Identification of protected coral hotspots using species distribution modelling

Report prepared for Project POP2021-02, Conservation Services Programme, Department of Conservation

August 2023

Prepared by:
Owen Anderson, Kareen Schnabel, David Bowden, Niki Davey, Alan Hart

For any information regarding this report please contact:

Owen Anderson
Group Manager, Deepsea Ecology and Fisheries
NIWA
Wellington

+64 4 386 0863

National Institute of Water & Atmospheric Research Ltd
Private Bag 14901
Kilbirnie
Wellington 6241

Phone +64 4 386 0300

Conservation Services Programme: csp@doc.govt.nz

NIWA CLIENT REPORT No: 2023230WN
Report date: August 2023
NIWA Project: DOC21304

Revision	Description	Date
Version 1.0	Final version sent to client	18 August 2023

Quality Assurance Statement		
	Reviewed by:	Judy Sutherland
	Formatting checked by:	Jess Moffat
	Approved for release by:	Alison MacDiarmid

Contents

Tables

Figures

Executive summary

A broad range of corals are protected in New Zealand seas by legislation that prohibits intentional damage and removal, and several protected areas have been designated in which bottom-contact fishing methods such as bottom trawling are prohibited. To understand the degree to which the distributions of these corals overlap with bottom trawling, and therefore to begin to assess the effectiveness of the current network of protected areas, it is necessary to have reliable maps of species distributions for these corals and identify key areas of high abundance across all taxa.

This report presents maps of abundance within the New Zealand region for eleven coral taxa, based on species distribution modelling using abundance values measured at 949 sample sites from image data collected by towed camera or remotely operated vehicle (ROV) systems. Most of these sites were sampled using NIWAs Deep-towed Imaging System (DTIS) and abundance values were based on archived analyses of video data. Further analysis of video data from three surveys was undertaken during this study to provide abundance values for selected locations not covered by archived data.

The species distribution models were constructed as an ensemble of predictions using two tree-based methods, Random Forests and Boosted Regression Trees. For each method abundance was estimated using a hurdle model – the combination of a presence-absence model predicting probability of presence and a regression model predicting abundance at locations of species presence. Uncertainty in the predictions was estimated using a bootstrap resampling technique to measure variability in predictions among 100 models built from random subsets of the sample data for each taxon.

The environmental predictor variables offered to the models represented a combination of seafloor characteristics, water chemistry, and productivity, tailored slightly for each taxon according to the chemical composition of their skeleton. In addition, and for the first time in any New Zealand species distribution modelling study, a date-specific trawl-contact variable representing bottom-trawling on the seafloor prior to the sample date was also included, based on compilation of recorded fishing effort in the region since 1990.

Models were successfully constructed for 11 of the 15 taxa considered, with observations of the remaining 4 taxa proving too rare in the data for models to be produced. Model performance metrics indicated good fits to the input data and a high level of agreement between observed and predicted values; these were backed up by metrics based on comparison of predicted values with values from data withheld from the model in each bootstrap sample.

A map of overall abundance for the set of protected corals modelled was made by adding abundance estimates across all taxa. Hotspots of coral abundance, nominally defined as areas where overall abundance was predicted to be over 40 individuals per 1000 m^2, represented about 2% of the modelled area (the area of the seafloor between 50 m and 3000 m). These hotspots were spread widely around the region, most notably on the Norfolk Ridges, Challenger Plateau, Macquarie Ridge, Eastern Chatham Rise, and Kermadec Ridge.

The influence of the fishing impact variable was low in models for most taxa. For two taxa where it was important, *Radicipes* spp. and *Goniocorella dumosa*, maps were constructed to estimate pre-fishing distributions. However, these proved to differ very little from those predicting current distributions, probably because of insufficient temporal spread in the coral sample data for models to detect negative effects from fishing disturbance (i.e., most historical fishing effort and impacts precede collection of coral data in certain areas).

1 Introduction

Corals (Phylum Cnidaria) constitute a prominent component of species bycatch in bottom-contact trawl fisheries in New Zealand. Despite most of the species caught being afforded legal protection under the Wildlife Act 1953 (which prohibits intentional damage, and retention of accidental catch), there is limited knowledge of their spatial distributions, particularly in relation to where they occur at highest densities and how bottom-contact fishing affects their abundance and diversity.

Deep-sea corals are a structurally and ecologically important component of deep-sea habitats in the New Zealand region and beyond, and most species are highly fragile and therefore vulnerable to physical disturbance from human activities that impact the seafloor; primarily bottom-contact trawling but also potentially seabed mining. The occurrence of protected coral species in fisheries bycatch has been examined in several studies (e.g., Baird et al. 2013, Anderson et al. 2014, Macpherson et al. 2022) providing strong evidence for the substantial overlap between coral habitat and fishing grounds.

Core strategic objectives of the Conservation Services Programme (CSP) of the Department of Conservation (DOC) emphasise the need for data to support evaluation of adverse effects of commercial fishing on protected species, not only to inform their status but also to ensure adequate mitigation measures are in place. To achieve these objectives, it is imperative to understand where these protected species live, and particularly to identify areas where they occur in high abundance and are therefore especially vulnerable. Several recent studies have used species distribution modelling tools to predict suitable habitat for a range of coral species within the New Zealand region and wider South Pacific Ocean (Tracey et al. 2011, Baird et al. 2013, Anderson et al. 2014, 2015, 2016a, 2016b, 2020, 2022, Georgian et al. 2019, Stephenson et al. 2021a) but only two studies (both based on DTIS image data) have attempted to predict the abundance of any coral species (Rowden et al. 2017, Bowden et al. 2019), and these were restricted to relatively small spatial extents within the New Zealand region.

The objective of the study undertaken here was to assemble image-based coral abundance and absence data for incorporation into species distribution models that can predict abundance levels and thus help identify areas likely to be of high conservation value (‘hotspots’) for protected corals within the New Zealand Exclusive Economic Zone and out to the edges of the wider New Zealand Region (as defined by Mackay et al. 2015). As the first component of the project involves collation and analysis of new seabed imagery data to inform the models, the project also serves as an audit of data sources for image-based coral research.

The protected coral taxa modelled here comprise a combination of individual species and groups of species at the Genus, Family, and Order level within the four protected groups: black corals (Antipatharia), gorgonian alcyonaceans (a subset of Alcyonacea), stony corals (Scleractinia), and stylasterids hydrocorals (Stylasteridae). In addition to these, the octocoral sea pens (Order Pennatulacea) were included opportunistically as they are a cnidarian group that is considered a vulnerable marine ecosystem (VME) indicator taxon (FAO 2009, Parker et al 2009, Parker & Bowden 2010), includes biogenic habitat formers (Anderson et al. 2018), and is consistently identifiable in seafloor imagery. Although they are not protected under the Wildlife Act 1953, sea pen fields (see distribution models by Stephenson et al 2021a) are considered a Sensitive Environment under Schedule 6 of the Exclusive Economic Zone and Continental Shelf Regulations (2013).

Outputs from the models produced can be used in risk assessments and to inform management strategies that consider ecological processes, coral biology, and the current and historical impact of fishing on ecosystem services provided by deep-sea corals and other taxa.

1.1 Project Objectives

1. To collate and analyse cold water coral records from existing seabed towed camera transects in the New Zealand region, with a particular focus on areas known to overlap with fishing effort.
2. To identify hotspots of abundance for selected protected coral species in the New Zealand region using predictions from abundance-based species distribution models.
3. To better understand the historical effects of fishing on observed patterns of coral distribution and relative abundances.

2 Methods

2.1 Image-based abundance data collation

2.1.1 Deep-towed Imaging System (DTIS) data

Seabed photographic and video transects have been carried out by NIWA using its Deep Towed Imaging System (DTIS) to record data on the type of seabed substrate and on the presence, number (count), and identification of benthic invertebrates, algae, and demersal and benthic fish species, since 2006. Standard DTIS transects are of 60 minutes duration at the seabed at a target speed of between 0.25 and 0.5 metres per second (m/s), resulting in average transect lengths of 1000–1400 m. The target altitude of the camera is 2–3 metres (m) above the seabed, and both video and still camera frames include two equidistant laser points (20 cm apart) as a reference, thus allowing for size and area estimates (Hill 2009, Bowden & Jones 2016). During each DTIS deployment the real-time low-resolution video feed was monitored by biologists and all conspicuous organisms and changes in substratum type were recorded as spatially referenced observations using the software Ocean Floor Observation Protocol (OFOP, www.emma-technologies.com). This program automatically logs the time, spatial co-ordinates, depth, and a range of other parameters associated with each observation, every still photograph taken during the transect, and the start and end point of the video transect.

The seafloor observations logged into the OFOP software at-sea and in real time (referred to as ship-board), are conducted at a coarse taxonomic level and are not typically suitable for quantitative analyses but are a useful indicator of a presence of generic taxa including protected corals and the archived files can be used to produce an initial summary of the fauna observed during a survey.

Video and/or still images are later analysed in much finer detail post-voyage (referred to as re-runs) by examining the high-definition files collected by DTIS/OFOP and therefore include more accurate counts and an improved taxonomic identification – typically in collaboration with NIWA and overseas taxonomic experts. Detailed methodology for these procedures is described in Bowden & Hewitt (2012) and Jones et al. (2018)

2.1.2 Examination and selection of available image data from DTIS and other sources

Using NIWA's DTIS shipboard benthic fauna observations database, we interrogated the distribution of DTIS transect stations deployed throughout the region between 2006 and 2020, from NIWA Research Vessels (RV) Tangaroa and/or Kaharoa, and the level of subsequent detailed faunal analyses conducted.

The locations of all video transects with full, partial, or no post-survey analysis of species abundance were mapped and summarised to inform a gap analysis. Data from transects with full post-processing were re-formatted to match existing taxonomic divisions of a dataset produced for an earlier analysis of benthic species abundance on the Chatham Rise and Campbell Plateau (Bowden et al. 2019, Stephenson et al. 2021b). These standardised taxonomic divisions were necessary because although identifications were often made to genus or species level, or as operational taxonomic units (OTUs), which were not used consistently within and between surveys. Of the surveys where post-processing of video data had not been undertaken, and in discussion of the gap analysis and location of surveys in relation to commercial fishing grounds with DOC CSP staff, transects in key areas from three surveys were selected for analysis and processing by experts to the required standard and format for model input. In addition, post-processed DTIS data became available for two further surveys (59 transects) from two concurrent studies by NIWA for DOC (voyages TAN0906 and TAN1612 under project BIO183, Bowden et al. 2022). These too were formatted as necessary to meet the requirements for inclusion in coral abundance models.

In addition to these video data analyses, post-processed still image data was available for two Chatham Rise surveys (CRP2012 and TAN1503), and these were also processed into a format suitable for inclusion in the models (Figure 3-1).

2.2 Selection of coral taxa for modelling and aggregation of DTIS data

The number of DTIS transect in which an individual taxon was recorded varied widely among taxa, from over 400 transects for two taxa that combined multiple species (gorgonian alcyonaceans, and pennatulaceans) down to 18–21 transects for individual reef-forming species (*Goniocorella dumosa* and *Madrepora oculata*) (Table 2-1). Species distribution models can be built based on relatively few presence/absence records if these records represent sufficient coverage of the species' tolerance range for each environmental parameter considered, but this can be difficult to assess. Therefore, consideration must be given to the associated spatial uncertainty (the uncertainty associated with the estimate of abundance in each cell of the map grid) when assessing model performance and the reliability of maps of abundance, as well as considering the number of presence records and their spatial distribution. In addition, models such as those used here may fail due to insufficient information to parameterise the model, especially when sub-sampling for estimation of model uncertainty and performance.

Table 2-1: Protected coral taxa modelled, and number of DTIS transects with positive abundance available for use in abundance models. Shading used to distinguish taxonomic groupings.

Order	Taxon	Description	Number of DTIS transects
Scleractinia	*Solenosmilia variabilis*	Reef-forming coral	108
	Goniocorella dumosa	Reef-forming coral	73
	Enallopsammia rostrata	Reef-forming coral	21
	Madrepora oculata	Reef-forming coral	18
	Caryophylliidae	Solitary coral	228
	Stephanocyathus spp.	Solitary coral	41
	Flabellum spp.	Solitary coral	301
Gorgonian alcyonaceans	All*	Sea-fans/sea-whips/bamboo & bubble-gum corals	405
	Keratoisididae/Mopseidae	Bamboo corals	144
	Primnoidae	Sea-fans and bottle brush corals	196
	Paragorgiidae	Bubble-gum corals	25
	Radicipes spp.	Sea whips	113
Antipatharia	Antipatharia (all)	Black corals	221
Anthoathecata	Stylasteridae	Hydrocorals	232
Pennatulacea	Pennatulacea	Sea-pens	415

*Includes the three other alcyonacean taxa in the list as well as any other species outside of these groups

2.3 Environmental variables

An initial set of environmental predictor variables was selected, based on those used in recent habitat suitability modelling of corals and other VME taxa in the South Pacific (Georgian et al. 2019). In that study, variables were selected from a larger set based on model fitting procedures that maximised model fit while avoiding using too many, or highly correlated, variables. These predictor variables were gridded at a 1 km^2 resolution and represent a combination of seafloor characteristics, water chemistry and productivity, and fishing impacts.

For this study, a core subset of these predictor variables was tested in initial models for all taxa (Table 2-2). This subset varied among taxa in that aragonite saturation state was only included for those taxa which incorporate the aragonite form of calcium carbonate into their skeletal materials (i.e., scleractinian taxa *Solenosmilia variabilis, Goniocorella dumosa*, Caryophylliidae, *Flabellum* sp., and Stylasteridae), and calcite saturation state was only included for taxa that incorporate calcite in their skeletons (Gorgonians, Keratoisididae/Mopseidae, Primnoidae, *Radicipes* spp., and Antipatharia).

The impacts from mobile bottom fishing gear, especially bottom trawls, can be severe for fragile invertebrates such as corals (Clark et al. 2019) but few studies have attempted to account for this factor when predicting species distributions within heavily fished regions. For this study we produced a predictor layer representing the degree of fishing in each 1x1 km cell of the study grid as the total accumulated swept area ratio (SAR) of bottom trawls between 1990 and 2020. The construction of this layer was based on commercial fisheries data extracts and analyses undertaken for a FNZ "Fishing footprint" assessment of the New Zealand EEZ (Baird & Mules 2021), refined to incorporate a broader range of vessel types and gridded to the finer 1 km^2 resolution (see Rowden et al. in press for full details of the procedure). For the New Zealand region outside of its EEZ the necessary data were obtained from an assessment of fishing impacts on VME taxa in the SPRFMO region of the South Pacific (SPRFMO 2020) and re-analysed to match the methods for the within-EEZ data analysis. Data from the two sources were then combined to provide estimates of SAR for the entire New Zealand region, then separate layers produced for each year for which DTIS data were available to provide estimates of the level of fishing effort up until the date of the survey. In this way each DTIS record could be assigned a level of fishing effort that represented the historic benthic fishing impact relevant to the date of sampling.

When fitting the final models, a data layer of fishing effort representing all historical bottom trawling was used, in order to predict the current abundance of corals, but it was also possible to estimate the distribution of corals prior to any recorded bottom trawling by fitting the model to a fishing effort grid in which all values were set to zero, as long as there was a meaningful influence of the SAR variable in the model parameters. The applicability of such a predictor variable in these Species Distribution Models (SDMs) is only possible due to the availability of real absence information from the image data. If pseudo-absence data locations (typically randomly-generated background locations) were to be used, as in most SDM models of deep-sea fauna, no date could be assigned and therefore nor could an appropriate year-specific value for prior fishing effort.

Table 2-2: Environmental variables used in species distribution models. (This is the same set used by Georgian et al. (2019), except for updates of particulate organic carbon export and seamount, and the addition of fishing effort). Shading used to distinguish variable classes.

Variable	Units	Source
Seafloor Characteristics		
Percent gravel	%	Bostock et al. (2018a, 2018b)
Percent mud	%	Bostock et al. (2018a, 2018b)
Ruggedness[1] (VRM)	–	Derived from bathymetry (Mackay et al., 2015)
Slope SD[1]	–	Derived from bathymetry
Bathymetric Position Index – broad	–	Derived from bathymetry
Seamounts	–	Clark et al. 2022
Water Chemistry		
Aragonite saturation state at depth	–	Bostock et al. (2013)
Calcite saturation state at depth	–	Bostock et al. (2013)
Dissolved oxygen at depth	ml l^{-1}	Garcia et al. (2013)
Temperature at depth	°C	Locarnini et al. (2013)
Productivity		
Particulate organic carbon export	mg C m^{-2} d^{-1}	Following methods described in Pinkerton et al. (2016) and using data from Cael et al. (2018).
Human impacts		
Fishing effort	Fraction (0–1)	Rowden et al (in press)

[1]-Terrain metrics calculated using window sizes of 5x5 cells.

2.4 Abundance modelling

2.4.1 Model description

The input data, models, and predictions of protected coral distributions were constrained to the New Zealand region (Mackay et al. 2015), an area bounded by 157°E, 167°W, 24°S, 57.5°S and depths of 50–3000 m (Figure 3-1).

Models predicting the abundances of individual protected coral taxa were based on an ensemble of the outputs from two decision tree-based modelling methods, Boosted Regression Trees (BRT) and Random Forests (RF), and generally followed the approach of Bowden et al. (2019), with all analyses conducted using the statistical computing software R (R Core Team 2022).

Tree-based methods such as BRT and RF have the advantage over traditional methods that they can easily handle missing data, outliers, categorical as well as continuous variables, and automatically handle interactions between predictors (Elith et al. 2008), and these methods have been shown to be a useful combination for similar modelling efforts in the region (e.g., Anderson et al. 2022, Bowden et al. 2019, Stephenson et al. 2021a, 2021b).

RF modelling is a non-parametric approach (i.e., one that does not rely on any specific assumptions about underlying distribution of the input data) which builds classification or regression trees using random subsets of the input data (Breiman 2001). All models were tuned using the *tuneRF* function in the *extendedForest* R package to select optimal values for complexity parameter *mtry* (the number of variables used in each tree node), with *ntree* (the number of trees to grow) initially set to 1000.

BRT modelling is an advanced form of additive regression based on decision trees, where the individual terms of the regression are simple trees, fitted in a stage-wise manner. Simple (short) trees are formed by relating a response to repetitive binary splits of the data, then combined (boosted) to improve predictive power by focussing each successive tree on model residuals. The BRT models were run with the tree-complexity (number of splits) set to 1, and the learning rate (which determines the weight given to each successive tree in the model) adjusted so that the number of trees in the final models exceeded 1500 (presence-absence models) or 500 (abundance models), following guidelines in Elith et al. (2008). Stochasticity is incorporated into BRT models by selecting at random only a portion (the bag-fraction) of the data to use at each step in the model. We used a value of 0.6 in all presence-absence models, but higher values were occasionally required in the abundance models to achieve model convergence for some compilations of resampled data.

For each model method (BRT and RF) two separate component models were constructed: 1) a presence-absence model based on a binary logistic regression, which predicts probability of presence and 2) a regression model based only on the positive (i.e., non-zero) observations of species abundance, which predicts abundance at locations of species presence. For the regression model component, abundances were log(x+1) transformed to provide a near-normal distribution of the response, and the predictions back-transformed (exp(x)-1). Final estimates of abundance were made by multiplying the probabilities from the first component by the abundances from the second component.

The predictor variables used in each model were selected from an initial set of variables tailored slightly to the specific requirements of each taxon (as described above), then the most influential and least correlated of those were determined using a process of conditional permutation through an initial presence-absence model run, using all available data, in the *extendedForest* extension of *randomForest* in R. This subset of variables was then used for all four component models (BRT/FR, presence-absence/abundance) for that taxon.

To estimate the uncertainty in model predictions in each cell, a bootstrap resampling technique was applied following the methods of Anderson et al. (2016b), Bowden et al. (2019), and Georgian et al. (2019). Applying this method, random samples of the input data were drawn with replacement and sets of presence-absence, abundance, and hurdle models constructed using the same settings as the original. Predictions of abundance were then made for each 1x1 km cell of the model extent. This process was repeated 100 times for each model type (BRT and RF) resulting in 100 estimates of abundance for each taxon in each cell. This then allows the model uncertainty in each cell to be calculated, as the standard deviation (SD) of the 100 bootstrap estimates of abundance.

To complete the process and provide a single model of abundance for each taxon, an ensemble of the outputs from the two hurdle models for each taxon was produced. This approach incorporates the predictions and underlying assumptions and modelling strategies of both methods (BRT and RF), thus limiting dependence on a single model type or structural assumption and enabling a more robust characterization of the predicted spatial variation and uncertainties (Robert et al. 2016). Ensemble models were constructed by taking weighted averages of the predictions of the two hurdle models, using methods adapted from Oppel et al. (2012), Anderson et al. (2016), Rowden et al. (2017), Bowden et al. (2019), and Georgian et al. (2019). This procedure uses a two-part weighting for the BRT and RF components of the ensemble model, taking equal contributions from the overall model performance and the uncertainty measure (SD) in each cell, as follows,

$$W_{BRT1} = \frac{MPS_{BRT}}{MPS_{BRT}+MPS_{RF}} \text{ and } W_{RF1} = \frac{MPS_{RF}}{MPS_{BRT}+MPS_{RF}}$$

$$W_{BRT2} = 1 - \frac{SD_{BRT}}{SD_{BRT}+SD_{RF}} \text{ and } W_{RF2} = 1 - \frac{SD_{RF}}{SD_{BRT}+SD_{RF}}$$

$$W_{BRT} = \frac{W_{BRT1}+W_{BRT2}}{2} \text{ and } W_{RF} = \frac{W_{RF1}+W_{RF2}}{2}$$

$$X_{ENS} = X_{BRT} * W_{BRT} + X_{RF} * W_{RF}$$

$$SD_{ENS} = \sqrt{\frac{(SD_{BRT} * X_{BRT})^2 * W_{BRT}^2 + (SD_{RF} * X_{RF})^2 * W_{RF}^2}{X_{ENS}^2}}$$

where MPS_{BRT} and MPS_{RF} are the hurdle model performance statistics; X_{BRT} and X_{RF} are the model predictions; and SD_{BRT} and SD_{RF} are the bootstrap SDs; and X_{ENS} and SD_{ENS} are the weighted ensemble predictions and weighted SDs, respectively, from which maps of predicted species distribution and model uncertainty were produced

2.4.2 Model performance

Performance of the binomial models was evaluated primarily using AUC, the Area Under the receiver operating characteristic Curve. AUC is calculated as the fraction of true positives versus the fraction of true negatives (comparing predictions with observations), where a score of 0.5 indicates a model with predictions no better than random, values over 0.7 indicate "adequate" performance, and values over 0.8 indicate "excellent" performance (Hosmer et al. 2013). In addition, the True Skill Statistic (TSS), a combination of specificity (% of absences correctly identified) and sensitivity (% of presences correctly identified) was calculated. This gives an index from −1 to +1, where +1 equals perfect agreement and −1 = no better than random, Allouche et al., 2006). As well as these performance-based statistics, a cut-off habitat suitability value for presence/absence was calculated as that which maximises sensitivity and specificity, useful for creating binary (i.e., presence or absence) model output.

For the abundance and hurdle models, performance was measured by calculating the Pearson correlation between predicted and observed values as well as the Normalized Root Mean Square Error (NRMSE), which measures deviation of predicted values from observed while allowing standardised comparison among models.

Each of these performance metrics was calculated separately for each of the 200 bootstrap model runs, and mean values then obtained. Metrics were calculated by comparing predictions with not only the data selected for each run (*training* data) but also with the data not selected (*evaluation* data), thus providing a test of the model against data independent of that used in the model itself.

2.5 Determination of coral hotspots

The core objective of this modelling exercise was to identify areas where there are predicted to be high concentrations of protected corals, "coral hotspots", within the New Zealand Region. To enable this, the level of predicted overall abundance for the protected corals modelled here was calculated by adding together the predicted abundance values in each cell for all of the coral models to produce an overall raster layer. This layer then represents, in each of its cells, the total estimated abundance (in numbers/1000m^2) of all 11 taxa modelled, including overlaps of taxa (i.e., Gorgonians (all) with separately modelled subsets of this group, Keratoisididae/Mopseidae, Primnoidae, Paragorgiidae and *Radicipes* spp.). Species richness (the number of taxa present) is not included in our definition of a hotspot as this information is not available from the models of combined taxa because taxonomic identifications based on video data do not allow sufficient resolution of the species diversity within taxa. Therefore, it is important to note that in this report a hotspot is not based upon a measure of coral diversity but on numbers of individuals of a given taxon. The hotspot layer was presented as a single map, with abundance shown in four categories defined to approximate low, medium, high, and hotspot locations.

3 Results

3.1 Selection and processing of video transect data

Examination of the DTIS image database, covering all DTIS transects between 2006–2020, revealed numerous biodiversity surveys for which post-survey analysis of image data was either incomplete or yet to be undertaken. Location data for these unprocessed transects surveys was examined and after consultation with DOC CSP, three surveys were selected for post-processing of video transects by NIWA benthic invertebrate experts, focussing on protected corals and (opportunistically) the non-protected sea pens but recording other taxa where possible, although at a somewhat coarser level. These surveys comprised one of seamounts on Macquarie Ridge in 2008 (TAN0803) and two biogenic habitat surveys that circumnavigated New Zealand in continental shelf depths in 2011 (TAN1105 & TAN1108) (Table 3-1, Figure 3-1). Due to resource constraints, not all transects in these three surveys could be processed, therefore transects were selected randomly, with a total of 44 transects ultimately processed:TAN110, 14 transects from a total of 74 available (60 remaining, 18 of which may have no coral); TAN 1108, 9 transects from a total of 119 available (110 remaining, 66 of which may have no coral); TAN0803, all 21 transects available were processed.

In addition to these transects, there were several surveys identified from the DTIS database that had fully post-processed video transect data but not used in the modelling of Bowden et al. (2019) and for which the data were not in a suitable format (i.e., arranged into compatible taxa) for modelling of coral taxa. These included a pre-drill survey (TAN1902) in the Great South Basin (off Southland) for an oil-prospecting company (OMV) who, during the course of this project, released all DTIS image data collected to NIWA for use in research activities, and nine other surveys covering a broad range of locations. Data from each of these surveys (Table 3-1) were therefore processed and formatted for inclusion in the dataset for modelling, providing a total of 438 additional sample sites.

Only one of the surveys in the final set collated for modelling was not undertaken using RV *Tangaroa*. This survey, CRP2012, on the Chatham Rise, used the RV *Dorado Discovery* (Odyssey Marine Explorations) and their *Zeus II* ROV to collect imagery. Processing however used OFOP as for all DTIS-based surveys. Additionally, analysis of image data from this survey, along with one RV *Tangaroa* survey (TAN1503 on the Graveyard seamounts of the Chatham Rise), was based on still images rather than video data.

Although sample sites were strongly concentrated on the Chatham Rise (Figure 3-1), the final dataset included a broad range of locations, covering many of the key geographical features of the New Zealand Region and additionally including: the Macquarie Ridge; the Challenger Plateau; the Louisville Seamount Chain; the east, south, and north coasts of the North Island; the West Norfolk Ridge, and the Kermadec Ridge. Areas notably absent from the final dataset include the eastern and southern regions of the Campbell Plateau, the South Island west coast, and the north-western quadrant of the region including the Lord Howe Rise and Norfolk Ridge.

Table 3-1: Location and year of voyages. Showing numbers of video transects assembled for this study including those with data already fully processed, data only requiring formatting of taxonomic categories and density measurements, and data from transects requiring full post-processing including formatting. Abundance estimation was derived from video camera data for all surveys except CRP2012 and TAN1503 (marked *) for which still image data was used.

Voyage code	Voyage location and year	Fully processed-previously	Formatted for this study	Fully processed-this study	Total
TAN0705	Chatham Rise 2007	108			108
TAN0707	Challenger Plateau 2007		46		46
TAN0803	Macquarie Ridge 2008			21	21
TAN0906	Northland Plateau 2009		38		38
TAN1004	Cook Strait and Hikurangi 2010		39		39
TAN1105	Cape Reinga, E coast N Island, N coast S Island 2011			14	14
TAN1108	E coast from Stewart Is. to East Cape 2011			9	9
TAN1206	Bay of Plenty 2012		61		61
CRP2012*	Chatham Rise 2012	39			39
TAN1306	Chatham Rise 2013	53			53
TAN1402	Louisville seamounts 2014		118		118
TAN1503*	Chatham Rise: Graveyard and Andes seamounts 2015	11			11
TAN1505	East Cape 2015		12		12
TAN1602	Stewart Island and Auckland Island shelves 2016	23			23
TAN1603	Project West, Northland 2016		11		11
TAN1612	Kermadec Islands 2016		21		21
TAN1701	Chatham Rise 2017	147			147
TAN1902	Great South Basin 2019		63		63
TAN1904	Cook Strait and Hikurangi 2019		29		29
TAN2004	Challenger Plateau 2020	86			86
All	All New Zealand Region	467	438	44	949

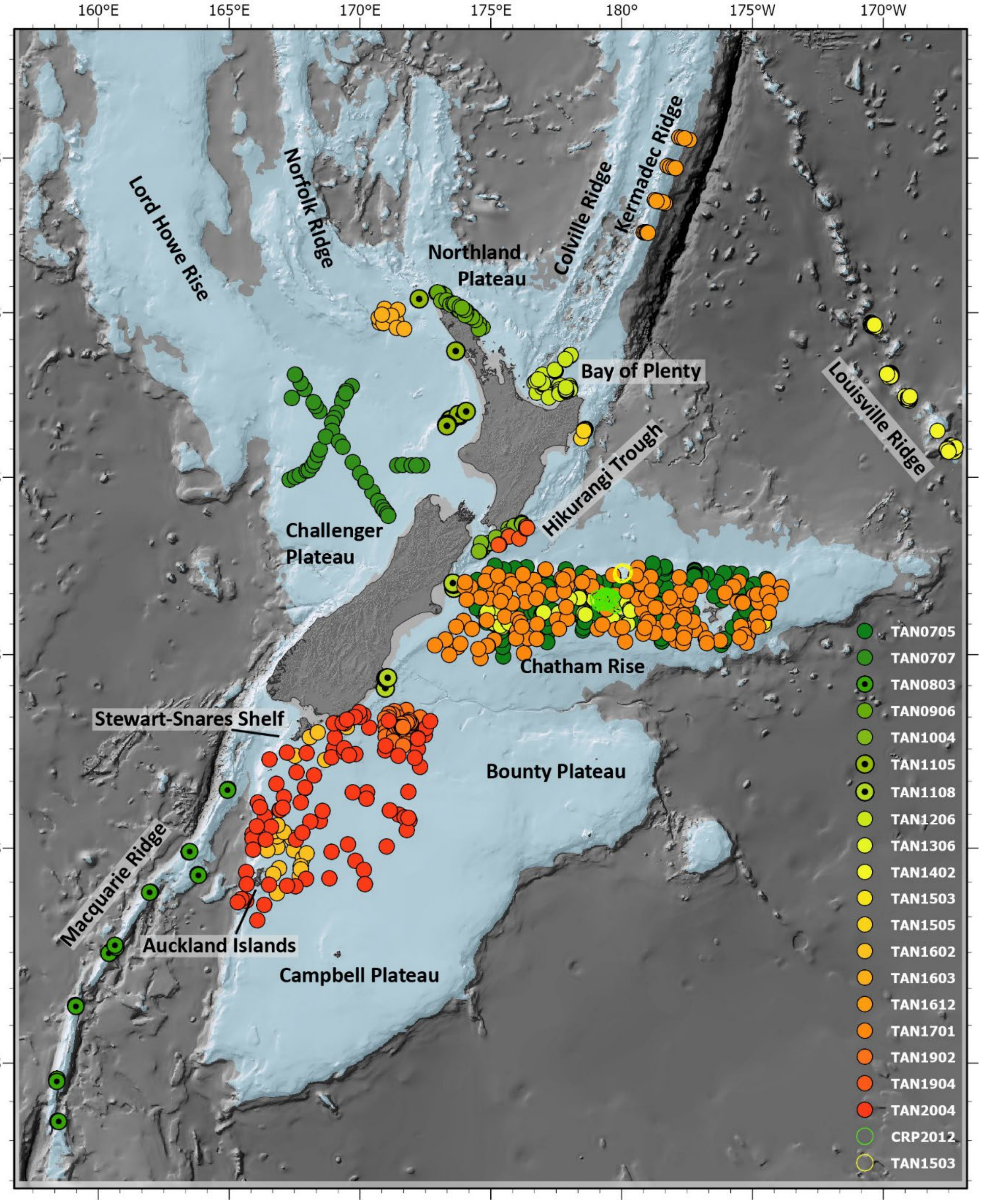

Figure 3-1: The New Zealand Region, with the location of Deep Towed Imaging System (DTIS) transects used in the coral abundance models. The first two numbers of the voyage codes in the legend signify the survey year (except for CRP2012 which took place in 2012); plotting symbol colours follow a colour ramp from older surveys (dark green) to more recent (red). The greyscale with hill-shading indicates the depth gradient across the region. The bathymetric boundary of the model predictions (50–3000 m) is shown in blue. Filled circles indicate surveys with video analysis, open circles indicate surveys with stills analysis. Stations newly analysed for this study are shown with black centre.

3.2 Development of fishing effort layer

The spatial representation of total swept area from all recorded bottom trawling between 1990 and 2019 (deepwater fleet) and between 2008 and 2019 (inshore fleet) shows a range in values from 0 to a maximum of over 250 (equal to 250 square km trawled per 1 km cell) and from this map (Figure 3-2) the regions of the highest fishing intensity can be identified. These occur mainly around the fringes and western end of the Chatham Rise, the edges of the Stewart-Snares shelf, west coast South Island, and coastal regions of the North Island (see Figure 3-1 for the location of these features). The peaks of some of the seamounts in the Louisville Seamount Chain have also been heavily fished.

Areas within the 50–3000 m depth range of the study area where little fishing effort is apparent include the Macquarie Ridge in the south, the shallow part of the central Chatham Rise, the Colville and Kermadec Ridges in the northeast, and much of the Lord Howe Rise and Norfolk ridge systems in the northwest. Much of the Challenger Plateau has also only been lightly trawled.

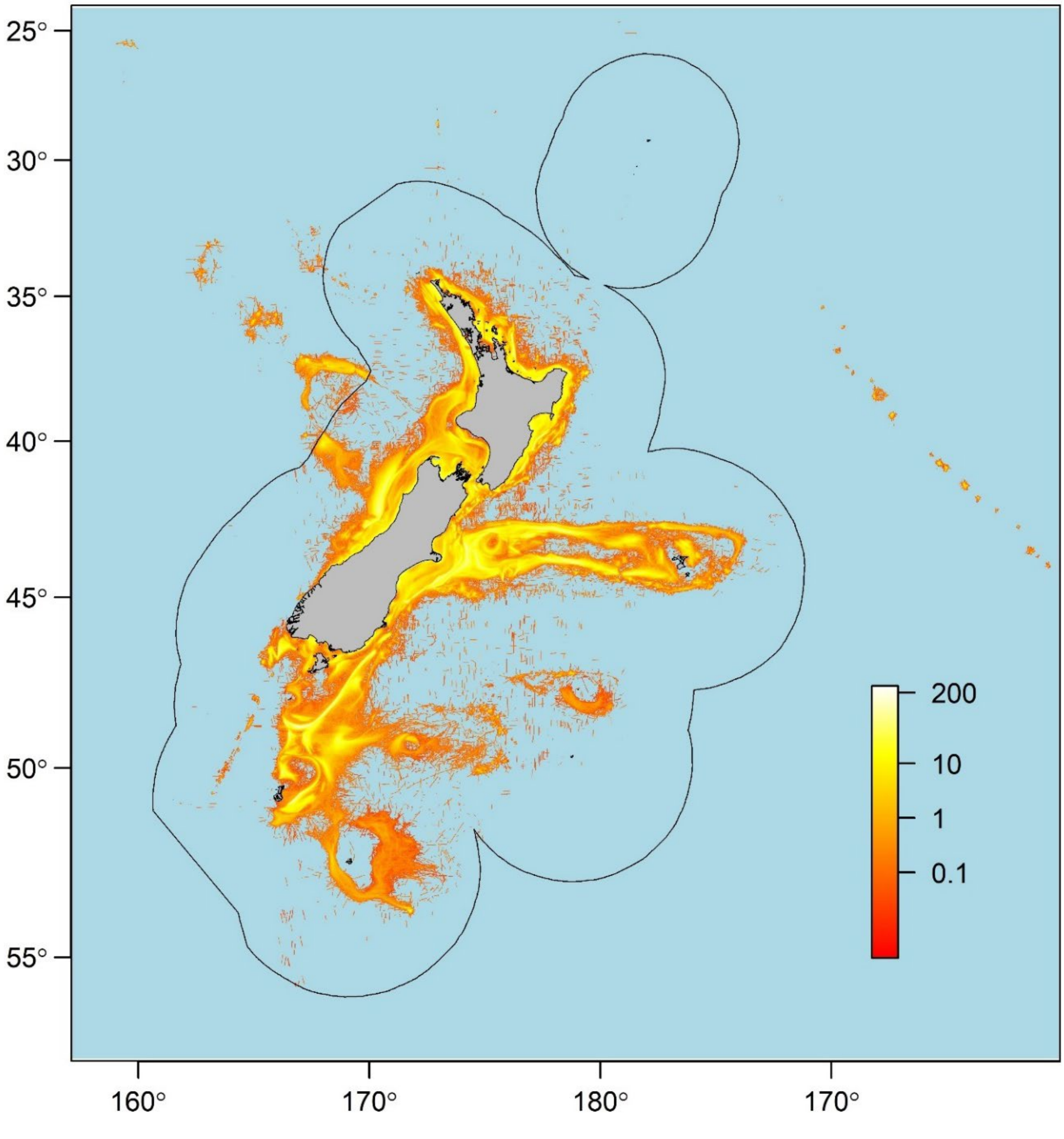

Figure 3-2: Spatial variation in the level of bottom trawl seafloor contact across the New Zealand region. The level of contact is presented as the Swept Area Ratio (SAR), the total accumulated swept area (km^2 per 1 km^2 grid cell) of all recorded bottom trawls between 1990 and 2020, shown on a log scale.

3.3 Abundance models

Ensemble hurdle models were successfully produced for 11 of the 15 taxa in Table 2-1. For the remaining taxa, *Enallopsammia rostrata*, *Madrepora oculata*, *Stephanocyathus* spp., and Paragorgiidae, there proved to be too few records for model parameters to be successfully developed. Maps of predicted abundance are shown in Figure 3-3–Figure 3-23, with model performance metrics and predictor importance provided in Appendix A and Appendix B.

3.3.1 Model performance

The key model performance statistic for the presence-absence part of the BRT and RF hurdle models is AUC (Area Under the receiver operating characteristic Curve). When calculated using the training data used to build the models the AUC values were mostly in the "excellent" category of Hosmer et al. (2013), ranging from 0.94 to 0.98 for RF models and from 0.69 to 0.95 for BRT models; only for Keratoisididae/Mopseidae, Primnoidae, and *Radicipes* spp. were any AUC values in a lower category, with *Radicipes* spp. (BRT) on the lower bound of "adequate". Values of AUC calculated from the evaluation data (data excluded from model building) were naturally lower than the training data-based values but were still mostly "adequate" or "excellent", ranging from 0.76 to 0.90 for RF models and 0.66 to 0.94 for BRT models (see Hosmer & Lemeshow (2013) for full details).

True Skill Statistic values were all positive and in almost all cases substantially greater than 0.5 for the training data and slightly lower for the evaluation data. Highest values were for *Solenosmilia variabilis* and lowest for Keratoisididae/Mopseidae. Presence/absence cut-off values for the training data were very similar between the BRT and RF binomial models for each taxon, ranging from about 0.14 for *Goniocorella dumosa* to about 0.45 for gorgonians and Pennatulacea; evaluation data cut-off values were similar to those for the training data, although usually slightly lower.

For the abundance models, NRMSE (Normalized Root Mean Square Error) values calculated with training data were lower in each case for the BRT models than the RF models, and correlations were higher. The opposite was true for the evaluation data. For the hurdle models, NRMSE values calculated with training data were lower for RF models than for BRT models, and correlations higher; this was also true when using the evaluation data although values were generally very similar between BRT and RF models.

Overall, the best performing hurdle models were those for *Flabellum* spp., *Goniocorella dumosa*, Primnoidae, and Stylasteridae, while the worst were for *Solenosmilia variabilis* and *Radicipes* spp.

3.3.2 Model predictors

For most taxa the conditional permutation "learning" process in the *extendedForests* models selected between 7 and 9 predictors for use in the main modelling process; the only exceptions to this were for *Solenosmilia variabilis* (3 predictors) and Stylasteridae (6 predictors) (Table 3-2).

The most influential predictor overall in the ensembled presence/absence models (Table 3-2) was ruggedness, ranked first for 5 of the 11 taxa. For the abundance models (Table 3-3), however, ruggedness was far less important, with BPI-broad easily the most influential of the seafloor physical-characteristic variables. Dissolved oxygen and POC were the two other key variables in both presence-absence and abundance models, with POC of considerable influence in abundance models in particular – ranking first in models for four taxa.

Carbonate ion concentration (aragonite or calcite depending on the taxon) were of only modest influence in most models, the exception being the presence-absence model for Pennatulacea in which it was the highest ranked variable. Seafloor material variables, mud and gravel were similarly of modest influence in the models, although were generally more important in the abundance models than the presence-absence models. This was especially true for models comprising gorgonian species in which they were ranked first or second for three of the four taxa modelled.

Fishing effort (SAR) and seamount were overall the least important of the variables offered to the models, with each only selected for use in models for two taxa. However, where fishing effort was used in the presence-absence models this predictor had substantial influence and was the most important predictor for *Radicipes* spp. and third ranked for *Goniocorella dumosa*. See Table A-1 for further details on predictor importance for individual models.

Table 3-2: Order of importance for each variable in the BRT/RF ensemble presence/absence models for each of the 11 protected coral taxa modelled. Variables are ordered from most to least important by their overall rank, calculated as the average of the individual ranks for each taxon where unselected variables (marked *) are given a number 1 greater than the least important selected variable. Variables marked "–" were not considered in the model; shading used to differentiate taxonomic groupings.

	Solenosmilia variabilis	*Goniocorella dumosa*	Caryophylliidae	*Flabellum*	Gorgonians	Keratoisididae/Mopseidae	Primnoidae	*Radicipes*	Antipatharia	Stylasteridae	Pennatulacea	Rank
Ruggedness	*	4	1	8	4	1	1	*	1	1	*	4.0
BPI-broad	1	6	4	4	2	9	5	4	4	2	4	4.1
Dissolved oxygen	*	1	3	2	1	4	7	8	3	*	6	4.2
POC	*	2	8	1	8	3	4	3	7	3	8	4.6
Slope SD	2	*	7	3	3	2	3	7	6	5	5	4.6
Aragonite	*	5	9	6	–	–	–	–	–	*	1	5.3
Bottom temperature	*	7	2	5	7	7	8	6	2	*	7	5.6
Percent mud	*	*	6	7	6	6	2	9	8	*	2	5.9
Percent gravel	*	*	5	*	*	5	9	2	*	4	3	6.0
Calcite	–	–	–	–	5	8	6	5	5	–	–	6.0
Fishing effort (SAR)	*	3	*	*	*	*	*	1	*	*	*	7.3
Seamount	3	*	*	*	*	*	*	*	*	6	*	8.4

Table 3-3: Order of importance for each variable in the BRT/RF ensemble abundance models for each of the 11 protected coral taxa modelled. Further details as per Table 3-2.

g	*Solenosmilia variabilis*	*Goniocorella dumosa*	Caryophyllidae	*Flabellum*	Gorgonians	Keratoisididae/Mopseidae	Primnoidae	*Radicipes*	Antipatharia	Stylasteridae	Pennatulacea	Rank
BPI-broad	1	2	1	4	2	4	6	8	1	2	4	3.2
POC	*	1	9	2	5	7	1	1	5	3	1	3.5
Dissolved oxygen	*	6	3	1	6	5	5	2	4	*	3	4.2
Bottom temperature	*	3	7	5	4	8	3	5	2	*	2	4.5
Percent mud	*	*	4	3	1	2	2	7	8	*	5	4.6
Slope SD	2	*	8	7	3	1	8	4	3	5	7	5.1
Percent gravel	*	*	2	*	*	3	4	3	*	1	8	5.4
Aragonite	*	5	5	6	–	–	–	–	–	*	6	5.5
Ruggedness	*	4	6	*	8	6	9	*	7	4	*	6.8
Calcite	–	–	–	–	7	9	7	6	6	–	–	7.0
Fishing effort (SAR)	*	7	*	*	*	*	*	9	*	*	*	8.4
Seamount	3	*	*	*	*	*	*	*	*	6	*	8.4

3.3.3 Model predictions, and uncertainty

Maps of predicted abundance and model uncertainty of protected coral taxa across the New Zealand region were produced from the ensembled hurdle models for each protected coral taxon. Typically, the distribution of abundance and uncertainty values across the model extent was non-normal and considerably more detail can be distinguished by displaying predictions on a log scale.

Species of the order Antipatharia (black corals) were predicted to be in high abundance on the Challenger Plateau, where they were frequently encountered on survey TAN0707, and in northern parts of the region more than the south. Predicted abundance was relatively low in shallow parts of the Chatham Rise and sub-Antarctic plateaus but higher in deeper fringes of these areas (Figure 3-3).

The cup coral family Caryophylliidae was predicted to be widespread around the region, with high abundance particularly on the Chatham Rise, around the northern ridge and Rise systems, the sub-Antarctic plateaus including Macquarie Ridge, and the Louisville Seamount Chain. Lower abundance was predicted for the Bounty Trough and Challenger Plateau (Figure 3-5).

The solitary cup coral genus *Flabellum* shows highest densities along the shallower regions of the Chatham Rise, sub-Antarctic plateaus, as well as the fringes and shelves surrounding both main islands. Abundance of this taxon is generally lower in northern areas, especially the Colville and Kermadec ridges and Louisville seamount chain (Figure 3-7).

Presence records for the combined gorgonian alcyonaceans were widely spread around the region, mainly in relatively deep water. High abundances were predicted around the fringes of the Chatham

Rise and sub-Antarctic plateaus, and Macquarie Ridge in the south, and in most parts of the northers ridge and rise features, including the Louisville Seamount Chain (Figure 3-11).

The bamboo corals (Keratoisididae & Mopseidae) were predicted to have highest abundance around the deeper fringes of the Chatham Rise, on parts of the Challenger Plateau and Lord Howe Rise, and around the Campbell Plateau in the south. Multiple records of relatively low abundance in the Bay of Plenty match with predicted moderate levels of abundance in that area (Figure 3-13).

Sea pens (order Pennatulacea) were predicted to be most abundant on the south Chatham Rise, the Bounty Trough, and in mostly deeper waters surrounding the Challenger Plateau and Lord Howe Rise. The highest observed abundances were around the south Chatham Rise, Wairarapa coast, and northwest North Island (Figure 3-15).

The sea fans and bottle brushes (primnoid alcyonaceans) were predicted to occur in highest abundances in deeper water around the Chatham Rise and Bounty Plateau, as well as the fringes of the Challenger Plateau extending north and south on the edge of the mainland shelf, on Lord Howe Rise, the Norfolk Ridges, and Macquarie Ridge in the south. The lowest predicted abundances were in the shallower parts of the Chatham Rise, sub-Antarctic plateaus, and shelf regions around the main islands (Figure 3-17).

High abundance of the sea whip genus *Radicipes* was predicted to be strongly focused around much of the perimeter of the Chatham Rise, along with the shelf edge off the northwest of the South Island and some areas on the Campbell Plateau. Lower abundances were predicted for deeper areas of the region, especially in the northeast (Figure 3-19).

Presence observations for the branching coral *Solenosmilia variabilis* were mostly limited to the Louisville Seamount Chain, with one or two very high abundances recorded around the Graveyard Hills on the north Chatham Rise. This led to predictions of uniformly low abundance across much of the region, with higher abundances predicted only for the Louisville Seamount Chain and Macquarie Ridge, with a few small patches of moderate abundance widely spread around the region (Figure 3-21).

Areas of high abundance of stylasterid hydrocorals were widespread around the modelled area. These included a number of locations along the Chatham Rise, especially east of the Chatham Islands where the highest observed densities were found, the ridge systems to the north, east, and south of New Zealand, and around the fringes of the individual features of the sub-Antarctic plateaus. Relatively lower abundance was predicted for the Challenger Plateau and much of the Lord Howe Rise (Figure 3-23).

Uncertainty as measured by standard deviation tends to be related to abundance, so that where abundance is predicted to be high, model standard deviations are also high; models tend to be more certain in predictions of areas where the coral taxon is absent or in low abundance, as may be expected for animals such as these which are sparsely distributed even when in relatively high abundance.

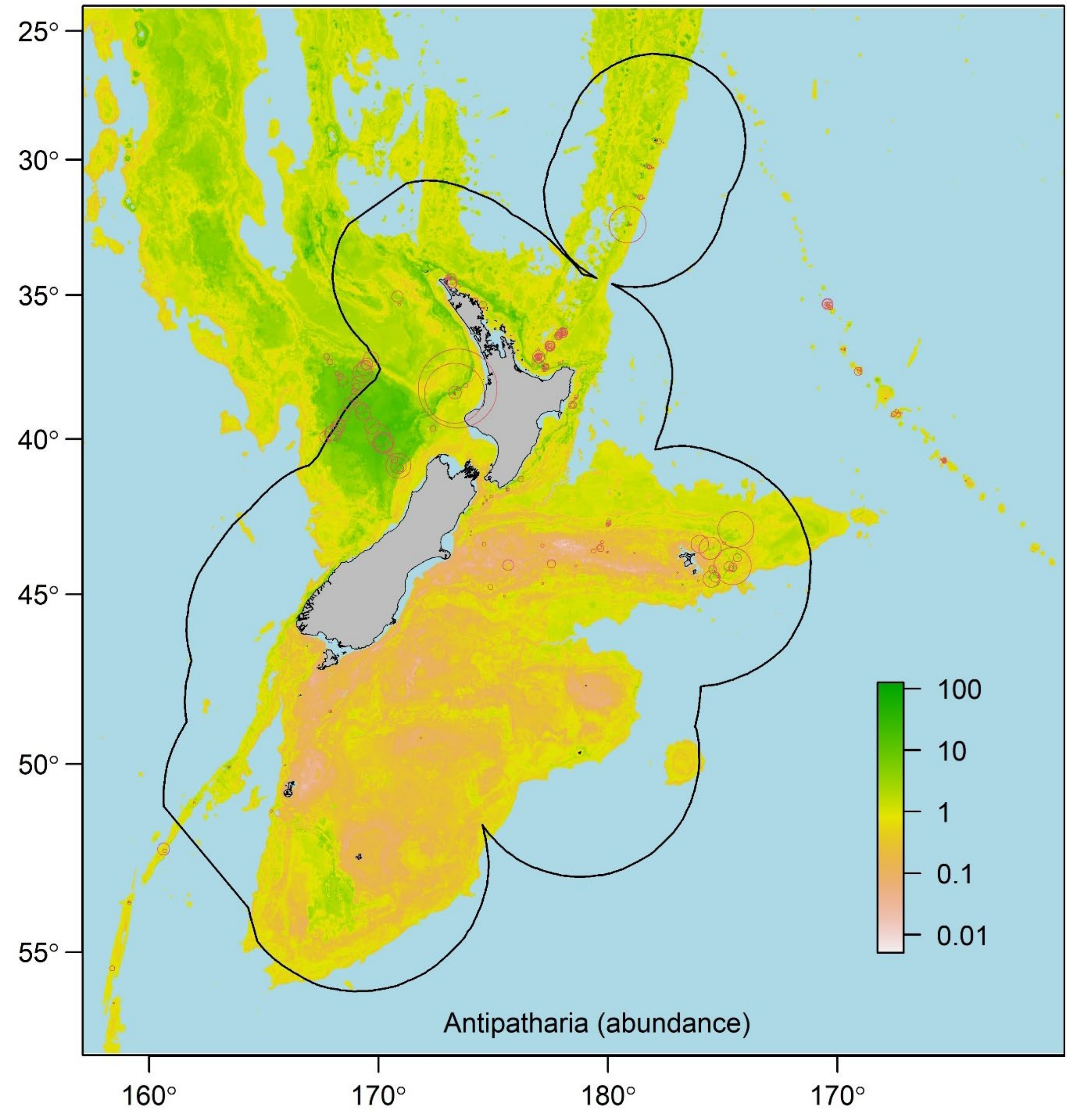

Figure 3-3: Antipatharia (black corals, all species). Predicted abundance (number of individuals per 1000 m^2). Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 733 individuals per 1000 m^2.

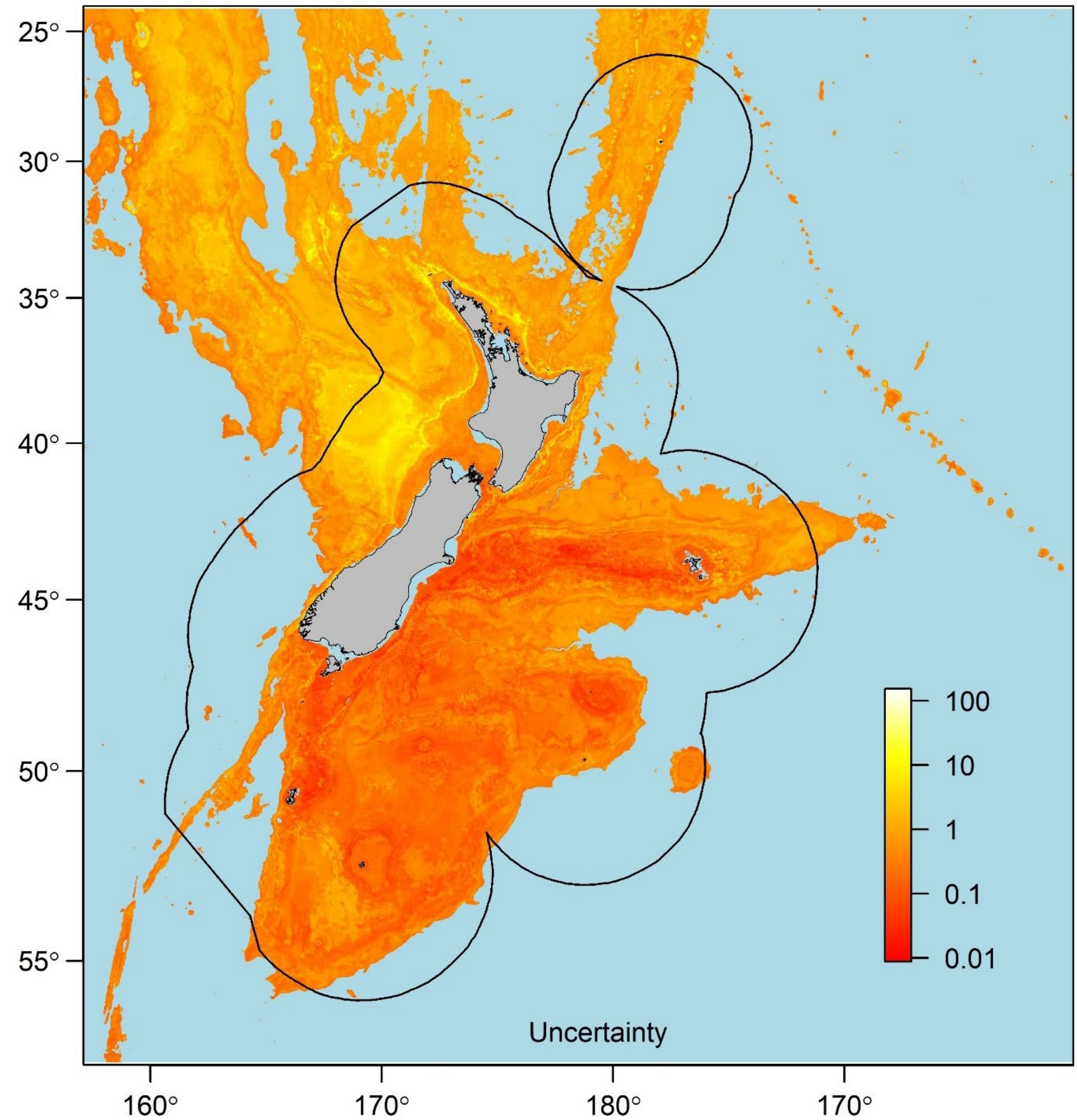

Figure 3-4: Antipatharia (black corals, all species). Model uncertainty (Standard Deviation).

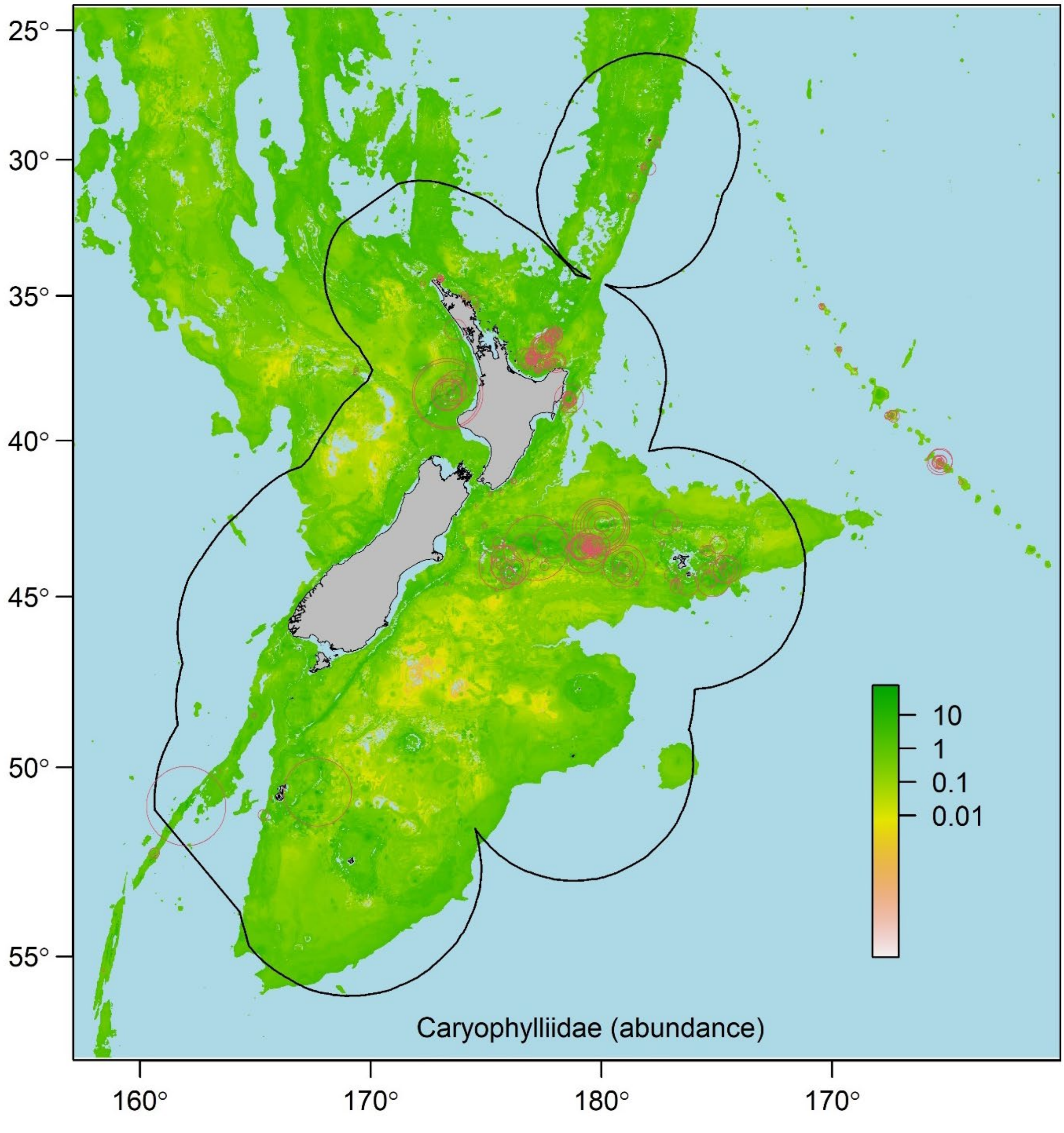

Figure 3-5: Caryophylliidae (cup corals). Predicted abundance (number of individuals per 1000 m^2). Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 262 individuals per 1000 m^2.

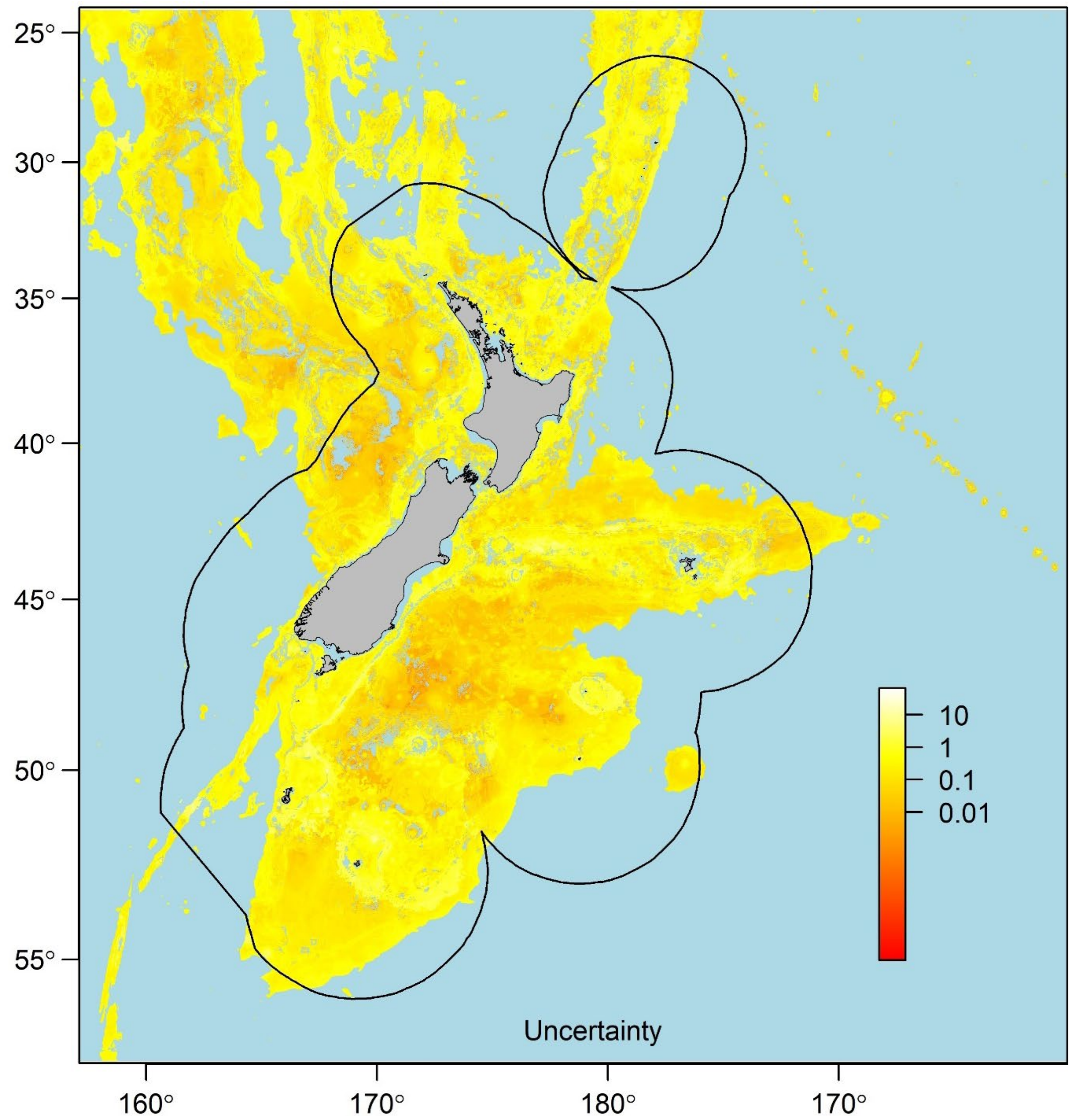

Figure 3-6: Caryophylliidae (cup corals). Model uncertainty (Standard Deviation).

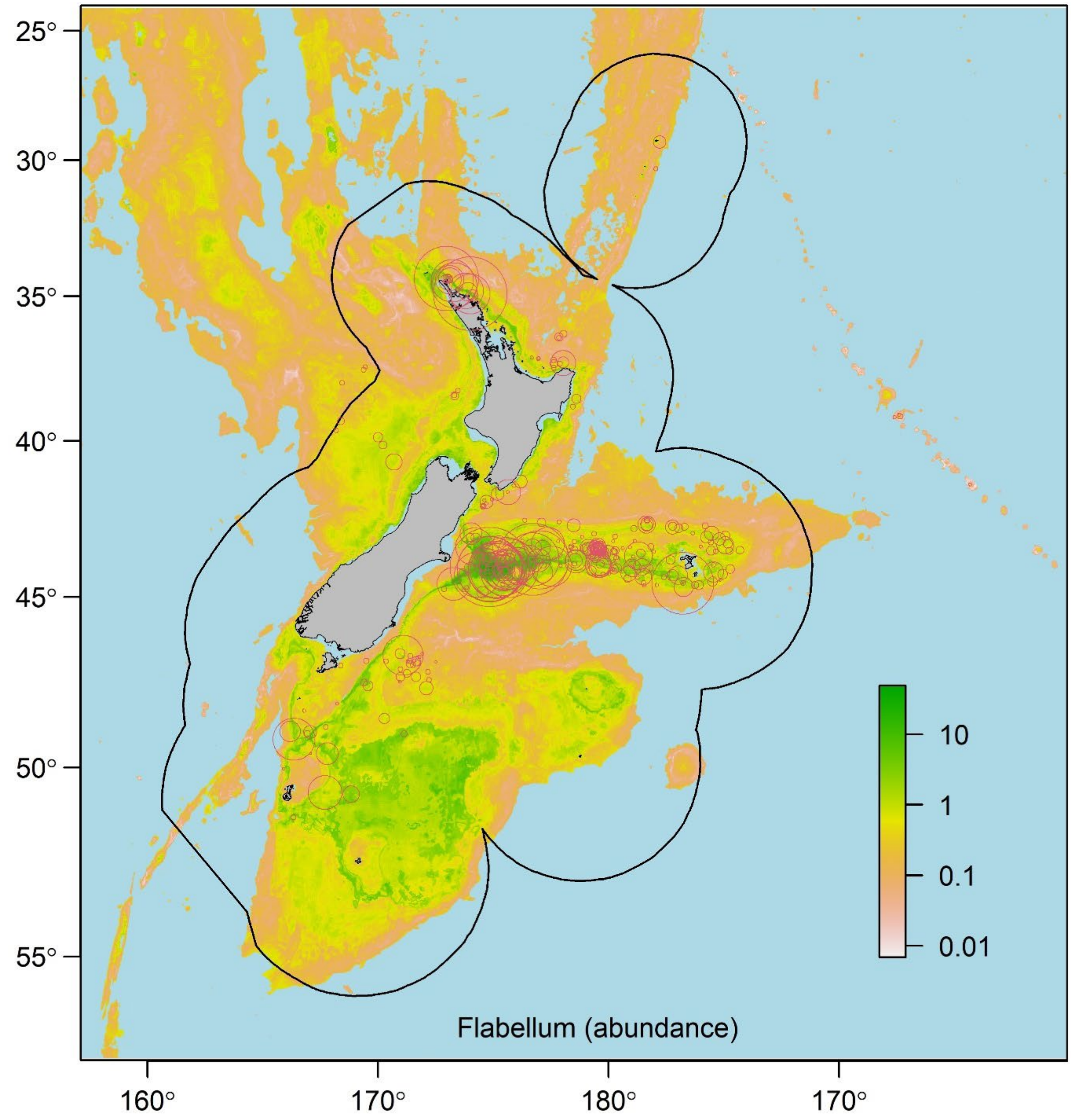

Figure 3-7: ***Flabellum*** **spp. (solitary corals). Predicted abundance (number of individuals per 1000 m^2).** Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 130 individuals per 1000 m^2.

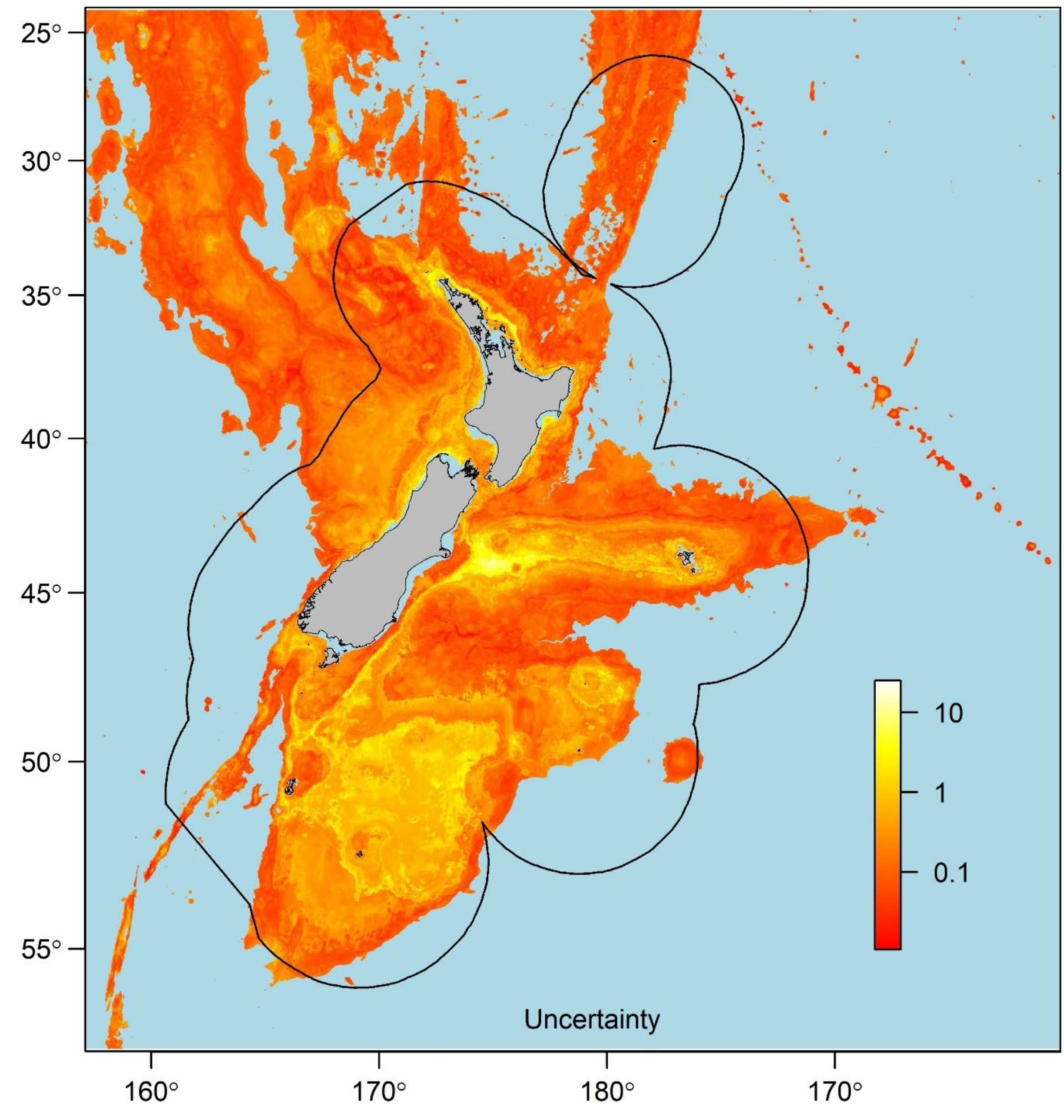

Figure 3-8: ***Flabellum*** **spp. (solitary corals). Model uncertainty (Standard Deviation).**

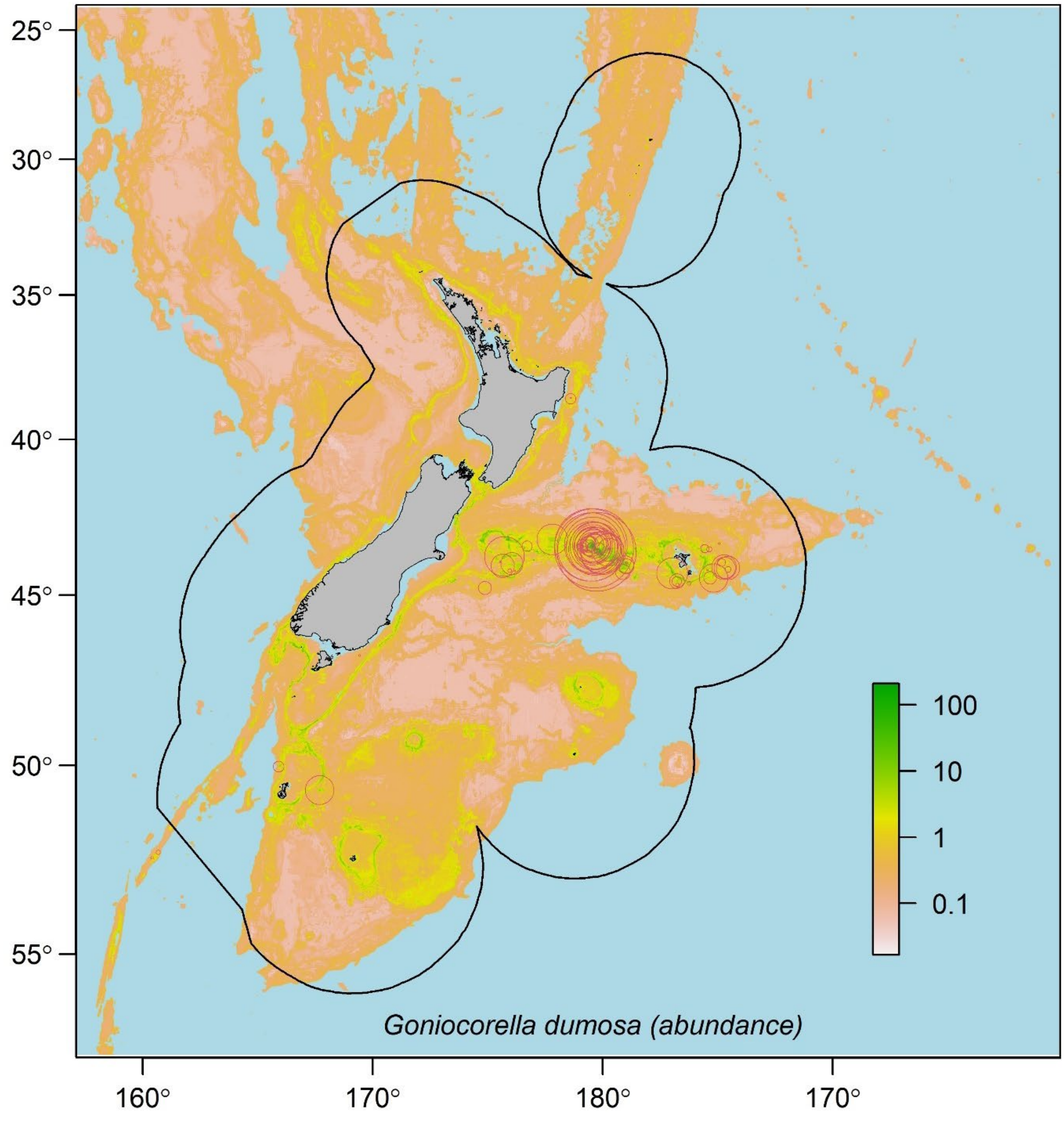

Figure 3-9: *Goniocorella dumosa* (stony coral). Predicted abundance (number of individuals per 1000 m²). Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 582 individuals per 1000 m².

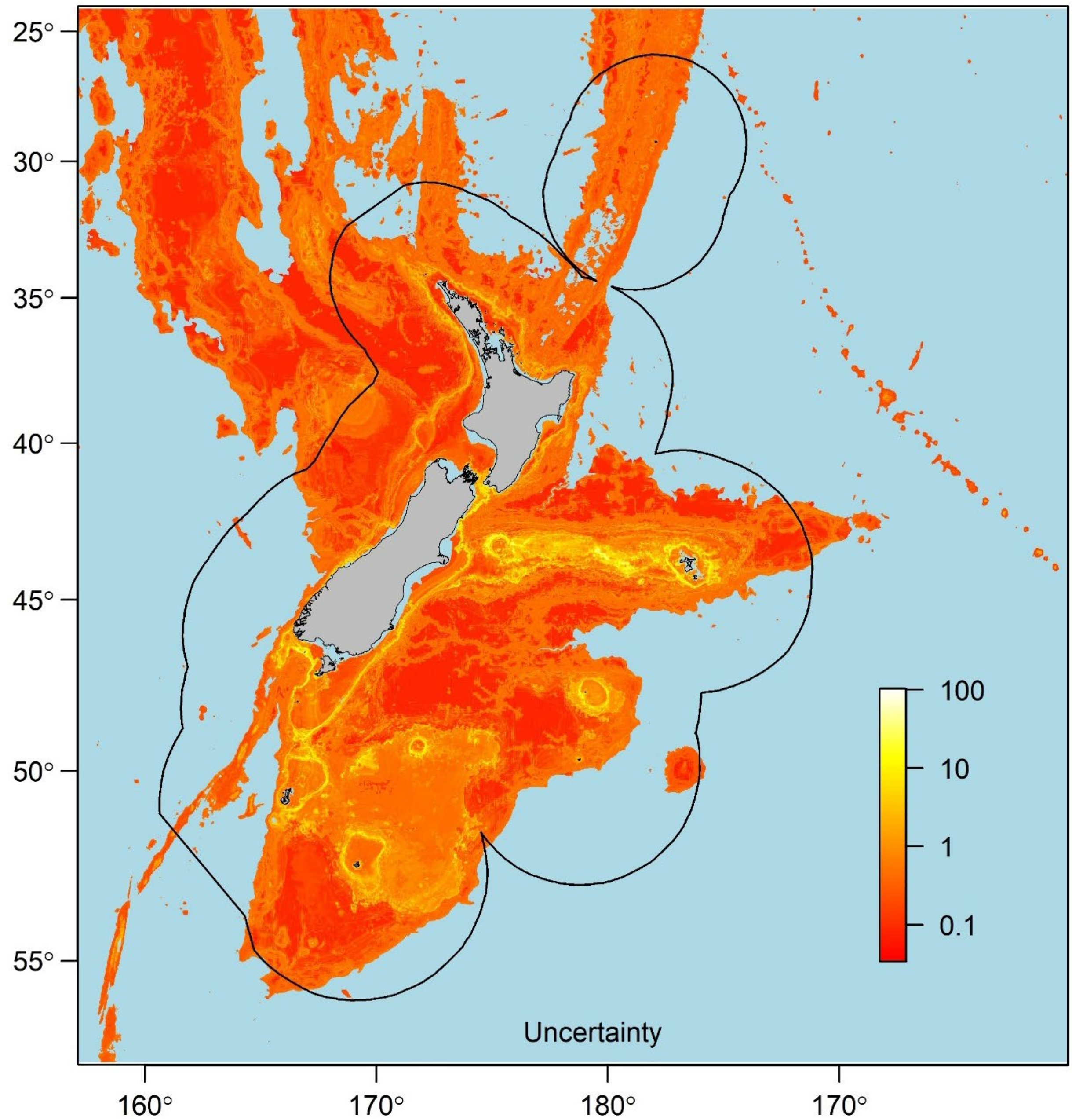

Figure 3-10: *Goniocorella dumosa* (stony coral). Model uncertainty (Standard Deviation).

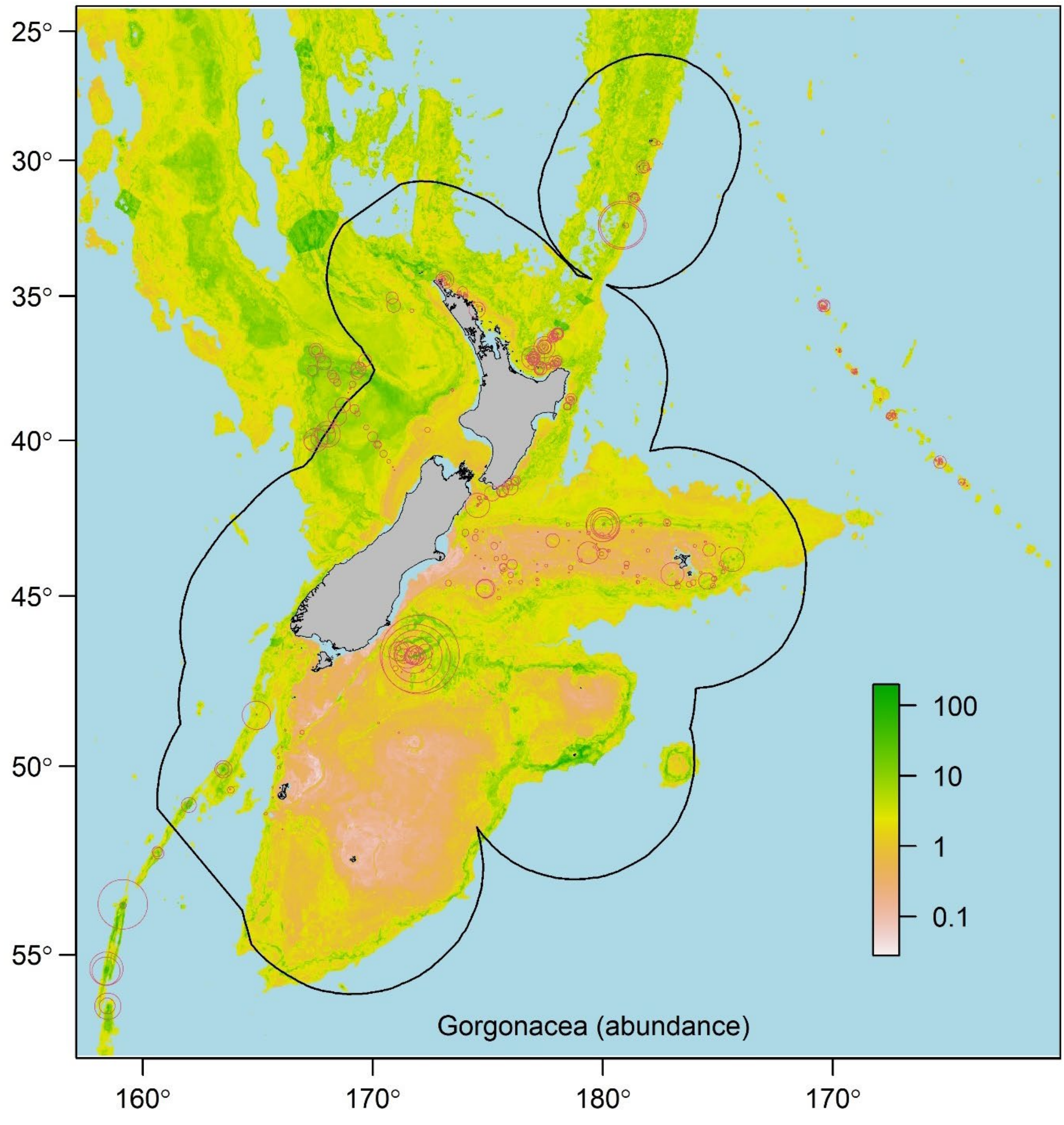

Figure 3-11: Gorgonian alcyonaceans (all). Predicted abundance (number of individuals per 1000 m^2). Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 1080 individuals per 1000 m^2.

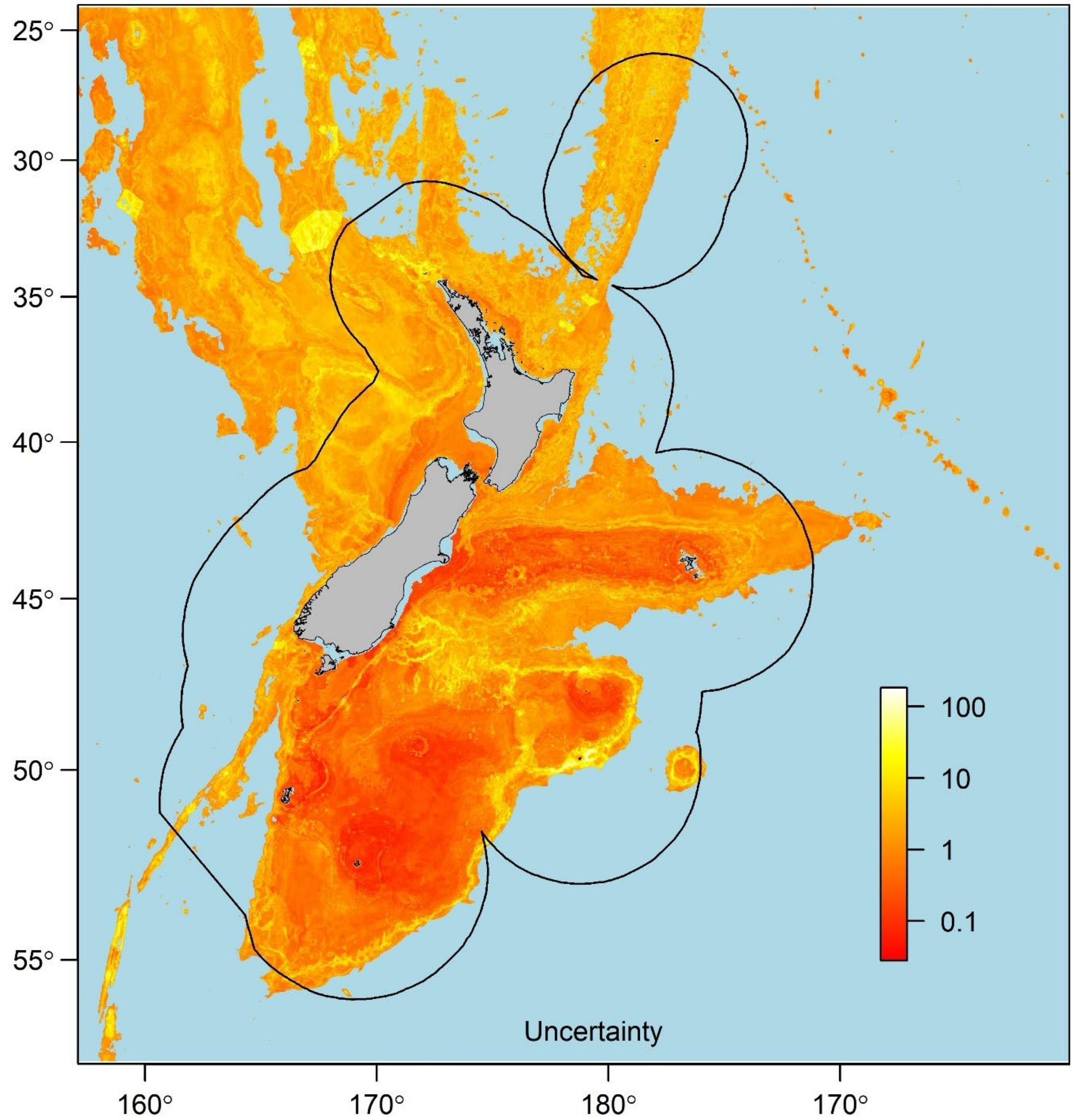

Figure 3-12: Gorgonian alcyonaceans (all). Model uncertainty (Standard Deviation).

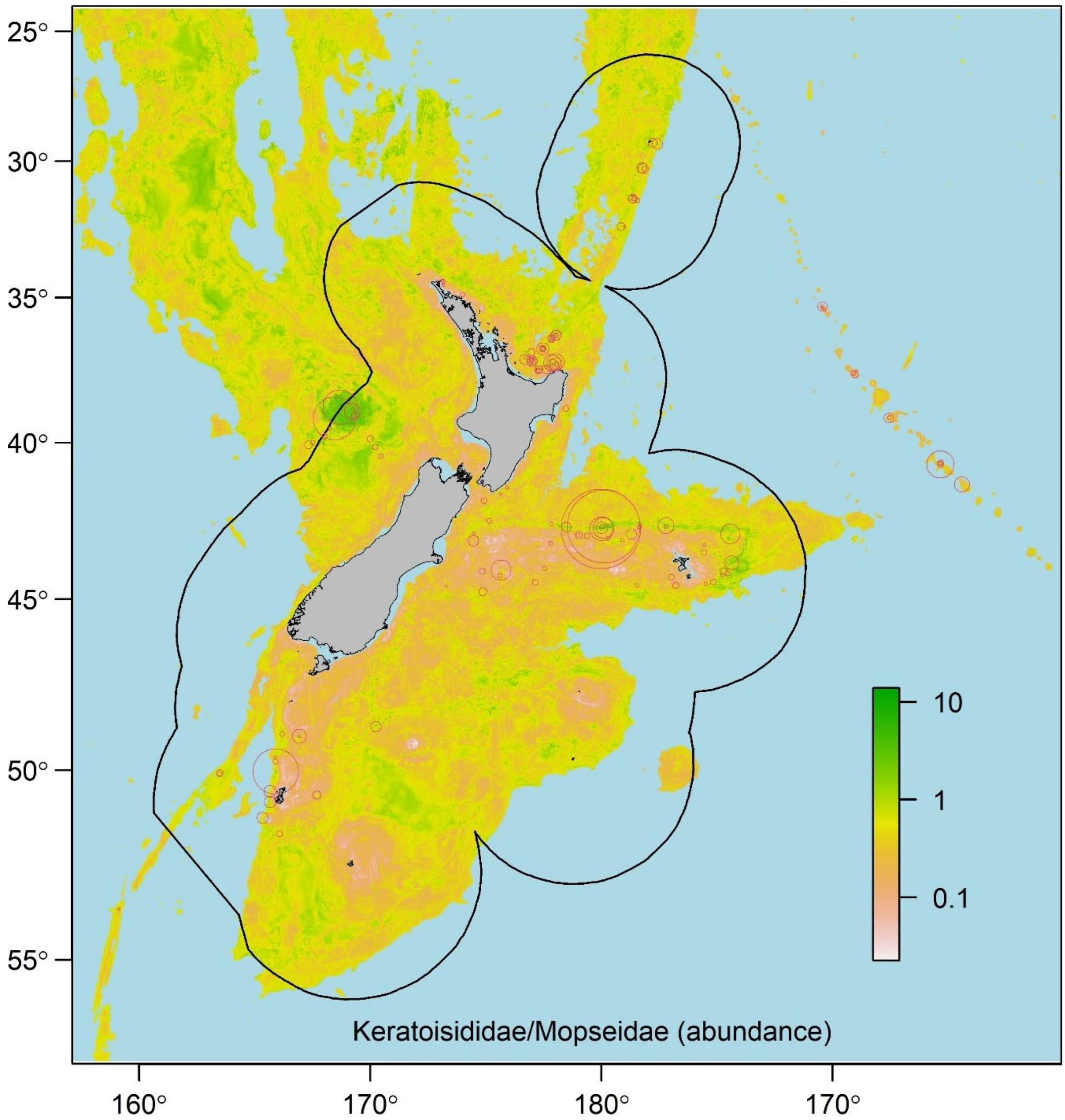

Figure 3-13: Keratoisididae/Mopseidae (bamboo corals). Predicted abundance (number of individuals per 1000 m²). Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 199 individuals per 1000 m².

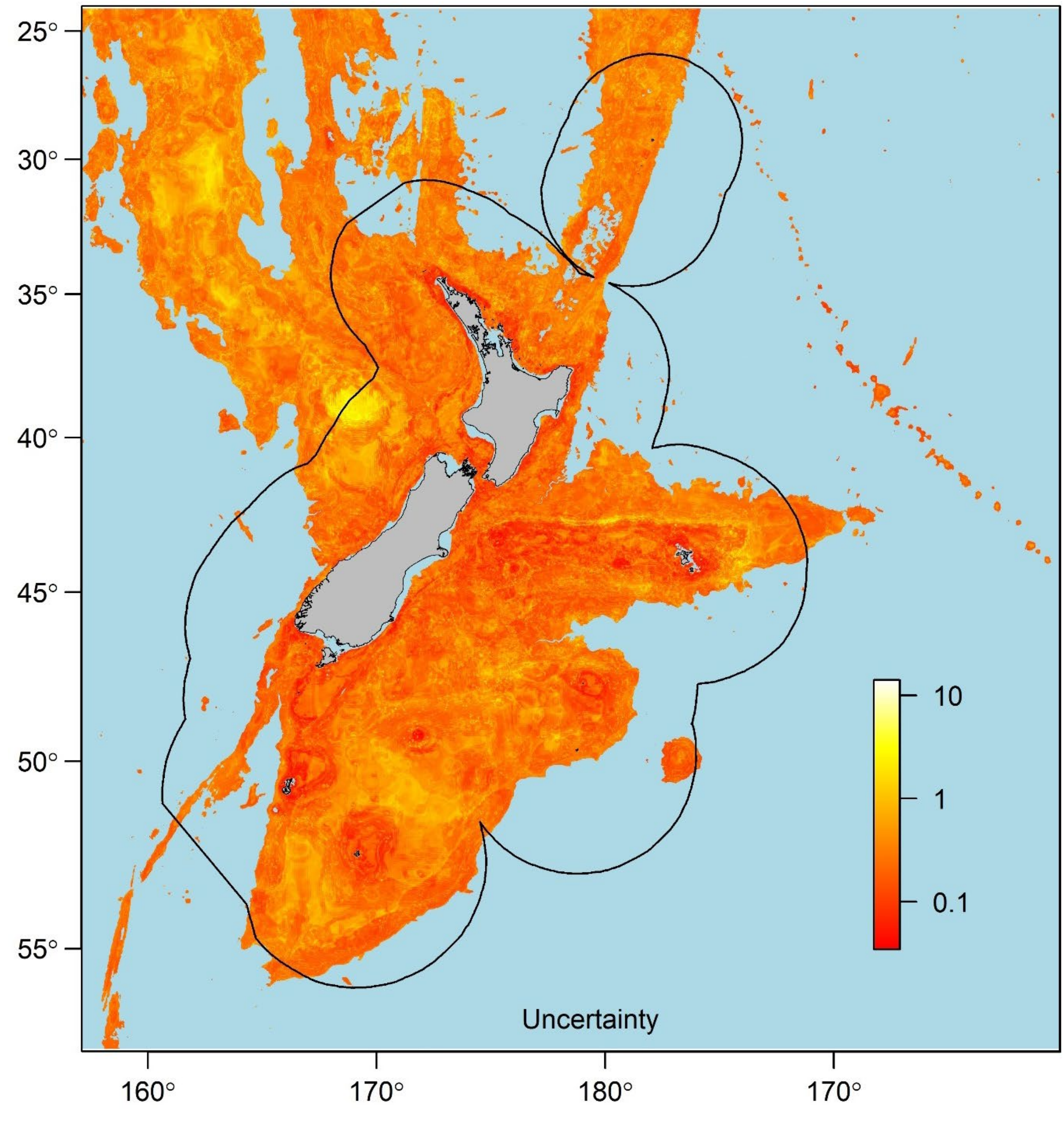

Figure 3-14: Keratoisididae/Mopseidae (bamboo corals). Model uncertainty (Standard Deviation).

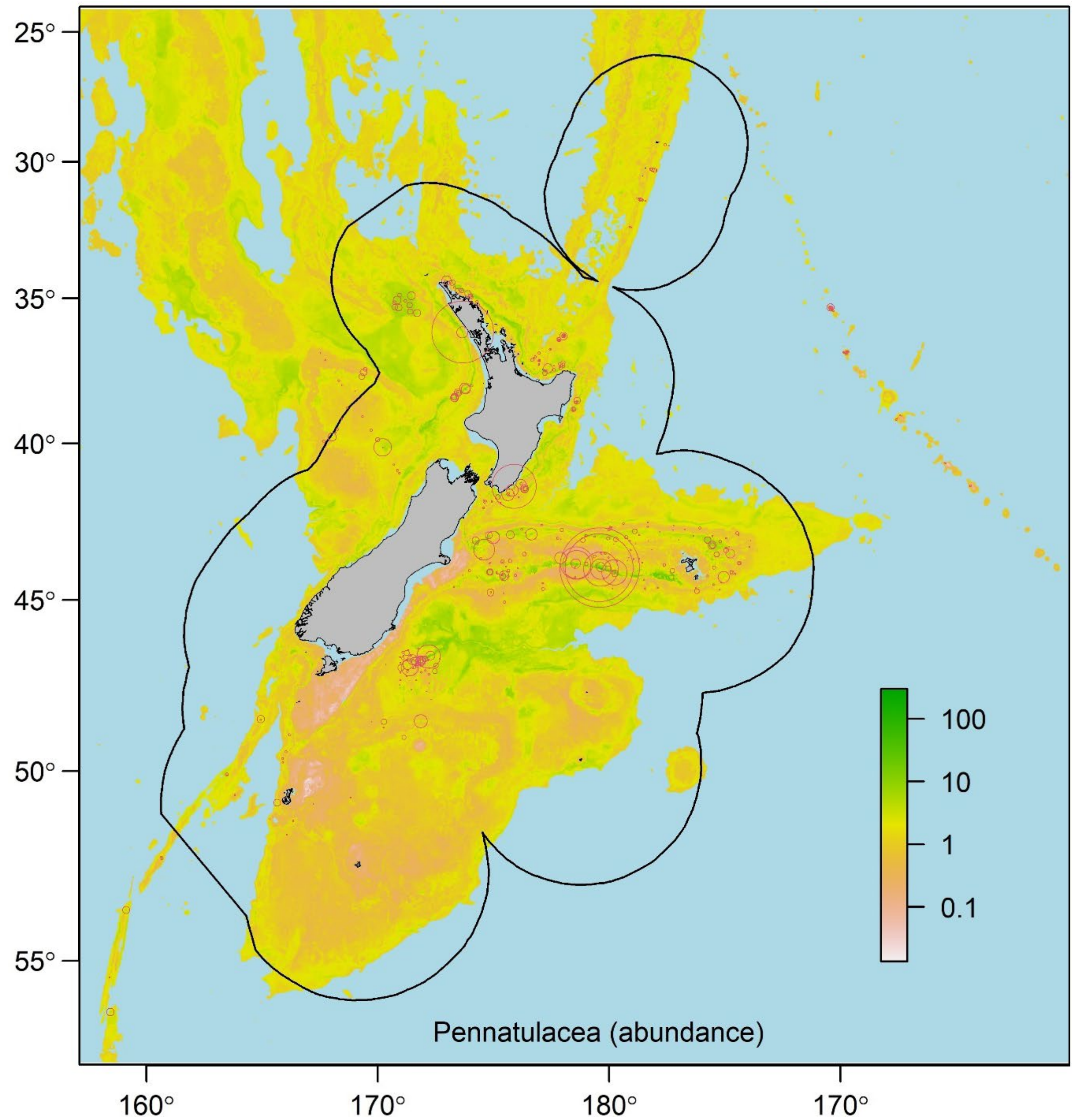

Figure 3-15: Pennatulacea (sea pens). Predicted abundance (number of individuals per 1000 m²). Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 1669 individuals per 1000 m².

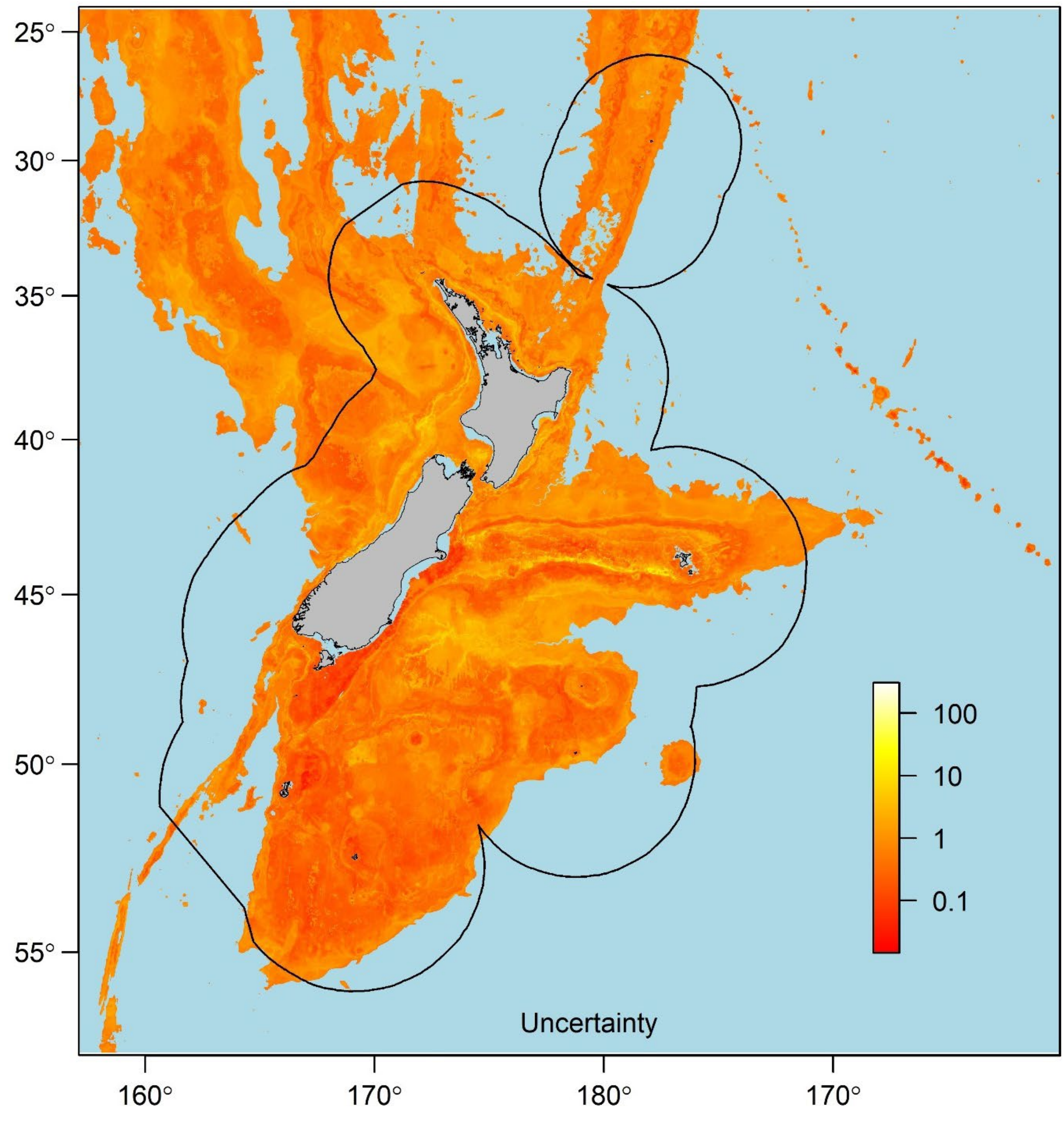

Figure 3-16: Pennatulacea (sea pens). Model uncertainty (Standard Deviation).

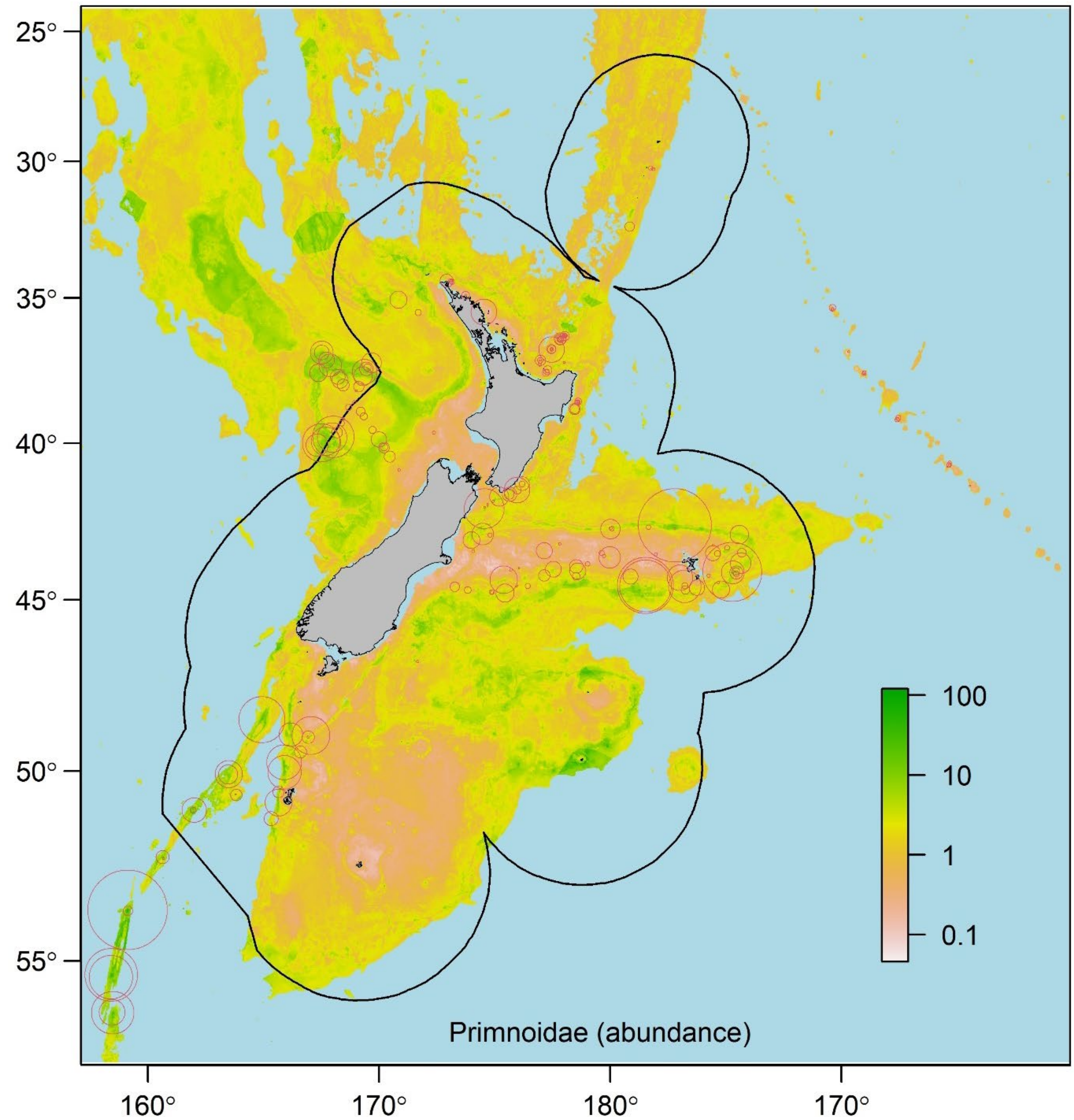

Figure 3-17: Primnoidae (sea-fans and bottle brush corals). Predicted abundance (number of individuals per 1000 m^2). Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 420 individuals per 1000 m^2.

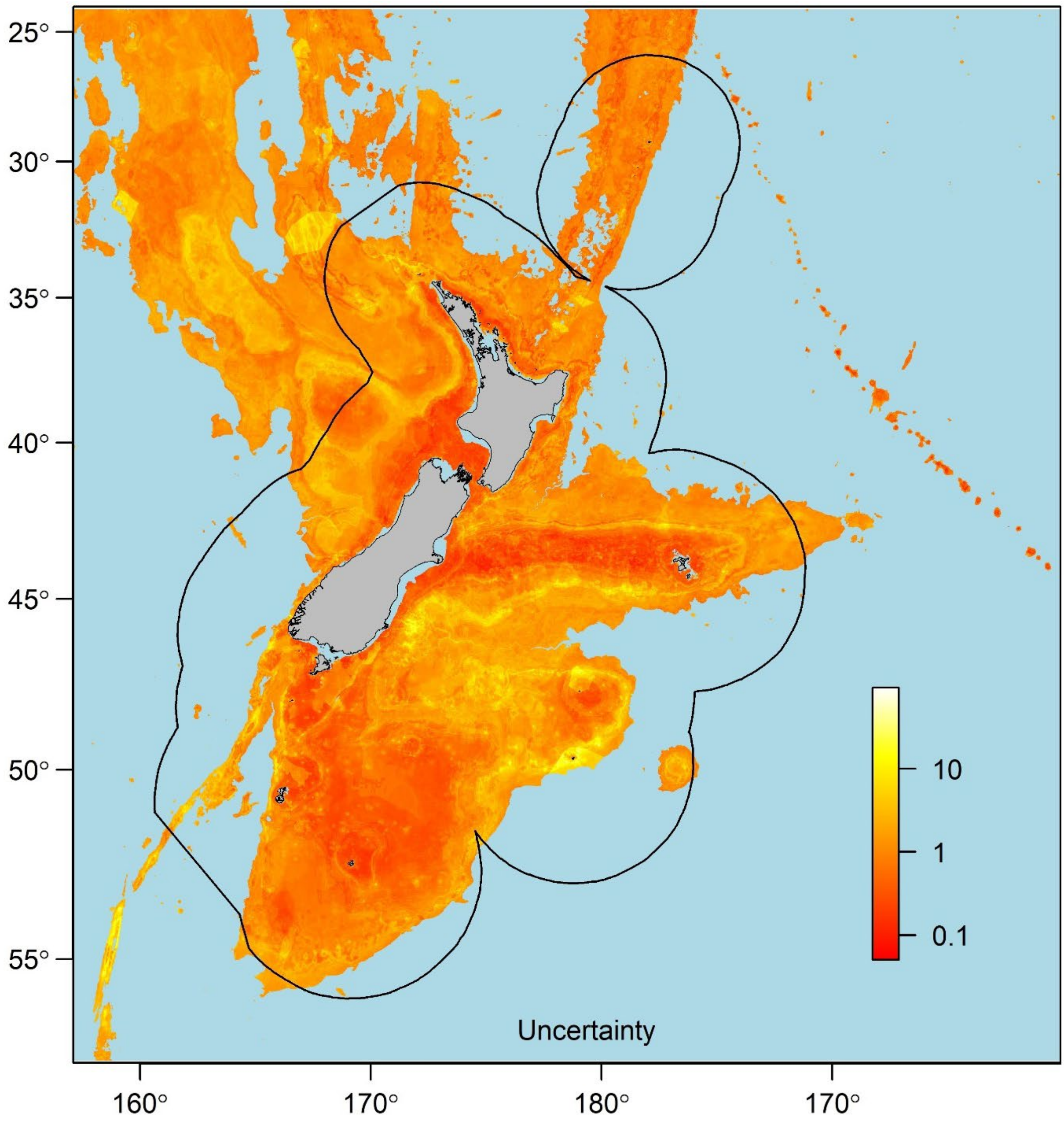

Figure 3-18: Primnoidae (sea-fans and bottle brush corals). Model uncertainty (Standard Deviation).

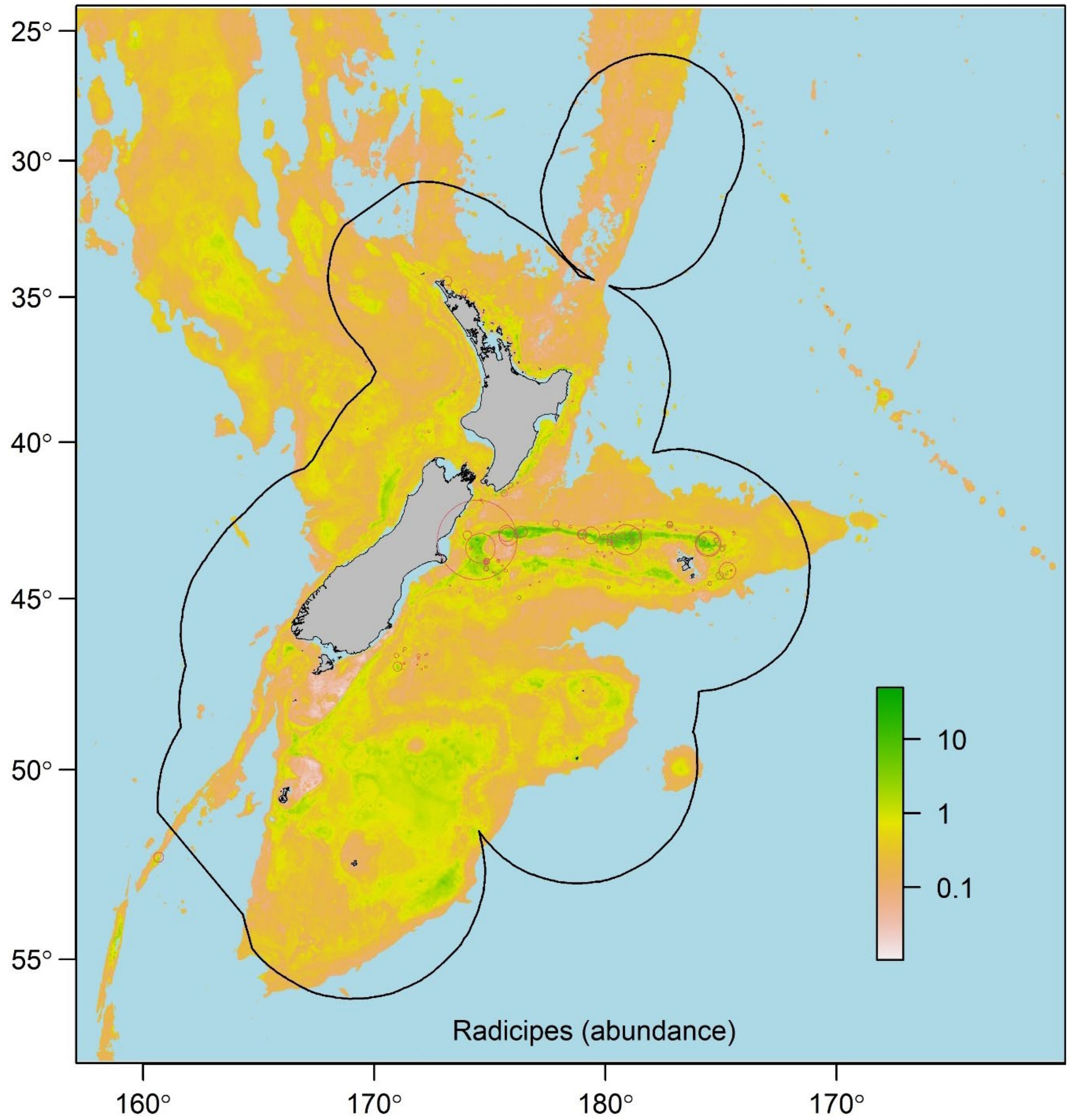

Figure 3-19: ***Radicipes*** **spp. (sea whips). Predicted abundance (number of individuals per 1000 m^2.** Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 1322 individuals per 1000 m^2.

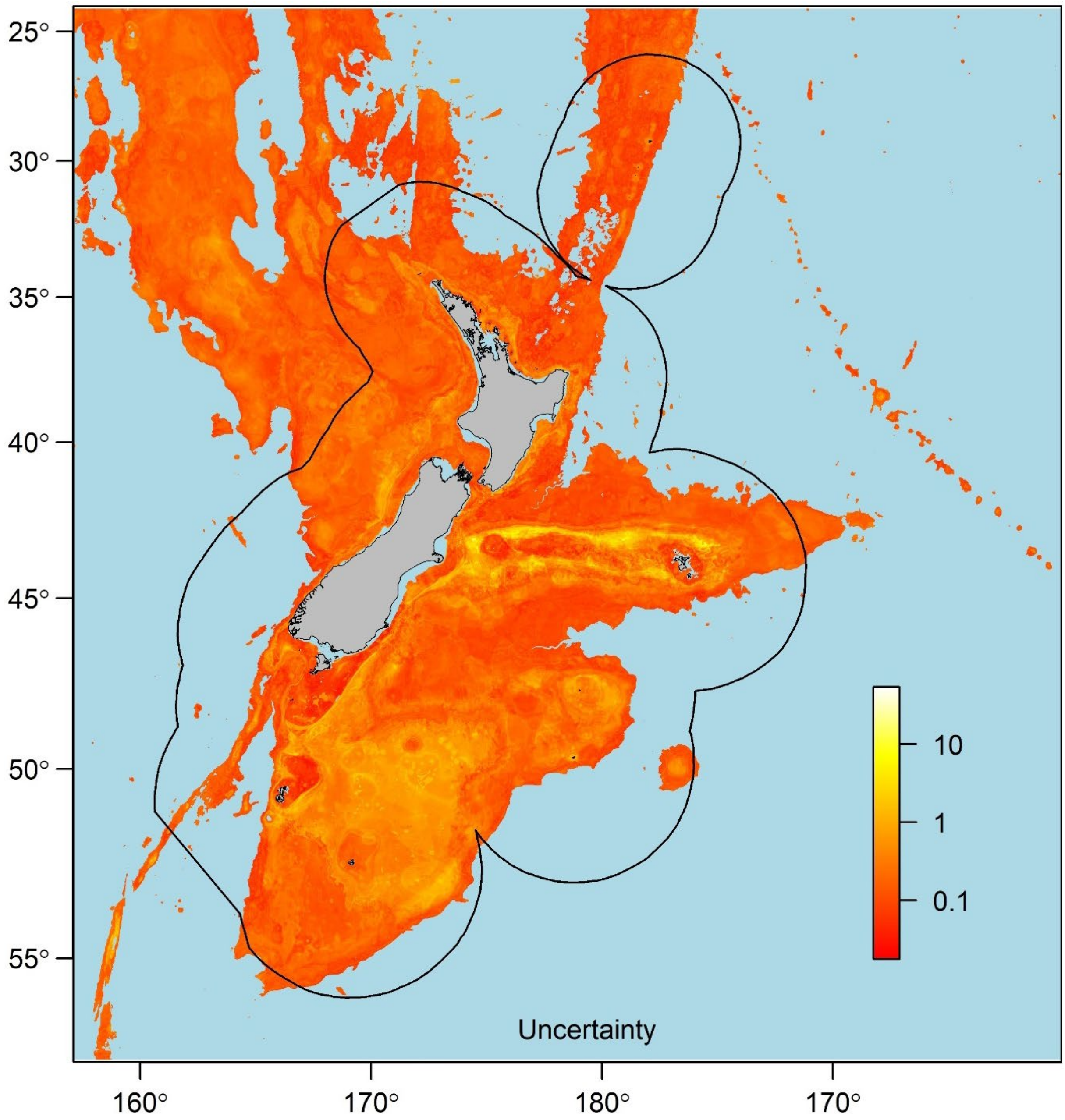

Figure 3-20: ***Radicipes*** **spp. (sea whips). Model uncertainty (Standard Deviation).**

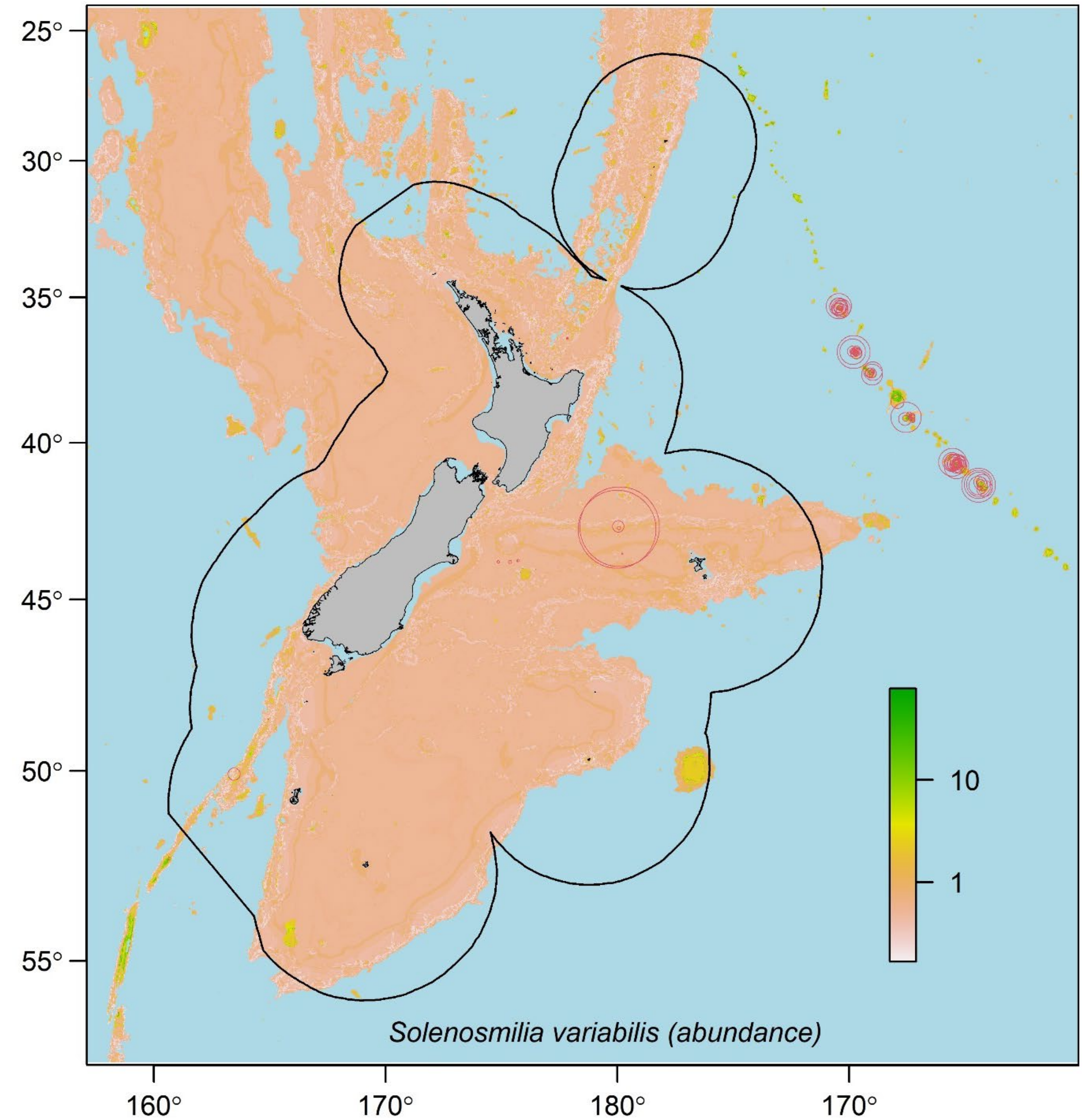

Figure 3-21: ***Solenosmilia variabilis*** **(stony coral). Predicted abundance (number of individuals per 1000 m²).** Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 1378 individuals per 1000 m².

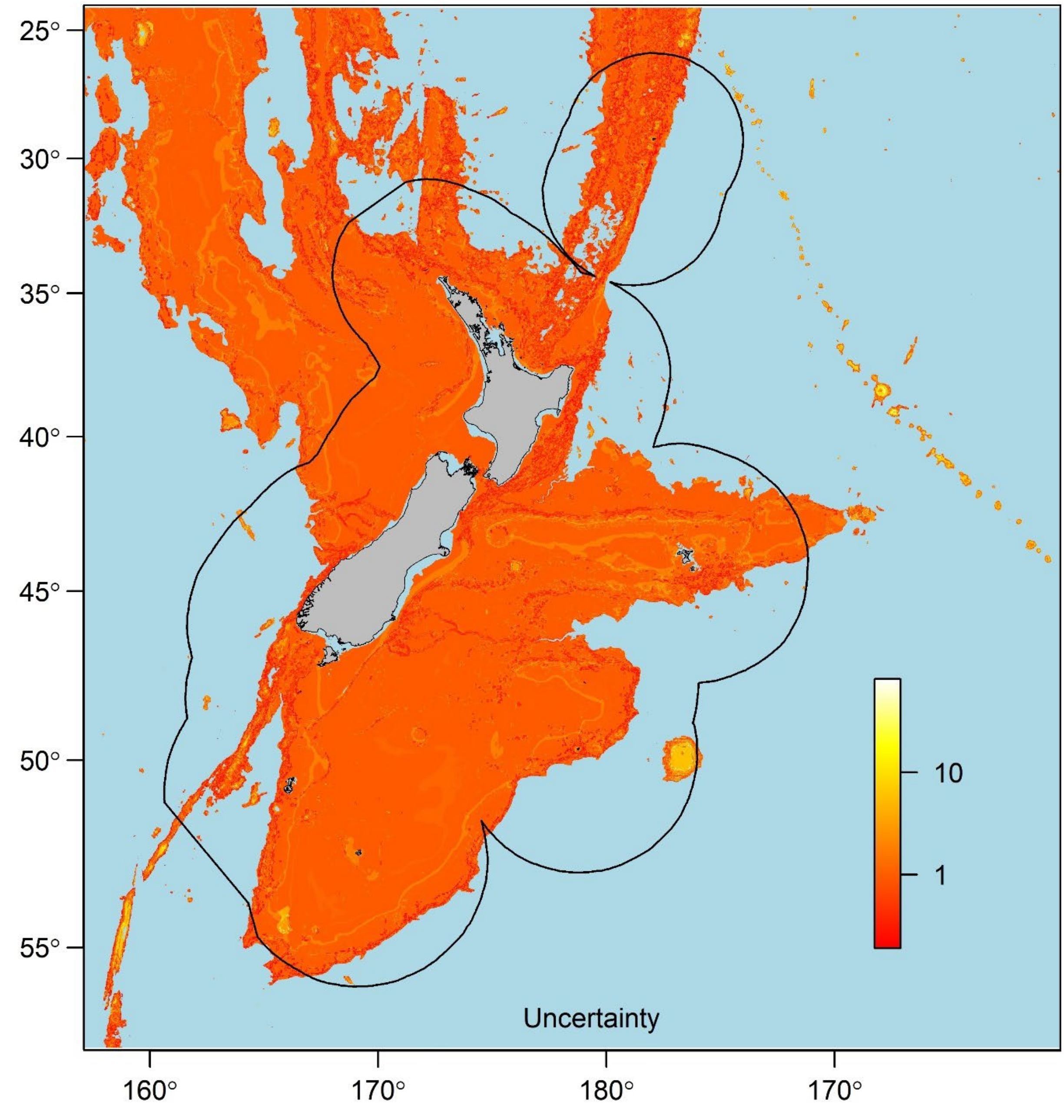

Figure 3-22: ***Solenosmilia variabilis*** **(stony coral). Model uncertainty (Standard Deviation).**

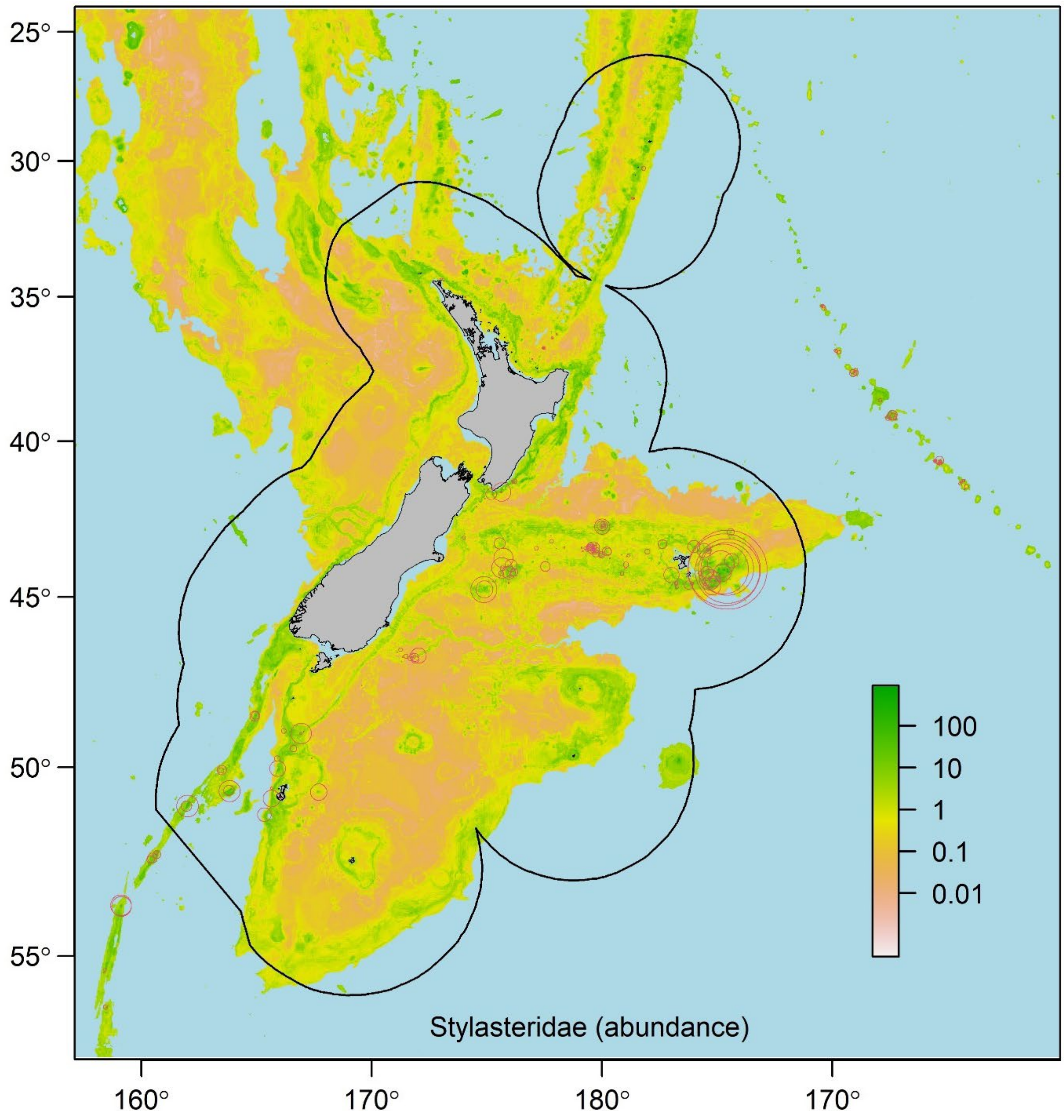

Figure 3-23: Stylasteridae (hydrocorals). Predicted abundance (number of individuals per 1000 m^2). Predictions are based on an ensemble of BRT and RF hurdle models. Note: values are shown on a log scale; red circles indicate observed abundance at presence locations, largest circle size = 3362 individuals per 1000 m^2.

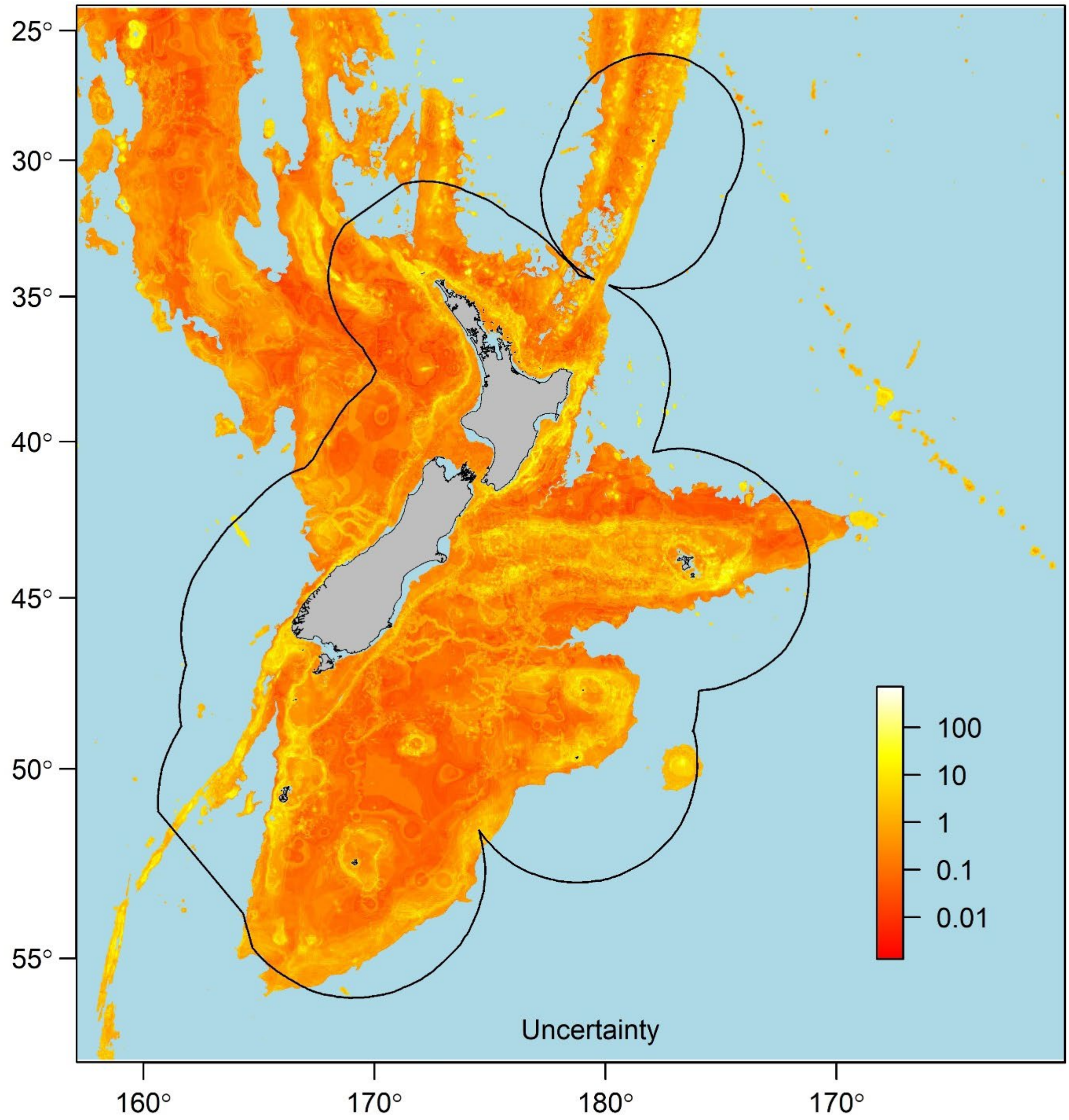

Figure 3-24: Stylasteridae (hydrocorals). Model uncertainty (Standard Deviation).

3.4 Hotspots of coral distribution

The raster layer of overall coral abundance produced from the individual model output layers was plotted as a map with abundance in four categories: 0–10, 10–20, 20–40, and >40 individuals per 1000 m^2 (Figure 3-25).

Abundance of the protected corals modelled were predicted to be less than 10 per 1000 m^2 across about 32% of the modelled area (the area within the New Zealand region in depths of 50–3000 m). Abundance of 10–20 per 1000 m^2 was predicted for about 54% of the area, and abundance of 20–40 per 1000 m^2 for about a further 12% of the area.

The highest predictions of abundance (over 40 per 1000 m^2) represent about 2% of the modelled area, with this (arbitrary) cut-off allowing the key areas to be highlighted. These key areas are spread widely around the region: in the northwest they occur in a band encircling the outer part of the

Challenger Plateau (Figure 3-26), in patches along the western margin of the Lord Howe Rise (strongly associated with the Lord Howe Seamount Chain), in a large area between the West Norfolk Ridge and Norfolk Ridge, and various small patches on other parts of these features as well as the Three Kings Ridge; in the northeast they occur in small patches dotted along the crests of the Colville and Kermadec ridges (Figure 3-26), and around numerous individual features in the northern part of the Louisville Seamount Chain; in the centre of the region, hotpots occur around numerous small features along the shelf of the North Island east coast, close to parts of the Fiordland coast, and on the Chatham Rise (especially adjacent to the Mernoo and Veryan Banks in the west, the rise centre, and east of the Chatham Islands (Figure 3-26)).

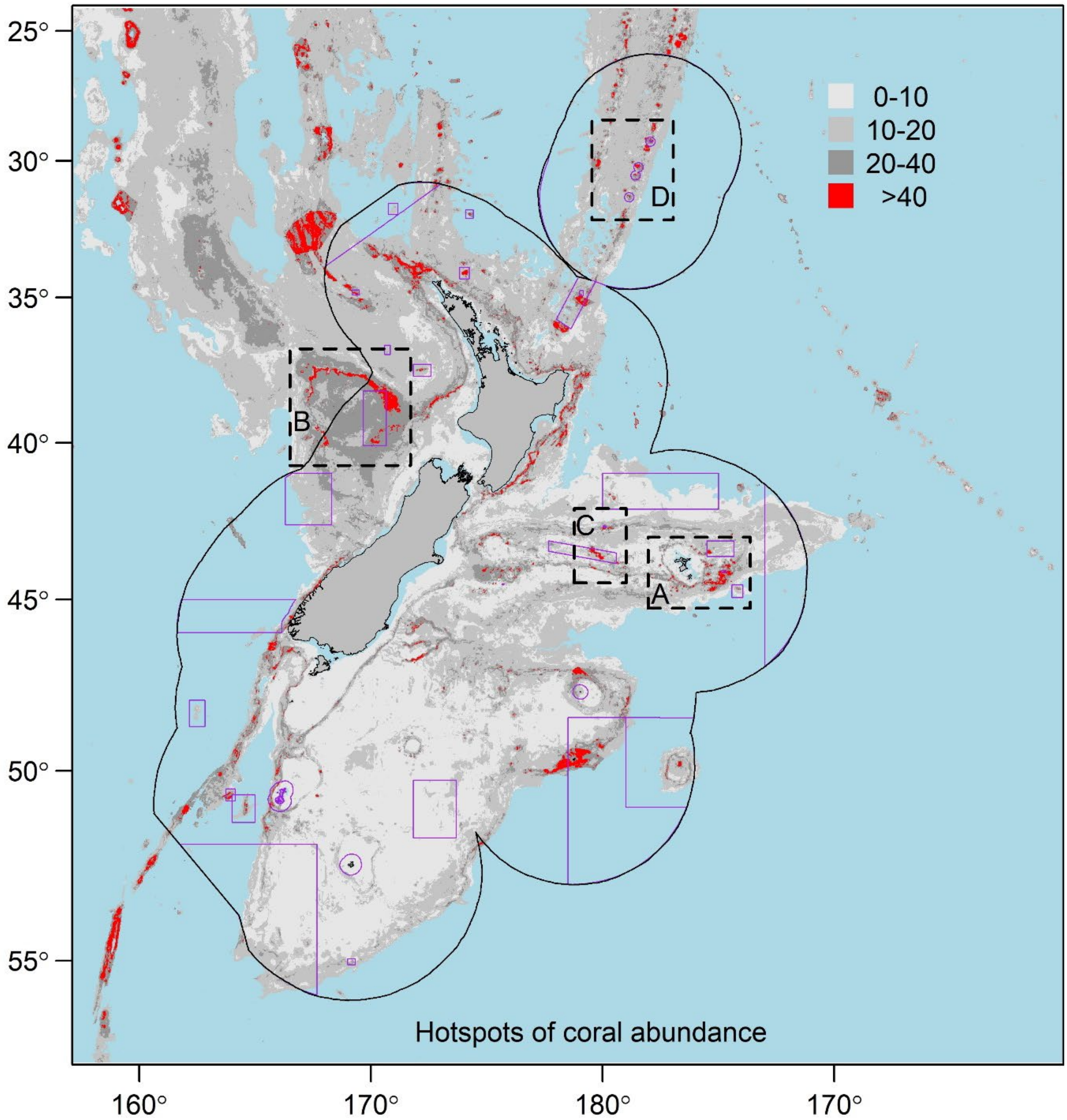

Figure 3-25: Hotspots of protected coral in the New Zealand region. Abundance was calculated by adding together estimates from each model, so that the plot represents the total number of individuals per 1000 m^2 across all modelled taxa. Areas of greatest abundance of protected corals (>40/1000 m^2) are highlighted in red. Boundaries of existing Benthic Protection Areas, Marine Reserves, and closed seamounts are shown in purple. Focus boxes for Figure 3–26 delineated with dashed lines.

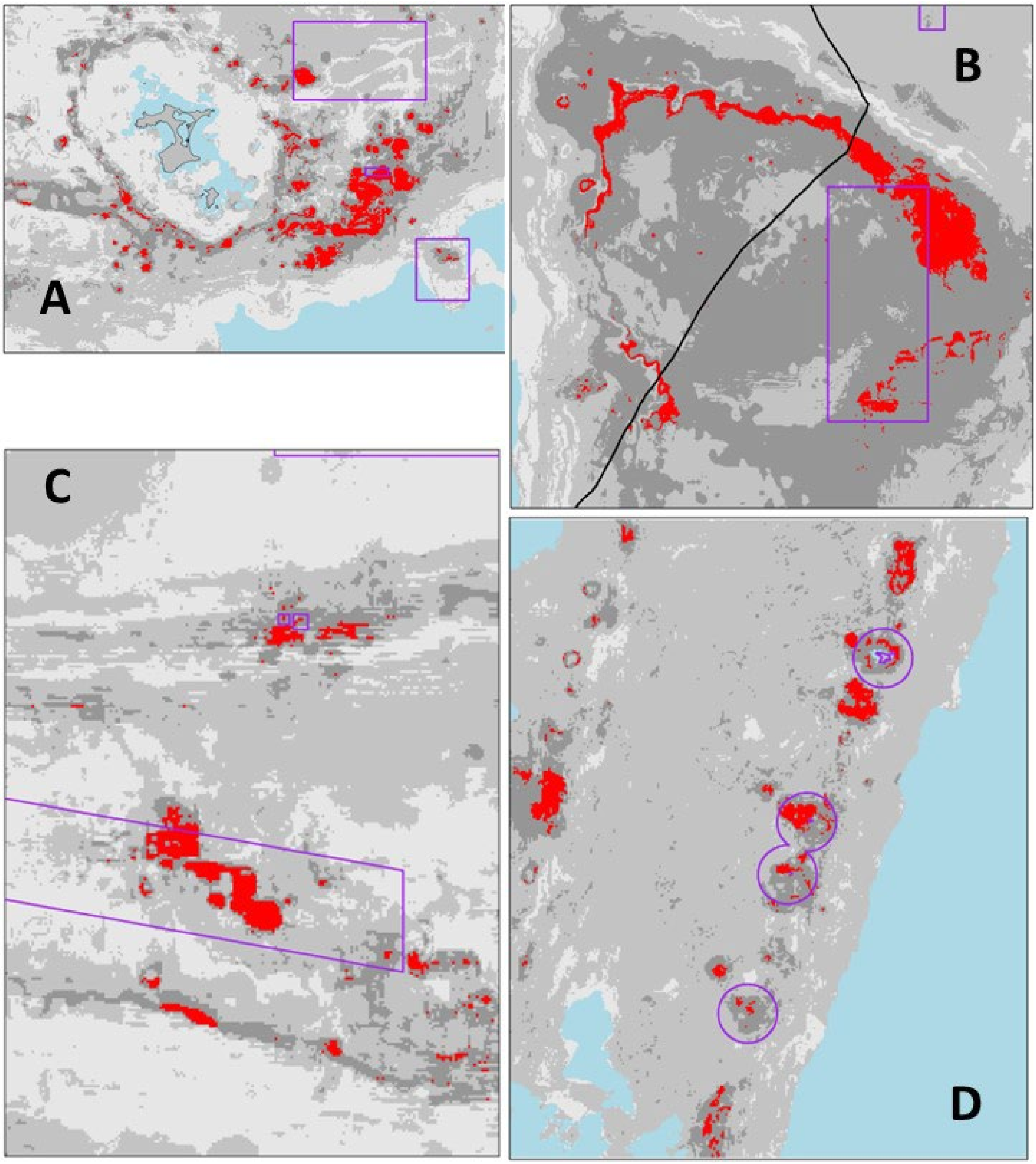

Figure 3-26: Hotspots of protected coral, close-ups. Hotspots of protected coral in the New Zealand region. Close-up maps for: A, Southeast Chatham Rise (including Andes Seamounts); B, Challenger Plateau; C, Central Chatham Rise (including Graveyard Seamounts in the north); D, Kermadec Ridge. Other details as for Figure 3-25.

3.5 Influence of fishing effort – predicting pre-1990-fishing distributions

The fishing effort variable (SAR) was excluded from the models for many taxa during the *randomforest* variable selection process due to its lack of predictive power, and for other taxa this variable had only a small influence in the models, ranking as the least or next to least most influential predictor. However, for two taxa, *Radicipes* spp. and *Goniocorella dumosa*, SAR ranked within the top three most important predictors in the presence/absence models (see Table 3-2). Predictions were therefore made for these two taxa by fitting the ensemble hurdle to the predictor data with SAR set to zero, to simulate levels of fishing disturbance at pre-1990 levels (Figure 3-27).

There were few obvious differences between the two maps of predicted abundance for *Radicipes* spp., with very slightly greater abundance predicted in eastern parts of the Campbell Plateau and Chatham Rise, but lower abundance predicted in the western Chatham Rise and off the west coast of the South Island (Figure 3-27, top). For *Goniocorella dumosa* too there were few differences between the two maps, with slightly greater abundance in parts of the western Chatham Rise and fringes of the Stewart-Snares and Auckland Islands shelves but lower abundances in some northern areas (Figure 3-27, bottom).

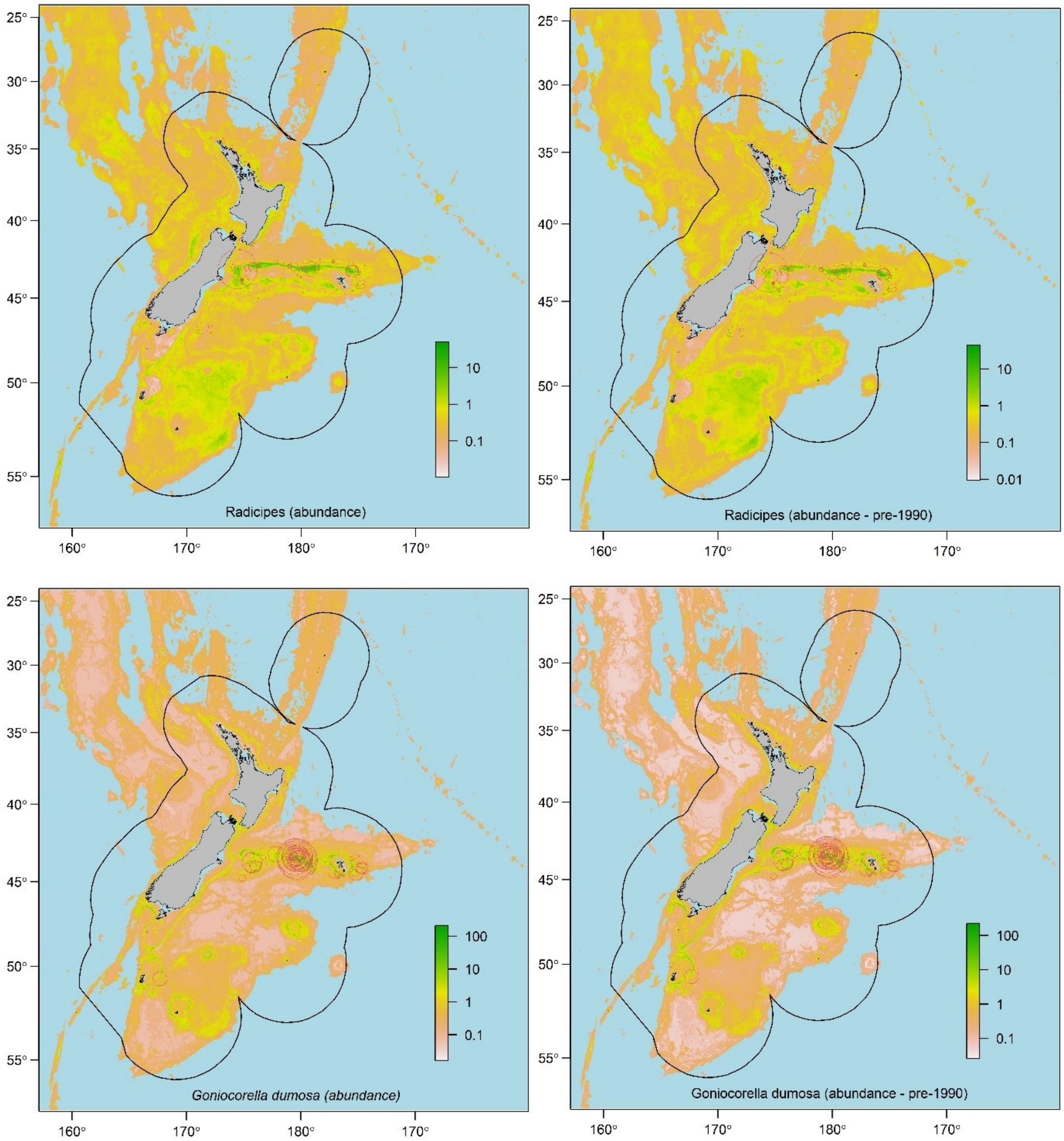

Figure 3-27: Comparison of abundance predictions before and after recorded fishing effort. Top, *Radicipes* spp., bottom, *Goniocorella dumosa*. Model coefficients are fitted to data with SAR set at values representing total historic fishing effort (left), and to data with SAR set at 0 to represent disturbance before recorded fishing effort (pre-1990) (right).

4 Discussion

This study represents the first attempt to build species distribution models for predicting abundance of a suite of protected corals across the entire New Zealand region, and the first attempt to produce abundance models across this extent using solely image-based sample data. Presence data were limited for several taxa, yet it was possible to build acceptable models for 11 of the 15 taxa attempted.

Performance metrics indicated that the models generally fitted very well to the input data, with AUC values all in the “excellent” or “adequate” category even when testing with evaluation data excluded from model building, and comparisons of observed vs predicted values for the hurdle models showed generally strong correlations.

The primary objective of the study, to predict the distribution of hot-spots of protected corals, was achieved using a somewhat arbitrary definition of a hot-spot (locations where predicted densities of all taxa combined are predicted to be over 40 individuals per 1000 m^2). Although it was necessary for descriptive purposes to make such a definition, the underlying raster of combined abundance can be used to identify hotspot locations at any lower or higher threshold, with maximum predicted densities of over 500 individuals 1000 m^{-2} in a few cells east of the Chatham Rise. It should also be noted that this representation of hot-spots is specific to the particular set of coral taxa modelled at the taxonomic resolution that it was possible to achieve.

A comprehensive comparison of the species abundance distributions produced here with habitat suitability model-based species distributions from earlier studies is not attempted here; noting that this is complicated by the uncertain relationship between relative habitat suitability and abundance estimates from species distribution modelling (SPRFMO 2020). However, some broad, descriptive comparisons can be made.

For the branching coral taxa, there is general alignment between areas of high abundance (as defined here in terms of numbers of individuals/1000 m^2) and high habitat suitability (as values between 0 and 1) for *Goniocorella dumosa* as predicted by Anderson et al. (2014) and Anderson et al. (2015) (with the Chatham Rise a key region for this species in each study), but similarities are less clear for *Solenosmilia variabilis* for which presence locations in the abundance models were almost exclusively limited to the Louisville Seamount Chain – an area not covered in the earlier studies. There was broad agreement between the three models for the bamboo corals too, with areas of high abundance/habitat suitability located on the eastern Chatham Rise and on the Challenger Plateau in particular. The predictions from each study highlighted the Bounty Trough, south Chatham Rise, and Challenger Plateau as prime locations for Primnoidae; and northern regions in general, along with the north-eastern Chatham Rise, for Antipatharia.

Cumulative historical fishing effort (SAR) proved to be of little predictive value for most of the models, with the RF variable selection procedure excluding this variable from models for five of the 11 taxa and ranking it in the bottom two predictors for four other taxa. For two taxa however, *Radicipes* spp. and *Goniocorella dumosa*, SAR was the first or second ranked predictor and therefore it was reasonable to attempt estimation of abundance for these taxa in an environment where disturbance from fishing was at levels pre-dating recorded spatial effort data (1990). However, comparison of these abundance estimates with the main results for these taxa showed only weak differences in spatial patterns. One possible reason for this may be that areas where conditions are highly suitable for these corals may only weakly overlap with areas that are highly productive for commercial fish species or, perhaps more likely, a lack of a temporal spread in the coral sample data may be limiting the model's ability to detect negative effects from fishing disturbance, with DTIS surveys spanning only the period from 2007 to 2020.

4.1 Recommendations for future model development

The effects of spatial autocorrelation were not specifically accounted for in the models constructed here, with methods used previously for this purpose (e.g., Rowden et al. 2017) not easily incorporated into a bootstrap resampling approach. The strong performance metrics may suggest that this was not a substantial issue here but nevertheless development of methods to address this issue would be worthwhile in future modelling efforts, especially given the uneven geographic distribution of sample data.

Methods for incorporating specimen-based and fisheries bycatch-based presence records into abundance models should continue to be investigated. If this can be done satisfactorily then the spatial spread of input data can be much improved, as fishing grounds and research surveys cover a wider area than the current set of camera-based surveys.

Further improvements in model performance and reliability, particularly in currently unsampled predictor space, will be possible as sample data from unprocessed DTIS video transects become available. There are many such transects, from several surveys (including the remaining stations from the three surveys from which videos were analysed for this study) for which post-processing of video data was not required at the time.

Additionally, some key areas have not been covered using video transects, notably the lower Lord Howe, Norfolk and Three Kings and Colville ridges in the north, the west coast of the South Island and the southern areas of Bounty Rise and Campbell Plateau in the south. Some of these regions are clearly predicted in this study to be hotspots of coral abundance for the region, based on favourable environmental conditions, and the collection of additional seafloor image data from these locations would be highly valuable for both model validation and improvement of model predictions.

Finally, the comprehensive dataset acquired in this project provides further opportunities to model abundance of other vulnerable, non-coral, taxa such as biogenic habitat forming species (e.g. bryozoans and sponges). The data-compilation procedures and modelling R code developed during this study mean that this can now be achieved with much lower effort and cost than previously.

5 Acknowledgements

We thank the many research voyage participants and image analysis experts who have contributed to the compilation of abundance data used in these models. We are also grateful to Matt Bennion for his advice on developing the model structure and providing updated scripts to help automate the process. Thank you also to Judy Sutherland for review and useful comments for improving the report. We acknowledge the support of Lyndsey Holland, Conservation Services Programme, Department of Conservation — Te Papa Atawhai, this work being undertaken through Project POP2021-02.

6 References

Allouche, O., Tsoar, A., Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology, 43*: 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Anderson, O., Tracey, D., Bostock, H., Williams, M., Clark, M. (2014) Refined habitat suitability modelling for protected coral species in the New Zealand EEZ. *NIWA Client Report* WLG2014-69. Prepared for Marine Species and Threats, Department of Conservation, December 2014.

Anderson, O.F., Mikaloff-Fletcher, S.E., Bostock, H.C. (2015) Development of models for predicting future distributions of protected coral species in the New Zealand region. *NIWA Client Report* WLG2015-65. Prepared for Department of Conservation.

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 broadscale models in fisheries management. *Ocean & Coastal Management,* 120: 110–126. https://doi.org/10.1016/j.ocecoaman.2015.11.025

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 Research Part I: Oceanographic Research Papers,* 115: 265–292. https://doi.org/10.1016/j.dsr.2016.07.006

Anderson, O., Stevenson, F., Behrens, E. (2020) Updated habitat suitability modelling for protected corals in New Zealand waters. FINAL REPORT Prepared for Department of Conservation. *NIWA Client Report* 2020174WN

Anderson, O.F., Stephenson, F., Behrens, E., Rowden, A.A. (2022) Predicting the effects of climate change on deep-water coral distribution around New Zealand—Will there be suitable refuges for protection at the end of the 21st century? *Global Change Biology*. DOI: 10.1111/gcb.16389

Anderson, T.J., Morrison, M., MacDiarmid, A., D'Archino, R., Nelson, W., Tracey, D., Clark, M., Gordon, D., Read, G., Morrisey, D., Kettles, H., Wood, A., Anderson, O., Smith, A.M., Page, M., Paul-Burke, K., Schnabel, K., Wadhwa, S. (2018) Review of New Zealand's Key Biogenic Habitats. NIWA Client Report 2018139WN: 184.

Baird, S.J., Tracey, D., Mormede, S., Clark, M. (2013). The distribution of protected corals in New Zealand waters. *NIWA Client Report* WLG2012-43

Baird, S.J., Mules, R. (2021) Extent of bottom contact by commercial trawling and dredging in New Zealand waters, 1989–90 to 2018–19. *New Zealand Aquatic Environment and Biodiversity Report No. 260.* 157 p. https://fs.fish.govt.nz/Doc/24909/AEBR-260-Extent-Of-Commercial-Trawl-And-Dredge-Bottom-Contact-1990-2019-4152.pdf.ashx

Bostock, H., Fletcher, S.M., Williams, M. (2013) Estimating carbonate parameters from hydrographic data for the intermediate and deep waters of the Southern Hemisphere oceans. *Biogeosciences, 10*: 6199. https://doi.org/10.5194/bg-10-6199-2013

Bostock, B., Jenkins, C., Mackay, K., Carter, L., Nodder, S., Orpin, A., Pallentin, A., Wysoczanski, R. (2018a) Distribution of surficial sediments in the ocean around New Zealand/Aotearoa. Part A: Continental slope and deep ocean. *New Zealand Journal of Geology and Geophysics*. https://doi.org/10.1080/00288 306.2018.1523198

Bostock, B., Jenkins, C., Mackay, K., Carter, L., Nodder, S., Orpin, A., Pallentin, A., & Wysoczanski, R. (2018b). Distribution of surficial sediments in the ocean around New Zealand/Aotearoa. Part B: Continental shelf. *New Zealand Journal of Geology and Geophysics*. https://doi.org/10.1080/00288 306.2018.1523199

Bowden, D.A., Hewitt, J. (2012) Recommendations for surveys of marine benthic biodiversity: outcomes from the Chatham-Challenger Ocean Survey 20/20 Post-Voyage Analyses Project. *New Zealand Aquatic Environment and Biodiversity Report No. 91*. 34 p. https://fs.fish.govt.nz/Doc/22979/AEBR_91.pdf.ashx

Bowden, D.A., Jones, D.O.B. (2016) Towed Cameras. In: Clark, M.R., Rowden, A.A., Consalvey, M. (eds). Biological Sampling in The Deep Sea, pp. 260-284. Wiley & Sons.

Bowden, D.A., Anderson, O., Escobar-Flores, P., Rowden, A., Clark, M. (2019) Quantifying benthic biodiversity: using seafloor image data to build single-taxon and community distribution models for Chatham Rise. Wellington, New Zealand: *New Zealand Aquatic Environment and Biodiversity 235*.

Bowden, Chin, & Carter. (2022) Analysing existing seafloor imagery to inform biodiversity mapping and conservation. *NIWA Client Report* 2022218WN

Breiman, L. (2001) Random forests. *Machine Learning, 45*: 5–32.

Cael, B., Bisson, K., and Follett, C. L. (2018) Can rates of ocean primary production and biological carbon export be related through their probability distributions? *Global Biogeochemical Cycles, 32*: 954–970. https://doi.org/10.1029/2017GB005797

Clark, M.R., Bowden, D.A., Rowden, A.A., Stewart, R. (2019) Little Evidence of Benthic Community Resilience to Bottom Trawling on Seamounts After 15 Years. *Front. Mar. Sci. 6:63*. doi: 10.3389/fmars.2019.00063

Clark, M.R., Wood, B., Mackay, K., Anderson, O.F., Hart, A., Rickard, G., Rowden, A.A. (2022) Underwater Topographic Features in the New Zealand region: development of an updated 'SEAMOUNT' database and information on the extent and intensity of deep-sea trawl fisheries on them. *New Zealand Aquatic Environment and Biodiversity Report No. 291*. 28 p. https://www.mpi.govt.nz/dmsdocument/53304-AEBR-291-Underwater-Topographic-Features-in-the-New-Zealand-region-development-of-an-updated-SEAMOUNT-database-and-information-on-the-extent-and-intensity-of-deep-sea-trawl-fisheries-on-them

Elith, J., Leathwick, J. R., Hastie, T. (2008) A working guide to boosted regression trees. *J. Anim. Ecol. 77*, 802–813. doi: 10.1111/j.1365-2656.2008.01390.x

FAO, (2009) International Guidelines for the Management of Deep-Sea Fisheries in the High-Seas. Food and Agriculture Organisation of the United Nations, Rome, Italy, 73 p.

Garcia, H.E., Boyer, T.P., Locarnini, R.A., Antonov, J.I., Mishonov, A.V., Baranova, O.K., Zweng, M.M., Reagan, J.R., Johnson, D.R., Levitus, S. (2013) World ocean atlas 2013. Volume 3, Dissolved oxygen, apparent oxygen utilization, and oxygen saturation. https://www.nodc.noaa.gov/OC5/woa13/

Georgian, S.E., Anderson, O.F., Rowden, A.A. (2019) Ensemble habitat suitability modelling of vulnerable marine ecosystem indicator taxa to inform deep-sea fisheries management in the south Pacific Ocean. *Fisheries Res. 211*, 256–274. doi: 10.1016/j.fishres.2018.11.020

Hill, P. (2009) Designing a deep-towed camera vehicle using single conductor cable. *Sea Technology, 50*: 49–51.

Hosmer Jr, D.W., Lemeshow, S., Sturdivant, R.X. (2013) Applied Logistic Regression, John Wiley b& Sons.

Jones, E.G., Morrison, M.A., Davey, N., Mills, S., Pallentin, A., George, S., Kelly, M., Tuck, I. (2018) Biogenic habitats on New Zealand's continental shelf. Part II: National field survey and analysis. New Zealand Aquatic Environment and Biodiversity Report 202. 261 p. https://fs.fish.govt.nz/Doc/24624/AEBR-2018-202-Biogenic-habitats-on-continental-shelf-Part-II.pdf.ashx

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. https://www.nodc.noaa.gov/OC5/woa13/

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

Macpherson, D., Mills, S., Tracey, D. (2022) INT2019-04 Identification and storage of cold-water coral bycatch specimens 1 July 2020 – 30 June 2021. Final Annual Report Prepared for Conservation Services Programme, Department of Conservation, May 2022.

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. *Biological Conservation, 156*: 94–104. https://doi.org/10.1016/j.biocon.2011.11.013

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. *Marine Ecology Progress Series 397*: 309-317. http://dx.doi.org/10.3354/meps08115

Parker, S.J., Bowden, D.A. (2010) Identifying taxonomic groups vulnerable to bottom longline fishing gear in the Ross Sea region. *CCAMLR Science 17*: 105-127.

Pinkerton, M., Rickard, G., Nodder, S., Macdonald, H. (2016) Climate change impacts and implications: net primary production, particulate flux and impacts on fish species in the New Zealand region. NIWA report 62 pp.

R Core Team (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

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 Research Part I: Oceanographic Research Papers, 113*: 80–89. https://doi.org/10.1016/j.dsr.2016.04.008

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. *Frontiers in Marine Science,* 4: 335. https://doi.org/10.3389/fmars.2017.00335

Rowden, A.A., Anderson, O.F., Neubauer, P., Hamill, J., Bowden, D.A., Tremblay-Boyer, L., Charsley, A., MacGibbon, D. (in press). Spatially explicit benthic impact assessments for bottom trawling in New Zealand. *New Zealand Aquatic Environment and Biodiversity Report*.

SPRFMO (2020) Cumulative Bottom Fishery Impact Assessment for Australian and New Zealand bottom fisheries in the SPRFMO Convention Area, 2020 Australia – New Zealand. Paper SC8-DW07_rev_1, 8th MEETING OF THE SCIENTIFIC COMMITTEE, New Zealand, 3 to 8 October 2020. https://www.sprfmo.int/meetings/scientific-committee/8th-sc-2020/

Stephenson, F., Rowden, A.A., Anderson, O.F., Pitcher, C.R., Pinkerton, M.H., Petersen, G., Bowden, D.A. (2021a) Presence-only habitat suitability models for vulnerable marine ecosystem indicator taxa in the South Pacific have reached their predictive limit. *ICES Journal of Marine Science, 78*: 2830–2843. https://doi.org/10.1093/icesjms/fsab162

Stephenson, F., Bowden, D.A., Finucci, B., Anderson, O.F., Rowden, A.A. (2021b) Developing updated predictive models for benthic taxa and communities across Chatham Rise and Campbell Plateau using photographic survey data. *New Zealand Aquatic Environment and Biodiversity Report No. 276*. 82 p. https://docs.niwa.co.nz/library/public/NZAEBR-276.pdf

Tracey, D.M., Rowden, A.A., Mackay, K.A., Compton, T. (2011) Habitat-forming Coldwater corals show affinity for seamounts in the New Zealand region. *Marine Ecology Progress Series, 430*: 1–22. https://doi.org/10.3354/meps09164

Appendix A Predictor importance

Table A-1: Relative importance of model predictors. Predictors are ordered from most to least important for the presence-absence ensemble model.

Antipatharia	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
Ruggedness	16.0	19.7	17.8	13.3	8.7	11.0
Bottom temperature	12.9	18.0	15.4	13.0	16.2	14.6
Dissolved oxygen	13.8	16.2	15.0	12.1	12.6	12.4
BPI-broad	12.7	13.9	13.3	15.6	17.4	16.5
Calcite	11.1	10.7	10.9	11.2	11.2	11.2
Slope SD	12.2	9.5	10.9	13.2	12.0	12.6
POC	10.9	6.0	8.5	12.2	12.0	12.1
Percent mud	10.5	6.1	8.3	9.4	9.8	9.6

Caryophylliidae	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
Ruggedness	13.5	20.3	16.9	7.8	10.0	8.9
Bottom temperature	11.2	14.7	12.9	9.1	8.7	8.9
Dissolved oxygen	10.9	11.8	11.3	13.0	14.7	13.8
BPI-broad	11.0	11.2	11.1	13.4	18.1	15.8
Percent gravel	10.9	11.3	11.1	16.8	14.4	15.6
Percent mud	11.7	10.2	10.9	14.5	11.2	12.8
Slope SD	11.4	8.5	9.9	8.2	7.5	7.8
POC	9.8	7.1	8.5	7.5	4.8	6.2
Aragonite	9.6	4.8	7.2	9.7	10.7	10.2

Flabellum	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
POC	18.1	31.0	24.5	13.90	16.79	15.35
Dissolved oxygen	16.2	19.1	17.7	21.29	29.43	25.36
Slope SD	13.5	13.0	13.2	9.09	7.59	8.34
BPI-broad	13.6	11.1	12.4	15.01	12.98	14.00
Bottom temperature	13.3	10.7	12.0	12.54	11.80	12.17
Aragonite	14.2	9.1	11.6	11.77	7.16	9.46
Percent mud	11.2	6.0	8.6	16.39	14.26	15.32

Goniocorella dumosa	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
Dissolved oxygen	17.3	43.0	30.1	12.0	9.7	10.8
POC	14.5	12.6	13.6	18.5	23.3	20.9
Fishing effort (SAR)	14.4	10.8	12.6	4.7	5.6	5.2
Ruggedness	14.4	7.9	11.1	16.7	15.7	16.2
Aragonite	13.8	8.0	10.9	14.4	11.2	12.8
BPI-broad	12.5	9.2	10.9	16.5	17.9	17.2
Bottom temperature	13.1	8.6	10.8	17.2	16.5	16.9

Gorgonians	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
Dissolved oxygen	16.0	21.7	18.8	12.0	10.5	11.2
BPI-broad	11.9	15.1	13.5	14.7	17.9	16.3
Slope SD	13.2	12.7	13.0	12.4	11.5	11.9
Ruggedness	12.7	13.1	12.9	10.6	9.6	10.1
Calcite	12.1	12.4	12.3	10.0	10.6	10.3
Percent mud	11.7	9.2	10.4	16.1	17.1	16.6
Bottom temperature	10.7	9.0	9.8	12.4	11.4	11.9
POC	11.7	6.8	9.3	11.7	11.4	11.6

Keratoisididae/Mopseidae	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
Ruggedness	11.3	37.2	24.2	11.7	8.4	10.0
Slope SD	11.2	16.0	13.6	16.0	15.1	15.6
POC	11.6	15.3	13.4	9.7	10.1	9.9
Dissolved oxygen	11.0	10.3	10.6	10.8	11.3	11.1
Percent gravel	10.0	9.8	9.9	12.1	13.2	12.6
Percent mud	12.9	2.3	7.6	10.9	15.3	13.1
Bottom temperature	11.0	3.9	7.5	9.0	7.8	8.4
Calcite	10.5	3.2	6.9	8.4	5.2	6.8
BPI-broad	10.6	1.9	6.3	11.5	13.5	12.5

Pennatulacea	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
Aragonite	12.9	13.6	13.3	12.7	8.7	10.7
Percent mud	11.3	9.5	10.4	11.8	11.2	11.5
Percent gravel	10.8	10.0	10.4	10.9	7.9	9.4
BPI-broad	14.7	17.1	15.9	14.1	14.2	14.1
Slope SD	11.0	12.3	11.6	10.7	8.7	9.7
Dissolved oxygen	13.6	12.9	13.3	12.2	17.2	14.7
Bottom temperature	12.5	14.0	13.3	14.6	15.1	14.8
POC	13.3	10.5	11.9	13.1	17.0	15.1

Primnoidae	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
Ruggedness	13.2	40.6	26.9	8.1	5.3	6.7
Percent mud	11.9	11.9	11.9	14.1	14.7	14.4
Slope SD	11.6	11.9	11.8	9.0	8.1	8.6
POC	10.9	9.1	10.0	13.1	16.2	14.6
BPI-broad	10.5	6.5	8.5	10.6	9.4	10.0
Calcite	10.7	6.3	8.5	9.4	7.8	8.6
Dissolved oxygen	11.3	4.4	7.8	11.3	10.8	11.1
Bottom temperature	10.8	4.3	7.5	13.4	14.2	13.8
Percent gravel	9.1	5.0	7.0	11.0	13.5	12.3

Radicipes spp.	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
Fishing effort (SAR)	10.2	57.0	33.6	8.7	8.0	8.4
Percent gravel	10.4	30.6	20.5	10.2	14.4	12.3
POC	11.7	7.0	9.3	14.1	18.9	16.5
BPI-broad	12.5	1.9	7.2	11.2	7.8	9.5
Calcite	11.3	1.4	6.4	11.1	8.6	9.9
Bottom temperature	11.6	1.0	6.3	11.0	8.9	9.9
Slope SD	10.7	1.1	5.9	10.2	10.3	10.3
Dissolved oxygen	10.9	0.0	5.5	13.3	13.9	13.6
Percent mud	10.6	0.1	5.3	10.1	9.1	9.6

Solenosmilia variabilis	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
BPI-broad	46.6	86.1	66.3	46.4	52.3	49.3
Slope SD	29.7	10.2	19.9	46.5	47.6	47.0
Seamount	23.7	3.7	13.7	7.1	0.1	3.6

Stylasteridae	Presence/absence			Abundance		
	RF	BRT	Ensemble	RF	BRT	Ensemble
Ruggedness	21.6	28.2	24.9	16.8	14.1	15.4
BPI-broad	19.7	18.4	19.1	20.7	21.7	21.2
POC	17.0	17.6	17.3	21.1	20.4	20.8
Percent gravel	16.7	12.8	14.7	23.4	29.9	26.7
Slope SD	16.6	8.5	12.6	14.6	11.7	13.1
Seamount	8.4	14.5	11.4	3.4	2.1	2.7

Appendix B Model performance

Table B-1: Model performance metrics.

Antipatharia		**Presence/absence**			**Abundance**		**Hurdle**	
		RF	**BRT**		**RF**	**BRT**	**RF**	**BRT**
Training data	RMSE	0.23	0.25	Rsq	0.61	0.84	0.46	0.36
	AUC	0.97	0.90	NRMSE	0.50	0.32	6.00	6.56
	TSS	0.83	0.83	Correlation	0.78	0.92	0.79	0.74
	Cutoff	0.27	0.29					
Evaluation data	RMSE	0.35	0.36	Rsq	0.17	0.13	0.07	0.03
	AUC	0.85	0.84	NRMSE	0.76	0.82	7.52	7.48
	TSS	0.59	0.57	Correlation	0.41	0.35	0.18	0.19
	Cutoff	0.23	0.20					

Caryophyllidae		**Presence/absence**			**Abundance**		**Hurdle**	
		RF	**BRT**		**RF**	**BRT**	**RF**	**BRT**
Training data	RMSE	0.25	0.29	Rsq	0.63	0.86	0.70	0.40
	AUC	0.96	0.87	NRMSE	0.46	0.29	2.73	3.89
	TSS	0.80	0.75	Correlation	0.79	0.93	0.91	0.75
	Cutoff	0.29	0.29					
Evaluation data	RMSE	0.37	0.39	Rsq	0.19	0.15	0.07	-0.04
	AUC	0.81	0.79	NRMSE	0.69	0.74	4.93	5.16
	TSS	0.52	0.48	Correlation	0.42	0.37	0.35	0.27
	Cutoff	0.25	0.25					

***Flabellum* spp.**		**Presence/absence**			**Abundance**		**Hurdle**	
		RF	**BRT**		**RF**	**BRT**	**RF**	**BRT**
Training data	RMSE	0.28	0.35	Rsq	0.68	0.81	0.80	0.52
	AUC	0.95	0.84	NRMSE	0.49	0.39	1.92	2.99
	TSS	0.77	0.64	Correlation	0.82	0.90	0.93	0.85
	Cutoff	0.39	0.36					
Evaluation data	RMSE	0.42	0.42	Rsq	0.31	0.22	0.31	0.18
	AUC	0.77	0.77	NRMSE	0.73	0.81	3.57	3.91
	TSS	0.43	0.44	Correlation	0.55	0.46	0.63	0.50
	Cutoff	0.34	0.33					

Goniocorella dumosa		Presence/absence			Abundance		Hurdle	
		RF	BRT		RF	BRT	RF	BRT
Training data	RMSE	0.14	0.21	Rsq	0.62	0.87	0.81	0.44
	AUC	0.98	0.89	NRMSE	0.36	0.21	2.88	4.95
	TSS	0.87	0.73	Correlation	0.79	0.93	0.94	0.85
	Cutoff	0.14	0.13					
Evaluation data	RMSE	0.22	0.23	Rsq	0.23	0.12	0.31	0.15
	AUC	0.90	0.86	NRMSE	0.55	0.64	5.40	6.02
	TSS	0.69	0.64	Correlation	0.46	0.31	0.64	0.53
	Cutoff	0.11	0.11					

Gorgonians		Presence/absence			Abundance		Hurdle	
		RF	BRT		RF	BRT	RF	BRT
Training data	RMSE	0.27	0.33	Rsq	0.65	0.83	0.59	0.38
	AUC	0.96	0.86	NRMSE	0.46	0.33	3.31	4.14
	TSS	0.81	0.73	Correlation	0.81	0.91	0.86	0.75
	Cutoff	0.45	0.44					
Evaluation data	RMSE	0.42	0.43	Rsq	0.22	0.15	0.09	0.06
	AUC	0.81	0.79	NRMSE	0.69	0.76	4.92	4.97
	TSS	0.52	0.48	Correlation	0.47	0.38	0.38	0.33
	Cutoff	0.44	0.41					

Keratoisididae/ Mopseidae		Presence/absence			Abundance		Hurdle	
		RF	BRT		RF	BRT	RF	BRT
Training data	RMSE	0.23	0.35	Rsq	0.49	0.84	0.50	0.11
	AUC	0.94	0.74	NRMSE	0.61	0.34	6.52	8.81
	TSS	0.77	0.44	Correlation	0.69	0.91	0.86	0.70
	Cutoff	0.21	0.16					
Evaluation data	RMSE	0.34	0.35	Rsq	0.04	0.04	0.00	0.01
	AUC	0.77	0.71	NRMSE	0.89	0.97	8.40	8.42
	TSS	0.46	0.36	Correlation	0.13	0.12	0.14	0.10
	Cutoff	0.17	0.15					

Pennatulacea		**Presence/absence**			**Abundance**		**Hurdle**	
		RF	**BRT**		**RF**	**BRT**	**RF**	**BRT**
Training data	RMSE	0.29	0.34	Rsq	0.64	0.81	0.52	0.27
	AUC	0.95	0.83	NRMSE	0.55	0.42	6.30	7.88
	TSS	0.78	0.73	Correlation	0.80	0.90	0.88	0.82
	Cutoff	0.43	0.47					
Evaluation data	RMSE	0.44	0.46	Rsq	0.18	0.12	0.07	-0.01
	AUC	0.77	0.73	NRMSE	0.82	0.89	7.48	7.70
	TSS	0.44	0.37	Correlation	0.42	0.34	0.41	0.20
	Cutoff	0.44	0.41					

Primnoidae		**Presence/absence**			**Abundance**		**Hurdle**	
		RF	**BRT**		**RF**	**BRT**	**RF**	**BRT**
Training data	RMSE	0.23	0.36	Rsq	0.63	0.85	0.72	0.31
	AUC	0.96	0.77	NRMSE	0.45	0.29	2.78	4.42
	TSS	0.82	0.51	Correlation	0.79	0.92	0.92	0.72
	Cutoff	0.26	0.21					
Evaluation data	RMSE	0.35	0.38	Rsq	0.17	0.09	0.11	0.04
	AUC	0.82	0.71	NRMSE	0.69	0.78	5.04	5.21
	TSS	0.52	0.36	Correlation	0.41	0.28	0.43	0.28
	Cutoff	0.23	0.20					

***Radicipes* spp.**		**Presence/absence**			**Abundance**		**Hurdle**	
		RF	**BRT**		**RF**	**BRT**	**RF**	**BRT**
Training data	RMSE	0.20	0.32	Rsq	0.60	0.92	0.39	0.08
	AUC	0.94	0.69	NRMSE	0.59	0.25	9.95	11.91
	TSS	0.76	0.37	Correlation	0.77	0.96	0.77	0.70
	Cutoff	0.20	0.12					
Evaluation data	RMSE	0.30	0.32	Rsq	0.14	0.10	0.10	0.02
	AUC	0.76	0.66	NRMSE	0.90	1.02	10.03	10.38
	TSS	0.43	0.30	Correlation	0.34	0.27	0.28	0.19
	Cutoff	0.16	0.12					

Solenosmilia variabilis		Presence/absence			Abundance		Hurdle	
		RF	BRT		RF	BRT	RF	BRT
Training data	RMSE	0.17	0.19	Rsq	0.30	0.52	0.11	0.08
	AUC	0.98	0.95	NRMSE	0.49	0.40	8.60	8.80
	TSS	0.91	0.86	Correlation	0.53	0.71	0.54	0.28
	Cutoff	0.15	0.13					
Evaluation data	RMSE	0.21	0.21	Rsq	0.07	0.06	0.08	-0.14
	AUC	0.94	0.94	NRMSE	0.57	0.62	7.61	7.88
	TSS	0.87	0.85	Correlation	0.22	0.17	0.28	0.23
	Cutoff	0.17	0.11					

Stylasteridae		Presence/absence			Abundance		Hurdle	
		RF	BRT		RF	BRT	RF	BRT
Training data	RMSE	0.24	0.29	Rsq	0.69	0.87	0.67	0.49
	AUC	0.97	0.88	NRMSE	0.40	0.26	4.71	6.03
	TSS	0.83	0.73	Correlation	0.83	0.93	0.92	0.90
	Cutoff	0.30	0.30					
Evaluation data	RMSE	0.35	0.36	Rsq	0.31	0.24	0.21	0.19
	AUC	0.85	0.83	NRMSE	0.60	0.67	7.04	7.10
	TSS	0.59	0.55	Correlation	0.55	0.48	0.67	0.64
	Cutoff	0.25	0.24					