Exploring the drivers of spatial distributions of basking shark (*Cetorhinus maximus*) in the South Pacific

B. Finucci, F. Stephenson, G. Petersen, M.P. Francis, & M.H. Pinkerton

POP2020-03

Climate, Freshwater & Ocean Science

NIWA
Taihoro Nukurangi

Basking sharks in New Zealand

- Historically, basking sharks widely reported throughout NZ between 39°S and 51°S; most records south of Cook Strait

- Early 1990s, groups of 100+ from aerial surveys off Banks Peninsula. Subsequent surveys failed to find sharks

- In NZ fisheries, bycatch peaked 1988-1991

- Protected in NZ since 2010

- Small number individuals reported annually as fisheries bycatch

- Basking shark observations highly variable across years, with gaps in regional sightings up to 20 years

- Northern Hemisphere distribution and occurrence strongly linked to prey abundance (zooplankton), sea surface temperature, thermal fronts, chl-a concentration

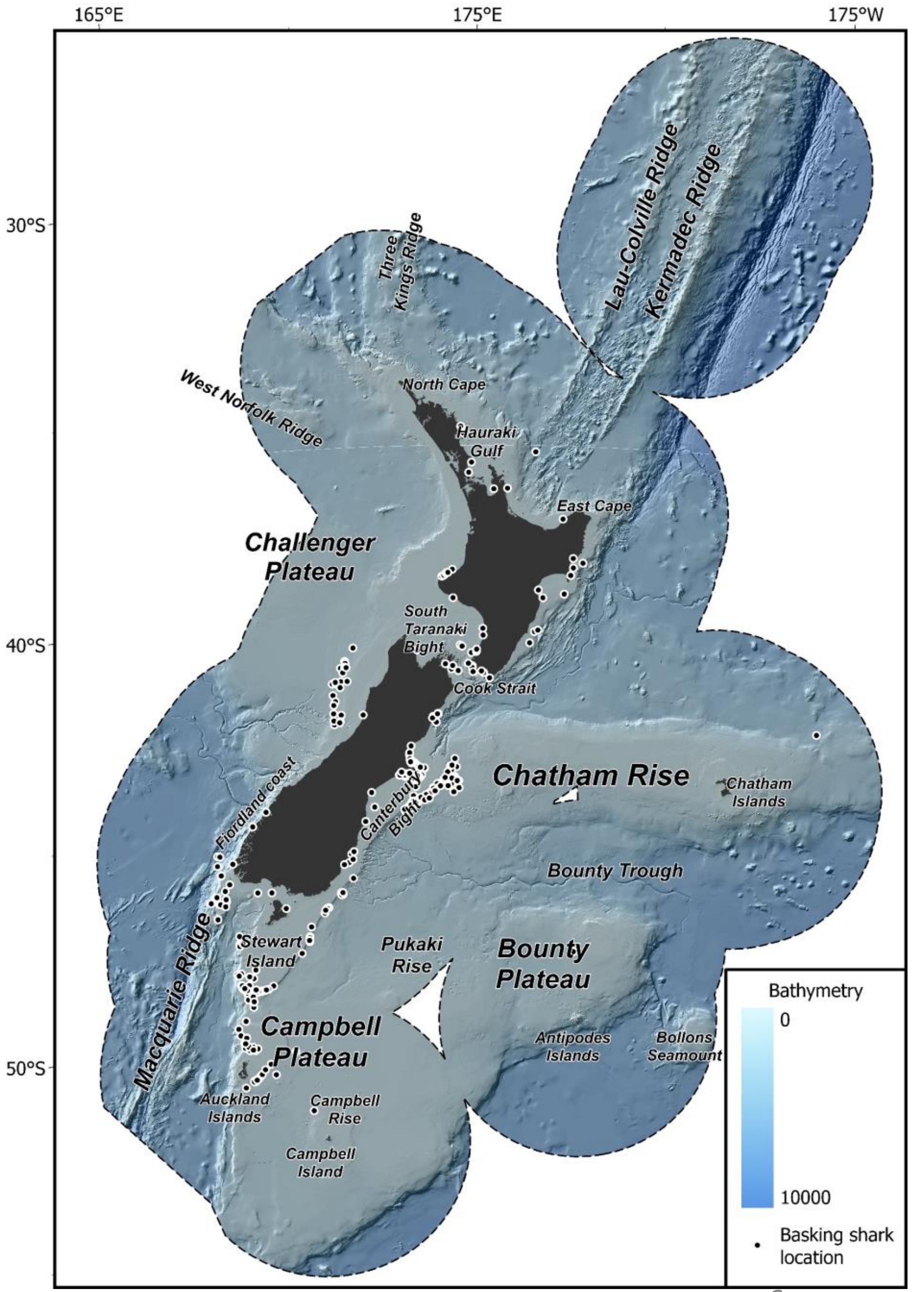
- Need sufficient information on distribution, habitat use, migratory patterns to determine cause in abundance variability

Project Aim CSP POP2020-03

Use existing data on species' occurrence to improve our understanding of basking shark distribution in New Zealand waters, identify environmental factors that potentially drive changes in species' distribution

Study area and NZ basking shark records

- Basking shark records collated from various sources, including commercial fisheries, public sightings, media report, museum records, scientific surveys, and beach cast specimens
- All catch records were converted into presence records
- To minimize effect of spatial bias in occurrence data, records aggregated spatially to a 1 km2 grid resolution
- Final dataset of 369 unique sampling locations

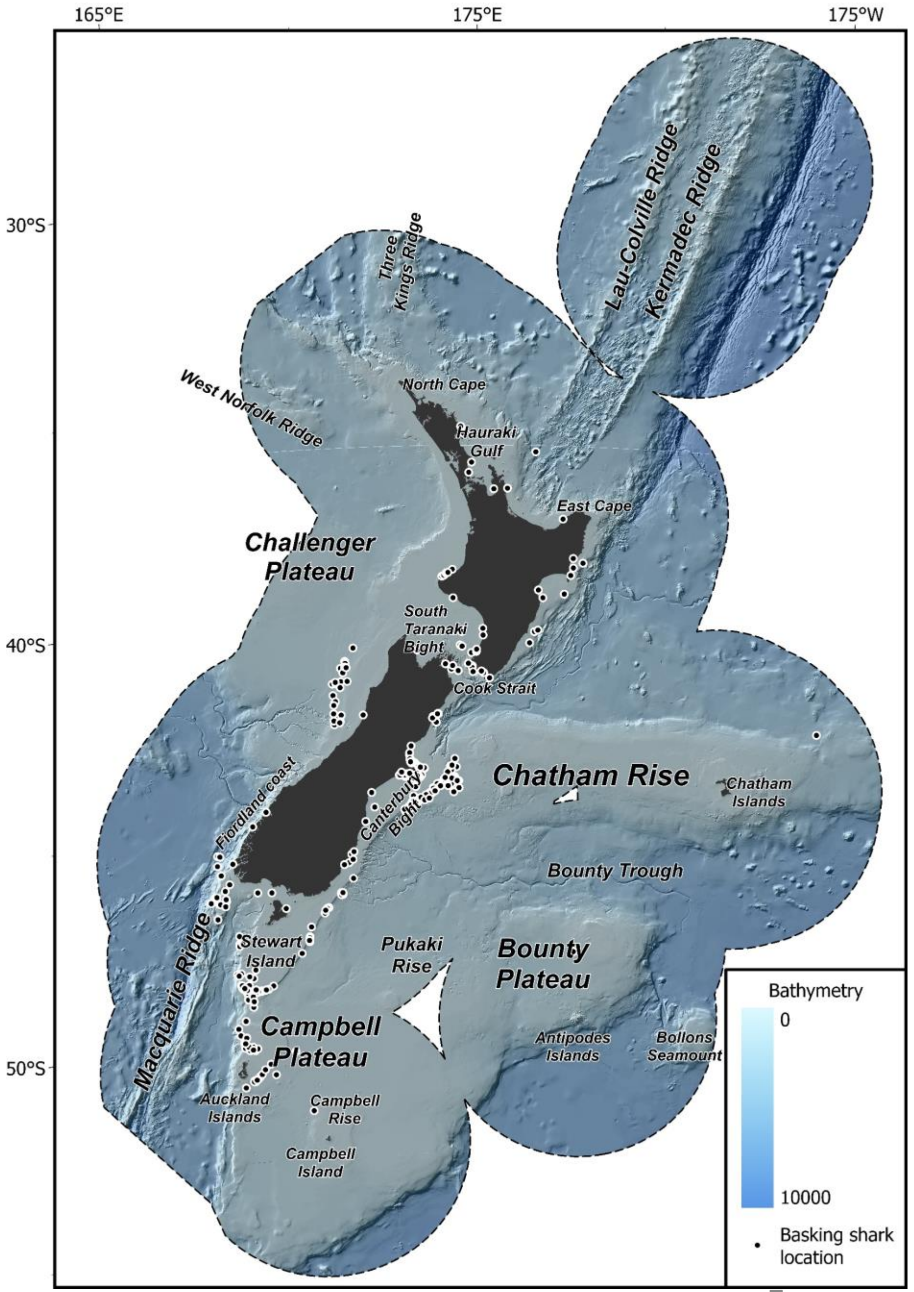


Habitat Suitability Modelling (HSM)

- Ensemble predictions (Ensemble HSM) from **Boosted Regression Tree** (BRT) and **Random Forest** (RF) models

- **BRT**: combines many individual regression trees and boosting, fitted with tree complexity of 2, learning rate of 0.01, bag fraction of 0.7 and random 10-fold cross evaluation

- **RF**: fits ensemble of regression or classification tree models describing relationship between species' distribution and environmental variables, optimal values for complexity parameters (mtry, maxnodes, ntrees) selected by train function in R package 'caret'

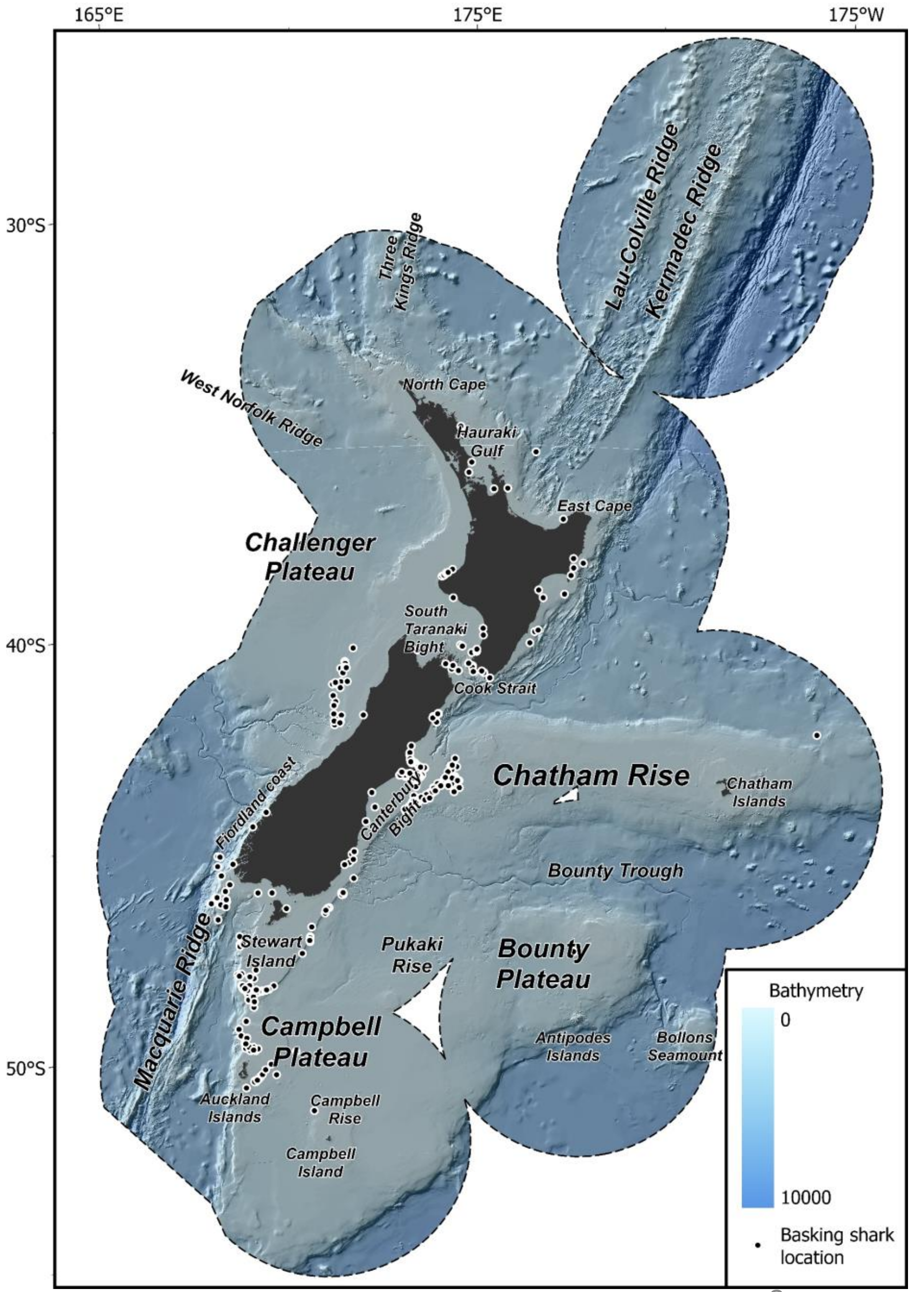


Habitat Suitability Modelling (HSM)

- Models require presence (occurrence records) and absences

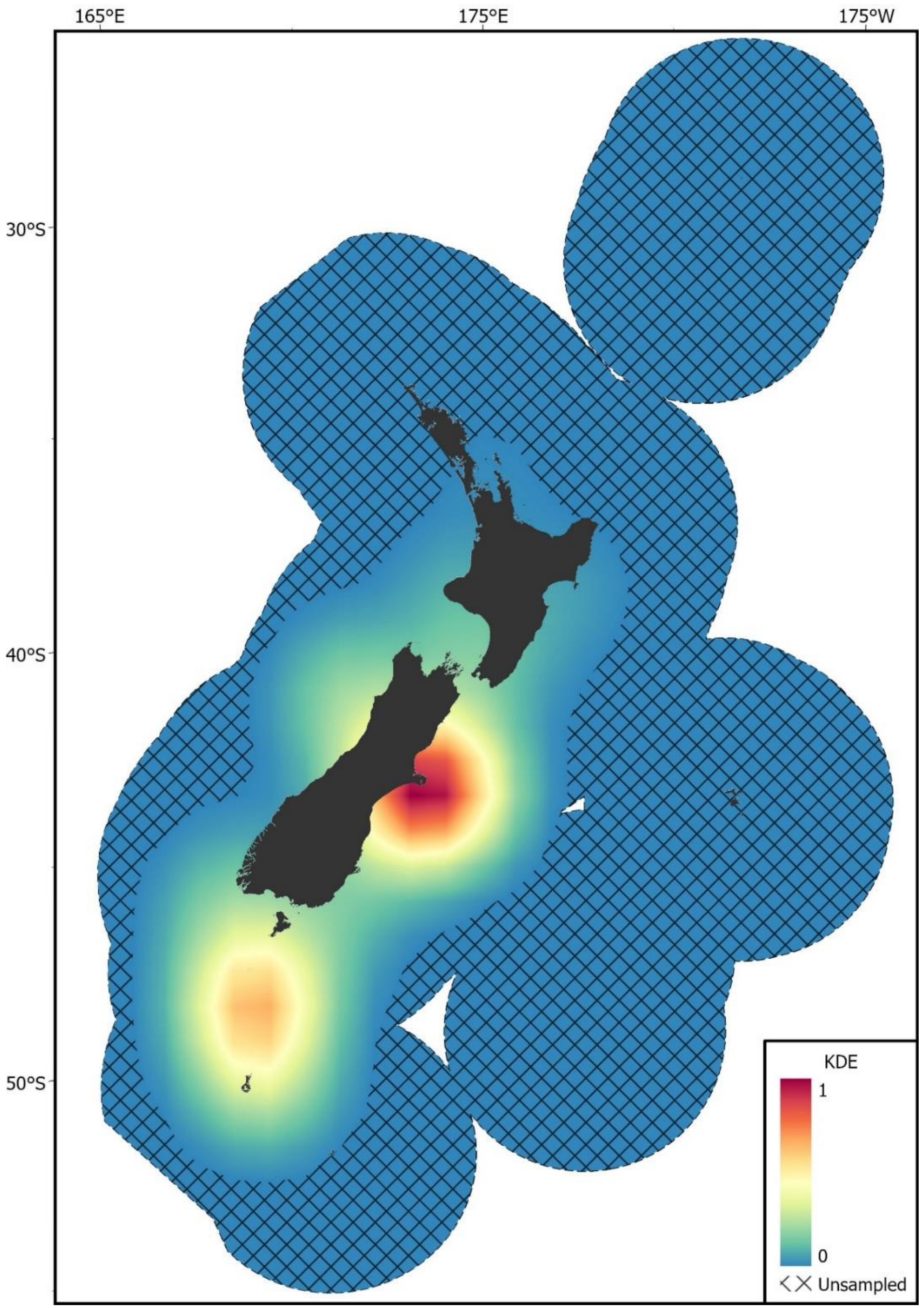
- True absences (sample location where no shark recorded) not easily available for sampling methods (observer records) or not available for opportunistic records (public sighting)

- Creation of pseudo absences instead



Spatial coverage of species occurrence, pseudo-absences

- A two-dimensional kernel density estimate (KDE) was produced using presence data, cell size 1 km2
- Within KDE, 95% volume contour selected (where 95% data located)
- Probability grid created from KDE to sample pseudo absences according to probability of grid weights; random selection
- Number of pseudo absences selected by month equal to number of monthly presences



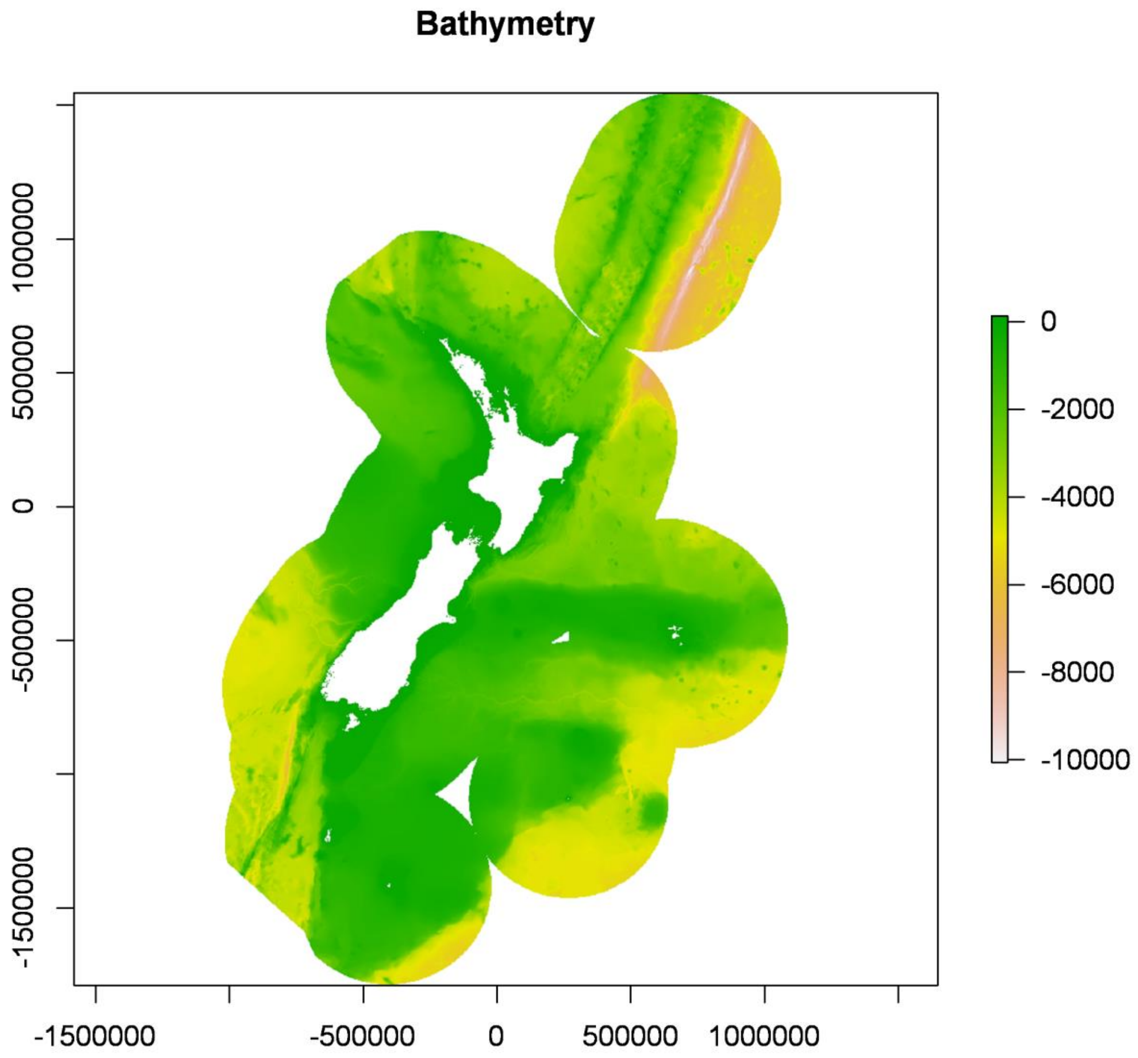
Environmental and Biotic Predictors

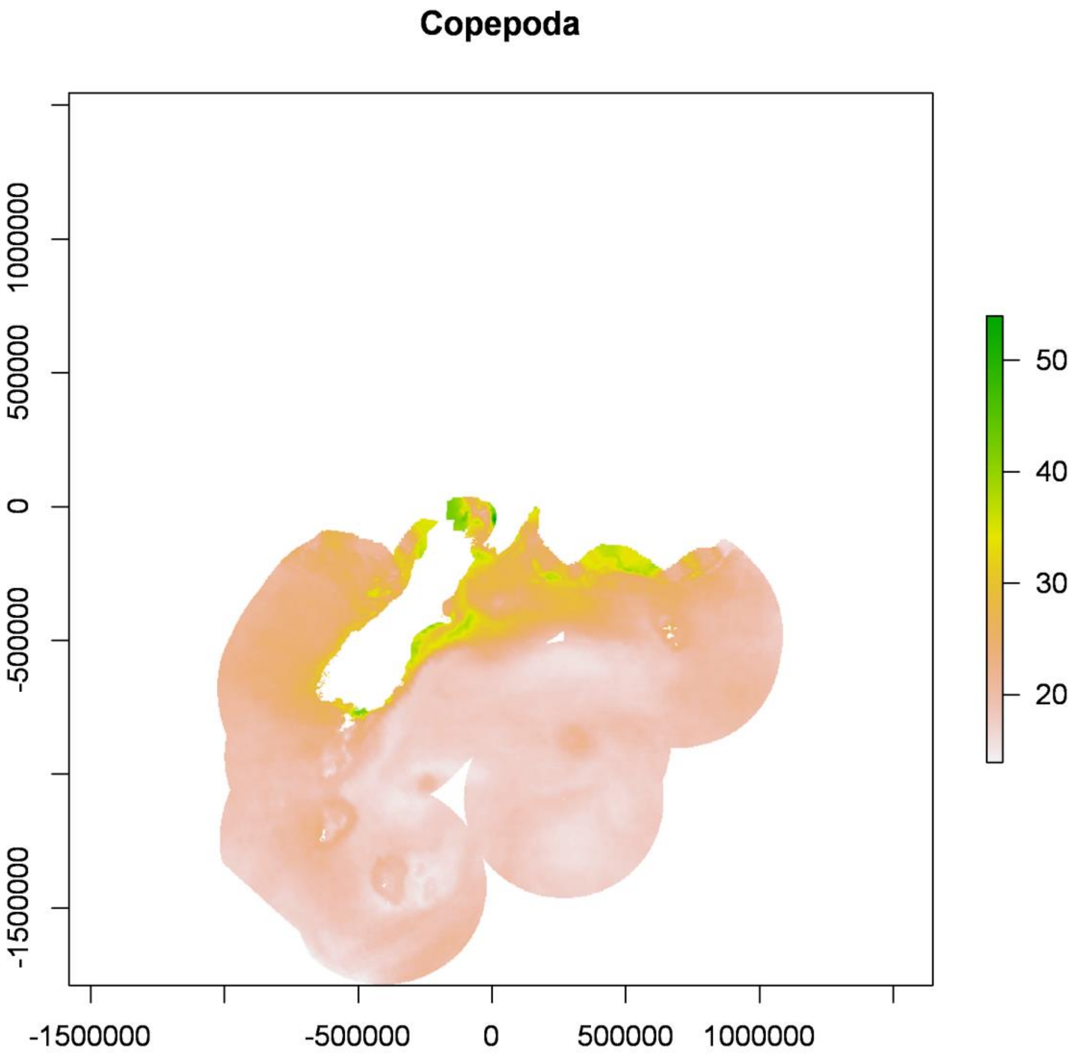
Environmental predictors

- 1 km grid resolution
- Cover entire NZ EEZ
- Static (bathymetry) and dynamic (chl-a)
- Dynamic predictors mean monthly temporal resolution
- Full list reported in Stephenson et al. 2020

Biotic predictors

- 1 km grid resolution
- Available south of 40
- Static (counts per 5 nautical mile Continuous Plankton Recorder (CPR) segment
- Full list reported in Pinkerton et al. 2020


Bathymetry
1000000
500000
0
-500000
-1500000
-1500000
-500000
0
500000
1000000
0
-2000
-4000
-6000
-8000
-10000


Copepoda
1000000
500000
0
-500000
-1500000
-1500000
-500000
0
500000
1000000
50
40
30
20

Environmental and Biotic Predictors – Model tuning

- Model tuning to reduce number of variables in order to produce parsimonious model

- Fitted BRT model with all variables, removed one at a time using "simplify" function by assessing relative contribution of each term (deviance explained) and remove lowest contributing variables

- Refit model and repeat process until deviance explained decreased by >1% between removal of predictor variables

Environmental and Biotic Predictors – Final predictors

Final Environmental predictors (8)

- Bathymetry
- Bathymetric position index (BPI-broad)
- Chlorophyll-a concentration (chl-a)
- Mixed layer depth (MLD)
- Downward vertical flux of particulate organic matter at seabed
- Turbidity
- Slope
- Sea surface temperature (SST)

Final Biotic predictor (1)

- Copepoda

Some co-linearity observed, but considered acceptable for tree-based machine learning methods (Pearson correlation < 0.75)

Model application

- BRT and RF models bootstrapped 200 times
- Random ‘training’ sample of total presence records drawn with replacement. Random sample of pseudo absence of equal number drawn without replacement
- Presence records not randomly selected combined with a random number of pseudo absences and set aside for independent assessment of model performance (‘evaluation’ data)
- Geographic prediction made using environmental predictor variables to 1 km2 grid
- Ensemble models produced with weighted average of predictions from each model type

Model performance and uncertainty

- Model performance evaluated using mean **AUC** (range 0 - 1, >0.7 good) and **TSS** (range -1 – 1, >0.6 good) scores calculated for ensemble model

- Two measures of spatially explicit uncertainty: estimate of spatial coverage of species occurrence (**95% KDE**) and SD of predicted shark distribution (**model uncertainty**)

- Partial dependence plots were made for the BRT and RF models to evaluate the effect of each predictor on species' distribution by plotting the effect of the predictor on the response (basking shark presence) after accounting for the average effects of all other model predictors

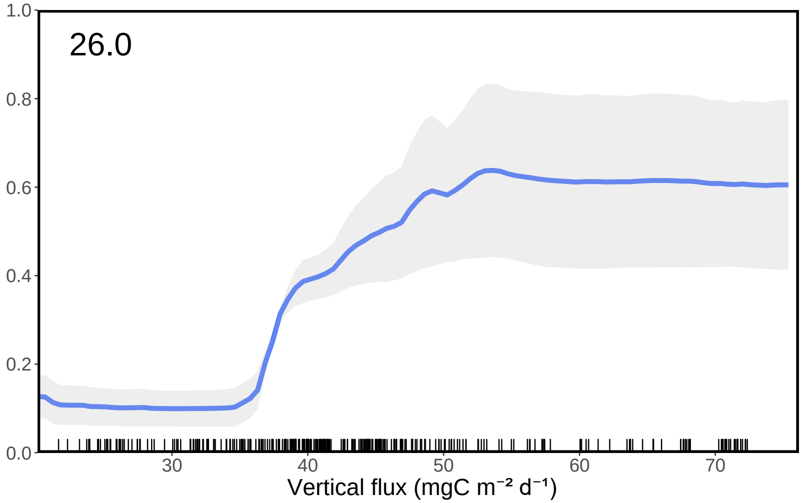
Results: Model performance

- AUC and TSS scores similar across models, RF slightly better performance
- Both indices considered useful for predicting shark occurrence (>0.7)
- Model performance had low variability (R-squared) – models consistently performed across bootstrap samples

	BRT model	RF model
Deviance explained (training data)	0.60 ± 0.03	0.75 ± 0.02
Deviance explained (evaluation data)	0.36 ± 0.10	0.52 ± 0.07
TSS (training data)	0.92 ± 0.02	0.88 ± 0.02
TSS (evaluation data)	0.69 ± 0.05	0.72 ± 0.04
AUC (training data)	0.95 ± 0.01	0.98 ± 0.00
AUC (evaluation data)	0.89 ± 0.03	0.92 ± 0.02

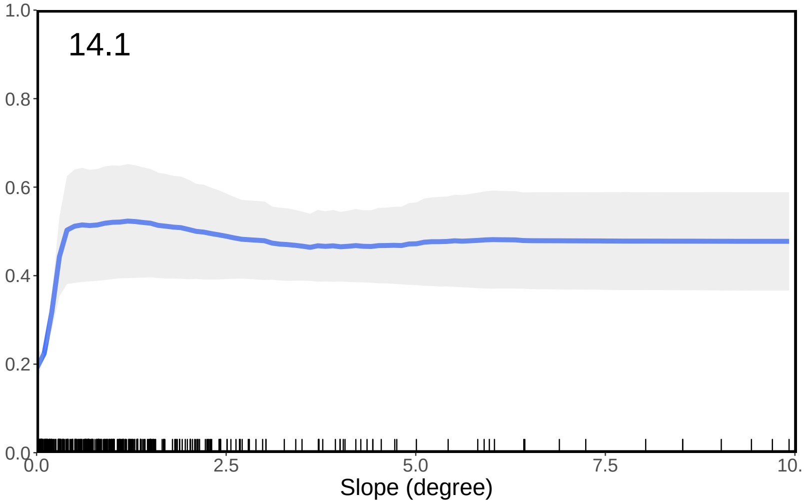
Results: Predictor selection and contribution

Vertical Flux



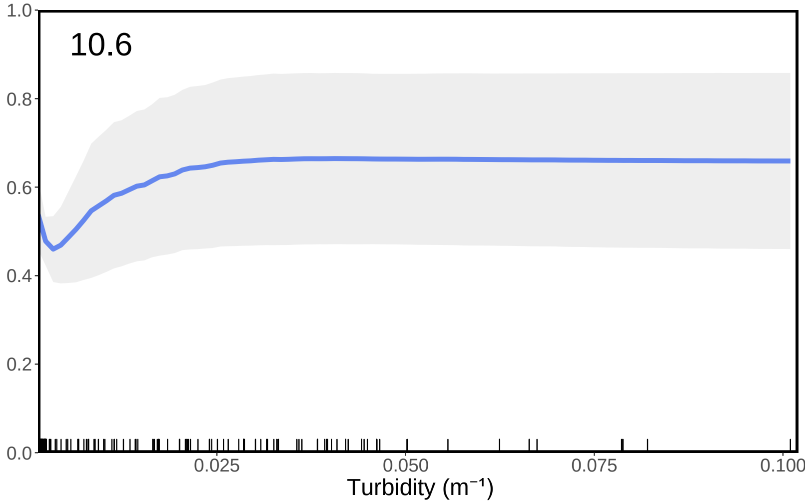
High HSI with high vertical flux >20 mgC m-2 d-1

Slope



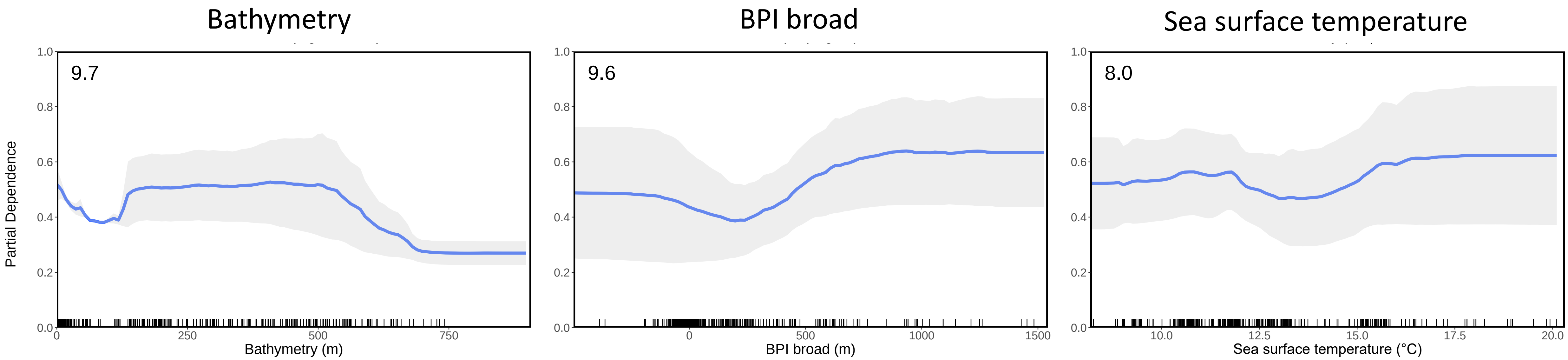
High HSI with gently sloping areas

Turbidity



High HSI with high turbidity

Results: Predictor selection and contribution



High HSI nearshore and again 200-500 m

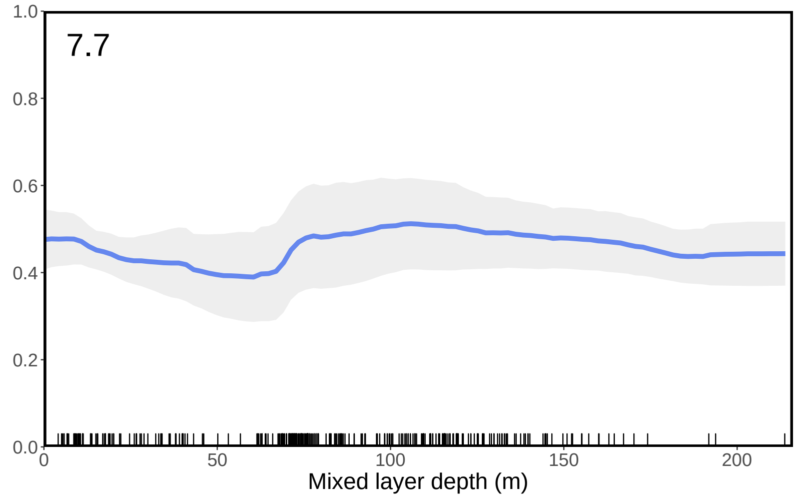
High HSI on less complex seafloor topologies

HSI highest < 12.5°C and >15°C

Climate, Freshwater & Ocean Science

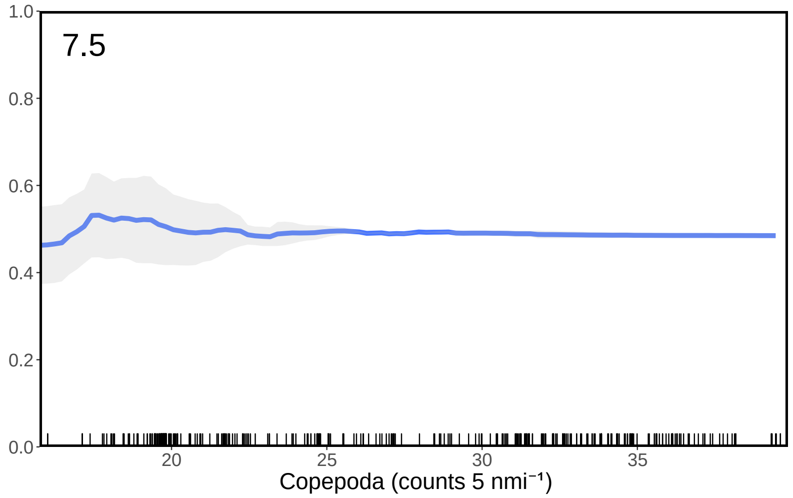
Results: Predictor selection and contribution

Mixed layer depth



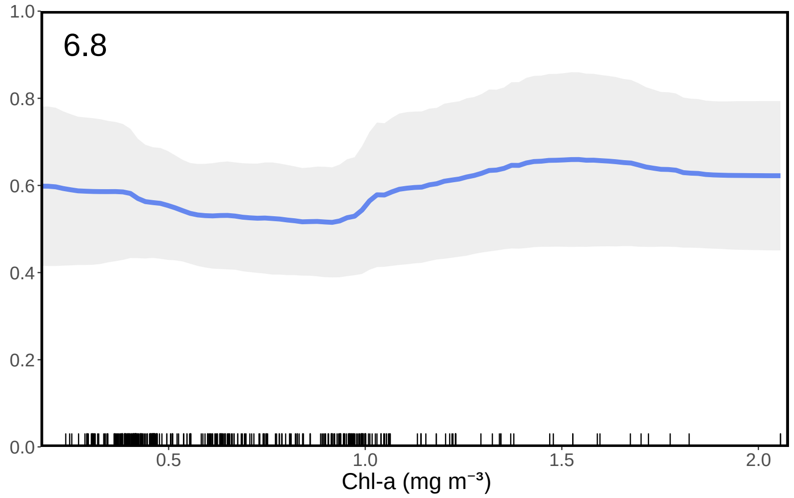
HSI highest ~75 m

Copepods



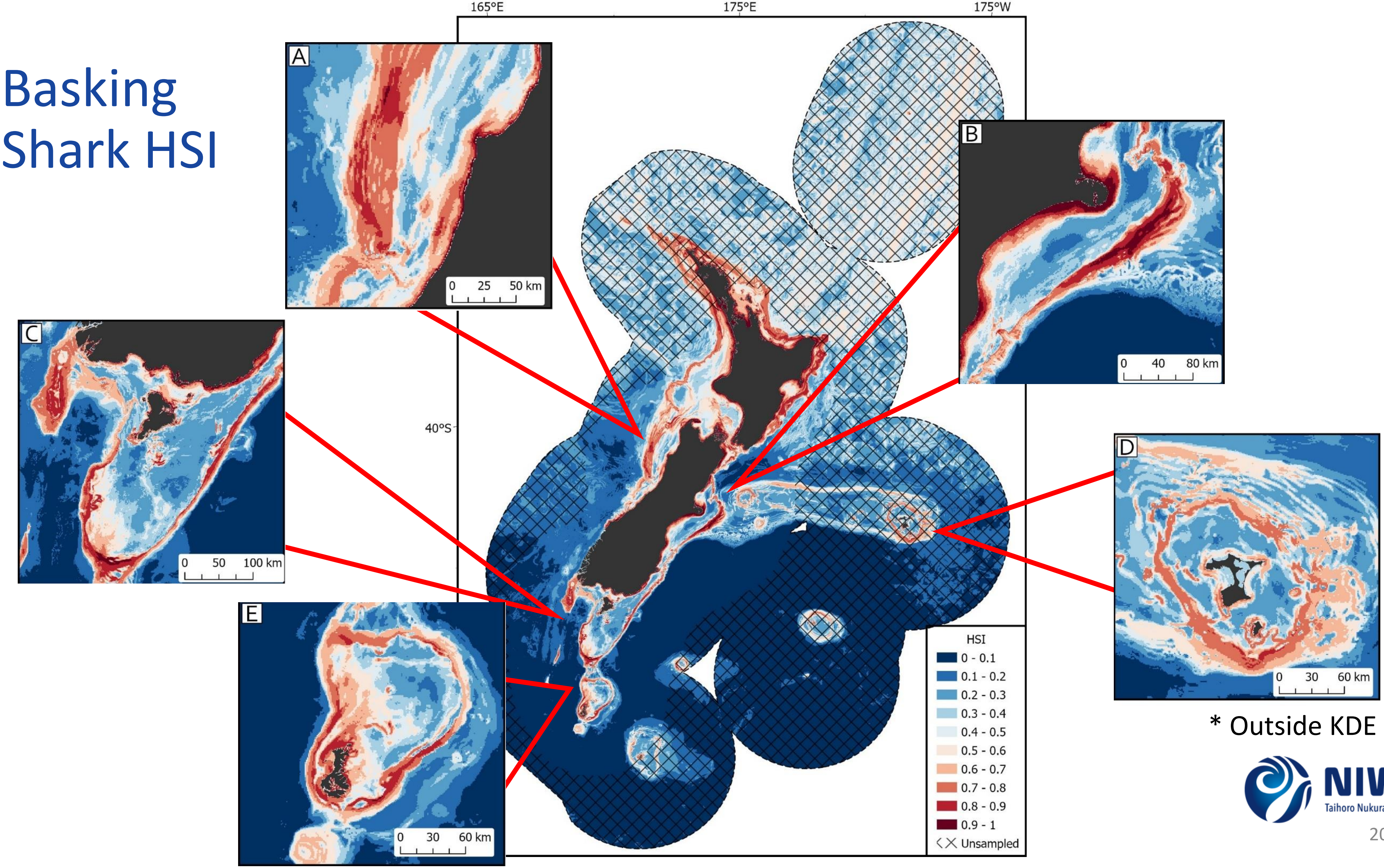
Highest HSI at moderate levels of copepod density (10-20 counts per 5 nautical miles)

Chl-a



HSI highest with high chl-a concentration (>1.2 mg m-3)

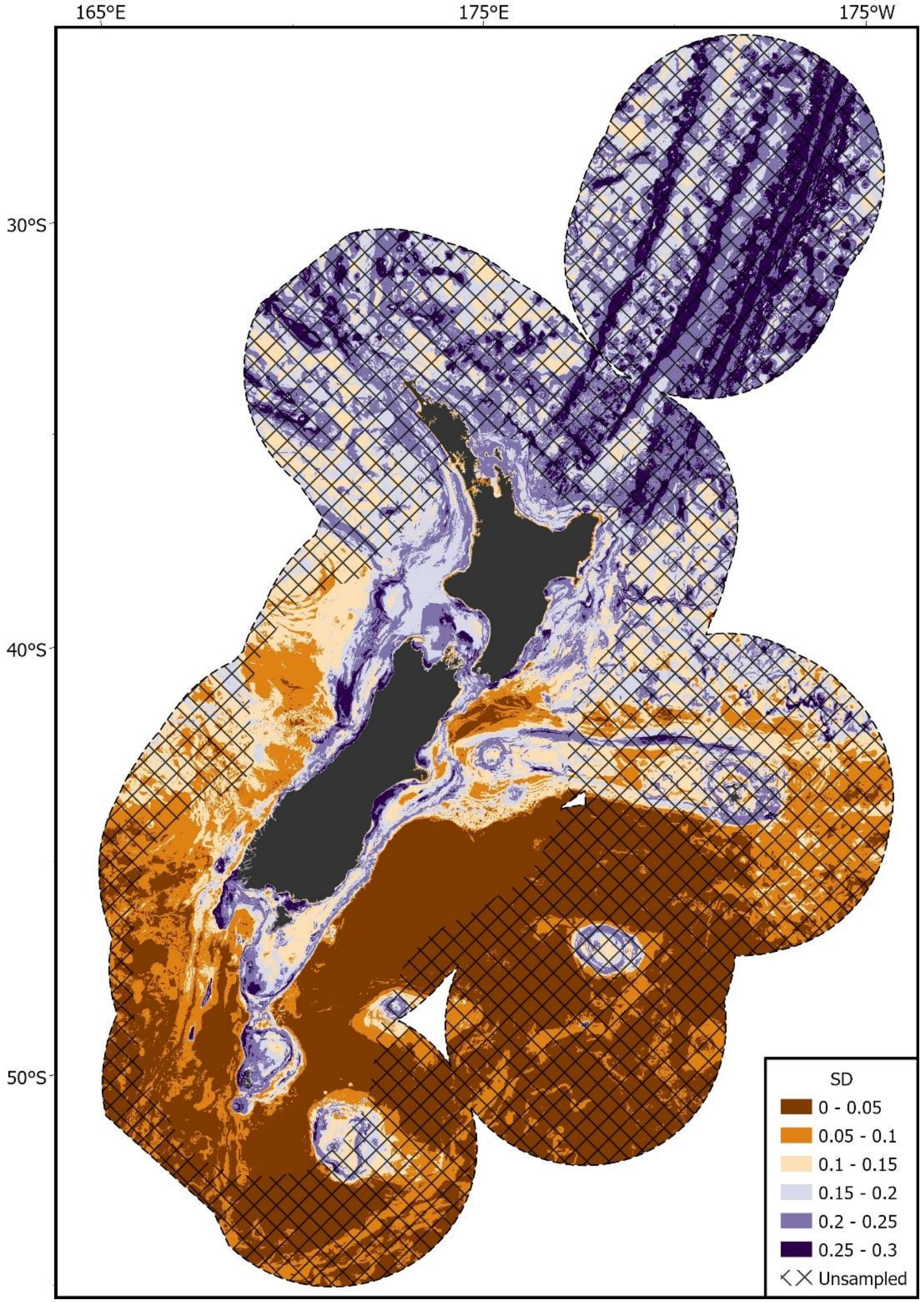
Basking Shark HSI



* Outside KDE

NIWA Taihoro Nukurangi

Basking Shark HSI uncertainty



NIWA Taihoro Nukurangi

Basking Shark Habitat Suitability in NZ

- First insight into habitat suitability for basking sharks in South Pacific, Southern Hemisphere

- BRT and RF models had good predictive power (AUC and TSS > 0.7)

- HSI largely influenced by variables representing ocean processes: areas with high levels of vertical flux of particulate organic matter at the seabed (net primary production in the surface mixed layer) were most influential; bathymetry and slope

- High HSI: ECSI, WCSI, Puysegur, southern Campbell Plateau, offshore islands

Basking Shark Habitat Suitability in NZ

Limitations

- Small sample size
- Lack of true absences
- Most current observations fisheries-dependent
- No estimates of abundance

- Long temporal span (121 years) - model predictions may be more representative of past, not current, suitable habitat

- No prey information north of 40°S

- Predicted distribution outside 95% KDE area should be treated with caution

Basking Shark Habitat Suitability in NZ

- Dynamic environmental variables may indicate seasonal/behavioural patterns of distribution: inshore and offshore regions highlighted as areas of high HSI

- Biotic predictor inclusion is important in understanding species' relationship with the marine environment in unobserved space, potential link in understanding effects in climate change

- Assist in assessing risk to fishing activities, incorporation into management frameworks (spatially explicit risk assessment)

Acknowledgements

This project was carried out under the Department of Conservation project POP2020-03

Thanks to

Jade Maggs (NIWA) and RDM (MPI) for the data extractions

Clinton Duffy (DOC) and Malcolm Francis (NIWA) for data collation

Clinton Duffy, Ben Sharp, and Marco Milardi (FNZ) for their feedback on the methods

Tom Brough (NIWA) for reviewing the draft report

Fisheries observers for collecting data