Ensemble habitat suitability modeling of vulnerable marine ecosystem indicator taxa to inform deep-sea fisheries management in the South Pacific Ocean

Samuel E. Georgian[a,*], Owen F. Anderson[b], Ashley A. Rowden[b]

[a] Marine Conservation Institute, 4010 Stone Way N, Suite 2010, Seattle, WA, USA
[b] National Institute of Water & Atmospheric Research, 301 Evans Bay Parade, Wellington, New Zealand

ABSTRACT

Resolutions of the United Nations General Assembly (UNGA) require states and competent authorities to protect vulnerable marine ecosystems (VMEs), ecologically important habitats in the deep sea that are considered to be especially at risk from anthropogenic disturbances such as fishing. The lack of data concerning the location and extent of VMEs poses a significant obstacle to their protection. Habitat suitability modeling is increasingly used in spatial management planning due to its ability to predict the distribution and niche of marine organisms based on limited input data. We generated broad-scale, medium-resolution (1 km^2) ensemble models for ten VME indicator taxa within the New Zealand Exclusive Economic Zone and a portion of the South Pacific Regional Fishery Management Organisation (SPRFMO) convention area. Ensemble models were constructed using a weighted average of three habitat suitability model types: Boosted Regression Trees, Maximum Entropy, and Random Forest. All models performed well, with area under the curve scores above 0.9, and ensemble models marginally outperformed any of the individual modeling approaches. Highly suitable habitat for each VME indicator taxa was predicted to occur in relatively small areas throughout the region, typically associated with elevated seafloor features with steep slopes. Determining the spatial distribution of VME indicator taxa is critical for assessing the current and historical extent of bottom trawling impacts on benthic communities, and for supporting the improved spatial management of fisheries in the South Pacific Ocean. Given the additional threats of climate change and ocean acidification to VME indicator taxa throughout the deep sea, habitat suitability modeling is likely to play an increasing role in designing effective, long-term protection measures for cumulative impacts on VMEs.

1. Introduction

The demand for resources and advances in technology have greatly accelerated the rate of anthropogenic disturbances to even the remotest of deep-sea environments (reviewed in Ramirez-Llodra et al., 2011). Historically, the largest documented disturbances to deep-sea communities of the seafloor have occurred as the result of fishery activities (reviewed by Clark et al., 2016). The bulk of impacts have stemmed from bottom trawling practices (Puig et al., 2012; Pusceddu et al., 2014), which can cause extensive physical damage to benthic environments. However, long-line fishing gear has also been observed to disturb deep-sea habitats (Orejas et al., 2009). Bottom trawls were found to have severely damaged 30–50% of cold-water coral reef structures in some regions (Fosså et al., 2002), and have also been shown to considerably reduce the richness, diversity, and abundance of

deep-sea soft sediment communities (Cryer et al., 2002). In addition to direct physical damage, bottom trawling resuspends sediments that can smother benthic filter feeding organisms even outside of the trawled area (Palanques et al., 2001).

In areas beyond national jurisdictions (the high seas) in the South Pacific Ocean, bottom trawl fisheries targeting deep-sea species such as orange roughy (*Hoplostethus atlanticus*), cardinalfish (*Epigonus telescopus*), black oreo (*Allocyttus niger*), and alfonsino (*Beryx* spp.) occur in several broad areas. The South Pacific Regional Fisheries Management Organisation (SPRFMO) is an intergovernmental agency established by nations obligated under United Nations General Assembly (UNGA) resolutions to ensure the sustainability of high seas fisheries and to safeguard marine ecosystems from adverse impacts. To meet these two objectives, management approaches have been designed to protect vulnerable marine ecosystems (VMEs) while simultaneously ensuring

---

fishery access to areas suitable for sustainable fishing. VMEs are defined by their susceptibility to disturbance based on the fragility, functional significance, rarity, and life history traits of their component communities, populations, or species (FAO, 2009). Several VME indicator taxa have been identified for the South Pacific Ocean using these guidelines (Parker et al., 2009). The United Nations require that states and associated regional fisheries management organizations put in place adequate measures to prevent significant adverse impacts on VMEs, which can include closing areas to fishing (UNGA, 2007). In the SPRFMO convention area, two main interim management measures have been implemented for New Zealand- and Australian-flagged vessels to protect VMEs: a move-on rule that requires vessels to cease bottom trawling within 5 nautical miles of an encountered VME, and a network of large (20 min latitude by 20 min longitude) closures in previously unfished or low-impacted areas (Penney et al., 2009). However, the Australian and New Zealand implementations of the interim measures were different and not always complementary, and the efficacy of these measures has been questioned (Penney and Guinotte, 2013). The SPRFMO Commission has, therefore, indicated a desire for the interim measures to be replaced by a consistent and improved measure as soon as possible, principally through the design and implementation of more effective spatial closures. The design of spatial management measures requires information about the distribution or likely distribution of VME indicator taxa in the South Pacific Ocean.

Habitat suitability modeling (also known as species distribution modeling) is increasingly used to predict the niche and distribution of species based on limited data. This statistical approach relates field observations to environmental predictor variables, yielding predictions of the abundance or probability of presence of taxa and the underlying environmental drivers of their geographic distribution. Habitat suitability models are particularly useful for characterizing VMEs in the deep sea, where field surveys are logistically difficult and expensive and it is not feasible to directly observe the entire area of interest (Vierod et al., 2014). In this study, we build on previous habitat suitability modeling work (Anderson et al., 2016a, 2016b; Rowden et al., 2017), using an ensemble habitat suitability modeling approach to predict the niche and distribution of ten VME indicator taxa over an area that covers the fishing interests of both New Zealand and Australia in the western South Pacific. The model outputs were designed for use in a multi-country and multi-stakeholder spatial management planning process to design and implement new spatial closures in the region, for eventual adoption by SPRFMO.

2. Materials and methods

2.1. Site description

The study area extends over a 21.6 million km^2 area of the South Pacific Ocean from approximately 24 °S to 58 °S and 143 °E to 146 °W. This area encompasses the New Zealand EEZ, and a large portion of the SPRFMO convention area that represents the main fishing areas of New Zealand and Australian fleets (Fig. 1). We set the spatial resolution of the models to 1 km^2 (see Table 1) to match that of the main sampling methods for the species presence records; trawls, sleds and dredges – which had a median tow length of about 0.8 km (Anderson et al., 2016a).

2.2. Occurrence sampling

Ten VME indicator taxa were selected for modeling. These included two sponge taxa; glass sponges (Hexactinellida) and demosponges (Demospongiae), and four species of framework-building scleractinian corals; *Enallopsammia rostrata*, *Goniocorella dumosa*, *Madrepora oculata*, and *Solenosmilia variabilis*. Soft corals (Alcyonacea), sea pens (Pennatulacea), and black corals (Antipatharia) were modeled as larger taxonomic groups rather than at the species level. Hydrozoans in the

family Stylasteridae, commonly known as hydrocorals, were modeled as a single taxonomic group. These taxa have all been identified as VME indicator taxa based on their vulnerability to trawls (fragility, rareness, or low productivity), ecological significance (functionally significant via habitat creation or ecological services), and ability to be readily identified from trawl collections (Parker et al., 2009). Presence data for the ten modeled taxa were obtained from New Zealand and Australian museum/collection records, fisheries research databases, and online biodiversity databases. Over 24,000 individual records were compiled, but these were unequally distributed among the taxa with numbers ranging from 203 (*Madrepora oculata*) to 5345 (Alcyonacea) (Table 2). Duplicate occurrences were removed prior to further analyses.

2.3. Background sampling

For the background (or pseudo-absence) data locations used in the models, we used a combination of the concepts of random selection and target-group background sampling (Phillips et al., 2009) to generate a spatially structured set of background points. While commonly not accounted for in habitat suitability models, spatial bias in the sampling of occurrence data considerably weakens the performance and interpretability of models (Phillips et al., 2009). However, it is possible to reduce the effects of sampling bias by selecting or creating a set of background data that reflect the same bias as the occurrence data (Phillips et al., 2009). We were unable to construct target-group background data directly (i.e., from the locations of a broad set of species presence records based on similar observation methods) due to the wide range of data sources supplying the presence data. Given the relatively high sampling bias observable in the observed presences, we generated a structured sample of background points for each taxon, based on the distribution of presence records. To preferentially generate background points in areas with more presence observations, we first created a two-dimensional kernel density estimate of sampling effort based on the presence locations for each taxon (e.g., Supplemental Fig. S1a). This estimate created a probability grid from which 10,000 background points were sampled according to the probability grid weights (e.g., Supplemental Fig. S1b). Previous studies using a similar approach to reduce the influence of sampling bias found that model performance was significantly improved (Elith et al., 2010; Huang et al., 2011; Fitzpatrick et al., 2013).

2.4. Environmental variables

A suite of environmental variables (Table 1) that may potentially influence marine benthic organism distributions were generated using a similar approach to that described by Davies and Guinotte (2011), based on depth data derived from a recently updated bathymetry grid for the New Zealand region, which we gridded at a resolution of 1 km^2 (Mackay et al., 2015). Additional seafloor terrain metrics were also derived from this bathymetry using the Benthic Terrain Modeler in ArcGIS 10.3.1.1 (Wright et al., 2012). Locations were defined as being on a seamount if they were within the footprint of the deepest complete depth contour around any of the seamounts recorded in the NIWA Seamounts database (Rowden et al., 2008) or identified by Yesson et al. (2011). The percent mud and percent gravel layers for the region were developed from > 30,000 raw sediment sample data compiled in dbseabed (Jenkins, 1997), which were then imported into ArcGIS and interpolated using Inverse Distance Weighting.

Dissolved oxygen (Garcia et al., 2014b), apparent oxygen utilization (Garcia et al., 2014b), percent oxygen saturation (Garcia et al., 2014b), temperature (Locarnini et al., 2013), salinity (Zweng et al., 2013), nitrate (Garcia et al., 2014a), phosphate (Garcia et al., 2014a), and silicate (Garcia et al., 2014a) data were obtained from the World Ocean Atlas (version 2, 2013). The saturation states of the aragonite and calcite polymorphs of calcium carbonate were obtained from Bostock et al. (2013). Oxygen and nutrient data were originally available at a
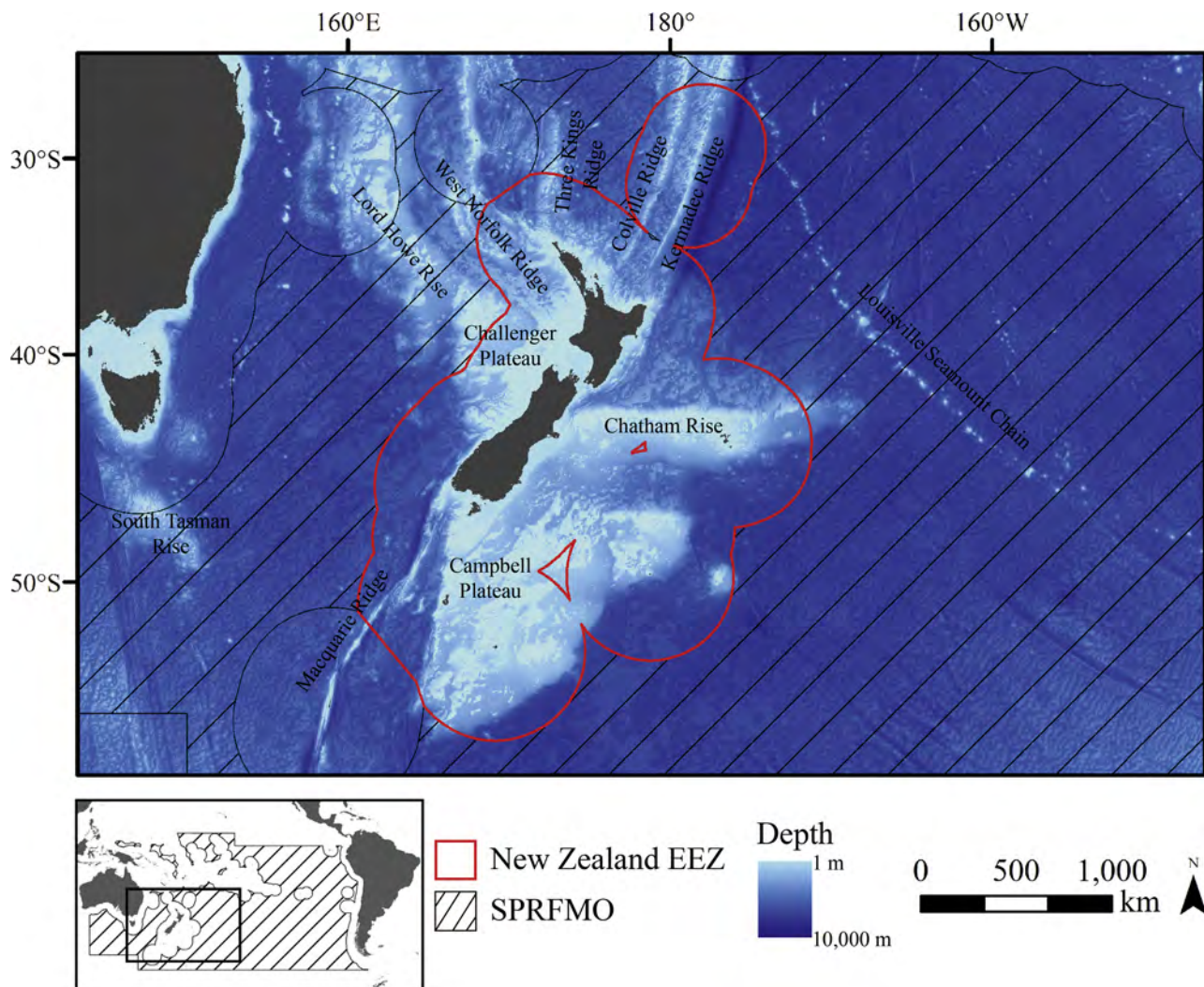
Fig. 1. Map of the study area. The red line indicates the extent of the New Zealand Exclusive Economic Zone (EEZ). Hatched lines indicate the extent of the South Pacific Regional Fisheries Management Organisation (SPRFMO) convention area. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

resolution of 1°. Salinity and temperature were available at a resolution of 0.25°. Carbonate data were available at a resolution of 0.5°. These variables were obtained as depth-gridded data ranging from 0 to 5,500 m, and were transformed to match the resolution of the bathymetry data (1 km^2) using an upscaling approach that approximates conditions at the seafloor (*sensu* Davies and Guinotte, 2011).

Particulate organic carbon (POC) flux at the seafloor was obtained from Lutz et al. (2007) at a native resolution of 0.08°. Surface productivity data were obtained from the Oregon State Ocean Productivity site (http://www.science.oregonstate.edu/ocean.productivity) as the Standard Vertically Generalized Production Model (VGPM; Behrenfeld and Falkowski, 1997), the Eppley-VGPM (Eppley, 1972), and the updated Carbon-based Production Model (CbPM, Westberry et al., 2008). All products were derived from NOAA's Aqua MODIS data ranging from mid-2002 through 2016, and calculated as the mean, minimum, maximum, and standard deviation over that period (Aqua MODIS, 2014). These layers had a native resolution of 0.167°, and were interpolated to a slightly higher resolution (0.05°) through kriging and then resampled to match the extent and resolution of the bathymetry data (1 km^2).

2.5. Variable selection

Our aim in selecting variables was to identify a base set of predictors

for use in each model, modified by specific biological requirements of some taxa, and among which correlations were minimized. The inclusion of correlated environmental variables may inhibit model performance and interpretation (e.g., Huang et al., 2011). Therefore, highly correlated variables were removed based on their relationship with other variables and on their performance in preliminary random forest models. First, all variables were placed into ecologically relevant groupings (see Table 1): seafloor characteristics (terrain metrics and sediment type), water chemistry (nutrients, carbonate chemistry, temperature, salinity, sigma theta, and oxygen metrics), and productivity (POC export and surface productivity). Then a random forest model was constructed for each taxon using the variables in each grouping, and variable performance was assessed as the mean decrease in accuracy (% Inc. MSE) scaled to 100% (Supplemental Fig. S2). Among highly correlated (> 0.7 Pearson's Correlation; Supplemental Fig. S3) variables that clustered together (Supplemental Fig. S4), the variable with the highest performance across all taxa was retained. If performance was equivalent (i.e., within approximately 10%) within a group of highly correlated variables, we also took into account the clustering and correlation of each variable with the remaining variable set, as well as biological relevance. Equivalent variables were preferentially retained if they clustered separately and had lower correlations with the remaining variable set. Depth was removed prior to these preliminary

Table 1
Environmental variables considered for use in habitat suitability models.

Variable	Units	Native Resolution	Reference
Seafloor Characteristics	meters	1 km^2	Mackay et al. (2015)
Depth	%	–	NIWA
Percent gravel	%	–	NIWA
Percent mud	–	–	Derived from bathymetry
Ruggedness [2]	degrees	–	Derived from bathymetry
Slope[2]	–	–	Derived from bathymetry
Slope SD[2]	degrees	–	Derived from bathymetry
Aspect	–	–	Derived from bathymetry
Range[2]	–	–	Derived from bathymetry
Standard deviation[2]	–	–	Derived from bathymetry
Profile curvature	–	–	Derived from bathymetry
Plan curvature	–	–	Derived from bathymetry
Curvature	–	–	Derived from bathymetry
Bathymetric Position Index – fine	–	–	Derived from bathymetry
Bathymetric Position Index – broad	–	–	Rowden et al. (2008), Yesson et al. (2011)
Seamounts			
Water Chemistry			
Apparent oxygen utilization	ml l^{-1}	1°	Garcia et al. (2014b)
Aragonite saturation state	–	0.5°	Bostock et al. (2013)
Calcite saturation state	–	0.5°	Bostock et al. (2013)
Dissolved oxygen	ml l^{-1}	1°	Garcia et al. (2014b)
Nitrate	μmol l^{-1}	1°	Garcia et al. (2014a)
Percent oxygen saturation	%	1°	Garcia et al. (2014b)
Phosphate	μmol l^{-1}	1°	Garcia et al. (2014a)
Salinity	–	0.25°	Zweng et al. (2013)
Sigma theta	kg m^{-3}	0.25°	Derived from temperature and depth
Silicate	μmol l^{-1}	1°	Garcia et al. (2014a)
Temperature	°C	0.25°	Locarnini et al. (2013)
Productivity			
Particulate organic carbon export	mg C m^{-2} d^{-1}	0.08°	Lutz et al. (2007)
Vertically Generalized Production Model[1]	mg C m^{-2} d^{-1}	0.167°	Oregon State University[3]
Eppley-VGPM[1]	mg C m^{-2} d^{-1}	0.167°	Oregon State University[3]
Carbon-Based Productivity Model-2[1]	mg C m^{-2} d^{-1}	0.167°	Oregon State University[3]

1-Surface data derived from MODIS–Aqua (NASA) as the mean, minimum, maximum, and standard deviation from mid-2002–2016.
2-Terrain metrics calculated using window sizes of 3, 5, 7, and 15 cells.
3-Data obtained from http://www.science.oregonstate.edu/ocean.productivity.

Table 2
Taxonomic groups modeled, sample size (number of presences), and the suite of environmental variables used to train models for each taxon. BPI Broad = bathymetric position index calculated at a broad scale. POC = particulate organic carbon flux to the seafloor. RAC = residual autocovariate. Slope SD = standard deviation of slope. See Table 1 and text for details about each variable.

Taxon	Group	Sample Size	Variable Set
Alcyonacea	Soft corals	5345	Calcite, BPI-Broad, Dissolved Oxygen, % Gravel, % Mud, POC, RAC, Ruggedness, Seamounts, Slope SD, Temperature
Antipatharia	Black corals	1,505	Calcite, BPI-Broad, Dissolved Oxygen, % Gravel, % Mud, POC, RAC, Ruggedness, Seamounts, Slope SD, Temperature
Demospongiae	Demosponges	3,166	Silicate, BPI-Broad, Dissolved Oxygen, % Gravel, % Mud, POC, RAC, Ruggedness, Seamounts, Slope SD, Temperature
E. rostrata	Stony corals	236	Aragonite, BPI-Broad, Dissolved Oxygen, % Gravel, % Mud, POC, RAC, Ruggedness, Seamounts, Slope SD, Temperature
G. dumosa	Stony corals	464	Aragonite, BPI-Broad, Dissolved Oxygen, % Gravel, % Mud, POC, RAC, Ruggedness, Seamounts, Slope SD, Temperature
Hexactinellida	Glass sponges	1,916	Silicate, BPI-Broad, Dissolved Oxygen, % Gravel, % Mud, POC, RAC, Ruggedness, Seamounts, Slope SD, Temperature
M. oculata	Stony corals	203	Aragonite, BPI-Broad, Dissolved Oxygen, % Gravel, % Mud, POC, RAC, Ruggedness, Seamounts, Slope SD, Temperature
Pennatulacea	Sea pens	695	Aragonite, BPI-Broad, Dissolved Oxygen, % Gravel, % Mud, POC, RAC, Ruggedness, Seamounts, Slope SD, Temperature
S. variabilis	Stony corals	463	Aragonite, BPI-Broad, Dissolved Oxygen, % Gravel, % Mud, POC, RAC, Ruggedness, Seamounts, Slope SD, Temperature
Stylasteridae	Hydrocorals	1,481	Aragonite, BPI-Broad, Dissolved Oxygen, % Gravel, % Mud, POC, RAC, Ruggedness, Seamounts, Slope SD, Temperature

model runs due to its lack of biological relevance apart from strong correlations with other variables. Aragonite saturation state was only included for taxa that incorporate calcium carbonate in the form of aragonite into skeletal materials (i.e., Stylasteridae, *Enallopsammia rostrata, Goniocorella dumosa, Madrepora oculata,* Pennatulacea, and *Solenosmilia variabilis*), and calcite saturation state was only included for taxa that incorporate calcite (Alcyonacea and Antipatharia). Given the strong correlation between aragonite and calcite (Pearson's correlation, r > 0.999, p < 0.001), these variables can be expected to perform identically in models, but were distinguished among taxa to make biological interpretations easier. Silicate was included for Hexactinellida and Demospongiae because these taxa incorporate silicate into their skeletons, but was omitted for all other taxa. Temperature

was included in all models due to its biological relevance and high performance in previous cold-water coral modeling studies (e.g., Davies and Guinotte, 2011). The correlations and clustering of variables in the final retained set are shown in Fig. 2 (see Supplemental Table S1 for Pearson's correlation coefficient values). Supplemental Fig. S5 shows the relationship of occurrence points, targeted background points, and random background points with each environmental variable in the final set.

2.6. Boosted regression tree models

Boosted Regression Tree (BRT) is a presence-absence (as opposed to presence-background) data modeling approach, that can also be used to
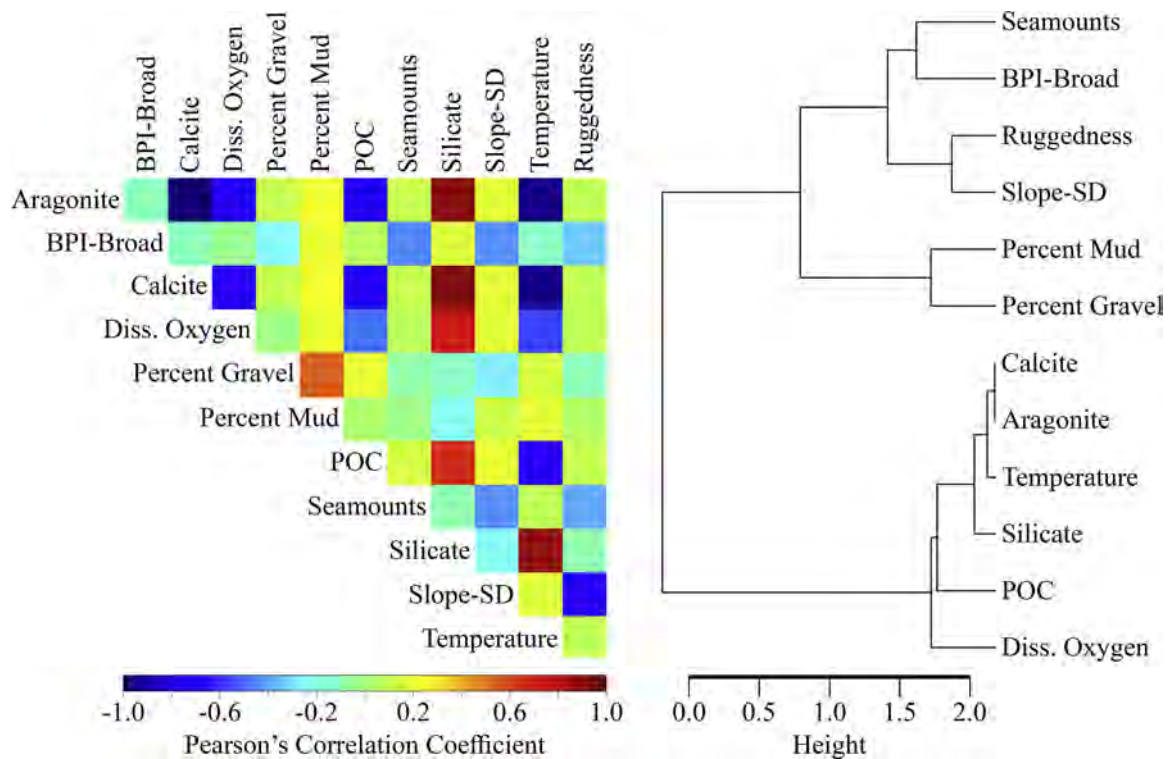
Fig. 2. Relationships among environmental variables used to train the habitat suitability models. A) Pearson's correlation coefficients among the final variable set. B) Cluster dendrogram showing the conceptual relationships among the final variable set. Not every variable was included in the model for each taxon (see Table 2). See Supplemental Table S1 for actual Pearson's correlation coefficient values.

model abundance, and is often used to predict distributions of deep-sea taxa (e.g., Rooper et al., 2017). The model calculates habitat suitability probabilities by incorporating recursive binary splits within a regression tree structure to explore the relationship between response and predictor variables (Elith et al., 2008). We fitted BRT models using a standard approach, including optimization of the learning rate and number of trees by internal cross validation (Elith et al., 2008), and adjustment of the learning rate as necessary to ensure a minimum number of 1000 trees. Tree-complexity was set to three to allow a moderate level of interactions between terms. A random selection of absence records, equal in number to the number of presence records, was selected from the sampling bias-corrected set of 10,000 background points generated for each taxon, following the recommendations of Barbet-Massin et al. (2012). The relative influence of each variable was determined from the number of times it was selected for splitting, weighted by the squared improvement to the model as a result of each split (using in-bag data) (Friedman and Meulman, 2003).

2.7. Maximum entropy models

The Maximum Entropy (Maxent) modeling algorithm (Phillips et al., 2006) is a machine learning, presence-only model that is based on the statistical assumption that within known environmental constraints, maximizing entropy is the optimal method to address remaining uncertainty in a distribution. Maxent has been shown to outperform other presence-only models (Elith et al., 2006; Tittensor et al., 2009) and presence-absence models (Reiss et al., 2011). All Maxent models were run in an R environment using the 'dismo' package (Hijmans et al., 2015). Models were built using default Maxent settings (convergent threshold = 10^{-5}, prevalence = 0.5, regularization multiplier = 1) that have performed well in previous studies (see Phillips and Dudik, 2008). However, the number of maximum iterations was increased to 5000 to ensure model convergence, as preliminary models did not always converge within the default 500 iterations. In addition, instead of

generating random background points, models were trained using the sampling bias-corrected set of 10,000 background points. The standard jackknifing procedure was used to determine the percent contribution of environmental variables to the model.

2.8. Random forest models

Random forest (RF) modeling (Breiman, 2001) is a classification and regression tree-based approach similar to boosted regression trees. However, it chooses a random selection of explanatory values at each split and employs an internal bootstrapping approach to produce and combine a large number of individual trees. Models were built using the 'randomForest' package (Liaw and Wiener, 2002) in R. In most cases, the error rate stabilized after approximately 100 trees, however 501 trees were generated for each taxon to ensure stabilization in all model runs. Optimal values of 'mtry', which determines the number of explanatory variables selected at each tree split, were explored using the 'tuneRF' function and preliminary models. Ultimately, the default value (mtry = 3, the square root of the number of explanatory variables) was used due to its consistent performance across all taxa. Classification trees are sensitive to imbalances in the response variable (Chen et al., 2004; Evans et al., 2011) so, as in the BRT models, the number of background points was set equal to the number of occurrences for each taxon, randomly selected from the 10,000 set of sampling bias-corrected background points. This sampling approach has been found to optimize model performance for both BRT and RF models (Barbet-Massin et al., 2012).

2.9. Spatial autocorrelation

Spatial autocorrelation occurs when there is a relationship between observations and geographic distance (Legendre, 1993). Environmental variables are frequently spatially autocorrelated (Lichstein et al., 2002), and observed species distributions often exhibit strong spatial

autocorrelation due to sampling bias or natural processes such as dispersal limitations (Lennon, 2000). The majority of spatial analyses do not account for spatial autocorrelation (Dormann, 2007) despite the fact that it typically violates the underlying assumptions of many models (Kühn, 2007) and can considerably affect model performance and the importance of explanatory variables (Tognelli and Kelt, 2004).

To account for spatial autocorrelation, we calculated a residual autocovariate (RAC) variable for each taxon (*sensu* Crase et al., 2012). The RAC method derives a term representing the spatial autocorrelation present in the residuals of a preliminary model. This term is then added as an explanatory variable in subsequent models, accounting for spatial autocorrelation present in the initial occurrence data. The RAC approach has been used to correct for spatial autocorrelation in classification models including boosted regression trees and random forest (Crase et al., 2012; Rowden et al., 2017) as well as in Maxent (Václavík et al., 2012).

To calculate the RAC variable, preliminary models were built using environmental variables alone, and the model residuals were tested for spatial autocorrelation using Moran's I (Moran, 1948). Model residuals were rasterized at the same resolution of the prediction grids (1 km²). Next, the autocovariate was calculated based on the focal mean of these residuals, in which the value for each output cell was the average of the input cells within the specified window size. The resulting RAC was then included in models as an explanatory variable. In Maxent models, RAC was calculated for *E. rostrata*, Hexactinellida, Pennatulacea, and *S. variabilis* using a 5 × 5 focal mean window (indicating the number of grid cells used to calculate the focal mean) rather than the default 3 × 3 window (Crase et al., 2012), because initial tests revealed highly inflated importance (∼80–90%) of the RAC variable in these models. Increasing the size of the focal mean window reduced the importance of RAC while still resulting in the reduction of spatial autocorrelation among model residuals. The same was true for BRT models, where a 3 × 3 window was used for Antipatharia, Hexactinellida, and Pennatulacea, a 7 × 7 window was used for *E. rostrata* and *M. oculata*, and a 5 × 5 window was used for the remaining taxa. For random forest models, a 3 × 3 focal window was used for all taxa, as the highest percent contribution of RAC was relatively low (∼30%) without adjusting the window size. In BRT and RF models, the inclusion of the RAC variable considerably reduced spatial autocorrelation in model residuals as measured by Moran's I. In Maxent models, spatial autocorrelation was originally lower and the RAC variable reduced spatial autocorrelation to a lesser extent than in BRT and RF models, consistent with previous studies showing that Maxent is naturally robust to underlying spatial structures in the input data (De Marco et al., 2008; Václavík et al., 2012).

2.10. Model performance

The BRT, Maxent, and RF models were verified using a ten-fold cross-validation procedure in which input data were randomly partitioned into training (70%) and testing datasets (30%) for ten model runs. The area under the curve (AUC) metric was calculated for each model and averaged across all ten runs. AUC is a threshold-independent assessment of the ability of the model to correctly rank occurrences above background locations. A random model has a theoretical AUC of 0.5, while AUC values greater than 0.7 indicate adequate performance, and values greater than 0.8 indicate excellent performance (Wiley et al., 2003; Hosmer et al., 2013). Final models were constructed using all available data. A separate AUC score was calculated for each final model using the entire training dataset to facilitate direct comparisons to the ensemble models (see below). As an additional measure to compare each model to the ensemble model, we also calculated the correlation between predicted habitat suitability and expected values based on the entire training dataset.

To ascertain the spatial variation in model uncertainty, we employed a bootstrap approach to generate coefficient of variation

estimates for every cell within the study area (*sensu* Anderson et al., 2016b). For each modeling type, 200 models were run by randomly sampling occurrences and background points with replacement equal to the original number of occurrences and background points. The mean and standard deviation of each cell was calculated from this set of 200 models, and the coefficient of variation was calculated as the mean divided by standard deviation, reported as a percentage. This approach yields a spatial estimate of uncertainty based on how sensitive the models are to changes in the input presence and background data.

2.11. Ensemble modeling

An ensemble of models with fundamentally different approaches to the same problem is a practical way of avoiding dependence on a single model type or structural assumption, and enables a more robust characterization of the predicted spatial variation and uncertainties (Robert et al., 2016). To incorporate the predictions and underlying assumptions of BRT, Maxent, and RF into a single output grid, we produced ensemble models by calculating weighted averages of the three models (after Oppel et al., 2012; Anderson et al., 2016b; Rowden et al., 2017) as follows:

$$W_B = \frac{AUC_B}{AUC_B + AUC_M + AUC_R}, \; W_M = \frac{AUC_M}{AUC_B + AUC_M + AUC_R}, \text{ and } W_R$$
$$= \frac{AUC_R}{AUC_B + AUC_M + AUC_R}$$

$$X_E = X_B * W_B + X_M * W_M + X_R * W_R$$

$$CV_E = \sqrt{\frac{(CV_B * X_B)^2 * W_B^2 + (CV_M * X_M)^2 * W_M^2 + (CV_R + X_R)^2 * W_R^2}{X_E^2}}$$

where AUC_B, AUC_M, and AUC_R are the model performance statistics (average cross-validated AUC values); X_B, X_M, X_R, and X_E the model predictions; and CV_B, CV_M, CV_R, and CV_E the bootstrap CVs from the BRT, Maxent, RF, and ensemble models respectively. The performance of ensemble models was calculated as AUC scores calculated using the entire training dataset rather than through a cross-validation procedure using test data, and as the correlation between expected (observed presence or assumed absence) and predicted values.

2.12. Model outputs

Model outputs were limited to depths of 300 to 3000 m, approximately equivalent to the combined bathymetric range of all modeled taxa, covering a reduced area of 4.5 million km². For each modeling type (BRT, Maxent, RF, and ensemble), habitat suitability scores were calculated across the study area at a resolution of 1 km². In the context of our modeling approach, these values represent the relative predicted suitability for each taxon ranging from 0 to 1, with higher values indicating a more suitable habitat. While these scores are commonly and at times erroneously referred to as the 'probability of occurrence' in the literature, it is important to note that our habitat suitability scores do not represent the actual probability of occurrence due to the lack of true absence data and unknown efficiency of the sampling gear. In the following text, we use the following arbitrary divisions for interpreting habitat suitability scores: < 0.4 = low suitability, 0.4–0.8 = moderate suitability, and > 0.8 = high suitability.

3. Results

3.1. Final variable set

Of the 30 variables considered for inclusion in models (Table 1), 13 were ultimately retained by the variable selection process (Fig. 2; Table 2). Nine of these variables: broad-scale bathymetric position index (BPI), dissolved oxygen, percent gravel, percent mud, particulate

organic carbon flux to the seafloor (POC), occurrence on or off seamounts, standard deviation of slope (slope-SD) calculated with a window size of 15 × 15, temperature, and vector ruggedness measure (ruggedness) calculated with a window size of 7 × 7 were included in all models. The saturation states of aragonite and calcite were included only for coral taxa that incorporate one of these calcium carbonate polymorphs, and silicate was included only for sponge taxa that incorporate silicate into their skeleton (Table 2). In addition, the RAC variable, although not ecologically interpretable, was included in all models to reduce the spatial autocorrelation of model residuals. Therefore, each model was built using a total of 11 variables. The final set of variables generally had lower correlations than the initial variable set (see Supplemental Figs S3 and S4), however some of the retained variables still had correlations greater than 0.7 (Fig. 2; Supplemental Table S1). Aragonite, calcite, and silicate were all highly correlated. However, only one of these variables was included in each model in accordance with its biological relevance for each taxon. Temperature was highly correlated with aragonite (Pearson's correlation, r = 0.960, p < 0.001), calcite (Pearson's correlation, r = 0.959, p < 0.001), and silicate (Pearson's correlation, r = -0.876, p < 0.001), but was retained in the final set due to its known biological relevance. Dissolved oxygen and POC were also correlated with other variables including aragonite, calcite, silicate, and temperature, but were retained because the correlations were lower (< 0.8), and because they not grouped together in the cluster analysis (Fig. 2).

3.2. Model performance

Cross-validation testing suggested that all models for all taxa had excellent performance, with AUC scores above 0.9. However, R^2 values were more variable, ranging from 0.14 to 0.64 for the ensemble models (Table 3). Unlike AUC, R^2 is dependent on scale and will tend to be lower where probabilities are close to zero. Accordingly, R^2 values were higher for taxa with more presence records (e.g., Alcyonacea and Antipatharia) than for taxa with fewer presences (e.g. *M. oculata* and *G. dumosa*). AUC scores for BRT models ranged from 0.961 to 0.998, with a mean of 0.978 ± 0.014 across all taxa. Maxent model AUC scores were consistently the lowest of the three modeling approaches, ranging from 0.913 to 0.990 with a mean of 0.964 ± 0.026 across all taxa. RF model AUC scores ranged from 0.951 to 0.993 with a mean of 0.976 ± 0.014 across all taxa. The worst performing model for an individual taxon across the three modeling approaches was Pennatulacea, with a minimum AUC of 0.913 and a mean of 0.948 ± 0.031. Models for *S. variabilis* consistently performed the best, with a mean AUC of 0.994 ± 0.004 across the three modeling approaches. The sample size (number of presences) for each taxon (Table 2) was not significantly correlated with AUC (Linear regression, R^2 = 0.03, p = 0.33).

Table 3
Performance metrics for Boosted Regression Trees (BRT), Maximum Entropy (Maxent), and Random Forest (RF) for each taxon, assessed by the area under the curve (AUC) score measured using ten-fold cross-validation on a 30% test dataset withheld from model creation. Larger AUC scores (closer to 1) indicate better model performance.

Taxa	BRT	Maxent	RF
Alcyonacea	0.967	0.958	0.970
Antipatharia	0.992	0.984	0.987
Demospongiae	0.974	0.966	0.964
E. rostrata	0.997	0.983	0.990
G. dumosa	0.961	0.944	0.973
Hexactinellida	0.965	0.939	0.951
M. oculata	0.975	0.983	0.989
Pennatulacea	0.971	0.913	0.960
S. variabilis	0.998	0.990	0.993
Stylasteridae	0.984	0.984	0.981

3.3. Variable contributions

The BRT and RF models had similar percent contributions of predictor variables, while the Maxent models typically had higher percent contributions from variables associated with skeletal chemistry (i.e., calcite, aragonite, and silicate) (Table 4). Variables describing physical characteristics of the seafloor, the saturation state of aragonite or calcite, and RAC were typically the most important variables. While seafloor characteristics (BPI, seamount, slope SD, and ruggedness) generally did not contribute a large percentage of information to the models on their own, their combined contribution when considered together was substantial, ranging between 8.5% and 51.8%. Ruggedness was frequently the best predictor in this group, with percent contributions as high as 36.3%, followed by BPI with a maximum contribution of 23.7%. Taxa were generally located in areas with more complex, steep, and elevated terrain features than expected by chance (Supplemental Figs. S5, S11–S20). When the saturation state of calcite and aragonite where included in models, they typically had a large percent contribution, with species demonstrating a preference for higher saturation states (Supplemental Figs. S11–S20). Calcite contributed an average of 28.2% and 21.0% in models for Alcyonacea and Antipatharia, respectively, across all model types. The average percent contribution of aragonite, when included, ranged from 6.7 to 17.7% across all model types, and contributed an overall average of 12.0% across all taxa and model types. Models for the aragonite-forming taxa that had a relatively low percent contribution from aragonite often had an elevated percent contribution from temperature. Aragonite and temperature were highly correlated across the study region (Pearson's correlation, r = 0.96, p < 0.001; Fig. 2; Supplemental Table S1), making it difficult to tease apart their influence in models. Temperature contributed 4.0–18.3% of total explanatory power when averaged across all model types, and was most important for *G. dumosa*, Pennatulacea, and *S. variabilis*. Most taxa occurred predominantly within a narrower temperature range than expected by chance based on the available habitat space (Supplemental Fig. S5), with Demospongiae, Pennatulacea, and Stylasteridae having the largest observed range. For the two sponge taxa, silicate contributed an average of 19.8% (Demospongiae) and 14.0% (Hexactinellida), with both taxa preferentially located in areas with intermediate concentrations (Supplemental Figs. S5, S13, S16). However, detection of the real drivers of sponge distribution may be confounded by the relatively high correlation of silicate with dissolved oxygen (Pearson's correlation, r = −0.72, p < 0.001), POC (Pearson's correlation, r = −0.68, p < 0.001), or temperature (Pearson's correlation, r = −88, p < 0.001). Both dissolved oxygen and POC generally contributed between 3–7% when averaged across all model types. However, these variables were somewhat correlated with each other as well as with aragonite, calcite, and temperature (Fig. 2). In general, taxa occurred in areas with higher POC concentrations than expected by chance, and at intermediate dissolved oxygen concentrations (Supplemental Figs. S5, S11–S20). RAC contributed between 9.1–75.2% of information to the models. BRT models generally had the largest percent contribution of RAC, with a mean of 46.3% across all taxa and the largest percent contribution of 75.2%. Maxent and RF models both had lower contributions of RAC, with means of 24.8% and 22.6%, respectively. Substrate type, quantified as percent gravel and percent mud, generally contributed less than 10% of information to the models for all taxa. While taxa generally occurred across all values of these predictors (Supplemental Figs. S5, S11–S20), there was a general trend towards areas with a lower percentage of mud and a higher percentage of gravel than expected by chance. Being located on or off a seamount generally contributed negligibly to models, with an average percent contribution of 3.0% across all taxa and model types. Taxa were most frequently predicted to be distributed on or near seamounts, although this may partially reflect sampling bias in the occurrence data.

Table 4

Percent variable contributions for Boosted Regression Trees (BRT), Maximum Entropy (Maxent), and Random Forest (RF) models. A value of 'na' indicates that the variable was not used in the model for that taxon. BPI = broad scale bathymetric position index. POC = particulate organic carbon flux. RAC = residual auto-covariate. Slope SD = standard deviation of slope.

Taxa	Model	Arag.	BPI	Calcite	Diss. Oxygen	Percent Gravel	Percent Mud	POC	RAC	Seamount	Silicate	Slope SD	Temp.	Ruggedness
Alcyonacea	BRT	na	6.3	17.5	3.9	4.1	2.7	5.6	9.1	1.6	na	7.6	5.4	36.3
	Maxent	na	3.0	58.3	1.5	0.7	0.3	0.9	20.2	0.1	na	2.1	0.9	11.9
	RF	na	8.7	8.8	10.3	9.1	10.2	8.7	17.3	3.6	na	6.6	6.4	10.2
Antipatharia	BRT	na	1.8	7.4	2.6	2.0	1.1	4.8	61.1	0.9	na	1.6	9.5	7.4
	Maxent	na	3.1	47.9	2.1	0.8	0.7	1.4	32.1	1.4	na	1.7	0.8	8.2
	RF	na	9.4	7.7	8.3	6.9	8.7	7.1	24.7	2.1	na	6.2	5.6	13.3
Demospongiae	BRT	na	17.9	na	4.7	4.1	4.6	9.6	19.1	5.1	14.3	4.2	4.0	12.3
	Maxent	na	23.7	na	0.8	0.9	2.7	1.8	23.5	2.3	36.3	2.0	2.0	4.2
	RF	na	13.0	na	8.4	8.1	7.1	7.4	16.5	4.4	8.7	8.2	8.1	10.2
E. rostrata	BRT	4.4	3.0	na	0.6	1.2	1.9	1.5	75.2	2.0	na	1.6	4.7	4.0
	Maxent	24.4	4.9	na	2.1	3.4	0.6	5.5	30.6	1.6	na	0.3	14.2	12.5
	RF	10.3	9.1	na	8.7	6.6	7.2	7.9	20.4	2.5	na	8.5	9.3	9.5
G. dumosa	BRT	13.1	5.5	na	4.6	2.9	4.7	6.3	37.9	0.3	na	2.9	14.8	6.9
	Maxent	7.8	6.6	na	1.0	7.7	2.4	0.5	16.0	6.8	na	2.1	31.8	17.1
	RF	8.9	9.9	na	8.7	6.7	6.9	7.2	20.7	5.7	na	8.2	8.3	8.8
Hexactinellida	BRT	na	4.4	na	3.5	2.3	3.6	7.3	44.3	6.3	5.4	2.5	6.9	13.6
	Maxent	na	0.5	na	1.0	2.9	0.7	2.9	15.6	12.9	29.0	4.9	11.1	18.5
	RF	na	8.1	na	6.2	6.1	6.8	7.7	30.1	4.4	7.7	6.1	7.9	8.9
M. oculata	BRT	12.5	14.1	na	3.8	5.9	2.2	3.5	32.6	1.0	na	4.3	10.4	9.8
	Maxent	30.4	10.8	na	3.4	10.2	3.4	2.0	21.6	4.4	na	1.8	0.2	11.7
	RF	10.3	9.7	na	6.1	6.1	6.7	7.3	24.2	4.8	na	7.5	7.7	9.6
Pennatulacea	BRT	4.1	5.9	na	2.6	1.2	1.8	8.0	68.6	0.1	na	0.9	5.3	1.6
	Maxent	7.8	20.6	na	7.3	2.2	2.6	8.4	16.5	1.7	na	2.6	30.0	0.3
	RF	8.1	8.8	na	7.1	7.6	7.6	7.5	27.6	4.0	na	6.5	8.7	6.6
S. variabilis	BRT	5.3	1.8	na	0.8	4.6	0.9	1.6	72.6	0.0	na	2.4	3.2	6.7
	Maxent	6.7	6.7	na	3.1	3.3	0.5	1.8	36.5	0.0	na	7.1	24.0	10.2
	RF	9.8	9.3	na	7.5	5.0	5.1	7.2	22.9	3.6	na	9.0	10.6	10.0
Stylasteridae	BRT	10.3	14.0	na	2.2	2.0	7.9	3.7	42.8	2.5	na	2.9	4.7	7.0
	Maxent	32.9	11.4	na	0.6	1.7	3.7	0.4	35.2	0.4	na	1.8	1.5	10.2
	RF	9.2	11.8	na	7.9	7.7	9.1	6.8	21.6	3.3	na	7.4	5.7	9.7

3.4. Ensemble modeling

A single ensemble model was constructed using a weighted average of the BRT, Maxent, and RF models for each taxon (Figs. 3–6). The weighted ensemble models were generally a nearly equal mixture of the three model types for each taxon, as each approach yielded similar AUC values for each taxon (Table 3). The ensemble models produced similar patterns, as the three modeling approaches yielded comparable habitat suitability predictions, with the Pearson's correlation between model types for the same taxon generally being 0.7 or higher (Supplemental Table S2). In general, the BRT and Maxent models were more highly correlated with an average correlation of 0.894 across all taxa, and the RF and Maxent models were less correlated with an average correlation of 0.812. The three modeling approaches exhibited similar spatial patterns, with the biggest absolute differences between each approach typically occurring within areas of predicted high suitability (see Supplemental Fig. S9). However, the model types differed considerably in the scale of predicted habitat suitability scores (Supplemental Fig. S10). BRT consistently predicted higher suitability scores across all taxa, with more extensive areas predicted to have high suitability scores (> 0.8) than the other models (Fig. 3). Maxent generally predicted the lowest habitat suitability scores (Fig. 4), and RF frequently predicted a larger area of intermediate suitability (Fig. 5).

Spatial uncertainty (measured as the coefficient of variation, CV) differed somewhat among the three modeling approaches (see Supplemental Figs, S6–S8). While the models generally exhibited a similar spatial pattern in uncertainty (therefore leading to similar patterns in ensemble models), BRT models typically had the largest CVs, followed by RF and then Maxent. The ensemble model for each taxon generally outperformed any of the individual modeling approaches as assessed by AUC scores calculated using the entire dataset and the correlation between predicted and expected values. However, the RF models marginally outperformed the ensemble model in some cases (Table 5).

3.5. Spatial distributions

The ensemble models (Fig. 6) predicted that highly suitable (scores > 0.8) habitat for each taxon occurs in various regions throughout the New Zealand EEZ as well as in portions of the SPRFMO convention area that were included in the study area (Fig. 1). For most taxa, predicted concentrations of highly suitable habitat were located on areas of the Chatham Rise, Kermadec Ridge, Colville Ridge, South Tasman Rise, Macquarie Ridge, Norfolk and West Norfolk Ridges, Lord Howe Rise, Three Kings Ridge, Louisville Seamount Chain, and along much of the continental shelf edge of Australia and the east coast of New Zealand's North Island (Fig. 6). Sea pens (Pennatulacea) had the largest predicted distribution of highly suitable habitat (Fig. 6h), with suitability predicted to be above 0.6 in most of these areas, the Chatham Rise in particular. Soft corals (Alcyonacea) and black corals (Antipatharia) generally exhibited a spatial pattern comparable to sea pens, but with lower predicted habitat suitability scores (Fig. 6a,b and Supplemental Fig. S10) and slightly higher CVs (Fig. 7a,b). The two sponge taxa (Demospongiae and Hexactinellida) had the lowest CVs across the study area (Fig. 7c,f), with habitat suitability predictions similar to but not quite as high as the predictions for sea pens (Fig. 6h). Models for the four species of reef-forming stony corals (*E. rostrata*, *G. dumosa*, *M. oculata*, and *S. variabilis*) typically predicted lower habitat suitability than for other taxa (Supplemental Fig. S10), with highly suitable habitats clustered on the southeast corner of Chatham Rise, Kermadec Ridge, Colville Ridge, and West Norfolk Ridge (Fig. 6). On the Louisville Seamount Chain, stony corals typically had lower habitat suitability scores than other taxa, with predictions generally showing that only the upper flanks and tops of seamounts were moderately or highly suitable (Fig. 8). In addition, the stony coral taxa exhibited the highest CVs, although model predictions were typically more robust in areas predicted to be more suitable (Fig. 7). Hydrocorals (Stylasteridae) followed a similar spatial pattern to the stony corals, but with generally
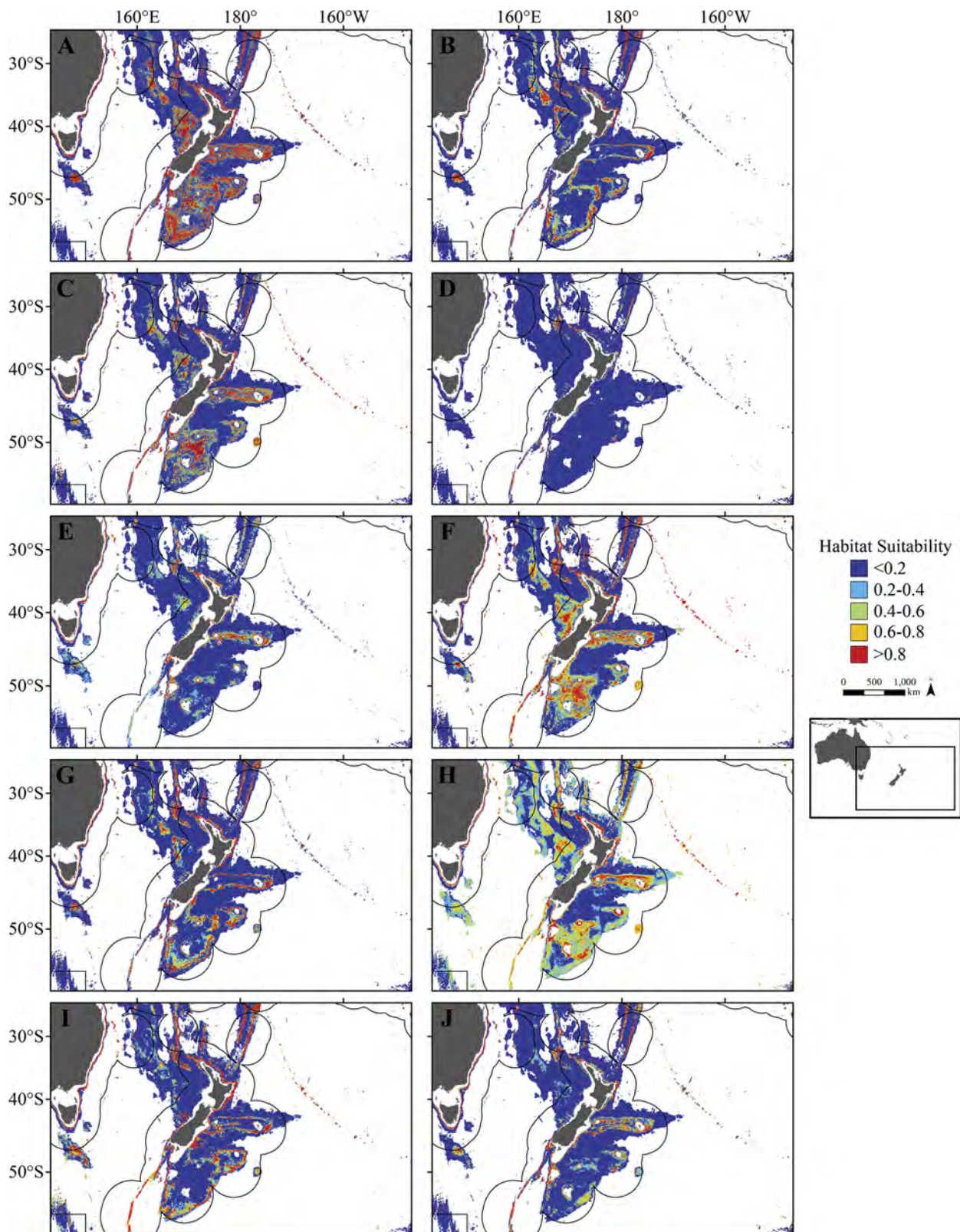
Fig. 3. Boosted Regression Tree models for each taxonomic group showing predicted habitat suitability. Warmer colors indicate more suitable habitat. The black line indicates the extent of the SPRFMO convention area. A = Alcyonacea, B = Antipatharia, C = Demospongiae, D = *E. rostrata*, E = *G. dumosa*, F = Hexactinellida, G = *M. oculata*, H = Pennatulacea, I = *S. variabilis*, J = Stylasteridae.
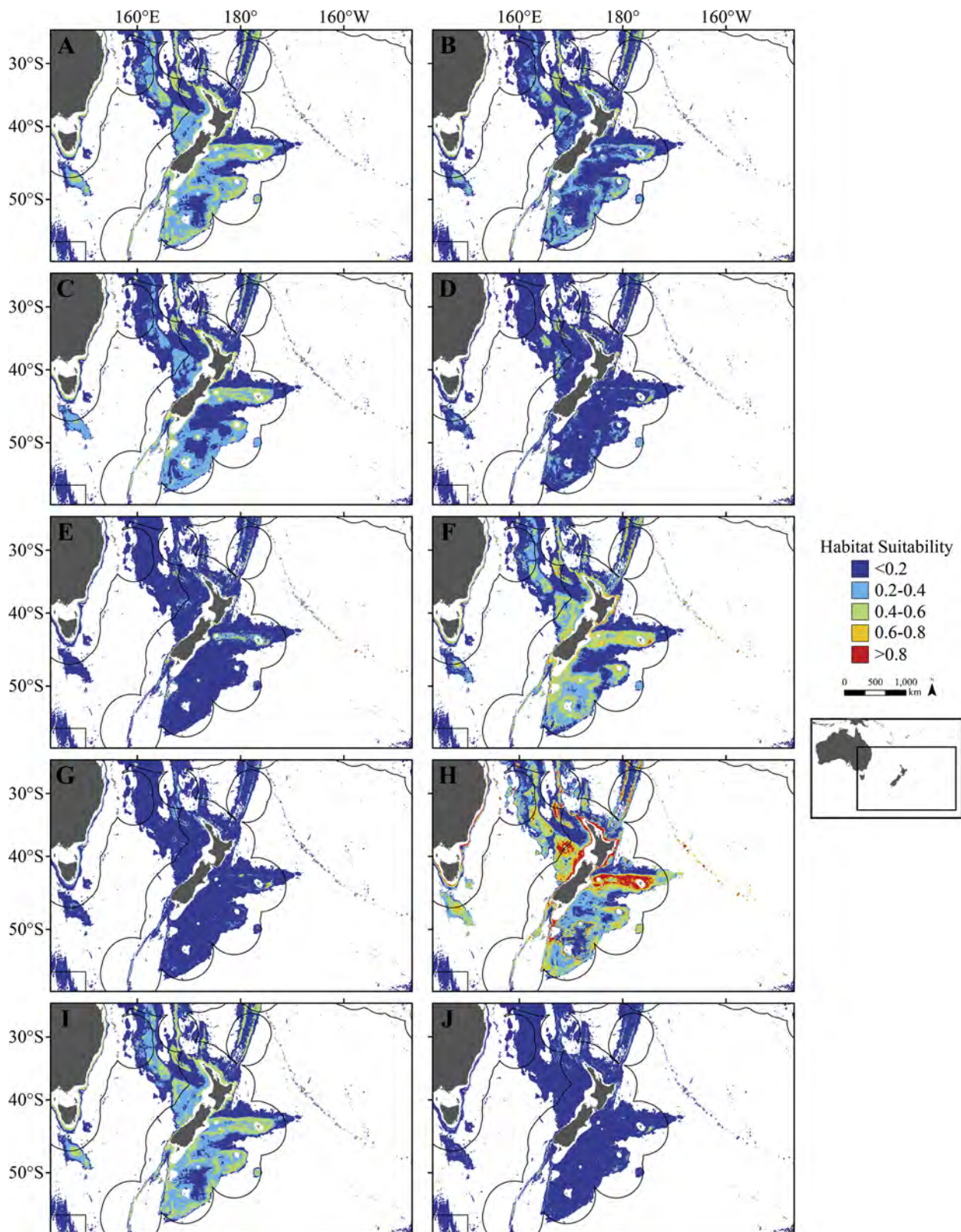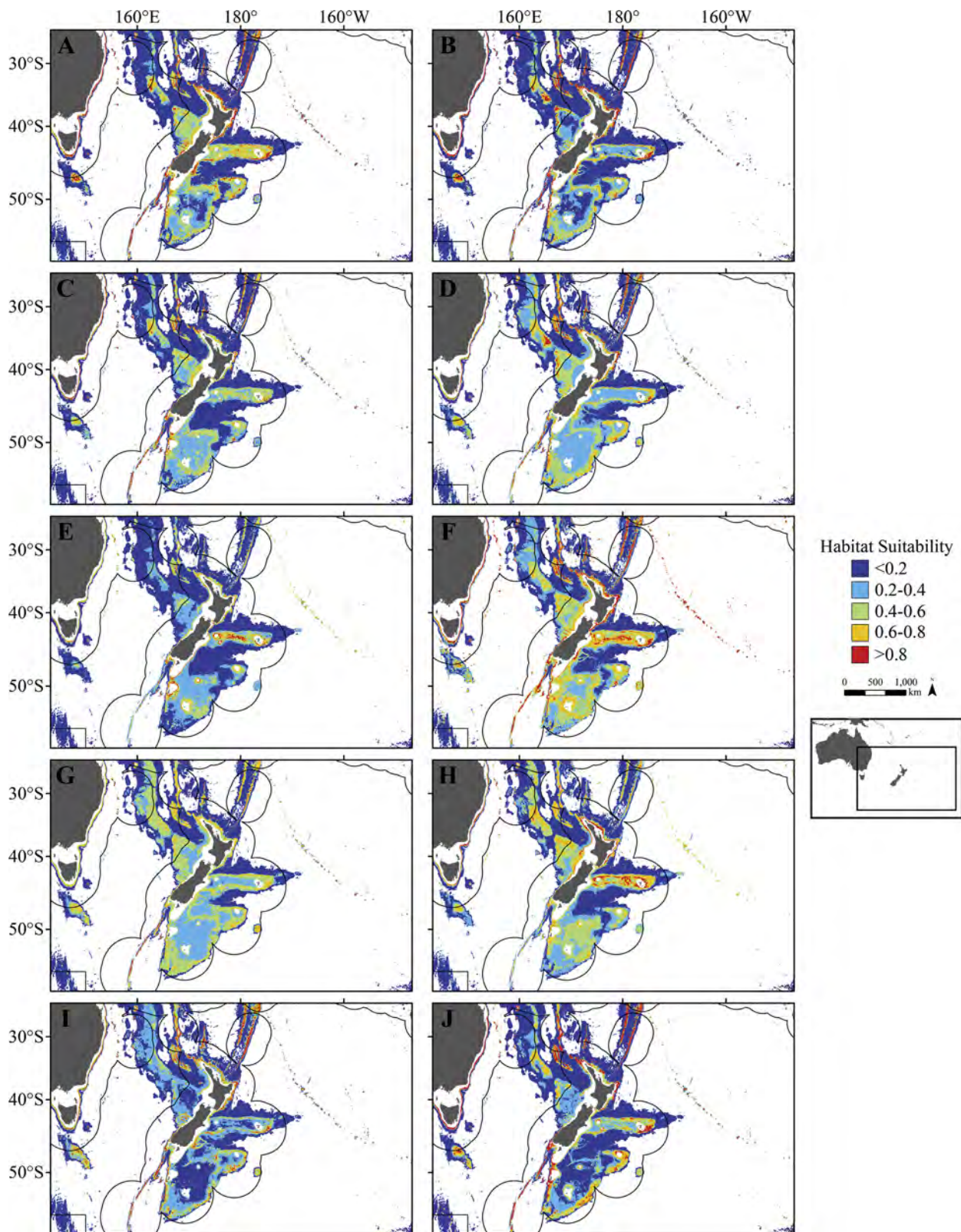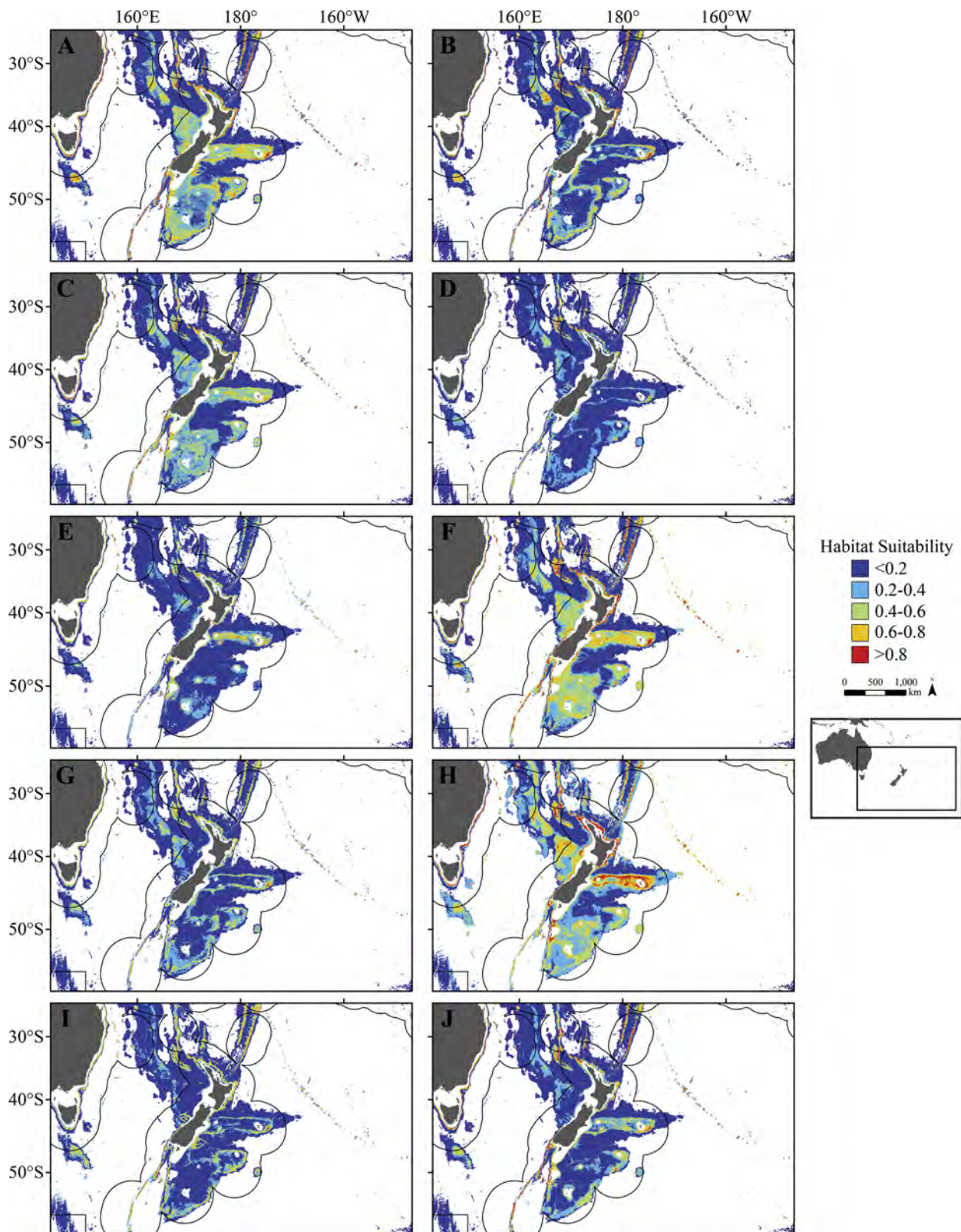
Fig. 4. Maximum Entropy models for each taxonomic group showing predicted habitat suitability. Warmer colors indicate more suitable habitat. The black line indicates the extent of the SPRFMO convention area. A = Alcyonacea, B = Antipatharia, C = Demospongiae, D = *E. rostrata*, E = *G. dumosa*, F = Hexactinellida, G = *M. oculata*, H = Pennatulacea, I = *S. variabilis*, J = Stylasteridae.

Fig. 5. Random Forest models for each taxonomic group showing predicted habitat suitability. Warmer colors indicate more suitable habitat. The black line indicates the extent of the SPRFMO convention area. A = Alcyonacea, B = Antipatharia, C = Demospongiae, D = *E. rostrata*, E = *G. dumosa*, F = Hexactinellida, G = *M. oculata*, H = Pennatulacea, I = *S. variabilis*, J = *Stylasteridae*.

Fig. 6. Ensemble models for each taxonomic group showing the weighted habitat suitability mean calculated using Maximum Entropy, Boosted Regression Trees, and Random Forest models. Warmer colors indicate more suitable habitat. The black line indicates the extent of the SPRFMO convention area. A = Alcyonacea, B = Antipatharia, C = Demospongiae, D = *E. rostrata*, E = *G. dumosa*, F = Hexactinellida, G = *M. oculata*, H = Pennatulacea, I = *S. variabilis*, J = Stylasteridae.

Table 5
Performance metrics for Boosted Regression Trees (BRT), Maximum Entropy (Maxent), Random Forest (RF), and ensemble models for each taxon, assessed by the area under the curve (AUC) score calculated using the entire occurrence and background dataset, and by the correlation (R^2) between expected and predicted values. Larger AUC and R^2 values (closer to 1) indicate better model performance.

Taxa	AUC				R^2			
	BRT	Maxent	RF	Ensemble	BRT	Maxent	RF	Ensemble
Alcyonacea	0.967	0.897	0.974	0.972	0.61	0.42	0.67	0.64
Antipatharia	0.947	0.936	0.970	0.959	0.53	0.49	0.57	0.56
Demospongiae	0.958	0.890	0.960	0.962	0.57	0.39	0.59	0.59
E. rostrata	0.941	0.962	0.962	0.961	0.23	0.32	0.17	0.27
G. dumosa	0.924	0.927	0.935	0.942	0.17	0.24	0.15	0.19
Hexactinellida	0.901	0.867	0.910	0.914	0.32	0.29	0.33	0.34
M. oculata	0.954	0.948	0.959	0.963	0.17	0.35	0.11	0.22
Pennatulacea	0.845	0.847	0.884	0.872	0.14	0.11	0.15	0.14
S. variabilis	0.955	0.959	0.959	0.962	0.31	0.44	0.32	0.37
Stylasteridae	0.952	0.932	0.955	0.956	0.52	0.46	0.47	0.51

higher suitability scores (Supplemental Fig. S10), especially along the Louisville Seamount Chain (Fig. 8j), and lower CVs (Fig. 7j).

4. Discussion

4.1. Spatial distribution patterns and niche space

We produced habitat suitability models for ten VME indicator taxa in order to inform the spatial management of bottom trawl fisheries in the SPRFMO convention area. Areas of suitable habitat were predicted to occur throughout the region for each taxon, but the most highly suitable habitat was generally clustered in patches covering a relatively small fraction of the total area studied. The patchy distribution of suitability predicted for each taxon highlights the difficulties in drawing broad conclusions from limited field observations over vast areas of seafloor, and emphasizes the role of modeling to fill in key knowledge gaps.

While there were variations among the modeled taxa, habitat suitability maps generally predicted high suitability of at least one of the ten VME indicator taxa across a range of distinct areas, many of which support important bottom trawl fisheries, distributed across the SPRFMO convention area and the New Zealand EEZ. These areas included the Chatham Rise, Kermadec Ridge, Colville Ridge, South Tasman Rise, Macquarie Ridge, West Norfolk Ridge, and Louisville Seamount Chain. Terrain metrics, when considered in concert, generally contributed considerably to the models for each taxon. Topographical features have long been known to be important drivers of cold-water coral and sponge distributions, and many modeling studies have relied solely on terrain metrics as predictor variables (e.g., Dolan et al., 2008; Rengstorf et al., 2012; Tong et al., 2013; Rowden et al., 2017). As filter feeders, cold-water corals and sponges generally exhibit strong preferences for complex and elevated seafloor features that experience locally accelerated currents (Masson et al., 2003; Bryan and Metaxas, 2006). Amplified current speeds improve habitat quality by increasing food supply and capture rate (Thiem et al., 2006; Purser et al., 2010), larval delivery (Piepenburg and Müller, 2004), and the removal of sediments (Rogers, 1994). The availability of hard bottom substrate is generally considered critical for the successful recruitment of many coral species (Gass and Roberts, 2006). Therefore, it is surprising that the substrate variables (percent gravel and percent mud) were not more important in the models, as has been observed in previous studies (e.g., Georgian et al., 2014). However, it is possible that the use of broad scale, interpolated sediment data concealed the confluence of substrate and topographic features at finer scales, for example, small carbonate outcrops and boulders that are known to support cold-water corals (e.g., Baco et al., 2010; Ross and Howell, 2012). The 1 km^2 grid resolution used in our analysis may have also been too broad to capture small-scale variation in substrate type and

topographical features, even in areas with high sampling intensity.

Particulate organic carbon (POC) export to the seafloor, the saturation state of aragonite and calcite, temperature, and dissolved oxygen all contributed a small to moderate amount of explanatory power to the models for each taxon. POC flux to the seafloor serves as a proxy for the supply of food to cold-water corals and sponges, which are most frequently found in highly productive waters (Roberts et al., 2009; Wagner et al., 2011). The saturation states of aragonite and calcite generally had moderate to high importance when included in models, particularly for the four species of stony corals, which tended to occur in areas with higher saturation states. High saturation states make it energetically easier for organisms to secrete calcium carbonate skeletal materials, and have been demonstrated to be vital for the growth and survival of cold-water corals in experimental (Georgian et al., 2016b; Kurman et al., 2017), field (Lunden et al., 2013; Georgian et al., 2016a), and modeling (Tittensor et al., 2009; Davies and Guinotte, 2011; Quattrini et al., 2013) studies. Given that most occurrence records occurred within a relatively narrow temperature range, as has been commonly observed in other cold-water coral (Roberts et al., 2005; Cohen et al., 2006; Georgian et al., 2014) and sponge (Leys et al., 2004) studies, it is not surprising that temperature was also an important driver of predicted distributions, especially notable in the four branching coral species. Our models also indicated a general preference for moderate to high levels of dissolved oxygen, indicating that regions with low dissolved oxygen concentrations may be untenable for recruitment or long-term survival. Previous studies have also found that cold-water corals and sponges are generally absent in regions with low concentrations of dissolved oxygen (Freiwald, 2002; Leys et al., 2004; Wisshak et al., 2005). Experimental observations confirm that cold-water corals are sensitive to changes in both temperature and dissolved oxygen (Dodds et al., 2007; Lunden et al., 2014; Naumann et al., 2014). However, it is still uncertain to what extent dissolved oxygen structures their spatial distribution, especially given the extensive cold-water coral populations that occur in oxygen minimum zones in the Gulf of Mexico (Georgian et al., 2016a) and the North Pacific (Baco, 2007). It is important to note however, that due to the relatively high correlations of each of these variables with each other, as well as with depth, it is difficult to discern their biological relevance based solely on field observations and modeling results.

4.2. Previous modeling studies in the region

The models presented here expand and improve on previous efforts to model VME indicator taxa in the South Pacific Ocean. Global habitat suitability models have predicted that extensive suitable habitat exists throughout the New Zealand region for stony corals (Tittensor et al., 2009; Davies and Guinotte, 2011), black corals (Yesson et al., 2017), and octocorals (Yesson et al., 2012). On the regional scale, Tracey et al.
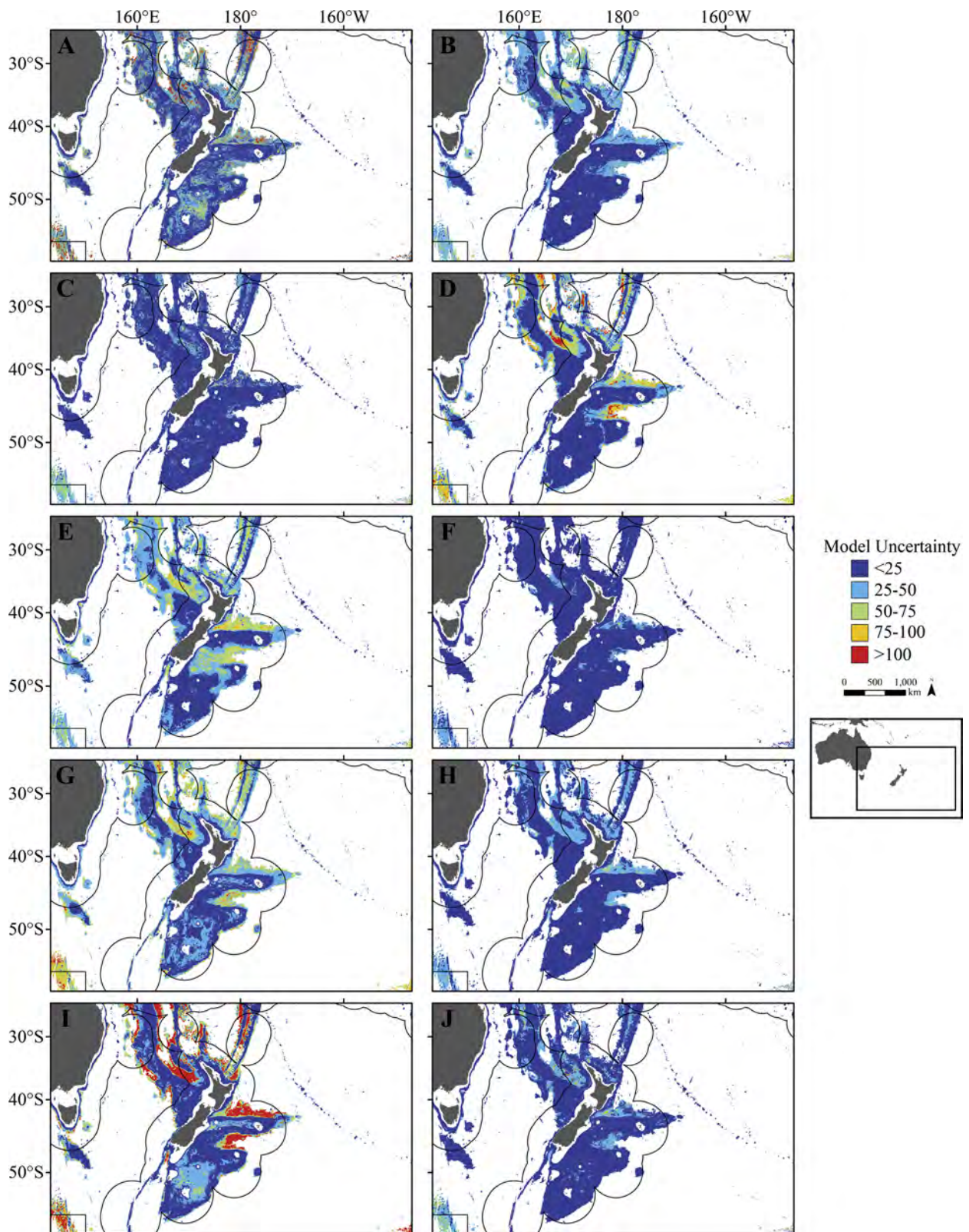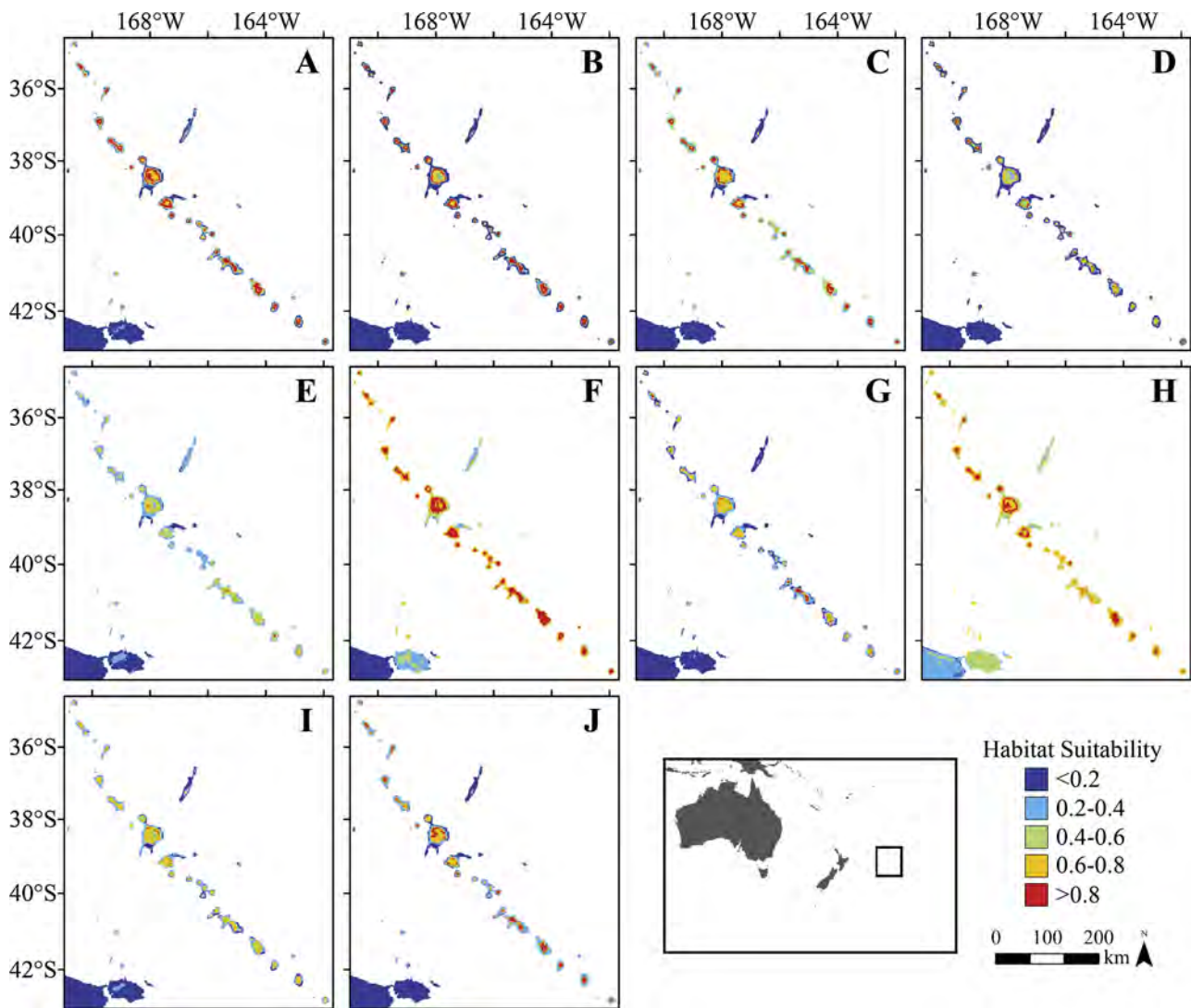
Fig. 7. Model uncertainty (coefficient of variation) for the ensemble model of each taxonomic group calculated using Boosted Regression Trees, Maximum Entropy, and Random Forest models. Warmer colors indicate areas with higher uncertainty. The black line indicates the extent of the SPRFMO convention area. A = Alcyonacea, B = Antipatharia, C = Demospongiae, D = *E. rostrata*, E = *G. dumosa*, F = Hexactinellida, G = *M. oculata*, H = Pennatulacea, I = *S. variabilis*, J = Stylasteridae.

Fig. 8. Ensemble models for the central Louisville Seamount Chain for each taxonomic group showing the weighted habitat suitability mean calculated using Boosted Regression Trees, Maximum Entropy, and Random Forest models. Warmer colors indicate more suitable habitat. A = Alcyonacea, B = Antipatharia, C = Demospongiae, D = *E. rostrata*, E = *G. dumosa*, F = Hexactinellida, G = *M. oculata*, H = Pennatulacea, I = *S. variabilis*, J = Stylasteridae.

(2011) used BRT to model the distribution of *E. rostrata*, *G. dumosa*, *M. oculata*, *Oculina virgosa*, and *S. variabilis* on the New Zealand extended continental shelf. Predicted habitat suitability generally exhibited similar spatial patterns but lower suitability scores than those calculated in the present study. However, *S. variabilis* had notably different predictions, with a much larger distribution of highly suitable habitat occurring on the deeper slopes of the Campbell Plateau. Anderson et al. (2016a) constructed Maxent and BRT models for four species of stony corals (*E. rostrata*, *G. dumosa*, *M. oculata*, and *S. variabilis* combined) over the entire SPFRMO area and New Zealand EEZ. Their results showed that suitable habitat for stony corals was generally restricted to the Kermadec Ridge, Colville Ridge, West Norfolk Ridge, Lord Howe Rise, Tasman Rise, and Louisville Seamount Chain, a spatial pattern similar to our results for the same species. The most important predictors in their models were aragonite saturation state, dissolved oxygen, particulate organic carbon, and temperature, which also contributed moderate to large amounts of explanatory power to the models in our study. Interestingly, as also seen in our models, their BRT models predicted higher suitability scores than the Maxent models, but exhibited similar spatial patterns, suggesting that this may be a trait inherent to these modeling types and highlighting the benefits of using an ensemble modeling approach. However, the authors noted that errors in

the SRTM30 bathymetry data, sampling approach, lack of substrate data, and spatial bias may have negatively affected their models - as demonstrated by subsequent field validation efforts. Building from that study, Anderson et al. (2016b) used BRT and Maxent to model suitable habitat for a wide range of VME indicator taxa in the New Zealand EEZ and immediately adjacent areas. Their habitat suitability models generally show similar spatial patterns as the results in this study for the same taxa, but typically predicted a much larger area of highly suitable habitat for most taxa. For example, highly suitable habitats for *S. variabilis*, *E. rostrata*, *G. dumosa*, *M. oculata*, Stylasteridae, and Antipatharia were predicted to occur extensively on and around the Campbell and Bounty plateaus, while in our study suitable habitats in this region were generally constrained to small patches. Interestingly, *Pennatulacea* also exhibited the lowest model performance in their study, while similarly having the largest predicted distribution of suitable habitat across the region. Previous studies have found that more generalized species with large areas of occupancy (in this case, likely driven by the ability of *Pennatulacea* to occupy soft bottom substrate on a variety of terrain features), corresponded to worse model performance (Araújo and Williams, 2000; Segurado and Araújo, 2004). However, neither Anderson et al. (2016b) or our study found a significant correlation between sample size and model performance, indicating that this result

was not simply an artefact of our available data. An important difference between the two studies (which used a similar set of presence data) was the use of a much larger set of background/absence records by Anderson et al. (2016b) (10,000) based on true target-group observations.

A more recent habitat suitability modeling study focused efforts on the Louisville Seamount Chain, an area historically and currently subjected to a bottom trawl fishery that largely targets orange roughy. Based on a camera and bathymetry survey, a suite of high-resolution (25 m) ensemble models were constructed for *S. variabilis* and two VME habitat indicator taxa, Brisingida and Crinoidea, for six of the chain's 80+ seamounts, using true presence-absence data (Rowden et al., 2017). The authors predicted that only small, rare patches (< 0.1% of the total modeled area) on the flanks of seamounts could be characterized as VMEs, based on the density thresholds specified by New Zealand's Environmental Protection Agency through domestic regulations (see Penney et al., 2009). The abundance and probability of presence of *S. variabilis* was predicted to be low across much of the seamounts. In contrast, the *S. variabilis* model produced in our study predicted extensive suitable habitat throughout the Louisville Seamount Chain, with the most suitable habitat frequently occurring on both the tops and sides of seamounts. This discrepancy may indicate that our models generally over-predict VME species occurrence on smaller scale topographical features and that, in general, habitat suitability values from models using pseudo-absence background data should be treated as relative rather than absolute quantities. It is also likely that their use of abundance data and true absence data contributed to an improved ability to classify an appropriate threshold for VME designation on the seamount scale. Previous models (Anderson et al., 2016a) also had apparent accuracy issues in the Louisville Seamount Chain based on subsequent field validation surveys, suggested to have resulted not only from errors in the bathymetry data but also from the lack of true absences, lack of substrate data, and the unique topography of the seamounts compared to the areas that predominantly influenced model training. In particular, the lack of high quality substrate data may be hampering modeling efforts at the seamount scale, as may be a lack of precision in upscaled transforms of critical predictor variables with low native resolutions (e.g., silicate (1˚) and carbonate parameters (0.5˚); Table 1). Field surveys have found that seamounts in the Louisville Seamount Chain frequently have relatively flat tops with sandy substrate that would not be suitable for larval recruitment of stony corals (Anderson et al., 2016a). While our models were trained using substrate data (percent gravel and percent mud), the raw data were generally concentrated within the New Zealand EEZ and interpolated elsewhere. Therefore, we likely failed to accurately depict substrate at a biologically relevant scale in the Louisville Seamount Chain. Regardless of the causal mechanism, these results highlight the importance of pairing broad-scale regional models with high-resolution models in order to better inform spatial management and produce protection recommendations that can be implemented at appropriate scales.

The habitat suitability modeling in this study represents a considerable improvement over previous regional efforts. In particular, many of the potential sources of error outlined in Anderson et al. (2016a) have been addressed, and we followed many of their recommendations for future studies. Environmental variables were constructed using new bathymetry data that is not prone to the errors and bias found in the SRTM30 data used in previous studies. In addition, we were able to include variables that characterize substrate, which is often a critical factor in determining the distribution of cold-water corals (e.g., Georgian et al., 2014). The modeling techniques used were also more sophisticated. We used a targeted background sampling approach designed to eliminate the deleterious influence of spatial bias in the occurrence datasets. Spatial bias reduces the performance and interpretability of models (Phillips et al., 2009), but can largely be accounted for using our targeted sampling approach (Elith et al., 2010). In

addition, we adopted the residual autocovariate approach to reduce spatial autocorrelation (*sensu* Crase et al., 2012), a common feature of sampling collection that violates the assumption of independent data in many models and has been shown to increase error rates and make it difficult to ecologically interpret the influence of environmental variables (Dormann, 2007; Dormann et al., 2007). Finally, our use of ensemble models constructed using three algorithms (BRT, Maxent, and RF models) reduces the risk that structural inadequacies, inappropriate parameter specifications, or bias in each approach may unduly influence the final outputs (Robert et al., 2016). While each of the three individual modeling approaches predicted similar spatial patterns in habitat suitability, ensemble models are preferred for spatial management applications as they typically produce more robust predictions (Araújo and New, 2007) and can improve model accuracy outside of the primary training region (Randin et al., 2006). Despite these improvements, future studies in the region would benefit from using true absence and abundance data, modeling strategically important areas at a high resolution, and including higher quality substrate data. Furthermore, our models did not account for any uncertainty in the environmental predictor variables, and future modelling efforts should seek to address this issue.

4.3. The role of habitat suitability models in the conservation and management of vulnerable marine ecosystems

The predicted areas of suitable habitat for VME indicator taxa overlap with historically trawled regions or places of current fishery interest in the South Pacific Ocean (Clark and O'Driscoll, 2003; Clark et al., 2016; Black and Tilney, 2017; Roux et al., 2017). The apparent overlap of VMEs and prime fish habitat is unsurprising given the strong association of many demersal fish species with seamounts and other elevated seafloor features, and the complex structures that can be created by cold-water corals and sponges on such topography (Costello et al., 2005; Baillon et al., 2012; Kutti et al., 2014). This overlap presents a significant obstacle to limiting adverse impacts to benthic habitats while still maintaining sustainable fisheries. However, it is possible to optimize closures by spatially quantifying the trade-off between biodiversity protection and fishery catch, including with the combined use of habitat suitability models and decision-support tools (e.g., Leathwick et al., 2008; Penney and Guinotte, 2013; Lagasse et al., 2015).

Bottom trawling has known detrimental impacts on benthic communities (Watling and Norse, 1998; Fosså et al., 2002; Clark and Koslow, 2007; Sampaio et al., 2012; Pusceddu et al., 2014). Interim protection measures for VMEs in the SPRFMO convention area include the move-on rule that requires vessels to move bottom trawl operations at least five nautical miles away if a VME taxon is caught as bycatch. While limiting fishing activity in areas known to contain VMEs is paramount, move-on rules are increasingly being questioned due to the risk of expanding the footprint of trawling into new areas and the potential for damage by even a low level of trawling (Williams et al., 2010). Spatial closures are an important tool for effectively managing bottom-trawling fisheries and preventing significant adverse impacts to VMEs in the deep sea. However, designating the location and extent of closures is challenging because the distribution of VMEs is generally poorly known, and there is often industry and political opposition to restricting fishing opportunities. The models presented here will aid in assessments of the effectiveness of current and historical closures, and will inform the creation of improved, long-term management plans for the region. The spatial predictions generated by these models will allow managers to move beyond the current interim protections by enabling future selection of closed areas with habitats that are predicted to be highly suitable for VME indicator taxa. Given the broad range of taxa modeled, our results will also help protect areas that are predicted to be highly suitable for a large diversity of associated taxa, both acting to reduce VME indicator taxa bycatch in trawls and protect key

biodiversity hotspots. In addition, models can be used to help ensure that future closures are representative of key seafloor features and biogeographic areas, ensuring adequate protection for all VME associated taxa (Penney et al., 2009).

Habitat suitability models also provide a framework for assessing and responding to other anthropogenic stressors that threaten VMEs in addition to bottom trawling. For example, climate change and ocean acidification are projected to impact the distribution of cold-water corals around the world (Guinotte et al., 2006; Tittensor et al., 2010). Cold-water corals in the Pacific Ocean already grow in waters with low aragonite and calcite saturation states (this study and Thresher et al., 2011; McCulloch et al., 2012; Jantzen et al., 2013; Baco et al., 2017). Ocean acidification is projected to significantly reduce aragonite and calcite saturation states (Feely et al., 2012), making it difficult or impossible for corals to continue to calcify and create reef structures (Turley et al., 2007). In concert with reduced pH, temperatures in the deep sea are projected to rise due to ongoing global warming (Mora et al., 2013). As deeper waters acidify and warm, cold-water corals and sponges may exhibit reduced calcification rates (Lunden et al., 2014; Georgian et al., 2016b), altered metabolic activity (Dodds et al., 2007), and changes in gene expression (Carreiro-Silva et al., 2014). Climate change is also expected to result in the decreased export of POC to the seafloor and reduced dissolved oxygen levels (Mora et al., 2013), which may hamper the ability of VME indicator taxa to metabolically compensate for other stressors (Georgian et al., 2016b). Thus, for spatial management measures to be able to accommodate the cumulative impact of fishing and climate change on VMEs, habitat suitability models that use environmental variables adjusted for future climate change scenarios should also be used to inform spatial management processes.

5. Conclusions

Given the likelihood that benthic communities in the deep sea will be adversely affected by several anthropogenic stressors, it is necessary to devise effective, science-based management strategies for ensuring their continued survival. Restoration of deep-sea environments is generally prohibitively expensive or impossible with current technology, making preemptive management activities critical for preserving these vulnerable habitats (Van Dover et al., 2014). While global stressors stemming from the effects of climate change will be extremely difficult to mitigate, efforts to reduce perturbations from acute stressors such as bottom trawling may prove to be both practical and fruitful. Recently, effort has been put into establishing closed areas for protecting deep-sea ecosystems from the impacts of trawling, and habitat suitability modeling can and has assisted in the design and evaluation of these conservation efforts (e.g., Knudby et al., 2013; Guinotte and Davies, 2014; Rooper et al., 2014). We recommend that regional and local habitat suitability models be used in conjunction to identify and characterize the distribution of VMEs and inform the spatial management planning of fisheries organizations.

Declarations of interest

None.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fishres.2018.11.020.

References

Anderson, O.F., Guinotte, J.M., Rowden, A.A., Clark, M.R., Mormede, S., Davies, A.J., Bowden, D.A., 2016a. Field validation of habitat suitability models for vulnerable marine ecosystems in the South Pacific Ocean: implications for the use of broad-scale models in fisheries management. Ocean Coast. Manage. 120, 110–126.

Anderson, O.F., Guinotte, J.M., Rowden, A.A., Tracey, D.M., Mackay, K.A., Clark, M.R., 2016b. Habitat suitability models for predicting the occurrence of vulnerable marine ecosystems in the seas around New Zealand. Deep-Sea Res. Pt. I 115, 265–292.

Aqua MODIS, 2014. NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Ocean Color Data; 2014 Reprocessing. NASA OB.DAAC, Greenbelt, MD, USA. https://doi.org/10.5067/AQUA/MODIS_OC.2014.0. [dataset].

Araújo, M.B., New, M., 2007. Ensemble forecasting of species distributions. Trends Ecol. Evol. 22 (1), 42–47.

Araújo, M.B., Williams, P.H., 2000. Selecting areas for species persistence using occurrence data. Biol. Conserv. 96, 331–345.

Baco, A.R., 2007. Exploration for deep-sea corals on North Pacific seamounts and islands. Oceanography 20 (4), 108–117.

Baco, A.R., Morgan, N., Roark, E.B., Silva, M., Shamberger, K.E.F., Miller, K., 2017. Defying dissolution: discovery of deep-sea scleractinian coral reefs in the North Pacific. Sci. Rep. 7, 5436. https://doi.org/10.1038/s41598-017-05492-w.

Baco, A.R., Rowden, A.A., Levin, L.A., Smith, C.R., Bowden, D.A., 2010. Initial characterization of cold seep faunal communities on the New Zealand Hikurangi margin. Mar. Geol. 272 (1–4), 251–259.

Baillon, S., Hamel, J.-F., Wareham, V.E., Mercier, A., 2012. Deep cold-water corals as nurseries for fish larvae. Front. Ecol. Environ. 10 (7), 351–356.

Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W., 2012. Selecting pseudo-absences for species distribution models: how, where and how many? Methods Ecol. Evol. 3 (2), 327–338.

Behrenfeld, M.J., Falkowski, P.G., 1997. Photosynthetic rates derived from satellite-based chlorophyll concentration. Limnol. Oceanogr. 42, 1–20.

Black, J., Tilney, R., 2017. Monitoring New Zealand's Trawl Footprint for Deepwater Fisheries: 1989/90 to 2011/12 and 1989/90 to 2012/13. New Zealand Aquatic Environment and Biodiversity Report No. 176. pp. 65 p. https://www.mpi.govt.nz/dmsdocument/5353/send.

Bostock, H.C., Fletcher, S.E.M., Williams, M.J.M., 2013. Estimating carbonate parameters from hydrographic data for the intermediate and deep waters of the Southern Hemisphere oceans. Biogeosciences 10 (10), 6199–6213.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Bryan, T.L., Metaxas, A., 2006. Distribution of deep-water corals along the North American continental margins: relationships with environmental factors. Deep-Sea Res. Pt. I 53 (12), 1865–1879.

Carreiro-Silva, M., Cerqueira, T., Godinho, A., Caetano, M., Santos, R.S., Bettencourt, R., 2014. Molecular mechanisms underlying the physiological responses of the cold-water coral *Desmophyllum dianthus* to ocean acidification. Coral Reefs 33, 465–476.

Chen, C., Liaw, A., Breiman, L., 2004. Using Random Forest to Learn Imbalanced Data. Technical Report 666. University of California, Berkeley, pp. 12 p. http://statistics.berkeley.edu/sites/default/files/tech-reports/666.pdf.

Clark, M., O'Driscoll, R., 2003. Deepwater fisheries and aspects of their impact on seamount habitat in New Zealand. J. Northwest Atl. Fish. Sci. 31, 441–458.

Clark, M.R., Koslow, J.A., 2007. Impacts of fisheries on seamounts. In: Pitcher, T.J., Morato, T., Hart, P.J.B., Clark, M.R., Haggan, N., Santos, R.S. (Eds.), Seamounts: Ecology, Fisheries and Conservation. Blackwell Publishing Ltd, Oxford, UK pp. 413–441.

Clark, M.R., McMillan, P.J., Anderson, O.F., Roux, M.J., 2016. Stock Management Areas for Orange roughy (*Hoplostethus atlanticus*) in the Tasman Sea and western South Pacific Ocean. New Zealand Fisheries Assessment Report 2016/19. pp. 27 p. https://www.researchgate.net/publication/312154577.

Cohen, A.L., Gaetani, G.A., Lundälv, T., Corliss, B.H., George, R.Y., 2006. Compositional variability in a cold-water scleractinian, *Lophelia pertusa*: new insights into "vital effects". Geochem. Geophys. Geosyst. 7, Q12004. https://doi.org/10.1029/2006GC001354.

Costello, M.J., McCrea, M., Freiwald, A., Lundälv, T., Jonsson, L., Bett, B.J., van Weering, T.C., de Haas, H., Roberts, J.M., Allen, D., 2005. Role of cold-water *Lophelia pertusa* coral reefs as fish habitat in the NE Atlantic. In: Freiwald, A., Roberts, J.M. (Eds.), Cold-Water Corals and Ecosystems. Springer-Verlag, Berlin Heidelberg, pp. 771–805.

Crase, B., Liedloff, A.C., Wintle, B.A., 2012. A new method for dealing with residual spatial autocorrelation in species distribution models. Ecography 35 (10), 879–888.

Cryer, M., Hartill, B., O'Shea, S., 2002. Modification of marine benthos by trawling: toward a generalization for the deep ocean? Ecol. Appl. 12 (6), 1824–1839.

Davies, A.J., Guinotte, J.M., 2011. Global habitat suitability for framework-forming cold-water corals. PLoS One 6, e18483. https://doi.org/10.1371/journal.pone.0018483.

De Marco, P., Diniz-Filho, J.A., Bini, L.M., 2008. Spatial analysis improves species distribution modelling during range expansion. Med. Microbiol. Lett. 4 (5), 577–580.

Dodds, L.A., Roberts, J.M., Taylor, A.C., Marubini, F., 2007. Metabolic tolerance of the cold-water coral *Lophelia pertusa* (Scleractinia) to temperature and dissolved oxygen change. J. Exp. Mar. Biol. Ecol. 349 (2), 205–214.

Dolan, M.F., Grehan, A.J., Guinan, J.C., Brown, C., 2008. Modelling the local distribution of cold-water corals in relation to bathymetric variables: adding spatial context to deep-sea video data. Deep-Sea Res. Pt. I 55 (11), 1564–1579.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W., Kühn, I., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30 (5), 609–628.

Dormann, C.F., 2007. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. Glob. Ecol. Biogeogr. 16 (2), 129–138.

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., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 1, 129–151.

Elith, J., Kearney, M., Phillips, S., 2010. The art of modelling range-shifting species. Methods Ecol. Evol. 1, 330–342.

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77 (4), 802–813.

Eppley, R.W., 1972. Temperature and phytoplankton growth in the sea. Fish. Bull. 70, 1063–1085.

Evans, J.S., Murphy, M.A., Holden, Z.A., Cushman, S.A., 2011. Modeling species distribution and change using random forest. In: Drew, A.C., Wiersma, Y., Huettmann, F. (Eds.), Predictive Species and Habitat Modeling in Landscape Ecology. Springer, New York, NY pp. 139–159.

FAO, 2009. International Guidelines for the Management of Deep-sea Fisheries in the High Seas. Technical Consultation Document. Food and Agricultural Organization (FAO), Rome, pp. 73 p. http://www.fao.org/docrep/011/i0816t/i0816t00.htm.

Feely, R.A., Sabine, C.L., Byrne, R.H., Millero, F.J., Dickson, A.G., Wanninkhof, R., Murata, A., Miller, L.A., Greeley, D., 2012. Decadal changes in the aragonite and calcite saturation state of the Pacific Ocean. Glob. Biogeochem. Cycles 26, GB3001. https://doi.org/10.1029/2011GB004157.

Fitzpatrick, M.C., Gotelli, N.J., Ellison, A.M., 2013. MaxEnt versus MaxLike: empirical comparisons with ant species distributions. Ecosphere 4 (5), 1–15.

Fosså, J.H., Mortensen, P.B., Furevik, D.M., 2002. The deep-water coral *Lophelia pertusa* in Norwegian waters: distribution and fishery impacts. Hydrobiologia 471, 1–12.

Freiwald, A., 2002. Reef-forming cold-water corals. In: Wefer, G., Billett, D., Hebbeln, D., Jorgensen, B.B., Schluter, M., Van Weering, T. (Eds.), Ocean Margin Systems. Springer-Verlag, Berlin Heidelberg, Berlin pp. 365–385.

Friedman, J.H., Meulman, J.J., 2003. Multiple additive regression trees with application in epidemiology. Stat. Med. 22 (9), 1365–1381.

Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Baranova, O.K., Zweng, M.M., Reagan, J.R., Johnson, D.R., 2014a. World Ocean atlas 2013, volume 3: dissolved oxygen, apparent oxygen utilization, and oxygen saturation. In: Levitus, S., Mishonov, A. (Eds.), NOAA Atlas NESDIS 75, pp. 27 pp. http://data.nodc.noaa.gov/woa/WOA13/DOC/woa13_vol3.pdf.

Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Baranova, O.K., Zweng, M.M., Reagan, J.R., Johnson, D.R., 2014b. World Ocean atlas 2013, volume 4: dissolved inorganic nutrients (phosphate, nitrate, silicate). In: Levitus, S., Mishonov, A. (Eds.), NOAA Atlas NESDIS 76, pp. 25 pp. http://data.nodc.noaa.gov/woa/WOA13/DOC/woa13_vol4.pdf.

Gass, S.E., Roberts, J.M., 2006. The occurrence of the cold-water coral *Lophelia pertusa* (Scleractinia) on oil and gas platforms in the North Sea: colony growth, recruitment and environmental controls on distribution. Mar. Pollut. Bull. 52 (5), 549–559.

Georgian, S.E., DeLeo, D., Durkin, A., Gomez, C.E., Kurman, M., Lunden, J.J., Cordes, E.E., 2016a. Oceanographic patterns and carbonate chemistry in the vicinity of cold-water coral reefs in the Gulf of Mexico: implications for resilience in a changing ocean. Limnol. Oceanogr. 61 (2), 648–665.

Georgian, S.E., Dupont, S., Kurman, M., Butler, A., Strömberg, S.M., Larsson, A.I., Cordes, E.E., 2016b. Biogeographic variability in the physiological response of the cold-water coral *Lophelia pertusa* to ocean acidification. Mar. Ecol. 37 (6), 1345–1359.

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. Mar. Ecol. Prog. Ser. 506, 145–161.

Guinotte, J.M., Davies, A.J., 2014. Predicted deep-sea coral habitat suitability for the U.S. West Coast. PLoS One 9 (4), e93918. https://doi.org/10.1371/journal.pone.0093918.

Guinotte, J.M., Orr, J., Cairns, S., Freiwald, A., Morgan, L., George, R., 2006. Will human-induced changes in seawater chemistry alter the distribution of deep-sea scleractinian corals? Front. Ecol. Environ. 4 (3), 141–146.

Hijmans, R.J., Phillips, S., Leathwick, J., Elith, J., 2015. Dismo: Species Distribution Modeling. R Package Version 1.0-12. The R Foundation for Statistical Computing, Vienna. http://CRAN.R-project.org/package=dismo.

Hosmer, D.W., Lemeshow, S., Sturdivant, R.X., 2013. Applied Logistic Regression, third

ed. John Wiley & Sons, New York, NY pp. 153–225.

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. Ecol. Inform. 6 (3–4), 205–216.

Jantzen, C., Häussermann, V., Försterra, G., Laudien, J., Ardelan, M., Maier, S., Richter, C., 2013. Occurrence of a cold-water coral along natural pH gradients (Patagonia, Chile). Mar. Biol. 160, 2597–2607.

Jenkins, C.J., 1997. Building a national scale offshore soils database from both word-based and numeric datasets. Sea Technol. 38 (12), 25–28.

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), e82306. https://doi.org/10.1371/journal.pone.0082306.

Kühn, I., 2007. Incorporating spatial autocorrelation may invert observed patterns. Divers. Distrib. 13 (1), 66–69.

Kurman, M.D., Gómez, C.E., Georgian, S.E., Lunden, J.J., Cordes, E.E., 2017. Intra-specific variation reveals potential for adaptation to ocean acidification in a cold-water coral from the Gulf of Mexico. Front. Mar. Sci. 4, 111. https://doi.org/10.3389/fmars.2017.00111.

Kutti, T., Bergstad, O.A., Fosså, J.H., Helle, K., 2014. Cold-water coral mounds and sponge-beds as habitats for demersal fish on the Norwegian shelf. Deep-Sea Res. Pt. II 99, 122–133.

Lagasse, C.R., Knudby, A., Curtis, J., Finney, J.L., Cox, S.P., 2015. Spatial analyses reveal conservation benefits for cold-water corals and sponges from small changes in a trawl fishery footprint. Mar. Ecol. Prog. Ser. 528, 161–172.

Leathwick, J., Moilanen, A., Francis, M., Elith, J., Taylor, P., Julian, K., Hastie, T., Duffy, C., 2008. Novel methods for the design and evaluation of marine protected areas in offshore waters. Conserv. Lett. 1 (2), 91–102.

Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74 (6), 1659–1673.

Lennon, J.J., 2000. Red-shifts and red herrings in geographical ecology. Ecography 23 (1), 101–113.

Leys, S.P., Wilson, K., Holeton, C., Reiswig, H.M., Austin, W.C., Tunnicliffe, V., 2004. Patterns of glass sponge (Porifera, Hexactinellida) distribution in coastal waters of British Columbia, Canada. Mar. Ecol. Prog. Ser. 283, 133–149.

Liaw, A., Wiener, M., 2002. Classification and regression by random forest. R News 2 (3), 18–22.

Lichstein, J.W., Simons, T.R., Shriner, S.A., Franzreb, K.E., 2002. Spatial autocorrelation and autoregressive models in ecology. Ecol. Monogr. 72 (3), 445–463.

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. In: Levitus, S., Mishonov, A. (Eds.), NOAA Atlas NESDIS 73, pp. 40 pp. http://data.nodc.noaa.gov/woa/WOA13/DOC/woa13_vol1.pdf.

Lunden, J.J., Georgian, S.E., Cordes, E.E., 2013. Aragonite saturation states at cold-water coral reefs structured by *Lophelia pertusa* in the northern Gulf of Mexico. Limnol. Oceanogr. 58 (1), 354–362.

Lunden, J.J., McNicholl, C.G., Sears, C.R., Morrison, C.L., Cordes, E.E., 2014. Acute survivorship of the deep-sea coral *Lophelia pertusa* from the Gulf of Mexico under acidification, warming, and deoxygenation. Front. Mar. Sci. 1, 78. https://doi.org/10.3389/fmars.2014.00078.

Lutz, M.J., Caldeira, K., Dunbar, R.B., Behrenfeld, M.J., 2007. Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. J. Geophys. Res. 112, C10011. https://doi.org/10.1029/2006JC003706.

Mackay, K.A., Mitchell, J.S., Neil, H.L., Mackay, E.J., 2015. New Zealand's Marine Realm, NIWA Chart, Miscellaneous Series No. 94. [dataset].

Masson, D.G., Bett, B.J., Billett, D.S.M., Jacobs, C.L., Wheeler, A.J., Wynn, R.B., 2003. The origin of deep-water, coral-topped mounds in the northern Rockall Trough, Northeast Atlantic. Mar. Geol. 194 (3), 159–180.

McCulloch, M., Trotter, J., Montagna, P., Falter, J., Dunbar, R., Freiwald, A., Försterra, G., López Correa, M., Maier, C., Rüggeberg, A., Taviani, M., 2012. Resilience of cold-water scleractinian corals to ocean acidification: boron isotopic systematics of pH and saturation state up-regulation. Geochim. Cosmochim. Acta 87, 21–34.

Mora, C., Wei, C., Rollo, A., Amaro, T., Baco, A.R., Billett, D., Bopp, L., Chen, C., Collier, M., Danovaro, R., Gooday, A.J., Grupe, B.M., Halloran, P.R., Ingels, J., Jones, D.O.B., Levin, L.A., Nakano, H., Norling, K., Ramirez-Llodra, E., Rex, M., Ruhl, R.A., Smith, C.R., Sweetman, A.K., Thurber, A.R., Tjiputra, J.F., Usseglio, P., Watling, L., Wu, T., Yasuhara, M., 2013. Biotic and human vulnerability to projected changes in ocean biogeochemistry over the 21st century. PLoS Biol. 11, e1001682. https://doi.org/10.1371/journal.pbio.1001682.

Moran, P., 1948. The interpretation of statistical maps. J. R. Stat. Soc. Ser. B Methodol. 10 (2), 243–251.

Naumann, M.S., Orejas, C., Ferrier-Pagès, C., 2014. Species-specific physiological response by the cold-water corals *Lophelia pertusa* and *Madrepora oculata* to variations within their natural temperature range. Deep-Sea Res. Pt. II 99, 36–41.

Oppel, S., Meirinho, A., Ramírez, I., Gardner, B., O'Connell, A.F., Miller, P.I., Louzao, M., 2012. Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. Biol. Conserv. 156, 94–104.

Orejas, C., Gori, A., Lo Iacono, C., Puig, P., Gili, J.M., Dale, M.R., 2009. Cold-water corals in the Cap de Creus canyon, northwestern Mediterranean: spatial distribution, density and anthropogenic impact. Mar. Ecol. Prog. Ser. 397, 37–51.

Palanques, A., Guillén, J., Puig, P., 2001. Impact of bottom trawling on water turbidity and muddy sediment of an unfished continental shelf. Limnol. Oceanogr. 46 (5), 1100–1110.

Parker, S.J., Penney, A.J., Clark, M.R., 2009. Detection criteria for managing trawl impacts on vulnerable marine ecosystems in high seas fisheries of the South Pacific

Ocean. Mar. Ecol. Prog. Ser. 397, 309–317.

Penney, A.J., Parker, S.J., Brown, J.H., 2009. Protection measures implemented by New Zealand for vulnerable marine ecosystems in the South Pacific Ocean. Mar. Ecol. Prog. Ser. 397, 341–354.

Penney, A.J., Guinotte, J.M., 2013. Evaluation of New Zealand's high-seas bottom trawl closures using predictive habitat models and quantitative risk assessment. PLoS One 8, e82273. https://doi.org/10.1371/journal.pone.0082273.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190 (3–4), 231–259.

Phillips, S.J., Dudík, 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. Ecol. Appl. 19 (1), 181–197.

Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 1 (2), 161–175.

Piepenburg, D., Müller, B., 2004. Distribution of epibenthic communities on the Great Meteor Seamount (North-East Atlantic) mirrors pelagic processes. Arch. Fish. Mar. Res. 51, 55–70.

Puig, P., Canals, M., Martín, J., Amblas, D., Lastras, G., Palanques, A., Calafat, A.M., 2012. Ploughing the deep sea floor. Nature 489 (7415), 286–289.

Purser, A., Larsson, A.I., Thomsen, L., van Oevelen, D., 2010. The influence of flow velocity and food concentration on Lophelia pertusa (Scleractinia) zooplankton capture rates. J. Exp. Mar. Biol. Ecol. 395 (1–2), 55–62.

Pusceddu, A., Bianchelli, S., Martín, J., Puig, P., Palanques, A., Masqué, P., Danovaro, R., 2014. Chronic and intensive bottom trawling impairs deep-sea biodiversity and ecosystem functioning. Proc. Natl. Acad. Sci. 111 (24), 8861–8866.

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. 15, 4123–4140.

Ramirez-Llodra, E., Tyler, P.A., Baker, M.C., Bergstad, O.A., Clark, M.R., Escobar, E., Levin, L.A., Menot, L., Rowden, A.A., Smith, C.R., Van Dover, C.L., 2011. Man and the last great wilderness: human impact on the deep sea. PLoS One 6 (8), e22588. https://doi.org/10.1371/journal.pone.0022588.

Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M., Guisan, A., 2006. Are niche-based species distribution models transferable in space? J. Biogeogr. 33 (10), 1689–1703.

Reiss, H., Cunze, S., König, K., Neumann, H., Kröncke, I., 2011. Species distribution modelling of marine benthos: a North Sea case study. Mar. Ecol. Prog. Ser. 442, 71–86.

Rengstorf, A.M., Grehan, A., Yesson, C., Brown, C., 2012. Towards high-resolution habitat suitability modeling of vulnerable marine ecosystems in the deep-sea: resolving terrain attribute dependencies. Mar. Geod. 35 (4), 343–361.

Robert, K., Jones, D.O., Roberts, J.M., Huvenne, V.A., 2016. Improving predictive mapping of deep-water habitats: considering multiple model outputs and ensemble techniques. Deep-Sea Res. Pt. I 113, 80–89.

Roberts, J.M., Brown, C.J., Long, D., Bates, C.R., 2005. Acoustic mapping using a multibeam echosounder reveals cold-water coral reefs and surrounding habitats. Coral Reefs 24, 654–669.

Roberts, J.M., Wheeler, A.J., Freiwald, A., Cairns, S.D., 2009. Cold-Water Corals: The Biology and Geology of Deep-Sea Coral Habitats. Cambridge University Press, Cambridge.

Rogers, A.D., 1994. The biology of seamounts. Adv. Mar. Biol. 30, 305–350.

Rooper, C.N., Zimmermann, M., Prescott, M.M., Hermann, A.J., 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.

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 Res. Pt. I 126, 148–161.

Ross, R.E., Howell, K.L., 2012. Use of predictive habitat modelling to assess the distribution and extent of the current protection of 'listed' deep-sea habitats. Divers. Distrib. 19, 433–445.

Roux, M.J., Doonan, I., Edwards, C., Clark, M.R., 2017. Low Information Stock Assessment of Orange Roughy Hoplostethus atlanticus in the South Pacific Regional Fisheries Management Organisation Convention Area. New Zealand Fisheries Assessment Report 2017/01. pp. 62 pp. https://fs.fish.govt.nz/Doc/24213/FAR-2017-01-Orange-roughy-SPRFMO-area.pdf.ashx.

Rowden, A.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. Front. Mar. Sci. 4, 335. https://doi.org/10.3389/fmars.2017.00335.

Rowden, A.A., Oliver, M., Clark, M.R., Mackay, K., 2008. New Zealand's 'SEAMOUNT' Database: Recent Updates and Its Potential Use for Ecological Risk Assessment. New Zealand Aquatic Environment and Biodiversity Report 27. New Zealand Ministry of Fisheries, Wellington, NZ. https://fs.fish.govt.nz/Doc/21591/AEBR_27.pdf.ashx.

Sampaio, I., Braga-Henriques, A., Pham, C., Ocaña, O., De Matos, V., Morato, T., Porteiro,

F.M., 2012. Cold-water corals landed by bottom longline fisheries in the Azores (north-eastern Atlantic). J. Mar. Biol. Assoc. U. K. 92 (7), 1547–1555.

Segurado, P., Araújo, M.B., 2004. An evaluation of methods for modelling species distributions. J. Biogeogr. 31 (10), 1555–1568.

Thiem, Ø., Ravagnan, E., Fosså, J.H., Berntsen, J., 2006. Food supply mechanisms for cold-water corals along a continental shelf edge. J. Mar. Syst. 60 (3-4), 207–219.

Thresher, R.E., Tilbrook, B., Fallon, S., Wilson, N.C., Adkins, J., 2011. Effects of chronic low carbonate saturation levels on the distribution, growth and skeletal chemistry of deep-sea corals and other seamount megabenthos. Mar. Ecol. Prog. Ser. 442, 87–99.

Tittensor, D.P., Baco, A.R., Brewin, P.E., Clark, M.R., Consalvey, M., Hall-Spencer, J., Rowden, A.A., Schlacher, T., Stocks, K.I., Rogers, A.D., 2009. Predicting global habitat suitability for stony corals on seamounts. J. Biogeogr. 36 (6), 1111–1128.

Tittensor, D.P., Baco, A.R., Hall-Spencer, J.M., Orr, J.C., Rogers, A.D., 2010. Seamounts as refuge from ocean acidification for cold-water corals. Mar. Ecol. 31 (1), 212–225.

Tognelli, M.F., Kelt, D.A., 2004. Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. Ecography 27 (4), 427–436.

Tong, R., Purser, A., Guinan, J., Unnithan, V., 2013. Modeling the habitat suitability for deep-water gorgonian corals based on terrain variables. Ecol. Inform. 13, 123–132.

Tracey, D.M., Rowden, A.A., Mackay, K.A., Compton, T., 2011. Habitat-forming cold-water corals show affinity for seamounts in the New Zealand region. Mar. Ecol. Prog. Ser. 430, 1–22.

Turley, C.M., Roberts, J.M., Guinotte, J.M., 2007. Corals in deep-water: will the unseen hand of ocean acidification destroy cold-water ecosystems? Coral Reefs 26 (3), 445–448.

UNGA, 2007. Resolution 61/105. Sustainable Fisheries, Including Through the 1995 Agreement for the Implementation of the Provisions of the United Nations Convention on the Law of the Sea of 10 December 1982 Relating to the Conservation and Management of Straddling Fish Stocks and Highly Migratory Fish Stocks, and Related Instruments. United Nations General Assembly (UNGA) A/RES/61/105. www.un.org/Depts/los/general_assembly/general_assembly_resolutions.htm.

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

Van Dover, C.L., Aronson, J., Pendleton, L., Smith, S., Arnaud-Haond, S., Moreno-Mateos, D., Barbier, E., Billett, D., Bowers, K., Danovaro, R., Edwards, A., 2014. Ecological restoration in the deep sea: desiderata. Mar. Policy 44, 98–106.

Vierod, A.D., Guinotte, J.M., Davies, A.J., 2014. Predicting the distribution of vulnerable marine ecosystems in the deep sea using presence-background models. Deep-Sea Res. Pt. II 99, 6–18.

Wagner, H., Purser, A., Thomsen, L., Jesus, C.C., Lundälv, T., 2011. Particulate organic matter fluxes and hydrodynamics at the Tisler cold-water coral reef. J. Mar. Syst. 85, 19–29.

Watling, L., Norse, E.A., 1998. Disturbance of the seabed by mobile fishing gear: a comparison to forest clearcutting. Conserv. Biol. 12 (6), 1180–1197.

Westberry, T., Behrenfeld, M.J., Siegel, D.A., Boss, E., 2008. Carbon-based primary productivity modeling with vertically resolved photoacclimation. Glob. Biogeochem. Cycle 22 (2), GB2024. https://doi.org/10.1029/2007GB003078.

Wiley, E.O., McNyset, K.M., Peterson, A.T., Robins, C.R., Stewart, A.M., 2003. Niche modeling and geographic range predictions in the marine environment using a machine learning algorithm. Oceanography 16 (3), 120–127.

Williams, A., Schlacher, T.A., Rowden, A.A., Althaus, F., Clark, M.R., Bowden, D.A., Stewart, R., Bax, N.J., Consalvey, M., Kloser, R.J., 2010. Seamount megabenthic assemblages fail to recover from trawling impacts. Mar. Ecol. 31 (s1), 183–199.

Wisshak, M., Freiwald, A., Lundälv, T., Gektidis, M., 2005. The physical niche of the bathyal lophelia pertusa in a non-bathyal setting: environmental controls and palaeoecological implications. In: Freiwald, A., Roberts, J.M. (Eds.), Cold-Water Corals and Ecosystems. Erlangen Earth Conference Series. Springer, Berlin, Heidelberg pp 979–1001.

Wright, D.J., Pendleton, M., Boulware, J., Walbridge, S., Gerlt, B., Eslinger, D., Sampson, D., Huntley, E., 2012. ArcGISBenthic Terrain Modeler (BTM), v. 3.0, Environmental Systems Research Institute. NOAA Coastal Services Center, Massachusetts Office of Coastal Zone Management. http://esriurl.com/5754.

Yesson, C., Bedford, F., Rogers, A.D., Taylor, M.L., 2017. The global distribution of deep-water Antipatharia habitat. Deep-Sea Res. Pt. II 45, 79–86.

Yesson, C., Clark, M.R., Taylor, M.L., Rogers, A.D., 2011. The global distribution of seamounts based on 30 arc seconds bathymetry data. Deep-Sea Res. Pt. I 58 (4), 442–453.

Yesson, C., Taylor, M.L., Tittensor, D.P., Davies, A.J., Guinotte, J., Baco, A., Black, J., Hall-Spencer, J.M., Rogers, A.D., 2012. Global habitat suitability of cold-water octocorals. J. Biogeogr. 39 (7), 1278–1292.

Zweng, M.M., Reagan, J.R., Antonov, J.I., Locarnini, R.A., Mishonov, A.V., Boyer, T.P., Garcia, H.E., Baranova, O.K., Johnson, D.R., Seidov, D., Biddle, M.M., 2013. World Ocean atlas 2013, volume 2: salinity. In: Levitus, S., Mishonov, A. (Eds.), NOAA Atlas NESDIS 74, pp. 39 pp. http://data.nodc.noaa.gov/woa/WOA13/DOC/woa13_vol2.pdf.