:26923 (Hijmans, 2024). The SAGA 'Fill sinks' tool (Wang and Liu, 2006), accessed using RSAGA with a slope threshold of 0.005, was used to smooth out artefacts in the GEBCO DEM before calculating the derivative terrain layers. Further details are provided in Murillo et al. (2024).

All resulting terrain variable layers were then transformed to match the 1-km resolution and origin of other environmental data raster layers with the 'resample' function from the raster R package (Hijmans, 2023) using a bilinear interpolation method. Layers were then cropped and masked to the study area extent.

Table 2. Description of terrain variable layers calculated from GEBCO 2024 bathymetry data.

Variable	**Short name**	**Unit**	**R package**	**RSAGA library**	**SAGA module/MultiscaleDTM function**	**Arguments**
Fill-sink bathymetry*	FS005	m	RSAGA	ta_preprocessor	Fill Sinks (Wang & Liu, 2006)	MINSLOPE = 0.005
Slope	SLOPE	degrees	RSAGA	ta_morphometry	Slope, Aspect, Curvature	UNIT_SLOPE = 1
Bathymetric Position Index (fine-scale)	BPIF	index	MultiscaleDTM	N/A	BPI	w = c(4, 8)
Bathymetric Position Index (broad-scale)	BPIB	index	MultiscaleDTM	N/A	BPI	w = c(4,64)
Ruggedness	VRM	index	RSAGA	ta_morphometry	Vector Ruggedness Measure (VRM)	MODE = 0, RADIUS = 3
Eastness (aspect)	EAST	radians	RSAGA	ta_compound	Basic Terrain Analysis	
Northness (aspect)	NORTH	radians	RSAGA	ta_compound	Basic Terrain Analysis	
Channel Network Base Level (3 & 5)	CHNETBL3/5	m	RSAGA	ta_compound	Basic Terrain Analysis	THRESHOLD = 3 & 5
Channel Network Distance (3 & 5)	CHNETD3/5	m	RSAGA	ta_compound	Basic Terrain Analysis	THRESHOLD = 3 & 5
Valley Depth (3)	VALD	m	RSAGA	ta_compound	Basic Terrain Analysis	THRESHOLD = 3
Relative Slope Position (3 & 5)	RSP3/5	index	RSAGA	ta_compound	Basic Terrain Analysis	THRESHOLD = 3 & 5
LS-Factor	LSF	index	RSAGA	ta_compound	Basic Terrain Analysis	
Positive and Negative Openness	POP/NOP	radians	RSAGA	ta_lighting	Topographic Openness	
Wind Exposition Index	WEI	index	RSAGA	ta_morphometry	Wind Exposition Index	

*Used as the digital elevation model (DEM) for all the other variables requiring a DEM input layer.

Fishing effort variables

Two methods were used to produce fishing effort layers (NAFO, 2019; Murillo et al., 2024). The first method used both bottom trawling and bottom longline effort data resolved at a native resolution of 0.05 degrees (approximately 3.8 x 5.6 km^2) and represented effort as hours fished per grid cell. The second method used only bottom trawling data, was resolved at a native resolution of 1 km^2, and represented effort as km trawled per km^2 per year. Both fishing effort layers were used in the analyses as they capture distinct aspects of fishing pressure at different scales of resolution (Murillo et al., 2024).

Biological data

The data records used for the response data in the SDMs were drawn from the research vessel trawl surveys conducted by Canada and the European Union (EU) (Spain and Portugal) from trawl sets in the NRA on Flemish Cap and the Nose and Tail of Grand Bank to 2500 m (Table 3). Over time, the at-sea identification and coding of the VME Indicator taxa has evolved and different identifications were reviewed and consolidated for each modeled taxon (Appendix Table A1). The time frame for the identification of the lower level taxa differs among surveys and is indicated in the descriptions for each taxon below and in the Appendices (Appendix Tables A2 and A3). Initially, data from Canada and the EU were identified only by the functional group attribution and not the taxon name or species code, even if such data were recorded at sea. That was because the data were used for the KDE analysis and identification of VME polygons at the functional group level (Kenchington et al., 2014), and finer taxonomic resolution was not needed. Those earlier records could be reviewed to ensure that the taxon names were consistent with the functional groups used today, however, as there are sufficient records that have been validated with taxon names (Appendix Table A1) for SDMs, those earlier records were evaluated on a case-by-case basis for inclusion in the models (see below for details for each functional group). Data listed in Appendix Table A1 include the raw data provided by the NAFO contracting parties. These could include different taxa at different taxonomic resolutions (e.g., *Paragorgia arborea*, *Keratoisis*) in the same trawl set that can be grouped differently depending on the taxa modelled. Once the data were evaluated, some were combined as a single record for the response data if they belonged to the same set (e.g., *Acanella*, *Acanella arbuscula*). As a result, the sum of records from the rows in Appendix Tables A2 and A3 is not always equal to the number of records in Table 4 which presents the response data used in the models.

Survey data were used to record both species biomass (kg), and presence or absence. Absence data at the functional group level (i.e., Large Gorgonian Corals, Small Gorgonian Corals, Erect Bryozoans, Sea Squirts) were determined on a tow-by-tow basis for each survey that recorded the presence of the functional group amongst their trawl sets. The assumption was that if the functional group was recognized and recorded on the survey its absence was not likely due to identification issues. For some functional groups where presence was not consistently recorded in the earlier years the associated absence (null) data were excluded from the SDMs. The same procedure was used to identify null data for the SDMs of the lower level taxa within each functional group. All nulls for the functional group were used, in addition to null data where the subgroups were not observed. This was necessary to fill gaps within functional group distributions where particular taxa may not occur. Once this data set was produced, subgroup-specific nulls were extracted. A pivot table was created for each set and nulls were calculated for each subgroup, so the number of nulls will differ by taxon. As an additional check, the number of presences by year (Appendix Tables A2 and A3) was examined to ensure that there were no trends in recording the taxon prior to accepting subgroup null data. Additionally, there were some questionable records based on the known depth in the study area of Small Gorgonian Corals and Sea Squirts from both EU and Canadian Surveys that could be associated with contamination from previous trawl sets (Kenchington et al., 2016). Despite efforts to avoid this problem, such as the nets are sprayed down with a firehose as the total catch is dumped into the fish hole, and the deck crew are diligent about checking the net to ensure everything is deposited into the fish hole, contamination still can happen for a reduced number of sets and therefore some records were removed from the analysis due to uncertainty around their positional accuracy (Appendix Tables A4 and A5). The data has been archived on the NAFO SharePoint site.

Table 3. Research Vessel Survey Data from NAFO Contracting Parties (EU and Canada); EU, European Union; DFO, Department of Fisheries and Oceans; NL, Newfoundland and Labrador; IEO, Instituto Español de Oceanografía; IIM, Instituto de Investigaciones Marinas; IPMA, Instituto Português do Mar e da Atmosfera.

Data Source	**Period**	**NAFO Division**	**Gear**	**Mesh Size in Codend Liner (mm)**	**Trawl Duration (min)**	**Average Wingspread (m)**
Spanish 3NO Survey (IEO)	2002 - 2023	3NO	Campelen 1800	20	30	24.2 – 31.9
EU Flemish Cap Survey (IEO, IIM, IPMA)	2003 - 2023	3M	Lofoten	35	30	13.89
Spanish 3L Survey (IEO)	2003 - 2023	3L	Campelen 1800	20	30	24.2 – 31.9
DFO NL Multi-species Surveys (DFO)	1995 - 2022	3LNO	Campelen 1800	12.7	15	15 - 20

Large Gorgonian Corals

The available raw data for the SDM models for the VME Large Gorgonian Coral Functional Group included 337 presence records (Appendix Table A1) and 7454 null records obtained from the surveys shown in Table 3. Of the presence records, 134 were provided without a taxon name associated with the data. The largest proportion (37%) of the 203 records with a taxon name were members of the family Acanthogorgiidae, most likely *Acanthogorgia armata*. However, there were too few records of any individual taxon to warrant modeling by subgroup.

The oldest record was collected in 2000 in the Canadian survey, but corals were not reported again until 2005, also by Canada (Wareham and Edinger, 2007). Records from the EU surveys were received from 2006, although the taxon names were not provided. After 2011, taxon names were provided from the EU surveys, although some are missing from some sets in data provided by both the EU and Canada. The final biological data used for the response data in the Large Gorgonian Coral SDMs included 161 presence/biomass records from 2011-2023, and 5279 associated null data (Table 4).

Small Gorgonian Corals

The available raw data for the SDM models for the VME Small Gorgonian Coral Functional Group included 950 presence records (Appendix Table A1) and 7139 null records obtained from the surveys shown in Table 3. Of the presence records, 762 have taxon names associated with them and 188 recorded presences with no taxon name provided (Appendix Table A1). Those latter records were almost all recorded between 2005 and 2010, mostly by the EU, with one EU record from 2013. Data collection was inconsistent in the early years (Appendix Table A2) with one record from 2002 and no data collected in 2003 and 2004. Data are sparse until 2006 and are more consistent thereafter. However, to ensure a consistent timeframe was being considered for each of the data sources, Canadian records with an associated taxon name prior to 2011 were not included in the analyses.

Two subgroups have sufficient data to run separate models: the species *Acanella arbuscula* and *Radicipes gracilis*. Records for *Acanella arbuscula* included 560 records recorded as ‘Acanella’, ‘Acanella arbuscula’ or ‘ACANELLA ARBUSCULA’. Records for *Radicipes gracilis* included 153 records recorded as ‘Radicipes’, ‘Radicipes gracilis’, ‘RADICIPES SP’, ‘Radicipes sp.’ and ‘Radicipes spp’.

The breakdown of the subgroup and functional group records by year is provided in Appendix Table A3. There are very few records of *Radicipes gracilis* prior to 2011 when the number of records of named taxa also improved with the EU recording of taxon names. Additionally, some records were recorded at an unusual depth for the taxon (8 *Acanella arbuscula*, 2 *Anthothela*, and 2 *Radicipes*). These are likely erroneous records data from previous carried out in deeper water, and therefore were discarded (Appendix Tables A4 and A5). Additionally, 42 records of two taxa (Octocorallia sp. (SUBCLASS) and Isididae) were listed in the VME Small Gorgonian Coral Functional Group sent by the data providers. However, it was unclear if these could also include Large Gorgonian Corals which also fall under this subclass and family, and therefore they were excluded from the analyses. The final biological data used for the response variables in the Small Gorgonian Coral SDMs consisted of 574 presence/biomass records from 2011 to 2023, all with taxon names, and 4708 associated null records. Records for *Acanella arbuscula* included 488 presences and 4794 absences, and 147 presences and 5135 absences for *Radicipes gracilis* (Table 4).

Erect Bryozoans

The available raw data for the SDM models for the VME functional group Erect Bryozoans included 842 data presence/biomass records (Appendix Table A1) and 6249 null records obtained from the surveys shown in Table 3. These presence/biomass records were compiled from those identified at sea as 'BRYOZOAN ECT. OR ENT.', 'BRYOZOA', and 'Bryozoa', and also included the data for the functional group but with no taxon name provided (N=276). The majority of records with taxon names (N=561) were simply recorded as Bryozoa. The coding of 'BRYOZOAN ECT. OR ENT.' is nebulous as Ectoprocta is an unaccepted phylum name which was replaced with Bryozoa, and the Entoprocta are a separate phylum. Therefore, the code suggests that those records could be something other than bryozoans. As there were only 5 records with that taxon name, all provided by Canada prior to 2016, those data were not included in the analyses. The remaining data were all recorded after 2011 but no data were reported in 2020 (Appendix Table A1). As a result, the null data from that year were also excluded from the analyses.

Therefore, the biological data considered for the response data for the Erect Bryozoan SDMs included 561 presence/biomass records from 2011 to 2023, all with taxon names, and 4357 associated null data (Table 3). However, only one species of bryozoan is considered to be a VME indicator (NAFO, 2024). In order to ensure that the models reflect the VME Indicator *Eucratea loricata*, the KDE biomass threshold of catches ≥ 0.2 kg (Kenchington et al., 2019) was used for the models for Erect Bryozoan VME Habitat; something that was not done in the previous modeling of this group (Kenchington et al., 2019). *E. loricata* is associated with the larger catches and this approach will also focus on the bryozoan habitat known in the literature as 'bryozoan turf'. The final data set used in the model for Erect Bryozoan VME Habitat included 26 presence records with biomass above the 0.2 kg threshold, and 4892 null records recorded from 2011-2023.

Sea Squirts (*Boltenia ovifera*)

The biological raw data for the SDM models for the VME functional group Sea Squirts included 380 presence records (Appendix Table A1) and 7283 null records obtained from the surveys shown in Table 3. These presence/biomass records were compiled from those identified at sea as 'BOLTENIA OVIFERA', 'BOLTENIA SP.', and 'TUNICATE, SESSILE', and also included the data for the functional group but with no taxon name or biomass provided (N=64). Records with taxon names were first recorded in 2007 (Appendix Table A2) and appear in small numbers in each subsequent year of the time series.

Four records with a taxon name occurring deeper than usual for the taxon were recorded. After reviewing at-sea photos and annotations (Appendix Table A6), these records were discarded as they likely represented erroneous data from previous sets carried out in shallower waters. The final biological data used for the response data in the Sea Squirt SDM included 312 presence/biomass records from 2007-2023, with taxon names, and 7347 associated null data (Table 4). This data set has been archived on the NAFO SharePoint site.

Table 4. Summary of the Response Data Inputs to the Random Forest Species Distribution Models.

Response Group	Period	No. Presences	No. Absences
Large Gorgonian Corals	2011 - 2023	161	5279
Small Gorgonian Corals	2011 - 2023	574	4708
Acanella arbuscula	2011 - 2023	488	4794
Radicipes gracilis	2011 - 2023	147	5135
Erect Bryozoans	2011 - 2023	561	4358
Erect Bryozoan VME Habitat	2011 - 2023	26	4892
Sea Squirts (*Boltenia*)	2007 - 2023	312	7347

Variable Reduction

Following previous work (Murillo et al., 2024), preliminary SDMs were generated for each of the modelled taxa using the full suite of predictor variables to rank variable importance (Murillo et al., 2024; Appendix Table A5). Following this, an iterative approach was used to conduct model specific variable selection. First, Spearman correlations were calculated for variable pairs, and for those with correlation scores > 0.70, the least important variable was removed. Subsequently, the variance inflation factor (VIF), which measures the amount of inflation in the variance of a regression coefficient due to multicollinearity, was evaluated for the remaining uncorrelated variables. If VIFs > 10 were observed, the Spearman correlation scores were recomputed with progressively lower thresholds (decreased by increments of 0.05) until all remaining predictor variables achieved a VIF < 10.

It should be noted that the order of variable importance used in the selection of uncorrelated variables comes from the run of a preliminary full model (see page 94 in Murillo et al. (2024)). Random association with the Random Forest algorithm may produce slight differences in the order of variable importance which could influence the final selection of uncorrelated variables, and in consequence the extrapolated areas. We also reiterate that the primary purpose of these models was to generate accurate maps of the distribution, i.e. ‘Assessment and Prediction’, and not to describe the environmental niches of the VME taxa. The variables selected by the models are highly correlated with other predictors which may have greater biological relevance than the ones selected, although all were chosen based on their potential to influence distribution.

Model Fitting

Following Murillo et al. (2024), models predicting the probability of presence for each taxon were built using classification Random Forest models. Random Forest is an ensemble method, where a large number of decision trees (typically 500-1000) are built using random subsets of the data (Breiman, 2001; Cutler et al., 2007). The models were built in the free statistical computing software R (v. 3.5.1; R Development Core Team, 2018) using the ‘randomForest’ R package (Liaw and Wiener, 2002) modified to output desired maps and tables (Murillo et al., 2024; Appendix Table A5). The models were run using the default settings of the randomForest function, using 500 trees.

Predictor importance was investigated for each model using the decrease in end node impurity, measured by the Gini index for presence/absence. Partial response plots were used to visualize the relationship between each predictor variable and the response variables in turn, while accounting for the average effect of the other predictors in the model.

Models were validated using a bootstrap k-fold cross-validation procedure. For each response variable, the data was randomly subsampled into 10 folds and training sets constructed leaving each fold out in turn, to be used as test data (resulting in a 90/10 split, keeping balance of classes equal). Models were built using each train set, and validation statistics were calculated for each corresponding test dataset. A cross-validation approach, such as this, gives an average cross-validation score, but also an estimate of variability around the mean. The variability can be used as an indicator of the stability of the model fit, and to check for the arbitrary effects caused by subsetting data to train and test a model. Accuracy measures used to validate the models included Sensitivity, Specificity, Kappa, True Skill Statistic (TSS; Allouche et al., 2006) and Balanced Accuracy, with the mean and standard deviation calculated across model runs (N=10).

Sensitivity, also referred to as the True Positive Rate, corresponds to the proportion of observed presences correctly predicted as such. Conversely, Specificity, or True Negative Rate, is the proportion of absences correctly predicted. These can be used to judge how likely a model is to detect presence and how specific the predictions are to the correct class. High sensitivity with a low specificity indicates a model that is overpredicting, whilst an underpredicting model shows high specificity and low sensitivity. The overall accuracy was additionally investigated using the Kappa Coefficient. Kappa evaluates how well the classification performed compared to randomly assigned values. The Kappa Coefficient can range from -1 to +1. A value of 0 indicates that the classification is no better than a random classification. A negative number indicates the classification is significantly worse than random. A value close to +1 indicates that the classification is significantly better than random. While Kappa is useful, it can be sensitive to imbalanced datasets such as those used here. Also computed were the TSS (Sensitivity + Specificity – 1) and Balanced Accuracy (average of Sensitivity and Specificity) which, unlike Kappa, are both independent of prevalence and can give a much better estimate of overall model performance where the classes are unbalanced. TSS ranges from -1 to +1, with +1 indicating perfect agreement and 0 or less indicating no better than random chance.

Binary presence/absence maps were created by using two thresholds, the prevalence of the data and a threshold optimised to ensure that resulting Sensitivity and Specificity are afforded equal weight (Sensitivity=Specificity). The former was used in previous work (Kenchington et al., 2019), as a threshold to account for the class imbalance in data. However, in this model iteration the Sensitivity=Specificity threshold was used, as it affords equal weight to detection of presence and absence, minimising both false positives and false negatives.

The final model output was plotted as the class (presence/absence) with the majority vote of all 10 model runs, i.e., each cell was assigned to the class with the greatest number of outcomes from each of the 10 model runs. Two confidence map layers were also produced consisting of: 1) the frequency of the most common class (N/10), and 2) the average probability over all 10 model runs of the majority vote class referred to as the maximum frequency class in the figures and text hereafter. For example, a value of 0.6 associated with presence areas in the frequency of the most common class map, means that a presence was predicted for 6 runs, but an absence for the other 4 runs. And the average probability of presence from the 10 model runs would be shown in the average probability of the maximum frequency class map.

Areas of extrapolation (univariate and combinatorial) were characterized using the Extrapolation Detection (ExDet) tool, based on Euclidean and Mahalanobis distances, as implemented in the 'dsmextra' R package (Bouchet et al., 2020). Cells with non-analogous environments are novel because they are outside the range of individual covariates (univariate) or they are within the univariate range but constitute novel combinations between covariates not found in the reference data set (Mesgaran et al., 2014). As the extrapolated areas are subject to the selection of uncorrelated variables used in the model, these areas can vary across taxa.

These presentations of uncertainty associated with the distributions is an advancement over the previous work from 2019 (Kenchington et al., 2019) and earlier. Uncertainty is shown in three distinct ways: 1) inclusion of areas of model extrapolation (predictions occurring outside of the range of environmental conditions encountered by response variables) on all maps; 2) maps showing the frequency of presences/absences from the 10 cross-validation runs (values close to 1 give confidence in the presences/absences identified in the binary maps); and 3) maps of the average probability of the maximum frequency class (e.g., presence or absence) from the 10 cross-validation runs (areas with lower average probability within the same class can be associated with areas of uncertainty). For each model we provided maps of the predicted presences/absences based on a threshold of Sensitivity=Specificity, showing the areas of extrapolation and uncertainty from the 10 cross-validation runs. For all of the models there were areas of uncertainty at the border of the presence/absence prediction as was seen in the models of the Large-Size Sponges, Sea Pens and Black Corals (Murillo et al., 2024).

Results

Assessment and Prediction of the Large Gorgonian Coral Functional Group

Random Forest models predicting the probability of the presence of the Large Gorgonian Coral VME functional group generally performed with moderate accuracy across the validation statistics (Balanced Accuracy, Sensitivity and Specificity all > 0.7; Table 5). Kappa, which measures the extent to which the agreement between observed and predicted is higher than that expected by chance alone, was 0.13 which is considered 'fair' performance. The TSS was 0.45 which indicates good model performance.

Table 5. Model Validation Results for the Presence/Absence Random Forest Model for the Large Gorgonian Coral VME Functional Group. TSS=True Skill Statistic (Sensitivity + Specificity – 1).

Accuracy Measure	Mean ± SD
Sensitivity	0.74 ± 0.09
Specificity	0.71 ± 0.06
Kappa	0.13 ± 0.06
Balanced Accuracy	0.73 ± 0.07
TSS	0.45 ± 0.15

The most important variables were the averaged spring primary productivity, followed by the mean of the range in bottom salinity, the wind exposition index, the broad-scale bathymetric position index, and ruggedness (Figure 1). The models indicate that the Large Gorgonian Corals are found in areas with a mean spring primary productivity > 1200 mg C m^{-2} day^{-1}, a range of bottom salinity > 0.1, sheltered locations on broadscale depressions to 600 m, with rugged terrain (Figure 2).

The predicted distribution maps are shown in Figure 3, shown as binary plots of presence/absence based on the two thresholds (Prevalence and Sensitivity=Specificity). These two plots are very similar. The distribution of the data is shown overlain on the binary map of presence/absence based on Sensitivity=Specificity (Figure 4). Outside areas of model extrapolation, the Large Gorgonian Corals are distributed around the Flemish Cap except for the northern portion of Flemish Pass and on the slopes of the Tail of Grand Bank, south of Flemish Pass.

The uncertainty expressed as the frequency of presence/absence from the 10 cross-validation runs (Figure 4), the areas of extrapolation (Figures 4 and 5), and the average probability of the maximum frequency class (Figure 5) indicated high certainty within the fishing footprint for both presence and absence predictions outside of the deep slope areas. However there was increased uncertainty in the deeper slope waters (Figure 5) and in areas of transition between the presence and absence classes (Figures 4 and 5).

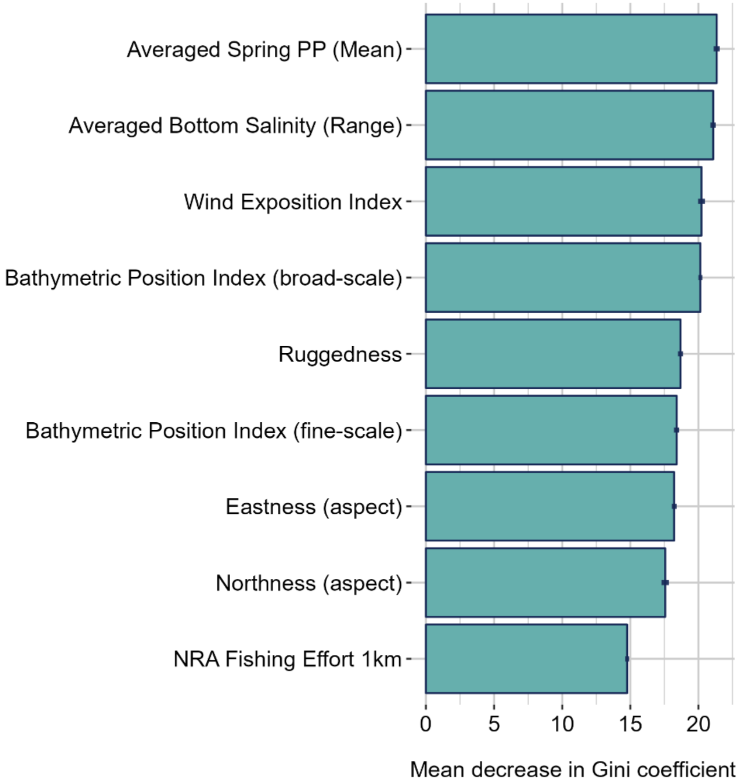

Figure 1. Plot of mean decrease and standard deviation in Gini Value for the 9 predictor variables in the Random Forest model for the Large Gorgonian Coral Functional Group, indicating their relative importance and variation across 10 model folds.

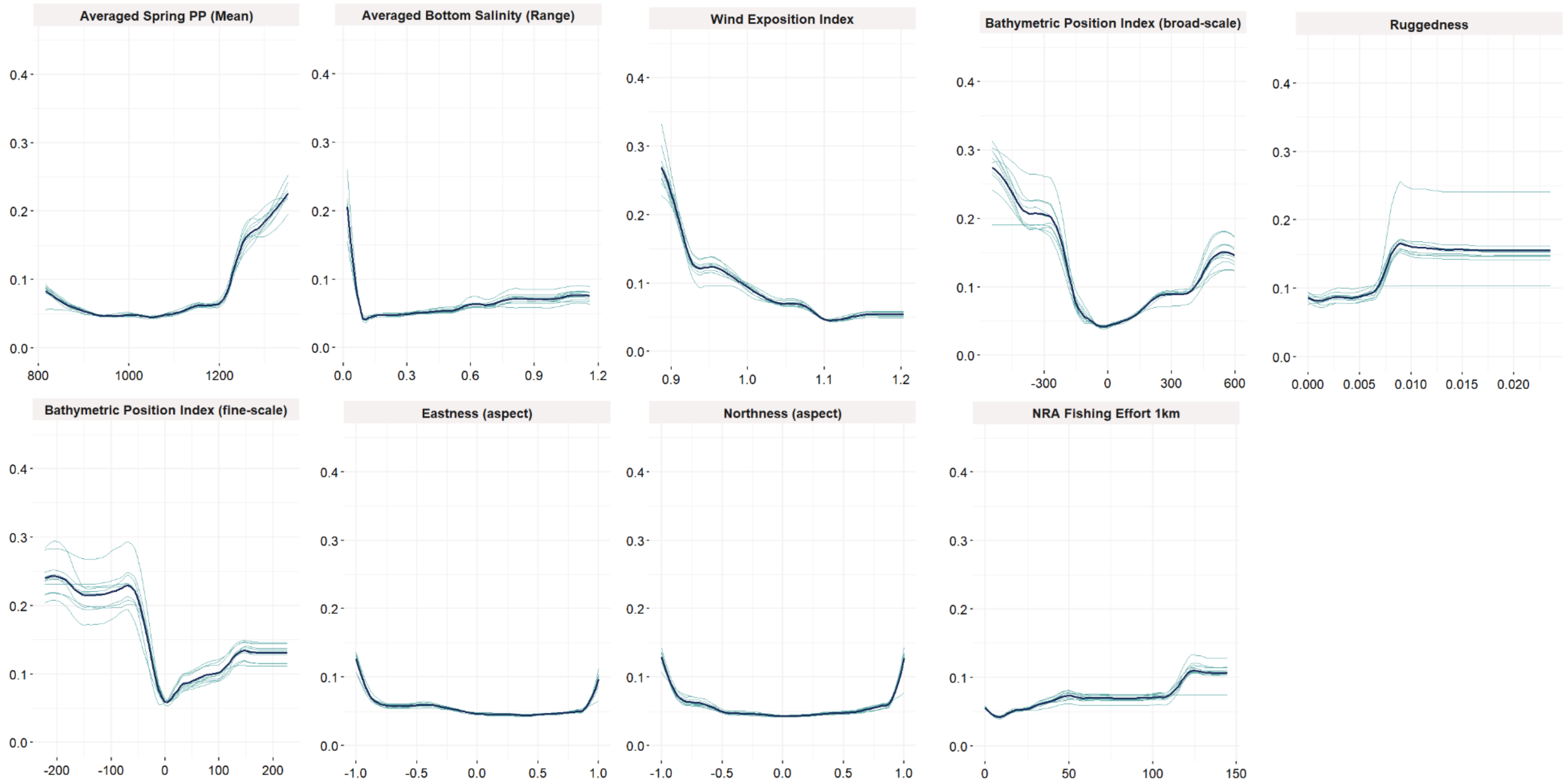

Figure 2. Response curves showing the partial dependence of the probability of presence on the predictors (Figure 1) identified in the Random Forest model for the Large Gorgonian Coral Functional Group. For each variable, the mean response and curves for each of the model folds are plotted. The plots show the predicted response to each predictor variable in turn, whilst other variables are held at their mean value.

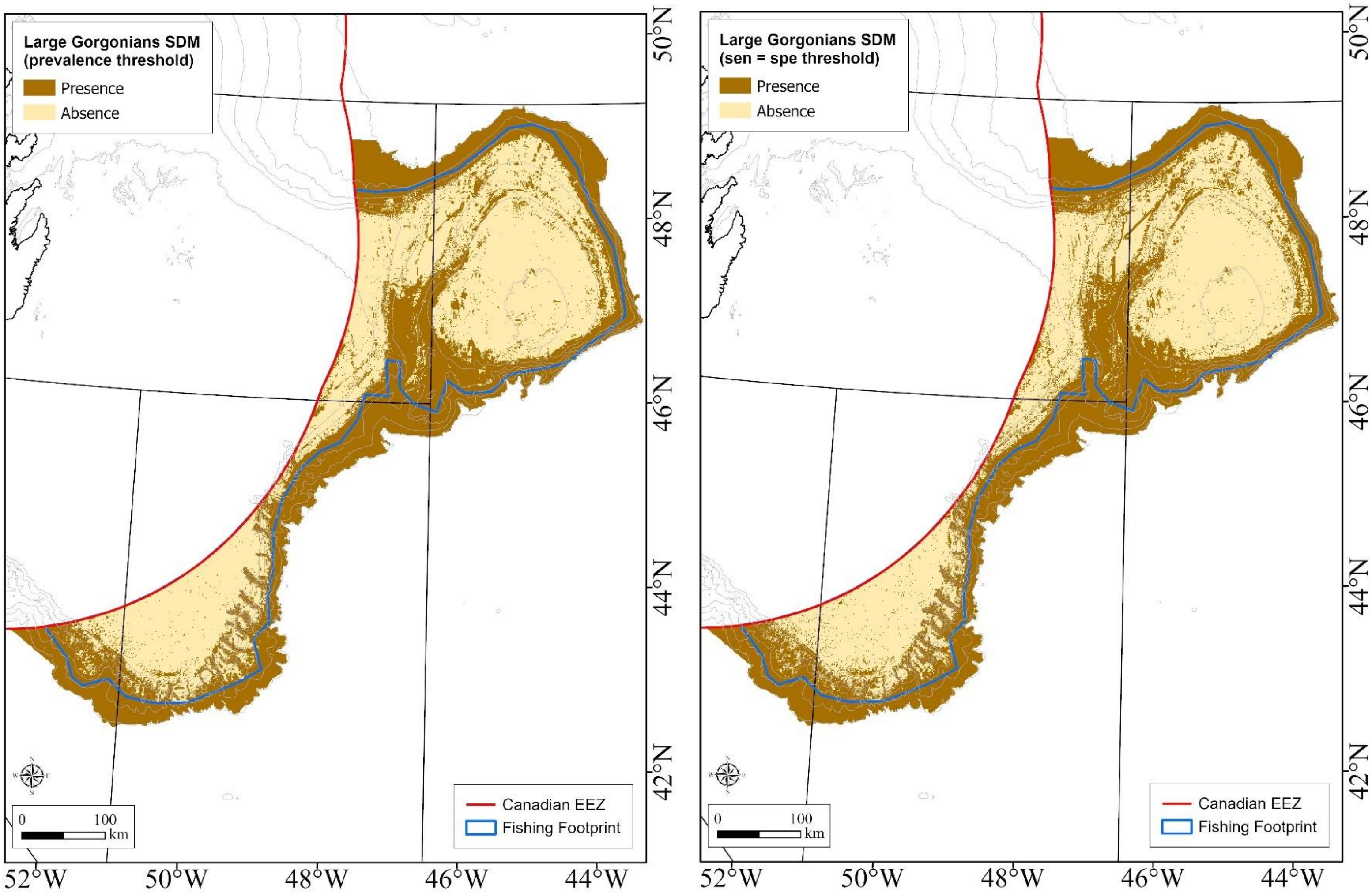

Figure 3. Random Forest species distribution model for the VME Large Gorgonian Coral Functional Group showing binary maps of VME presence thresholded using data prevalence (left panel) and a Sensitivity=Specificity threshold (right panel). The perimeter of the fishing footprint is shown on both maps in blue.

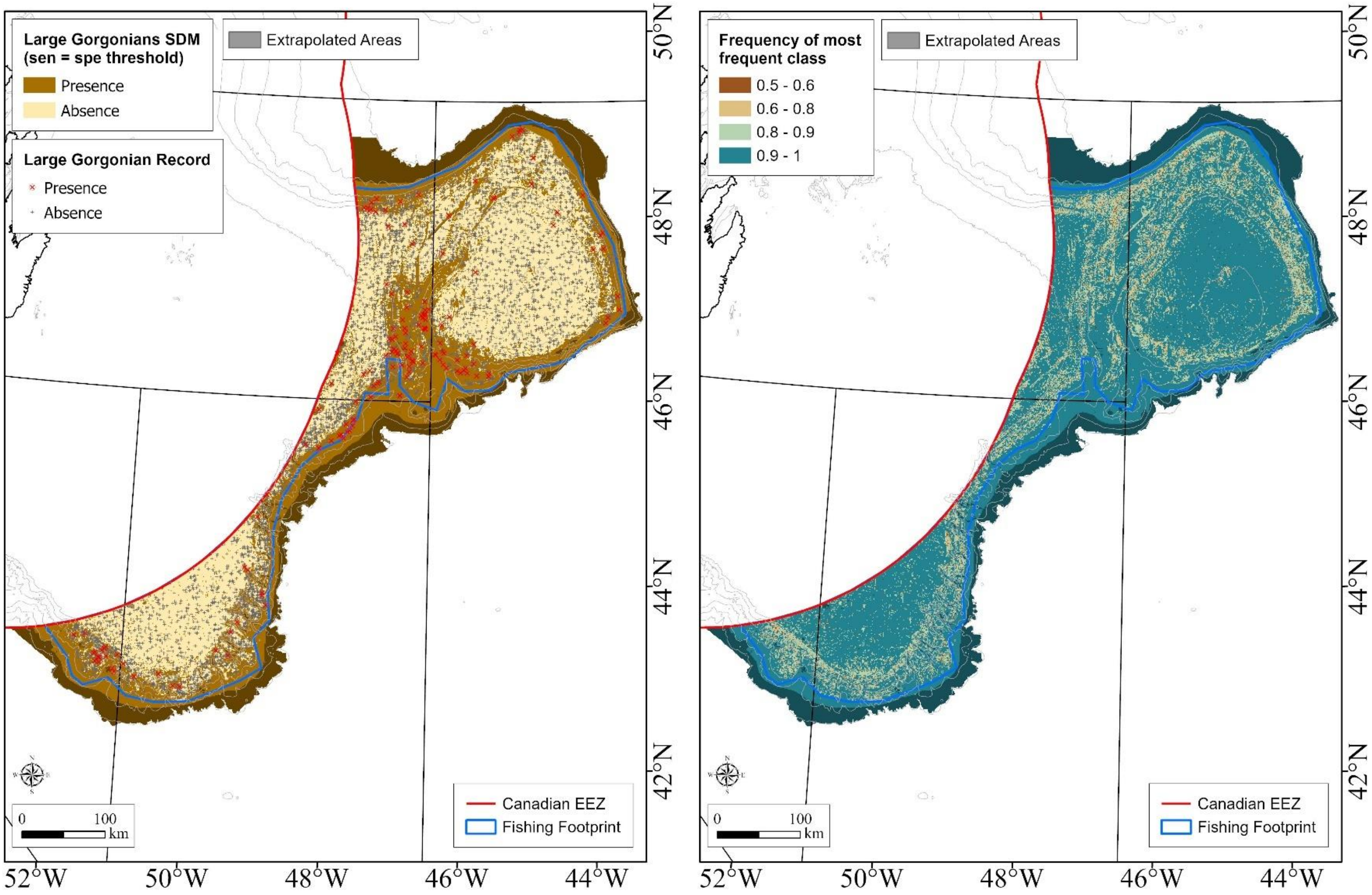

Figure 4. Random Forest species distribution model for the VME Large Gorgonian Coral Functional Group showing the distribution of the presence and absence data overlain on a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated by showing the frequency of P/A from the 10 cross-validation runs (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

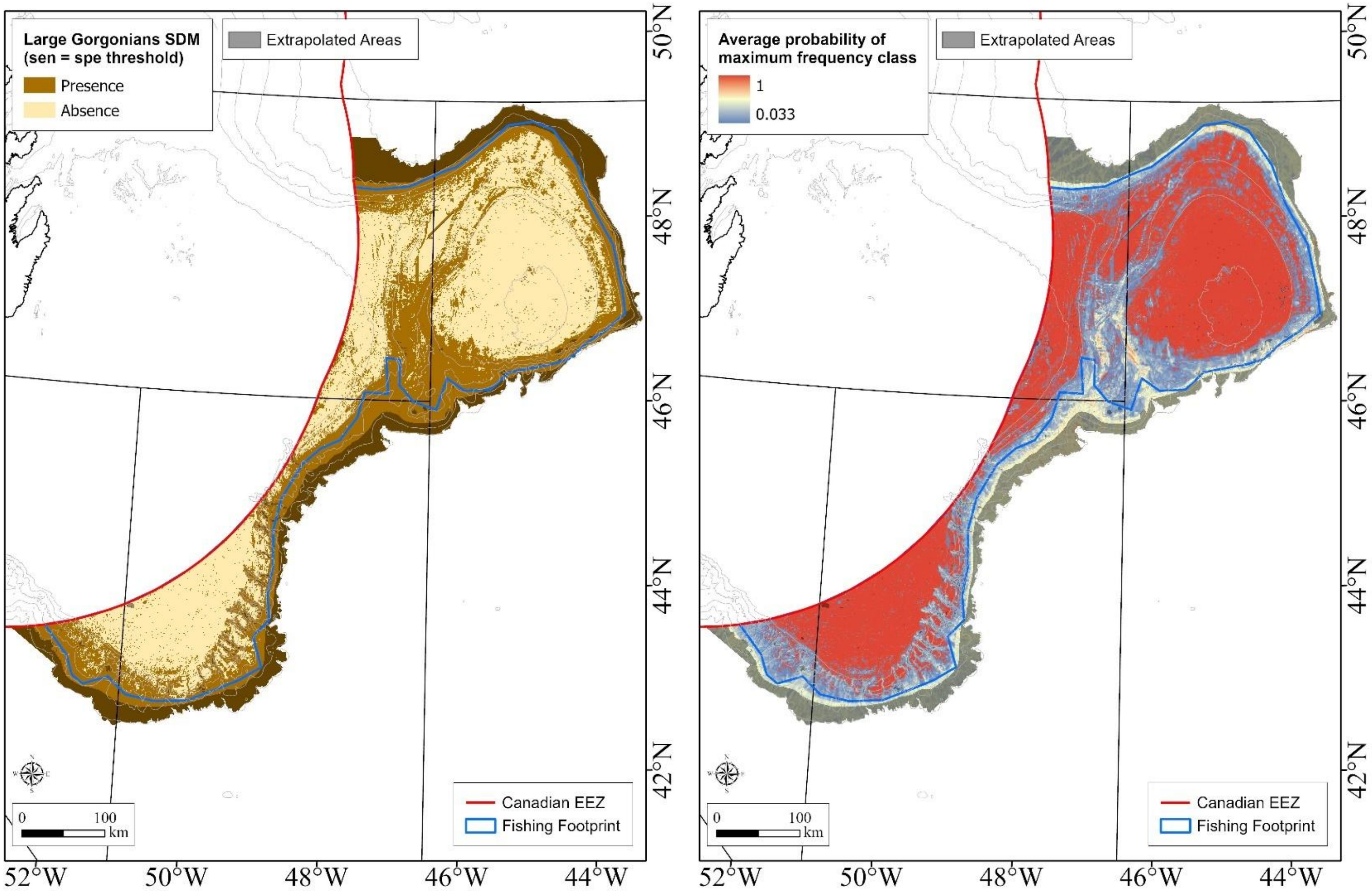

Figure 5. Random Forest species distribution model for the Large Gorgonian Coral Functional Group showing a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated as the average probability of the maximum frequency class (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

Assessment and Prediction of the Small Gorgonian Coral Functional Group

Three models for the Small Gorgonian Coral Functional Group were performed, each using both Prevalence and Sensitivity = Specificity thresholds for the data, although model performance statistics and predictive maps are only shown for the latter (Table 6). Balanced accuracy, sensitivity and specificity were all high for all three models (Small Gorgonian Coral Functional Group, *Acanella arbuscula* and *Radicipes gracilis*). Kappa was 0.18 for *R. gracilis*, which is considered a slight agreement between observed and predicted, 0.3 for *A. arbuscula* and 0.37 for the functional group which are considered 'fair' performance. The TSS was positive for all models and near or equal to 0.6 which indicates that the model is very accurate, with good predictive power for both presence and absence.

Table 6. Model Validation Results for the Presence/Absence Random Forest Model for the Small Gorgonian Coral Functional Group, and Subgroups. TSS=True Skill Statistic (Sensitivity + Specificity – 1).

	Small Gorgonian Coral Functional Group	***Acanella arbuscula***	***Radicipes gracilis***
Accuracy Measure	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**
Sensitivity	0.80 ± 0.03	0.78 ± 0.04	0.82 ± 0.06
Specificity	0.80 ± 0.02	0.78 ± 0.03	0.75 ± 0.04
Kappa	0.37 ± 0.04	0.30 ± 0.05	0.18 ± 0.04
Balanced Accuracy	0.80 ± 0.02	0.78 ± 0.03	0.79 ± 0.04
TSS	0.60 ± 0.05	0.57 ± 0.07	0.58 ± 0.07

Small Gorgonian Coral Functional Group

The most important variables for predicting the distribution of the Small Gorgonian Coral Functional Group were the fill-sink bathymetry, followed by the mean of the range in surface salinity, and the mean minimum summer primary productivity (Figure 6). The models indicate that the Small Gorgonian Corals are found in sink areas of -500 to -1500 m, with a mean of the range in surface salinity > 1.3 and the mean minimum summer primary productivity < 280 mg C m^{-2} day^{-1} (Figure 7).

The predicted distribution maps are shown in Figure 8, shown as binary plots of presence/absence based on the two thresholds (Prevalence and Sensitivity=Specificity). These two plots are very similar. The distribution of the data is shown overlain on the binary map of presence/absence based on Sensitivity=Specificity (Figure 9). Outside areas of model extrapolation, the Small Gorgonian Corals are distributed around the Flemish Cap and on the slopes of the Tail of Grand Bank, Flemish Pass, and Flemish Cap. They are absent from the shallow waters of both the Grand Bank and the Flemish Cap.

The uncertainty expressed as the frequency of P/A from the 10 cross-validation runs (Figure 9), the areas of extrapolation (Figures 9 and 10), and the average probability of the maximum frequency class (Figure 10) indicated high certainty within the fishing footprint for both presence and absence predictions outside of the deep slope areas. However there was increased uncertainty in the deeper slope waters (Figure 10) and in areas of transition between the presence and absence classes (Figures 9 and 10). Areas of predicted presence appear to have greater uncertainty than areas of predicted absence.

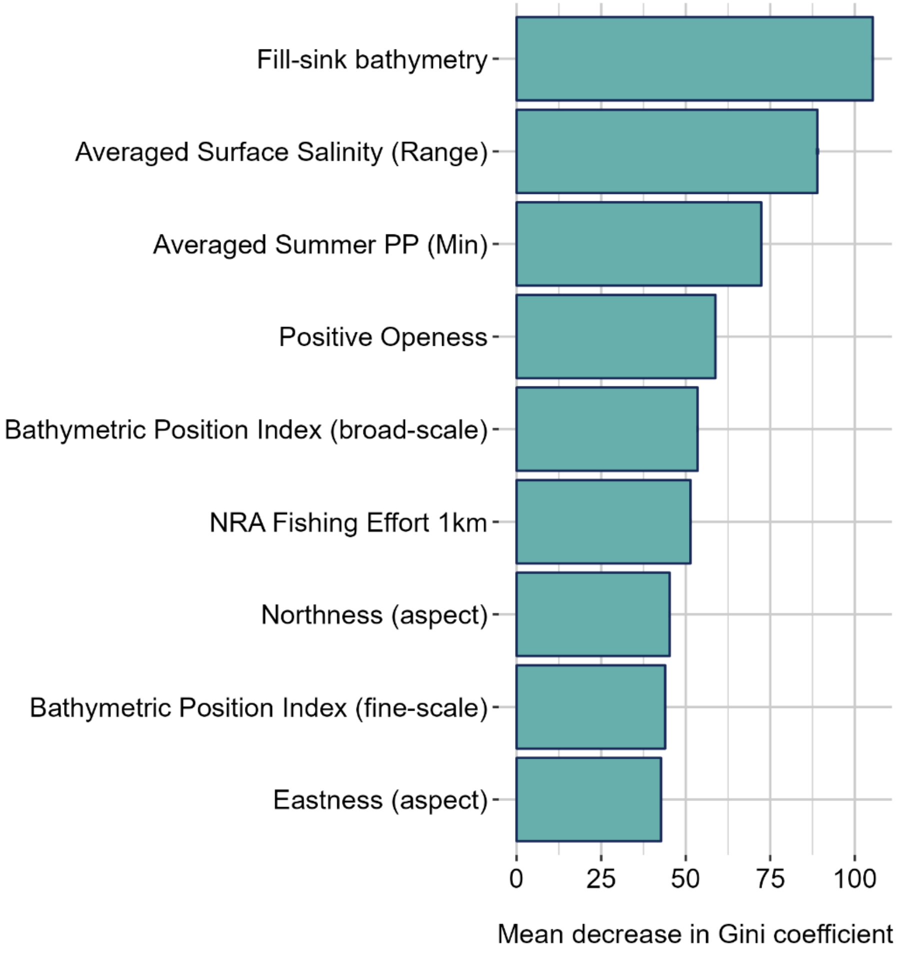

Figure 6. Plot of mean decrease and standard deviation in Gini Value for the 9 predictor variables in the Random Forest model for the Small Gorgonian Coral Functional Group, indicating their relative importance and variation across 10 model folds.

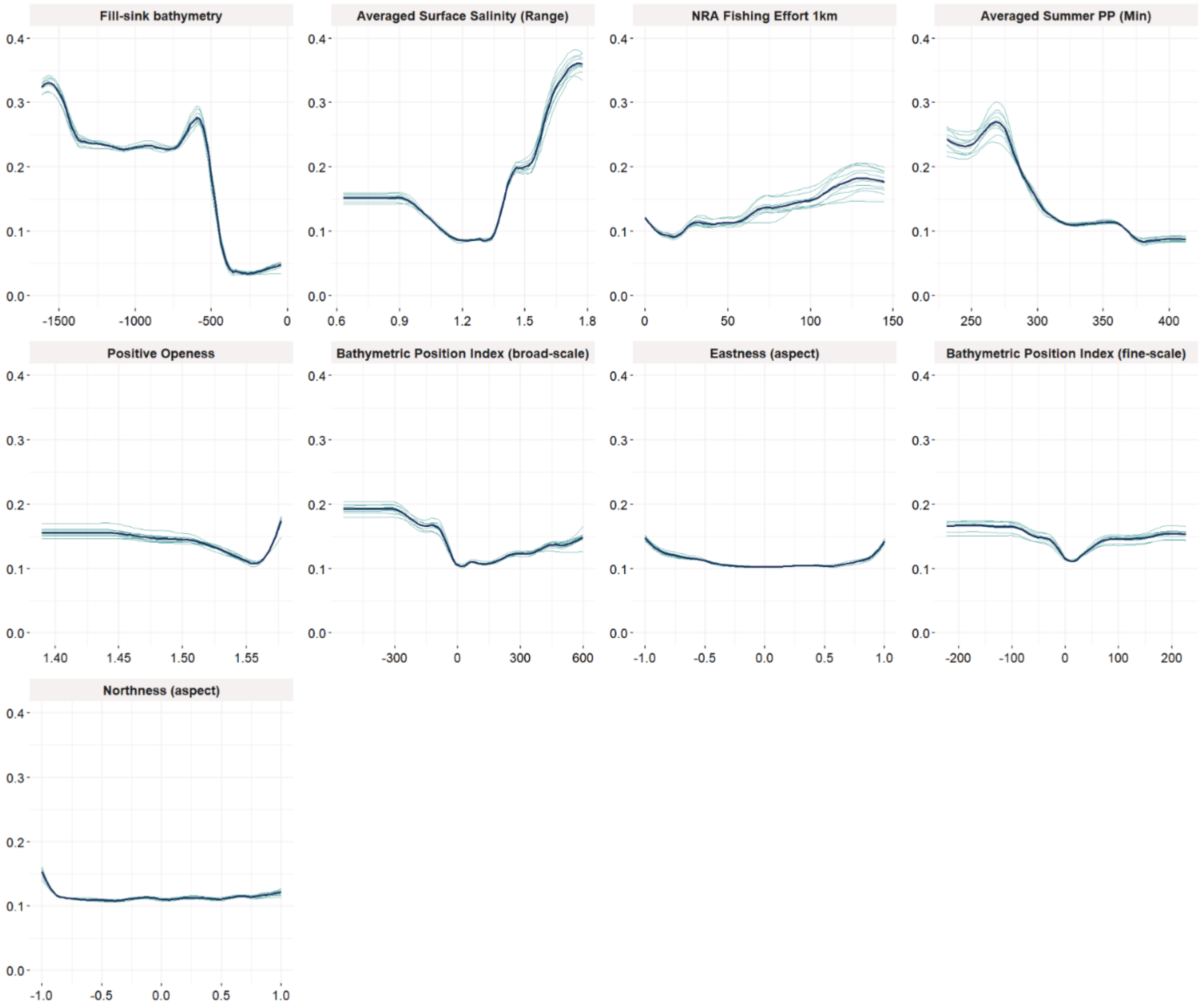

Figure 7. Response curves showing the partial dependence of the probability of presence on the predictors (Figure 6) identified in the Random Forest model for the Small Gorgonian Coral Functional Group. For each variable, the mean response and curves for each of the model folds are plotted. The plots show the predicted response to each predictor variable in turn, whilst other variables are held at their mean value.

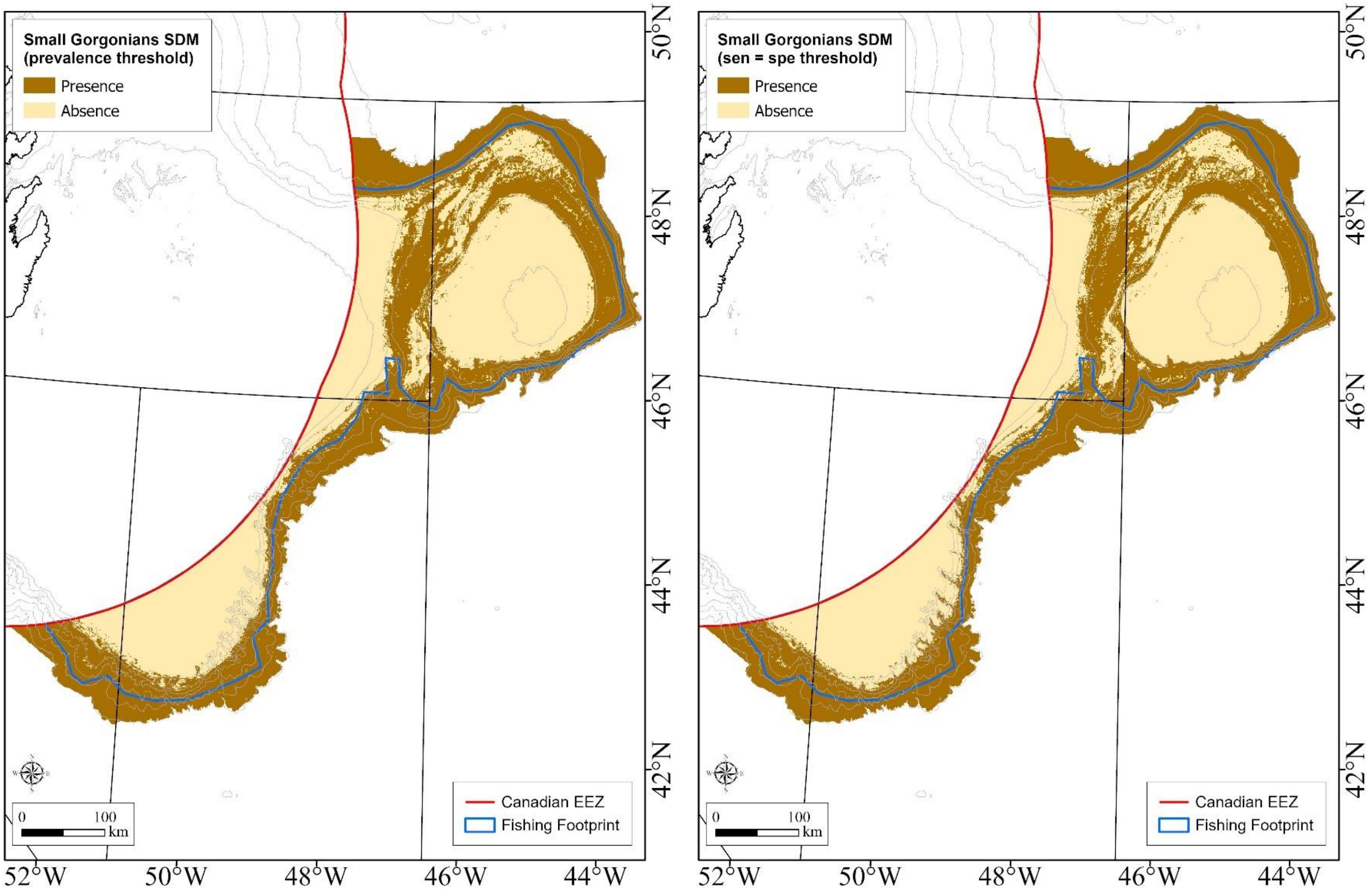

Figure 8. Random Forest species distribution model for the VME Small Gorgonian Coral Functional Group showing binary maps of VME presence thresholded using data prevalence (left panel) and a Sensitivity=Specificity threshold (right panel). The perimeter of the fishing footprint is shown on both maps.

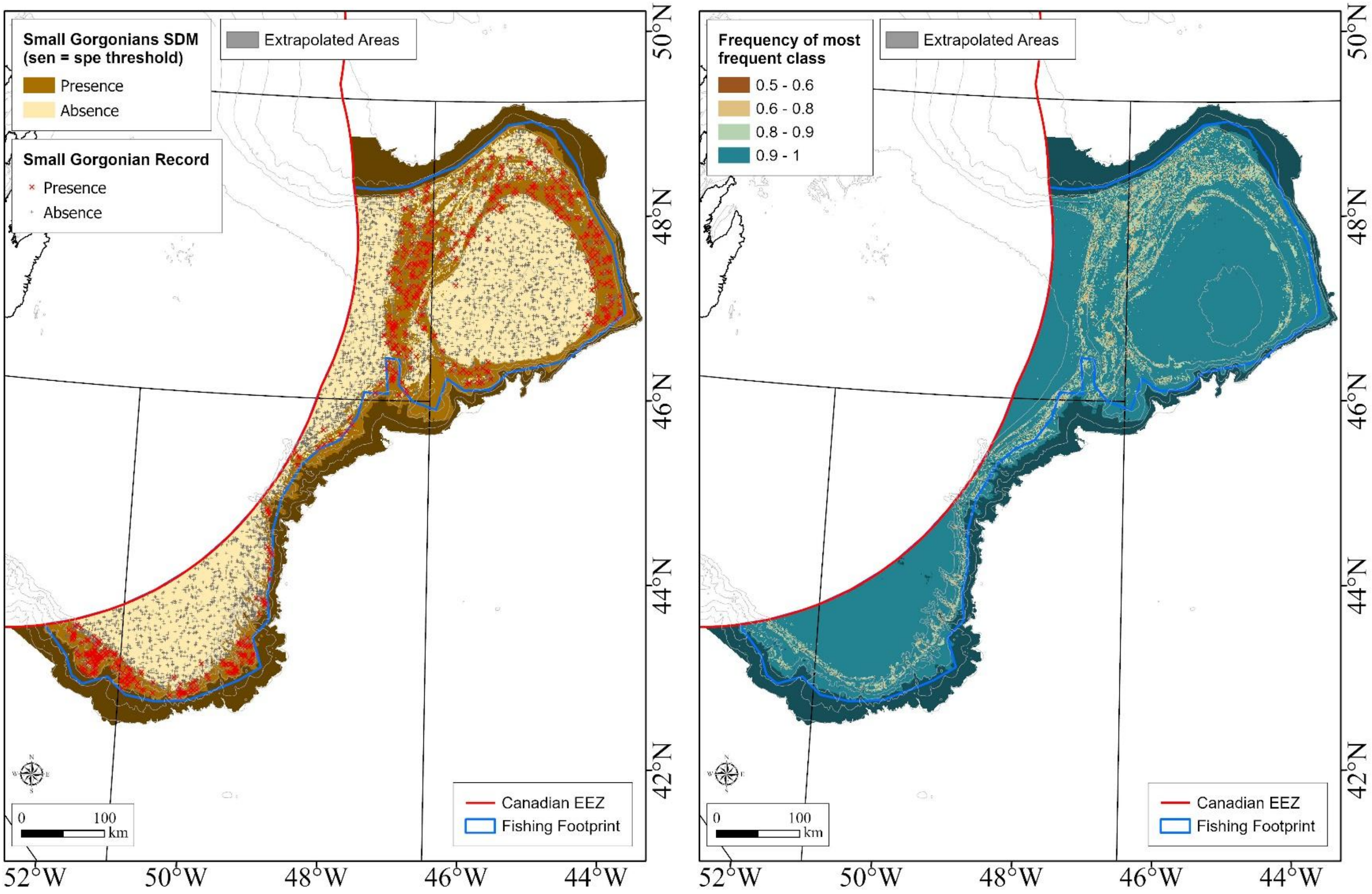

Figure 9. Random Forest species distribution model for the VME Small Gorgonian Coral Functional Group showing the distribution of the presence and absence data overlain on a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated by showing the frequency of P/A from the 10 cross-validation runs (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

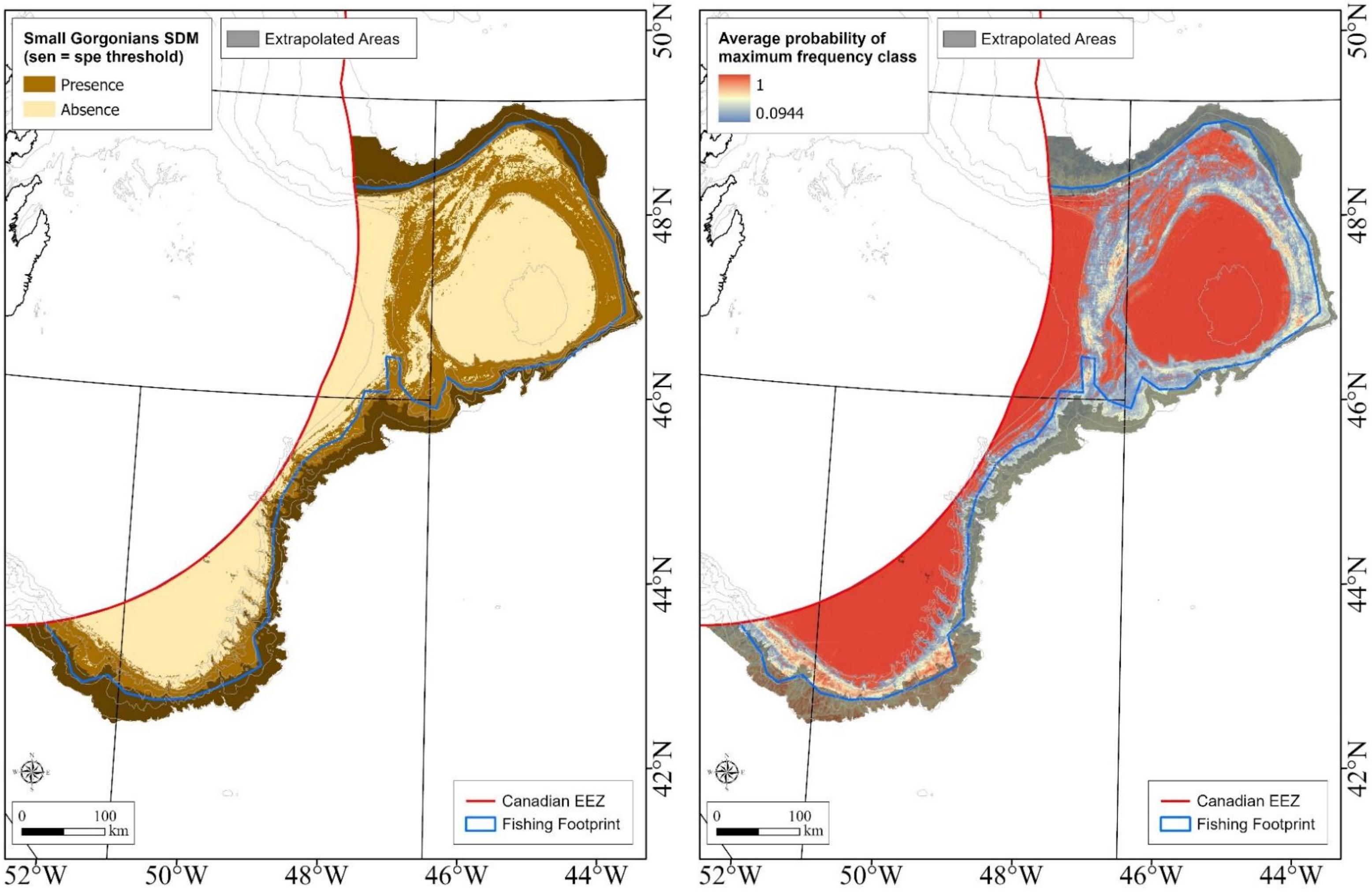

Figure 10. Random Forest species distribution model for the Small Gorgonian Coral Functional Group showing a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated as the average probability of the maximum frequency class (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

Acanella arbuscula

The most important variables for predicting the distribution of *Acanella arbuscula* were the mean of the maximum bottom salinity, mean of the range in surface salinity, the mean minimum summer primary productivity, and the mean of the averaged surface temperature (Figure 11). The models indicate that *Acanella arbuscula* are found in areas with a mean maximum bottom salinity > 34.8, a mean of the range in surface salinity of > 1.5, and the mean minimum summer primary productivity < 280 mg C m^{-2} day^{-1} (Figure 12).

The predicted distribution maps are shown in Figure 13, shown as binary plots of presence/absence based on the two thresholds (Prevalence and Sensitivity=Specificity). These two plots are very similar. The distribution of the data is shown overlain on the binary map of presence/absence based on Sensitivity=Specificity (Figure 14). Outside areas of model extrapolation which are minimal, *Acanella arbuscula* is distributed around the Flemish Cap and on the slopes of the Tail of Grand Bank, Flemish Pass, and Flemish Cap. They are absent from the shallow waters of both the Grand Bank and the Flemish Cap with low uncertainty (Figures 14 and 15). *Acanella arbuscula* presence is found on the slopes of the Tail of Grand Bank and in Flemish Pass with high certainty (Figures 14 and 15).

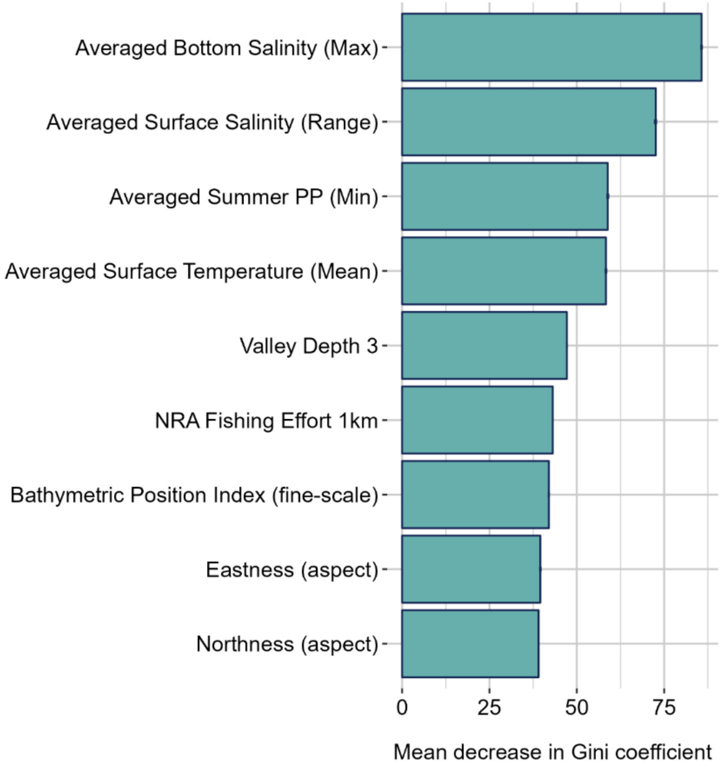

Figure 11. Plot of mean decrease and standard deviation in Gini Value for the 9 predictor variables in the Random Forest model for *Acanella arbuscula*, indicating their relative importance and variation across 10 model folds.

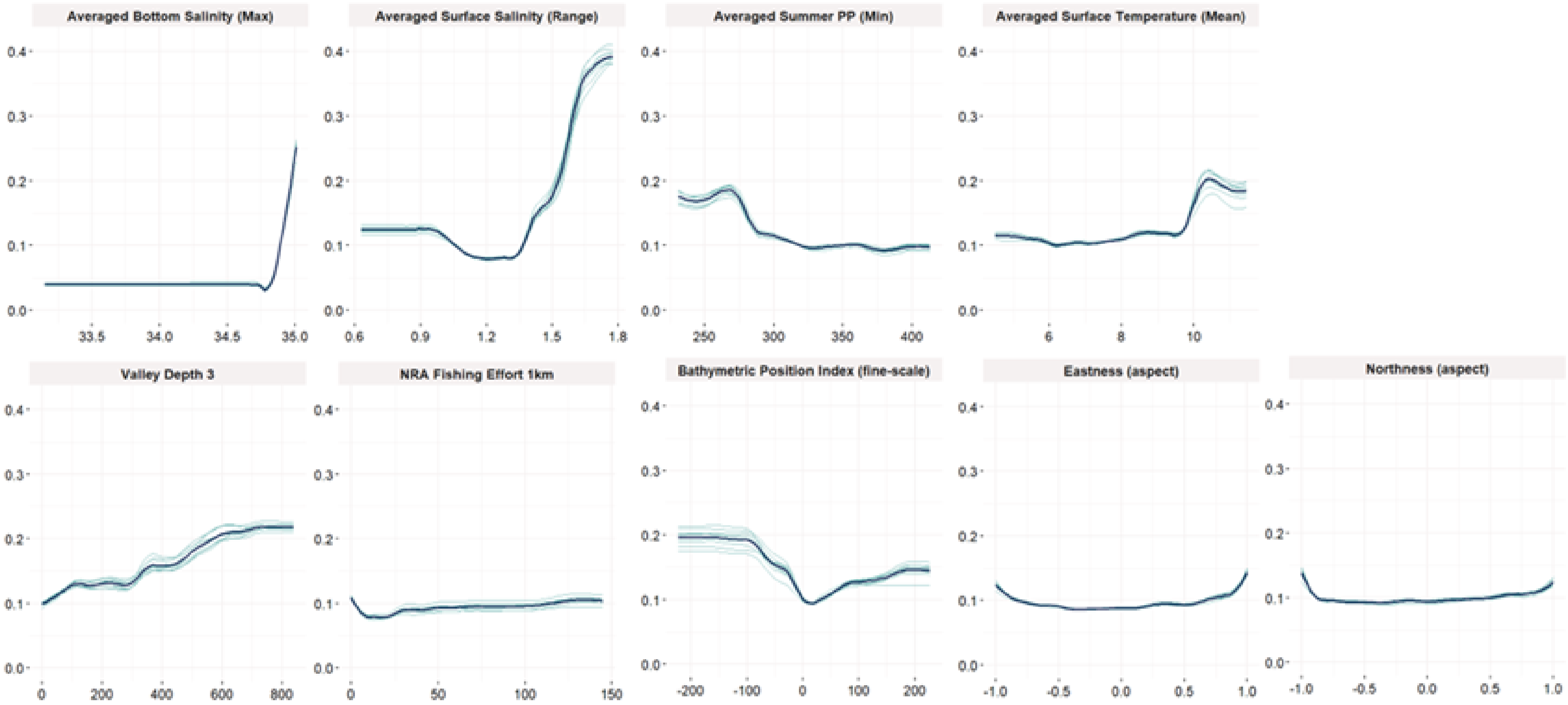

Figure 12. Response curves showing the partial dependence of the probability of presence on the predictors (Figure 11) identified in the Random Forest model for *Acanella arbuscula*. For each variable, the mean response and curves for each of the model folds are plotted. The plots show the predicted response to each predictor variable in turn, whilst other variables are held at their mean value.

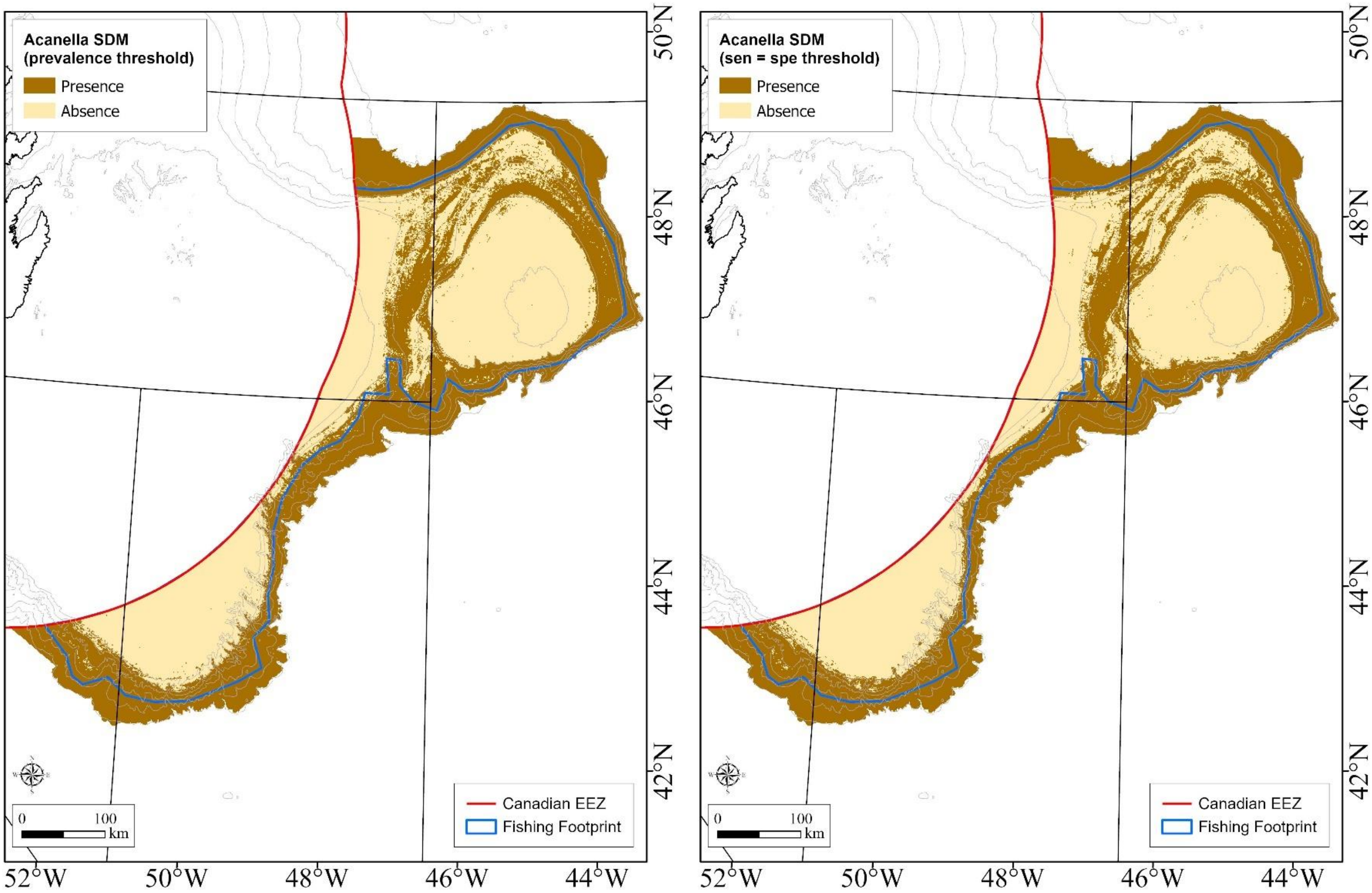

Figure 13. Random Forest species distribution model for *Acanella arbuscula* showing binary maps of VME presence thresholded using data prevalence (left panel) and a Sensitivity=Specificity threshold (right panel). The perimeter of the fishing footprint is shown on both maps.

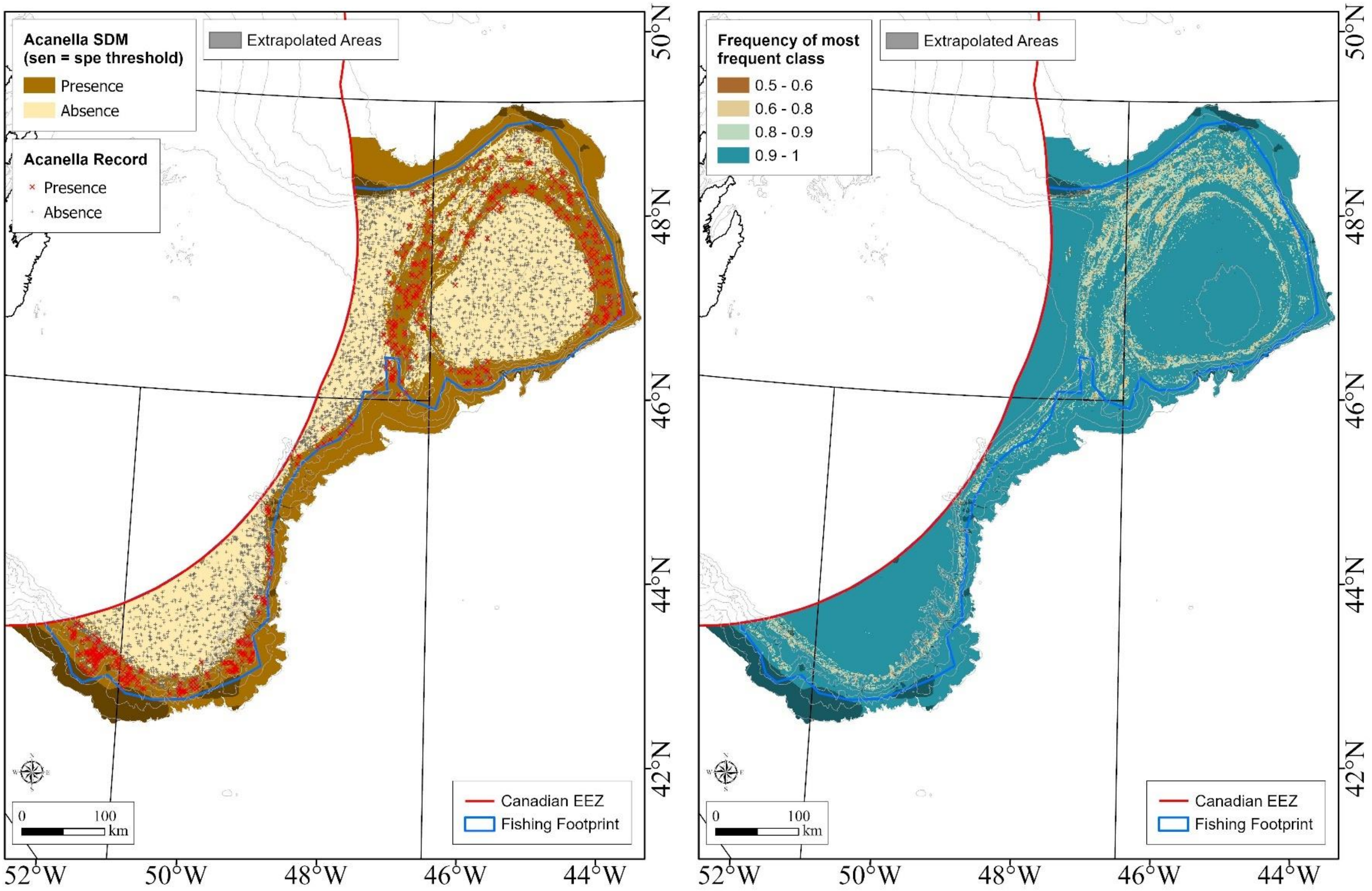

Figure 14. Random Forest species distribution model for *Acanella arbuscula* showing the distribution of the presence and absence data overlain on a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated by showing the frequency of P/A from the 10 cross-validation runs (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

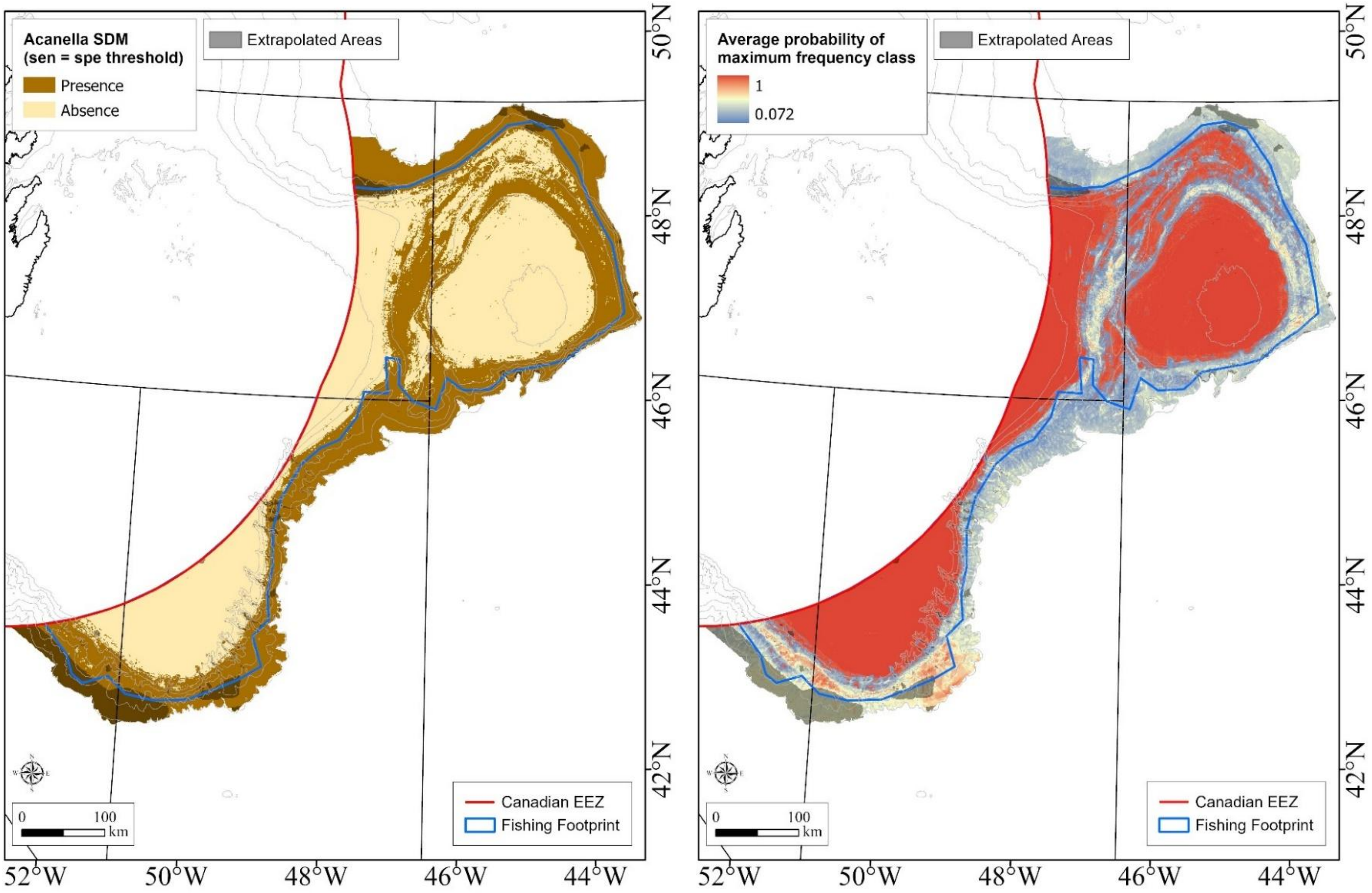

Figure 15. Random Forest species distribution model for *Acanella arbuscula* showing a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated as the average probability of the maximum frequency class (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

Radicipes gracilis

The most important variables for predicting the distribution of *Radicipes gracilis* were the channel network base (level 3), mean of the range in surface temperature, mean of the range in surface salinity, the mean minimum summer primary productivity, and positive openness (Figure 16). The models indicate that *Radicipes gracilis* are found areas with a lowest elevation point of -500 to -1500 m in the channel network, a mean of the range in surface temperature of < 10°C or > 15°C, mean of the range in surface salinity > 1.5, mean minimum summer primary productivity < 280 mg C m^{-2} day^{-1}, and an open landscape (Figure 17).

The predicted distribution maps are shown in Figure 18, shown as binary plots of presence/absence based on the two thresholds (Prevalence and Sensitivity=Specificity). These two plots are very similar. The distribution of the data is shown overlain on the binary map of presence/absence based on Sensitivity=Specificity (Figure 19). Outside areas of model extrapolation, *Radicipes gracilis* is distributed around the Flemish Cap and on the slopes of the Tail of Grand Bank, Flemish Pass, and Flemish Cap. They are absent from the shallow waters of both the Grand Bank and the Flemish Cap with low uncertainty (Figures 19 and 20). *Radicipes gracilis* presence is found on the slopes of the Tail of Grand Bank and in Flemish Pass where it associated with high certainty and with presence records. Lower certainty is found on the edges of the presence areas (Figures 19 and 20).

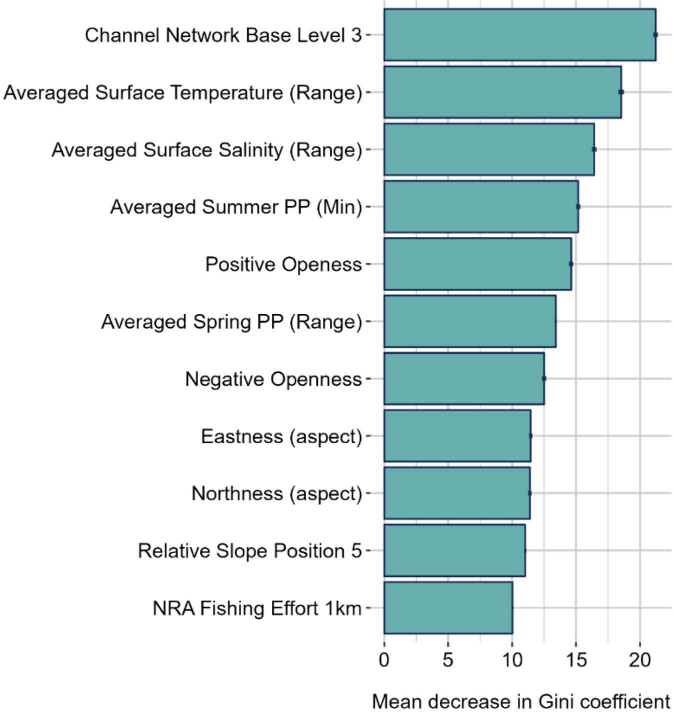

Figure 16. Plot of mean decrease and standard deviation in Gini Value for the 11 predictor variables in the Random Forest model for *Radicipes gracilis*, indicating their relative importance and variation across 10 model folds.

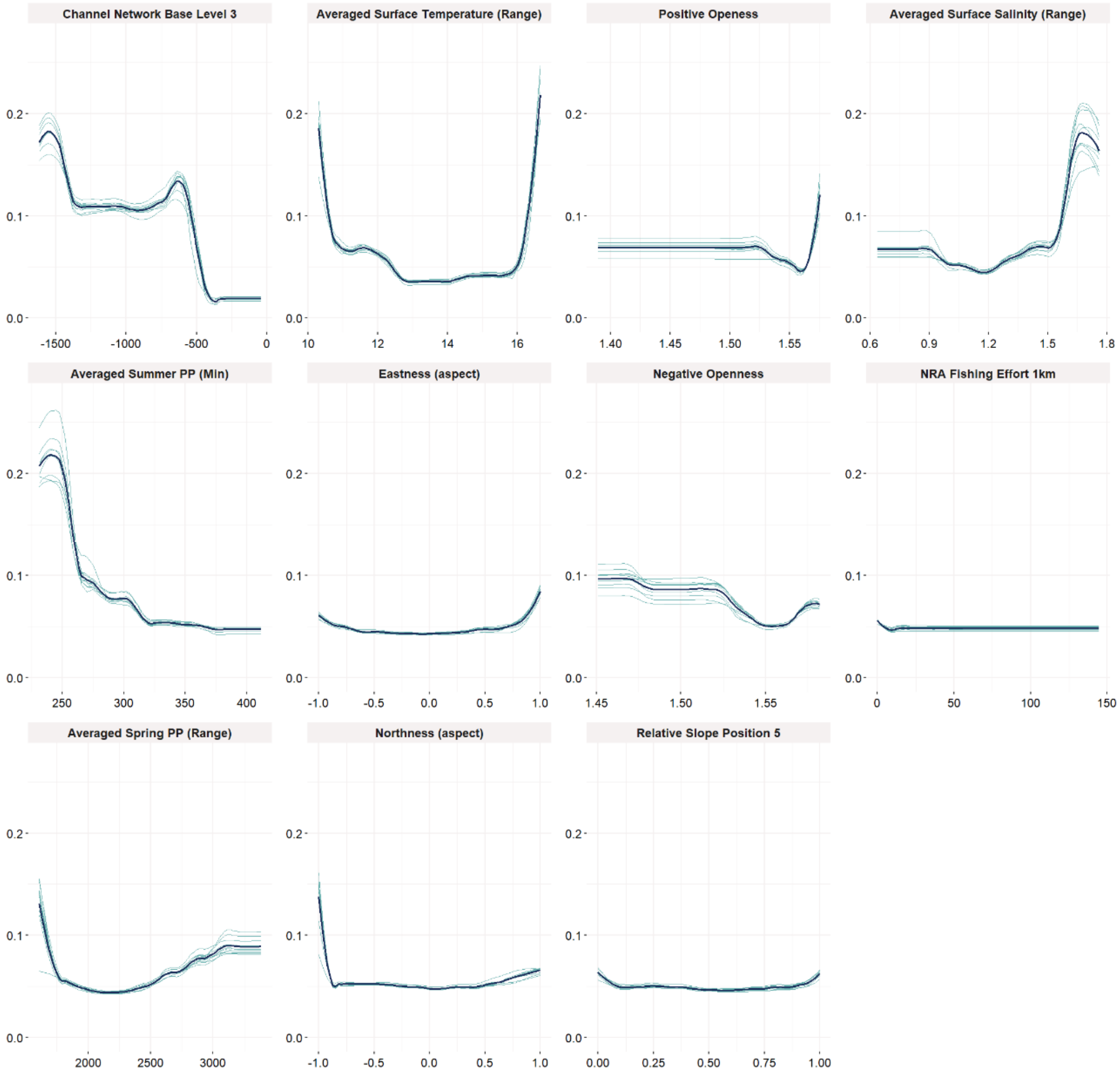

Figure 17. Response curves showing the partial dependence of the probability of presence on the predictors (Figure 16) identified in the Random Forest model for *Radicipes gracilis*. For each variable, the mean response and curves for each of the model folds are plotted. The plots show the predicted response to each predictor variable in turn, whilst other variables are held at their mean value.

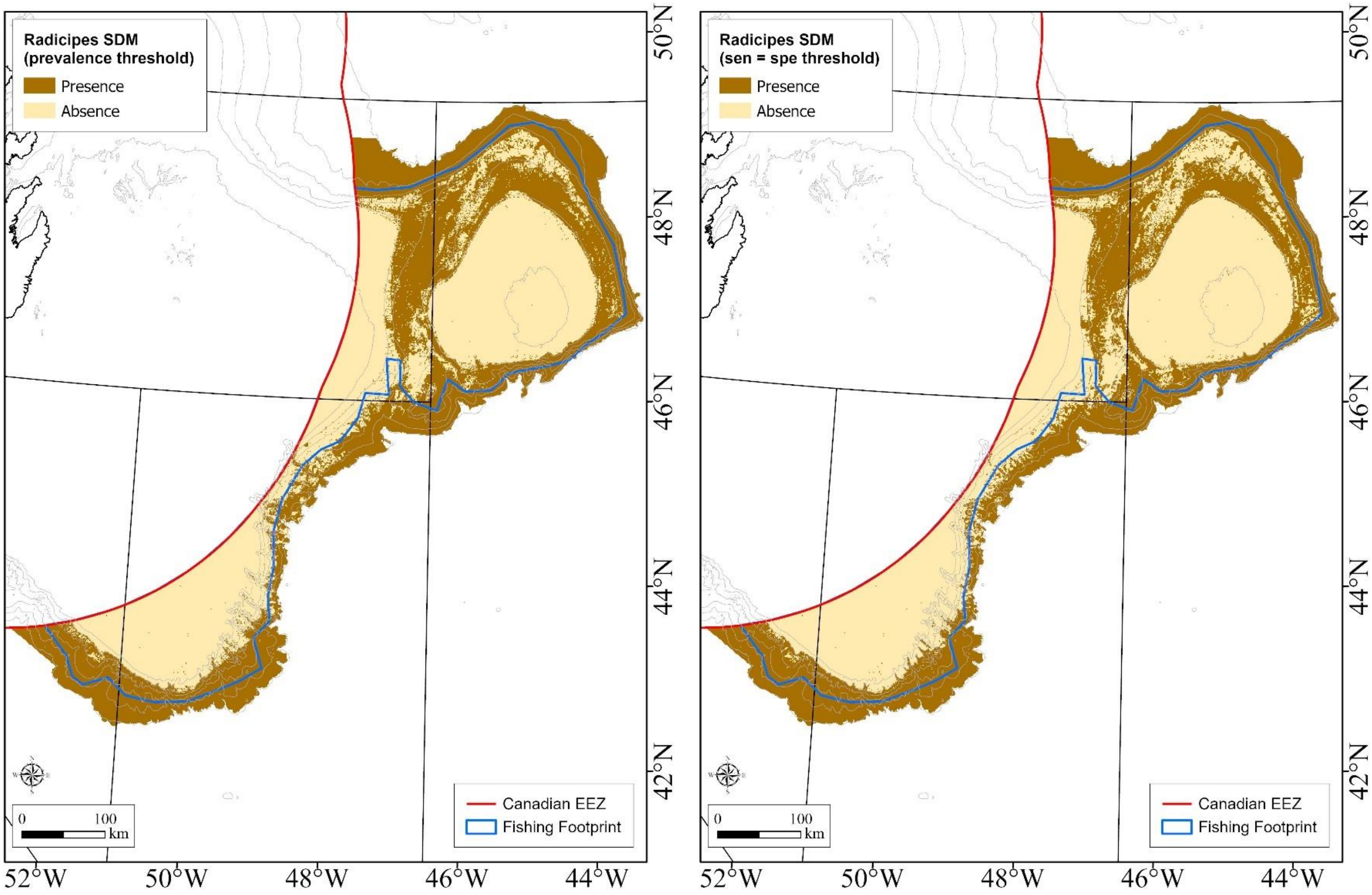

Figure 18. Random Forest species distribution model for *Radicipes gracilis* showing binary maps of VME presence thresholded using data prevalence (left panel) and a Sensitivity=Specificity threshold (right panel). The perimeter of the fishing footprint is shown on both maps.

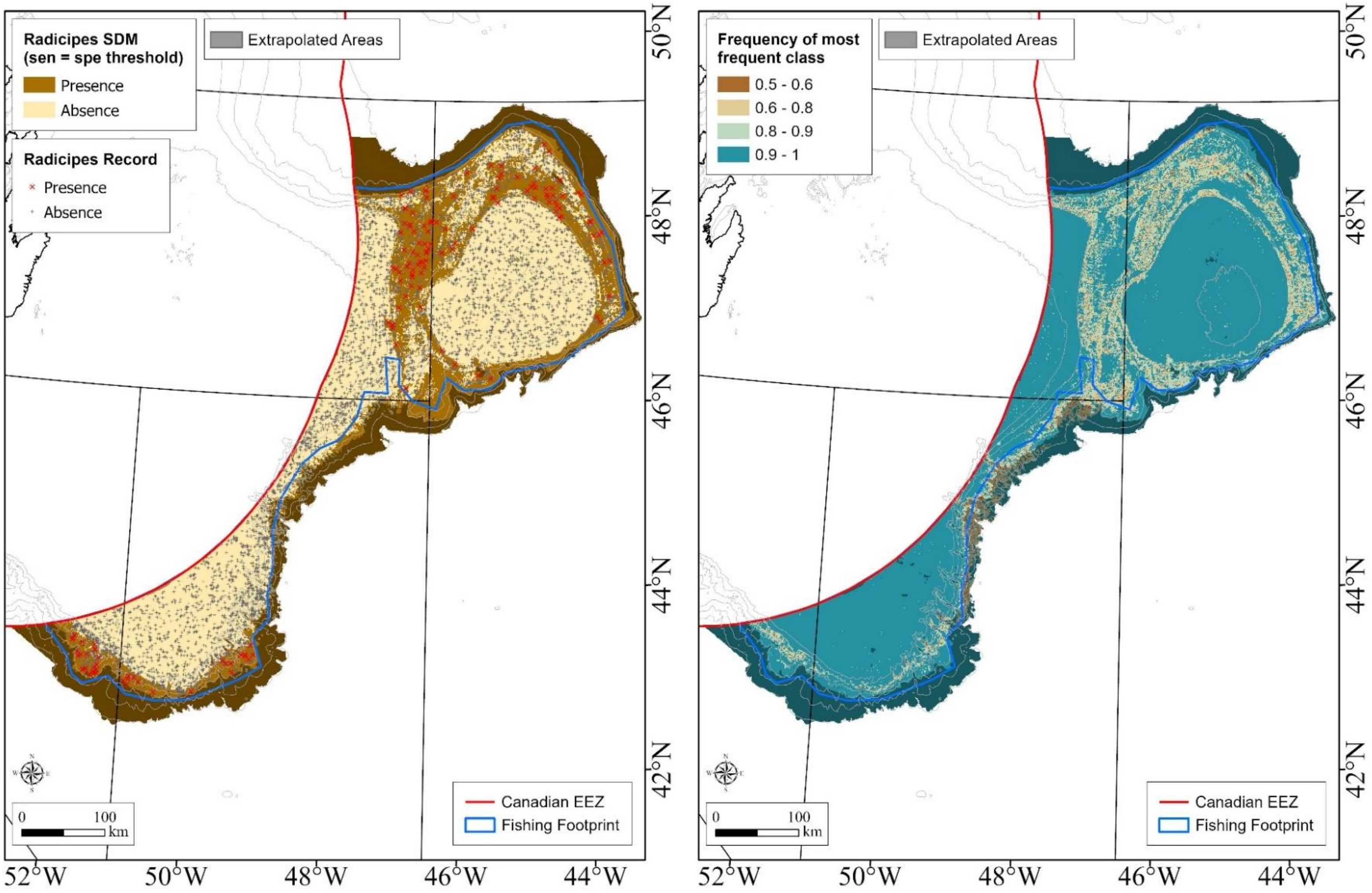

Figure 19. Random Forest species distribution model for *Radicipes gracilis* showing the distribution of the presence and absence data overlain on a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated by showing the frequency of P/A from the 10 cross-validation runs (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

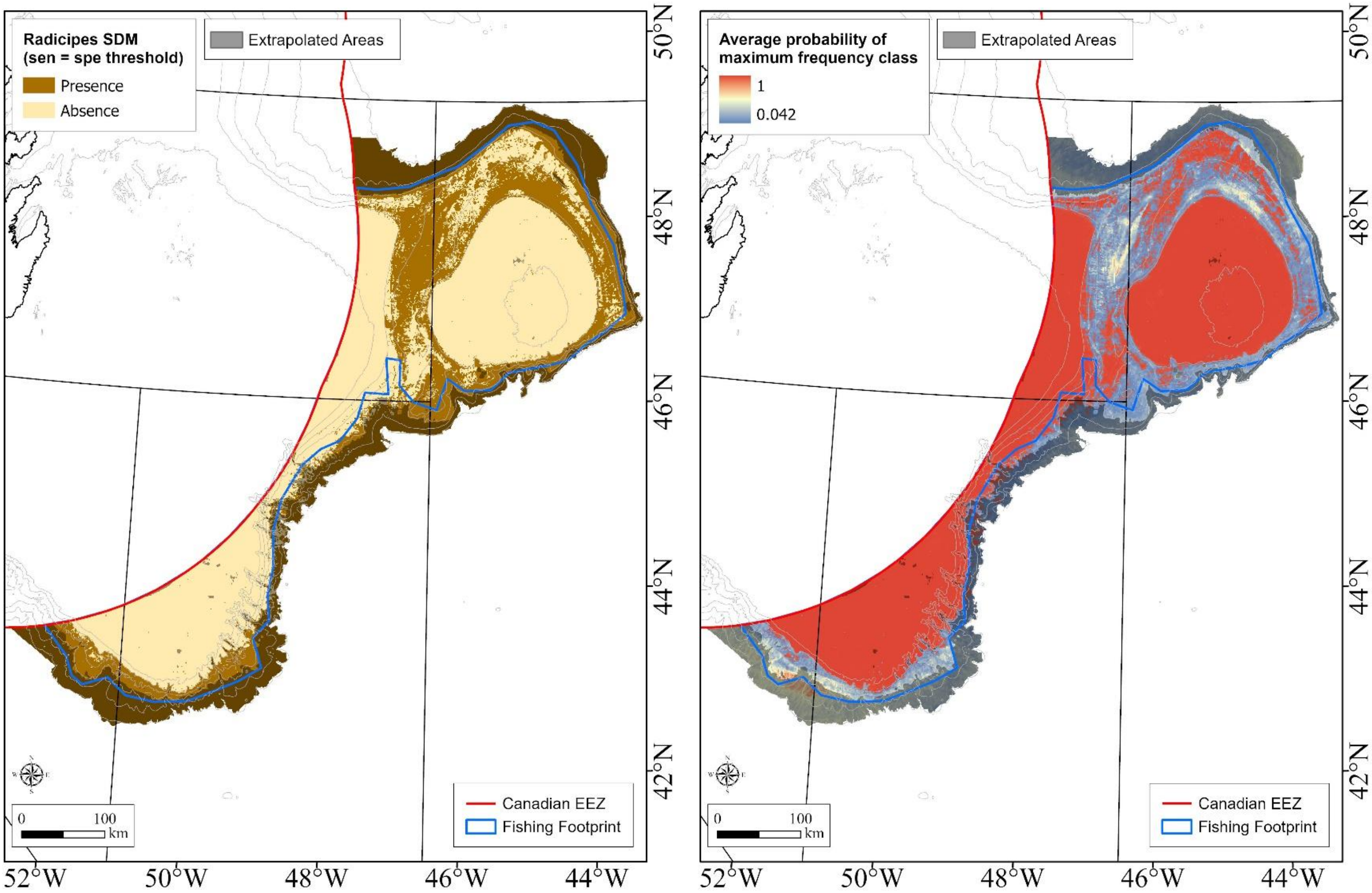

Figure 20. Random Forest species distribution model for *Radicipes gracilis* showing a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated as the average probability of the maximum frequency class (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

Assessment and Prediction of the Erect Bryozoan Functional Group

Random Forest models predicting the probability of the presence of the Erect Bryozoan Functional Group performed well across the validation statistics (Balanced Accuracy, Sensitivity and Specificity all > 0.7; Table 7) for both the functional group and habitat models. Statistics for the Erect Bryozoan VME Habitat model were higher than those for the functional group model (Table 7) with the TSS of 0.73 indicative of a highly accurate model.

Table 7. Model Validation Results for the Presence/Absence Random Forest Model for the Erect Bryozoan Functional Group and Erect Bryozoan VME Habitat. TSS=True Skill Statistic (Sensitivity + Specificity – 1).

	Erect Bryozoan Functional Group	**Erect Bryozoan VME Habitat**
Accuracy Measure	**Mean ± SD**	**Mean ± SD**
Sensitivity	0.72 ± 0.03	0.88 ± 0.19
Specificity	0.72 ± 0.03	0.85 ± 0.10
Kappa	0.24 ± 0.05	0.38 ± 0.23
Balanced Accuracy	0.72 ± 0.03	0.87 ± 0.14
TSS	0.44 ± 0.06	0.73 ± 0.28

Erect Bryozoan Functional Group

The mean of the average bottom temperature was the most influential variable in the model (Figure 21), followed by the average of the mean of the averaged bottom current speed and the mean of the minimum spring primary production and the mean of the averaged summer primary production. The Erect Bryozoan Functional Group were predicted to occur in areas of cold bottom temperatures averaging 0-3 °C and with average current speeds of > 0.1 m s^{-1} and high primary production in the spring and summer (Figure 22).

The predicted distribution maps are shown in Figure 23, shown as binary plots of presence/absence based on the two thresholds (Prevalence and Sensitivity=Specificity). These two plots are very similar, although the map based on prevalence has more restricted distribution. The distribution of the presence/absence data is shown overlain on the binary map based on Sensitivity=Specificity (Figure 24). Outside areas of model extrapolation, the Erect Bryozoan Functional Group is distributed in the shallow waters of the Tail of Grand Bank, along the outer slopes of the NRA and in the southern portion of Flemish Pass (Figures 24 and 25). The lowest uncertainty for Erect Bryozoan Functional Group presence is found on the shallow waters of the Nose and Tail of Grand Bank (Figures 24 and 25). They are not predicted to occur on the shallow waters of Flemish Cap with a high degree of certainty (Figures 24 and 25).

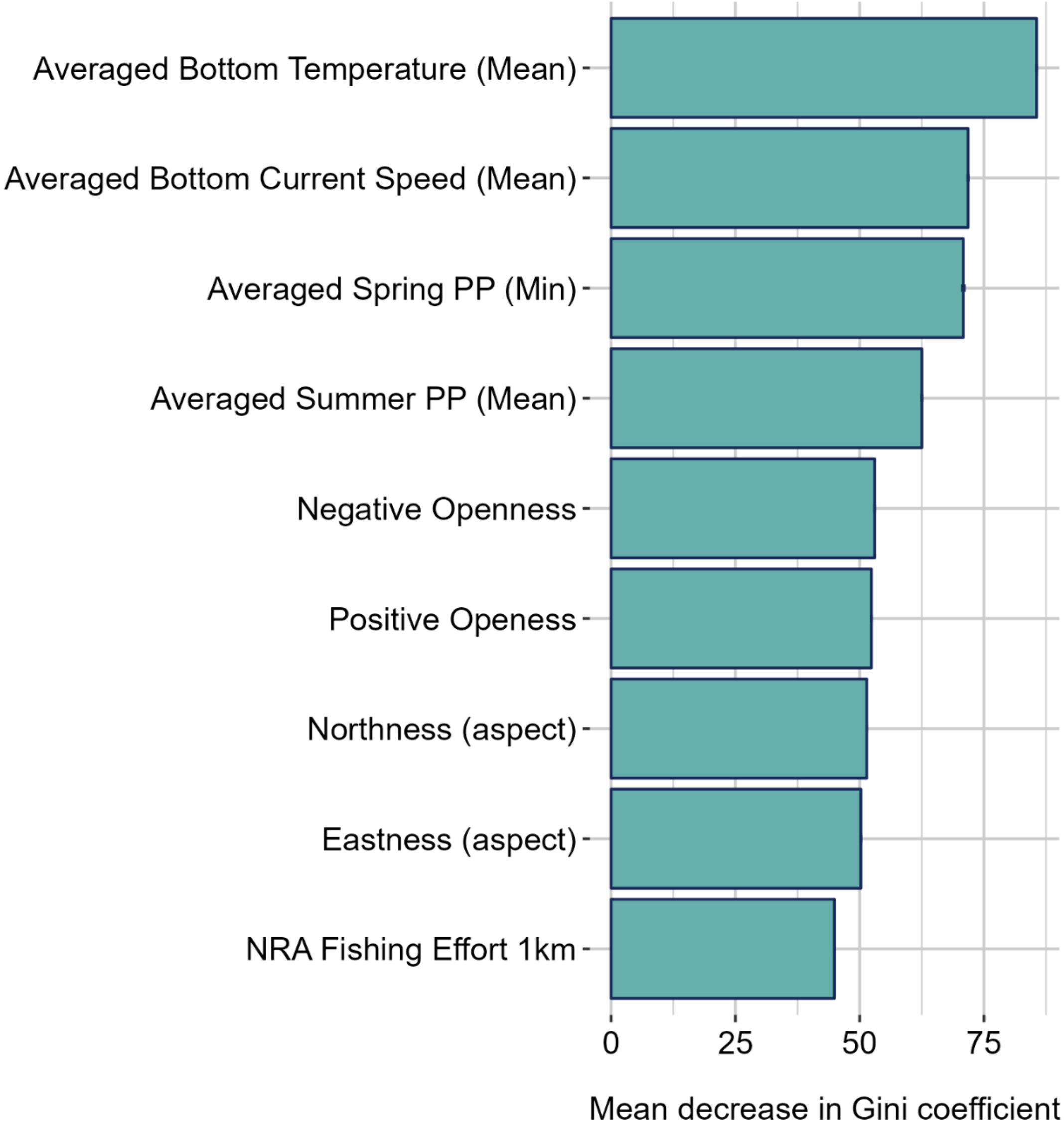

Figure 21. Plot of mean decrease and standard deviation in Gini Value for the 9 predictor variables in the Random Forest model for the Erect Bryozoan Functional Group, indicating their relative importance and variation across 10 model folds.

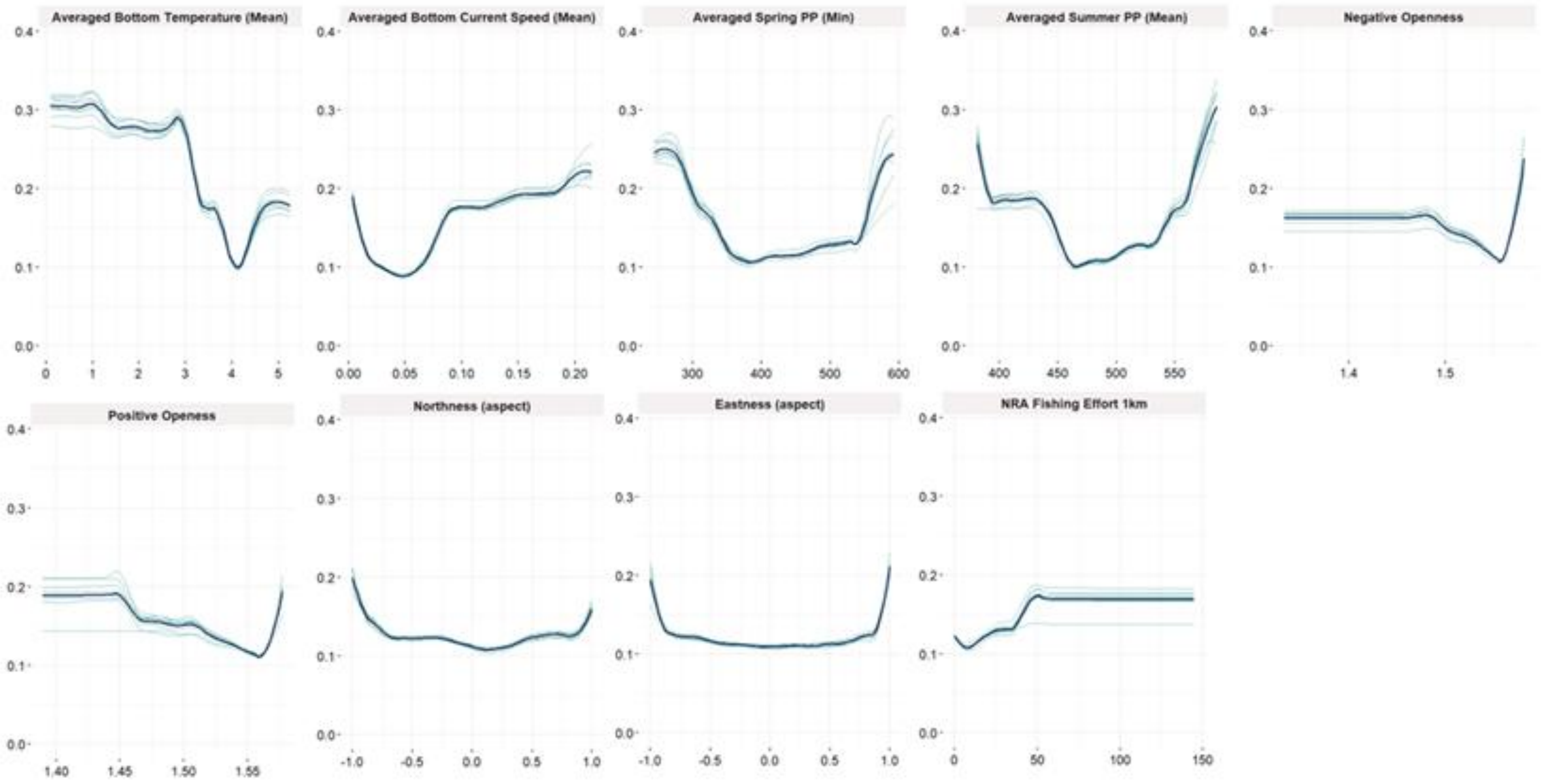

Figure 22. Response curves showing the partial dependence of the probability of presence on the predictors (Figure 21) identified in the Random Forest model for the Erect Bryozoan Functional Group. For each variable, the mean response and curves for each of the model folds are plotted. The plots show the predicted response to each predictor variable in turn, whilst other variables are held at their mean value.

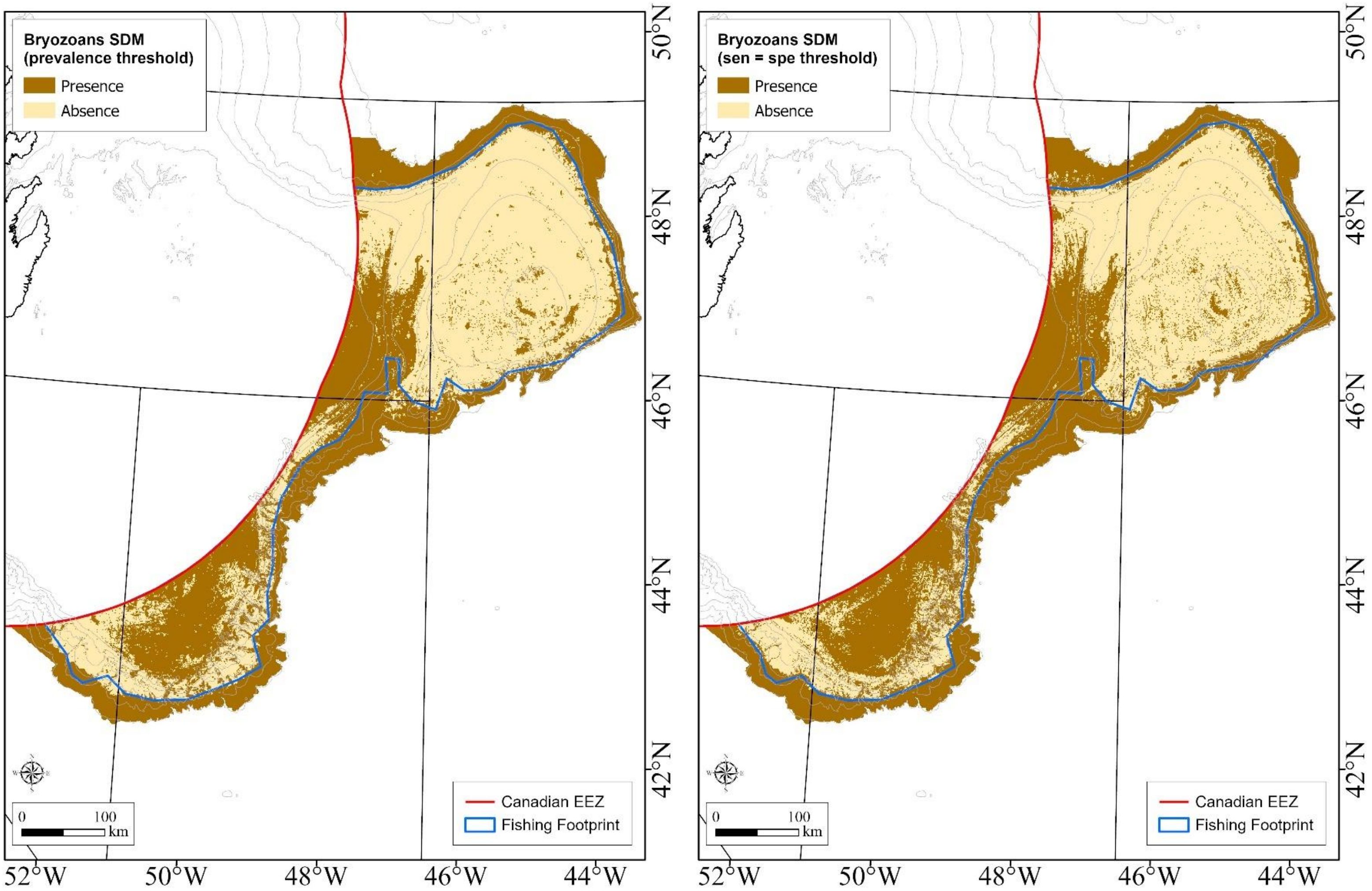

Figure 23. Random Forest species distribution model for the Erect Bryozoan Functional Group showing binary maps of VME presence thresholded using data prevalence (left panel) and a Sensitivity=Specificity threshold (right panel). The perimeter of the fishing footprint is shown on both maps.

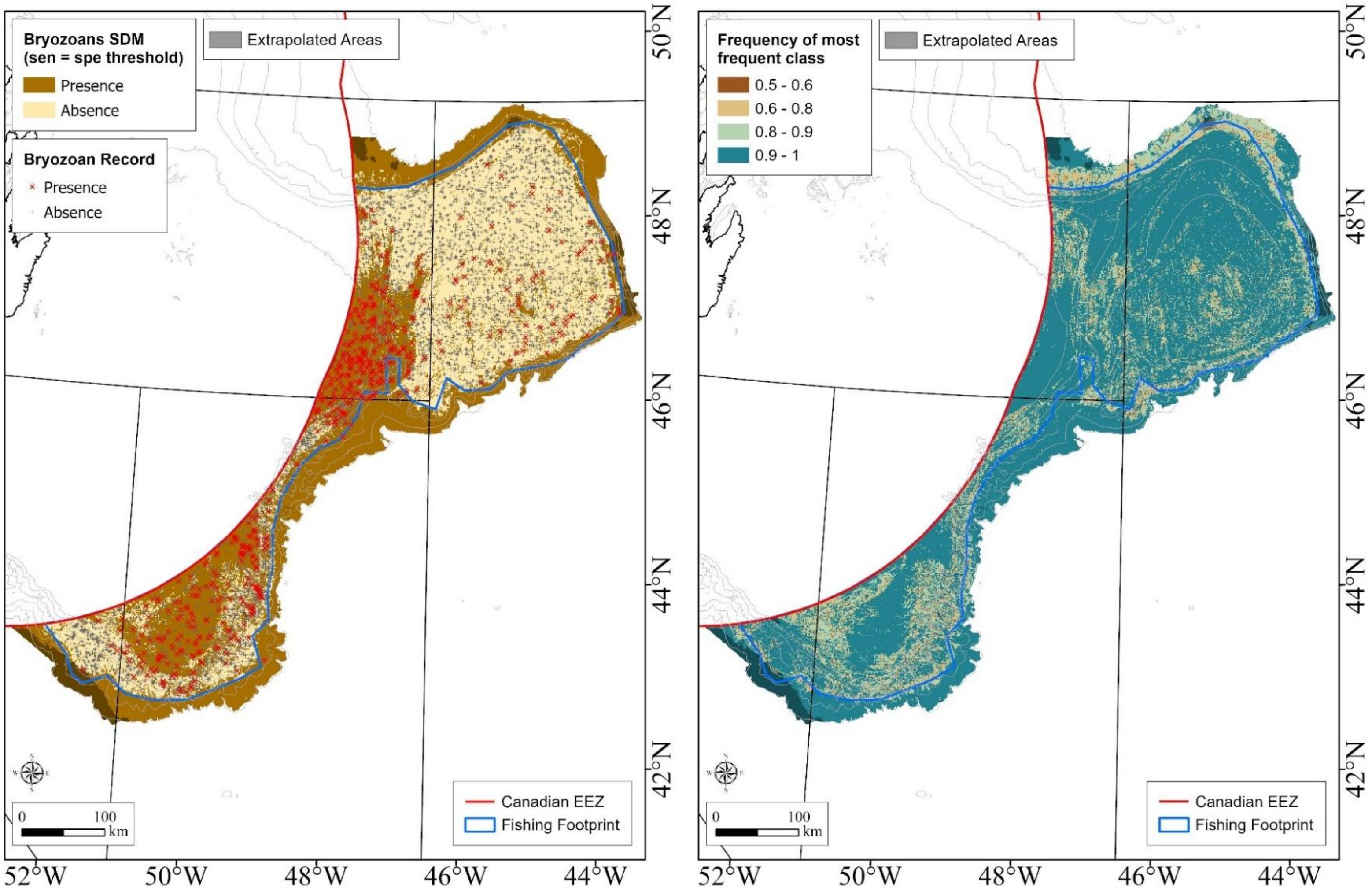

Figure 24. Random Forest species distribution model for the Erect Bryozoan Functional Group showing the distribution of the presence and absence data overlain on a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated by showing the frequency of P/A from the 10 cross-validation runs (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

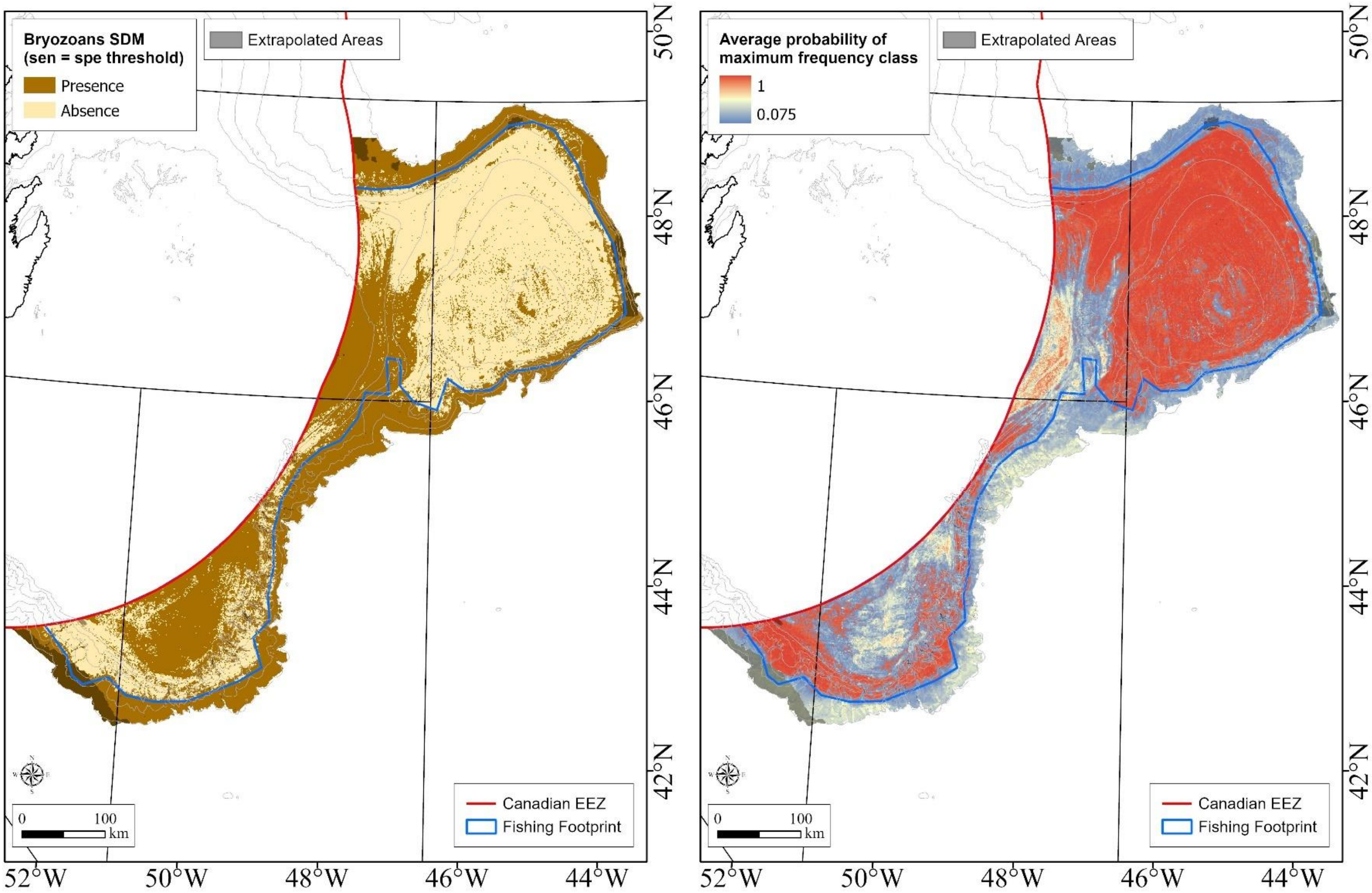

Figure 25. Random Forest species distribution model for the Erect Bryozoan Functional Group showing a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated as the average probability of the maximum frequency class (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

Erect Bryozoan VME Habitat

The SDMs for the Erect Bryozoan VME Habitat performed better than that of the functional group, likely because the functional group include several species whereas the VME habitat is mostly one single species (*Eucratea loricata*) suggesting that areas of high biomass can be more accurately predicted than areas of presence (Table 7). The mean of the average annual chlorophyll *a* was the most influential variable in the model, followed by the average of the mean of the fall primary production and the fill-sink bathymetry (Figure 26). The Erect Bryozoan VME Habitats were predicted to occur in areas of high mean annual chlorophyll *a* concentrations (> 0.9 mg m^{-3}) and fall primary production of > 500 mg C m^{-2} day^{-1} and relatively flat terrain (Figure 27).

The predicted distribution maps are shown in Figure 28, shown as binary plots of presence/absence based on the two thresholds (Prevalence and Sensitivity=Specificity). These two plots are very similar, although the Sensitivity=Specificity threshold map predicts less presence areas on the northern portion of Flemish Cap and along the slopes of the Grand Bank, both areas of model extrapolation (Figure 29). The distribution of the data is shown overlain on the binary map of presence/absence based on Sensitivity=Specificity (Figure 29). Outside areas of model extrapolation, the Erect Bryozoan VME Habitat is distributed on the shallow waters of the Tail of Grand Bank where it occurs with high certainty (Figures 29 and 30). They are not predicted to occur elsewhere with a high degree of certainty (Figures 29 and 30).

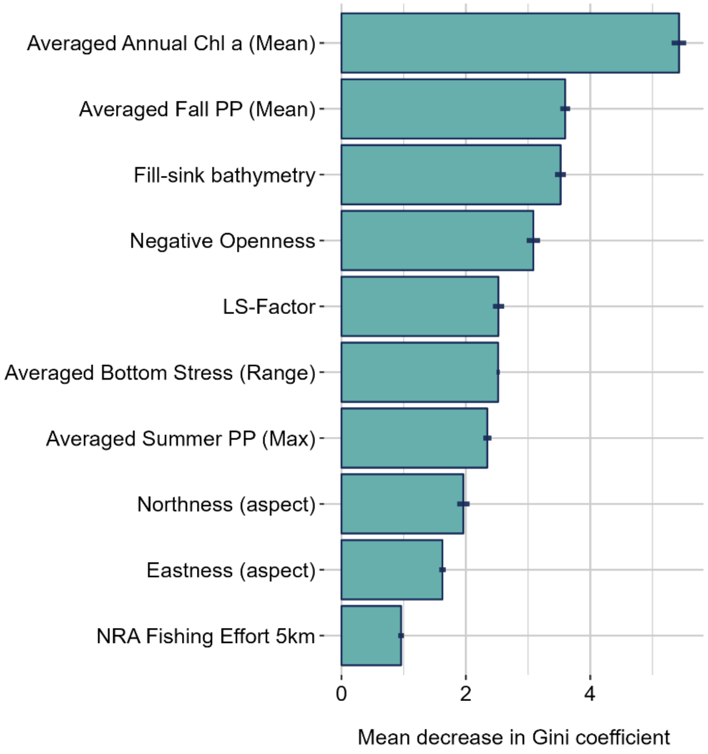

Figure 26. Plot of mean decrease and standard deviation in Gini Value for the 10 predictor variables in the Random Forest model for Erect Bryozoan VME Habitat, indicating their relative importance and variation across 10 model folds.

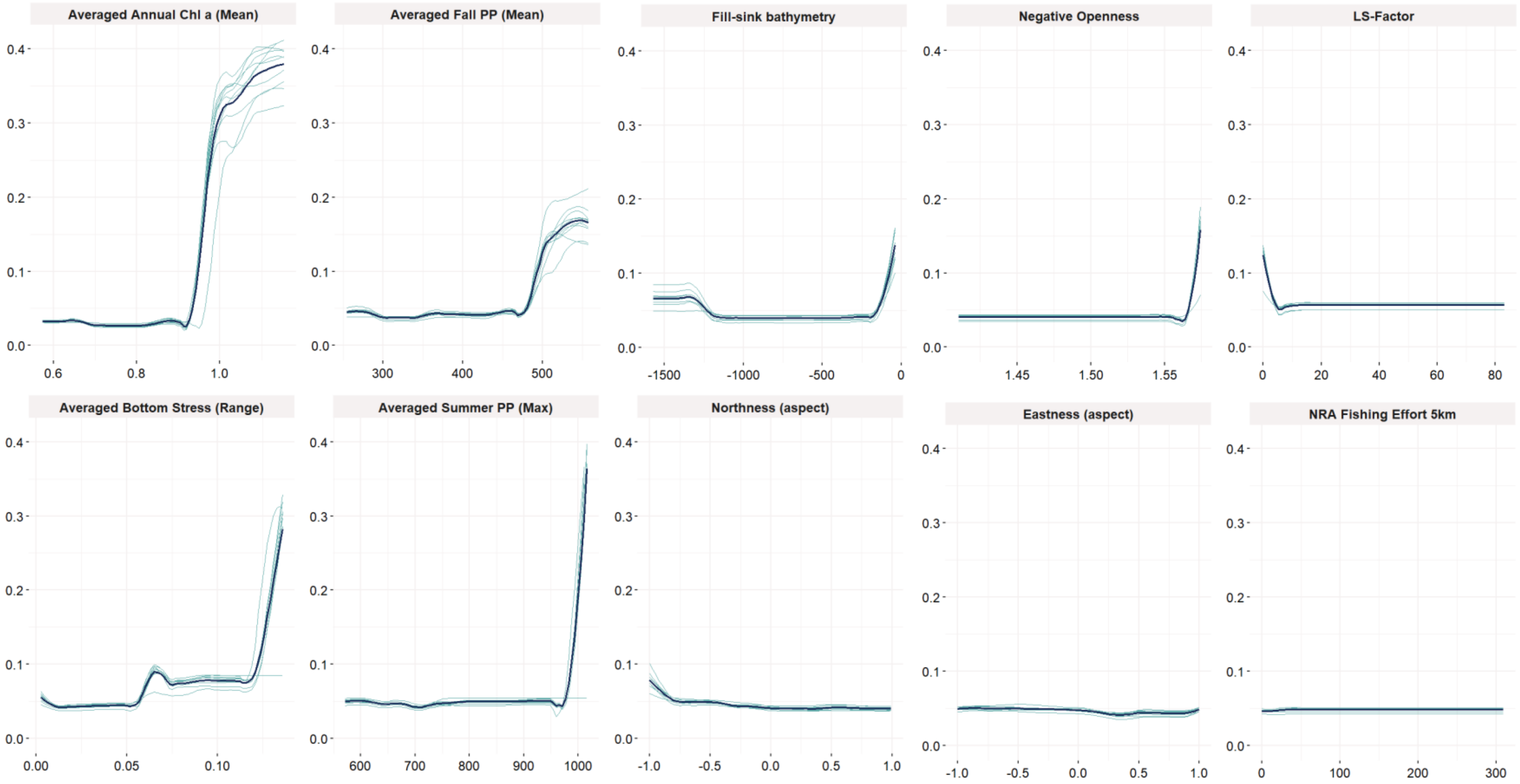

Figure 27. Response curves showing the partial dependence of the probability of presence on the predictors (Figure 26) identified in the Random Forest model for Erect Bryozoan VME Habitat. For each variable, the mean response and curves for each of the model folds are plotted. The plots show the predicted response to each predictor variable in turn, whilst other variables are held at their mean value.

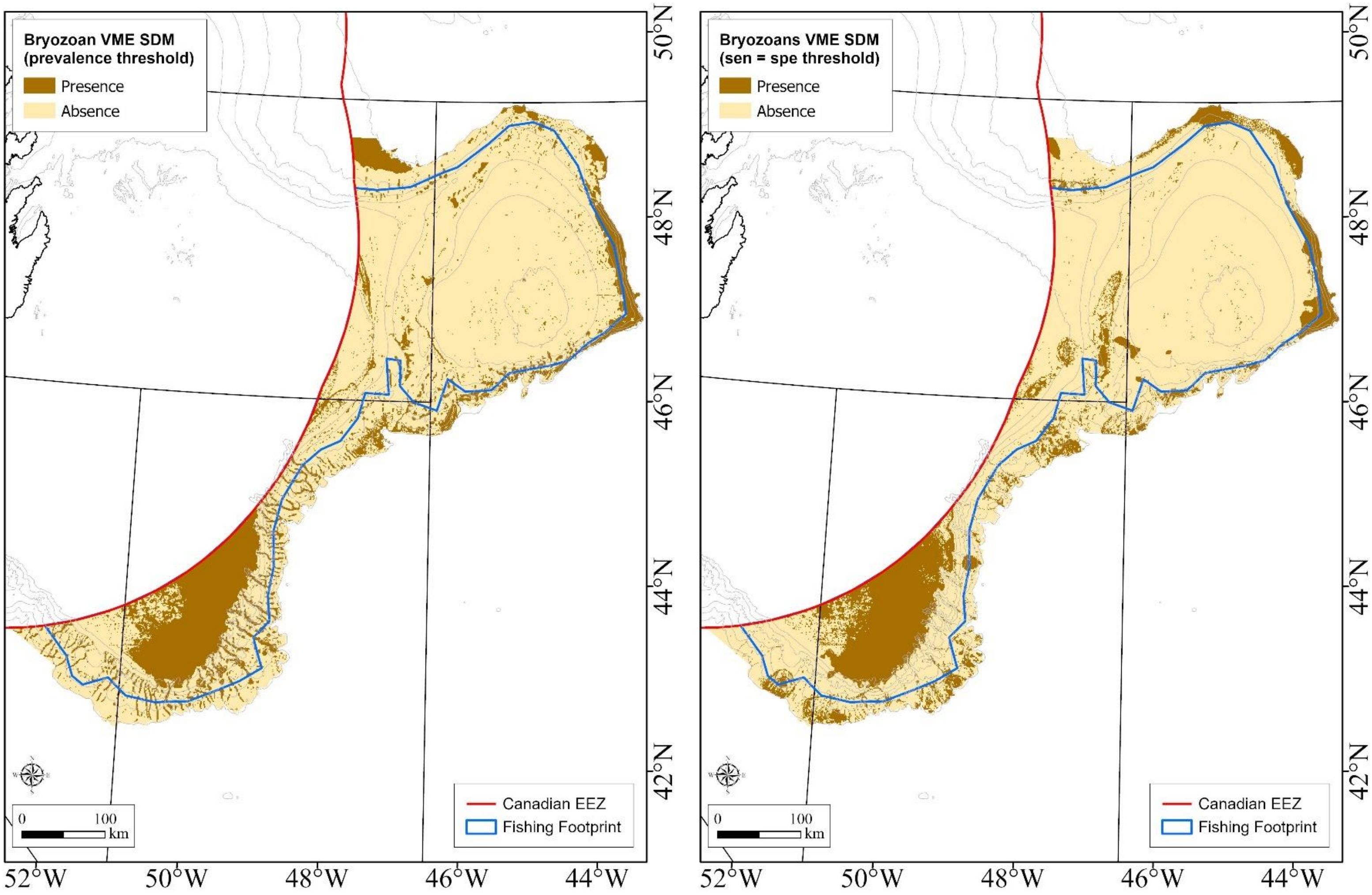

Figure 28. Random Forest species distribution model for Erect Bryozoan VME Habitat showing binary maps of VME presence thresholded using data prevalence (left panel) and a Sensitivity=Specificity threshold (right panel). The perimeter of the fishing footprint is shown on both maps.

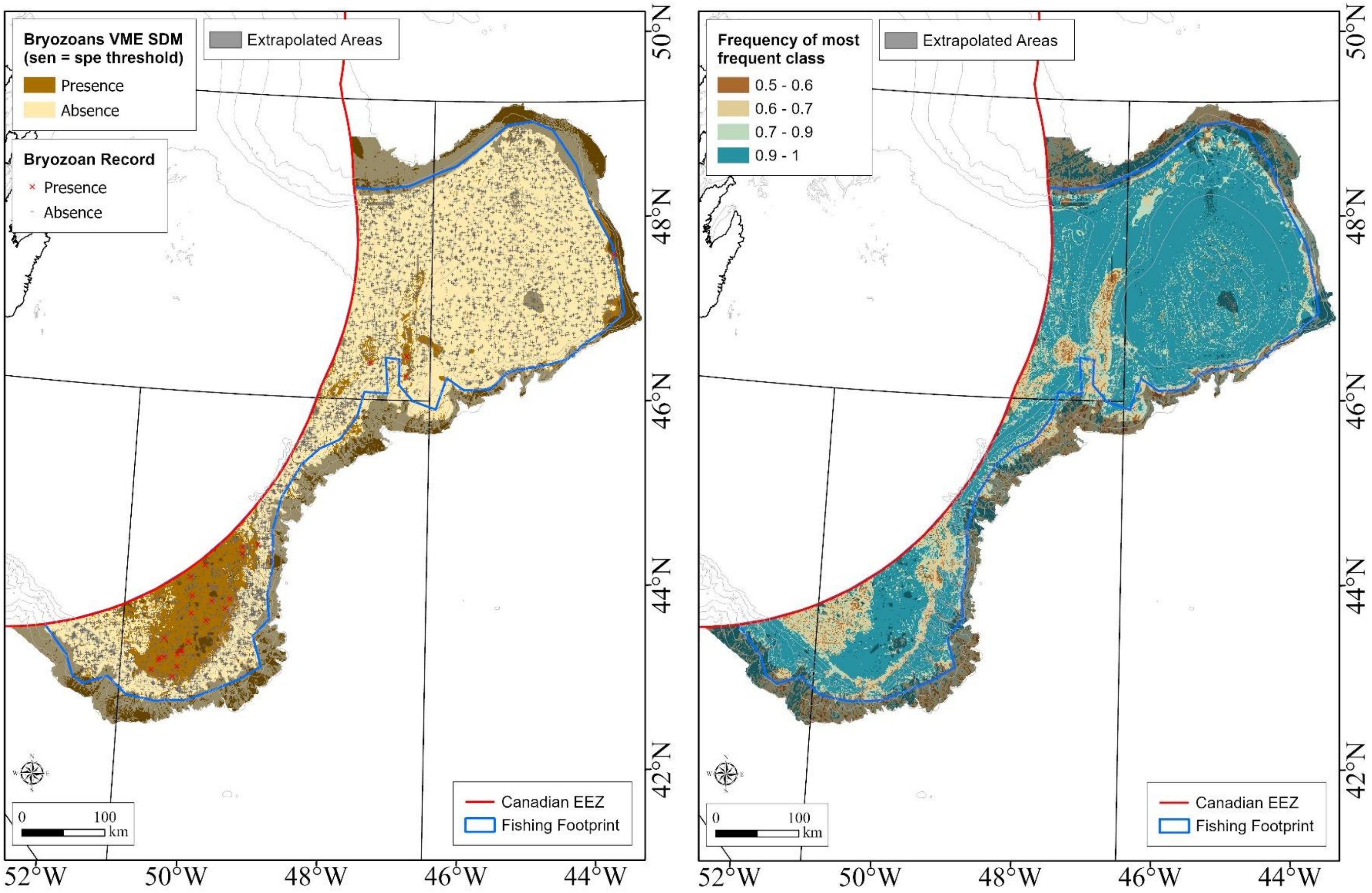

Figure 29. Random Forest species distribution model for Erect Bryozoan VME Habitat showing the distribution of the presence and absence data overlain on a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated by showing the frequency of P/A from the 10 cross-validation runs (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

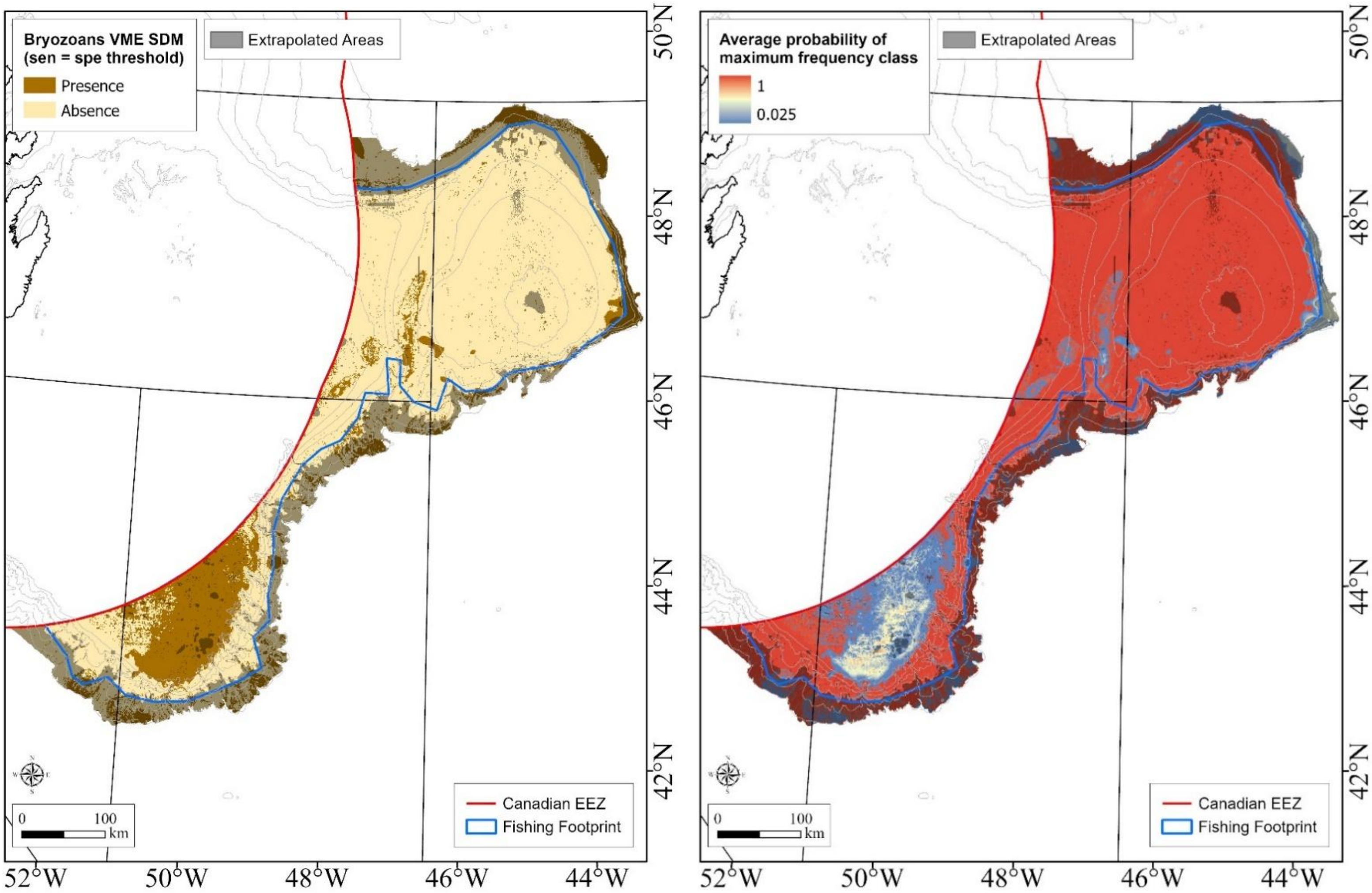

Figure 30. Random Forest species distribution model for Erect Bryozoan VME Habitat showing a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated as the average probability of the maximum frequency class (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

Assessment and Prediction of the Sea Squirt (Boltenia ovifera) Functional Group

Random Forest models predicting the probability of the presence of the Sea Squirt (*Boltenia ovifera*) Functional Group performed well across the validation statistics (Balanced Accuracy, Sensitivity and Specificity all > 0.83) and the TSS was 0.68 which indicates a highly accurate model (Table 8).

Table 8. Model Validation Results for the Presence/Absence Random Forest Model for the Sea Squirt (*Boltenia ovifera*) Functional Group. TSS=True Skill Statistic (Sensitivity + Specificity – 1).

	Sea Squirt Functional Group
Accuracy Measure	**Mean ± SD**
Sensitivity	0.85 ± 0.04
Specificity	0.83 ± 0.04
Kappa	0.25 ± 0.07
Balanced Accuracy	0.84 ± 0.04
TSS	0.68 ± 0.08

The SDMs for the presence of the Sea Squirt (*Boltenia ovifera*) Functional Group were most influenced by the maximum of the summer mixed layer depth, followed by the average of the minimum summer primary production and the fill-sink bathymetry (Figure 31). The Sea Squirt (*Boltenia ovifera*) Functional Group was predicted to occur in areas where the mixed layer depth is less than 7 m in summer, where the mean minimum summer primary production is > 375 mg C m^{-2} day^{-1} and shallower than 250 m depth (Figure 32).

The predicted distribution maps are shown in Figure 33, shown as binary plots of presence/absence based on the two thresholds (Prevalence and Sensitivity=Specificity). These two plots are very similar. The distribution of the data is shown overlain on the binary map of presence/absence based on Sensitivity=Specificity (Figure 34). Outside areas of model extrapolation, the Sea Squirt (*Boltenia ovifera*) Functional Group is distributed on the shallow waters of the Nose and Tail of Grand Bank. They occur with high certainty in the east and northeast portions of the presence area (Figures 34) where it reaches the higher probabilities (Figure 35). The rest of the presence areas are associated with higher uncertainty (Figures 34 and 35). They are not predicted to occur on the shallow waters of Flemish Cap or the northern portion of Flemish Pass.

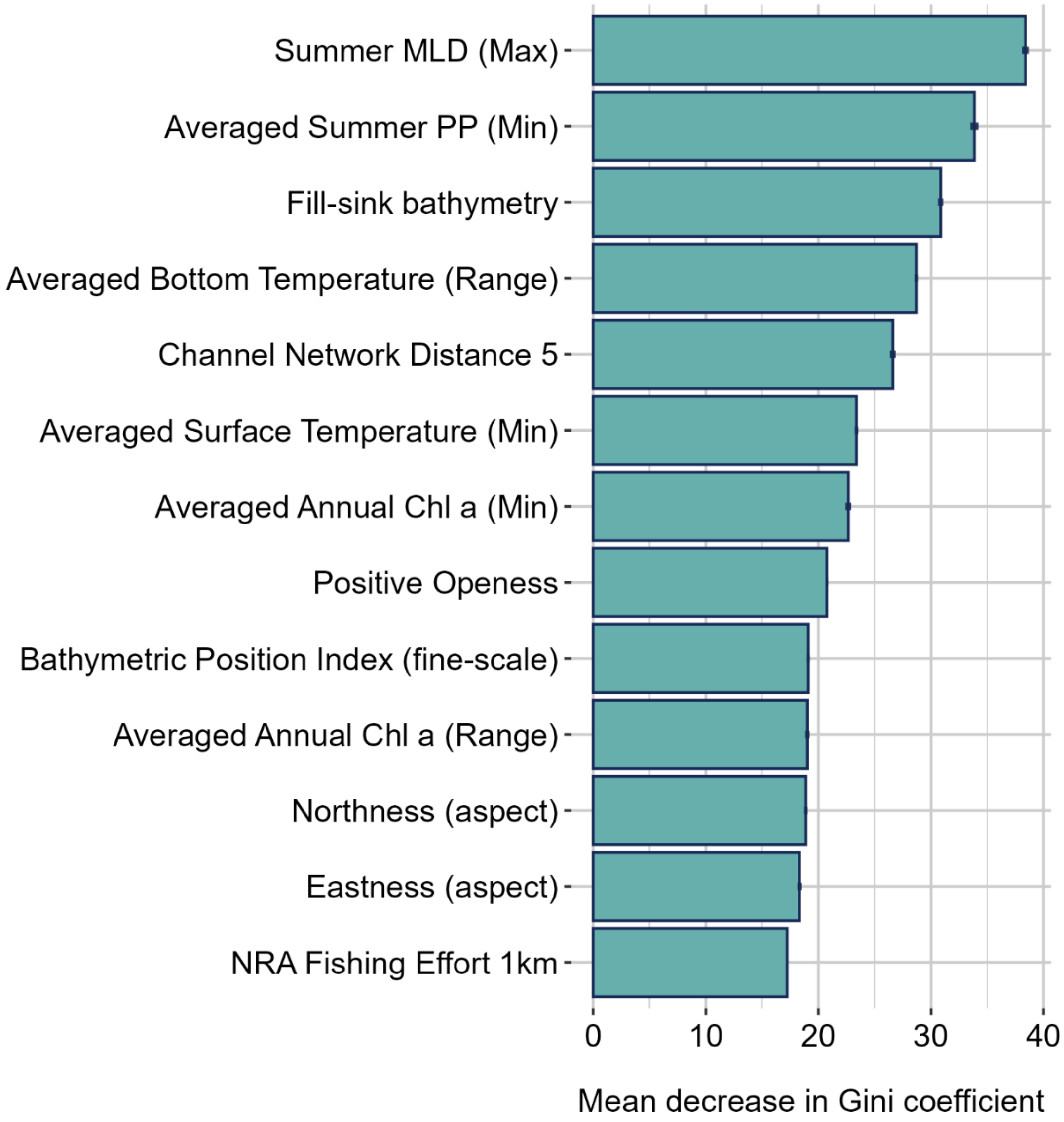

Figure 31. Plot of mean decrease and standard deviation in Gini Value for the 13 predictor variables in the Random Forest model for Sea Squirt (*Boltenia ovifera*) Functional Group, indicating their relative importance and variation across 10 model folds.

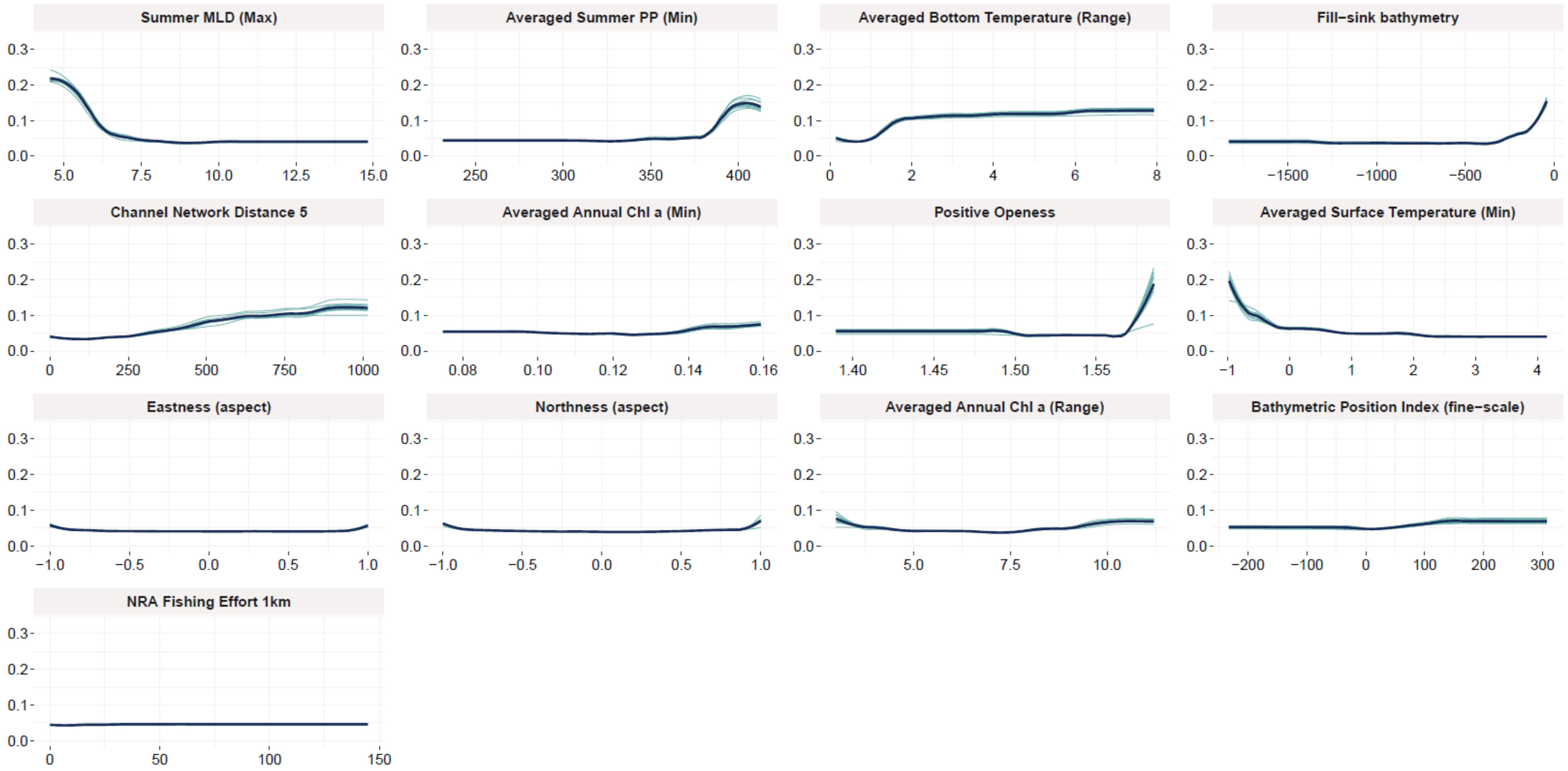

Figure 32. Response curves showing the partial dependence of the probability of presence on the predictors (Figure 31) identified in the Random Forest model for the Sea Squirt (*Boltenia ovifera*) Functional Group. For each variable, the mean response and curves for each of the model folds are plotted. The plots show the predicted response to each predictor variable in turn, whilst other variables are held at their mean value.

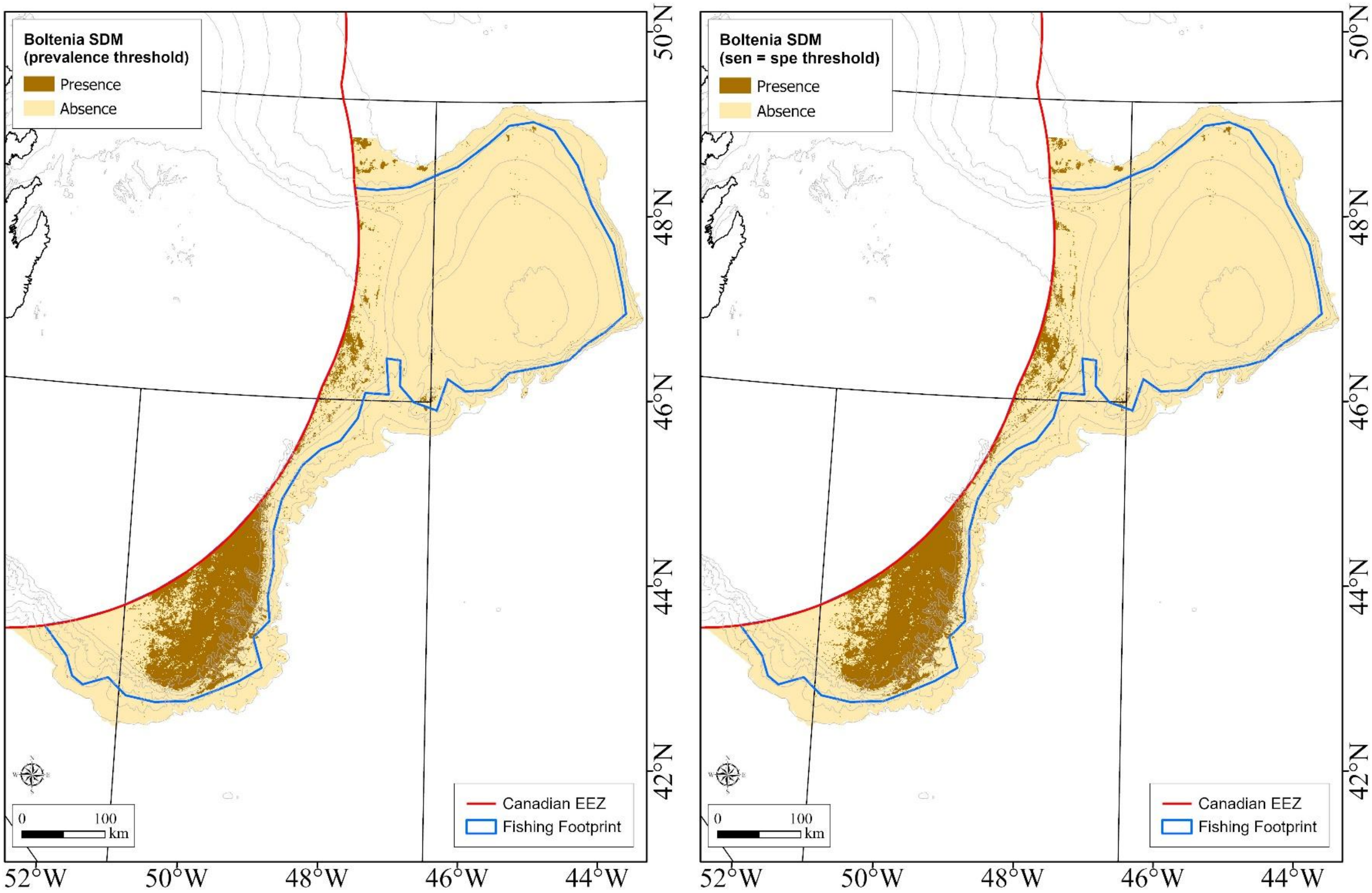

Figure 33. Random Forest species distribution model for the Sea Squirt (*Boltenia ovifera*) Functional Group showing binary maps of VME presence thresholded using data prevalence (left panel) and a Sensitivity=Specificity threshold (right panel). The perimeter of the fishing footprint is shown on both maps.

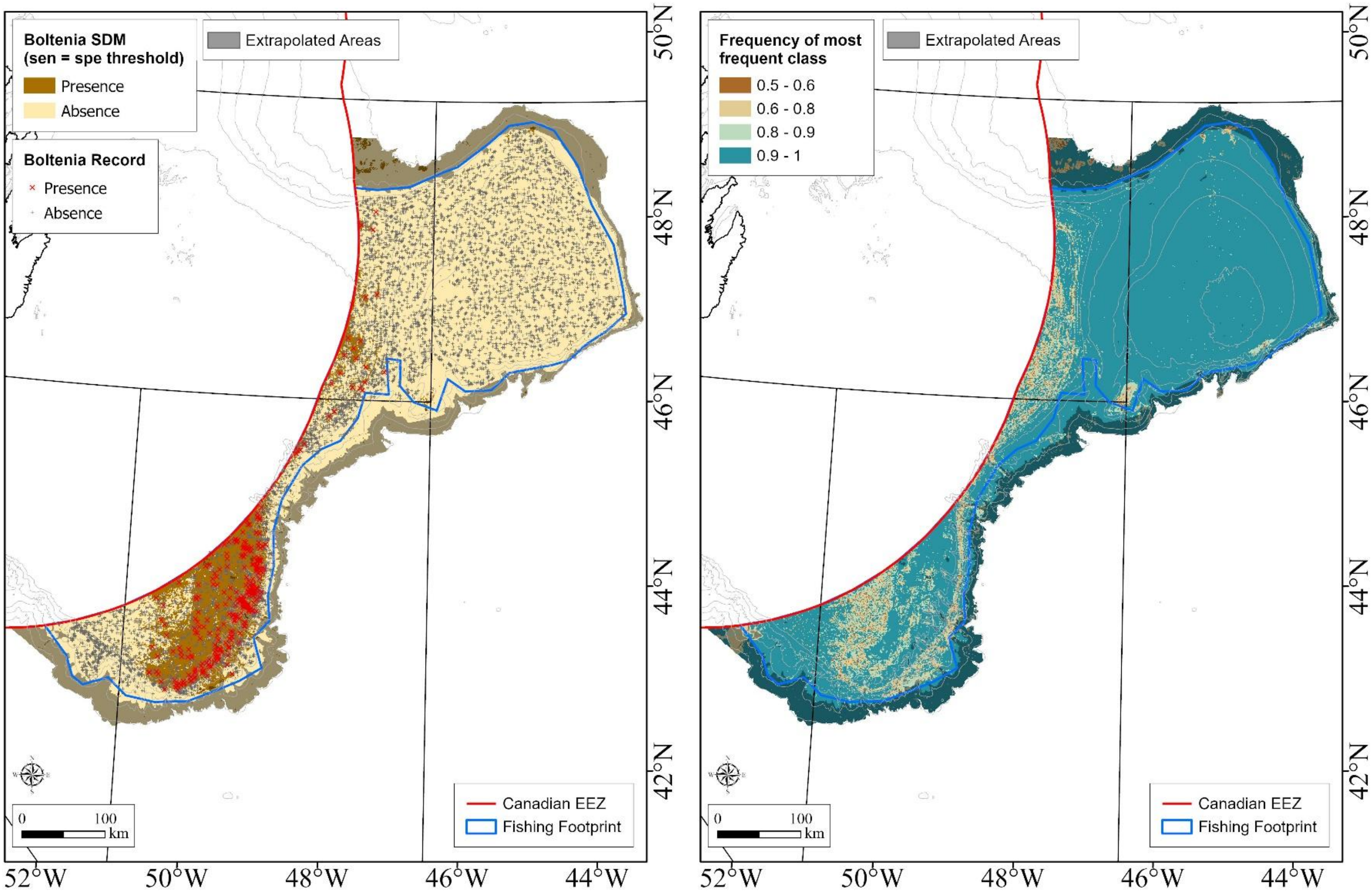

Figure 34. Random Forest species distribution model for the Sea Squirt (*Boltenia ovifera*) Functional Group showing the distribution of the presence and absence data overlain on a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated by showing the frequency of P/A from the 10 cross-validation runs (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

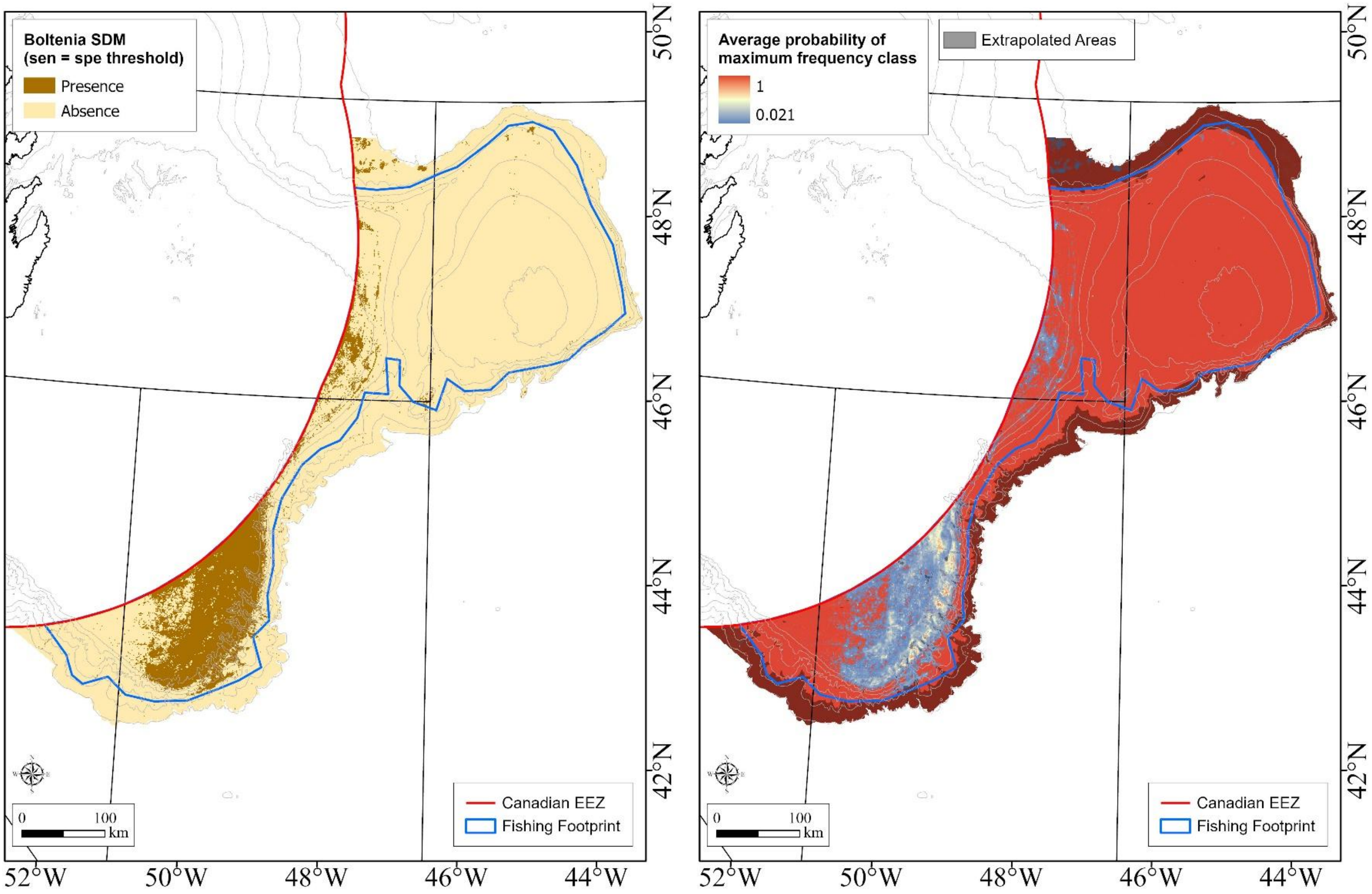

Figure 35. Random Forest species distribution model for the Sea Squirt (*Boltenia ovifera*) Functional Group showing a binary map thresholded using a Sensitivity=Specificity threshold (left panel). Model uncertainty is illustrated as the average probability of the maximum frequency class (right panel). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps.

Discussion

All models generally scored high accuracy across the validation statistics. The binary presences/absences maps are based on a threshold of Sensitivity=Specificity, which is the threshold where the chance of correctly predicting a positive or negative observation is the same. Previously, Prevalence (the ratio of presence/absence) was used which produced very similar outputs. However, a threshold of Sensitivity=Specificity will be used for the 2027 review of the closed areas to place equal emphasis on presence and absence areas since the latter is used to modify the VME polygons generated from the KDE analyses (Kenchington et al., 2019).

Models of the functional group and subgroups for the Small Gorgonian Corals illustrated that there is minimal potential for unequal protection of this VME Indicator taxa given the similarity of the areas of model prediction (Figure 36). The model performed on the Erect Bryozoan VME habitat (Figures 29 and 30), produced using only presence records above the KDE catch threshold for that taxon, performed much better than that of the functional group and was the only model to predict large areas of presence with high certainty.

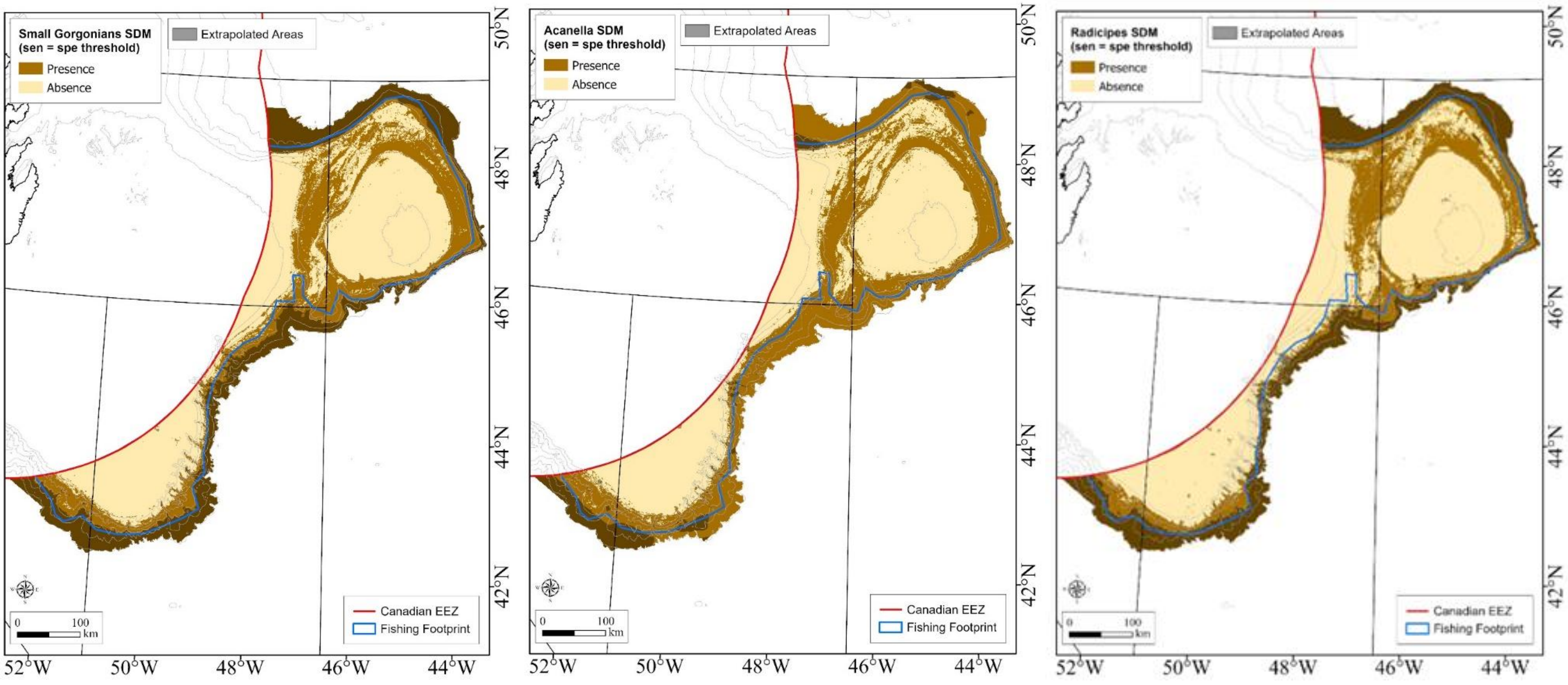

Figure 36. Random Forest species distribution model showing binary maps of VME presence thresholded using a Sensitivity=Specificity threshold for Small Gorgonian Coral Functional Group (left), *Acanella arbuscula* (middle) and *Radicipes gracilis* (right). The areas of extrapolation show where the model has predicted into areas outside of the environment for the presence and absence records. The perimeter of the fishing footprint is shown on both maps in blue.

Acknowledgements

The data collection of the EU Groundfish Surveys used in this paper has been funded by the EU through the European Maritime, Fisheries and Aquaculture Fund (EMFAF) within the Spanish Work Plan for the collection of data in the fisheries and aquaculture sectors in relation to the Common Fisheries Policy; and within the Portuguese Work Plan funded by the EU Data Collection Framework (DCF). The NEREIDA project was funded by the EU through the NAFO Secretariat. We also thank Dr. Bárbara Neves and Vonda Hayes, Department of Fisheries and Oceans, Canada based at the Northwest Atlantic Fisheries Centre in St. John's, NL, for provision of the data from the Canadian surveys. Contributions of EK, JM, CL, NP, ZW, ED, and CN were supported by DFO's Competitive Science Research Fund (CSRF). This report was prepared by the Species Distribution Modeling Subgroup of WG-ESA. We thank all participants in the 2024 WG-ESA meeting for the valuable comments and contributions to this work.

References

Allouche, O., Tsoar, A., and Kadmon, R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43: 1223–1232.

Bouchet, P. J., Miller, D. L., Roberts, J. J., Mannocci, L., Harris, C. M. and Thomas, L. 2020. Dsmextra: Extrapolation Assessment Tools for Density Surface Models. Methods Ecol. Evol. 11, 1464–1469. doi: 10.1111/2041-210X.13469.

Breiman, L. 2001. Random Forests. Mach. Learn. 45: 5–32.

Brenning, A., Bangs, D., and Becker, M. 2022. RSAGA: SAGA geoprocessing and terrain analysis. R package v. 1.4.0. https://CRAN.R-project.org/package=RSAGA .

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Böhner, J. 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4, Geosci. Model Dev. 8: 1991-2007.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., and Lawler, J.J. 2007. Random Forests for classification in ecology. Ecology 88: 2783–2792.

GEBCO Compilation Group. 2024. GEBCO 2024 Grid. doi:10.5285/1c44ce99-0a0d-5f4f-e063-7086abc0ea0f.

Hijmans, R. 2024. terra: Spatial Data Analysis. R package version 1.7-83. https://CRAN.R-project.org/package=terra .

Hijmans, R. 2023. raster: Geographic data analysis and modelling. R package v. 3.6-26. https://CRAN.R-project.org/package=raster .

Ilich, A. R., Misiuk, B., Lecours, V., and Murawski, S. A. 2023. MultiscaleDTM: An open-source R package for multiscale geomorphometric analysis. Trans. GIS. 27(4). https://doi.org/10.1111/tgis.13067.

Kenchington, E., Lirette, C., Murillo, F.J., Beazley, L., and Downie, A. L. 2019. Vulnerable Marine Ecosystems in the NAFO Regulatory Area: Updated Kernel Density Analyses of Vulnerable Marine Ecosystem Indicators. NAFO SCR Doc. 19/058, Serial No. N7030, 68 pp.

Kenchington, E., Beazley, L., Lirette, C., Murillo, F.J., Guijarro, J., Wareham, V., Gilkinson, K., Koen Alonso, M., Benoît, H., Bourdages, H., Sainte-Marie, B., Treble, M., and Siferd, T. 2016. Delineation of Coral and Sponge Significant Benthic Areas in Eastern Canada Using Kernel Density Analyses and Species Distribution Models. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/093. vi + 178 p.

Kenchington, E., Murillo, F.J., Lirette, C., Sacau, M., Koen-Alonso, M., Kenny, A., Ollerhead, N., Wareham, V. and Beazley, L. 2014. Kernel density surface modelling as a means to identify significant concentrations of vulnerable marine ecosystem indicators. PLoS ONE 9(10): e109365. doi:10.1371/journal.pone.0109365.

Knudby, A., Lirette, C., Kenchington, E., and Murillo, F.J. 2013. Species Distribution Models of Black Corals, Large Gorgonian Corals and Sea Pens in the NAFO Regulatory Area. NAFO SCR Doc. 13/78, Serial No. N6276, 17 pp.

Liaw, A., and Wiener, M. 2002. Classification and Regression by randomForest. R News 2: 18–22.

Mesgaran, M. B., Cousens, R. D. and Webber, B. L. 2014. Here be Dragons: A Tool for Quantifying Novelty Due to Covariate and Correlation Change When Projecting Species Distribution Models. Diversity Distrib. 20, 1147–1159. doi: 10.1111/ddi.12209

Murillo, F.J., Downie, A.-L., Abalo Morla, S., Lirette, C., Paulin, N., Wang, Z., Devred, E., Clay, S., Sacau, M., Nozères, C., Koen-Alonso, M., Gullage, L., and Kenchington, E. 2024. Vulnerable marine ecosystems in the NAFO

Regulatory Area: updated species distribution models of selected vulnerable marine ecosystem indicators (large-sized sponges, sea pens and black corals). NAFO SCR Doc. 24/063, Serial No. N7601, 105 pp.

NAFO. 2024. Northwest Atlantic Fisheries Organization Conservation and Enforcement Measures 2024. NAFO/COM Doc. 24-01. Serial No. N7490.

NAFO. 2019. Report of the 12th Meeting of the NAFO SC Working Group on Ecosystem Science and Assessment (WGESA) – November 2019. NAFO SCS Doc. 19/25. Serial No. N7027. https://www.nafo.int/Portals/0/PDFs/sc/2019/scs19-25.pdf

R Development Core Team. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

R Development Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Sofaer, H.R., Jarnevich, C.S., Pearse, I.S., Smyth, R.L., Auer, S., Cook, G.L., Edwards, T.C. Jr., Guala, G.F., Howard, T.G., Morisette, J.T., and Hamilton, H. 2019. Development and delivery of species distribution models to inform decision-making. BioScience 69: 544-557.

Wang, L., and Liu, H. 2006. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. Int. J. Geogr. Inf. Sci. 20: 193-213.

Wang, Z., Lu, Y., Greenan, B., Brickman, D., and DeTracey, B. 2018. An eddy-resolving North Atlantic model (BNAM) to support ocean monitoring. Can. Tech. Rep. Hydrogr. Ocean Sci. 327: vii + 18 pp.

Wareham, V.E., and Edinger, E.N. 2007. Distribution of deep-sea corals in the Newfoundland and Labrador region, Northwest Atlantic Ocean. Bulletin of Marine Science 81(3): 289-313.

Zurell, D., Franklin, J., Konig, C., Bouchet, P.J., Dormann, C.F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Fuisan, A., Lahoz-Monfort, J.J., Leitao, P.J., Park, D.S., Townsend Peterson, A., Rapacciuolo, G., Schmatz, D.R., Schroder, B., Serra-Diaz, J.M., Thuiller, W., Yates, K.L., Zimmermann, N.E., and Merow, C. 2020. A standard protocol for reporting species distribution models. Ecography (Cop.) 43: 1261-1277.

Appendix

Table A1. At-sea Identification Nomenclature and Corresponding Number of raw Records for Each of Large Gorgonian Coral, Small Gorgonian Coral, Erect Bryozoan and Sea Squirt (*Boltenia ovifera*) Functional Groups Considered for the Preparation of the Response Data (Table 4) in the Species Distribution Models. *Indicates taxon from the records of the Canadian DFO NL Multi-species Surveys (Table 3); All other taxa are as recorded from the EU Surveys undertaken by Spain and Portugal (Table 3).

At-Sea Identification for Large Gorgonian Corals	**Number of Records**
Acanthogorgia	18
Acanthogorgia armata*	11
ACANTHOGORGIA SP.	2
Acanthogorgia sp.	13
Acanthogorgiidae	32
Keratoisis	27
Keratoisis cf. flexibilis*	2
Octocorallia sp. (SUBCLASS)*	42
Paragorgia	7
PARAGORGIA ARBOREA	4
Paragorgia arborea	3
Paragorgia arborea*	15
Paragorgia spp	1
Paramuricea placomus	2
Paramuricea	9
PARAMURICEA SP	2
Paramuricea sp.	4
Paramuricea sp.*	4
Primnoa resedaeformis	2
Primnoa resedaeformis*	2
PRIMNOIDAE	1
Taxon Name Not Provided	134

At-Sea Identification for Small Gorgonian Corals	**Number of Records**
Acanella	2
Acanella arbuscula	425
Acanella arbuscula*	93
ACANELLA ARBUSCULA	40
Anthothela	2
Anthothelidae	2
Anthothela grandiflora*	3
Octocorallia sp. (SUBCLASS)*	22
Isididae	20
Radicipes	78
Radicipes gracilis	44
Radicipes gracilis*	4
RADICIPES SP	10
Radicipes sp.	13
Radicipes spp	4
Taxon Name Not Provided	188

At-Sea Identification for Erect Bryozoans	**Number of Records**
Bryozoa	508
BRYOZOA	53
BRYOZOAN ECT. OR ENT.*	5
Taxon Name Not Provided	276

At-Sea Identification Sea Squirts	**Number of Records**
BOLTENIA OVIFERA	128
BOLTENIA SP.	14
TUNICATE, SESSILE*	174
Taxon Name Not Provided	64

Table A2. The Number of Raw Records with Taxon Name Provided by VME Functional Group (Large-Gorgonian Corals, Small Gorgonian Corals, Erect Bryozoans, Sea Squirts) by Year.

Year	No. Records Large Gorgonian Corals	No. Records Small Gorgonian Corals	No. Records Erect Bryozoans	No. Records Sea Squirts
2000	1	-	-	-
2001	-	-	-	-
2002	-	1	-	-
2003	-	0	-	-
2004	-	0	-	-
2005	1	15	-	-
2006	2	9	-	-
2007	17	29	-	6
2008	1	7	-	13
2009	3	16	-	9
2010	7	7	-	19
2011	13	43	71	24
2012	18	42	27	31
2013	25	82	82	23
2014	12	61	62	11
2015	24	82	68	22
2016	20	51	65	39
2017	13	63	35	27
2018	11	48	51	28
2019	5	48	35	28
2020	4	17	0	7
2021	9	42	7	15
2022	10	51	27	7
2023	7	48	31	7

Table A3. The Number of Raw Records with Taxon Name Provided for *Acanella arbuscula*, *Radicipes gracilis* and the Small Gorgonian Coral Functional Group by Year. Shading marks data not included in the analyses.

Year	***Acanella arbuscula***	***Radicipes gracilis***	**Small Gorgonian Coral Functional Group**
2002	1	0	1
2003	0	0	0
2004	0	0	0
2005	3	0	15
2006	8	0	9
2007	25	2	29
2008	7	0	7
2009	13	2	16
2010	7	0	7
2011	33	10	43
2012	39	3	42
2013	53	24	82
2014	31	28	61
2015	66	7	82
2016	38	10	51
2017	36	25	63
2018	37	10	48
2019	38	5	48
2020	15	1	17
2021	36	6	42
2022	38	10	51
2023	36	10	48

Table A4. Review of Small Gorgonian Coral Records from the EU where Identification was Uncertain. Justification for inclusion or exclusion of individual records is provided in the comments section.

Survey	**Set**	**Year**	**Start Latitude (DD)**	**Start Longitude (DD)**	**Biomass (kg) *Acanella***	**Biomass (kg) Anthothelidae**	**Biomass (kg) *Radicipes***	**Photo Taken**	**Delete**
CAFC11	072	2011	47.65	-44.89733333	0.012	0	0	No	Yes
CAFC14	182	2014	47.259	-46.11716667	0.001	0	0	Yes	No
CAFC17	096	2017	47.7695	-45.61483333	0.006	0	0	Yes	No
CAFC18	182	2018	46.87066667	-46.324	0	0.002	0	No	Yes
CAFC19	082	2019	47.60016667	-45.43533333	0	0	0.002	Not sure	Yes
CAFC21	027	2021	47.92016667	-44.57333333	0.003	0	0	No	Yes
FN3L16	045	2016	47.4295	-47.57216667	0	0.001	0	Yes	Yes
FN3L16	052	2016	47.17266667	-47.2665	0.001	0	0	Yes	Yes
FN3L17	045	2017	47.38016667	-47.31183333	0	0	0.005	Yes	Yes
FN3L19	038	2019	47.57283333	-47.55733333	0.001	0	0	Yes	Yes
PLA21	047	2021	43.10333333	-50.79716667	0.001	0	0	Yes	Yes

Comments:

CAFC11, Set 72. No photo. Depth from GEBCO 249 m. No records of *Acanella* in previous sets. **Deleted it based on lack of photo confirmation and unusual depth.**	
CAFC14, Set 182. Depth from GEBCO 357 m. Good state in photo. **Kept it.**	
CAFC17, Set 96. Depth from GEBCO 311 m. Good state in photo. **Kept it.**	
CAFC18, Set 182. Depth from GEBCO 334 m. No photo. **Unsure, but seems shallow for *Anthothela* and we cannot be sure is this species. We decided to remove it due the uncertainty.**	
CAFC19, Set 82. Depth from GEBCO 265 m. Small piece. No clear polyps. **Unsure, but based on the bad state and being the shallowest record of *Radicipes* we deleted it.**	
FN3L16, Set 45, depth 257 m. It can be contamination from a deeper set. However, the deeper sets were done one day before. **Due to the bad state of the sample and unusual depth we decided to remove it.**	

FN3L16, Set 52, depth 436 m. It can be contamination from a deeper set. Set 50 from same day at 1107 m recorded *A. arbuscula*. **Due to the bad state of the sample and unusual depth (for the area – seems to be shallower in FC) we decided to remove it.**	
FN3L17, Set 45, depth 277 m. It can be contamination from a deeper set. Set 40 from same day at 1079 m recorded *Radicipes*. What is in the photo with the sea anemone and that was recorded as *Radicipes* seems another thing. **We decided to remove it.**	
FN3L19, Set 38, depth 275 m. It can be contamination from previous set. Last set with presence of *A. arbuscula* was Set 30 (2 days before). Sometimes a small piece can be entangled in the net, but it seems too much time. **However, due to the bad state of the sample and unusual depth we decided to remove it.**	
PLA21, Set 47. Depth from GEBCO 102 m. Bad state. Also sand dollars and *Gersemia* in the sample. This does not look in a good state in photo. **Due to the bad state of the sample and unusual depth we decided to remove it.**	

Table A5. Review of Small Gorgonian Coral records from Canada where Identification was Uncertain. Justification for inclusion or exclusion of individual records is provided in the comments section. DD=decimal degrees.

Survey	Set	Trawl ID	Gear	Year	Start Latitude (DD)	Start Longitude (DD)	Biomass (kg) *Acanella*	Delete
NED2013438	055	NED2013438055	Campelen	2013	43.295	-51.21666667	0.01	Yes
NED2019515	030	NED2019515030	Campelen	2019	43.51	-51.37	0.0063	Yes
NED2013438	056	NED2013438056	Campelen	2013	43.46666667	-51.49333333	0.013	Yes

Comments:

All of these are good records with physical samples; however, they are still very shallow, and may have arisen through contamination from previous sets leaving an uncertainty associated with their spatial position.

NED2013438, Set 55. Depth from GEBCO 112 m. 0.06 kg of *Acanella* recorded in Set 54 at 594 m. We think this could be contamination as it is very unusual depth. We decided to remove it.

NED2013438, Set 56. Depth from GEBCO 100 m. 0.06 kg of *Acanella* recorded in Set 54 at 594 m. We think this could be contamination as it is very unusual depth. We decided to remove it.

NED2019515, Set 30. Depth from GEBCO 82 m. Not *Acanella* recorded previously, but due to the shallow depth we are dubious about it. The two previous Sets 28 and 29 were done in Canadian waters at 579 m and 429 m, respectively. It could come from there. We decided to remove it.

Table A6. Review of Sea Squirt (*Boltenia ovifera*) records from the EU and Canada where identification was uncertain. Justification for inclusion or exclusion of individual records is provided in the comments section.

Country	Survey	Set	Year	Start Latitude (DD)	Start Longitude (DD)	Biomass (Kg)	Delete
Canada	TEL2010979	052	2010	46.89167	-47.1117	0.01	Yes
Canada	TEL2016170	058	2016	46.39167	-46.955	0.47	Yes
Canada	NED2014447	034	2014	46.31	-47.2333	0.09	No
EU Spain	PLA18	076	2018	43.6645	-48.9973	0.006	Yes
EU Spain	PLA18	055	2018	43.0755	-49.4493	0.016	Yes
EU Spain	PLA17	092	2017	44.4025	-48.9488	0.005	No

Comments:

TEL2010979, Set 52. Depth from GEBCO 1057 m. No tunicates/ascidians/sea potatoes documented on deck sheet, but e-data shows 0.01 (n=1); **delete it.**	
TEL2016170, Set 58. Depth from GEBCO 806 m. "tunicate" subsample 0.047 kg with 10% bumped up weight for Total Catch =0.47 kg. This is deep for *Boltenia*, and most likely another species. Due to uncertainty, **delete it.**	
NED2014447, Set 34. Depth from GEBCO 546 m. 3 *Boltenia* recorded. As no clear evidence and we are keeping one at 687 m, we should keep this as well. **Keep it.**	
PLA18, Set 76. Depth from GEBCO 845 m. Ok state in photo. 6.74 kg of *Boltenia* recorded in set 73, likely contamination. **Delete it.**	
PLA18, Set 55. Depth from GEBCO 881 m. Bad state in photo. *Boltenia* recorded in Set 51 (0.036 kg), likely contamination. **Delete it.**	
PLA17, Set 92. Depth from GEBCO 687 m. Ok state in photo. The closest *Boltenia* recorded was in Set 83 (0.048 kg) and 2.73 kg in Set 81. **Keep it.**