



Fisheries New Zealand

Tini a Tangaroa

Spatial distribution modelling of New Zealand cetacean species

New Zealand Aquatic Environment and Biodiversity Report No. 240

F. Stephenson,
K. Goetz,
T. Mouton,
F. Beets,
S. Hailes,
J. Roberts,
M. Pinkerton,
A. MacDiarmid

ISSN 1179-6480 (online)
ISBN 978-1-99-002502-0 (online)

May 2020



Requests for further copies should be directed to:

Publications Logistics Officer
Ministry for Primary Industries
PO Box 2526
WELLINGTON 6140

Email: brand@mpi.govt.nz

Telephone: 0800 00 83 33

Facsimile: 04-894 0300

This publication is also available on the Ministry for Primary Industries websites at:

<http://www.mpi.govt.nz/news-and-resources/publications>

<http://fs.fish.govt.nz> go to Document library/Research reports

© Crown Copyright – Fisheries New Zealand

TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
1. INTRODUCTION	2
1.1 Background	2
1.2 Aims and objectives	4
2. METHODS	4
2.1 Sampling data	4
2.2 Environmental predictors	7
2.3 Relative environmental suitability models	11
2.4 Boosted Regression Tree Models	12
2.5 Summary of analyses	13
2.6 Assessing model uncertainty	14
3. RESULTS	18
3.1 RES model outputs	18
3.2 BRT model outputs	20
4 DISCUSSION	82
4.1 Critical appraisal of model results	82
5 MANAGEMENT IMPLICATIONS	85
6 ACKNOWLEDGEMENTS	85
7. REFERENCES	86
APPENDIX 1: ENVIRONMENTAL PREDICTOR VARIABLES	89
APPENDIX 2: RES MODEL OUTPUTS	103
APPENDIX 3: BRT MODEL OUTPUTS	122

EXECUTIVE SUMMARY

Stephenson, F.; Goetz, K.; Mouton, T.; Beets, F.; Hailes, S.; Roberts, J.; Pinkerton, M.; MacDiarmid A. (2020). Spatial distribution modelling of New Zealand cetacean species.

New Zealand Aquatic Environment and Biodiversity Report No. 240. 217 p.

All cetaceans in New Zealand's territorial seas and Exclusive Economic Zone (EEZ) are protected under national law by the New Zealand Marine Mammal Protection Act (1978), which mandates that all physical or habitat disturbances to the animals must be avoided or mitigated. However, little is known about the distribution and habitat use patterns of cetaceans in the seas surrounding New Zealand, especially for those species that inhabit offshore waters. This general paucity of spatial information is a major limitation with respect to the management of potential threats to cetaceans, e.g., fishing, mineral extraction industries, and other threats that are heterogeneous in space.

Opportunistic sighting records ($n = 14\,207$) of 31 cetacean species complexes, species and subspecies, and 14 spatially comprehensive environmental data layers were combined to derive habitat preference models and, for a select number of species with sufficient data, to derive relative density using a two-step hurdle model. Depending on the number of records available for each species, different analyses were undertaken to estimate species distribution, i.e., Relative Environmental Suitability (RES) models for species with few sightings and Boosted Regression Tree (BRT) models for species with adequate sightings. Presence/absence models require negative events (i.e., the spatial locations where a species was not observed). BRT models developed for this assessment assumed that positive sightings of a species could be considered as absences for all other species.

Predicted distributions for rarely sighted species using RES were consistent with the spatial distribution of sightings and strandings available for those species. For more frequently sighted cetacean species, BRT model predictions were consistent with known distributions and the respective species sighting and stranding records. Species distribution models for certain key species, e.g., Hector's/Māui dolphins (*Cephalorhynchus hectori*) were corroborated by in-depth modelling by other assessments. For some species, e.g., common dolphin (*Delphinus delphis*), the spatial prediction obtained for this assessment was quite different from the more subjective spatial density analyses used for the most recent spatial risk assessment of the effects of fishing for cetacean species. The most consistently important environmental predictors for species presence were relative temperature at depth, bathymetry, distance to the 500 m isopleth, mixed layer depth, and water turbidity. Other, less consistently important environmental predictor variables included: slope, distance to shore, sea surface temperature, benthic sediment disturbance, tidal current speed, chlorophyll-a concentration, diffuse downwelling attenuation, and primary productivity estimates.

This assessment provides an objective cross-species assessment of the spatial distribution and relative density of New Zealand cetaceans that can be used for spatial risk assessment and spatial planning. Further improvements to the spatial predictions of New Zealand cetaceans could be achieved by accounting for sightings' effort (i.e., information on true species presence and absence). More in-depth assessments can also consider model development processes tailored to cetacean species groupings with similar spatial information and biology.

1. INTRODUCTION

1.1 Background

There are 47 cetacean species, subspecies, and/or ecotypes known to occur in New Zealand waters – three of which are listed as Nationally Critical (Bryde's whale (*Balaenoptera brydei*), Māui dolphin (*Cephalorhynchus hectori māui*), killer whale (*Orcinus orca*)); and another three as Nationally Endangered (Hector's dolphin (*Cephalorhynchus hectori hectori*), southern right whale (*Eubalaena australis*), bottlenose dolphin (*Tursiops truncatus*)) (Baker et al. 2010). All cetaceans in New Zealand's Territorial Sea and Exclusive Economic Zone (EEZ) are protected under national law by the New Zealand Marine Mammal Protection Act (1978), which mandates that all physical or habitat disturbances to the animals must be avoided or mitigated. However, little is known about the distribution and habitat use patterns of cetaceans in the seas surrounding New Zealand, especially for those species that inhabit offshore waters. This general paucity of spatial information is a major limitation with respect to the management of potential threats to cetaceans, e.g., fishing, mineral extraction industries, and other threats that are heterogeneous in space.

The Ministry for Primary Industries (MPI) has led the development of the spatially explicit risk assessment (SEFRA) framework for assessing the impact of direct fisheries mortality on protected species (Sharp 2017). Spatial risk models can then be used to estimate the impact and population-level consequences of past and current fishing effort, and also to evaluate the consequences of alternate spatial management strategies on threat-specific risk to marine mammals and other protected species. However the utility of these assessments relies critically on the accuracy of the species' distribution layers used to estimate spatial overlap with fishery threats.

Outputs of a multi-species assessment of the commercial fisheries risk to marine mammals are shown in Figure 1-1 (from Abraham et al. 2017). In this figure a risk ratio equal to or greater than one equates to a fisheries impact sufficiently high to prevent the population from recovering to an average level at or above 75% of the environmental carrying capacity in the long term. Broadly, large and small dolphin species had the highest species risk ratio (> 0.1 , Figure 1-1). However, because robust spatial densities were not available for most cetacean species, Abraham et al. (2017) used the Delphi method of using expert judgement to derive spatial distributions for most marine mammal species. A quantitative assessment of the distribution of New Zealand marine mammals was highlighted as an important step to improve the estimation of fisheries-related fatalities (Lonergan et al. 2017). The project reported on here produced cetacean species distribution maps, with the intent that these maps will improve spatially explicit and quantitative marine mammal risk assessments. In addition, these distribution maps may also be of use for other spatial planning applications (e.g., as inputs to systematic conservation planning tools).

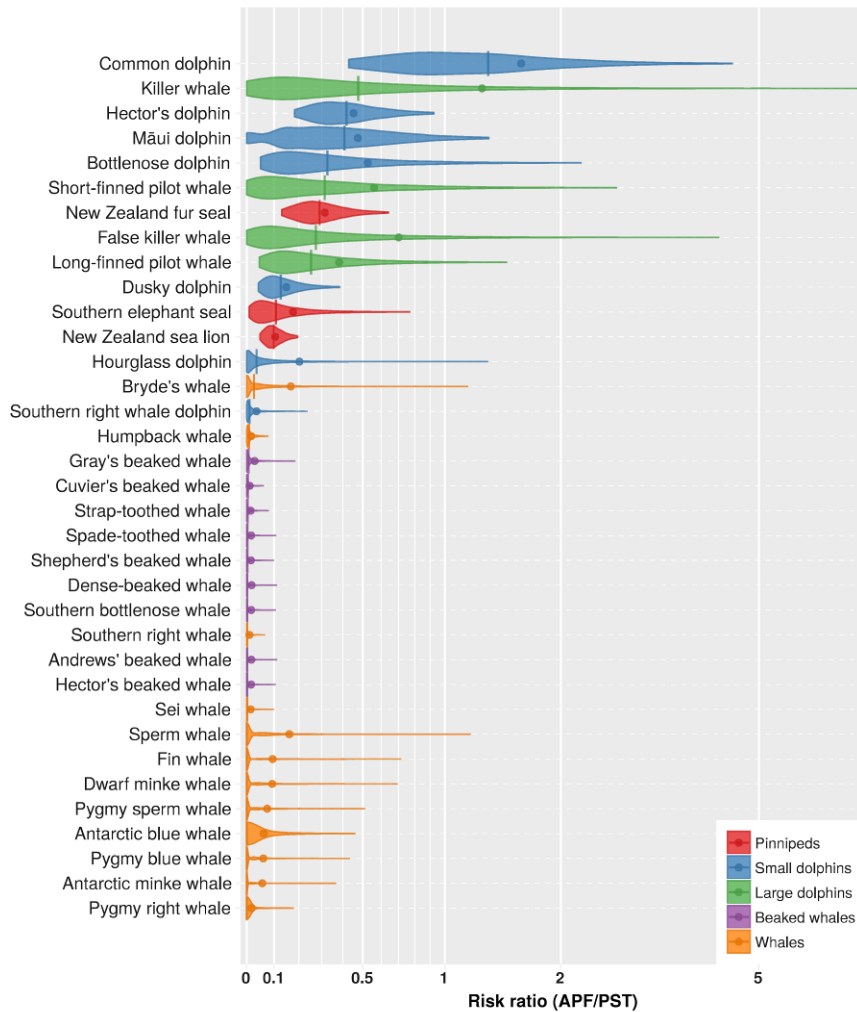


Figure 1-1: Risk ratio for New Zealand marine mammals, calculated as the ratio of the annual potential fatalities (APF) to the Population Sustainability Threshold (PST). Values are displayed on a logarithmic scale, and the distribution of the risk ratios within their 95% credible interval is indicated by the coloured shapes, including the median risk ratio (vertical line). Species are listed in decreasing order of the median risk ratio (from Abraham et al. 2017).

Because marine mammal survey data across the entire EEZ do not exist, a database compiling all robust cetacean sightings records (including from both fisheries and non-fisheries sources) and spatially comprehensive, high resolution, and functionally relevant environmental data layers were combined to derive empirically based habitat preference models for selected cetacean species. For select species with sufficient sightings records, these were then used to estimate relative spatial density around New Zealand. Species distribution models have become a reliable and recognised method of predicting species probability of presence across vast areas and are an integral part of resource management and conservation biology (Guisan and Thuiller 2005, Elith et al. 2006). Spatial information, such as from incidental cetacean sightings, can be used to model a species' ecological niche based on the assumption that the distribution of known encounters reflects the species' environmental preferences (Guisan and Zimmermann 2000, Hirzel et al. 2006). To reduce uncertainty in future risk assessments, the analysis was spatially and temporally structured (reflecting seasonal trends in sightings and habitat when possible). For species with sparse or unrepresentative sightings data, species distributions were inferred using a different method for species habitat preference based on a review of the latest knowledge.

All final spatial layers produced for this report will be made available on the MPI hosted and managed NABIS website which will allow closer inspection of areas of interest by all interested stakeholders and managers.

1.2 Aims and objectives

The overall aim of this research was to produce a digital atlas mapping the distribution of selected cetacean species via synthesis of existing information and reflecting any seasonal temporal changes. Specific objectives were:

1. To produce an agreed list of cetacean species for inclusion and compile all available spatial data for these species; and
2. To model and map species' distributions and relative density of species identified in objective 1 from available spatial data, reflecting any seasonal variability in sightings and/or species habitat.

2. METHODS

2.1 Sampling data

Multiple databases of cetacean sightings records from five organisations/institutes were compiled by the Department of Conservation (DOC) (Table 2-1). The data were groomed to remove any errors. This included: removing all records with locations on land; fixing latitudes that should have been south but were entered as north; removing records located outside the New Zealand EEZ; removing records of pinnipeds; removing duplicated records from within and between databases; checking and standardising all spelling of species names; removing all species records labelled as unknown; removing all records with estimated locations or dates; and removing records where no group sizes were recorded.

Table 2-1: Summary of sightings data contained in the various databases. The oldest and most recent records for each database are shown in columns 'Start' 'End' respectively. The number of sightings in each database is shown in the final column.

Database	Source	Start	End	Number
Cawthorn	Martin Cawthorn (independent consultant)	January 1980	November 1999	1 150
COD	Central Observer Database (COD): NIWA maintained database on behalf of Fisheries New Zealand	January 2009	April 2017	4 768
DOC	Department of Conservation (DOC)	January 1970	July 2017	7 867
NIWA	National Institute of Water and Atmospheric Research (NIWA)	April 2011	July 2016	6
OMV	OMV limited	March 2005	April 2015	416

Following grooming to remove errors in location and duplicate sightings among others, a total of 14 207 records of 31 cetacean species were retained for further analyses (Table 2-2, Figure 2-1). Each record included a group size estimate, date, and location (latitude, longitude). The number of records differed between species and seasons (here, winter is defined as from 1 May to 31 October and summer from 1 November to 30 April) (Table 2-2).

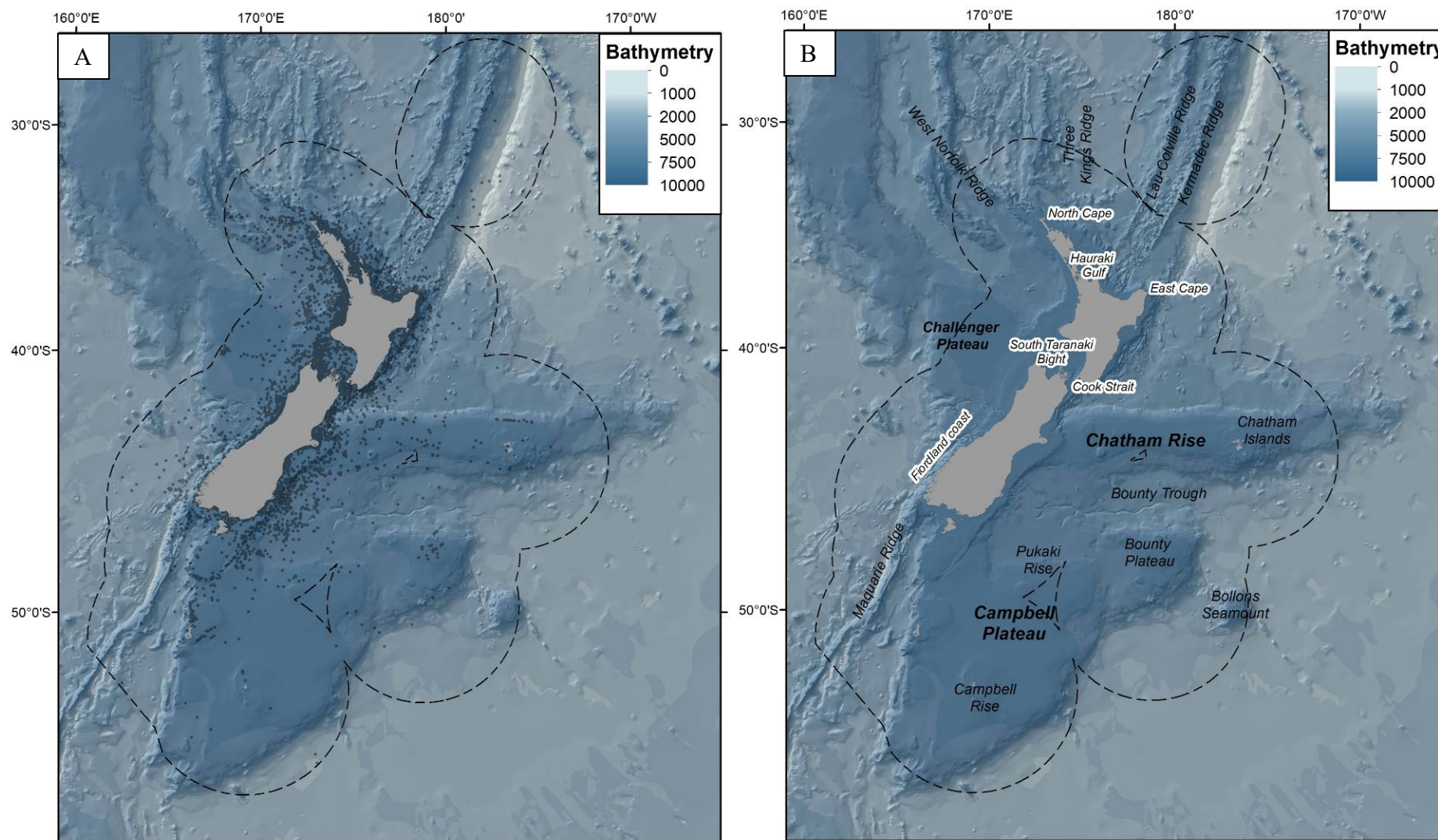


Figure 2-1: Map of the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line), bathymetry (m) and in A) species sightings (grey dots) and B) feature names used in the text.

Depending on the number of records available for each species, different analyses were undertaken to estimate distributions and abundances. For those species with fewer than 50 recorded sightings, Relative Environmental Suitability (RES) models (Kaschner et al. 2006) were used to predict probability of presence (see section 2.3 for further details). For species with more than 50 recorded sightings, Boosted Regression Tree (BRT) models (Elith et al. 2006) were fitted with a Bernoulli error distribution to predict the average year-round probability of presence (see section 2.4 for further details). To predict the average year-round probability of presence, BRTs require locations of both presences (sightings records) and absences. Here, true absences (i.e., location records for where cetacean species were not sighted) were not available. True absences require a record of sighting effort which is not available for opportunistic sightings data compiled from multiple sources. BRTs are sometimes instead fitted using randomly generated ‘background’ zeroes (sometimes referred to as pseudo-absences), but this approach implicitly assumes that sighting effort is uniform in space, which is rarely the case. Here NIWA utilised a multi-species sightings database in which a positive sighting record for any cetacean species was counted as a negative record (i.e., a ‘zero’ sighting) for all other species, described here as ‘relative absence’. Individual species models were then fitted, based on these presence/ relative absence records. Although not as robust as true absences, relative absences provided the advantage of a zero only being recorded if a human observer was present in that time and place, and sufficiently motivated to report a cetacean sighting. With this approach there is an implicit assumption that if that observer had sighted two species simultaneously they would have reported both species, so the recorded presence of one species can be regarded as a verified absence of all others.

In addition to predicting probability of presence, for those species with more than 300 records, BRT models were also fitted with a Poisson error distribution to predict average year-round species relative densities (see section 2.4 for further details). Further, for those species with more than 75 records in each season, BRT models fitted with a Bernoulli error distribution were used to predict seasonal probability of presence (see section 2.4 for further details).

Table 2-2: Number of cetacean sighting records per species included in the analyses. White indicates species for which RES models were run; grey shading indicates species for which BRT average year-round probability of presence models were fitted; blue shading indicates species for which average year-round BRT probability of presence and relative density distribution models were fitted. Orange shading indicates species for which average year-round and seasonal probability of presence and average year-round relative density distribution models were fitted.

Species names	Species	Summer	Winter	Total
Blainville's beaked whale	<i>Mesoplodon densirostris</i>	1	0	1
Dwarf minke whale	<i>Balaenoptera acutorostrata</i>	0	1	1
Spectacled porpoise	<i>Phocoena dioptrica</i>	1	0	1
Striped dolphin	<i>Stenella coeruleoalba</i>	0	1	1
Andrew's beaked whale	<i>Mesoplodon bowdoini</i>	2	0	2
Hourglass dolphin	<i>Lagenorhynchus cruciger</i>	2	0	2
Pygmy sperm whale	<i>Kogia breviceps</i>	2	0	2
Southern bottlenose whale	<i>Hyperoodon planifrons</i>	4	0	4
Risso's dolphin	<i>Grampus griseus</i>	3	2	5
Shepherd's beaked whale	<i>Tasmacetus shepherdi</i>	1	4	5
Cuvier's beaked whale	<i>Ziphius cavirostris</i>	2	5	7
Gray's beaked whale	<i>Mesoplodon grayi</i>	8	1	9
Southern right whale dolphin	<i>Lissodelphis peronii</i>	26	1	27
False killer whale	<i>Pseudorca crassidens</i>	26	2	28
Arnoux's beaked whale	<i>Berardius arnuxii</i>	30	1	31
Minke whale	<i>Balaenoptera acutorostrata</i>	38	19	57
Fin whale	<i>Balaenoptera physalus</i>	46	15	61
Sei whale	<i>Balaenoptera borealis</i>	51	19	70
Blue whale	<i>Balaenoptera musculus</i>	304	50	354
Right whale	<i>Eubalaena australis</i>	75	402	477
Sperm whale	<i>Physeter macrocephalus</i>	337	160	497
Bottlenose dolphin	<i>Tursiops truncatus</i>	327	171	498
Killer whale	<i>Orcinus orca</i>	368	201	569
Bryde's whale	<i>Balaenoptera brydei</i>	309	284	593
Humpback whale	<i>Megaptera novaeangliae</i>	239	390	629
Pilot whale (2 spp.)	<i>Globicephala melas</i> <i>Globicephala macrorhynchus</i>	509	170	679
Dusky dolphin	<i>Lagenorhynchus obscurus</i>	654	169	823
Māui dolphin	<i>Cephalorhynchus hectori māui</i>	875	176	1 051
Hector's dolphin	<i>Cephalorhynchus hectori hectori</i>	3 394	294	3 688
Common dolphin	<i>Delphinus delphis</i>	3 128	1 283	4 411

2.2 Environmental predictors

New Zealand's EEZ encompasses a diverse range of environmental conditions (Bradford-Grieve et al. 2006). To capture this variability, 14 high resolution gridded environmental predictors, at a native resolution of 1 km, were collated and imported into ArcGIS (version 10.6) (described in Table 2-3, spatial layers are shown in Appendix 1). These variables were selected based on prior information with respect to their likely influence on cetacean presence and distribution (Table 2-3). Although it is

unlikely that the majority of these variables directly affect cetacean presence and distribution, physical processes and oceanographic features such as bathymetry, thermal layers, shelf breaks, and productivity are known to aggregate prey, attracting cetaceans for foraging (Tynan et al. 2005, Etnoyer et al. 2006, Bluhm et al. 2007). Although most of the chosen ocean climate variables were static (e.g., Bathymetry), several variables described mean monthly statistics (e.g., Mixed layer depth, MLD, column ‘Timescale’ in Table 2-3).

Several environmental variables exhibited co-linearity, e.g., between Sea Surface Temperature and Sea Surface Temperature Gradient (not shown). Although BRT modelling is reasonably robust to correlated variables (Elith and Leathwick 2009, Guisan et al. 2013), the use of highly correlated variables generally provides only minimal improvement in prediction accuracy and complicates interpretation of model outcomes (Leathwick et al. 2006). Consequently, several variables were excluded: ocean temperature at depth (100 m, 200 m, and 500 m), sea surface temperature gradient, sea surface height, sea surface height anomaly, and a benthic species classification layer (BOMECC). The remaining 14 variables (see Table 2-3) were retained for model selection, ignoring moderate correlations (≤ 0.75), because all were deemed to be potentially influential for cetacean species distribution).

Prior to BRT model fitting, values for each environmental variable were extracted for cetacean sighting location by overlaying these records onto each of the environmental variable layers using the ‘raster’ package in R (Hijmans & van Etten 2012). For monthly environmental variables, recorded dates of cetacean sightings were used to extract respective values at the time the observation was made.

Table 2-3: Environmental variables used as predictors in Boosted Regression Tree analyses. Variable abbreviations, names (whether annual or monthly means), units, and descriptions are provided.

Variable abbreviation	Variable name	Timescale	Unit	Description	Source
Slope	Slope	Annual	Degree	Bathymetric slope was calculated from bathymetric depth and is the degree of change from one depth value to the next.	NIWA unpublished
MLD	Mixed layer depth	Monthly	m	The depth that separates the homogenised mixed water above from the denser stratified water below.	Calculated from the CARS climatology, NIWA unpublished
Bathy	Bathymetry	Annual	m	Depth at the seafloor was interpolated from contours generated from various sources, including multi-beam and single-beam echo sounders, satellite gravimetric inversion, and others (Mitchell et al. 2012).	NIWA
Dist.Shore	Distance to shore	Annual	km	Using a NIWA sourced polygon of the New Zealand coastline, distance from shore was calculated using the spatial analysis extension tool in ArcGIS.	NIWA unpublished
SST	Sea surface temperature	Monthly	°C	MODIS-Aqua SST product, calculated as longterm (2002–2017) average values at 1000 m spatial resolution.	NIWA unpublished, based on processing described in Pinkerton et al. (2018)
Dist.Iso500	Distance to 500 m isobath	Annual	m	The 500 m bathymetric contour was used to denote the shelf break. Distance from this isobath was calculated using the spatial analysis extension in ArcGIS.	NIWA unpublished
BedDist	Benthic sediment disturbance	Annual	unitless	Combination of seabed orbital velocities (estimates the average mixing at the seafloor as a consequence of orbital wave action, calculated from a wave climatology derived hindcast (1979 to 1998) of swell-wave conditions in the New Zealand region (Gorman et al. 2003)) and friction velocity for seabed types (based on grain size). Benthic sediment disturbance from wave action was assumed to be zero where depth ≥ 200 m.	Leathwick et al. (2012)
TempRes	Temperature residuals		°C	Residuals from a Generalised Linear Model relating temperature to depth using natural splines – this highlights areas where average temperature is higher or lower than would be expected for any given depth.	Leathwick et al. (2006)

Variable abbreviation	Variable name	Timescale	Unit	Description	Source
TC	Tidal current speed	Annual	ms ⁻¹	Maximum depth-averaged (New Zealand region bathymetry) flows from tidal currents calculated from a tidal model for New Zealand waters (Walters et al. 2001).	Leathwick et al. (2012)
Turb	Turbidity	Monthly	NTU	Optical backscatter as measured by turbidity sensor. Estimated using quasi-analytic inversion algorithm applied to MODIS-Aqua data. Result calculated based on longterm (2002–2017) average values of particulate backscatter bbp(555) at 500 m resolution, converted to normalised turbidity units (NTU) using in situ turbidity measurements in the New Zealand coastal zone.	NIWA unpublished, based on processing described by Pinkerton et al. (2018)
ChlA	Chlorophyll-a concentration	Monthly	mg m ⁻³	A proxy for the amount of photosynthetic plankton, or phytoplankton, present in the ocean. Estimated using quasi-analytic inversion algorithm applied to MODIS-Aqua data. Result calculated based on longterm (2002–2017) average values of phytoplankton absorption aph(555) at 500 m spatial resolution.	NIWA unpublished, based on processing described by Pinkerton et al. (2018)
Kpar	Diffuse downwelling attenuation	Monthly	m ⁻¹	Attenuation of broadband irradiance (Photosynthetically Available Radiation, PAR) with depth. Estimated using quasi-analytic inversion algorithm applied to MODIS-Aqua data. Result calculated based on longterm (2002–2017) average values at 500 m spatial resolution.	NIWA unpublished, based on processing described by Pinkerton et al. (2018)
VGPM	Productivity Model	Monthly	mg C m ⁻² d ⁻¹	Provides estimates of surface water primary productivity based on the Vertically generalized productivity model of Behrenfeld & Falkowski (1997). Net primary productivity by phytoplankton (mean daily rate of water column carbon fixation) is estimated as a function of merged remotely sensed chlorophyll concentration, irradiance, and photosynthetic efficiency estimated from remotely sensed Sea-Viewing Wide-Field-of-view Sensor (SeaWiFS) and MODIS-Aqua satellite imagery (M. Pinkerton, NIWA, pers. comm.).	NIWA unpublished, Oregon State University (www.science.oregonstate.edu/ocean.productivity/)
DOM	Coloured dissolved organic matter (CDOM)	Annual	Indicative of CDOM absorption at 440 nm ag(440) (m ⁻¹)	Detrital absorption at 440 nm, including due to coloured dissolved organic matter (CDOM) and particulate detrital absorption. Estimated using quasi-analytic inversion algorithm applied to MODIS-Aqua data. Result calculated based on longterm (2002–2017) average values of detrital absorption coefficient ag(443) at 500 m spatial resolution.	NIWA unpublished, based on processing described by Pinkerton et al. (2018)

2.3 Relative environmental suitability models

Relative environmental suitability (RES) models predict the average year-round geographical ranges of species using basic descriptive data that are available for most species, including those for which few (or no) recorded locations are available (Kaschner et al. 2006). The average year-round geographical ranges of species was defined as “the maximum area between the known outer-most limits of a species’ regular or periodic presence” (Kaschner et al. 2006). This definition is inclusive of areas covered during annual migrations, dispersal of juveniles etc., but is exclusive of extralimital sightings.

Three environmental variables were selected to describe species geographic ranges in RES models following methods described by Kaschner et al. (2006): sea surface temperature, water depth, and distance to shore. The RES method does not require cetacean location records, but rather relies on the generic relationships between species and each of the environmental layers developed from the scientific literature and expert knowledge. The relationships between species and the three environmental variables are described using a trapezoidal response curve based on four parameters: Min_A , Min_p , Max_p , and Max_A (Figure 2-2). Min_A and Max_A refer to absolute minimum and maximum variable ranges (i.e., beyond these values, the species are not expected to occur), and Min_p and Max_p describe the ‘preferred’ range, in terms of habitat usage of a given species (Kaschner et al. 2006).

Min_A , Min_p , Max_p , and Max_A were defined for each species using values from Kaschner et al. (2006) and supplemented by a literature search and expert advice (all the values used and a qualitative description of the ranges for each species are provided in Appendix 2). For each species, RES scores for the three environmental variables (SST, Bathy, Dist.Shore) were produced by transforming the gridded data layers of the variables following the trapezoidal response curve. For each species, final RES scores were produced by multiplying the three suitability layers assigned to the individual attributes (e.g., SST, Bathy, Dist.Shore) resulting in spatial estimates of environmental suitability ranging from 0 (not present) to 1 (highly representative of the species preferred or overall habitat range).

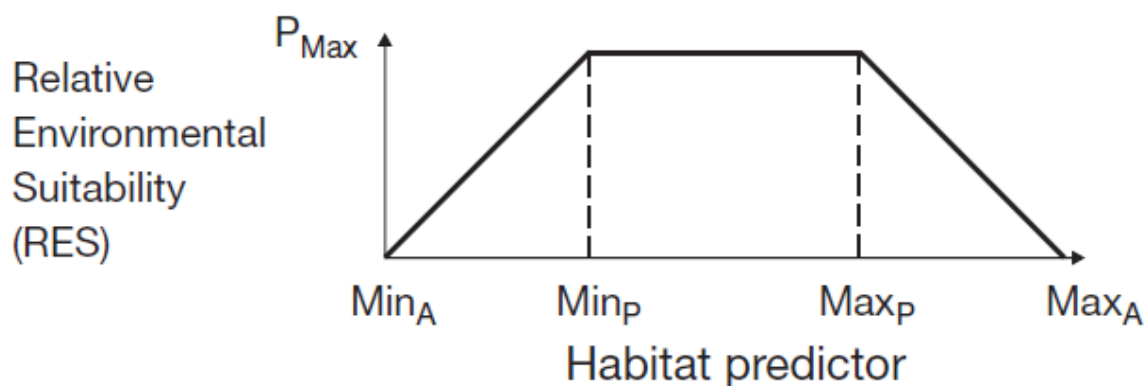


Figure 2-2: Trapezoidal species response curve describing the niche categories used in the RES model (from Kaschner et al. 2006). Min_A and Max_A refer to absolute minimum and maximum predictor ranges, and Min_p and Max_p describe the ‘preferred’ range, in terms of habitat usage of a given species.

2.4 Boosted Regression Tree Models

BRTs (Leathwick et al. 2006) were used to model average year-round and seasonal species probability of presence by relating species presence and relative absence data to environmental predictor variables (see sections 2.4.1 and 2.4.3 respectively). BRTs were also used to model average year-round species distribution of relative densities by linking species group size data to environmental predictor variables (see section 2.4.2).

BRT modelling combines many individual regression trees (models that relate a response to their predictors by recursive binary splits) and boosting (an adaptive method for combining many simple models to give improved predictive performance) to form a single ensemble model (Elith et al. 2008). Regression trees are produced iteratively, gradually improving the overall fit by giving more weight to those sites that are poorly fitted by the previous trees (Smith et al. 2013). Descriptions of the BRT method are available (see Ridgeway 2007, Elith et al. 2008). All statistical analyses were undertaken in R (R Core Team 2013) using the ‘Dismo’ package (Hijmans et al. 2017).

2.4.1 Predicted average year-round species probability of presence

Relationships between average year-round cetacean presence and relative absence (for 15 species, see Table 2-2) and 14 environmental predictor variables were investigated using BRTs fitted with a Bernoulli error distribution, a tree complexity of 5, a learning rate of between 0.01 and 0.0001 (with the parameter selected to fit between 1000 and 3000 trees for each model), a bag fraction of 0.6 and 10-fold cross evaluation (following recommendations from Elith et al. (2008) and Leathwick et al. (2006)).

For all assessed species, average year-round probability of presence models were originally fitted using all available predictor variables. The cross-validation process ensures that models are parsimonious, however, over-fitting can also occur by including more predictor variables than necessary (Leathwick et al. 2006). To reduce the risk of overfitting, the global models (those with all predictor variables included) for each species in turn were subjected to a simplification process whereby environmental variables were removed from the models, one at a time, using the ‘simplify’ function (Elith et al. 2006). This simplification process firstly assesses the relative contributions of each variable in terms of deviance explained, with the lowest contributing variables removed from the model, before the model is refitted with the remaining environmental variables. The change in deviance explained that resulted from removing the variable was then examined and the process repeated until all variables were sequentially removed. The final models were created by refitting the model with a reduced variable set that balanced the deviance explained with a reasonable number of predictor variables (Table 3-2 in section 3.2 provides information on the predictor variables used for each species).

BRT models were assessed using cross-validated measures of model performance (Elith et al. 2008, Compton et al. 2012). Model performance measures included the deviance explained and the area under the receiver operating characteristic curve (AUC). The explained deviance provides a measure of the goodness-of-fit between the predicted and raw values (total deviance) (Compton et al. 2012). The AUC measures the model’s ability to discriminate between presence and absence points; AUC ranges from 0 to 1, where a value of 1 indicates that the presences and absences are perfectly discriminated, and a value of 0.5 indicates that discrimination is no better than random chance (Elith et al. 2006). An AUC value of > 0.7 is considered a useful model (Derville et al. 2016). The variance explained by environmental predictor variables for each model was calculated to show their relative importance. The association between species probability of presence and the four most influential environmental predictor variables was illustrated using partial dependence plots (i.e., predicted response curve of species probability presence across the gradient of the variable of interest when all other variables are held at their means).

BRT models were bootstrapped 100 times for each species. That is, a random sample of the presence/relative absence data was drawn with replacement (i.e., 75% of the presence data and twice as many relative absences which were randomly selected from the database), and a model was constructed

with the same settings as the original. This process was repeated 100 times, and, at each iteration, predictions were made to the evaluation data (the remaining 25% of the presence data and twice as many randomly selected relative absences from the remaining absences) allowing model fits to be examined both on the training model and on the evaluation data.

Finally, each bootstrapped BRT model was predicted geographically using the annual mean environmental predictor variables to a 5 km grid. The mean estimated probability of presence and a spatially explicit measure of uncertainty (measured as the coefficient of variation (CV)) were calculated for each grid cell using the 100 bootstrap BRT layers. The CV of the bootstrap output – i.e., the standard deviation divided by the mean – was used because previous studies have found that its correlation with model predictions was substantially less than for unadjusted standard deviation (Anderson et al. 2016). In this report, all mapped outputs are provided at a 5 km grid resolution. However, all digital outputs were provided to Fisheries New Zealand at a 1 km grid resolution as part of a separate contract.

2.4.2 Predicted average year-round species group size and relative density

Spatial estimates of average year-round species relative density were investigated using a two-step BRT model (sometimes called a ‘hurdle model’) for 12 species (see Table 2-2) (Dedman et al. 2015). The first step was to model the probability of presence for each species (see section 2.4.1). The second step consisted of separately modelling species group size using the same parameters as those described in section 2.4.1, but using a Poisson error distribution. Finally, geographic predictions from both models were combined by multiplying both spatial layers to produce an estimate of average year-round species relative density distribution.

Group size BRT models were assessed using explained deviance and Pearson’s correlation between predicted group size (fitted using 75% of samples) and observed group size (evaluation data consisted of 25% of group size records) for each bootstrap.

2.4.3 Predicted seasonal species probability presence

Relationships between seasonal cetacean presence and relative absence (sightings records from winter and summer) and 14 environmental predictor variables were investigated using BRTs fitted with the same parameters and methods as those described in section 2.4.1. Geographic predictions were made using the same methods as those described in section 2.4.1, however, mean seasonal averages of the environmental predictor variables were used for variables where this information was available (see column 2 in Table 2-3).

2.5 Summary of analyses

A summary of the analyses undertaken, with key terminology used for each model (bold text) is provided in a flow diagram (Figure 2-3). Note the terminology used for BRT models fitted with group size data: results of model performance measures and importance of predictors are referred to in the text as “average year-round group size” model outputs. However, spatial predictions from these models are not shown in this report; rather the output from the two-step hurdle model is shown, referred to here as “average year-round relative density”.

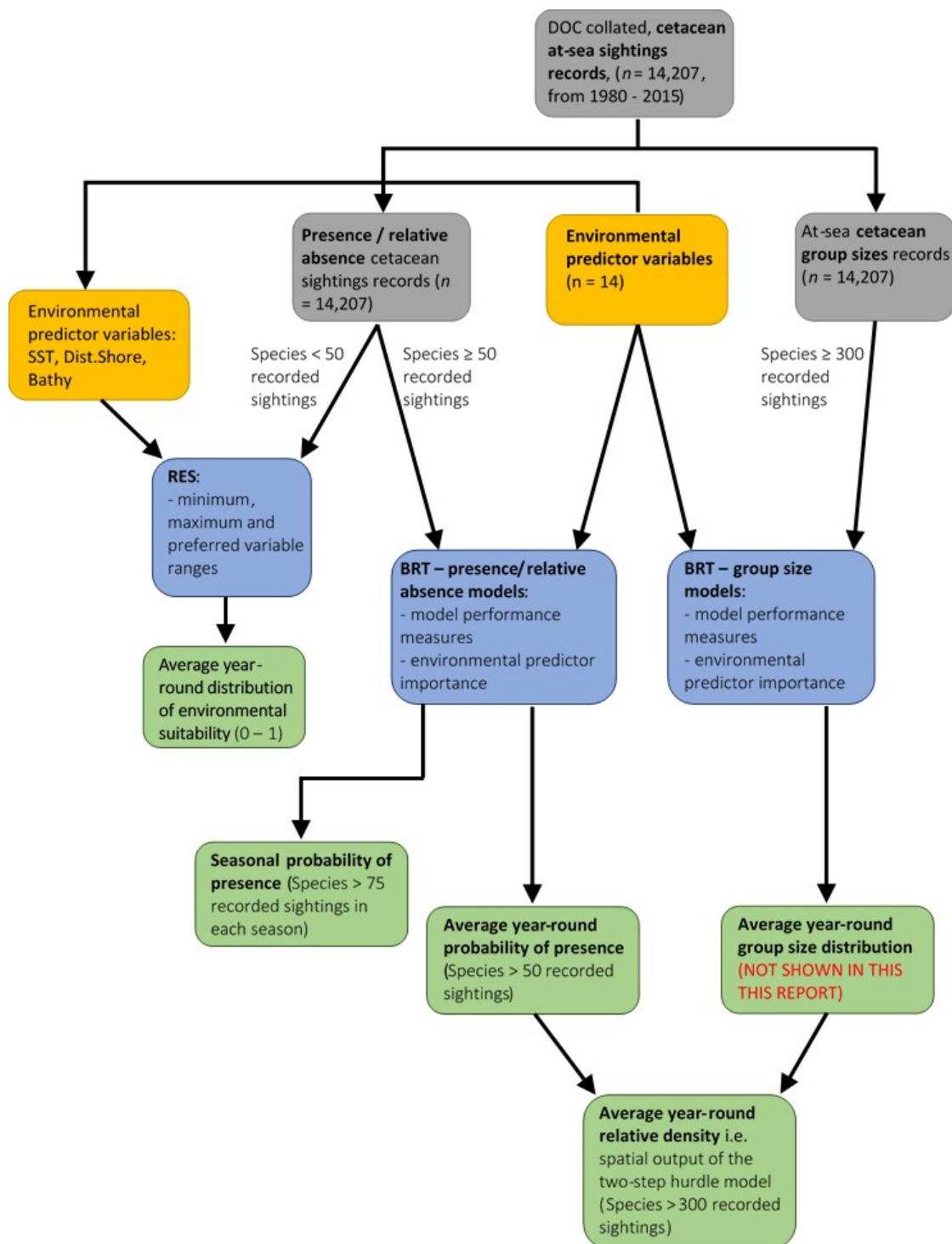


Figure 2-3: Summary of data inputs, analyses undertaken, key outputs, and terminology used (**bold**). Cetacean sightings records (grey boxes), environmental predictors (orange boxes), models used with overview of outputs (blue boxes), and spatial predictions (green boxes) are shown.

2.6 Assessing model uncertainty

A range of model evaluation statistics and methods were used to describe and evaluate model uncertainty.

2.6.1 Utilisation distribution

Utilisation Distribution (UD) is a simple method for determining cetacean probability distribution and was used as a visual comparison with the geographical predictions from BRT models. The UD of all cetacean sightings was described by a bivariate probability density function (Worton 1989) using a kernel method (Silverman 1986, Wand & Jones 1995). The bandwidth was selected using a smoothing parameter derived from Least Square Cross Evaluation (LSCV). The top 90% of the kernel distribution was selected to represent the UD (Calenge 2006). Because UD is determined using sightings point data only, this method also provides an indication of the geographic range of records (often constrained to areas that were close to shore due to sampling bias). All UD layers are shown overlaid on the geographic predictions from BRT models for visual comparison.

2.6.2 Coverage of the environmental space

As an added measure of model uncertainty ‘coverage of the environmental space by samples’ was estimated (Smith et al. 2013). The ‘environmental space’ is the multidimensional space when each variable is treated as a dimension. Cetacean location data can be projected into this space, where some parts of this environmental space will contain many samples (and are therefore well covered by the biological data) and other parts of this environmental space will contain few samples (and therefore the relationship between the environment and the biological samples are poorly understood resulting in potentially less certain predictions). For the latter, where predictions are considered less reliable, these should be treated with more scepticism than areas that are well covered by samples (Smith et al. 2013). Here we followed methods described by Smith et al. (2013) to model the coverage of the environmental space.

The degree to which the environmental conditions of each predictive site was covered by the samples was quantified by randomly sampling 50 000 values from the environmental space and assigning a ‘sample’ value of 0 to these, indicating that these were ‘absent’ sample sites. These were combined with the true samples ($n = 14\,602$), to which a ‘sample’ value of 1 was assigned (‘present’). A BRT was then used to model the relationship between ‘absent’ (random) samples and ‘present’ (true) samples for the 14 environmental predictor variables, using a Bernoulli error distribution. Predictions using this model yielded estimates of the probability of a site occurring in each part of the environmental space. A learning rate that yielded 2000 trees with an interaction depth of 2 was used (so that only pair-wise combinations of the environmental variables were considered). Predictions were then made spatially, generating values between 0 and 1 (where 0 indicated little understanding of the environmental space and 1 a perfect understanding), according to how well each cell was represented by the samples.

2.6.3 Spatial estimates of uncertainty

As described in sections 2.4.1, 2.4.2, and 2.4.3, spatially explicit measures of uncertainty were produced by bootstrapping the BRT models. For probability of presence models, uncertainty was measured as the CV. For predicted species relative density, uncertainty was measured as the 5th and 95th percentile values (i.e., 95% prediction interval) for each grid cell. However, an important consideration for these uncertainty layers is that BRT model predictions are not well extrapolated into unsampled areas (i.e., the predicted values shown will simply be those of the closest environmental space). In addition, the confidence estimates may remain low in poorly sampled areas because the bootstrapping requires variability between samples to produce estimates of error. Examining the coverage of the environmental space (section 2.6.2) and the species’ recorded locations is therefore particularly useful; the predicted species distribution and associated uncertainty in areas with low environmental coverage values and where no species presence records exist should be treated with greater caution.

2.6.4 Independent evaluation

Independently collected presence and (true) absence data have been collected by Fisheries New Zealand inshore fishery observers using “Trimble Nomad” GPS-based data loggers since 2009. Trained fishery

observers were required to conduct an inspection of the waters surrounding the vessel approximately every 20 minutes. These data were collated by DOC into a single NOMAD database (unpublished database, held by Fisheries New Zealand/DOC) with records for all on-effort periods and positive sightings of all marine mammal species observed (i.e., on-effort events with no sightings can be considered true zeros for all species not observed). In this study, NOMAD records were used to validate the average year-round mean probability of presence maps for bottlenose dolphin (*Tursiops truncatus*), Dusky dolphin (*Lagenorhynchus obscurus*), Hector's dolphin (*Cephalorhynchus hectori hectori*), and common dolphin (*Delphinus delphis*). Only models for which there were 50+ positive species records in the NOMAD database were validated. NOMAD presence/absence data were used to calculate the AUC by comparing the predicted values from the average year-round mean probability of presence models at locations of observed presence/absence.

2.6.5 Data presentation

For clarity, results are provided for a select number of species. Results for all other species are provided in Appendix 2 (RES predicted distributions) and Appendix 3 (BRT model outputs). In the first part of the results section we provide two examples of RES mapped distribution: one for species where the results of the RES are not fully corroborated by the (very limited) recorded sightings and strandings data (i.e., beach cast data used as a visual validation); and for another species where the data better support RES predictions. The subsequent section provides a broad overview of the BRT model results for all species. Finally, detailed BRT model results, and mapped distributions are provided for five species (bottlenose dolphin, dusky dolphin, Hector's dolphin, common dolphin, and Bryde's whale) which were selected on the basis of either i) being classified as Nationally Critical / Nationally Endangered and/or ii) highlighted as being potentially at risk from commercial fishing (see Figure 1-1).

For species where BRT models were fitted, we provide a qualitative (and subjectively defined) description of sample number, species distribution, assessment of model fits for probability of presence, group size models, and a brief description of any spatial differences between seasons. These descriptors aim to provide a qualitative overview of the totality of the model results and spatial distributions. Table 2-4 provides categories, qualitative descriptions, and associated colour scheme used in summary/highlight tables throughout this report. For each species where BRT models were fitted, model fit statistics, drivers of spatial distribution, geographical predictions, and associated uncertainty are also provided. Where available, the locations of cetacean strandings and NOMAD cetacean recorded sightings at sea were overlaid on the geographic predictions for visual evaluation.

Table 2-4: Model summaries/highlights: categories, qualitative descriptions, and associated colour schemes used to describe model outputs for each species.

Sample number	Distribution	P/RA		Group size model		Descriptive changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
None (0)	Unknown	No predictive power (< 0.5)	No predictive power (0)	No predictive power (0)	No predictive power (0)	Unknown
Very Low (0 – 60)	Cosmopolitan: species range extends across all or most of the study area	Poor (0.5 – 0.65)	Poor (0.0 – 0.1)	Poor (0.0 – 0.1)	Poor (0 – 0.1)	Short description of geographic changes
Low (60 – 100)		Fair (0.65 – 0.8)	Fair (0.1 – 0.2)	Fair (0.1 – 0.2)	Fair (0.1 – 0.2)	
Moderate (100 – 300)		Good (0.8 – 0.9)	Good (0.2 – 0.3)	Good (0.2 – 0.3)	Good (0.2 – 0.3)	
High (300 – 1000)	Localised: species range is restricted across the study area	Excellent (0.9 – 1)	Very good (0.3 – 0.5)	Very good (0.3 – 0.5)	Very good (0.3 – 0.5)	
Very high (1000+)			Excellent model fit (0.5 – 1.0)	Excellent model fit (0.5 – 1.0)	Excellent model fit (0.5 – 1.0)	

3. RESULTS

3.1 RES model outputs

It was not possible to quantitatively evaluate RES model performance. However, visual comparison of sightings and strandings data with the predicted RES scores suggests that for the majority of these rarely observed species, these distributions appeared reasonable (Appendix 2). Predicted RES scores with sightings and stranding records (overlaid for visual validation) for false killer whale (*Pseudorca crassidens*) and southern right whale dolphin (*Lissodelphis peronii*) are shown in Figure 3-1. For both species, the highest RES scores (score of 1) covered extensive parts of the study area (Figure 3-1). Sightings records ($n = 28$) and stranding records for false killer whale were largely inshore around both the North Island and South Island, providing some evidence for the cosmopolitan distribution described by the RES estimates (Figure 3-1). However, the majority of sightings records were in areas with low RES estimates close to shore. It is unclear whether the sightings records simply reflect the sampling bias which favours sightings inshore, or whether the RES did not effectively reflect the average year-round geographical ranges of this species. In contrast, most of the sightings records for southern right whale dolphin ($n = 27$) were located in a cluster offshore of the south-east coast of the South Island in areas with high RES values providing some evidence that this species has a preference for deeper, offshore waters (Figure 3-1). Further, the stranding records were located on both the North Island and South Island (as far north as the Bay of Islands), providing some evidence that these species may also use the offshore areas of the North Island despite no sightings recorded further north than the Taranaki Bight (Figure 3-1).

RES predictions for all other species assessed using this method (Table 2-2) are displayed in Appendix 2.

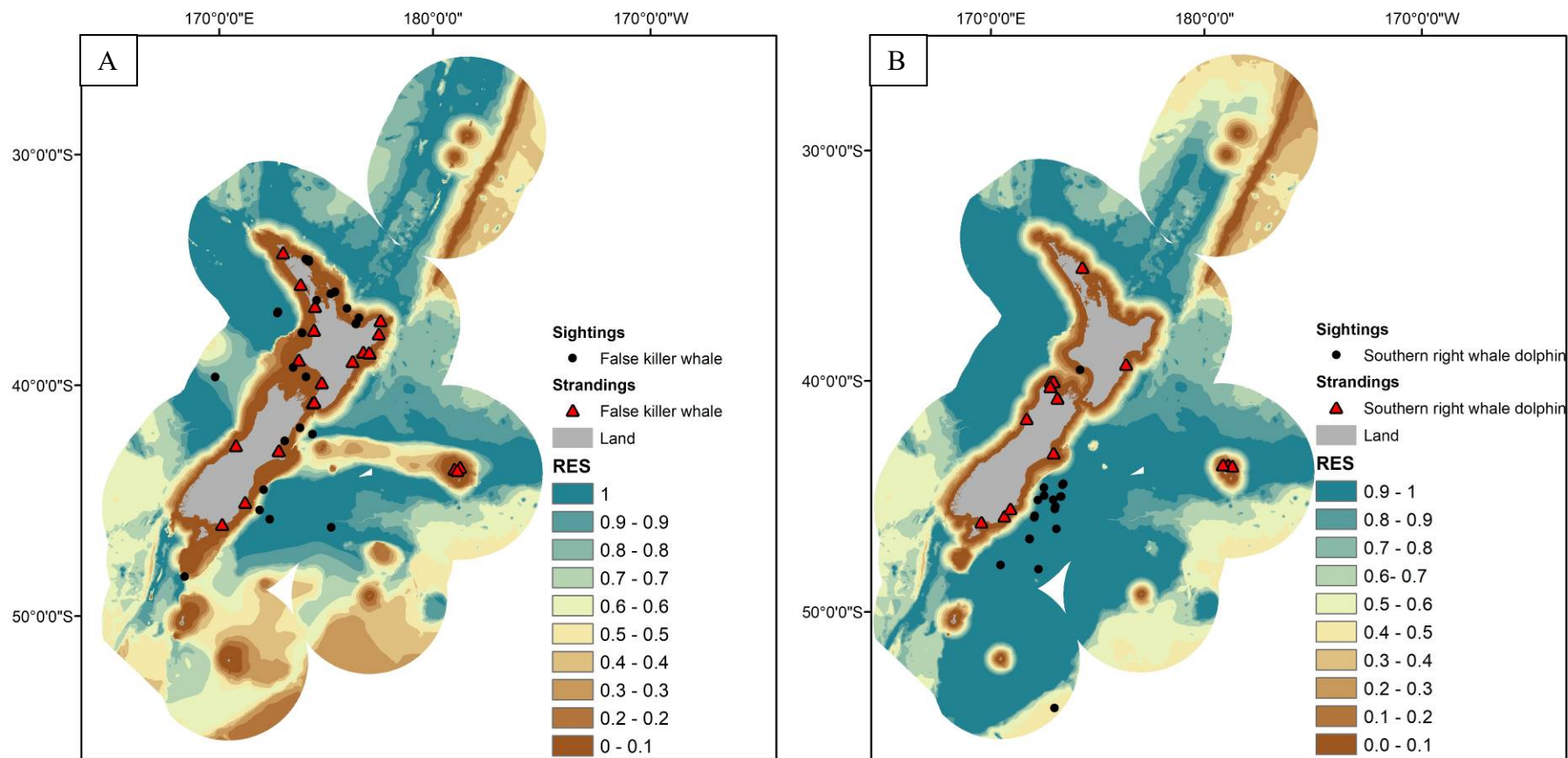


Figure 3-1: Predicted RES scores for A) false killer whale (*Pseudorca crassidens*) and B) southern right whale dolphin (*Lissodelphis peronii*), ranging from less suitable (brown) to very suitable (blue). Predicted RES scores are shown with sightings at sea and location of recorded strandings (from the DOC marine mammal strandings database).

3.2 BRT model outputs

3.2.1 Overview: average year-round presence/relative absence model outputs

For each species, two measures of model fit were used to evaluate the validity of average year-round presence/relative absence models: the mean deviance in average year-round presence/relative absence explained by the model, and the mean area under the AUC (Table 3-1). The mean deviance in presence/relative absence explained by the model and AUC were evaluated using data to fit the model (training data) and cross-validated data (evaluation data) that were systematically withheld from the modelling process when fitting the model. The latter is considered a more robust and conservative method of evaluating goodness-of-fit of a model than using the same data with which the model was trained (Friedman et al. 2001). As assessed by model fit measures using evaluation data, the models were able to explain between 16% (killer whale – qualitatively described as ‘fair’) and 88% (Māui dolphin – qualitatively described as ‘excellent’) and a mean 43% of the deviance in species presence/relative absence (Table 3-1). The AUC scores (measuring the ability for binary classification between presence and absence points) ranged from 0.79 (killer whale – qualitatively described as ‘fair’) to 0.99 (Hector's and Māui dolphins – qualitatively described as ‘excellent’) with a mean of 0.9 (Table 3-1).

Broadly, model goodness-of-fit increased with increasing number of records (e.g., see number of records in Table 2-2). In addition, model fit appeared to be linked to whether species records were widely distributed across the study area (poorer fitting models, e.g., killer whale, bottlenose dolphin – further details in section 3.2.5.1) or more tightly grouped in smaller areas of the study area (e.g., Māui dolphin, Hector's dolphin, Bryde's whale – further details in sections 3.2.5.3 and 3.2.5.5).

The model fit metrics between species training data and evaluation data were in most cases very similar (see Table 3-1) providing some evidence that models were not overly fitted to the training data. For those species with adequate sample numbers for bootstrapping, standard deviations for each model fitting metric are also provided in Table 3-1). Generally, these standard deviations were low (except for fin, minke, and sei whales where these were not available), providing evidence that models performed consistently across bootstrap samples.

Table 3-1: Cross-validated estimates of model performance for the bootstrapped BRT models fitted with presence/relative absence sightings per species. Model performance was assessed using the proportion of the mean total deviance explained \pm standard deviation (StDev) and mean AUC \pm standard deviation for models fitted with the training data (75% of sightings records) and the evaluation data (25% sightings records). Models for fin, minke, and sei whales were not bootstrapped and therefore only the results from a single model are g. Table colours refer to qualitative descriptions provided in Table 2.4.

Species		Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)
Fin whale		0.20	0.19	0.81	0.90
Minke whale		0.18	0.25	0.79	0.88
Sei whale		0.24	0.24	0.81	0.88
Killer whale	Mean	0.18 \pm 0.02	0.16 \pm 0.02	0.79 \pm 0.01	0.79 \pm 0.02
Bottlenose dolphin	Mean	0.20 \pm 0.02	0.19 \pm 0.02	0.81 \pm 0.01	0.81 \pm 0.02
Humpback whale	Mean	0.29 \pm 0.02	0.28 \pm 0.03	0.85 \pm 0.01	0.85 \pm 0.02
Common dolphin	Mean	0.39 \pm 0.00	0.39 \pm 0.01	0.90 \pm 0.00	0.90 \pm 0.00
Pilot whale	Mean	0.43 \pm 0.02	0.43 \pm 0.03	0.91 \pm 0.01	0.91 \pm 0.01
Sperm whale	Mean	0.46 \pm 0.02	0.46 \pm 0.04	0.92 \pm 0.01	0.92 \pm 0.01
Bryde's whale	Mean	0.49 \pm 0.02	0.50 \pm 0.04	0.93 \pm 0.01	0.93 \pm 0.01
Right whale	Mean	0.53 \pm 0.02	0.53 \pm 0.04	0.94 \pm 0.01	0.94 \pm 0.01
Dusky dolphin	Mean	0.55 \pm 0.01	0.55 \pm 0.03	0.95 \pm 0.00	0.95 \pm 0.01
Blue whale	Mean	0.58 \pm 0.02	0.58 \pm 0.05	0.95 \pm 0.01	0.95 \pm 0.01
Māui dolphin	Mean	0.87 \pm 0.01	0.88 \pm 0.02	0.99 \pm 0.00	0.99 \pm 0.00
Hector's dolphin	Mean	0.83 \pm 0.01	0.83 \pm 0.01	0.99 \pm 0.00	0.99 \pm 0.00

The number of environmental predictor variables retained for species modelled with average year-round presence/relative absence ranged from 6 to 13, with a median of 8 across all species (Table 3-2). The relative importance of each environmental variable (% deviance explained) varied across average year-round presence/relative absence models for each species (Table 3-2). The most consistently important variable for predicting the presence/relative absence was Temperature Residuals (TempRes) which was selected in 14 out of 15 models with an average contribution of 23.8% (Table 3-2). The next most important variables across all species models were bathymetry (Bathy), distance to the 500 m isopleth (Dist.Iso500), mixed layer depth (MLD), and water turbidity (Turb) – each of which was selected in 13 of the 15 species models (Table 3-2). Benthic sediment disturbance (BedDist) was seldom included in the models due to its poor predictive power relative to other environmental variables (Table 3-2).

Table 3-2: Variable contributions for bootstrapped BRT models fitted with presence/relative absence sightings per cetacean species. The percent contribution (%) \pm standard error (se) is shown for each environmental variable (light grey cells represent low percent contribution; darker grey cells represent higher percent contributions). For those variables not used in the final BRT models, table cells are blank. Models for fin, minke, and sei whales were not bootstrapped and therefore only the results from a single model are given. Environmental variables are described in Table 2-3.

Environmental Variables	Fin Whale	Minke whale	Sei whale	Killer whale mean	Bottlenose dolphin mean	Humpback whale mean	Common dolphin mean	Pilot whale mean
Bathy	16.1	14.1	5.3	16.9 \pm 0.4	12.1 \pm 0.3		0.3 \pm 0.0	25.3 \pm 0.8
BedDist			3.6					
ChlA	9.0	4.0	5.8	6.6 \pm 0.2		7.9 \pm 0.2	0.5 \pm 0.0	
Dist.Iso500	1.7	18.2	8.4	7.0 \pm 0.2		8.5 \pm 0.2	2.3 \pm 0.0	2.2 \pm 0.1
Dist.Shore	1.9	2.8	7.3	11.4 \pm 0.3	11.9 \pm 0.3	8.4 \pm 0.2	17.1 \pm 0.1	52.7 \pm 0.8
DOM	17.5	33.5	5.0	3.3 \pm 0.1			1.5 \pm 0.0	6.5 \pm 0.3
Kpar	3.9					5.6 \pm 0.2	1.5 \pm 0.1	2.6 \pm 0.1
MLD	5.3	21.8	3.9	9.4 \pm 0.2	9.6 \pm 0.2	27.9 \pm 0.4	0.7 \pm 0.0	
Slope	3.4		3.2	10.3 \pm 0.2	10.7 \pm 0.3	15.1 \pm 0.3		
SST	13.7		4.6	9.8 \pm 0.3		11.1 \pm 0.3	9.0 \pm 0.1	5.3 \pm 0.1
TC	2.6		6.1	6.6 \pm 0.3			0.6 \pm 0.0	
TempRes	5.1		6.2	18.6 \pm 0.3	26.5 \pm 0.3	15.5 \pm 0.2	42.9 \pm 0.1	2.0 \pm 0.1
Turb	12.0	5.6	40.7		15.3 \pm 0.2		21.5 \pm 0.1	3.5 \pm 0.1
VGPM	8.0				13.8 \pm 0.2		2.2 \pm 0.0	

Environmental Variables	Sperm whale mean	Bryde's whale mean	Right whale mean	Dusky dolphin mean	Blue whale mean	Māui dolphin mean	Hector's dolphin mean
Bathy	73.1 \pm 0.3	3.0 \pm 0.1	8.5 \pm 0.6	6.9 \pm 0.2	9.3 \pm 0.3		4.3 \pm 0.0
BedDist					4.2 \pm 0.2		
ChlA		6.7 \pm 0.2		0.8 \pm 0.0		1.0 \pm 0.0	
Dist.Iso500	3.7 \pm 0.1		2.6 \pm 0.2	16.4 \pm 0.2	11.1 \pm 0.4	4.7 \pm 0.2	3.3 \pm 0.0
Dist.Shore	2.8 \pm 0.1		15.8 \pm 0.8		48.8 \pm 0.4		
DOM	4.2 \pm 0.1		6.8 \pm 0.4	1.3 \pm 0.1	1.9 \pm 0.1		
Kpar			7.6 \pm 0.5	0.9 \pm 0.0	4.0 \pm 0.1	0.7 \pm 0.0	
MLD	1.8 \pm 0.1	1.6 \pm 0.1	27.4 \pm 0.8	3.4 \pm 0.1	2.4 \pm 0.1		29.5 \pm 0.3
Slope					3.2 \pm 0.2		
SST	2.5 \pm 0.1	6.2 \pm 0.2	26.4 \pm 1.0	10.5 \pm 0.2	7.2 \pm 0.2		10.9 \pm 0.1
TC	5.5 \pm 0.1	9.7 \pm 0.3		1.4 \pm 0.1		14.8 \pm 0.2	
TempRes	3.3 \pm 0.1	51.7 \pm 0.2	4.8 \pm 0.3	55.3 \pm 0.3	4.6 \pm 0.2	72.7 \pm 0.1	24.9 \pm 0.1
Turb		13.6 \pm 0.3		1.8 \pm 0.1	3.2 \pm 0.2	6.1 \pm 0.1	25.5 \pm 0.3
VGPM	3.0 \pm 0.1	7.6 \pm 0.2		1.4 \pm 0.1			1.6 \pm 0.0

3.2.2 Overview: average year-round group size model outputs

Two cross-validated measures of model fit were used to evaluate average year-round species group size models (Table 2-2): the mean deviance in relative group size explained and a Pearson's correlation between predicted group size and observed group size (Table 3-3). The average year-round group size models explained much lower percentages of the model deviance than average year-round presence/relative absence models; ranging from 0% (Bryde's whale and killer whale – no predictive power) to 26% (Hector's dolphin – qualitatively described as 'good'). On average, 10% of the deviance was explained by species group size models (Table 3-3). Pearson's correlations between predicted group size and observed group size ranged from 0.12 (Bryde's whale – qualitatively described as 'poor') to 0.43 (common dolphin – qualitatively described as 'very good') with a mean of 0.28 (Table 3-3). The relatively low correlation between predicted and observed group size was largely driven by under-prediction of higher (less frequently observed) group sizes for the majority of dolphin species and low variation in group sizes for some whale species (e.g., sperm whales and Bryde's whales) (see section 3.2.3).

In contrast to the models fitted with average year-round presence/relative absence data, average year-round species group size model metrics derived from training data differed from those derived from evaluation data. For example, mean explained deviance of sperm whale models using training data was 35% compared with the mean explained deviance using evaluation data of 10%, indicating that some of the group size models may be somewhat over-fitted to the training data (Table 3-3). Additionally, the standard deviations of these model fit metrics are much larger than for average year-round presence/relative absence models, particularly for model fit metrics using evaluation data (Table 3-3).

Table 3-3: Cross-validated estimates of model performance for the bootstrapped BRT models fitted with cetacean species group size records. Model performance was assessed using the proportion of the mean total deviance explained \pm standard deviation (StDev) and mean Pearson's correlation between predicted group sizes (trained using 75% of records) and observed (evaluation data - 25% of records). Table colours refer to qualitative descriptions provided in Table 2.4.

Species name	Statistic	Deviance explained (training data)	Deviance explained (evaluation data)	Correlation between predictions and evaluation data (R^2)
Bryde's whale	Mean	0.04 \pm 0.03	0.00 \pm 0.05	0.12 \pm 0.07
Killer whale	Mean	0.03 \pm 0.03	0.00 \pm 0.05	0.22 \pm 0.11
Dusky dolphin	Mean	0.10 \pm 0.06	0.09 \pm 0.08	0.25 \pm 0.12
Bottlenose dolphin	Mean	0.08 \pm 0.09	0.05 \pm 0.11	0.27 \pm 0.14
Right whale	Mean	0.14 \pm 0.06	0.08 \pm 0.12	0.27 \pm 0.14
Sperm whale	Mean	0.35 \pm 0.28	0.10 \pm 0.30	0.26 \pm 0.15
Māui dolphin	Mean	0.09 \pm 0.02	0.09 \pm 0.03	0.27 \pm 0.05
Hector's dolphin	Mean	0.26 \pm 0.04	0.26 \pm 0.06	0.42 \pm 0.12
Common dolphin	Mean	0.26 \pm 0.03	0.25 \pm 0.05	0.43 \pm 0.08

The number of environmental predictor variables selected for average year-round species group size models ranged from 8 to 12, with a median of 9 across all species (Table 3-4). The importance of environmental variables for species group size models varied between species (Table 3-4) and differed to those selected for average year-round presence/relative absence models for individual species. The most consistently important variable for predicting average year-round species group sizes was Kpar (a measure of diffuse light attenuation which is influenced by sedimentation close to the coast and chlorophyll in the open ocean); this variable was selected in all nine models albeit with a low average contribution of 8.5% (see Table 3-4). Coloured dissolved organic matter (DOM) had the highest average

contribution across all species (21%) and was selected in eight out of the nine average year-round species group size models. All other variables had more varying model contributions and their importance varied between species (see Table 3-4). Bathymetry (Bathy) and benthic sediment disturbance (BedDist) were seldom included in the models (included in 3 models) due to their poor predictive power (Table 3-4).

Table 3-4: Variable contributions for bootstrapped BRT models fitted with cetacean species group size records. For each species, the mean environmental variable percentage contribution (%) \pm standard error (se) is given. Light grey cells represent no percent contribution, darker cells represent higher percent contributions. For those variables not used in the final BRT models, table cells are blank.

	Bryde's whale	Killer whale	Dusky dolphin	Bottlenose dolphin	Right whale	Sperm whale	Māui dolphin	Hector's dolphin	Common dolphin
Environmental Variables	mean	mean	mean	mean	mean	mean	mean	mean	mean
Bathy						3.2 \pm 0.2	4.4 \pm 0.3	14.7 \pm 0.2	
BedDist	2.8 \pm 0.3						5.3 \pm 0.3		5.0 \pm 0.1
ChlA	43.5 \pm 2.8	1.4 \pm 0.2		7.6 \pm 0.5	9.2 \pm 0.8	2.5 \pm 0.5	3.1 \pm 0.1	6.1 \pm 0.5	4.9 \pm 0.1
Dist.Iso500		10.0 \pm 1.0	9.0 \pm 0.3	19.3 \pm 0.7			7.0 \pm 0.2	3.8 \pm 0.2	6.8 \pm 0.1
Dist.Shore			9.7 \pm 0.3		8.6 \pm 0.7				
DOM	18.3 \pm 1.3	65.5 \pm 2.8	8.0 \pm 0.3	10.4 \pm 0.6	5.7 \pm 0.3		8.1 \pm 0.3	36.4 \pm 1.5	15.1 \pm 0.3
Kpar	4.0 \pm 0.4	3.6 \pm 0.4	13.7 \pm 0.7	5.1 \pm 0.4	17.9 \pm 1.0	3.1 \pm 0.3	17.0 \pm 0.5	6.6 \pm 0.3	6.3 \pm 0.1
MLD	10.0 \pm 0.6		22.7 \pm 0.7	7.4 \pm 0.4	22.9 \pm 1.2		10.7 \pm 0.4		15.3 \pm 0.2
Slope		3.6 \pm 0.6	9.2 \pm 0.3	4.9 \pm 0.5	2.2 \pm 0.2	1.9 \pm 0.3	10.4 \pm 0.3	8.3 \pm 0.3	5.7 \pm 0.1
SST	7.3 \pm 0.4	1.1 \pm 0.1		10.1 \pm 0.4	4.2 \pm 0.4	45.3 \pm 1.7	16.9 \pm 0.3	8.0 \pm 0.2	10.7 \pm 0.1
TC	3.5 \pm 0.3	3.0 \pm 0.4	19.8 \pm 0.8	5.2 \pm 0.3		5.5 \pm 0.4	6.1 \pm 0.2	13.5 \pm 0.4	9.7 \pm 0.1
TempRes			7.9 \pm 0.3	23.5 \pm 1.5	3.8 \pm 0.2	4.2 \pm 0.3	5.7 \pm 0.2	2.7 \pm 0.1	6.2 \pm 0.1
Turb				3.8 \pm 0.4	17.1 \pm 0.7	34.2 \pm 1.5	5.2 \pm 0.2		6.1 \pm 0.1
VGPM	10.6 \pm 0.8	11.6 \pm 0.9		2.7 \pm 0.2	8.3 \pm 0.6				8.2 \pm 0.1

3.2.3 Overview: average seasonal presence/relative absence model outputs

Seasonal presence/relative absence models for individual species had similar mean values of model goodness-of-fit to their respective average year-round presence/relative absence models, albeit in most cases one of the seasonal models had lower predictive power (Table 3-5). Similar to average year-round presence/relative absence models, seasonal model fit metrics differed little between those derived from training data and evaluation and mean estimates did not vary greatly between bootstrap samples (Table 3-5) providing evidence that these models were not overfitted to the training data. For several species (Killer whale, Dusky dolphin, Hector's dolphin and Common dolphin) there were marked differences in model fits (explained deviance and AUC) between winter and summer presence/relative absence models (Table 3-5). Unsurprisingly, seasonal differences in model fits were seemingly linked to differences in sample number (i.e., seasons with lower of number of records had lower predictive power).

Table 3-5: Cross-validated estimates of model performance for the bootstrapped BRT models fitted with seasonal (winter, summer) presence/relative absence cetacean species sightings. Model performance was assessed using the proportion of the mean total deviance explained \pm standard deviation (StDev) and mean AUC (area under the receiver operating curve) (AUC) \pm standard deviation for models fitted with the training data (75% of sightings records) and the evaluation data (25% sightings records). Table colours refer to qualitative descriptions provided in Table 2.4.

Species name	Season	Statistic	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)
Bryde's whale	Winter	Mean	0.51 \pm 0.05	0.51 \pm 0.10	0.93 \pm 0.01	0.93 \pm 0.02
	Summer	Mean	0.50 \pm 0.06	0.50 \pm 0.11	0.93 \pm 0.01	0.93 \pm 0.02
Killer whale	Winter	Mean	0.09 \pm 0.24	0.08 \pm 0.36	0.71 \pm 0.03	0.73 \pm 0.06
	Summer	Mean	0.27 \pm 0.09	0.26 \pm 0.16	0.84 \pm 0.02	0.84 \pm 0.03
Dusky dolphin	Winter	Mean	0.40 \pm 0.09	0.41 \pm 0.19	0.90 \pm 0.02	0.91 \pm 0.03
	Summer	Mean	0.64 \pm 0.03	0.64 \pm 0.06	0.96 \pm 0.00	0.96 \pm 0.01
Bottlenose dolphin	Winter	Mean	0.17 \pm 0.18	0.17 \pm 0.36	0.77 \pm 0.03	0.79 \pm 0.06
	Summer	Mean	0.29 \pm 0.08	0.28 \pm 0.16	0.85 \pm 0.02	0.85 \pm 0.03
Humpback whale	Winter	Mean	0.30 \pm 0.06	0.29 \pm 0.13	0.85 \pm 0.01	0.85 \pm 0.03
	Summer	Mean	0.27 \pm 0.11	0.27 \pm 0.19	0.84 \pm 0.02	0.84 \pm 0.03
Sperm whale	Winter	Mean	0.47 \pm 0.08	0.48 \pm 0.16	0.93 \pm 0.01	0.93 \pm 0.03
	Summer	Mean	0.48 \pm 0.06	0.48 \pm 0.10	0.92 \pm 0.01	0.92 \pm 0.02
Māui dolphin	Winter	Mean	0.83 \pm 0.03	0.85 \pm 0.08	0.99 \pm 0.00	0.99 \pm 0.01
	Summer	Mean	0.89 \pm 0.01	0.89 \pm 0.03	1.00 \pm 0.00	1.00 \pm 0.00
Hector's dolphin	Winter	Mean	0.62 \pm 0.04	0.61 \pm 0.09	0.96 \pm 0.01	0.95 \pm 0.02
	Summer	Mean	0.85 \pm 0.00	0.86 \pm 0.01	0.99 \pm 0.00	0.99 \pm 0.00
Common dolphin	Winter	Mean	0.33 \pm 0.02	0.33 \pm 0.06	0.87 \pm 0.00	0.89 \pm 0.01
	Summer	Mean	0.49 \pm 0.01	0.50 \pm 0.03	0.93 \pm 0.00	0.93 \pm 0.01
Pilot whale	Winter	Mean	0.47 \pm 0.09	0.49 \pm 0.14	0.91 \pm 0.02	0.92 \pm 0.03
	Summer	Mean	0.44 \pm 0.05	0.45 \pm 0.09	0.91 \pm 0.01	0.92 \pm 0.02

3.2.4 Coverage of the environmental space

Geographic prediction of the coverage of the environmental space by the samples (see section 2.6.2) provided a spatially explicit indication of areas where species model predictions were extrapolated beyond the environmental characteristics of the input data (Figure 3-2, A). Areas where the environmental space was well covered by samples were predominately located within the region 50 – 100 km from shore around the North Island, South Island and the Chatham Islands, as well as smaller

areas surrounding the Auckland, Campbell, Antipodes and Kermadec Islands (red areas in Figure 3-2, A). Poorly, covered areas included much of the study area further from shore (blue areas in Figure 3-2, A).

Here, we have recommended a subjectively defined threshold of 0.1 (i.e. in areas with lower values we have limited understanding of the environmental space and therefore predictions of species distributions are less certain). Areas with values lower than this cut-off are shown in the hashed grey areas in Figure 3-2, B. In these poorly covered areas of the environmental space, geographic predictions from the species distribution models should be treated with greater scepticism and caution.

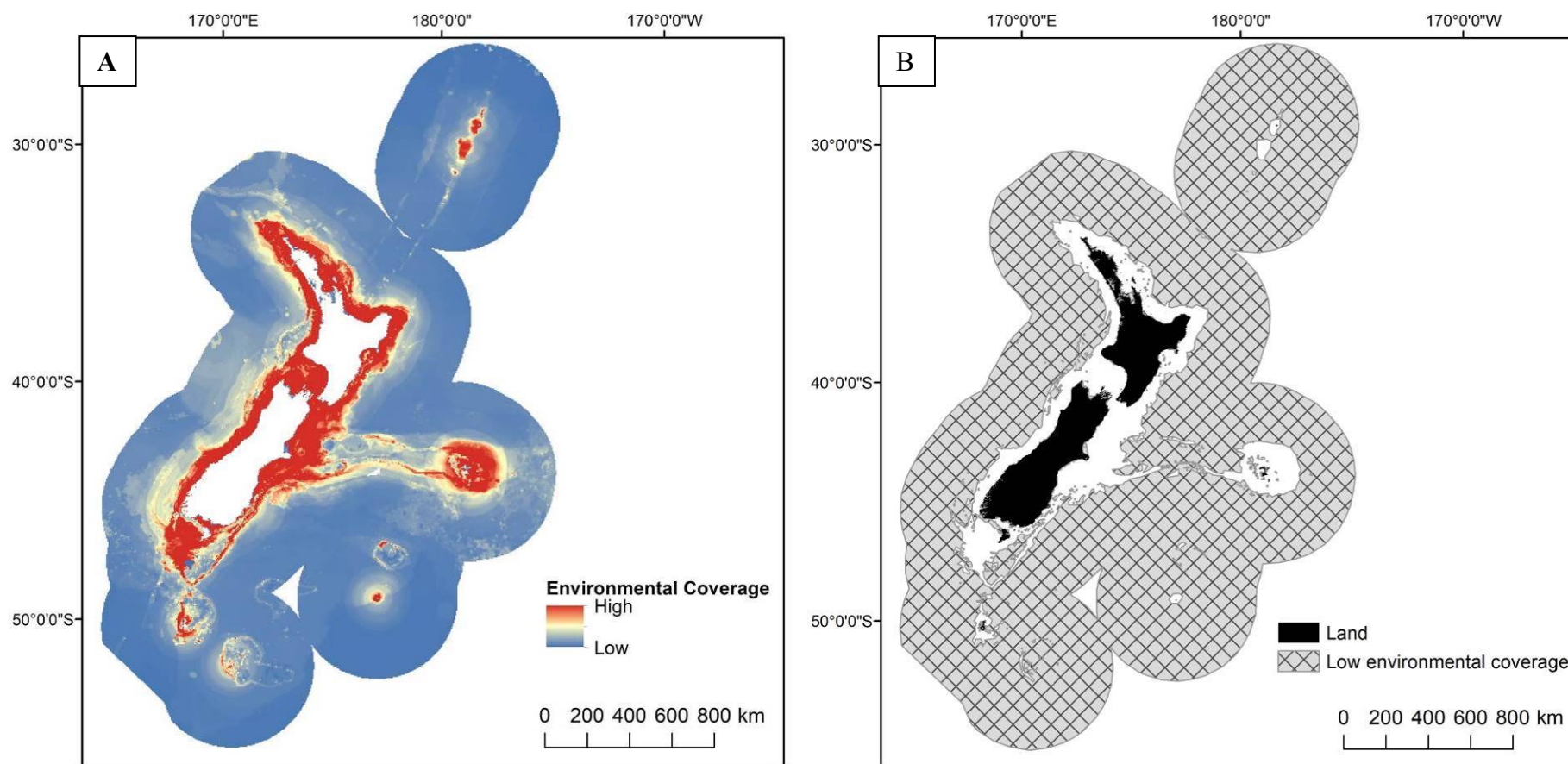


Figure 3-2: A) Predicted environmental coverage — index of how much confidence can be placed in the predictions, ranging from low (i.e., no samples in the dataset with those environmental conditions) to high (i.e., many samples with those environmental conditions). B) Subjectively defined areas of low environmental coverage (see section 2.6.2).

3.2.5 Species specific BRT model outputs and spatial predictions

In the following section detailed BRT model results, and mapped distributions are provided for bottlenose dolphin, dusky dolphin, Hector's dolphin, common dolphin, and Bryde's whale. Results for all other species are provided in Appendix 3.

3.2.5.1 Bottlenose dolphin (*Tursiops truncatus*)

3.2.5.1.1 Overview

Qualitative descriptions of bottlenose dolphin sample number, distribution, assessment of models fits (including independent validation using NOMAD data) for average year-round presence/relative absence and group size models, and a brief description of spatial differences between seasonal distributions are provided in the highlights table (Table 3-6).

Table 3-6: Bottlenose dolphin (*Tursiops truncatus*): highlights of model fits (including independent validation using NOMAD data) and geographic prediction. Table colours and qualitative categories used here are provided in Table 2.4.

Sample number	Distribution	P/RA			Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	NOMAD Evaluation (AUC)	Model fit (R ²)	Model fit (dev. Exp)	
High	Cosmopolitan	Good	Fair	Fair	Good	Poor	Expansion from northern distribution in winter to cosmopolitan distribution in summer

3.2.5.1.2 Average year-round presence/relative absence models

Deviance explained and AUC scores for bottlenose dolphin average year-round presence/relative absence models were reasonable and were consistent between training and evaluation data and between bootstrap samples (Table 3-7). Model validation with independently collected presence/(true) absence data showed that the model had some predictive power (NOMAD evaluation AUC: 0.66, 'fair') although this was substantially lower than for evaluation with non-independent data (Table 3-7).

Table 3-7: Mean presence/relative absence model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of bottlenose dolphin (*Tursiops truncatus*).

	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)	NOMAD Evaluation (AUC)
Mean	0.20	0.19	0.81	0.81	0.66
Standard Deviation	0.02	0.02	0.01	0.02	

The most important environmental predictor variables for common dolphin presence/relative absence models were TempRes (23.8%), Turb (15.3%), VGPM (14.6%) and Dist.Shore (11.2%) (Figure 3-3). Keeping all other variables constant, the probability of presence increased in areas with more than 1.5 °C bottom water temperatures expected for a given depth (Figure 3-3). Higher probability of presence was predicted for areas with low turbidity (Turb), low productivity (VGPM), and areas close to shore (Dist.Shore) (see Figure 3-3). Overall, there was low uncertainty in predicted relationships between predicted probability of presence and the environmental variables as shown by the 95%

prediction intervals (dashed lines in Figure 3-3). However, these relationships were not strong (i.e., no predictions were above 0.4 — Figure 3-3).

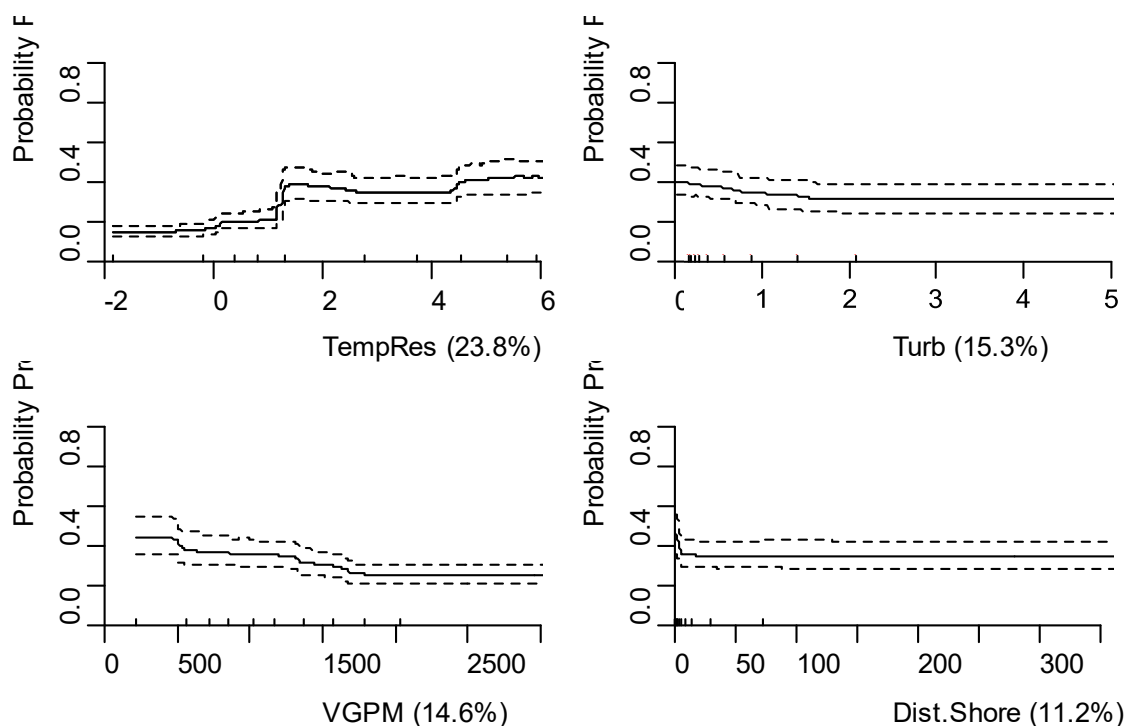


Figure 3-3: Partial dependence plots showing the relationships between predictor variables and the probability of bottlenose dolphin (*Tursiops truncatus*) presence modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown as ticks on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Bottlenose dolphin predicted probability of presence was moderate (0.3 – 0.5) across large parts of the study area (Figure 3-4). The highest probability of presence (0.79) was for areas very close to north-eastern coasts of the North Island and some large areas further offshore of the east coast of the North Island and the Kermadec Ridge (Figure 3-4) where few presence records exist (Figure 3-5).

Due to the cosmopolitan distribution of the bottlenose dolphin sightings records, the UD (Utilisation Distribution) covers a large portion of the study area (Figure 3-4); it is therefore difficult to comment on the congruence of the UD with the predicted probability of presence. Visually, presence/relative absence, strandings, and NOMAD location records overlapped with some of the highest predicted probability areas, but due to the moderate probability values across much of the study area, these do not provide firm evidence of the adequacy of the model (Figure 3-5).

Spatially explicit estimates of uncertainty (measured here as CV) were low across the study area (Figure 31 and 32 in Appendix 3). However, given the lack of spatial discrimination in the predicted probability presence (i.e., largely moderate probability presences predicted) and the wide distribution of recorded sightings across the study area, the interpretation of this uncertainty layer should be treated with some caution in areas where few samples exist (see section 2.6.3 for further details). Knowledge of the coverage of the environmental space and the recorded locations is therefore particularly useful and predictions of bottlenose dolphin in areas with low environmental coverage values and where no presences have been observed should be treated with caution. The cosmopolitan distribution of this

species and overlap with many of the other species presence records (which were used as relative absences here) may have resulted in the moderate predictive ability of the model.

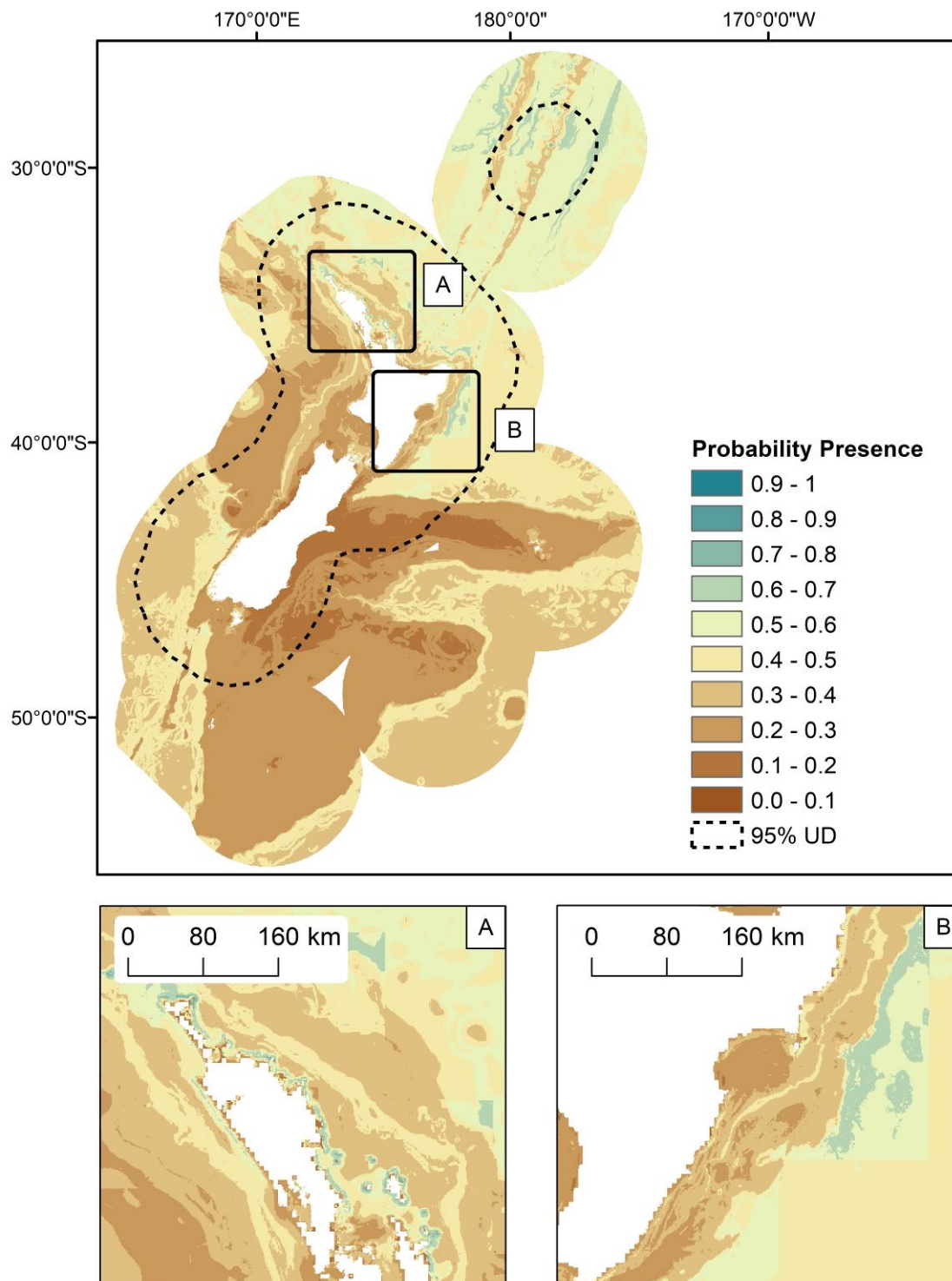


Figure 3-4: The predicted probability of bottlenose dolphin (*Tursiops truncatus*) presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is defined by the dashed line. Inset maps: A) north of North Island and Hauraki Gulf; B) south east coast of North Island.

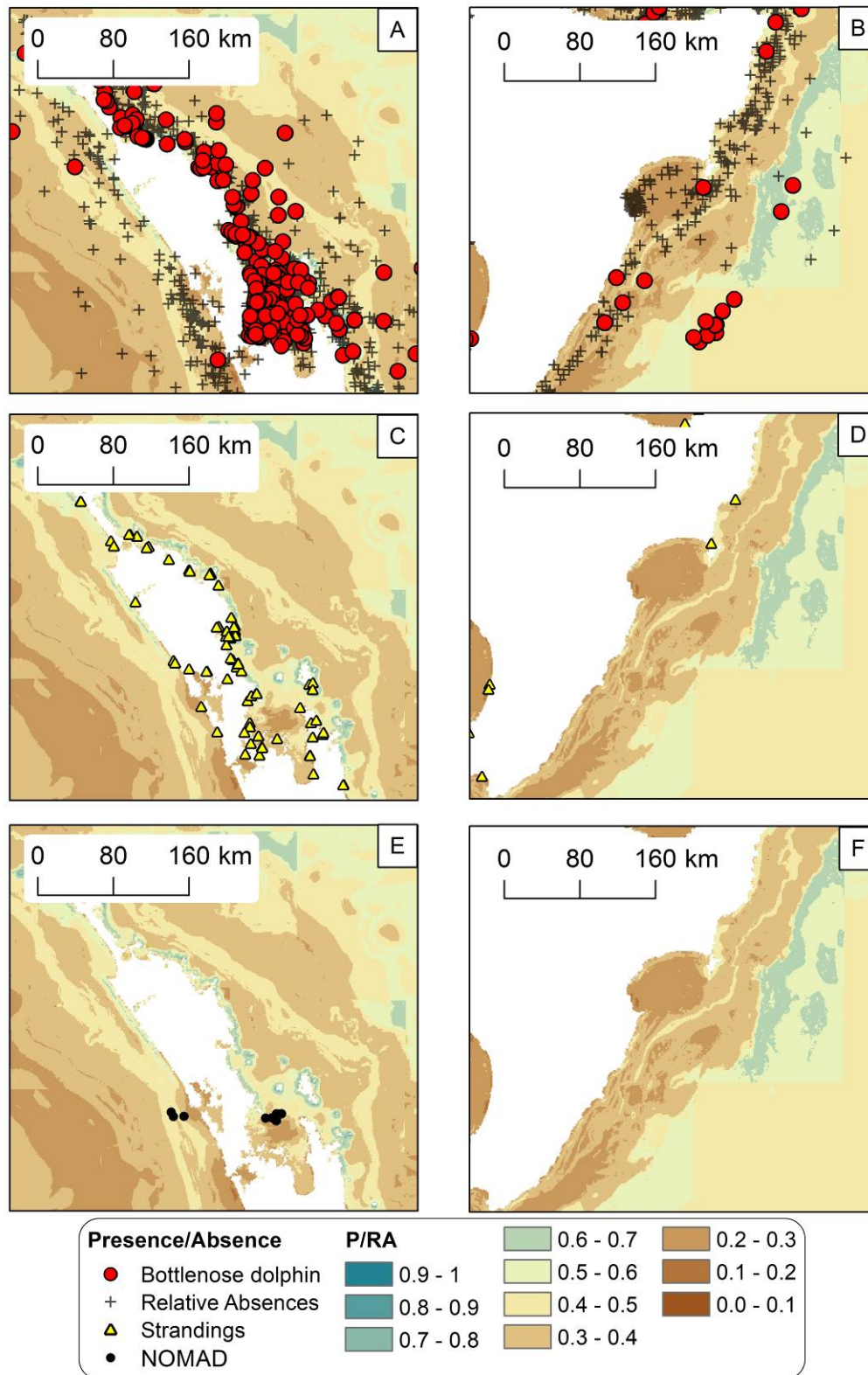


Figure 3-5: The predicted probability of bottlenose dolphin (*Tursiops truncatus*) presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted probability presence of Bottlenose dolphin in the north of North Island and Hauraki Gulf are shown (left panel) with presence/relative absence (red circles and black crosses respectively) (A), DOC stranding locations (C), and NOMAD sightings (E). Predicted probability presence of bottlenose dolphin on the south east coast of North Island are shown (right panel) with presence/relative absence (B), stranding locations (D), and NOMAD sightings (F).

3.2.5.1.3 Average year-round group size models and relative density predictions

Bottlenose dolphin average year-round group size models had low predictive power (0.05 mean deviance explained) and moderate correlative power for predicted vs observed species group size (Table 3-8). The very low model fits observed indicates that interpretation of predicted bottlenose dolphin average year-round group size should be treated with caution.

Table 3-8: Mean model performance measures (deviance explained and Pearson's correlation of predicted vs observed group sizes (R^2)) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of Bottlenose dolphin (*Tursiops truncatus*).

	Deviance explained (training data)	Deviance explained (evaluation data)	Pearson's correlation of predicted vs observed species group sizes (R^2)
Mean	0.08	0.05	0.27
Standard Deviation	0.09	0.11	0.14

The most important environmental predictor variables in average year-round bottlenose dolphin group size models were Dist.Iso500 (19.0%) and TempRes (24%), although relationships were unclear, weak, and had large predictive error (Figure 3-6).

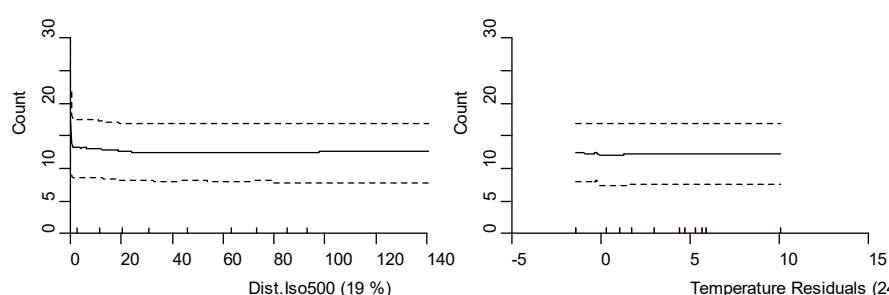


Figure 3-6: Partial dependence plots showing the relationships between predictor variables and bottlenose dolphin (*Tursiops truncatus*) group sizes, modelled using bootstrapped BRTs. The two most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Predicted average year-round bottlenose dolphin relative density distribution (the resulting geographical prediction from the two-step hurdle method) were predicted to be highest in offshore areas across the majority of the study area, in particular on the Kermadec Ridge (Figure 3-7). Uncertainty estimates (95% prediction interval) were relatively high and variable across the study area, likely reflecting the poor model fits (Figure 33, Appendix 3). Overall, the range of predicted relative density (Figure 3-7) were in line with the observed group sizes (Figure 3-8), although the low spatial correlation, the low model predictive power, and the occurrence of the highest predicted values in areas with low environmental coverage indicates the predicted average year-round relative density should be treated with caution.

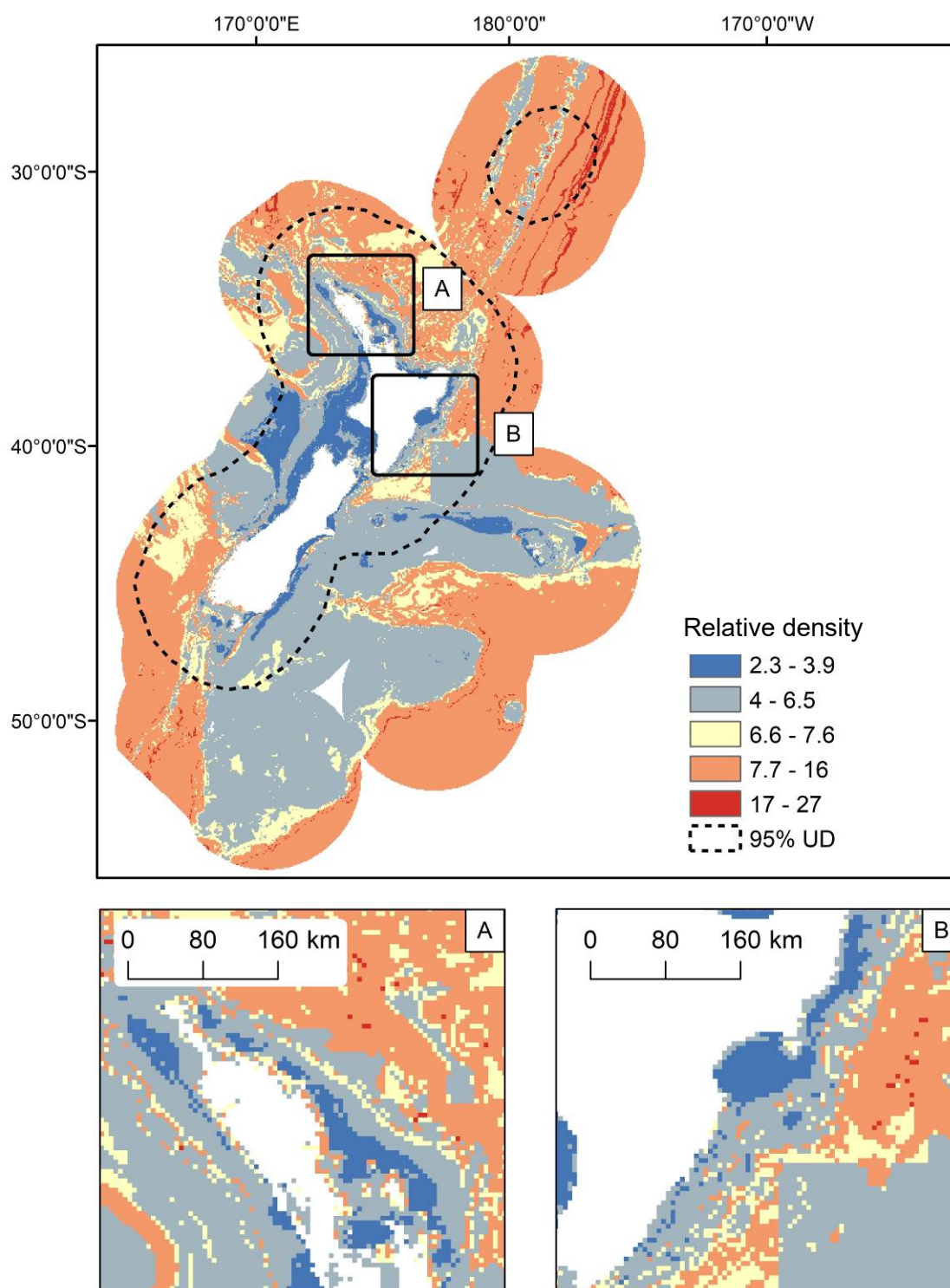


Figure 3-7: Predicted relative density of bottlenose dolphin (*Tursiops truncatus*) in the New Zealand EEZ, from bootstrapped BRT models. The predicted 95% utilisation distribution is defined by the dashed lines. Inset maps: A) north of North Island and Hauraki Gulf; B) south east coast of North Island.

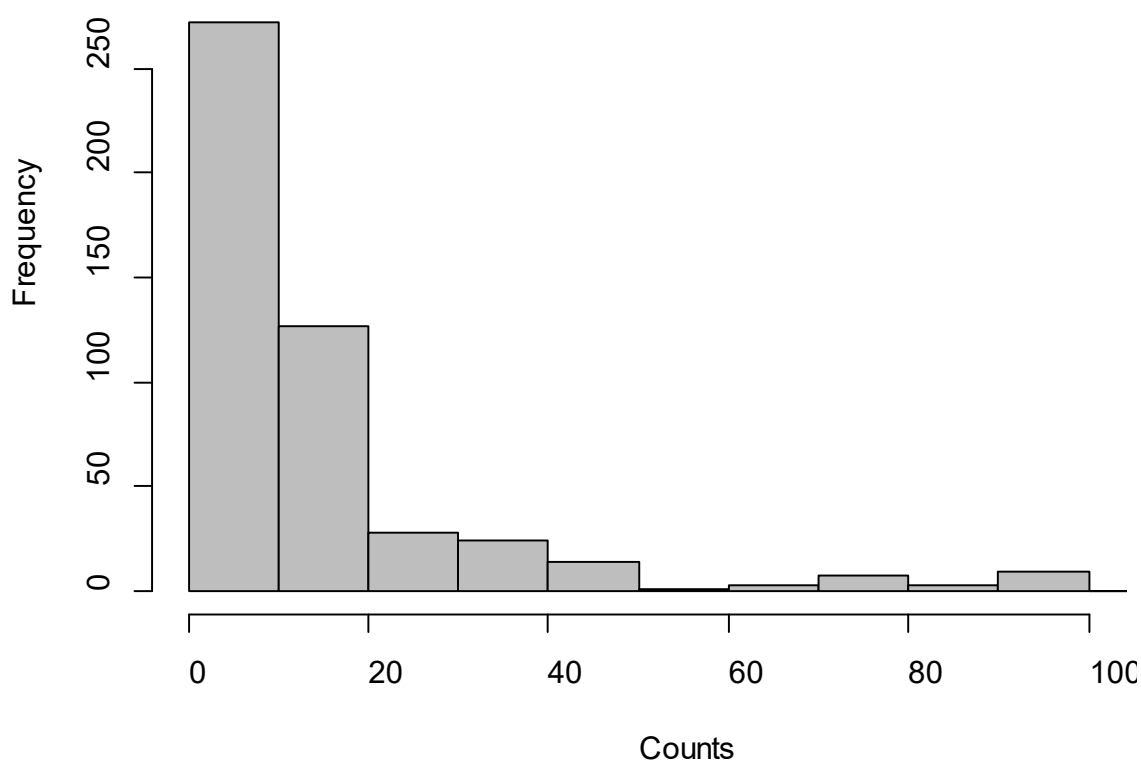


Figure 3-8: Histogram of bottlenose dolphin (*Tursiops truncatus*) group size recorded in the New Zealand EEZ.

3.2.5.1.4 Seasonal presence/relative absence models

Goodness-of-fit metrics for summer presence/relative absence models were higher (explained deviance: 0.28; AUC: 0.85) than both average year-round (explained deviance: 0.19; AUC: 0.81) and winter presence/relative absence models (explained deviance: 0.17; AUC: 0.79) (Table 3-9). Seasonal model fit metrics differed little between those derived from training data and evaluation, although mean estimates varied between bootstrap samples (particularly for winter evaluation data; see Table 3-7). The difference in model fits between seasonal models may be explained by differences in sample number, i.e., there were fewer species presence records in winter ($n = 169$) compared with summer ($n = 654$).

Table 3-9: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) from winter (May-Oct) and summer (Nov-Apr) sightings of bottlenose dolphin (*Tursiops truncatus*) records.

Season	Metric	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)
Winter	Mean	0.17	0.17	0.77	0.79
	Standard Deviation	0.18	0.36	0.03	0.06
	Mean	0.29	0.28	0.85	0.85
Summer	Standard Deviation	0.08	0.16	0.02	0.03

The four most important environmental predictors in bottlenose dolphin seasonal models differed between seasons and to those used in the average year-round presence/relative absence models (Figure 3-9). Stronger, albeit more variable, relationships between predicted probability of presence along environmental gradients were observed between seasons to those observed in average year-round presence/relative absence models (Figure 3-9).

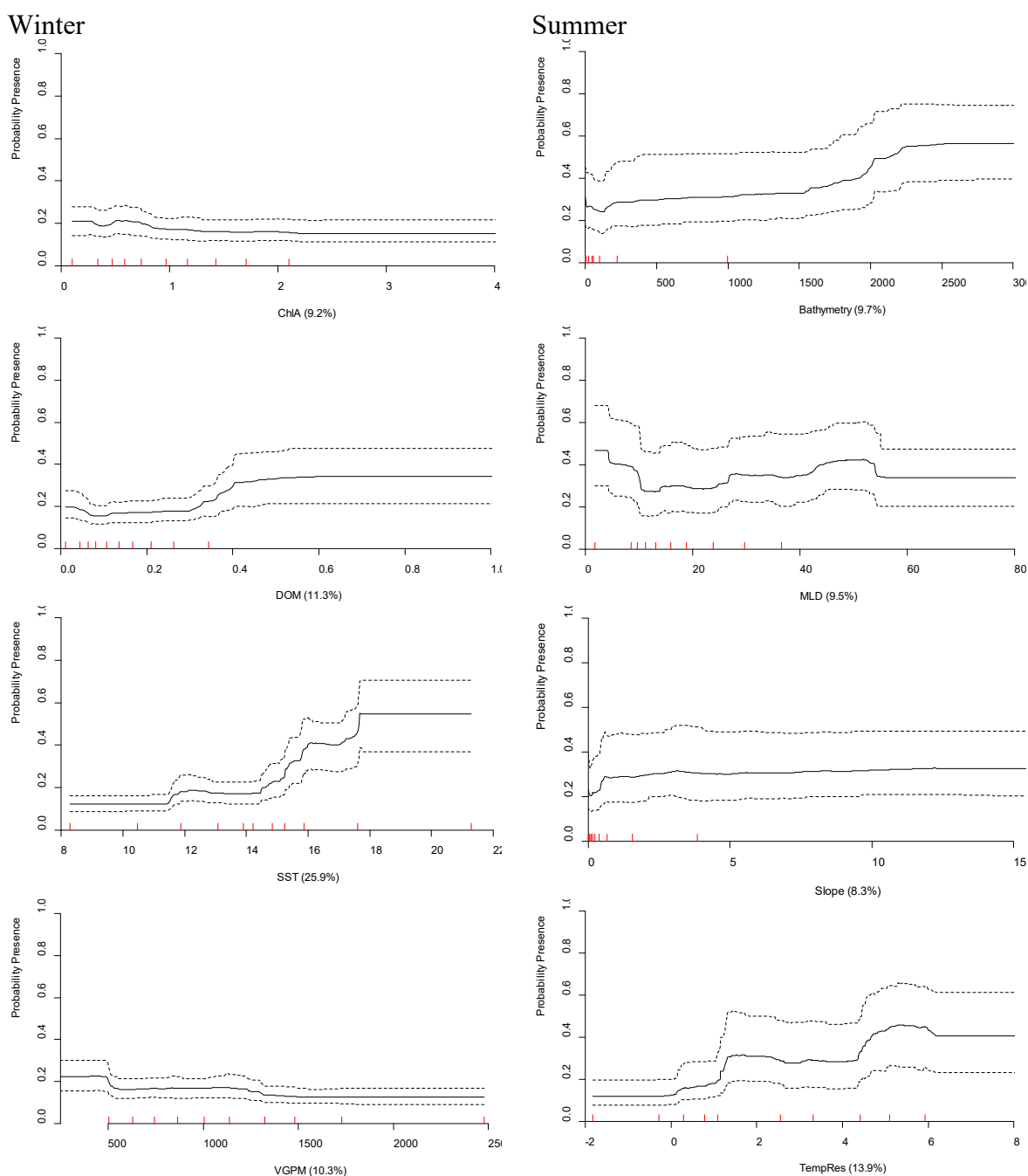


Figure 3-9: Partial dependence plots showing the relationships between predictor variables and predicted presence of bottlenose dolphin (*Tursiops truncatus*) modelled using bootstrapped BRTs for winter (left) and summer (right), showing the four most influential environmental predictors. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown as red tick marks on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Predicted distributions differed between winter and summer bottlenose dolphin presence/relative absence models (Figure 3-10). Both seasonal predictions had stronger discrimination in the predicted probability of presence than for those observed in average year-round probability of presence predictions (i.e., higher predicted probabilities and better-defined core species distributions) (Figure 3-10). Predicted winter presence of bottlenose dolphins were highest across inshore and offshore areas in the north of the study area; particularly close to the Kermadec Islands (Figure 3-10). Predicted

summer distributions of bottlenose dolphin were highest across a wider area with the areas of high predicted probability presence occurring further south in the study area compared with winter predictions (Figure 3-10). Uncertainty estimates (CV) for seasonal probability of presence models were moderate for both seasons, although these were higher and more uniform across the study area for winter predictions (Figure 34, Appendix 3).

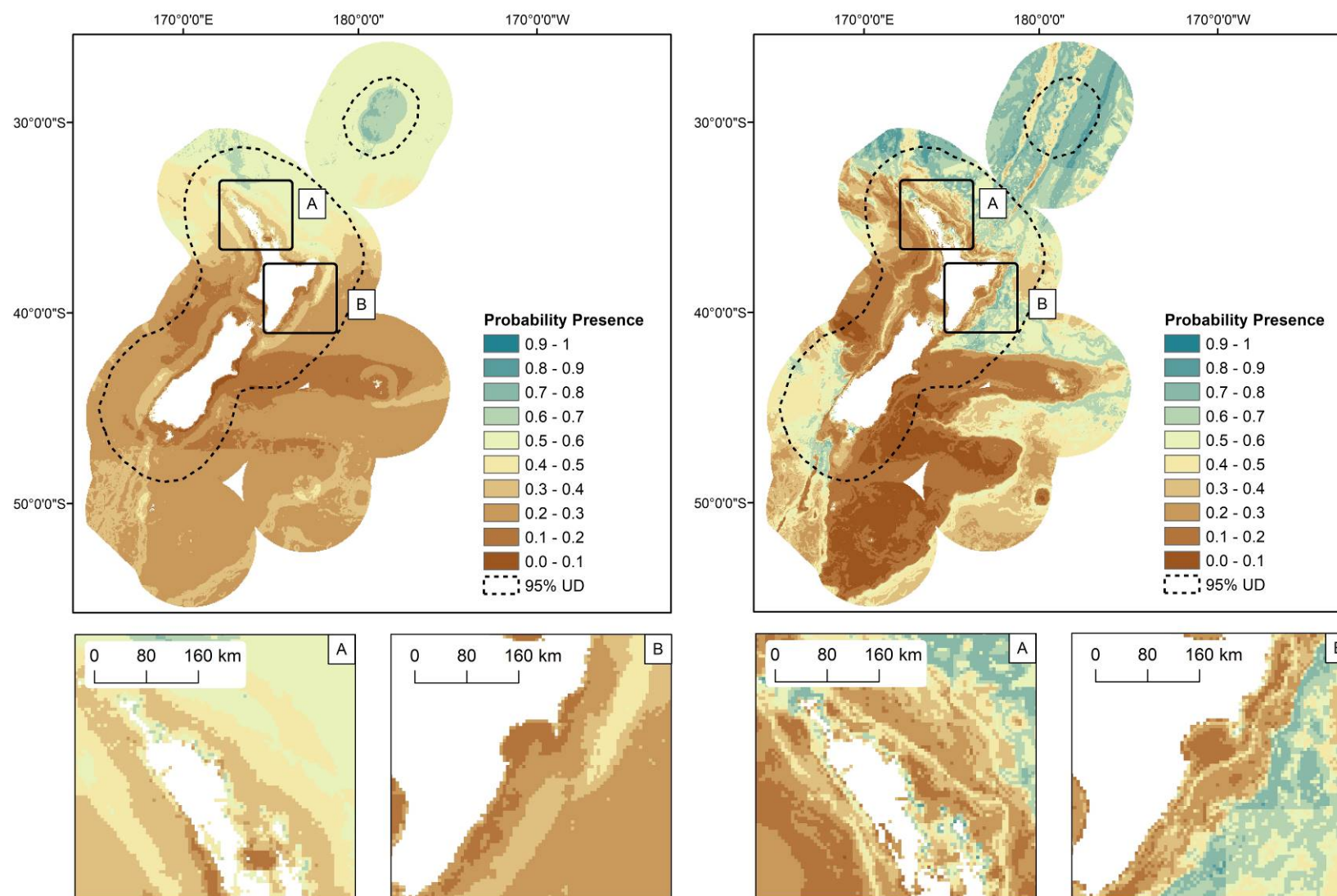


Figure 3-10: The predicted probability of bottlenose dolphin (*Tursiops truncatus*) presence modelled using bootstrapped BRTs fitted with winter (May-Oct, $n = 169$) sightings records (left) and summer (Nov-Apr, $n = 654$) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is defined by the dashed line. Inset maps: A) north of North Island and Hauraki Gulf; B) south east coast of North Island.

3.2.5.2 Common dolphin (*Delphinus delphis*)

3.2.5.2.1 Overview

Qualitative descriptions of common dolphin sample number, distribution, assessment of models fits (including independent validation using NOMAD data) for average year-round presence/relative absence and group size models and a brief description of spatial differences between seasonal distributions are provided in the highlights table (Table 3-10).

Table 3-10: Common dolphin (*Delphinus delphis*): highlights of model fits (including independent validation using NOMAD data) and geographic prediction. Table colours and qualitative categories used here are provided in Table 2.4.

Sample number	Distribution	P/RA			Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	NOMAD Evaluation (AUC)	Model fit (R ²)	Model fit (dev. Exp)	
Very high	Localised – within the first 50 km from shore; primarily North Island	Excellent	Very good	Fair	Very good	Good	No change in distribution, higher probability of presence in summer

3.2.5.2.2 Average year-round presence/relative absence models

Deviance explained and AUC scores for common dolphin average year-round presence/relative absence models were high, consistent between training and evaluation data, and between bootstrap samples (Table 3-11). Model validation with independently collected presence / (true) absence data showed that the model had some predictive power (NOMAD evaluation AUC: 0.68, ‘fair’), although this was substantially lower than for evaluation with non-independent data (Table 3-11).

Table 3-11: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of common dolphin (*Delphinus delphis*).

	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)	NOMAD Evaluation (AUC)
Mean	0.39	0.39	0.90	0.90	0.68
Standard Deviation	0.00	0.01	0.00	0.00	

The most important environmental predictor variables for common dolphin presence/relative absence models were TempRes (42.9%), Turb (21.1%), Dist.Shore (17.0%), and SST (8.9%) (Figure 3-11). Given all other variables kept at their means, there was a strong positive relationship of predicted common dolphin probability presence with TempRes, with increased probability of presence in areas with more than 1.5 °C bottom water temperatures expected for a given. Higher probability of presence was predicted for low turbidity areas and areas 1 – 50 km from shore (see Figure 3-11). A less clear (and weaker) relationship was observed between probability of common dolphin presence and SST, although a small increase in probability of presence was predicted for areas with SST greater than 17 °C.

Overall, there was low uncertainty in predictions, as shown by the 95% prediction intervals (dashed lines in Figure 3-11).

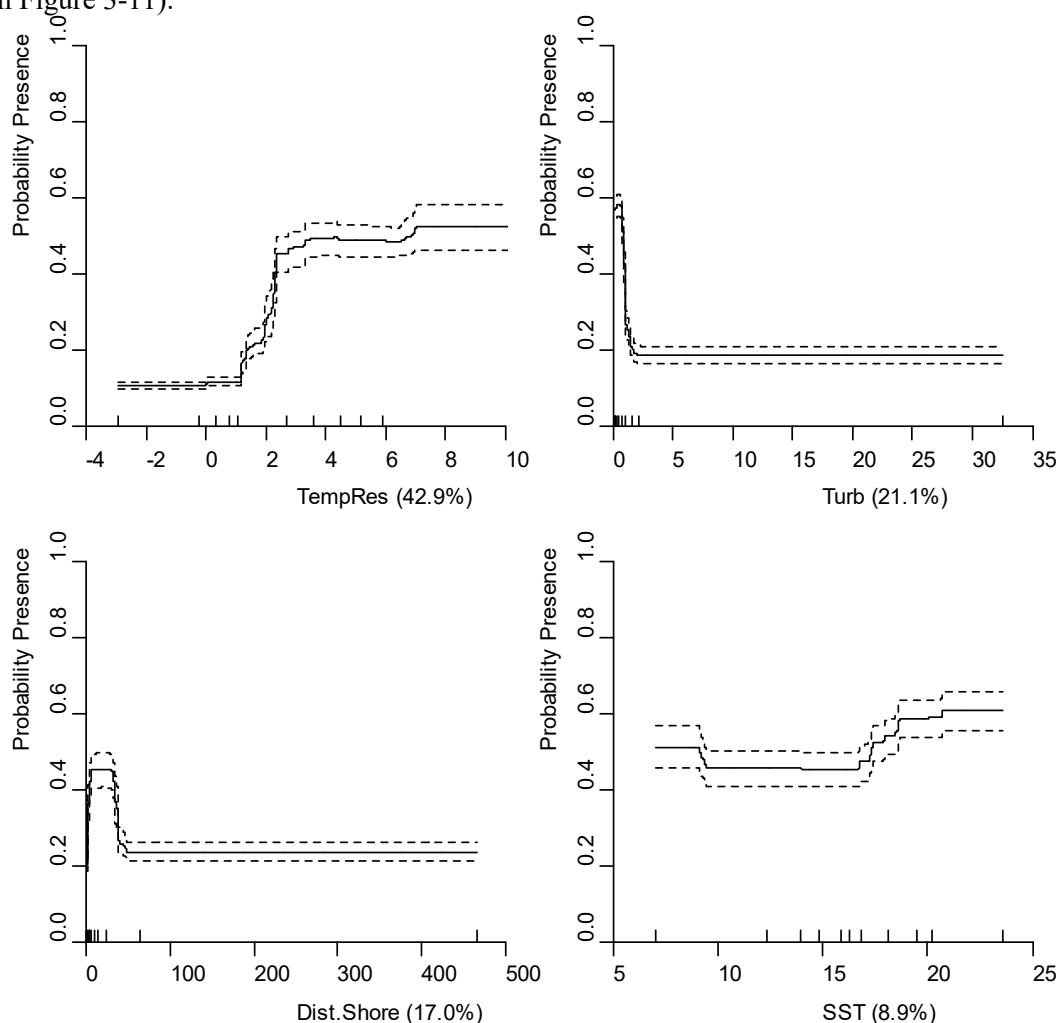


Figure 3-11: Partial dependence plots showing the relationships between predictor variables and probability of common dolphin (*Delphinus delphis*) presence modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown ticks on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Common dolphins are predicted to be mostly within the first 50 km from shore (although not within the first 5 km) and primarily in the seas surrounding the North Island and the north and west coasts of the South Island (blue areas in Figure 3-12). The highest probability of presence was predicted for areas between the North Cape and the Hauraki Gulf (inset A in Figure 3-12) in line with known distributions (Brager and Schneider 1998, Stockin et al. 2009).

Broadly, predicted probability of presence was congruent with the utilisation distribution (UD - Figure 3-12). Predicted probability of presence was also visually concurrent with: presence/relative absence, strandings, and NOMAD presence records (Figure 3-13).

Spatially explicit estimates of uncertainty (measured here as CV) were low across the study area (Figure 35 and 36 in Appendix 3). However, the interpretation of this uncertainty layer should be treated with some caution in areas where few samples exist (see section 2.6.3 for further details). Similarly, predictions of common dolphin with environmental coverage values of 0 and in areas where no presences have been observed should be treated with caution.

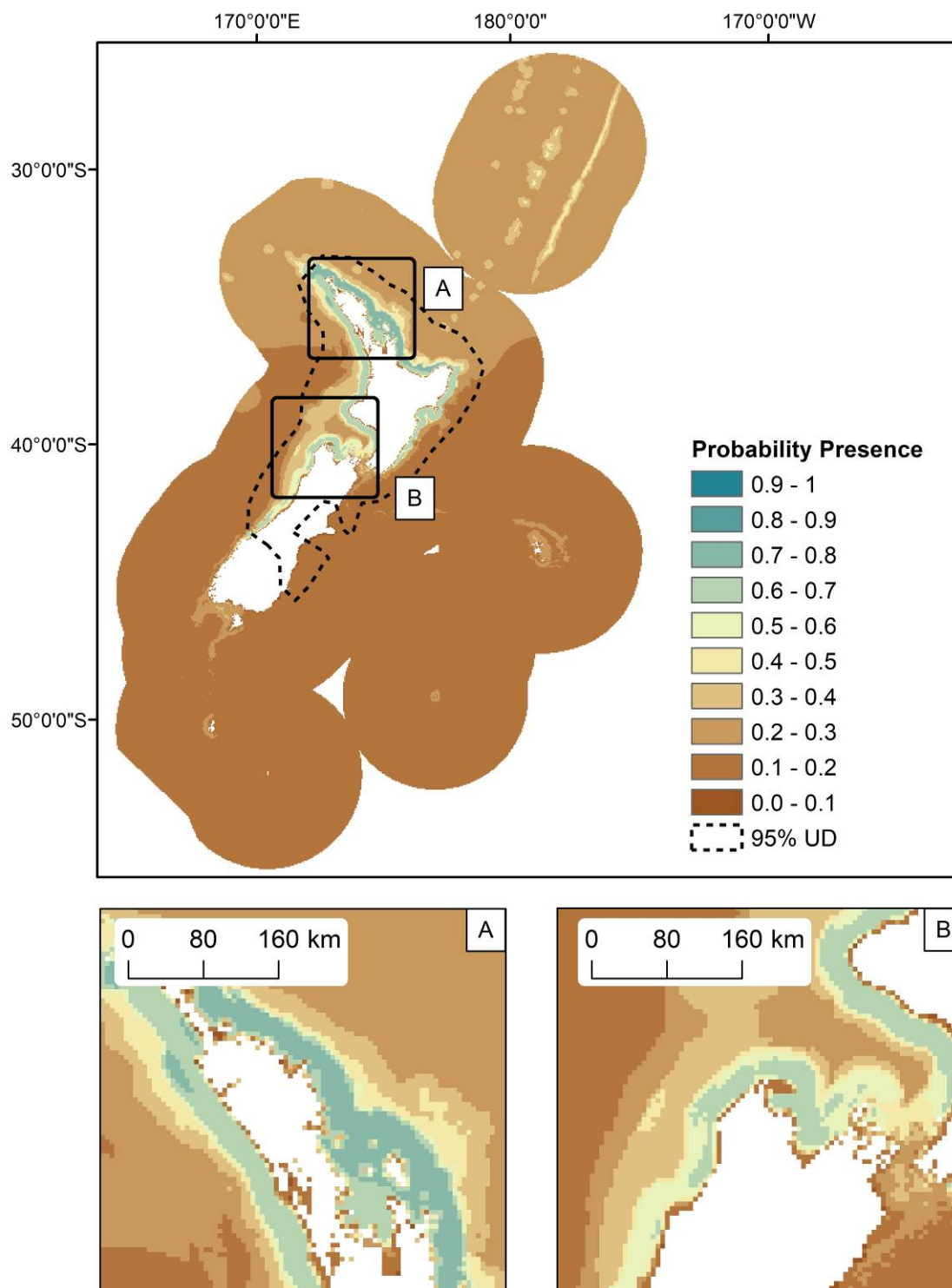


Figure 3-12: The predicted probability of common dolphin (*Delphinus delphis*) presence in the New Zealand EEZ, from bootstrapped BRT models. The predicted 95% utilisation distribution is defined by the dashed line. Inset maps: A) north of North Island; B) Taranaki Bight.

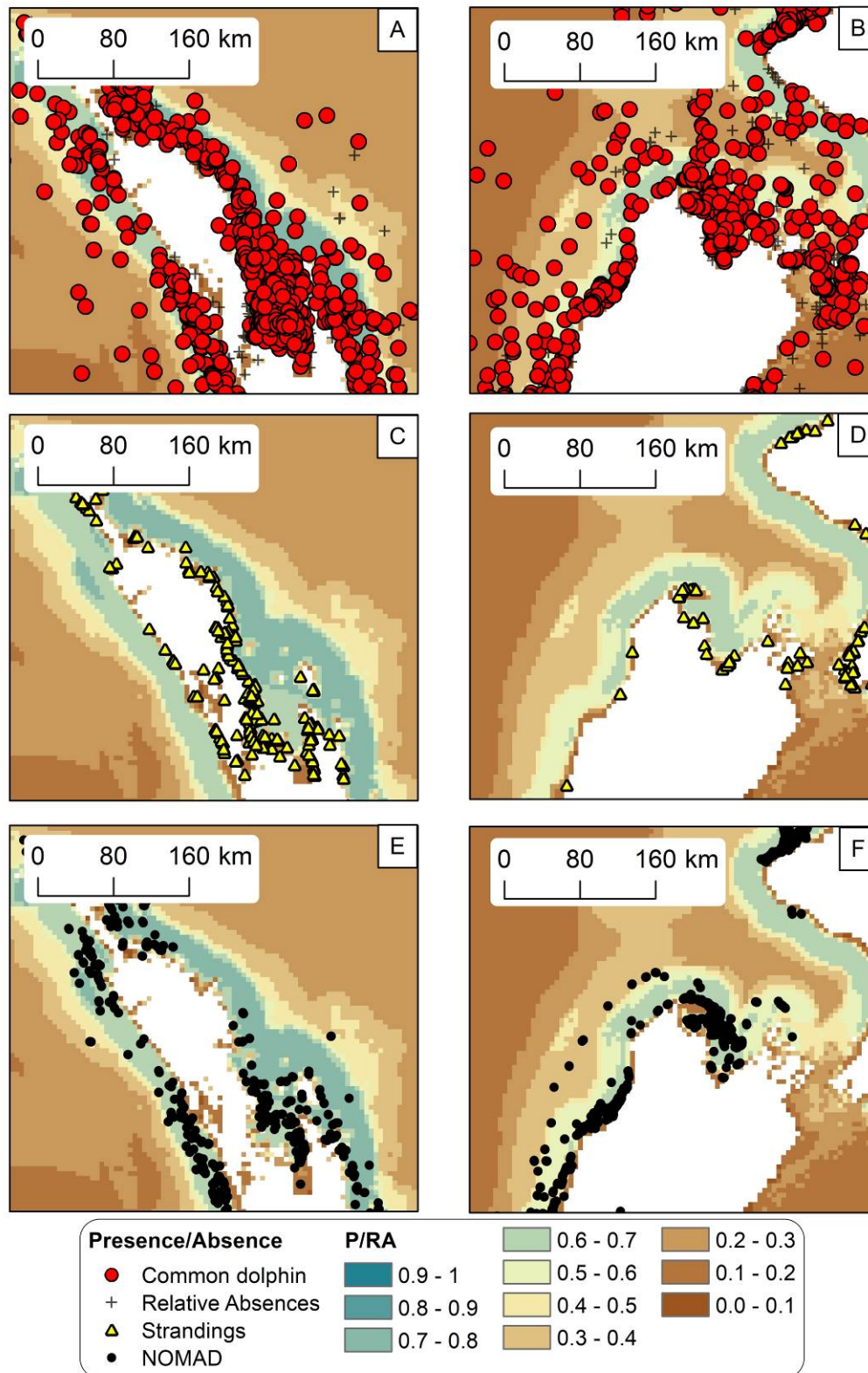


Figure 3-13: Predicted probability presence of common dolphin (*Delphinus delphis*) modelled using bootstrapped BRTs. Predicted probability presence of common dolphin around the north of North Island (left panel) are shown with presence/relative absence (red circles and black crosses respectively) (A), DOC stranding locations (C), and NOMAD sightings (E). Predicted probability presence of common dolphin in the Taranaki Bight (right panel) are shown with presence/relative absences (B), stranding locations (D), and NOMAD sightings (F).

3.2.5.2.3 Average year-round group size models and relative density predictions

Deviance explained for common dolphin average year-round group size models was reasonable, consistent between training and evaluation data, and between bootstrap samples (Table 3-12). Pearson's correlations between predicted and observed species group sizes were relatively high (0.43, Table 3-12). The highest values of predicted group sizes were lower than the observed group sizes (comparison of maximum predicted values in Figure 3-14 and Figure 3-15 to observed values in Figure 3-16).

Table 3-12: Mean model performance measures (deviance explained and Pearson's correlation of predicted vs observed group sizes (R^2)) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of common dolphin (*Delphinus delphis*).

	Deviance explained (training data)	Deviance explained (evaluation data)	Pearson's correlation of predicted vs observed species group sizes (R^2)
Mean	0.26	0.25	0.43
Standard Deviation	0.03	0.05	0.08

The most important environmental predictor variables in average year-round common dolphin group size models differed to those of the average year-round presence/relative absence model. In the average year-round group size models, the most important variables were MLD (15.3%), DOM (15.1%), SST (10.7%), and TC (9.7%) (Figure 3-14). Group sizes of common dolphin were predicted to increase with increasing SST (highest group sizes predicted in sea surface temperatures above 20°C), increasing MLD (highest predicted group sizes at mixed layer depths greater than 100 m), and increasing tidal current speeds (highest predicted group sizes at tidal current speeds greater than 1.5 ms⁻¹) (Figure 3-14). A less clear (and weaker) relationship was observed between common dolphin group sizes and DOM. Overall, there was higher uncertainty in average year-round group size predictions (95% prediction intervals — dashed lines in Figure 3-14) than for average year-round common dolphin predicted presence (see Figure 3-11).

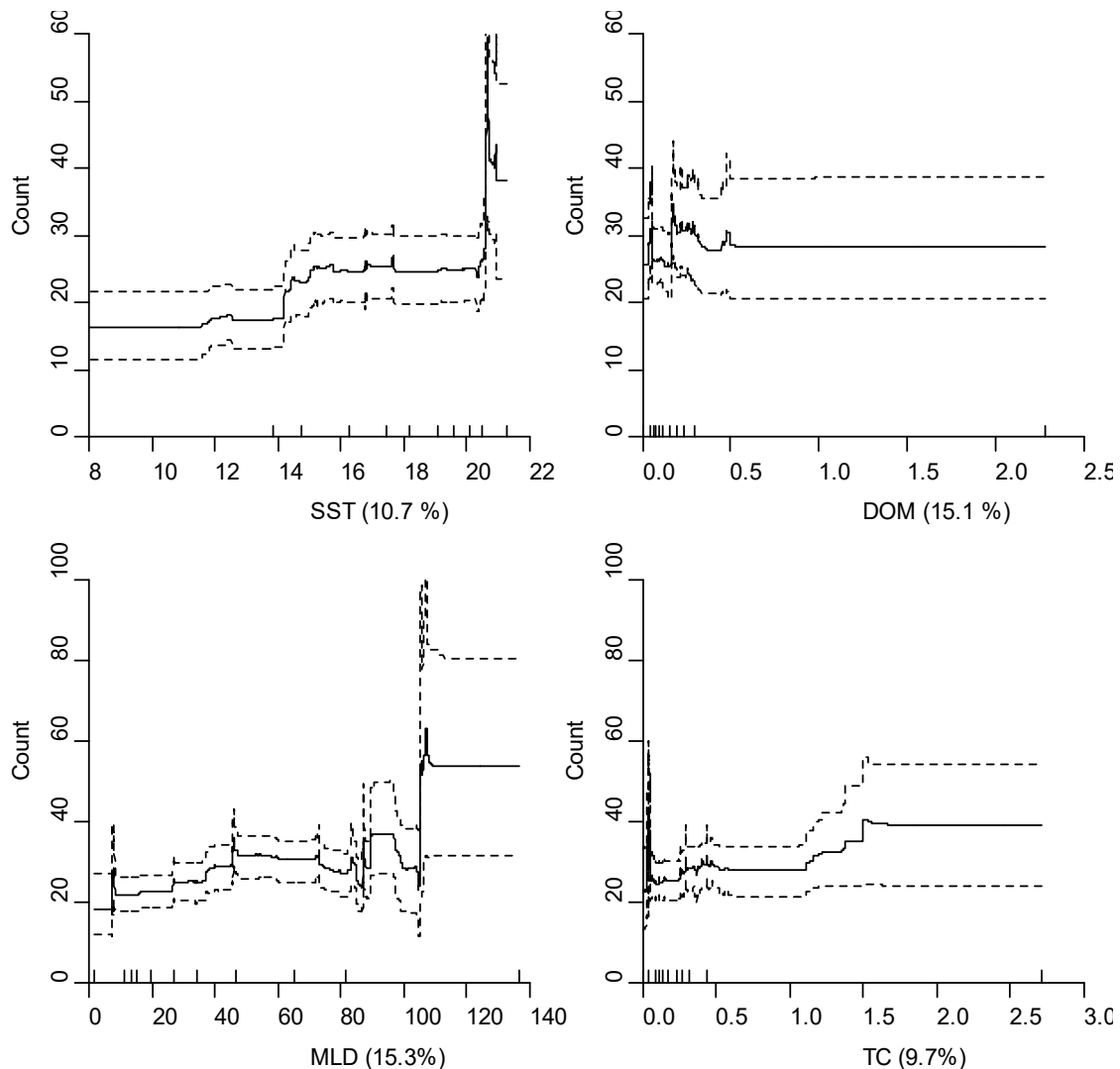


Figure 3-14: Partial dependence plots showing the relationships between predictor variables and common dolphin (*Delphinus delphis*) group sizes, modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Distribution of predicted average year-round relative density of common dolphin (the resulting prediction from the two-step hurdle method) largely followed geographic patterns described for average year-round probability of presence. The highest predicted relative densities were located in the Hauraki Gulf, Bay of Plenty, and off the north and east coasts of the North Island (Figure 3-15). The maximum predicted relative density (260 predicted individuals, Figure 3-15) was substantially lower than observed maximum of circa group size of 2000 individuals. However, broadly speaking, relative density was in line with observed group size (Figure 3-16). Uncertainty (95% prediction interval – Figure 37, Appendix 3) was relatively high for predicted common dolphin relative density in areas outside the core species distribution (i.e., areas with probability of presence values greater than 0.8 in Figure 3-12).

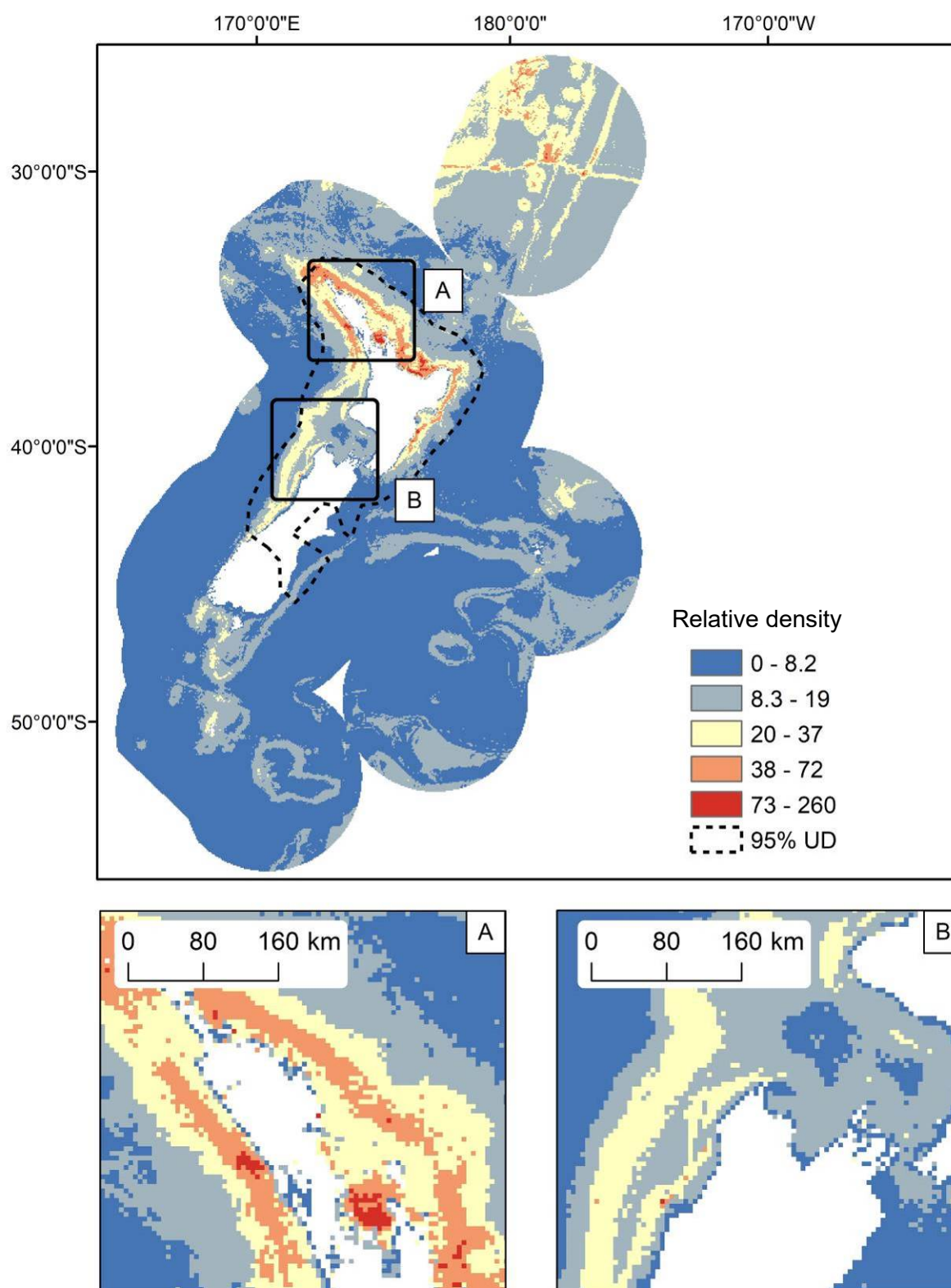


Figure 3-15: Predicted relative density of common dolphin (*Delphinus delphis*) in the New Zealand EEZ, from bootstrapped BRT models. The predicted 95% utilisation distribution is defined by the dashed line. Inset maps: A) north of North Island; B) Taranaki Bight.

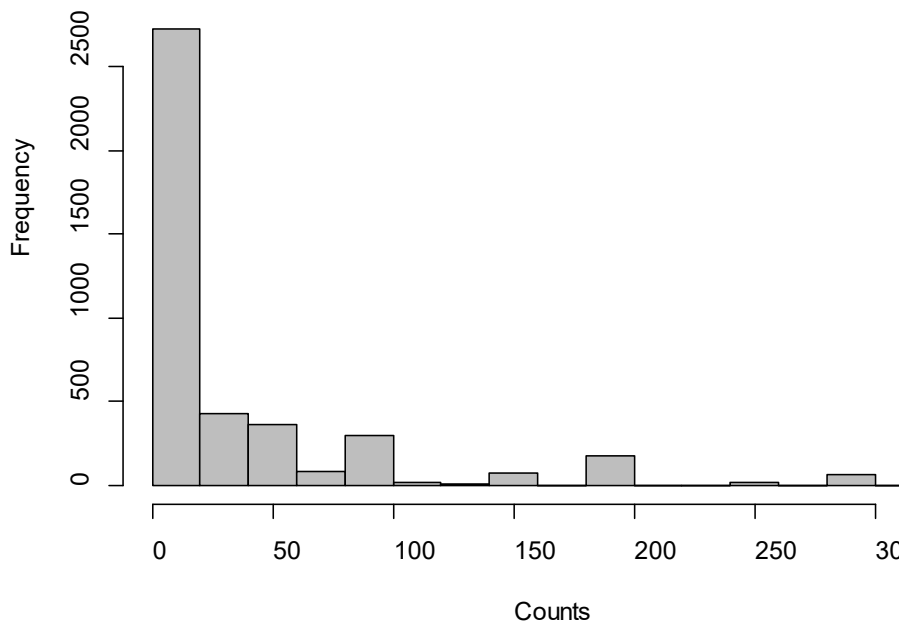


Figure 3-16: Histogram of common dolphin (*Delphinus delphis*) group sizes recorded in the New Zealand EEZ. Note: for clarity, the x-axis was truncated at 300; the maximum count was close to 2000.

3.2.5.2.4 Seasonal presence/relative absence models

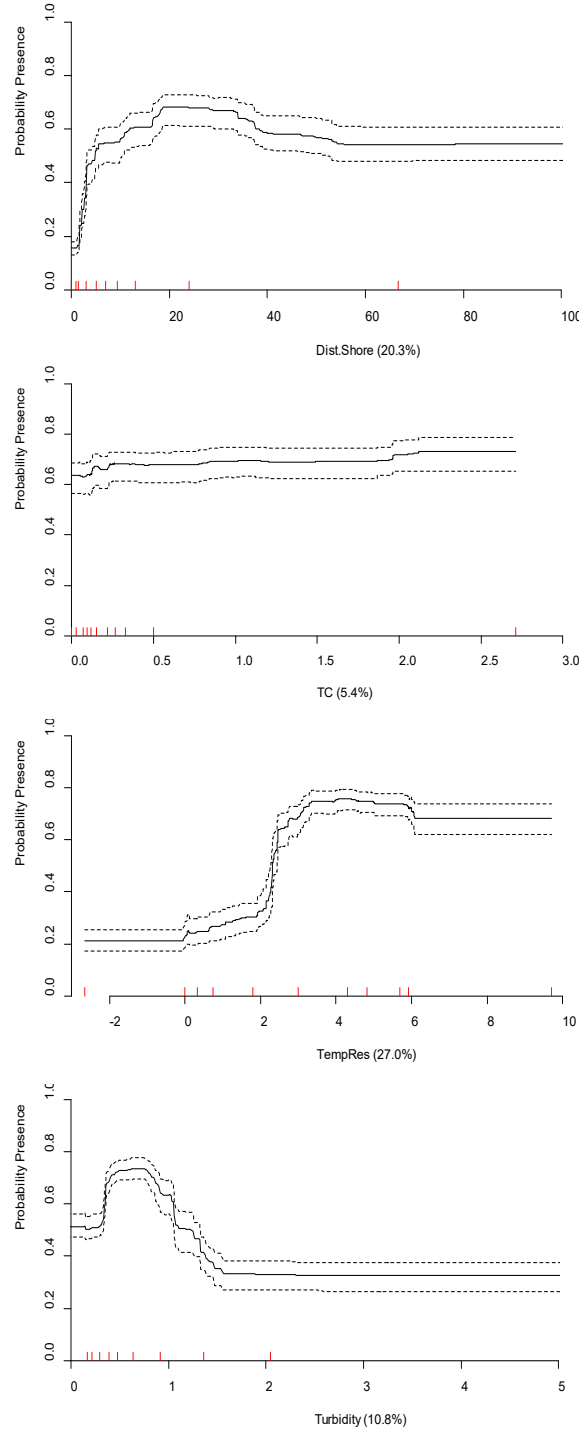
Model goodness-of-fit metrics differed for seasonal presence/relative absence models to those of average year-round presence/relative absence models (Table 3-13). Although AUC values were similar between the average year-round (0.9) and seasonal models (winter: 0.89, summer: 0.93), the explained deviance was considerably higher for summer models (50%) than for both the average year-round (39%) and winter models (33%) (Table 3-13). As with the average year-round presence/relative absence models, seasonal model fit metrics differed little between those derived from training data and evaluation data and mean estimates did not vary greatly between bootstrap samples. The difference in model fits between seasonal models may be explained by differences in sample number, i.e., winter had fewer species presence records ($n = 1283$), compared with summer ($n = 3128$) models.

Table 3-13: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) from winter (May-Oct) and summer (Nov-Apr) sightings of common dolphin (*Delphinus delphis*) records.

Season	Metric	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)
Winter	Mean	0.33	0.33	0.87	0.89
	Standard Deviation	0.02	0.06	0.00	0.01
Summer	Mean	0.49	0.50	0.93	0.93
	Standard Deviation	0.01	0.03	0.00	0.01

The same environmental predictors were important in common dolphin seasonal models as in the average year-round presence/relative absence models (albeit with differing percentage contributions). Similar patterns in predicted probability presence along environmental gradients were observed between seasons and average year-round presence/relative absence models (Figure 3-17).

Winter



Summer

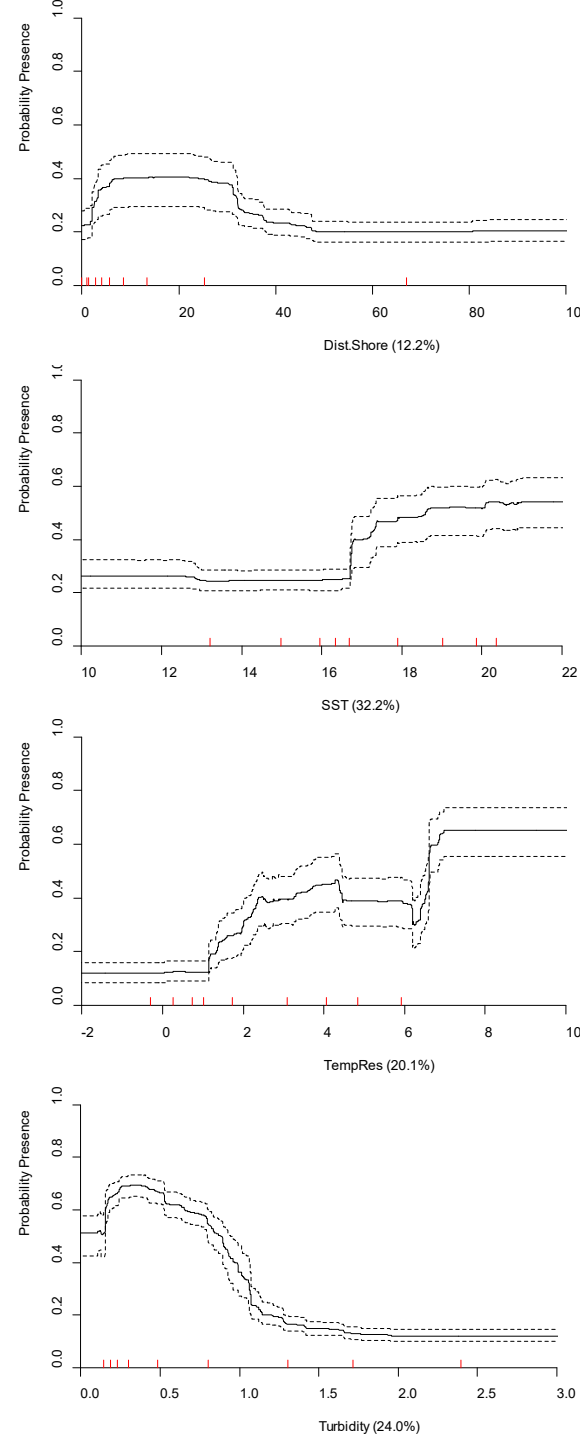


Figure 3-17: Partial dependence plots showing the relationships between predictor variables and predicted presence of common dolphin (*Delphinus delphis*) modelled using bootstrapped BRTs for winter (left) and summer (right). The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Similar predicted distributions were observed for seasonal probability of common dolphin presence models as were observed for average year-round probability of presence models (Figure 3-18). There was little difference in distribution patterns between winter and summer, although summer probability of presence predictions were higher in core species distribution areas (Figure 3-18). As with the average year-round models for common dolphin, uncertainty estimates (CV) for seasonal probability of presence models were low for both seasons (Figure 38, Appendix 3).

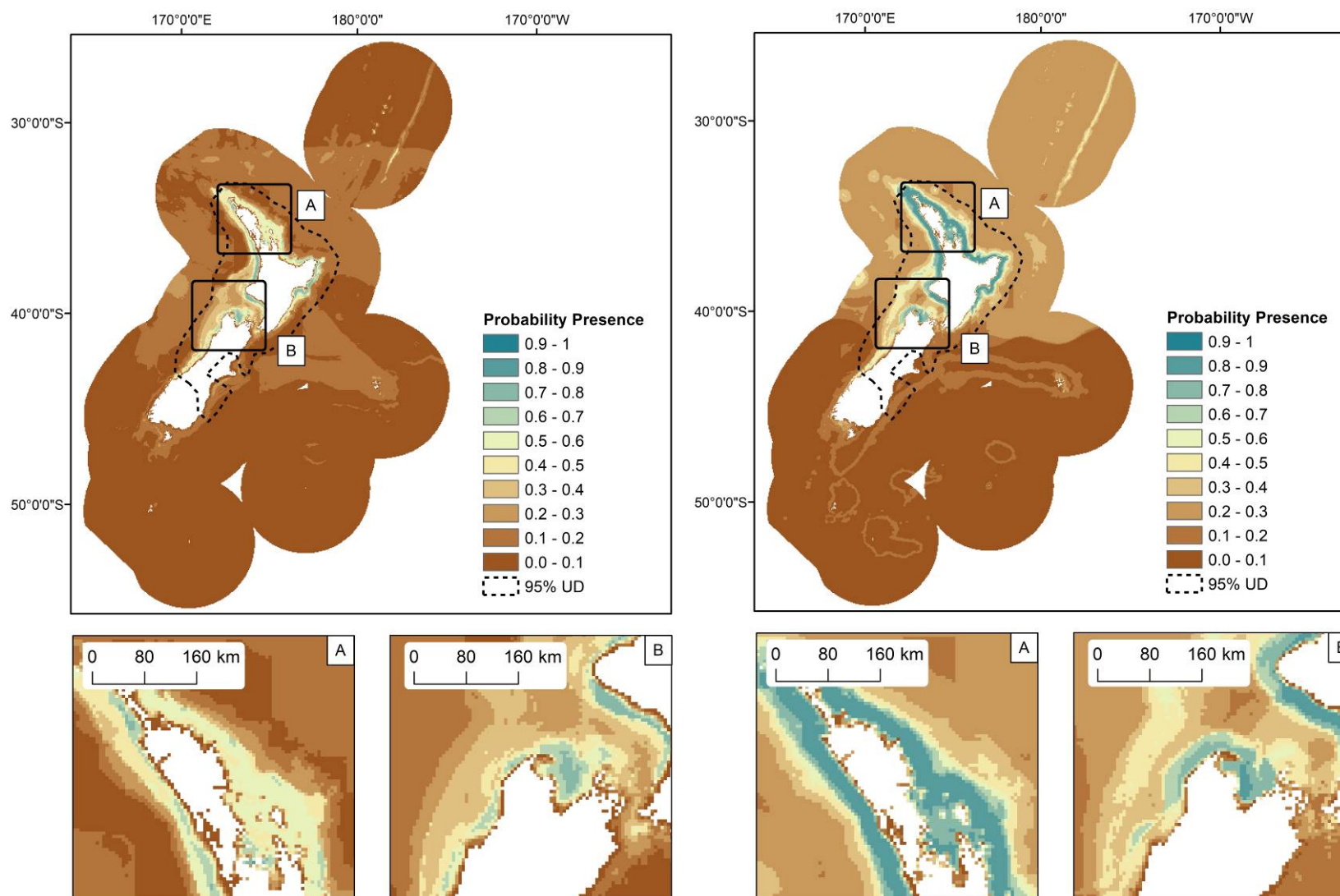


Figure 3-18: The predicted probability of common dolphin (*Delphinus delphis*) presence modelled using bootstrapped BRTs fitted with winter (May–Oct, $n = 1283$) sightings records (left) and summer (Nov–Apr, $n = 3128$) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island; B) Taranaki Bight.

3.2.5.3 Hector's dolphin (*Cephalorhynchus hectori hectori*)

3.2.5.3.1 Overview

Qualitative descriptions of Hector's dolphin sample number, distribution, assessment of model fits (including independent validation using NOMAD data) for average year-round presence/relative absence and group size models, and a brief description of spatial differences between seasonal distributions are provided in the highlights table (Table 3-14).

Table 3-14: Hector's dolphin (*Cephalorhynchus hectori hectori*): highlights of model fits (including independent validation using NOMAD data) and geographic prediction. Table colours and qualitative categories used here are provided in Table 2.4.

Sample number	Distribution	P/RA			Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	NOMAD Evaluation (AUC)	Model fit (R ²)	Model fit (dev. Exp)	
Very high	Localised – inshore South Island, particularly on east coast.	Excellent	Excellent	Excellent	Very good	Good	Similar distribution predicted across seasons

3.2.5.3.2 Average year-round presence/relative absence models

Deviance explained and AUC scores for Hector's dolphin average year-round presence/relative absence models were very high and were consistent between training and evaluation data, and between bootstrap samples (Table 3-15). Model validation with independently collected presence / (true) absence data showed that the model had high predictive power (NOMAD evaluation AUC: 0.96, 'excellent') providing strong evidence that models were performing well (Table 3-15).

Table 3-15 Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of Hector's dolphin (*Cephalorhynchus hectori hectori*).

	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)	NOMAD Evaluation (AUC)
Mean	0.83	0.83	0.99	0.99	0.96
Standard Deviation	0.01	0.01	0.00	0.00	NA

The most important environmental predictor variables for Hector's dolphin average year-round presence/relative absence models were MLD (29.5%), followed closely by Turb (25.5%) and TempRes (24.8%) and finally SST (10.9%) (Figure 3-19). There was a strong positive relationship of predicted Hector's dolphin presence with turbidity (Turb), with increased probability of presence for those areas with turbidity greater than 5 NTU. Higher probabilities of presences were predicted for MLD values between 0 – 40 m and for areas with lower than 4 °C bottom water temperatures expected for a given depth (TempRes) (Figure 3-19). A less clear (and weaker) relationship was observed between probability presence of Hector's dolphin and SST, although a small increase in predicted probability presence for areas with SST lower than 17 °C was observed. Overall, there was low uncertainty in predictions as shown by the 95% prediction intervals; although higher uncertainty was observed for high turbidity areas (dashed lines in Figure 3-19).

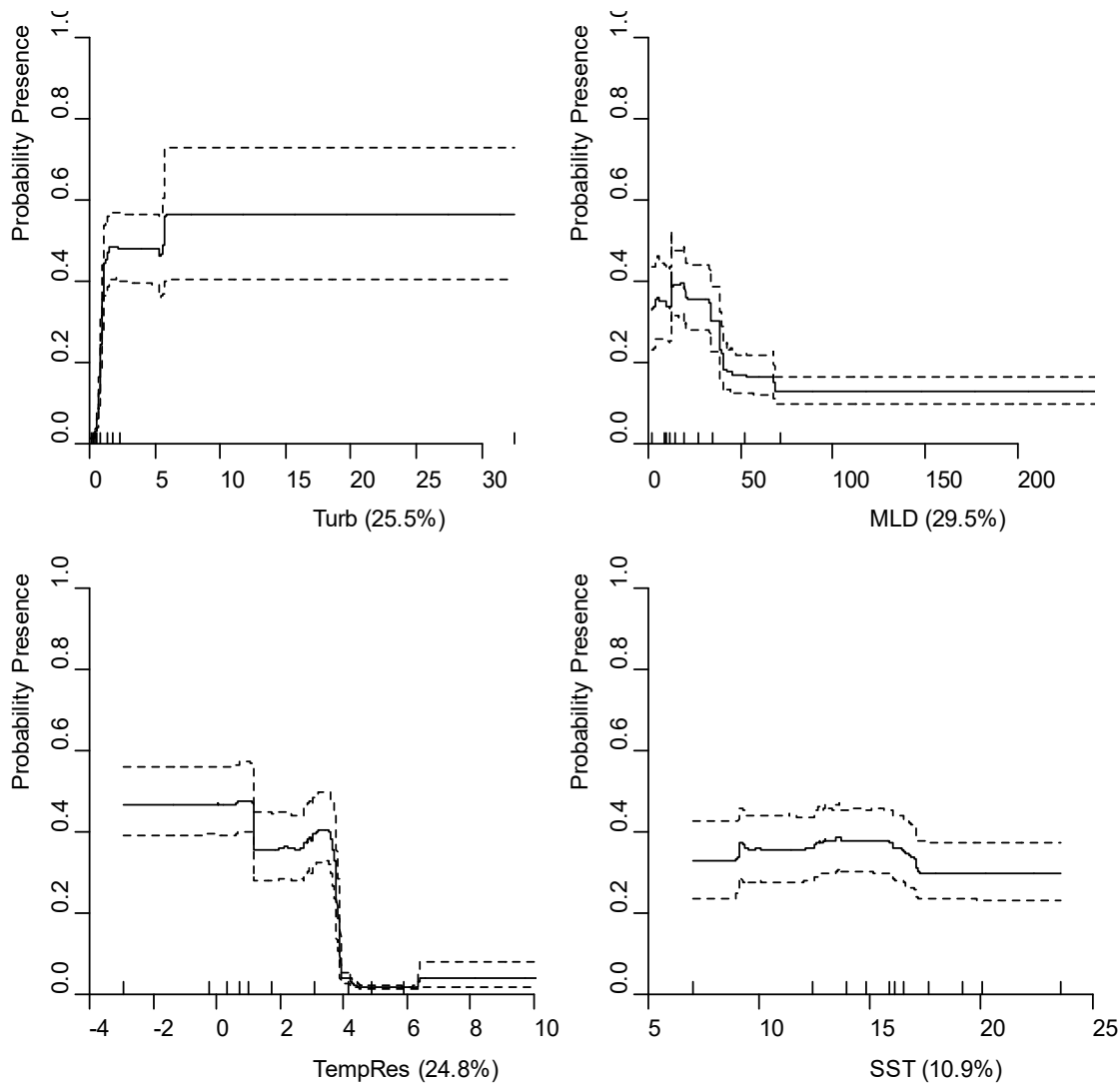


Figure 3-19: Partial dependence plots showing the relationships between predictor variables and probability presence of Hector's dolphin (*Cephalorhynchus hectori hectori*) modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Hector's dolphins were predicted to be present very close to shore (no further than 50 km) along the coastline of the South Island (blue areas in Figure 3-20). The highest probability of presence was located off the east coast of South Island (inset B in Figure 3-20) in agreement with expert knowledge of spatial distributions.

Predicted probability of presence was congruent with the utilisation distribution (UD, see Figure 3-20). Predicted probability of presence was also visually congruent with: presence/relative absence, strandings, and NOMAD presence location records (Figure 3-21).

Spatially explicit estimates of uncertainty (CV) were low across the study area (Figures 39 and 40, Appendix 3). Uncertainty estimates increased slightly for those areas immediately outside the core species distribution (slightly further from the coast) (see inset B in Figures 40, Appendix 3). Hector's dolphin records and model outputs fall within areas of the environmental space that are considered well

sampled (Figure 3-2), and thus uncertainty layers are likely to provide an accurate representation of spatial uncertainty.

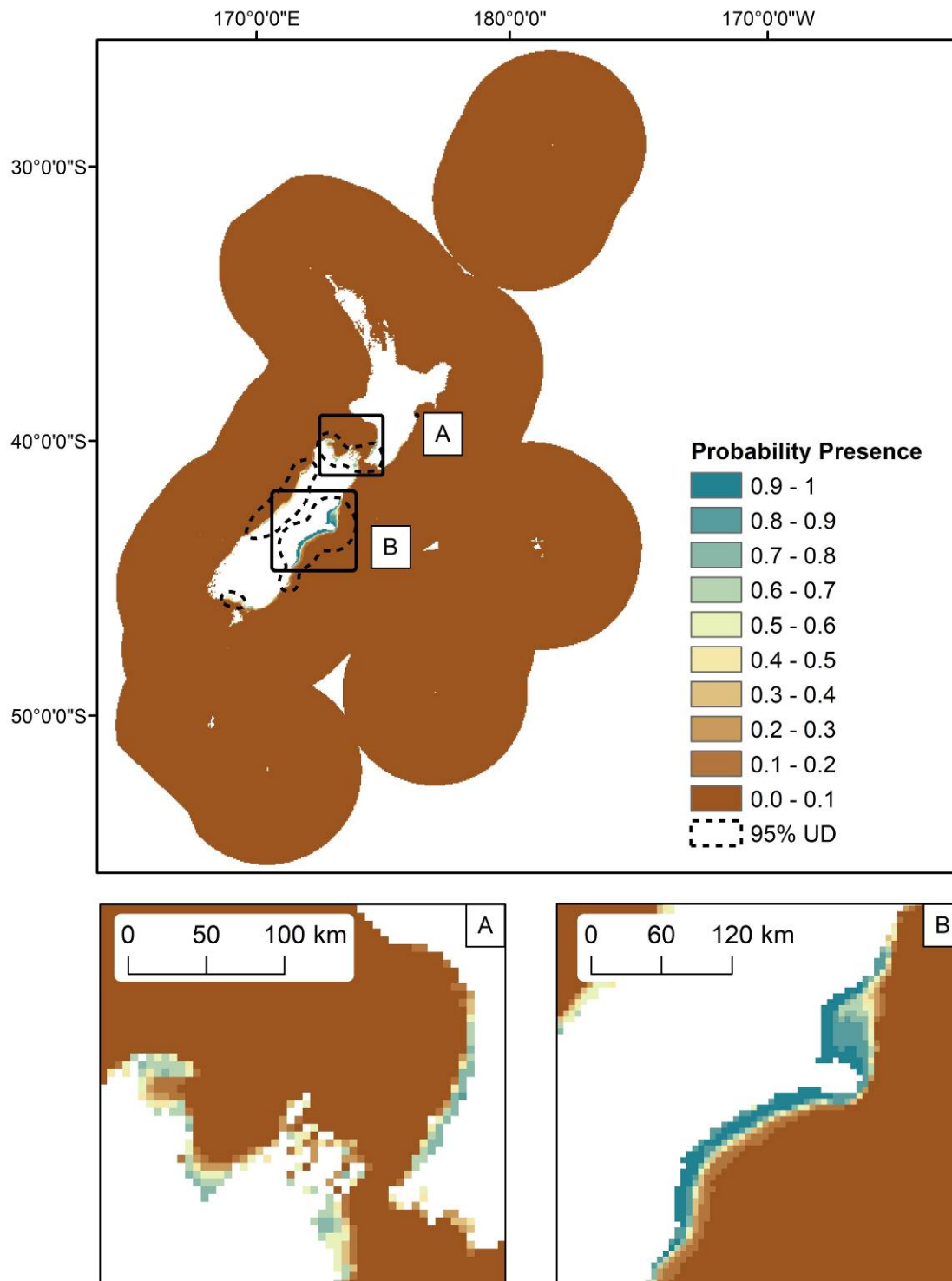


Figure 3-20: The predicted probability Hector's dolphin (*Cephalorhynchus hectori hectori*) presence in the New Zealand EEZ, from bootstrapped BRT models. The predicted 95% utilisation distribution (UD) is defined by the dashed line. Inset maps: A) Taranaki Bight; B) east coast of South Island.

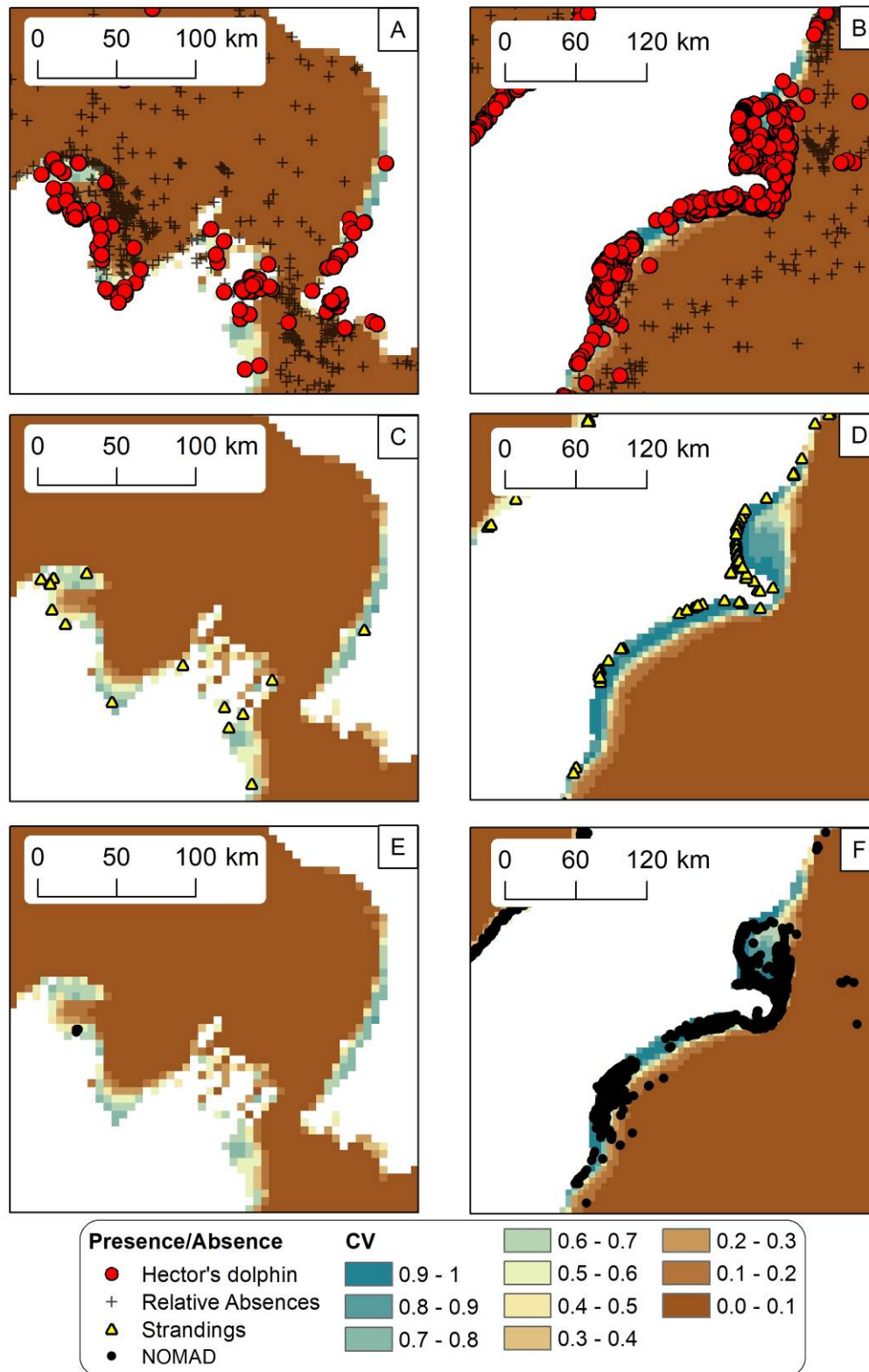


Figure 3-21: Predicted probability of Hector's dolphin (*Cephalorhynchus hectori hectori*) presence modelled using bootstrapped BRTs. Predicted probability of Hector's dolphin presence in the Taranaki Bight is shown (left panel) with presence/relative absence (red circles and black crosses respectively) (A), DOC stranding locations (C), and NOMAD sightings (E). Predicted probability presence of Hector's dolphin off the east coast of South Island is shown (right panel) with presence/relative absence (B), stranding locations (D), and NOMAD sightings (F).

3.2.5.3.3 Average year-round group size models and relative density predictions

Deviance explained for Hector's dolphin average year-round group size models was consistent between training and evaluation data, and between bootstrap samples (Table 3-16). Pearson's correlations between predicted and observed species group sizes were relatively high (0.42, Table 3-16). As for common dolphin, the highest values of predicted group sizes were lower than the observed group sizes (comparison of maximum predicted values shown in Figure 3-22 and Figure 3-23 with observed values in Figure 3-24).

Table 3-16: Mean model performance measures (deviance explained and Pearson's correlation of predicted vs observed group sizes (R^2)) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of Hector's dolphin (*Cephalorhynchus hectori hectori*).

	Deviance explained (training data)	Deviance explained (evaluation data)	Pearson's correlation of predicted vs observed species group sizes (R^2)
Mean	0.26	0.26	0.42
Standard Deviation	0.04	0.06	0.12

The most important environmental predictor variables in average year-round Hector's dolphin group size models differed from those of the average year-round presence/relative absence model. In the average year-round group size models, the most important variables were DOM (36.4%), Bathy (14.8%), TC (13.5%), and SST (8.0%) (Figure 3-22). The only clear trend in Hector's dolphin predicted group sizes was for DOM (higher group sizes predicted for values of DOM greater than 0.7) albeit with high uncertainty (Figure 3-22). All other trends in Hector's dolphin group sizes with environmental predictors were less clear (and weaker) (Figure 3-22).

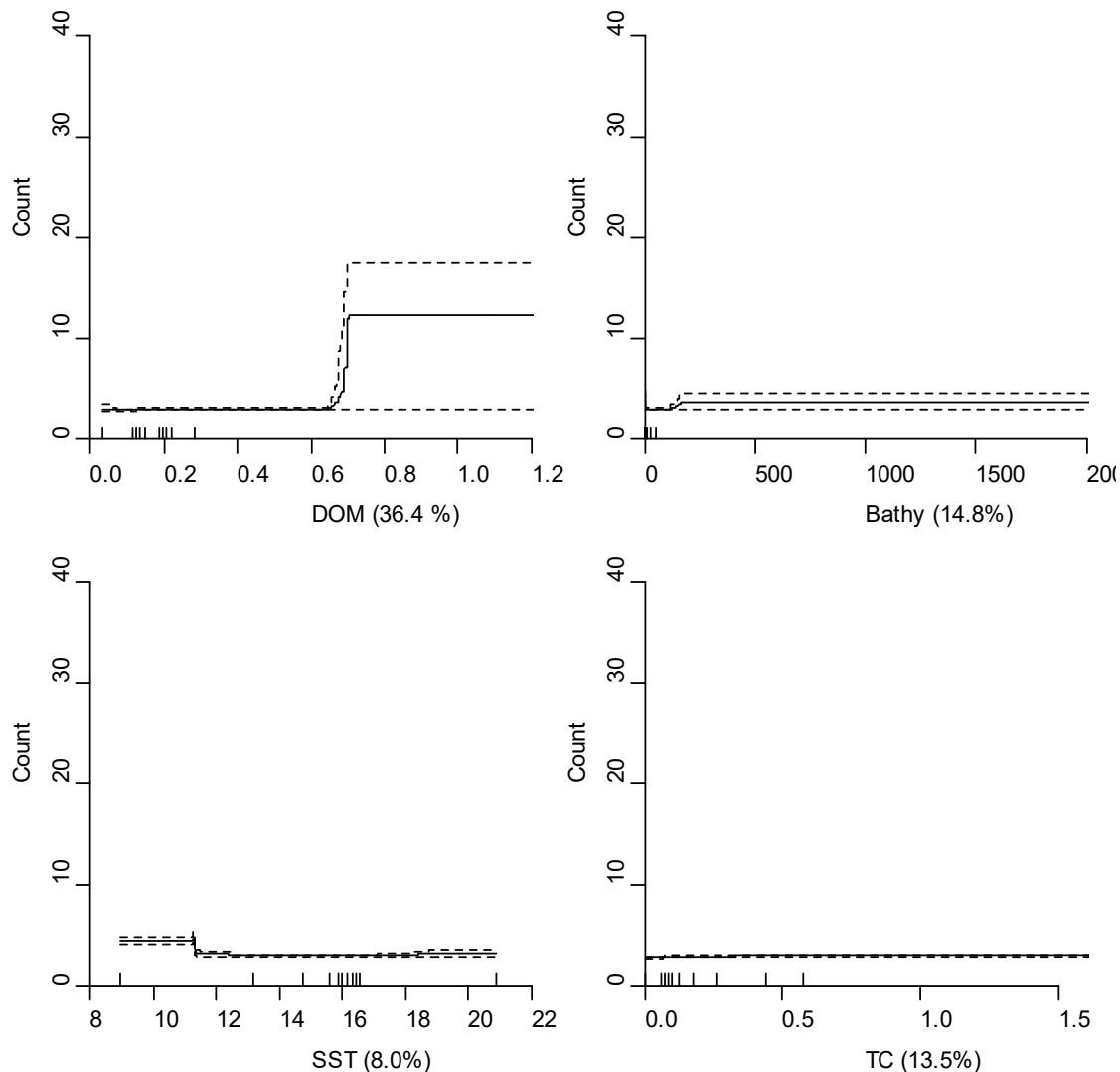


Figure 3-22: Partial dependence plots showing the relationships between predictor variables and Hector's dolphin (*Cephalorhynchus hectori hectori*) group sizes, modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Distribution of predicted average year-round Hector's dolphin relative density (the resulting prediction from the two-step hurdle method) tightly followed geographic patterns described for average year-round probability of presence (Figure 3-23). For the majority of areas, Hector's dolphin relative densities were predicted to be low (ranging between 1 and 5, Figure 3-23). The highest predicted relative densities (5–9) were located in very small areas in the north of Tasman Bay (inset A, Figure 3-23). Predicted relative densities were broadly in line with observed group sizes (comparison of relative density from Figure 3-23 with observed group sizes in Figure 3-24). Uncertainty (95% prediction interval – Figure 41, Appendix 3) was low across the study area.

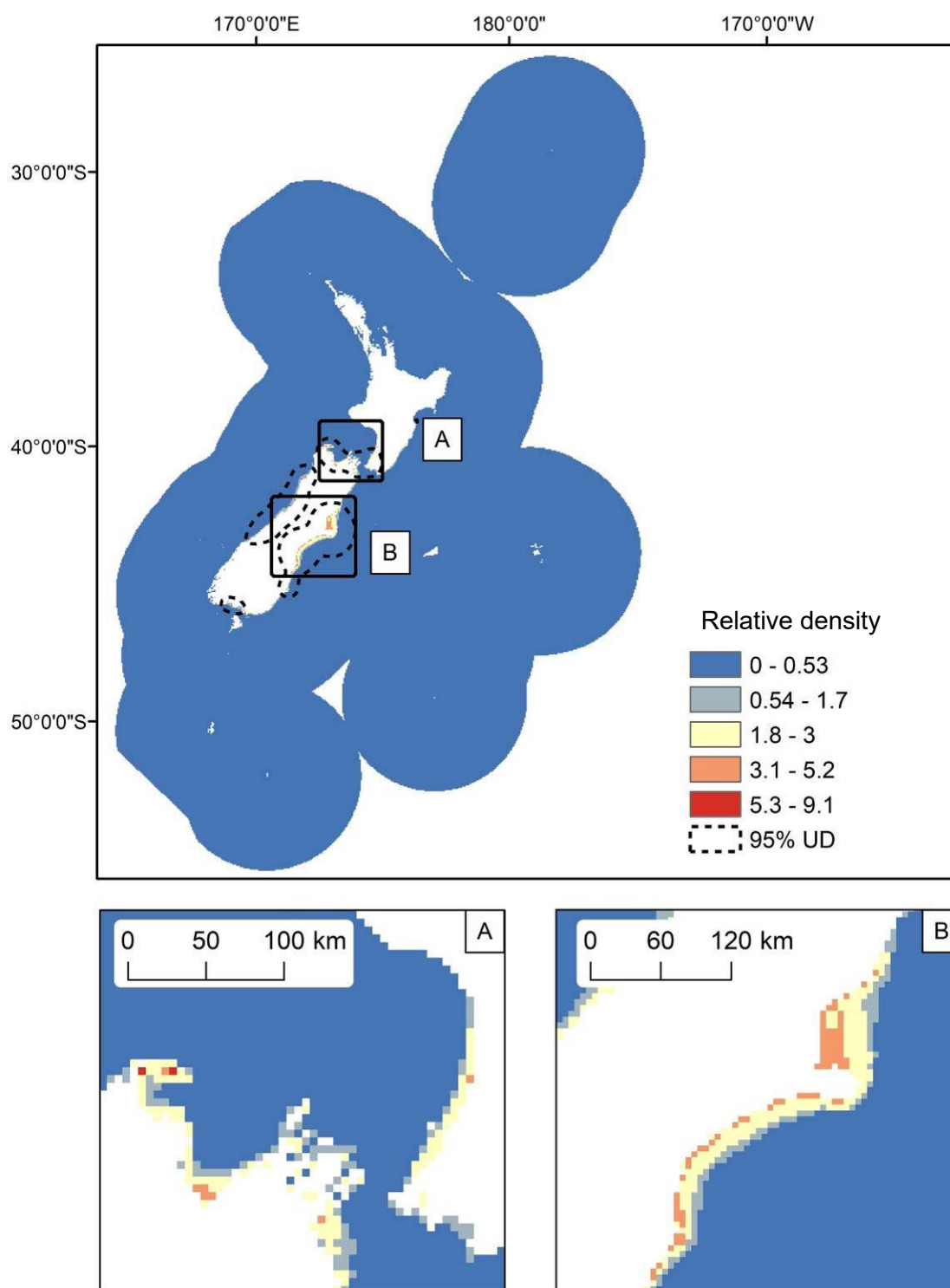


Figure 3-23: Predicted relative density of Hector's dolphin (*Cephalorhynchus hectori hectori*) in the New Zealand EEZ, from bootstrapped BRT models. The predicted 95% utilisation distribution is defined by the dashed line. Inset maps: A) Taranaki Bight; B) east coast of South Island.

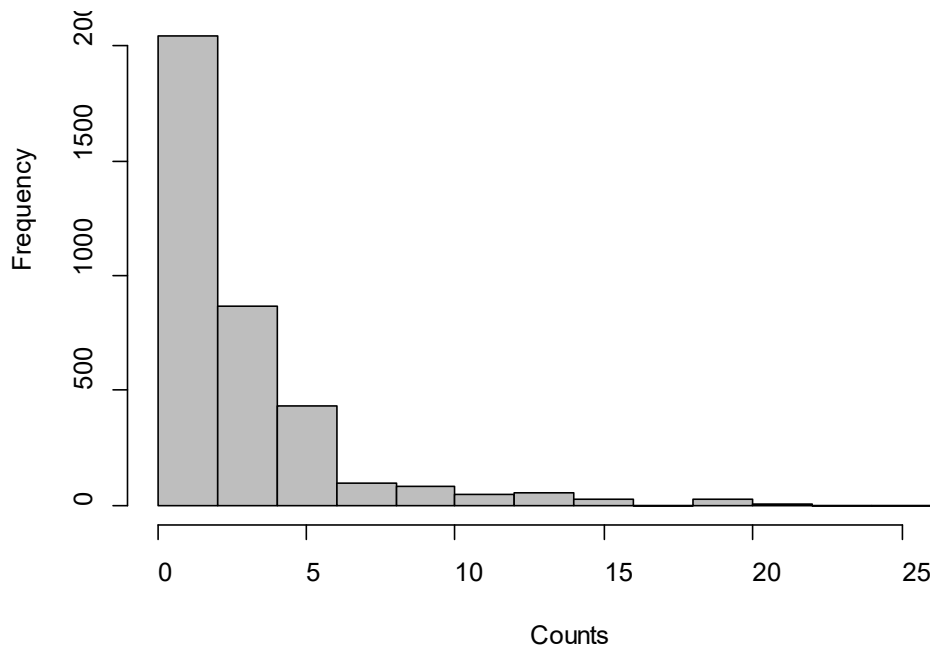


Figure 3-24: Histogram of Hector's dolphin (*Cephalorhynchus hectori hectori*) group sizes recorded in the New Zealand EEZ.

3.2.5.3.4 Seasonal presence/relative absence models

Goodness-of-fit metrics for winter presence/relative absence models were lower (explained deviance: 0.61; AUC: 0.96) than summer (explained deviance: 0.86; AUC: 0.99) and average year-round presence/relative absence models (explained deviance: 0.83; AUC: 0.99) which were similar (Table 3-17). As with the average year-round presence/relative absence models, seasonal model fit metrics differed little between those derived from training data and evaluation data, and mean estimates did not vary greatly between bootstrap samples (Table 3-17). The difference in model fits between seasonal models may be explained by the large difference in sample number, i.e., winter had fewer species presence records ($n = 294$) compared with summer ($n = 3394$) models. Regardless, winter presence/relative absence models still displayed excellent predictive power ($AUC > 0.9$).

Table 3-17: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) from winter (May-Oct) and summer (Nov-Apr) sightings of Hector's dolphin (*Cephalorhynchus hectori hectori*) records.

Season	Metric	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)
Winter	Mean	0.62	0.61	0.96	0.95
	Standard Deviation	0.04	0.09	0.01	0.02
Summer	Mean	0.85	0.86	0.99	0.99
	Standard Deviation	0.00	0.01	0.00	0.00

The same environmental predictors were important in Hector's dolphin seasonal models as were in the average year-round presence/relative absence models (albeit with differing percentage contributions). Similar patterns in predicted probability presence along environmental gradients were observed between seasons and average year-round presence/relative absence models, though with weaker relationships observed for the lower performing winter models (Figure 3-25).

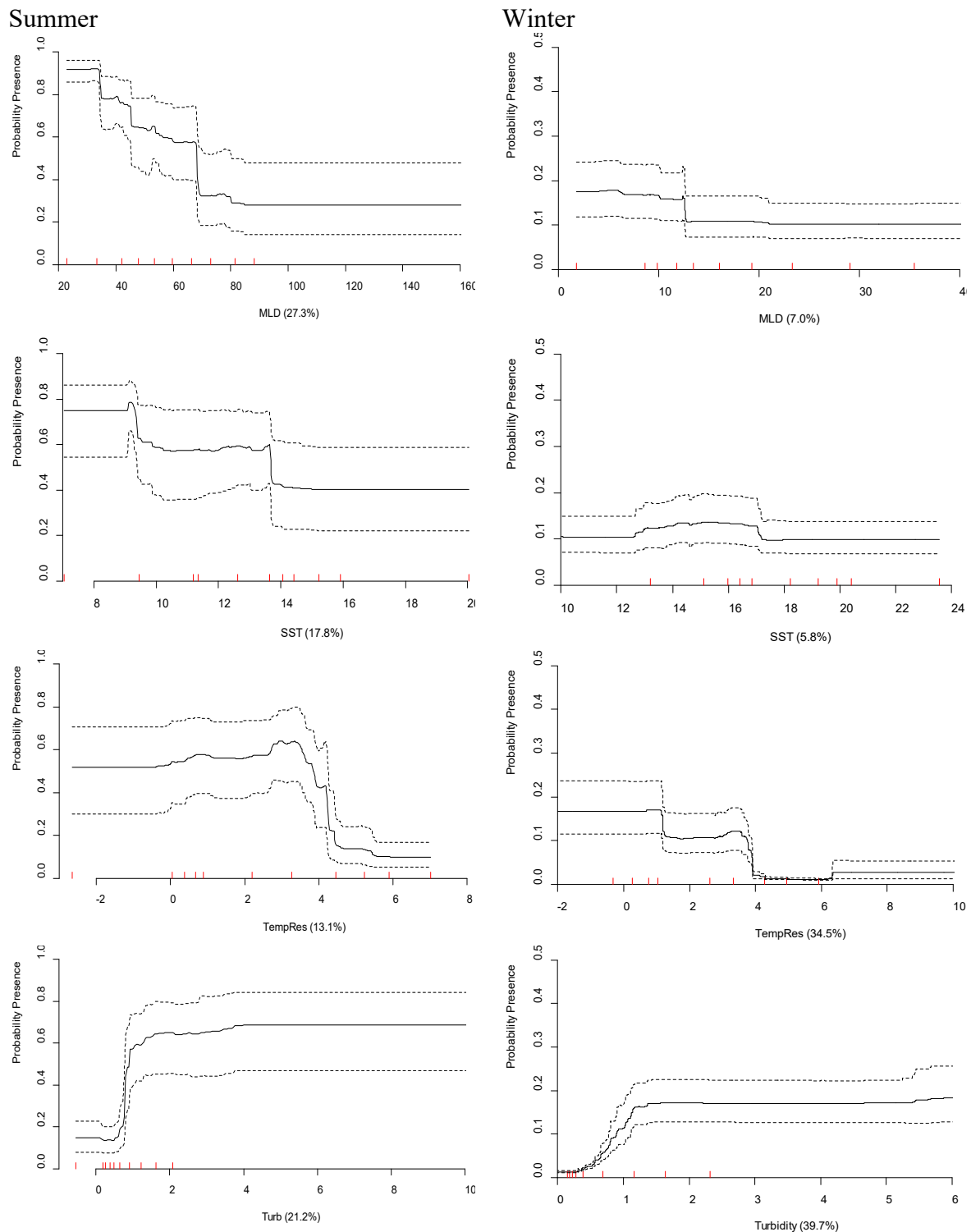


Figure 3-25: Partial dependence plots showing the relationships between predictor variables and predicted presence of Hector's dolphin (*Cephalorhynchus hectori hectori*) modelled using bootstrapped BRTs for winter (right) and summer (left). The four most influential environmental predictors in the model are shown for each season. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Similar predicted distributions were observed for seasonal Hector's dolphin presence/relative absence models as were observed for average year-round probability presence predictions; few differences were observed in distributions patterns between winter and summer (Figure 3-26). Uncertainty estimates (CV) for seasonal probability presence models were low for both seasons, although these were higher (particularly for areas outside the core species distribution) for the lower performing winter model (Figure 42, Appendix 3).

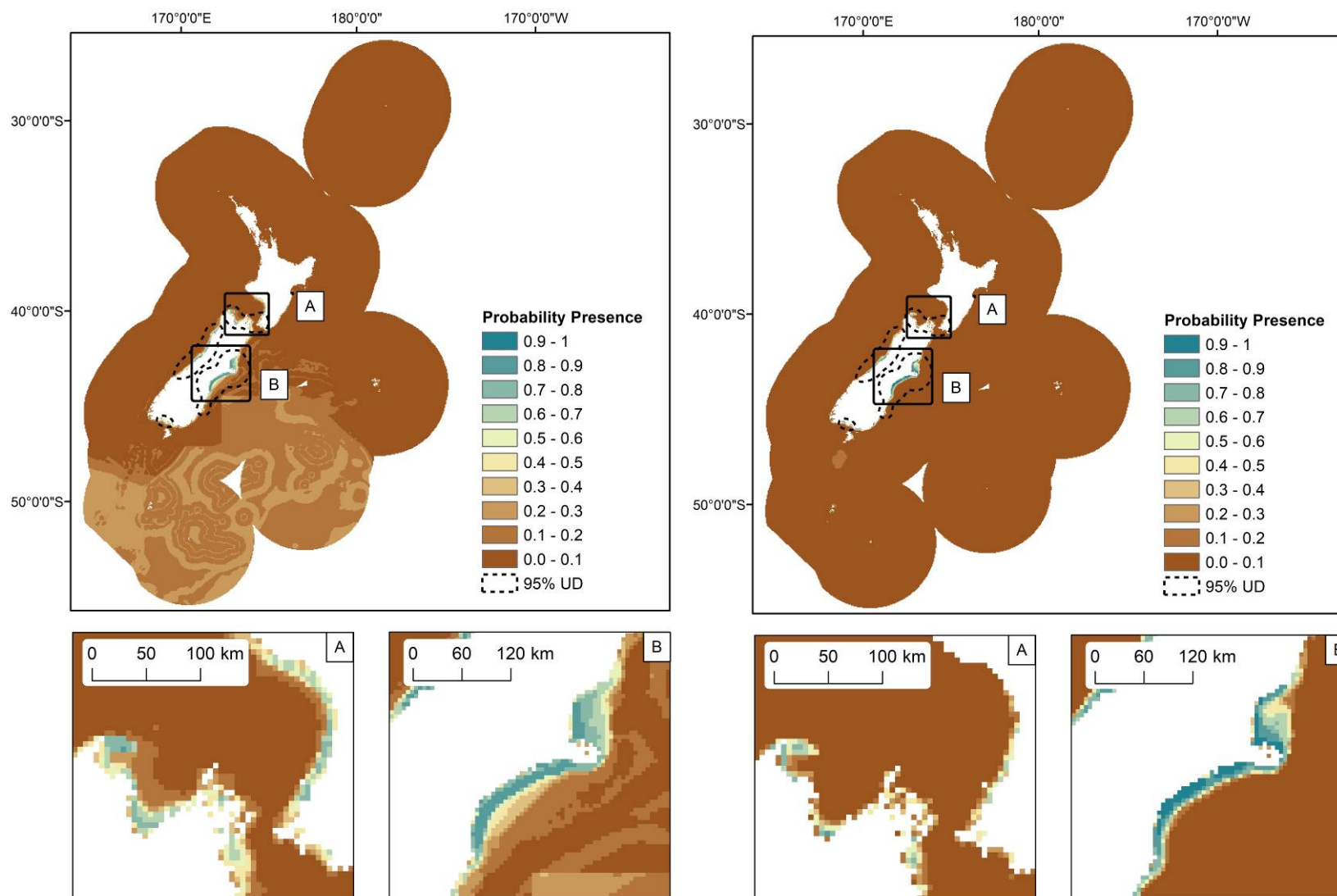


Figure 3-26: The predicted probability presence of Hector's dolphin (*Cephalorhynchus hectori hectori*) modelled using bootstrapped BRTs fitted with winter (May-Oct, $n = 294$) sightings records (left) and summer (Nov-Apr, $n = 3394$) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is defined by the dashed line. Inset maps: A) Taranaki Bight; B) east coast of South Island.

3.2.5.4 Dusky dolphin (*Lagenorhynchus obscurus*)

3.2.5.4.1 Overview

Qualitative descriptions of dusky dolphin sample number, distribution, assessment of models fits (including independent validation using NOMAD data) for average year-round presence/relative absence and group size models, and a brief description of spatial differences between seasonal distributions are provided in the highlights table (Table 3-18).

Table 3-18: Dusky dolphin (*Lagenorhynchus obscurus*): highlights of model fits (including independent validation using NOMAD data) and geographic prediction. Table colours and qualitative categories used here are provided in Table 2.4.

Sample number	Distribution	P/RA			Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	NOMAD Evaluation (AUC)	Model fit (R ²)	Model fit (dev. Exp)	
High	Localised – south-east study area	Excellent	Excellent	Excellent	Fair	Poor	Little seasonal variability in distribution, higher probability presence in summer

3.2.5.4.2 Average year-round presence/relative absence models

Deviance explained and AUC scores for dusky dolphin average year-round presence/relative absence models were high and showed consistency between training and evaluation data and between bootstrap samples (Table 3-19). Model validation with independently collected presence/(true) absence data showed that the model had high predictive power (NOMAD evaluation AUC: 0.91, ‘excellent’) providing strong evidence that models were performing well (Table 3-19).

Table 3-19: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of dusky dolphin (*Lagenorhynchus obscurus*).

	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)	NOMAD Evaluation (AUC)
Mean	0.55	0.55	0.95	0.95	0.91
Standard Deviation	0.01	0.03	0.00	0.01	NA

The most important environmental predictor variables for dusky dolphin presence/relative absence models were TempRes (55.2%), Dist.Iso500 (16.4%), SST (10.5%), and Bathy (6.9%) (Figure 3-27). Given all other variables kept at their means, there was a strong positive relationship of predicted dusky dolphin probability presence with TempRes, with increased probability presence for those areas with lower than 0 °C bottom water temperatures expected for any given depth (Figure 3-27). Decreasing probability presence of dusky dolphins was observed with increasing distance to the 500 m isopleth (Dist.Iso500), and higher probability presence was observed for areas deeper than ~75 m (Bathy) and with SST lower than 15 °C. Overall, there was low uncertainty in predictions as shown by the 95% prediction (dashed lines in Figure 3-27).

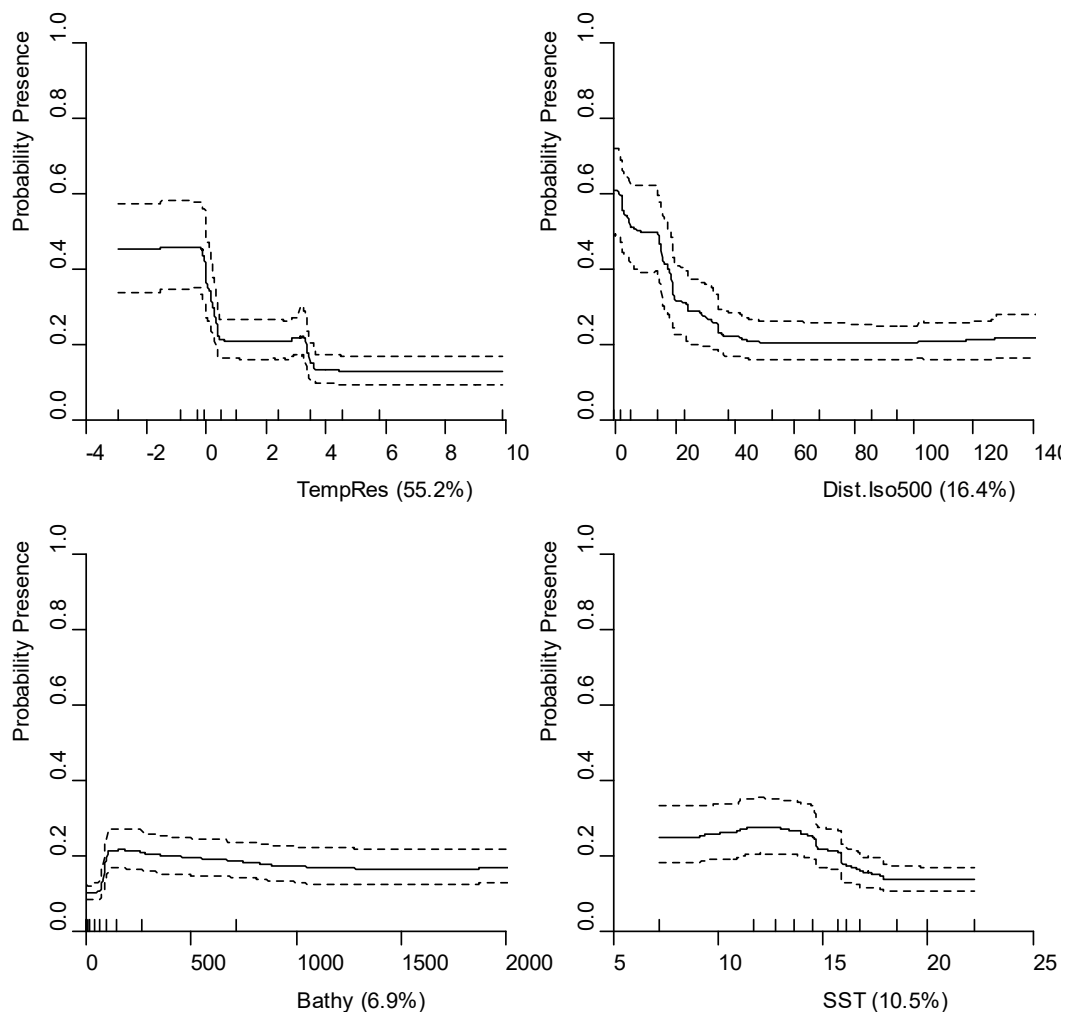


Figure 3-27: Partial dependence plots showing the relationships between predictor variables and probability presence of dusky dolphin (*Lagenorhynchus obscurus*) modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Dusky dolphin presence is predicted to be most likely in midwater depths along the Chatham Rise, Campbell Plateau and surrounding waters of Auckland, Campbell, and Antipodes islands (blue areas in Figure 3-28). The highest probability of presence was predicted across large areas on the west of the Chatham Rise (inset A in Figure 3-28) and off the east coast of the South Island (inset B in Figure 3-28), in agreement with expert knowledge of spatial distributions.

Predicted probability of presence was congruent with the utilisation distribution (UD - Figure 3-28). Predicted probability of presence was also visually congruent with: presence/relative absence, strandings, and NOMAD presence location records (Figure 3-29), although for large areas with high predicted probability of presence, no samples were available (e.g., inset A, C, and E in Figure 3-29).

Spatially explicit estimates of uncertainty (CV) were low across the study area particularly for areas with high predicted probability of presence (Figure 43 and 44, Appendix 3). Uncertainty estimates increased slightly for those areas immediately outside the core species distribution (see inset B in Figure 43, Appendix 3). A portion of dusky dolphin predicted presence distribution falls outside areas of the environmental space that are considered well sampled (see Figure 3-2) and there remains some

uncertainty as to the accuracy of these offshore predictions (particularly for those areas outside the UD — Figure 3-28).

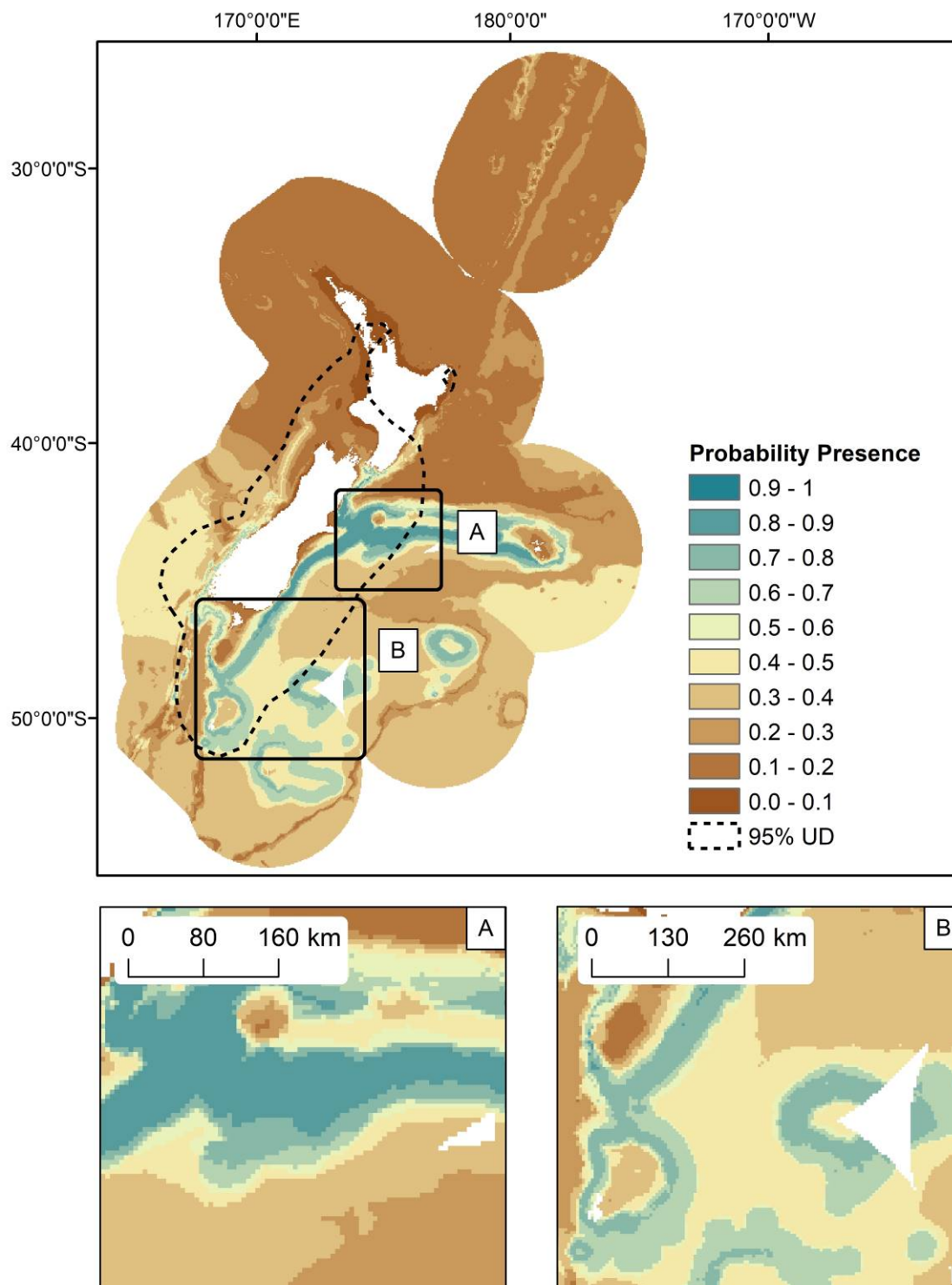


Figure 3-28: The predicted probability of dusky dolphin (*Lagenorhynchus obscurus*) presence in the New Zealand EEZ, from bootstrapped BRT models. The predicted 95% utilisation distribution is defined by the dashed line. Inset maps: A) west Chatham Rise; B) Campbell Plateau.

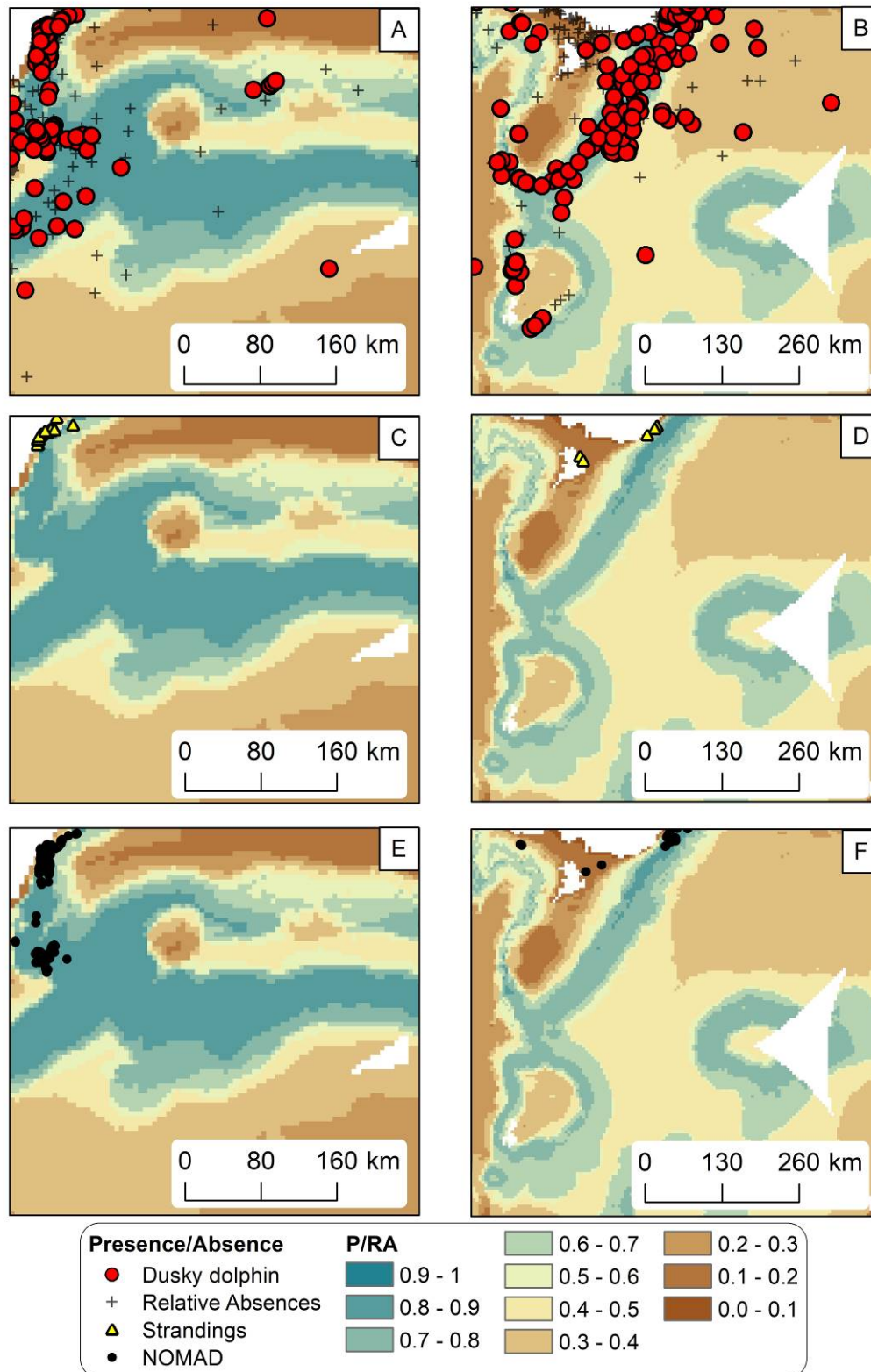


Figure 3-29: Predicted probability of dusky dolphin (*Lagenorhynchus obscurus*) presence modelled using bootstrapped BRTs. Predicted probability presence of dusky dolphin on the west Chatham Rise are shown (left panel) with presence/relative absence (red circles and black crosses respectively) (A), DOC stranding locations (C), and NOMAD sightings (E). Predicted probability presence of dusky dolphin on the Campbell Plateau are shown (right panel) with presence/relative absence (B), stranding locations (D), and NOMAD sightings (F).

3.2.5.4.3 Average year-round group size models and relative density predictions

Deviance explained and Pearson's correlation of predicted vs observed species group sizes for dusky dolphin average year-round group size models were low (Table 3-20). Estimates of these model fit metrics were variable between bootstrap samples (standard deviation in Table 3-20). With these low (and variable) model fits, caution should be used when interpreting dusky dolphin average year-round group size models.

Table 3-20: Mean model performance measures (deviance explained and Pearson's correlation of predicted vs observed group sizes (R^2)) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of dusky dolphin (*Lagenorhynchus obscurus*).

	Deviance explained (training data)	Deviance explained (evaluation data)	Pearson's correlation of predicted vs observed species group sizes (R^2)
Mean	0.06	0.08	0.12
Standard Deviation	0.08	0.05	0.27

The most important environmental predictor variables in average year-round dusky dolphin group size models differed to those of the average year-round presence/relative absence model. In the average year-round group size models, the most important variables were MLD (22.7%), TC (19.8%), Dist.Shore (9.8%), and Slope (9.5%) (Figure 3-30). Average year-round group sizes of dusky dolphin were predicted to be highest for areas with mixed layer depths greater than 40 m (MLD) and for areas with low tidal current speeds (less than 0.025 ms^{-1} – TC) (Figure 3-30). Trends in dusky dolphin group sizes for other environmental predictors were less clear (and weaker) (Figure 3-30).

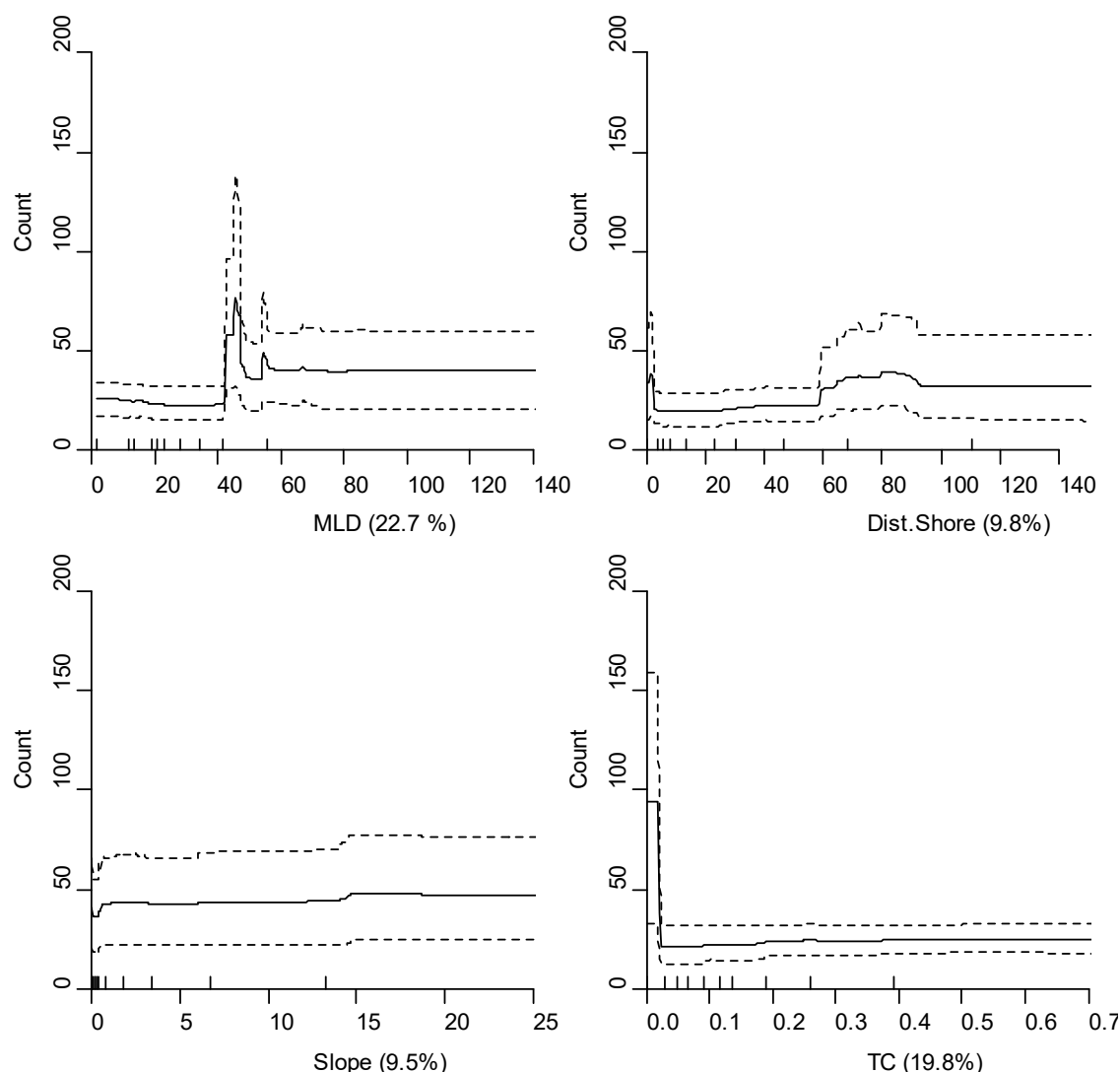


Figure 3-30: Partial dependence plots showing the relationships between predictor variables and dusky dolphin (*Lagenorhynchus obscurus*) group sizes, modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Distribution of predicted average year-round dusky dolphin relative densities (the resulting prediction from the two-step hurdle method) were widespread across the study area (Figure 3-31). The highest dusky dolphin relative densities were predicted to occur on the south-western side of the Chatham Rise and offshore areas on the Campbell Plateau (Figure 3-31). Predicted relative densities were broadly in agreement with observed group sizes (comparison of relative density estimates from Figure 3-31 with observed group size in Figure 3-32); however, the correlation of model predictions and observed group size was low (mean Pearson correlation 0.12). Uncertainty (95% prediction interval) was low for areas within ~300 km of the shore, but were higher for areas further from shore (Figure 44, Appendix 3).

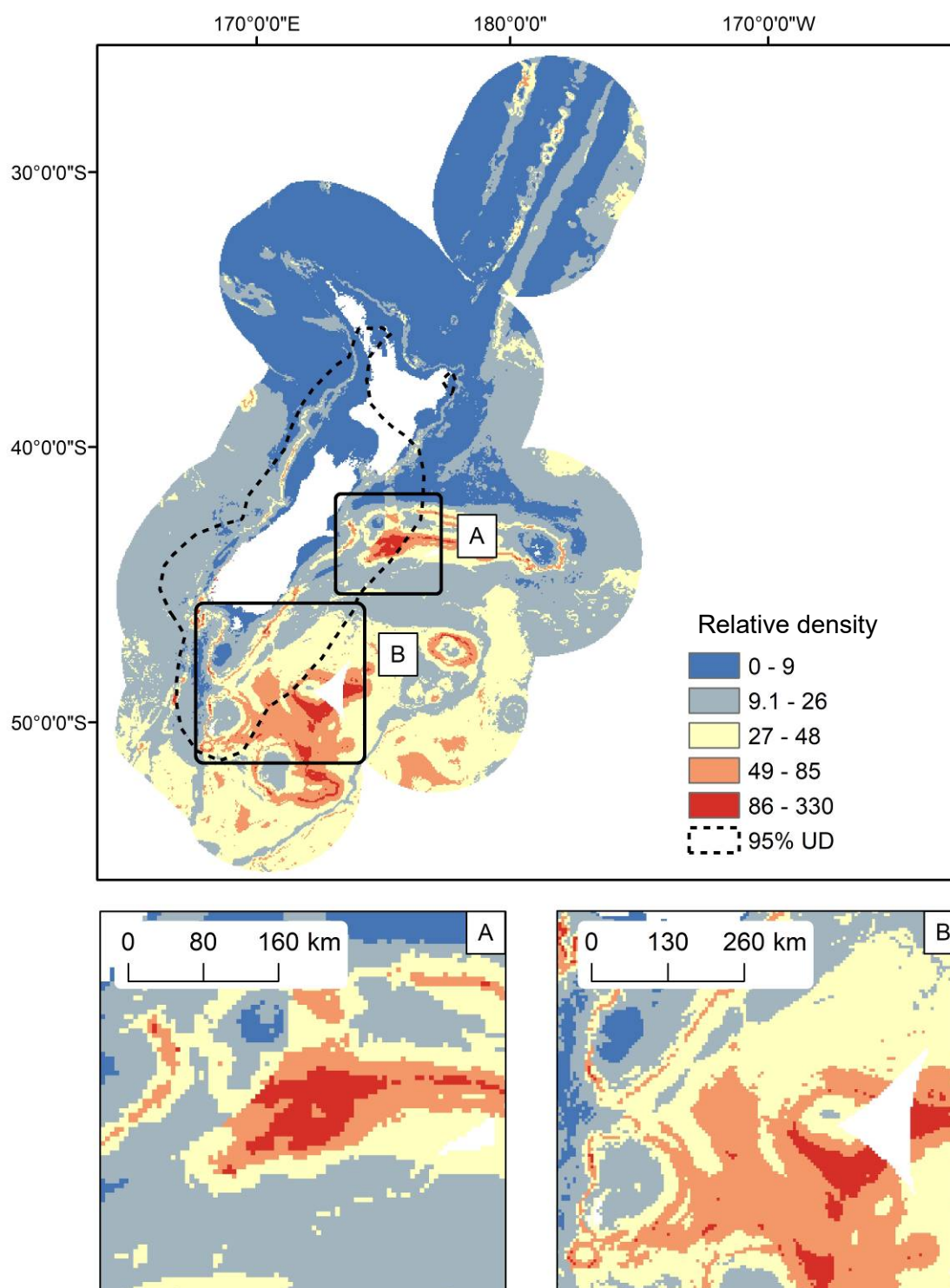


Figure 3-31: Predicted relative density of dusky dolphin (*Lagenorhynchus obscurus*) in the New Zealand EEZ, from bootstrapped BRT models. The predicted 95% utilisation distribution is defined by the dashed line. Inset maps: A) west Chatham Rise; B) Campbell Plateau.

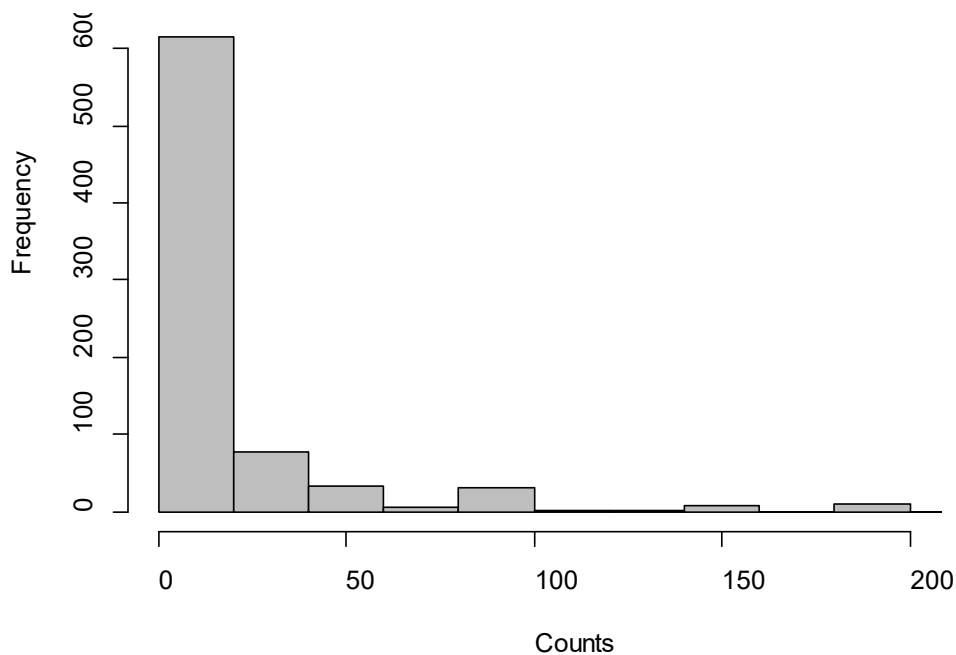


Figure 3-32: Histogram of dusky dolphin (*Lagenorhynchus obscurus*) group sizes recorded in the New Zealand EEZ.

3.2.5.4.4 Seasonal presence/relative absence models

Goodness-of-fit metrics for summer presence/relative absence models were higher (explained deviance: 0.64; AUC: 0.96) than both average year-round (explained deviance: 0.55; AUC: 0.91) and winter presence/relative absence models (explained deviance: 0.41; AUC: 0.91) (Table 3-21). As with the average year-round presence/relative absence models, seasonal model fit metrics differed little between those derived from training data and evaluation data, and mean estimates did not vary greatly between bootstrap samples (except for deviance explained using evaluation data in winter models; Table 3-21). The difference in model fits between seasonal models may be explained by the differences in sample number, i.e., winter had fewer species presence records ($n = 169$) compared with summer ($n = 654$) models.

Table 3-21: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) from winter (May-Oct) and Summer (Nov-Apr) sightings of dusky dolphin (*Lagenorhynchus obscurus*) records.

Season	Metric	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)
Winter	Mean	0.40	0.41	0.90	0.91
	Standard Deviation	0.09	0.19	0.02	0.03
	Mean	0.64	0.64	0.96	0.96
Summer	Standard Deviation	0.03	0.06	0.00	0.01

Similar environmental predictors were important in dusky dolphin seasonal models as were in the average year-round presence/relative absence models (albeit with differing percentage contributions). However, Bathy was not one of the most important variables in seasonal models and MLD was not the selected as one of the most important variables in average year-round models (Figure 3-33). Similar patterns in predicted probability of presence along environmental gradients were observed between

seasons and average year-round presence/relative absence models, though weaker relationships were observed for the lower performing winter models (note the difference in scale of the y-axis, probability of presence — Figure 3-33).

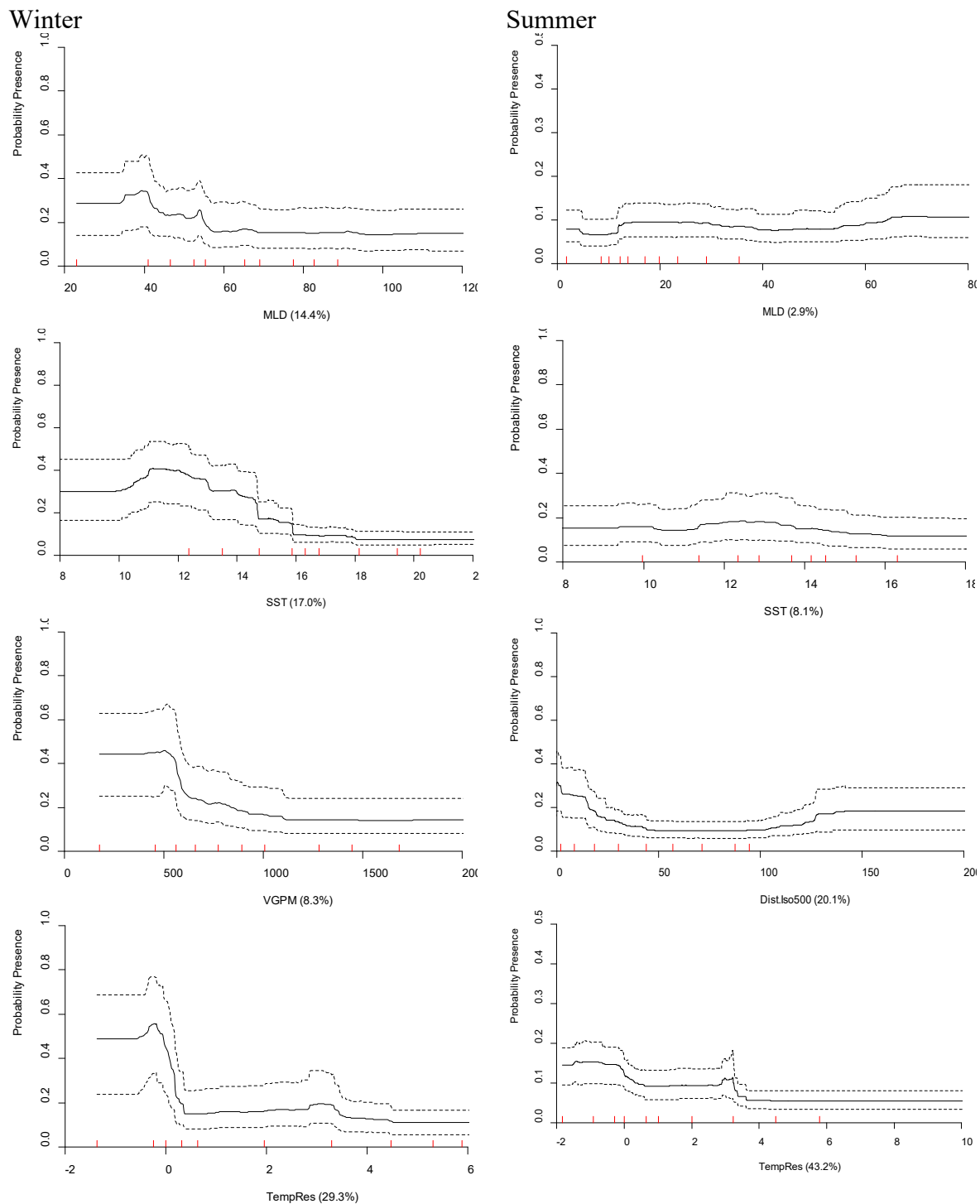


Figure 3-33: Partial dependence plots showing the relationships between predictor variables and predicted presence of dusky dolphin (*Lagenorhynchus obscurus*) modelled using bootstrapped BRTs for winter (left) and summer (right). The four most influential environmental predictors in the model are shown for each season. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x- axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Similar predicted distributions were observed for seasonal models as were observed for average year-round predicted probability of presence (Figure 3-34). There was little difference in distribution patterns between winter and summer, although summer probability of presence predictions were higher in core species distribution areas and were slightly more extensive than for winter predictions (Figure 3-34). Uncertainty estimates (CV) for summer probability of presence models were low (Figure 46, Appendix 3). Uncertainty estimates (CV) for winter probability of presence models were higher (particularly for areas outside the core species distribution and for areas surrounding the North Island where there was low predicted probability of presence) (Figure 46, Appendix 3).

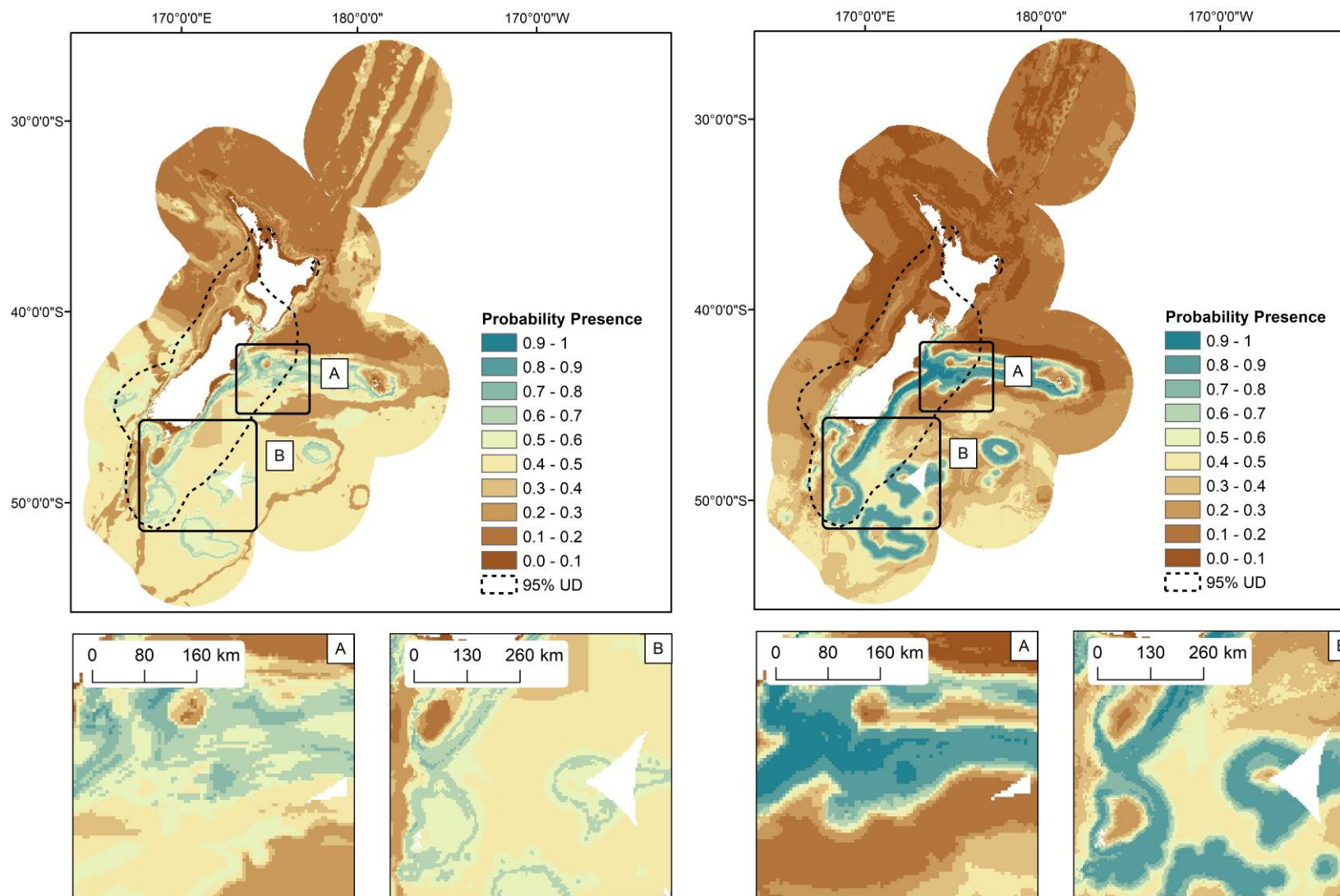


Figure 3-34: The predicted probability presence of dusky dolphin (*Lagenorhynchus obscurus*) modelled using bootstrapped BRTs fitted with winter (May-Oct, $n = 169$) sightings records (left) and summer (Nov-Apr, $n = 654$) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is defined by the dashed line. Inset maps: A) west Chatham Rise; B) Campbell Plateau.

3.2.5.5 Bryde's Whale (*Balaenoptera brydei*)

3.2.5.5.1 Overview

Qualitative descriptions of Bryde's whale sample number, distribution, assessment of models fits (including independent validation using NOMAD data) for average year-round presence/relative absence and group size models, and a brief description of spatial differences between seasonal distributions are provided in the highlights table (Table 3-22).

Table 3-22: Bryde's whale (*Balaenoptera brydei*): highlights of model fits and geographic prediction. Table colours and qualitative categories used here are provided in Table 2.4.

Sample number	Distribution	P/RA		Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
High	Localised – Hauraki Gulf and north of North Island	Excellent	Excellent	Poor	No predictive power	Constriction of range to Hauraki Gulf in winter

3.2.5.5.2 Average year-round presence/relative absence models

Deviance explained and AUC scores for Bryde's whale average year-round presence/relative absence models were high and were consistent between training and evaluation data, and between bootstrap samples (Table 3-23). Model validation with independently collected presence/(true) absence data was not possible due to limited sightings records in the NOMAD database.

Table 3-23: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of Bryde's whale (*Balaenoptera brydei*).

	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)
Mean	0.49	0.50	0.93	0.93
Standard Deviation	0.02	0.04	0.01	0.01

The most influential environmental predictor variables for Bryde's whale presence/relative absence models were TempRes (50.6%), Turb (13.2%), SST (10.4%), and TC (9.6%) (Figure 3-35). There was a strong positive relationship of predicted Bryde's whale probability of presence with TempRes, with increased probability of presence for those areas with higher than 4 °C bottom water temperatures expected for a given depth (Figure 3-35). Weaker relationships were observed between probability presence of Bryde's whale and SST (increased probability presence in areas with SST greater than 20 °C) and turbidity (increased probability of presence in areas with low turbidity) (Figure 3-35). There was no clear relationship between probability presence and TC (Figure 3-35).

Overall, there was low uncertainty in predictions as shown by the 95% prediction intervals; although higher uncertainty was observed for TempRes values greater than 4 °C (dashed lines in Figure 3-35).

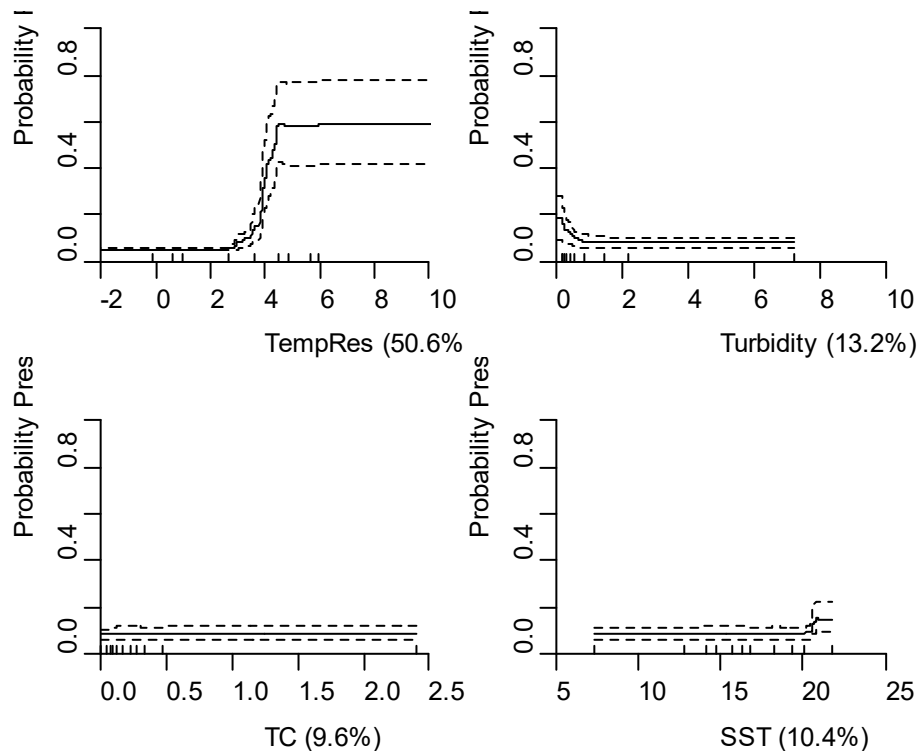


Figure 3-35: Partial dependence plots showing the relationships between predictor variables and probability presence of Bryde's whale (*Balaenoptera brydei*) modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

Bryde's whale predicted presence was highest close to shore (blue areas in Figure 3-36 and Figure 3-37), particularly in the Hauraki Gulf (inset A in Figure 3-36), in agreement with research (Wiseman et al. 2011).

Predicted probability of presence was congruent with the utilisation distribution (UD - Figure 3-36). Predicted probability of presence was also visually congruent with presence/relative absence and strandings location records (Figure 3-37).

Spatially explicit estimates of uncertainty (CV) were low across the study area (Figures 47 and 48, Appendix 3). Uncertainty estimates increased slightly for those areas with intermediate predicted probability of presence (e.g., 0.3 – 0.6 along the northern coastline of the North Islands) (see inset A in Figures 47 and 48, Appendix 3). Bryde's whale records and model outputs fall within areas of the environmental space that are considered well sampled (see Figure 3-2) and, thus, uncertainty layers are likely to provide an accurate representation of spatial uncertainty.

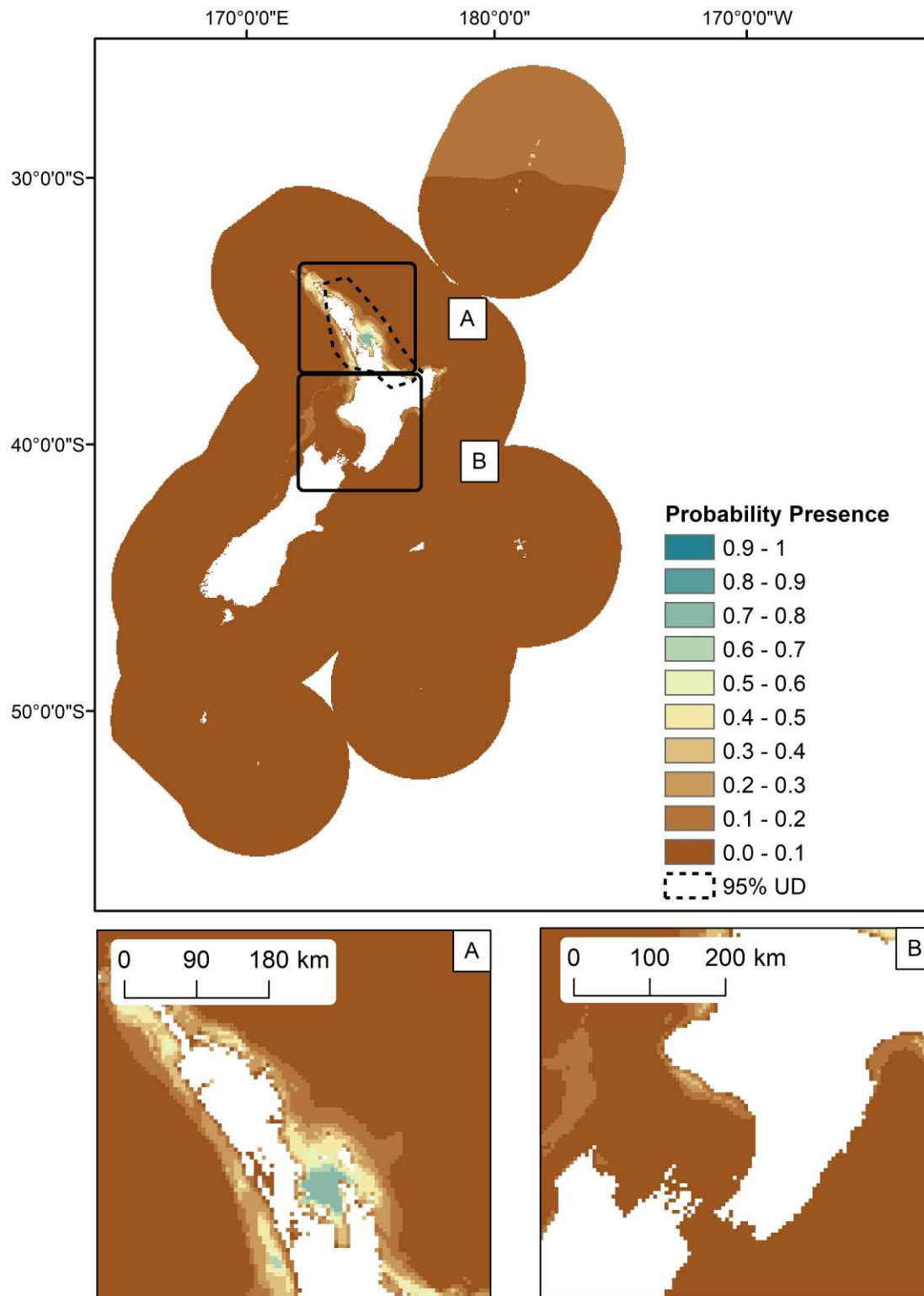


Figure 3-36: The predicted probability of Bryde's whale (*Balaenoptera brydei*) presence in the New Zealand EEZ, from bootstrapped BRT models. The predicted 95% utilisation distribution is defined by the dashed line. Inset maps: A) north of North Island; B) Taranaki Bight and south-east of North Island.

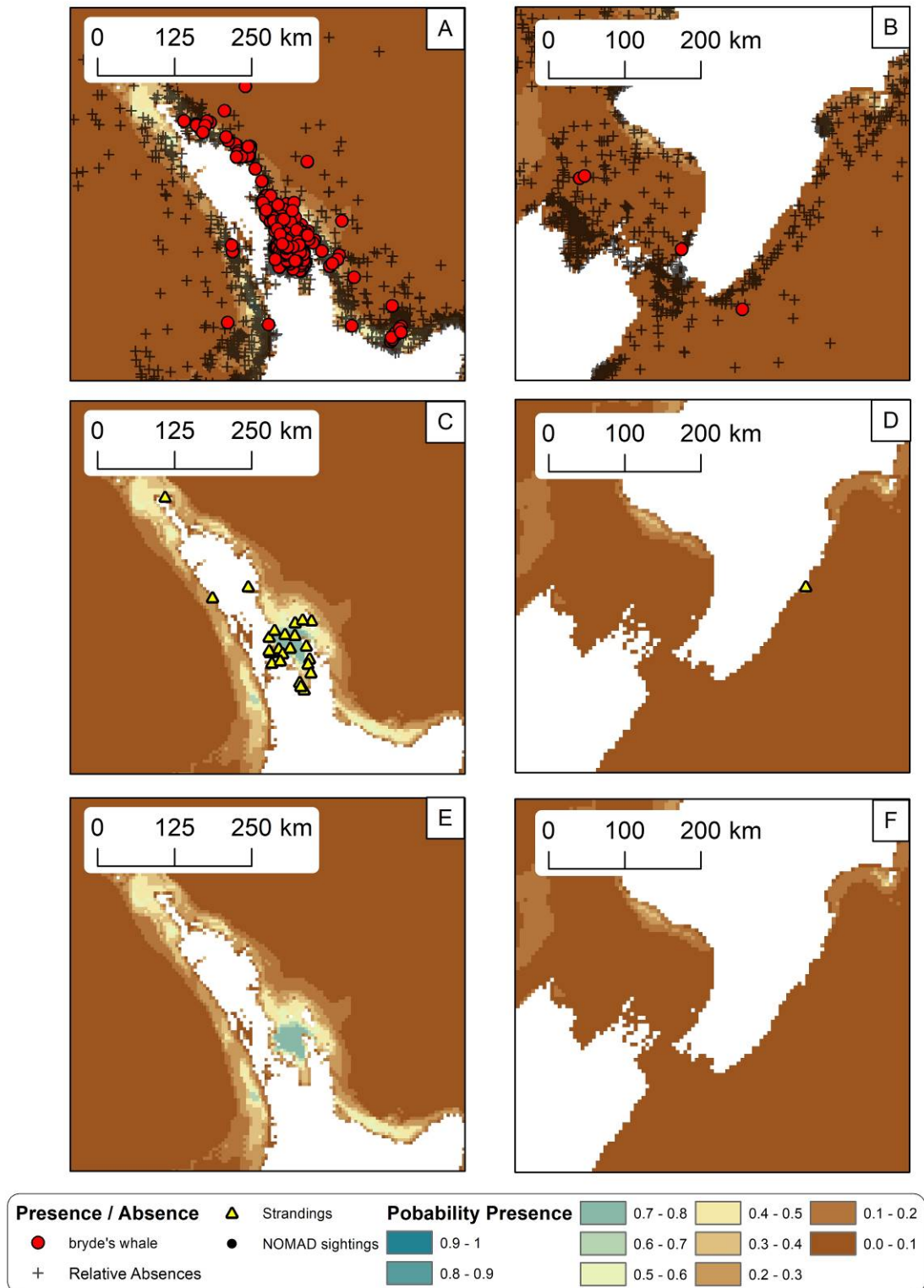


Figure 3-37: Predicted probability of Bryde's whale (*Balaenoptera brydei*) presence modelled using bootstrapped BRTs. Predicted probability of Bryde's whale presence north of North Island (left panel) are shown with presence/relative absence (red circles and black crosses respectively) (A), DOC stranding locations (C), and NOMAD sightings (E). Predicted probability of Bryde's whale presence in the Taranaki Bight and southeast of North Island (right panel) are shown with presence/relative absence (B), stranding locations (D), and NOMAD sightings (F).

3.2.5.5.3 Average year-round group size models and relative density predictions

Bryde's whale average year-round group size models had no predictive power (0 mean deviance explained from the evaluation data) and poor correlative power for predicted vs observed species group sizes (Table 3-24). The very low model fits observed indicates that little confidence should be placed on the prediction of Bryde's whale average year-round distribution of group sizes. The low model performance is most likely linked to the low group sizes observed (i.e., ~ 80% of group sizes were for 1–3 individuals, Figure 3-38). The low variability in group sizes likely resulted in the inability of the models to accurately determine the relationship, if any, between group sizes and the environmental variables.

Table 3-24: Mean model performance measures (deviance explained and Pearson's correlation of predicted vs observed group sizes (R^2)) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of Bryde's whale (*Balaenoptera brydei*).

	Deviance explained (training data)	Deviance explained (evaluation data)	Pearson's correlation of predicted vs observed species group sizes (R^2)
Mean	0.04	0.00	0.12
Standard Deviation	0.03	0.05	0.07

The most important environmental predictor variables in average year-round Bryde's whale group size models were ChlA (44.6%), VGPM (9.7%), SST (7.0%), and BedDist (2.7%) (Figure 3-39). The relationships between Bryde's whale group sizes and environmental predictor variables were unclear, except for the increased predicted group sizes for areas with ChlA values greater than 2 mg m^{-3} , though with high uncertainty (see Figure 3-39).

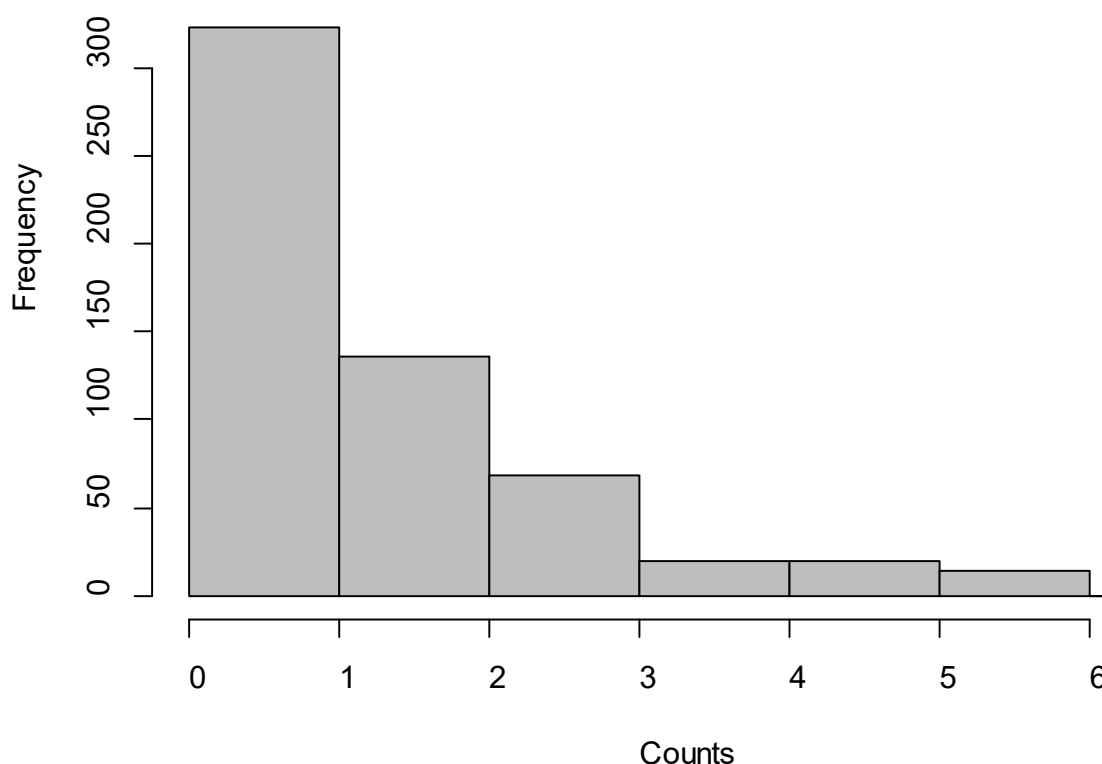


Figure 3-38 Histogram of Bryde's whale (*Balaenoptera brydei*) group sizes recorded in the New Zealand EEZ.

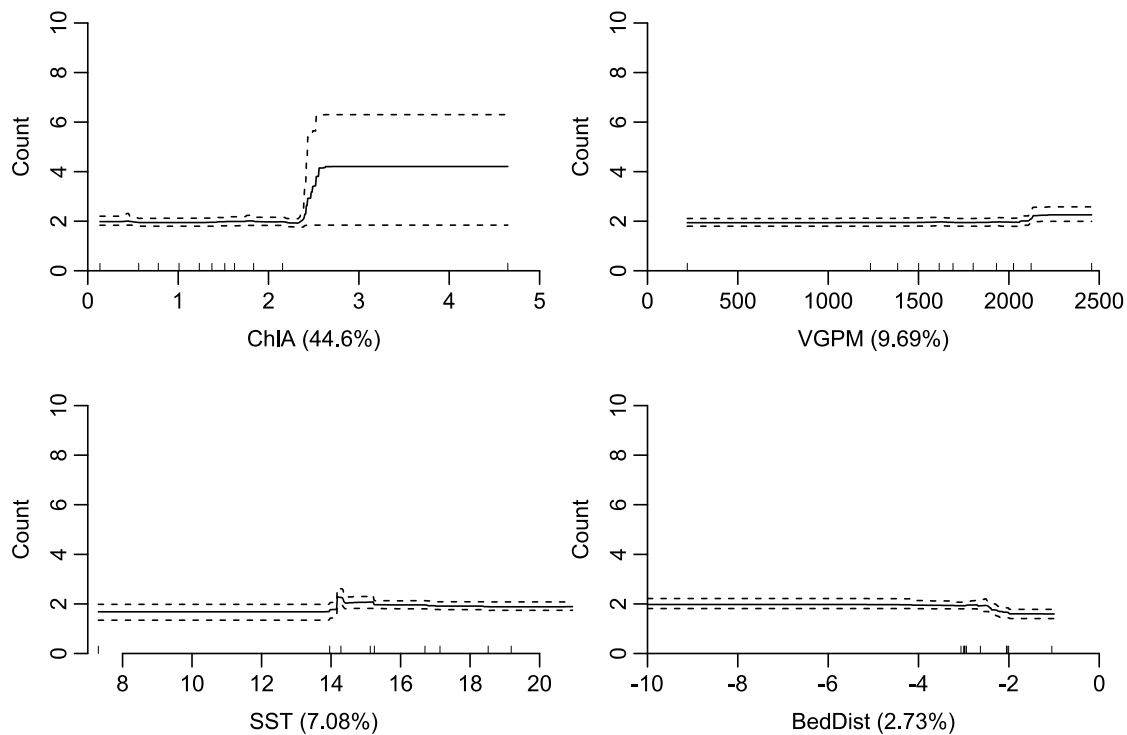


Figure 3-39: Partial dependence plots showing the relationships between predictor variables and Bryde's whale (*Balaenoptera brydei*) group sizes, modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

The distribution of predicted average year-round Bryde's whale relative density tightly followed geographic patterns described for average year-round probability presence with the only increase in predicted relative densities occurring in the Hauraki Gulf (Figure 3-40). For the majority of areas where Bryde's whale relative densities were predicted, these were low (ranging between 1 and 3; Figure 3-40). Uncertainty (95% prediction interval – Figure 49, Appendix 3) was variable across the study area, most likely a reflection of the poor model fits.

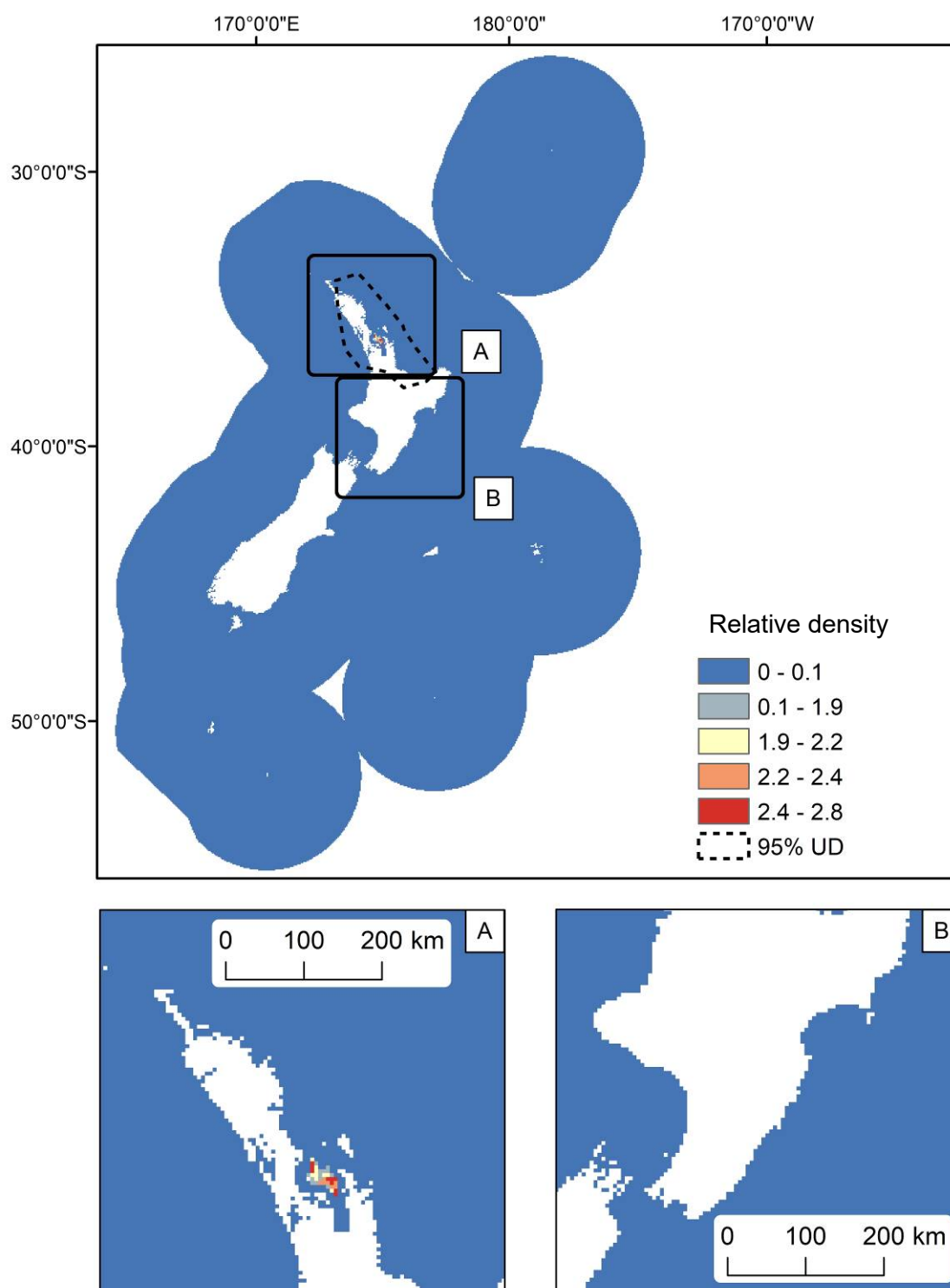


Figure 3-40: Predicted relative density of Bryde's whale (*Balaenoptera brydei*) in the New Zealand EEZ, from bootstrapped BRT models. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island; B) Taranaki Bight and south-east of North Island.

3.2.5.5.4 Seasonal presence/relative absence models

Goodness-of-fit metrics for seasonal presence/relative absence models were nearly identical to average year-round presence/relative absence models (Table 3-25). Seasonal model fit metrics differed little between those derived from training data and evaluation although mean estimates varied slightly more between bootstrap samples than those observed in the average year-round model fit metrics (Table 3-25). The similarity in model fits between seasonal models may be explained by the similar number of presence records in each season, i.e. winter models were tuned with 284 species presence records and summer models were tuned with 309 species presence records.

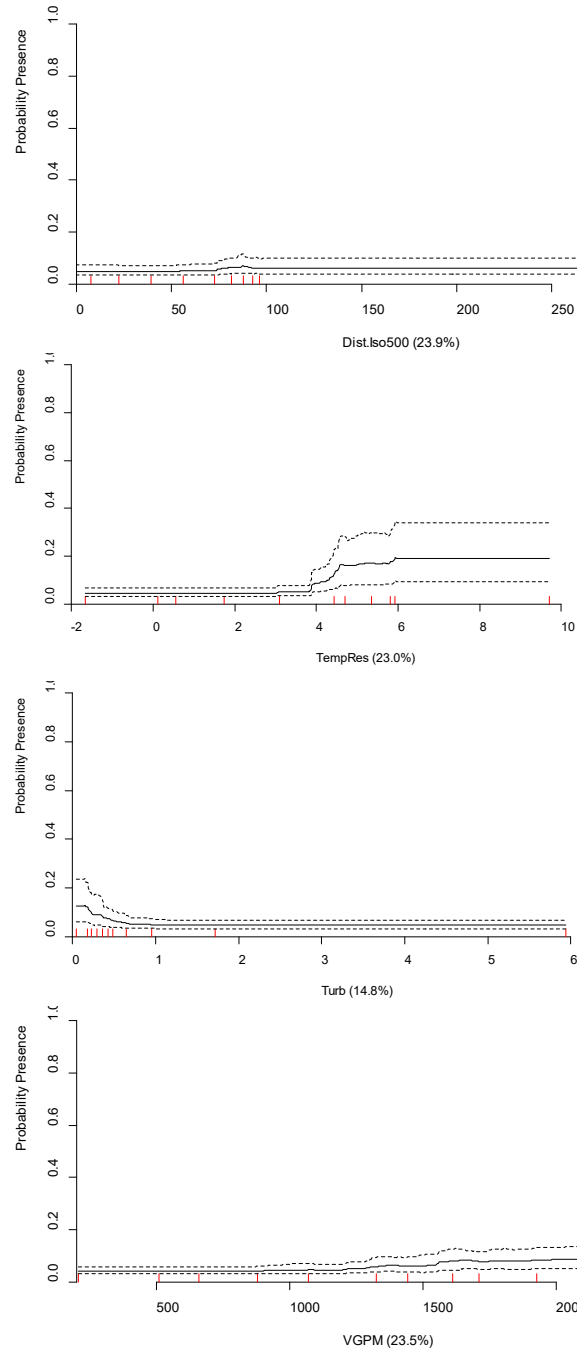
Table 3-25: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) from winter (May-Oct) and summer (Nov-Apr) sightings of Bryde’s whale (*Balaenoptera brydei*) records.

Season	Metric	Deviance explained (training data)	Deviance explained (evaluation data)	AUC (training data)	AUC (evaluation data)
Winter	Mean	0.51	0.51	0.93	0.93
	Standard Deviation	0.05	0.10	0.01	0.02
Summer	Mean	0.50	0.50	0.93	0.93
	Standard Deviation	0.06	0.11	0.01	0.02

The four most important environmental predictors in Bryde’s whale seasonal models differed between seasons and to those used in the average year-round presence/relative absence models (Figure 3-41). Weaker and less obvious relationships in the predicted probability presence along environmental gradients were observed between seasons compared with those observed in average year-round presence/relative absence models (Figure 3-41).

Despite the differences in predictor variable importance, broadly similar predicted distributions were observed for seasonal probability of Bryde’s whale presence as were observed for average year-round probability of presence predictions (Figure 3-42). Predicted winter distributions of Bryde’s whale presence were more constricted than those observed for summer (Figure 3-42). As with the average year-round models for Bryde’s whale probability of presence, uncertainty estimates (CV) for seasonal probability of presence models were moderate for both seasons, although these were higher for summer predictions (Figure 50, Appendix 3).

Winter



Summer

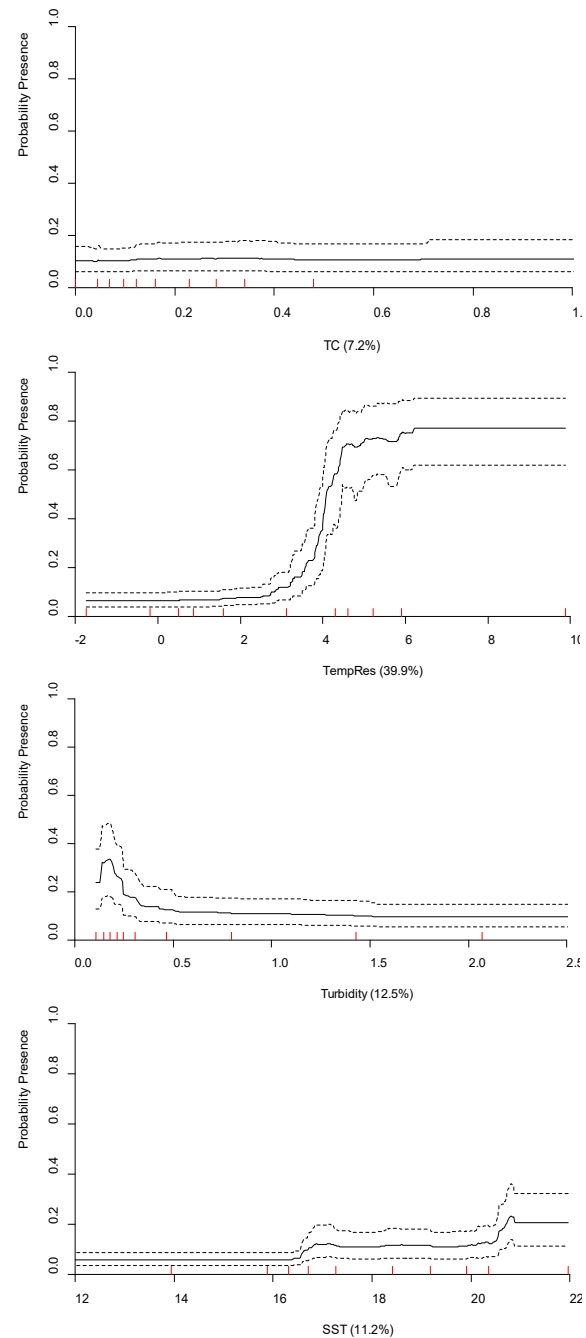


Figure 3-40: Partial dependence plots showing the relationships between predictor variables and predicted presence of Bryde's whale (*Balaenoptera brydei*) modelled using bootstrapped BRTs for winter (left) and summer (right). The four most influential environmental predictors in the model are shown for each season. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Deciles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

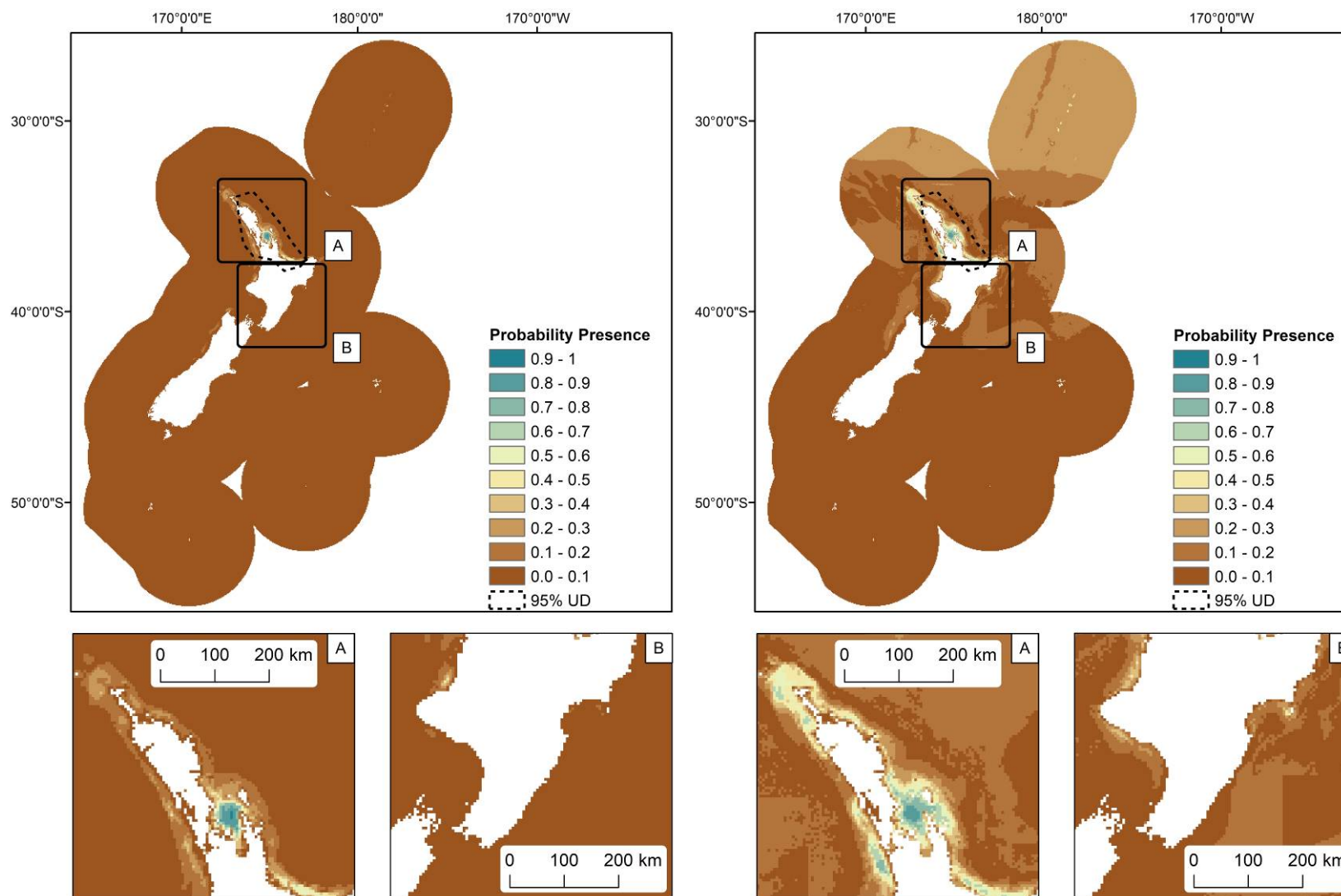


Figure 3-41: The predicted probability of Bryde's whale (*Balaenoptera brydei*) presence modelled using bootstrapped BRTs fitted with Winter (May - Oct, $n = 284$) sightings records (left) and Summer (Nov - Apr, $n = 309$) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island; B) Taranaki Bight and southeast of North Island.

4 DISCUSSION

Information on cetacean species distributions in New Zealand waters is limited, with many studies relying solely on observations to describe distributions and/or restricted to regional-scale study areas or restricted to species with high abundance or of particular conservation interest: e.g., distribution and density patterns of common dolphins, bottlenose dolphins, and Bryde's whales in the Hauraki Gulf (Dwyer et al. 2016); distribution and abundance of dolphins off the west coast of the South Island using observations only (Brager and Schneider 1998). More recently, species sighting records have been combined with environmental predictor variables to provide estimates of species distributions using both complex modelling approaches (e.g., BRTs as used by Derville et al. 2016) and simpler approaches (e.g., RES as described by Kaschner et al. 2006). These modelling approaches have the advantage of producing estimates of species presence and density for unsampled areas. However, there is still a paucity of information about cetacean densities, especially across the whole of the New Zealand EEZ. The best available estimates of cetacean species density estimates were previously produced using expert based estimates of species density for (large) fisheries management areas (Abraham et al. 2017).

This report describes the collation of a comprehensive set of spatial information for 31 cetacean species, subspecies, and species complexes occurring in New Zealand waters, addresses some of the knowledge gaps, and provides improved cetacean species distribution estimates by:

- 1) Using expert opinion to define environmental preferences and predict simple environmental habitat preference for rarely observed cetacean species using RES.
- 2) Modelling and mapping species distributions and relative densities (for those species with sufficient sightings records) using a comprehensive database of recorded cetacean sightings and spatially comprehensive, high-resolution, and functionally relevant environmental data layers.

In the following sections, these species' distributions and relative density estimates are critically appraised and management implications of using these layers for spatially-explicit fisheries risk assessments are discussed.

4.1 Critical appraisal of model results

4.1.1 Relative Environmental Suitability (RES) models

RES estimates for rarely sighted cetacean species (Andrew's beaked whale, Arnoux's beaked whale, Blainville's beaked whale, Cuvier's beaked whale, dwarf minke whale, false killer whale; Gray's beaked whale, hourglass dolphin, pygmy sperm whale; Risso's dolphin; Shepherd's beaked whale; southern bottlenose whale, southern right whale dolphin, spectacled porpoise, and striped dolphin) were produced using a combination of expert opinion and information available from the literature (see Appendix 2 for further information). The RES approach is a conceptually simple modelling approach, although it is a well-established method that has been successfully applied to predict global distributions of cetacean species and other marine species (e.g., AquaMaps: Standardised distribution maps for over 25 000 species of fishes, marine mammals, and invertebrates, <https://www.aquamaps.org>; Kesner-Reyes et al. 2016). RES has the advantage of correcting for biases in occurrence data such as non-representative coverage of a species' distribution, biases in sampling effort and data provision, and species misidentifications (since it does not directly use sightings records to inform environmental relationships) (Kaschner et al. 2011, Kesner-Reyes et al. 2016). Additionally, RES predictions have been validated using independent and effort-corrected survey data (Ready et al. 2010, Kaschner et al. 2011). In previous studies, the performance of RES models were found to be comparable with presence only species distribution models such as GARP - Genetic Algorithm for Rule Set Production, MaxEnt - Maximum Entropy Modeling, GLMs - generalised linear models, and GAMs - generalized additive models (Kesner-Reyes et al. 2016).

For the majority of species examined here, estimated probability distributions using RES were consistent with the (limited) sightings and strandings records available for these rarely sighted species, providing evidence that the broad-scale environmental niches of individual species were effectively captured. RES outputs provided finer spatial patterns of species distributions than the more subjective distributions used for the most recent spatial risk assessment of the effects of fishing for cetacean species (Abraham et al. 2017). However, RES methods do not produce density estimates, which are of particular interest for input into spatially explicit and quantitative marine mammal risk assessments.

Although beyond the scope of this report, future iterations of RES models may benefit from the inclusion of density estimates (although methods for doing so would require development), the selection of a greater number of environmental predictor variables, and/or, on an individual species level, the selection of additional (and potentially more ecologically meaningful) variables. Here, we selected three environmental predictor variables previously used to model and map global distributions of most cetacean species (Kaschner et al. 2006). At a finer spatial scale (as was done here), differing environmental variables may be of importance for individual species and provide greater information on habitat preference and use, including any potential seasonal differences (i.e., seasonal differences could be informed by strandings data for rarely observed species, e.g., Thompson et al. 2013).

4.1.2 Boosted Regression Tree (BRT) models

Average year-round probability of presence BRT models generally performed well with measures of model fits ranging from ‘fair’ to ‘excellent’ (i.e., all these models were considered useful). Distributions of individual species probability of presence were in line with known distributions (e.g., Māui dolphins by Derville et al. 2016; common dolphins, bottlenose dolphins, and Bryde’s whales by Dwyer et al. 2016, Brager & Schneider 1998) and were visually congruent with recorded sightings data (used to train the models) and strandings data (visual validation). Of particular importance was the statistical validation of bottlenose, common, Hector’s, and dusky dolphins probability of presence models with independent true presence/absence NOMAD sightings records which ranged from ‘fair’ to ‘excellent’ highlighting that these models – trained using opportunistically collected species sightings records – had predictive power. Visually, the NOMAD cetacean sightings data were also congruent with predicted probability of presence for individual species. For certain species, e.g., common dolphin (*Delphinus delphis*), the spatial prediction obtained for this assessment was quite different from the more subjective Delphi-derived spatial density used for the most recent spatial risk assessment of the effects of fishing for cetacean species.

In most cases, seasonal predictions of probability of presence were similar to overall predictions for individual species, although unsurprisingly, model fits seemed heavily reliant on the number of samples. Seasonal models may be particularly useful if fishing effort is known to change throughout the year as seasonal patterns in fishing effort distribution are commonly understood to occur in many fisheries (Baird et al. 2015, Stephenson et al. 2017).

Species group size models had lower model fit metrics than the probability of presence models. Despite these lower and more variable model fit metrics, the geographic prediction of species relative density (from the two-step hurdle model) may nevertheless be useful for management purposes, but greater care should be taken in their interpretation and application. The geographic estimates of relative density were obtained by multiplying the probability of presence layer with that from the group size model in a two-step approach (Dedman et al. 2015). Thus, for those species with excellent estimates of probability of presence but poor species group size estimates, the relative density estimates will still represent an indication of likely hotspots (e.g., dusky dolphin). Conversely, where the probability of presence is not as well understood, estimates of relative density should be treated with caution (e.g., bottlenose dolphin, killer whale). Note, that there was no ability or attempt to distinguish between different killer whale or bottlenose dolphin ecotypes in the available sightings data. Different killer whale ecotypes will have distinct behaviours and habitat preferences, which were obscured in this analysis which assigned all killer whale sightings to a single category. The inability to distinguish distinct ecotypes within the data may account for the relatively poor performance of the killer whale model.

Across average year-round and seasonal probability of presences, and average year-round group size BRT models, species with widely distributed recorded locations had poorer model fits, perhaps reflecting the cosmopolitan distribution of these species (e.g., low explained deviance, AUC scores, Pearson's correlation for killer whale and bottlenose dolphin) and the more aggregated nature of others (e.g., high explained deviance and AUC scores for Māui dolphin, Hector's dolphin, Bryde's whale). Evidence from previous studies have indicated that species with limited geographic ranges and/or environmental tolerances are generally better modelled than those with greater ranges (Thomson et al. 2014, Morán-Ordóñez et al. 2017, Stephenson et al. 2018) because widespread species are less likely to have sharp easily identifiable environmental thresholds that clearly delineate their environmental niche (Morán-Ordóñez et al. 2017).

Additionally, reduced model fit could be influenced by historical events, human activities, population and species dynamics (e.g., migration, competition, predation, and for many cetacean species, social interactions) (Elith & Leathwick, 2009), and temporal environmental patterns (e.g., diurnal, tidal, annual patterns, fluctuating weather patterns, and prey distributions) which were not accounted for here. Despite these factors not being considered in a quantitative manner, model outputs are still valid for management purposes, but it should be noted that the representations of species probability of presence and relative densities presented here are smoothed representations of the raw data (spatially and temporally) (Stephenson et al. 2018). Further work on integrating other predictors not accounted for here (e.g., prey distribution, population dynamics, species co-occurrences, social interactions, etc.) may improve model outputs. Prey distribution is likely to be of particular importance for the accurate prediction of relative densities and seasonal environmental differences in distribution (although the environmental predictor variables may act as proxies for prey). Defining prey target species and producing prey distributions for each cetacean taxa was outside the scope of this work.

The final dataset contained over 14 000 cetacean sightings records, which were unequally distributed across both the study area and among species. This spatial sample bias was somewhat addressed here by the subsampling of non-random absence/background data (e.g., relative absences) and through bootstrapping of the BRT models, but the effects of spatial autocorrelation (i.e., the tendency for clustering of records spatially) were not specifically accounted for here. Future model iterations could explore weighting of species presence based on environmental or spatial coverage (e.g., Smith et al. 2013, Anderson et al. 2016, respectively) and/or include a spatial autocorrelation term in the model (e.g., calculation of a residual autocovariate (RAC) variable for each species as by Crase et al. 2012) to investigate whether the inclusion of these weightings or terms provides higher spatial predictive accuracy. However, with the opportunistically collected sightings data, where no information of effort or true absences are available, it is unclear whether these inclusions would significantly improve model performance or transferability.

Here, relative absences were used (i.e., the presence of another species for which the model is not being predicted was used as evidence that the species of interest was absent). However, a relative absence at a particular location does not preclude the presence of the modelled species at another unsampled point in time (or the species use of that habitat / high environmental suitability) since at-sea cetacean sighting effort is low, and sightability varies substantially between species (Würsig et al. 1998, Barlow et al. 2001). The relative absence data are still informative (because we can be fairly certain that the modelled species, if present at the surface at that point in time, would have been recorded). However, interpretation of model outputs would ideally need to take into account the probability of sightings, because the value predicted by the BRT models relates to the probability of sighting as well as the probability of actual presence. This information is difficult to estimate and is currently not available.

Ideally, true presence/absence with known effort (such as in NOMAD database) would be used for all future models, however, these data are currently only available for a limited number of species and primarily in shallow/inshore locations. There is on-going work to produce species distribution and density estimates for Māui and Hector's dolphins using NOMAD data with promising results to date (J. Roberts, pers. comm.).

5 MANAGEMENT IMPLICATIONS

In this report, a comprehensive set of spatial information for 31 cetacean species occurring in New Zealand waters was collated. This included information on: predicted species distributions for all species; and for those with sufficient data, predicted distribution of relative density. The quality of the spatial information varied depending on the method used (RES and BRT) as well as the availability of species sightings records. However, predicted distributions of all species modelled here, could be considered the best available information (with a greater level of spatial detail than the predictions produced using the more subjective spatial density analyses used for the most recent spatial risk assessment of the effects of fishing for cetacean species). These predicted spatial distributions are a significant step towards providing the evidence necessary for spatial risk assessments either as direct inputs or as indicative of areas at potential risk and requiring further investigation.

Here, estimates of model uncertainty were produced using two methods: estimates of individual species uncertainty of the distributions (through bootstrapping of the BRT models); and estimates of the coverage of the environmental space from all species recorded sightings. These uncertainty estimates should not be considered in isolation and provide two complementary measures to be considered by managers when examining the overlap of fishing with species distribution and relative density. The former, provides important indication of the variability in the modelling estimates. The latter provides an indication of which areas of the distribution of probability of presences and relative density are likely to have been extrapolated into unsampled space (i.e., where there is limited information to validate the predicted relationships). Spatially explicit estimates of uncertainty have been used to account for uncertainty in spatial planning applications (e.g., Moilanen & Wintle 2006, Moilanen et al. 2006) and could be applied using similar methods to spatially explicit fisheries risk assessment analyses.

For those species where only predicted distribution of presences were available (i.e., RES models and presence/relative absence BRT models), direct use in spatially explicit fisheries risk assessments may not be possible. However, overlapping species and fishing effort distributions may still provide an initial assessment of whether a given species may be at risk and whether further information / assessment is necessary. In addition, for those species not thought to be at risk by fishing (i.e., species with values less than 1 in Figure 1-1), predicted species distributions may provide relevant information for other potential anthropogenic impacts (e.g., Bryde's whales and risk of collision with vessels in the busy shipping lanes in the Hauraki Gulf).

For clarity, only two areas of interest were shown for each species distribution. However, all final spatial layers produced for this report will be made available on the MPI hosted and managed NABIS website which will allow closer inspection of areas of interest by all interested stakeholders and managers.

6 ACKNOWLEDGEMENTS

We acknowledge the contributions of data and advice from many sources including: the Department of Conservation for the provision and permission to use the at-sea cetacean sightings database and the marine mammal strandings database; Fisheries New Zealand for the use of the at-sea cetacean sightings database and the NOMAD dataset; the National Institute of Water and Atmospheric Research (NIWA) for physical and biological data layers. We thank the members of the Aquatic Environment Working Group for advice and input on the methods and results: particular thanks go to Anton van Helden (Forest & Bird), Rochelle Constantine (University of Auckland), Deanna Clement (Cawthron Institute), and Krista Hupman (NIWA) for their advice and expert opinion.

We would like to thank Ben Sharp (Fisheries New Zealand) for providing advice and guidance throughout this project, the anonymous reviewers for their feedback, and Suze Baird (Fisheries New Zealand) for feedback and editing. We thank MPI for funding this project under Objectives 1-3,

PRO201401. We also thank Wellington Regional Council for providing \$5k of additional funding to help support the habitat use modelling.

7. REFERENCES

- Abraham, E.R.; Neubauer, P.; Berkenbusch, K.; Richard, Y. (2017). Assessment of the risk to New Zealand marine mammals from commercial fisheries. *New Zealand Aquatic Environment and Biodiversity Report No. 189*. 123 p.
- Anderson, O.F.; Guinotte, J.M.; Rowden, A.A.; Tracey, D.M.; Mackay, K.A.; Clark, M.R. (2016). Habitat suitability models for predicting the occurrence of vulnerable marine ecosystems in the seas around New Zealand. *Deep Sea Research Part I: Oceanographic Research Papers* 115: 265–292.
- Baird, S.J.; Hewitt, J.E.; Wood, B.A. (2015). Benthic habitat classes and trawl fishing disturbance in New Zealand waters shallower than 250 m. *New Zealand Aquatic Environment and Biodiversity Report No. 144*. 184 p.
- Baker, C.S.; Chilvers, B.L.; Constantine, R.; Dufresne, S.; Mattlin, R.H.; Van Helden, A.; Hitchmough, R. (2010). Conservation status of New Zealand marine mammals (suborders Cetacea and Pinnipedia), 2009. *New Zealand Journal of Marine and Freshwater Research* 44: 101–115.
- Barlow, J.; Gerodette, T.; Forcada, J. (2001). Factors affecting perpendicular sighting distances on shipboard line-transect surveys for cetaceans. *Journal of Cetacean Research and Management* 3: 201–212.
- Behrenfeld, M.J.; Falkowski, P.G. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and oceanography* 42: 1–20.
- Bluhm, B.; Coyle, K.; Konar, B.; Highsmith, R. (2007). High gray whale relative abundances associated with an oceanographic front in the south-central Chukchi Sea. *Deep Sea Research Part II: Topical Studies in Oceanography* 54: 2919–2933.
- Bradford-Grieve, J.; Probert, K.; Lewis, K.; Sutton, P.; Zeldis, J.; Orpin, A. (2006). *New Zealand shelf region*. In: Robinson A.R.; Brink, K.H. (eds.), pp. 1451–1492, *The sea*. Cambridge, MA, Harvard University Press.
- Brager, S.; Schneider, K. (1998). Near-shore distribution and abundance of dolphins along the West Coast of the South Island, New Zealand. *New Zealand Journal of Marine and Freshwater Research* 32: 105–112.
- Calenge, C. (2006). The package “adehabitat” for the R software: A tool for the analysis of space and habitat use by animals. *Ecological Modelling* 197: 516–519.
- Compton, T.J.; Morrison, M.A.; Leathwick, J.R.; Carbines, G.D. (2012). Ontogenetic habitat associations of a demersal fish species, *Pagrus auratus*, identified using boosted regression trees. *Marine Ecology Progress Series* 462: 219–230.
- Crane, B.; Liedloff, A.C.; Wintle, B.A. (2012). A new method for dealing with residual spatial autocorrelation in species distribution models. *Ecography* 35: 879–888.
- Dedman, S.; Officer, R.; Brophy, D.; Clarke, M.; Reid, D.G. (2015). Modelling abundance hotspots for data-poor Irish Sea rays. *Ecological modelling* 312: 77–90.
- Derville, S.; Constantine, R.; Baker, C.S.; Oremus, M.; Torres, L.G. (2016). Environmental correlates of nearshore habitat distribution by the Critically Endangered Māui dolphin. *Marine Ecology Progress Series* 551: 261–275.
- Dwyer, S.; Clement, D.; Pawley, M.; Stockin, K. (2016). Distribution and relative density of cetaceans in the Hauraki Gulf, New Zealand. *New Zealand Journal of Marine and Freshwater Research* 50: 457–480.
- Elith, J.; Graham, C.H.; Anderson, R.P.; Dudík, M.; Ferrier, S.; Guisan, A.; Hijmans, R.J.; Huettmann, F.; Leathwick, J.R.; Lehmann, A.; Li, J.; Lohmann, L.G.; Loiselle, B.A.; Manion, G.; Moritz, C.; Nakamura, M.; Nakazawa, Y.; Overton, J. McC. M.; Peterson, A.T.; Phillips, S.J.; Richardson, K.; Scachetti-Pereira, R.; Schapire, R.E.; Soberón, J.; Williams, S.; Wisz, M.S.; Zimmermann, N.E. (2006). Novel methods improve prediction of species’ distributions from occurrence data. *Ecography* 29: 129–151.

- Elith, J.; Leathwick, J.R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40: 677–697.
- Elith, J.; Leathwick, J.R.; Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology* 77: 802–813.
- Etnoyer, P.; Canny, D.; Mate, B.R.; Morgan, L.E.; Ortega-Ortiz, J.G.; Nichols, W.J. (2006). Sea-surface temperature gradients across blue whale and sea turtle foraging trajectories off the Baja California Peninsula, Mexico. *Deep Sea Research Part II: Topical Studies in Oceanography* 53: 340–358.
- Friedman, J.; Hastie, T.; Tibshirani, R. (2001). *The Elements of Statistical Learning*. Springer Series in Statistics New York.
- Gorman, R.M.; Bryan, K.R.; and Laing, A.K. (2003). Wave hindcast for the New Zealand region: Deep-water wave climate. *New Zealand Journal of Marine and Freshwater Research* 37, 589-612.
- Guisan, A.; and Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology letters* 8, 993-1009.
- Guisan, A.; Tingley, R.; Baumgartner, J.B.; Naujokaitis-Lewis, I.; Sutcliffe, P.R.; Tulloch, A.I.T.; Regan, T.J.; Brotons, L.; McDonald-Madden, E.; Mantyka-Pringle, C.; Martin, T.G.; Rhodes, J.R.; Maggini, R.; Setterfield, S.A.; Elith, J.; Schwartz, M.W.; Wintle, B.A.; Broennimann, O.; Austin, M.; Ferrier, S.; Kearney, M.R.; Possingham, H.P.; and Buckley, Y.M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters* 16, 1424-1435.
- Guisan, A.; and Zimmermann, N.E. (2000). Predictive habitat distribution models in ecology. *Ecological modelling* 135, 147-186.
- Hijmans, R.J.; Phillips, S.; Leathwick, J.; and Elith, J. (2017). dismo: Species Distribution Modeling R package version 1.1-4. <https://CRAN.R-project.org/package=dismo>.
- Hijmans, R.J.; and Van Etten, J. (2012). raster: Geographic analysis and modeling with raster data. R package version 2.0-12. Available at: <http://CRAN.R-project.org/package=raster>.
- Hirzel, A.H.; Le Lay, G.; Helfer, V.; Randin, C.; and Guisan, A. (2006). Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling* 199, 142-152.
- Kaschner, K.; Tittensor, D.P.; Ready, J.; Gerrodette, T.; and Worm, B. (2011). Current and future patterns of global marine mammal biodiversity. *PLoS ONE* 6(5): e19653. doi:10.1371/journal.pone.0019653
- Kaschner, K.; Watson, R.; Trites, A.; and Pauly, D. (2006). Mapping world-wide distributions of marine mammal species using a relative environmental suitability (RES) model. *Marine Ecology Progress Series* 316, 285-310.
- Kesner-Reyes, K.; Kaschner, K.; Kullander, S.; Garilao, C.; Barile, J.; and Froese, R. (2016). *AquaMaps: algorithm and data sources for aquatic organisms*. . www.fishbase.org, version (04/2012).
- Leathwick, J.; Elith, J.; Francis, M.; Hastie, T.; and Taylor, P. (2006). Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Marine Ecology Progress Series* 321, 267-281.
- Leathwick, J.; Rowden, A.; Nodder, S.; Gorman, R.; Bardsley, S.; Pinkerton, M.; Baird, S.J.; Hadfield, M.; Currie, K.; Goh, A. (2012). A Benthic-optimised Marine Environment Classification (BOMEC) for New Zealand waters. *New Zealand Aquatic Environment and Biodiversity Report No. 88*. 54 p.
- Lonergan, M.E.; Phillips, R.A.; Thomson, R.B.; Zhou, S. (2017). Independent review of New Zealand's Spatially Explicit Fisheries Risk Assessment approach – 2017. *New Zealand Fisheries Science Review* 2017/2. 36 p.
- Mitchell, J.S.; Mackay, K.A.; Neil, H.L.; Mackay, E.J.; Pallentin, A.; and Notman, P. (2012). Undersea New Zealand, 1:5,000,000. *NIWA Chart Miscellaneous Series No. 92*.
- Moilanen, A.; Wintle, B.A. (2006). Uncertainty analysis favours selection of spatially aggregated reserve networks. *Biological Conservation* 129: 427–434.
- Moilanen, A.; Wintle, B.A.; Elith, J.; Burgman, M. (2006). Uncertainty analysis for regional-scale reserve selection. *Conservation Biology* 20; 1688–1697.

- Morán-Ordóñez, A.; Lahoz-Monfort, J.J.; Elith, J.; Wintle, B.A. (2017). Evaluating 318 continental-scale species distribution models over a 60-year prediction horizon: what factors influence the reliability of predictions? *Global Ecology and Biogeography* 26: 371–384.
- Pinkerton, M.; Gall, M.; Wood, S.; Zeldis, J. (2018). Measuring the effects of bivalve mariculture on water quality in northern New Zealand using 15 years of MODIS-Aqua satellite observations. *Aquaculture Environment Interactions* 10: 529–545.
- R Core Team (2013). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.
- Ready, J.; Kaschner, K.; South, A.B.; Eastwood, P.D.; Rees, T.; Rius, J.; Agbayani, E.; Kullander, S.; Froese, R. (2010). Predicting the distributions of marine organisms at the global scale. *Ecological Modelling* 221: 467–478.
- Richard, Y.; Abraham, E.R. (2013). Risk of commercial fisheries to New Zealand seabird populations. *New Zealand Aquatic Environment and Biodiversity Report No. 109*. 58 p.
- Ridgeway, G. (2007). Generalized Boosted Models: A guide to the gbm package.
- Sharp, B.R. (2017). Spatially-explicit fisheries risk assessment (SEFRA): a framework for quantifying and managing incidental commercial fisheries impacts on non-target species. pp. 20–56. In: *Aquatic Environment and Biodiversity Annual Review 2017*. Ministry for Primary Industries. <https://www.mpi.govt.nz/dmsdocument/27471-aquatic-environment-and-biodiversity-annual-review-aebar-2017-a-summary-of-environmental-interactions-between-the-seafood-sector-and-the-aquatic-environment>
- Silverman, B.W. (1986). Density estimation for statistics and data analysis. *Monographs on Statistics and Applied Probability* 26. London: Chapman and Hall.
- Smith, A.N.; Duffy, C.a.J.; Leathwick, J.R. (2013). *Predicting the distribution and relative abundance of fishes on shallow subtidal reefs around New Zealand*. Science for Conservation 323. Department of Conservation, Wellington. 25 p. + 2 supplements.
- Stephenson, F.; Leathwick, J.R.; Geange, S.W.; Bulmer, R.H.; Hewitt, J.E.; Anderson, O.F.; Rowden, A.A.; Lundquist, C.J. (2018). Using Gradient Forests to summarize patterns in species turnover across large spatial scales and inform conservation planning. *Diversity and Distributions* 24:1641–1656.
- Stephenson, F.; Polunin, N.V.; Mill, A.C.; Scott, C.; Lightfoot, P.; Fitzsimmons, C. (2017). Spatial and temporal changes in pot-fishing effort and habitat use. *ICES Journal of Marine Science* 74(8): 2201–2212.
- Stockin, K.A.; Binedell, V.; Wiseman, N.; Brunton, D.H.; Orams, M.B. (2009). Behavior of free-ranging common dolphins (*Delphinus* sp.) in the Hauraki Gulf, New Zealand. *Marine Mammal Science* 25: 283–301.
- Thompson, K.F.; Millar, C.D.; Baker, C.S.; Dalebout, M.; Steel, D.; Van Helden, A.L.; Constantine, R. (2013). A novel conservation approach provides insights into the management of rare cetaceans. *Biological conservation* 157: 331–340.
- Thomson, R.J.; Hill, N.A.; Leaper, R.; Ellis, N.; Pitcher, C.R.; Barrett, N.S.; Edgar, G.J. (2014). Congruence in demersal fish, macroinvertebrate, and macroalgal community turnover on shallow temperate reefs. *Ecological Applications* 24: 287–299.
- Tynan, C.T.; Ainley, D.G.; Barth, J.A.; Cowles, T.J.; Pierce, S.D.; Spear, L.B. (2005). Cetacean distributions relative to ocean processes in the northern California Current System. *Deep Sea Research Part II: Topical studies in Oceanography* 52: 145–167.
- Walters, R.A.; Goring, D.G.; Bell, R.G. (2001). Ocean tides around New Zealand. *New Zealand Journal of Marine and Freshwater Research* 35: 567–579.
- Wand, M.P.; Jones, M.C. (1995). Kernel Smoothing. *Monographs on Statistics and Applied Probability* 60. Chapman and Hall, London.
- Wiseman, N.; Parsons, S.; Stockin, K.A.; Baker, C.S. (2011). Seasonal occurrence and distribution of Bryde's whales in the Hauraki Gulf, New Zealand. *Marine Mammal Science* 27(4). E253–E267.
- Worton, B.J. (1989). Kernel methods for estimating the utilization distribution in home-range studies. *Ecology* 70: 164–168.
- Würsig, S.K.L.; Jefferson, T.A.; Mullin, K.D. (1998). survey ships and aircraft. *Aquatic Mammals* 24: 41–50.

APPENDIX 1: ENVIRONMENTAL PREDICTOR VARIABLES

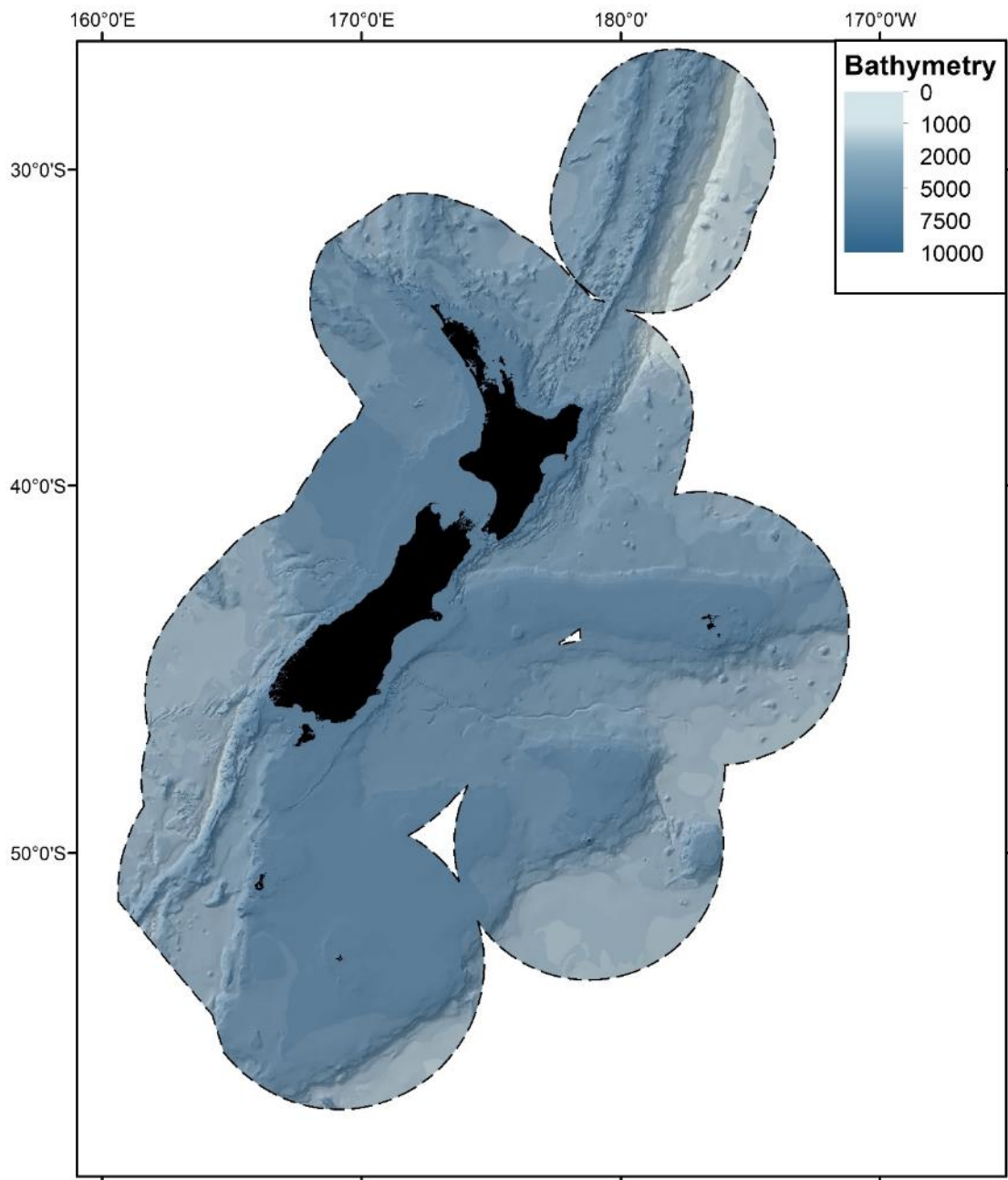


Figure 42: Bathymetry (Bathy) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

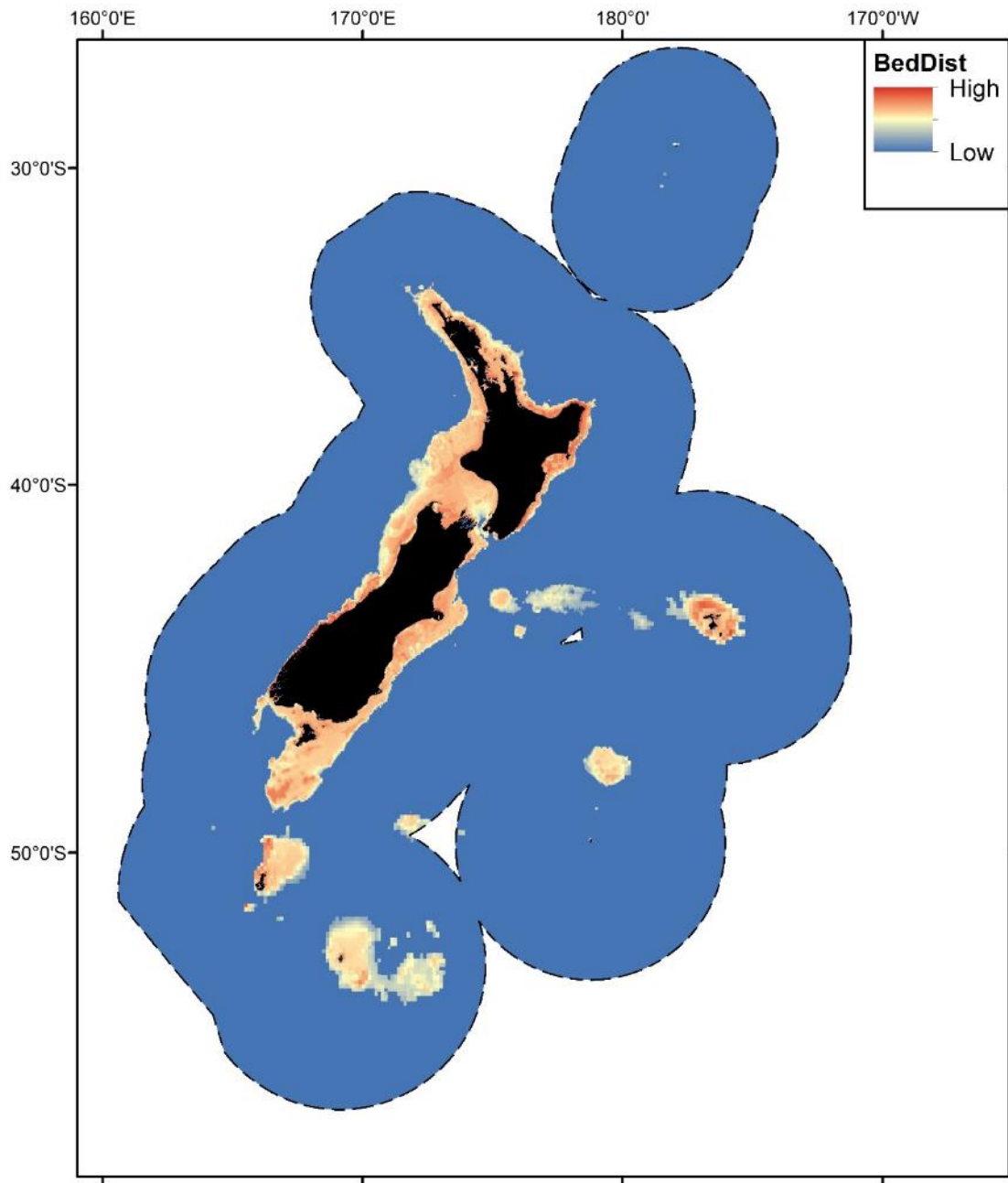


Figure 43: Benthic sediment disturbance (BedDist) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

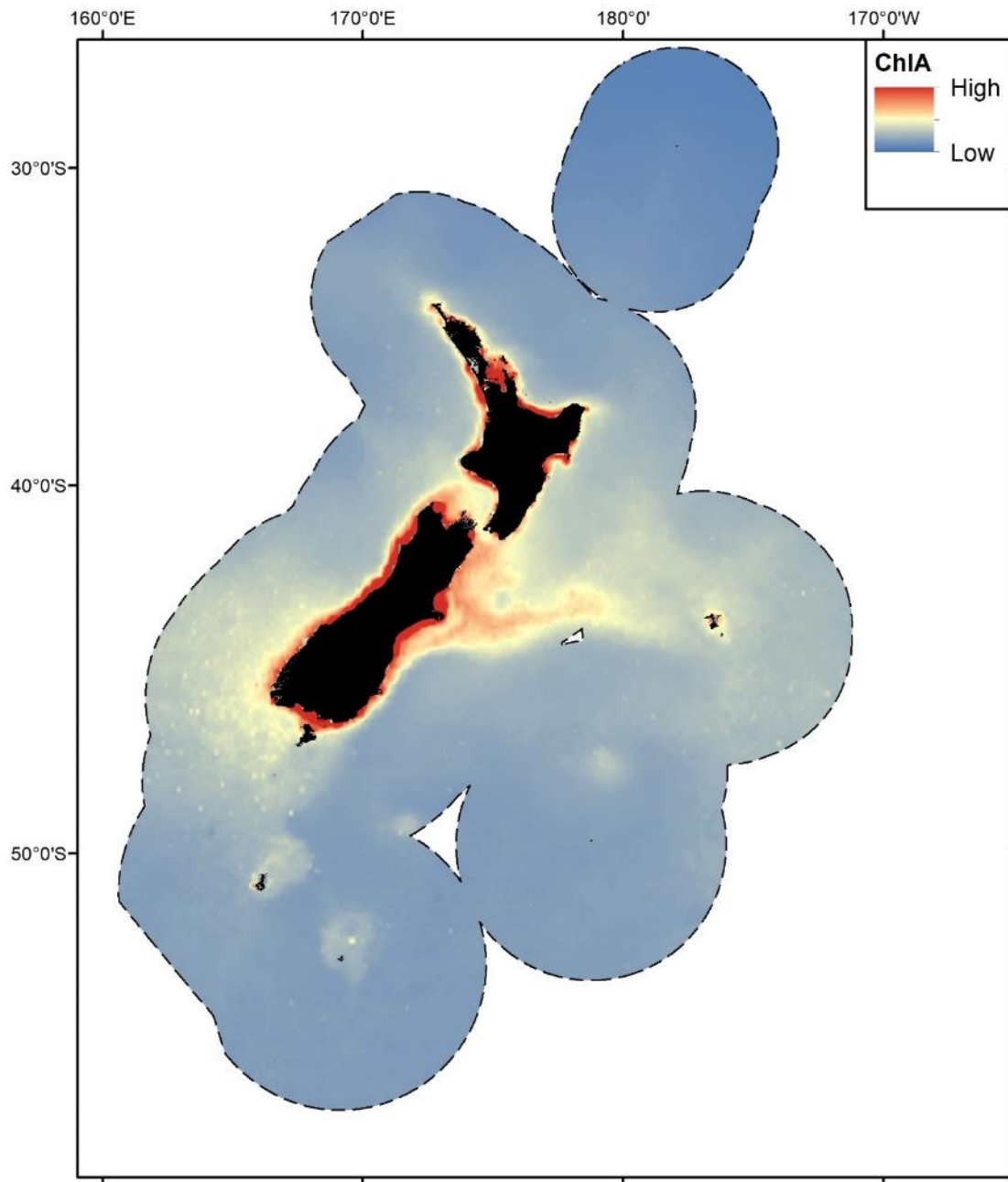


Figure 44: Annual mean chlorophyll-a concentration (ChlA) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

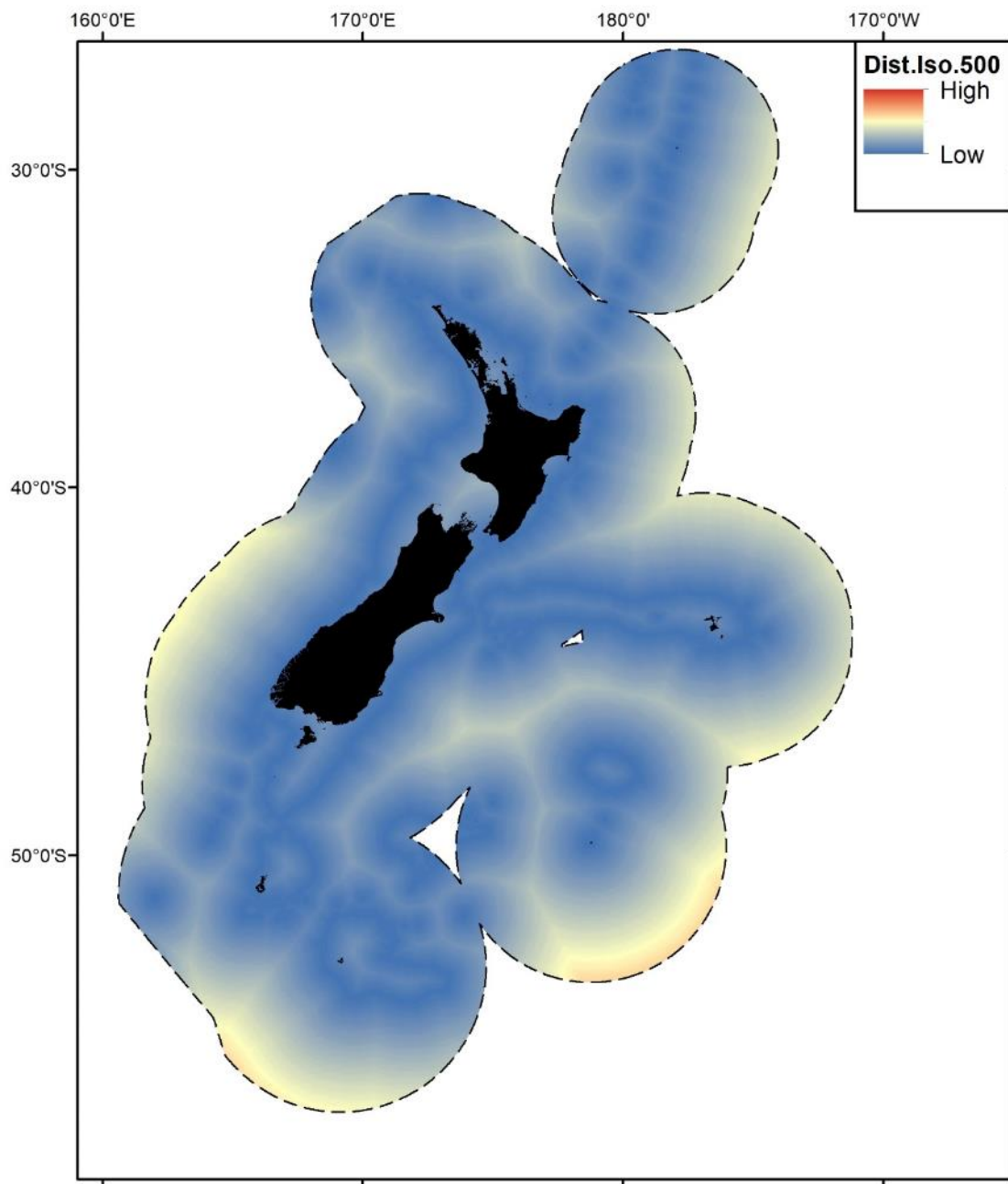


Figure 45: Distance to 500 m isobath (Dist.Iso.500) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

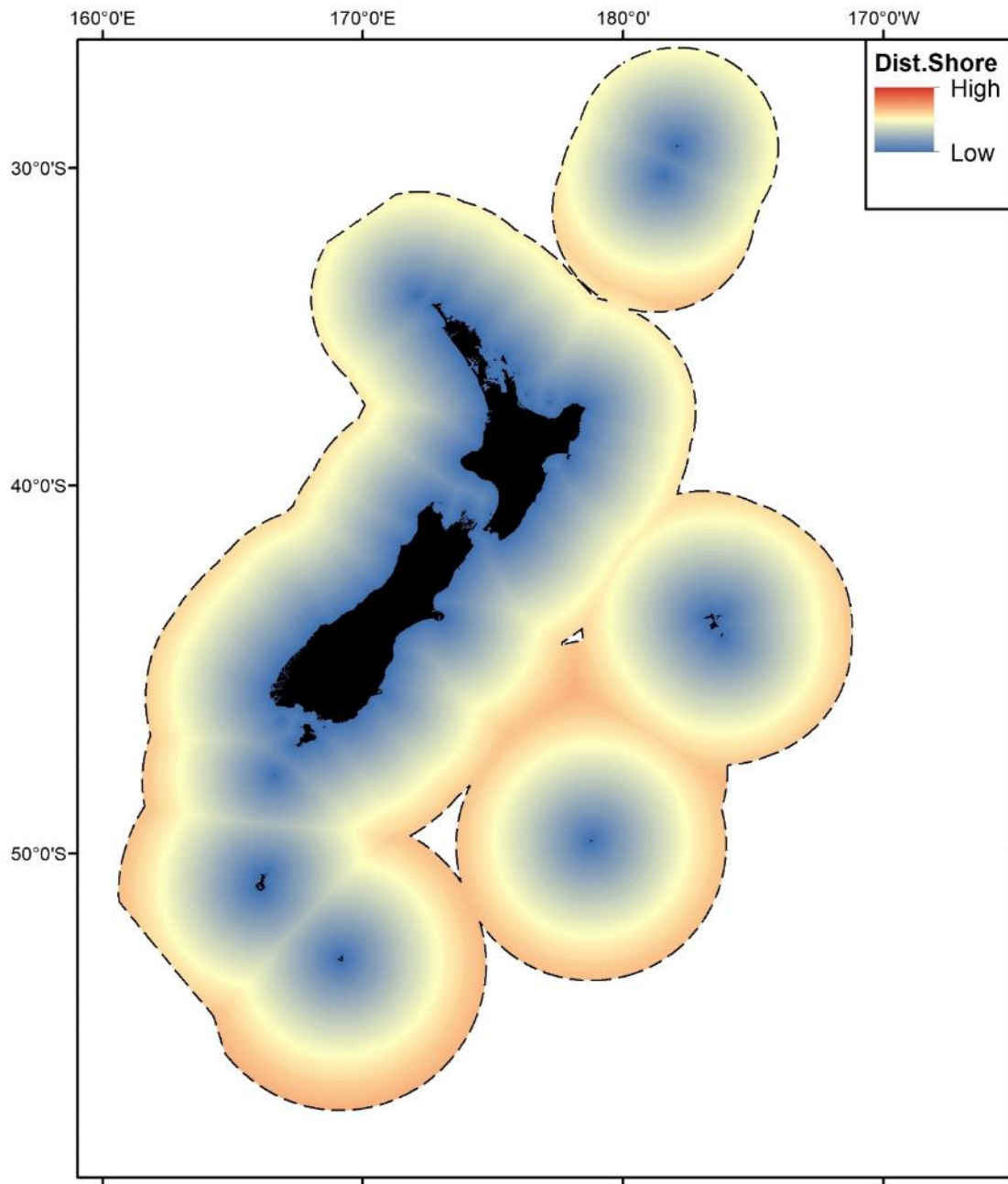


Figure 46: Distance to shore (Dist.Shore) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

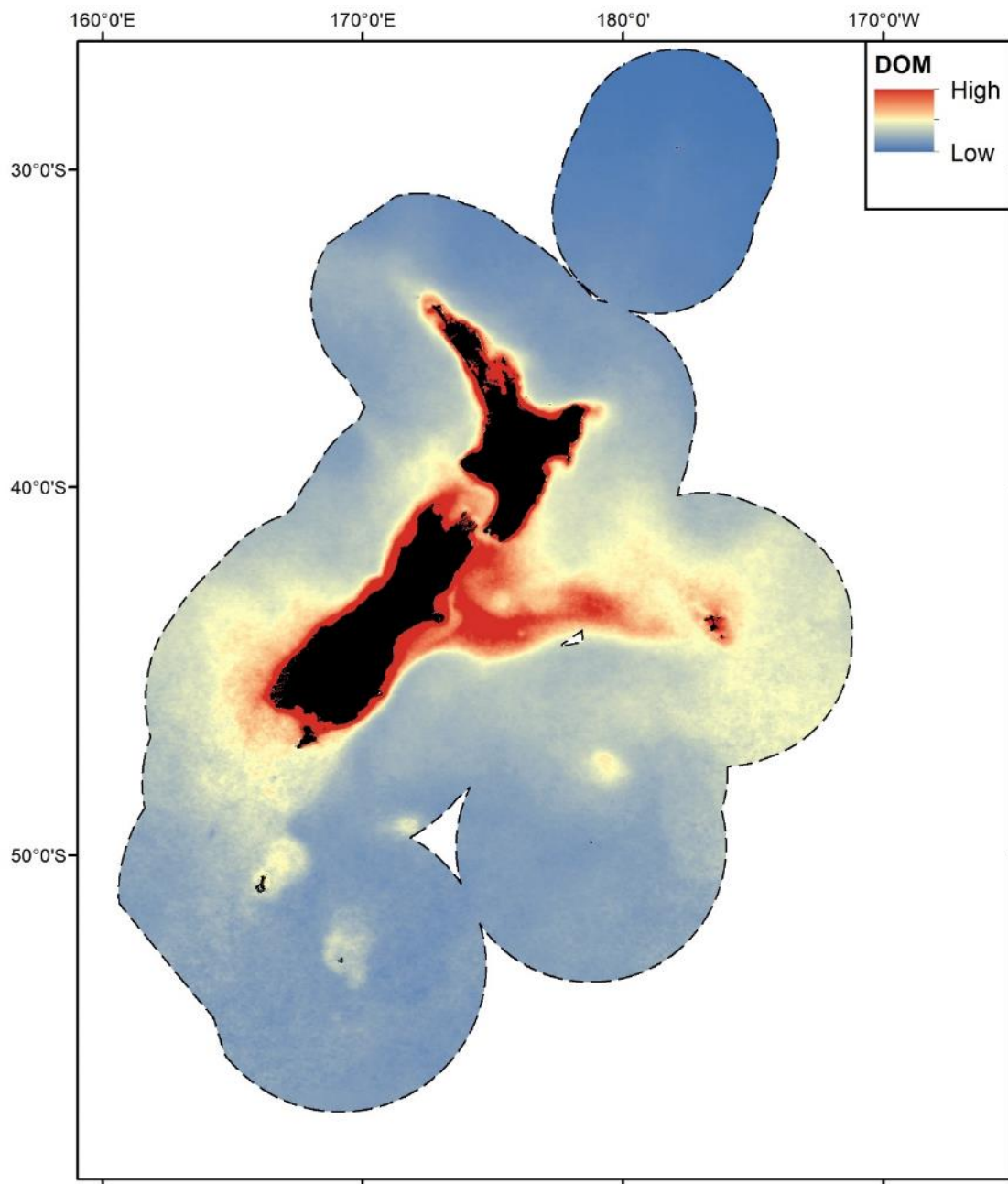


Figure 47: Coloured dissolved organic matter (DOM) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

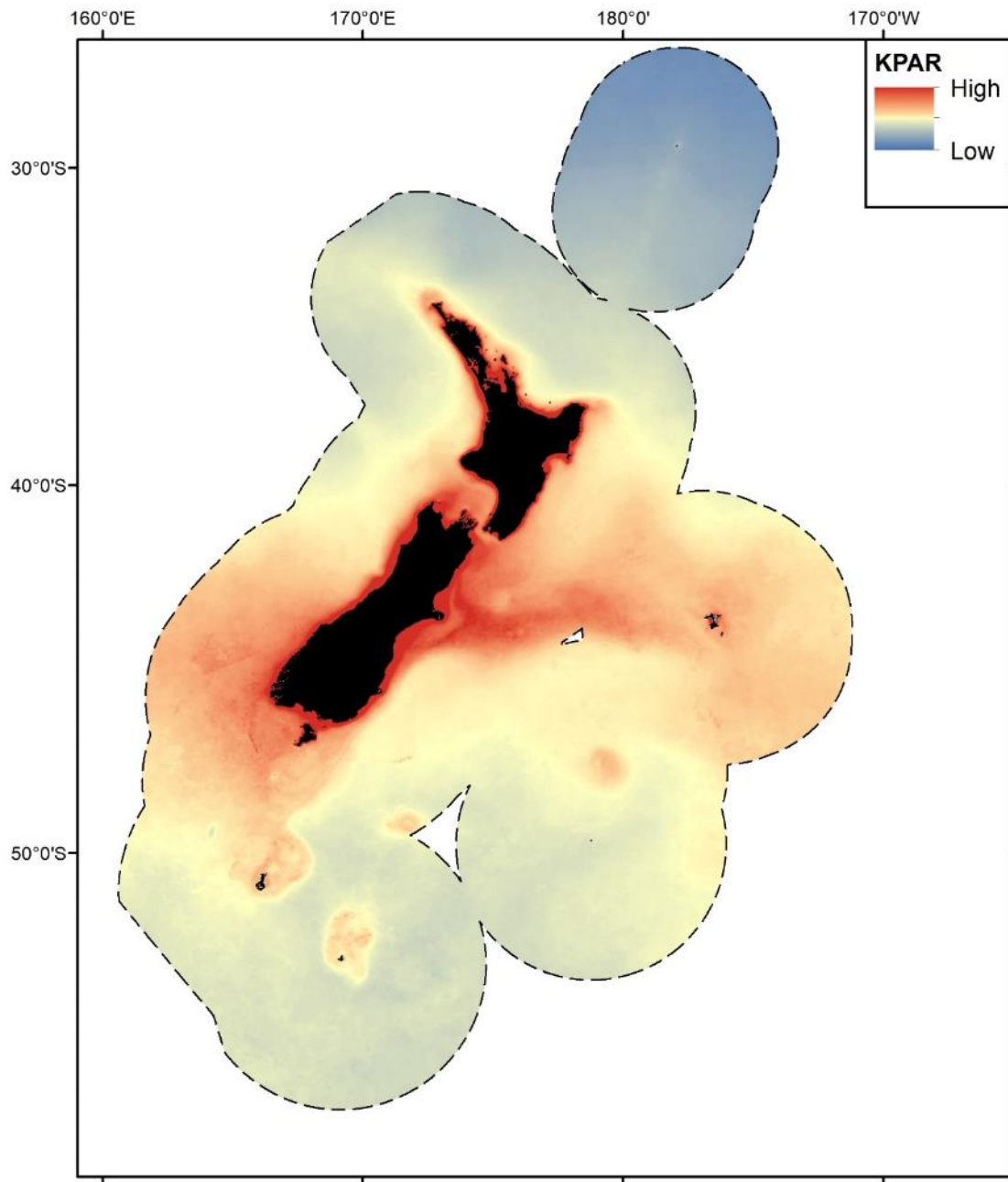


Figure 48: Annual mean KPAR (Kpar) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

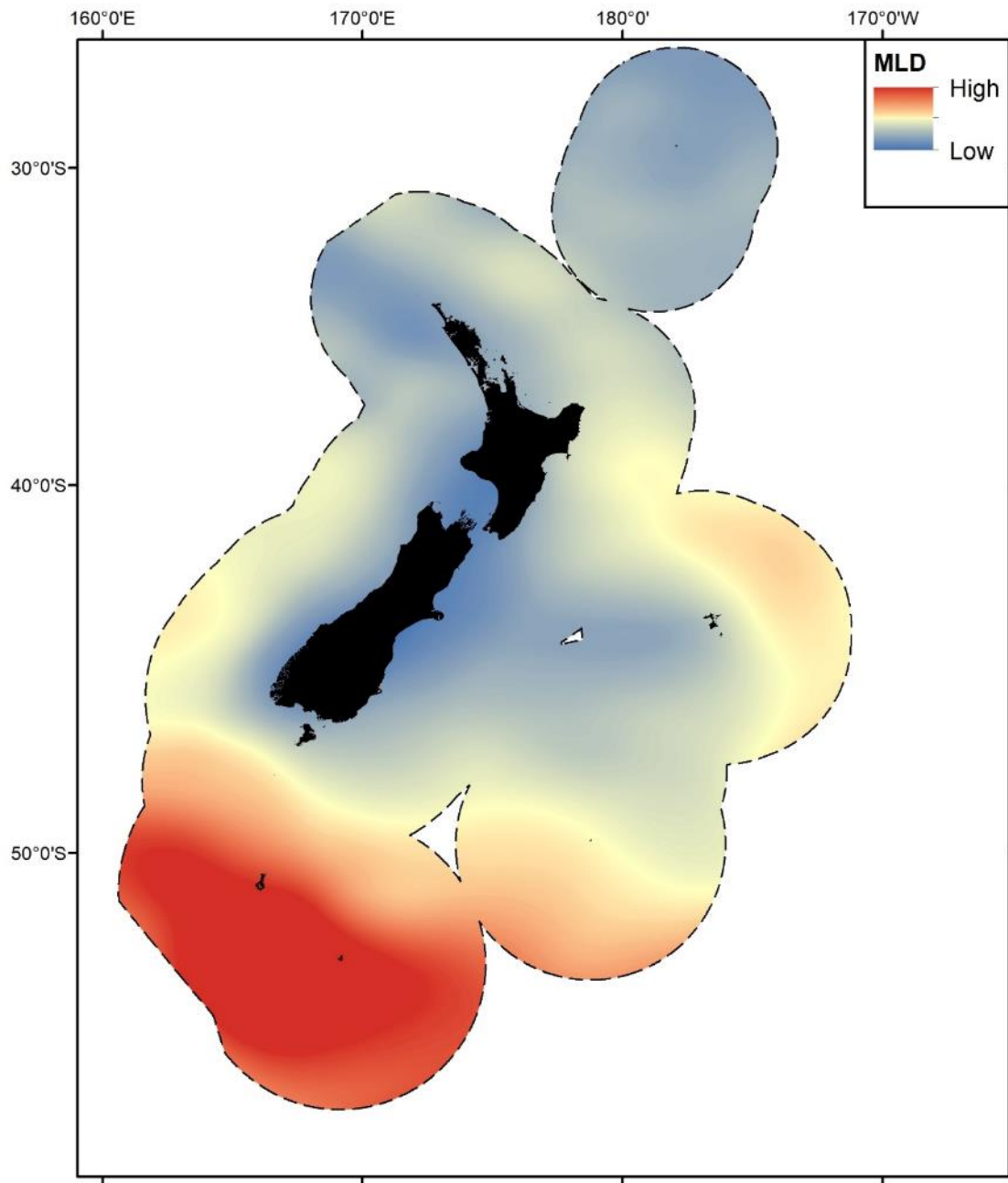


Figure 49: Annual mean mixed layer depth (MLD) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

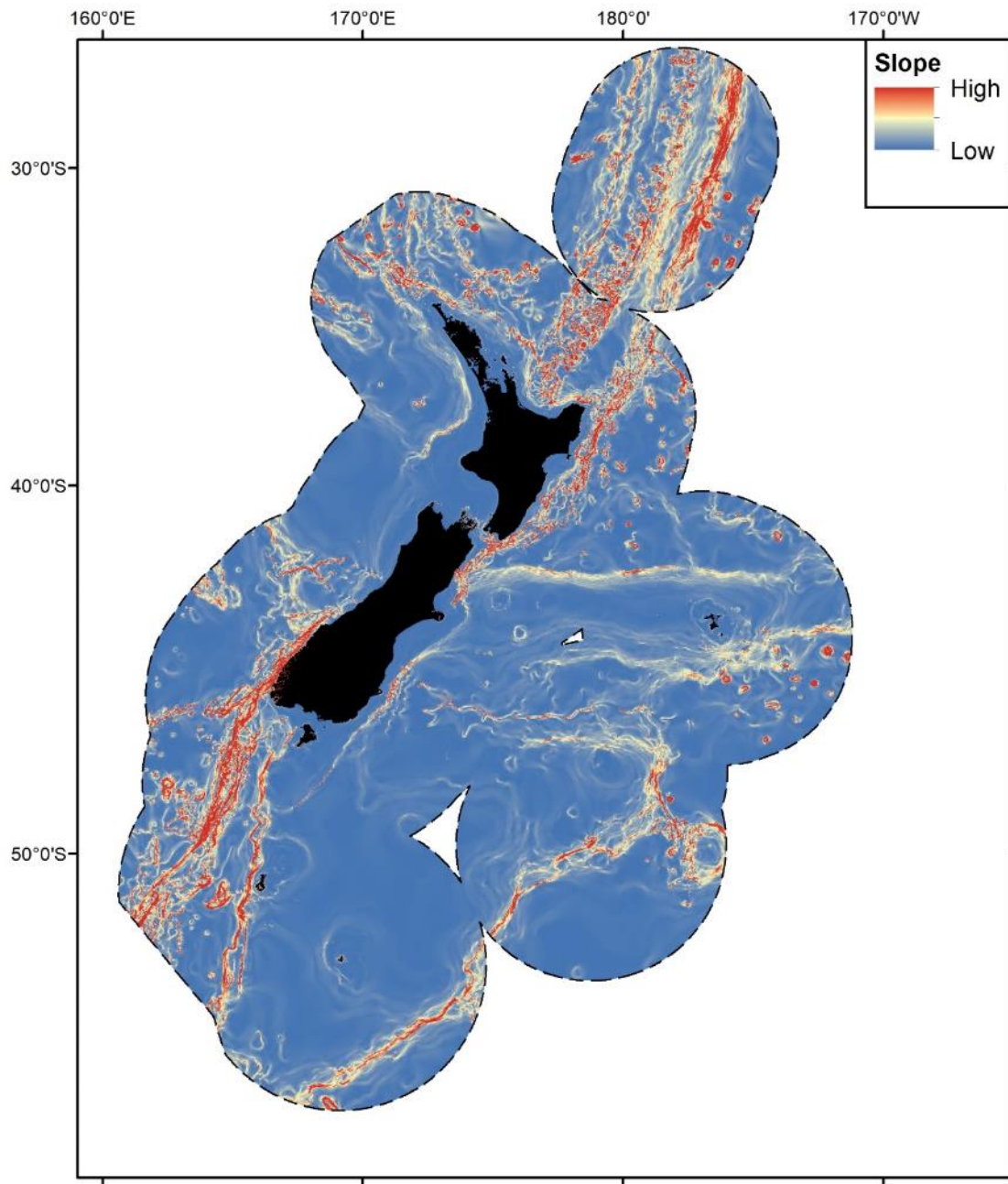


Figure 50: Slope (Slope) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

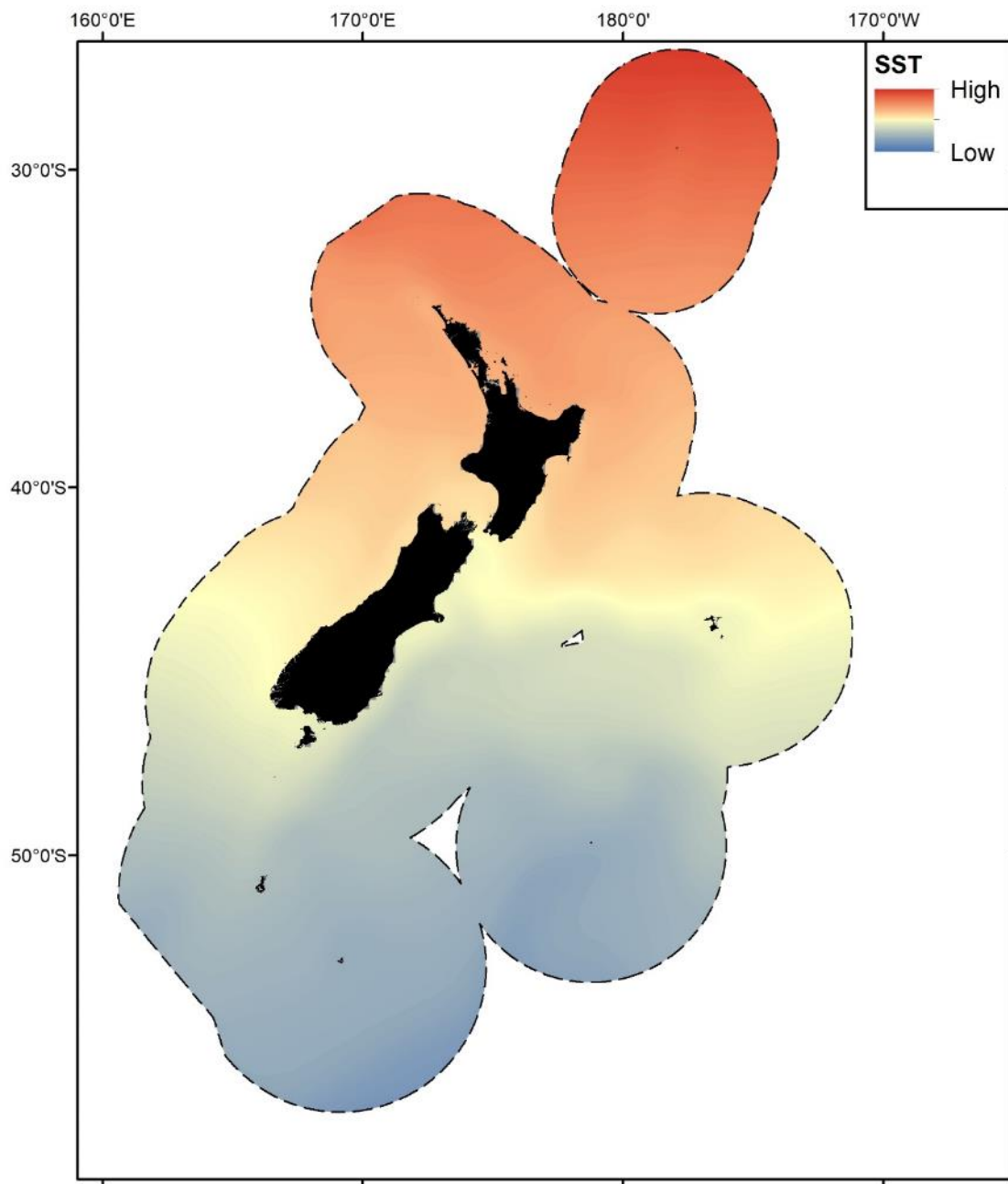


Figure 51: Annual mean sea surface temperature (SST) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

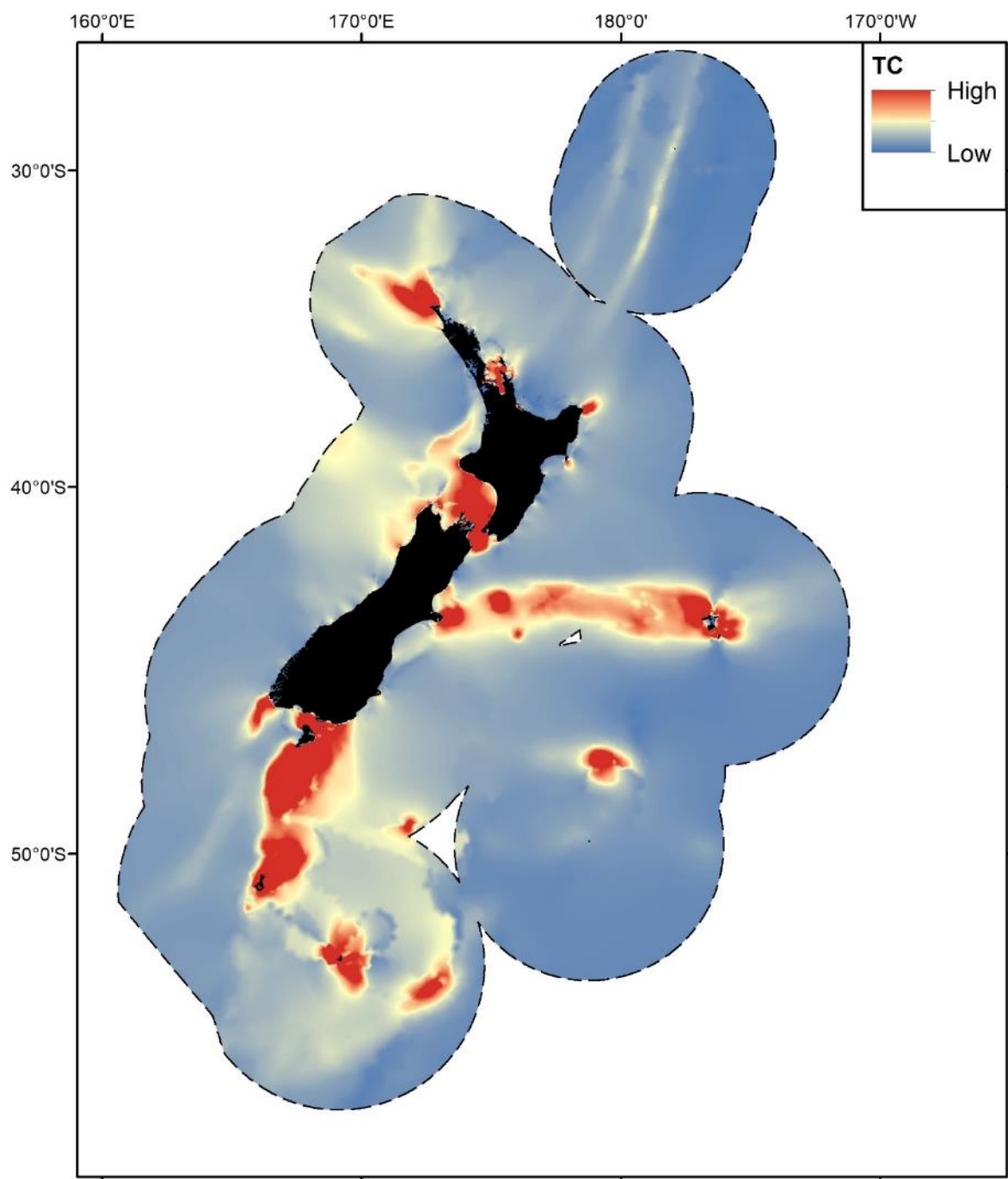


Figure 52: Tidal current speed (TC) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

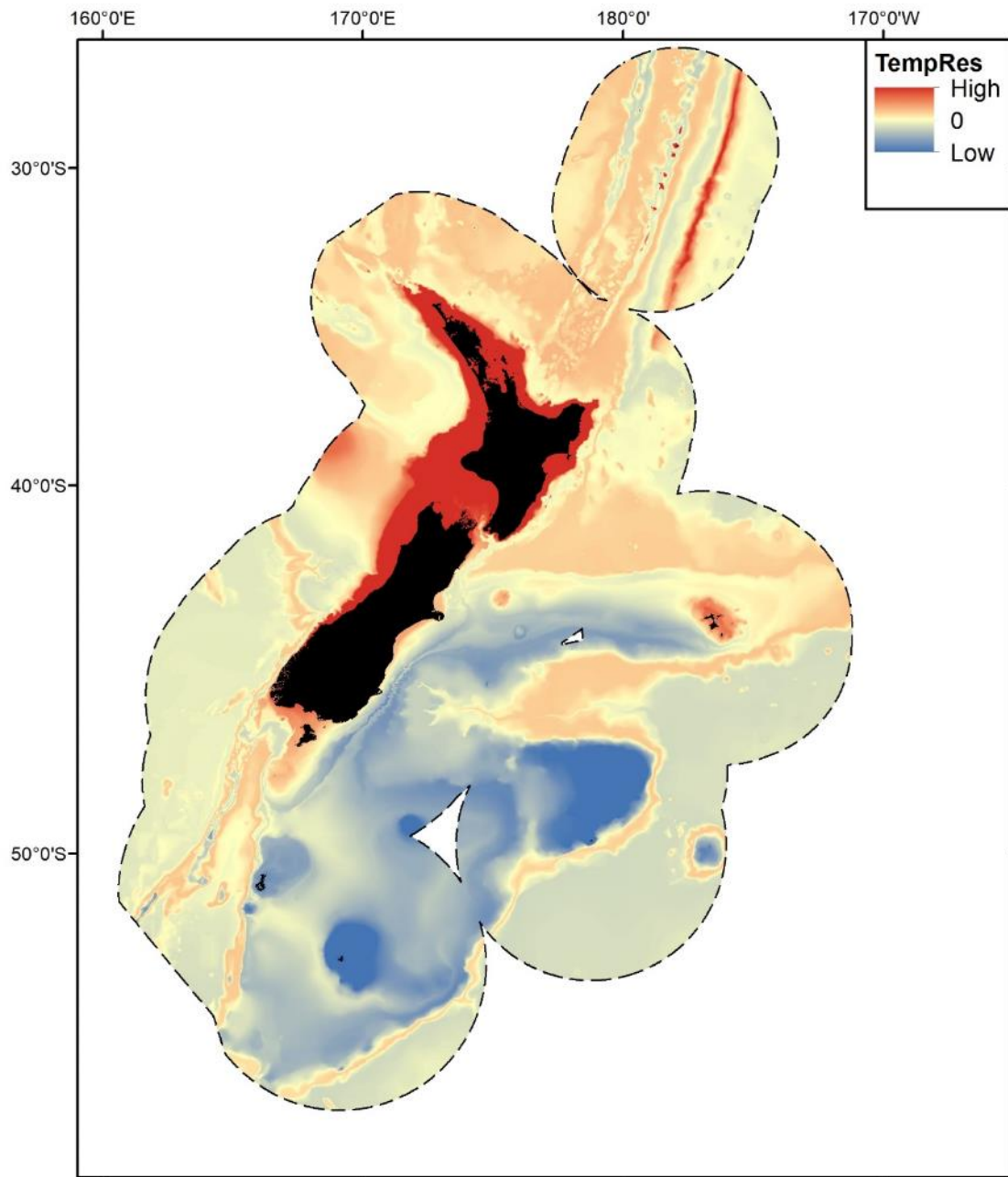


Figure 53: Temperature residuals (TempRes) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

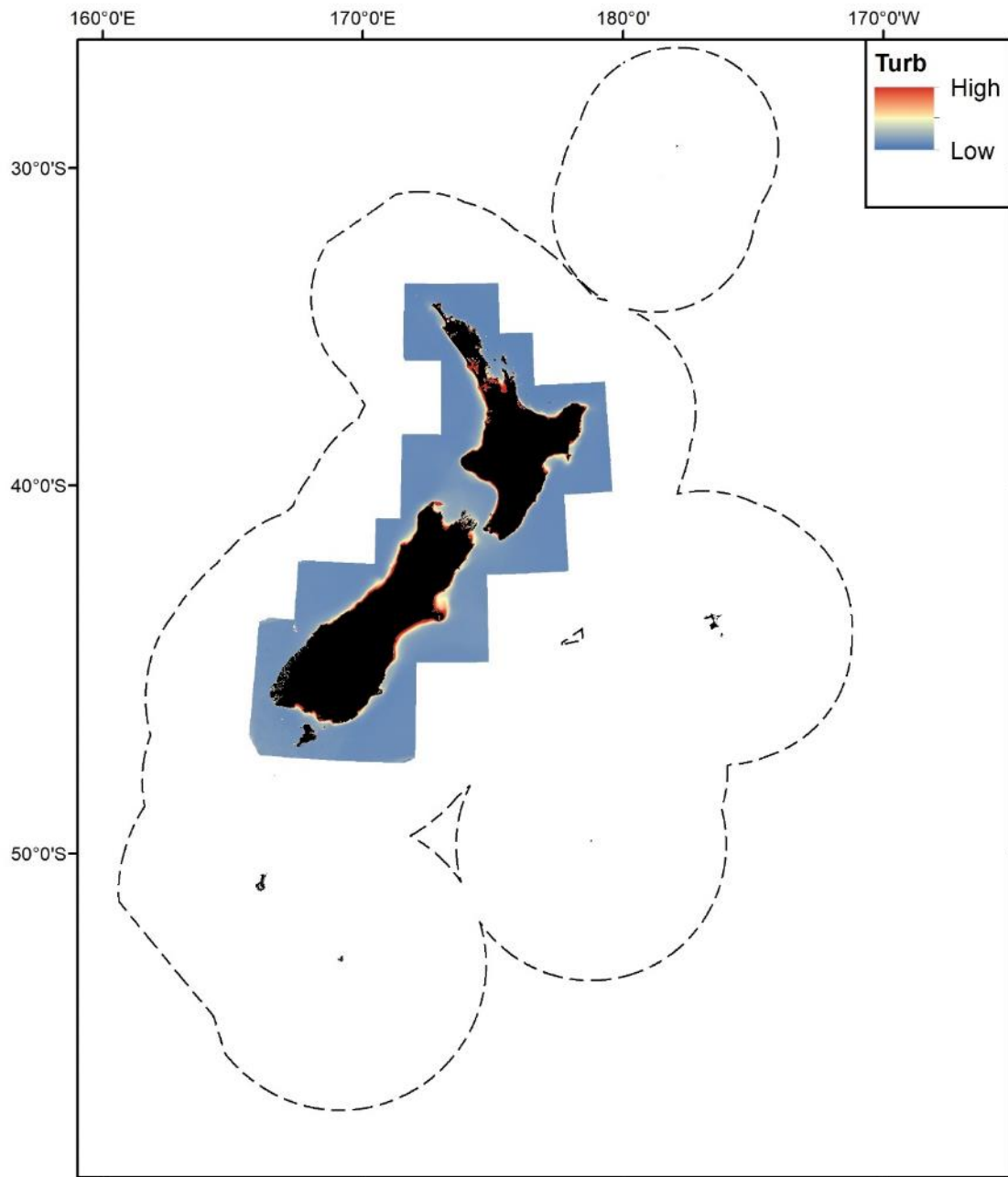


Figure 54: Annual mean turbidity (Turb) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

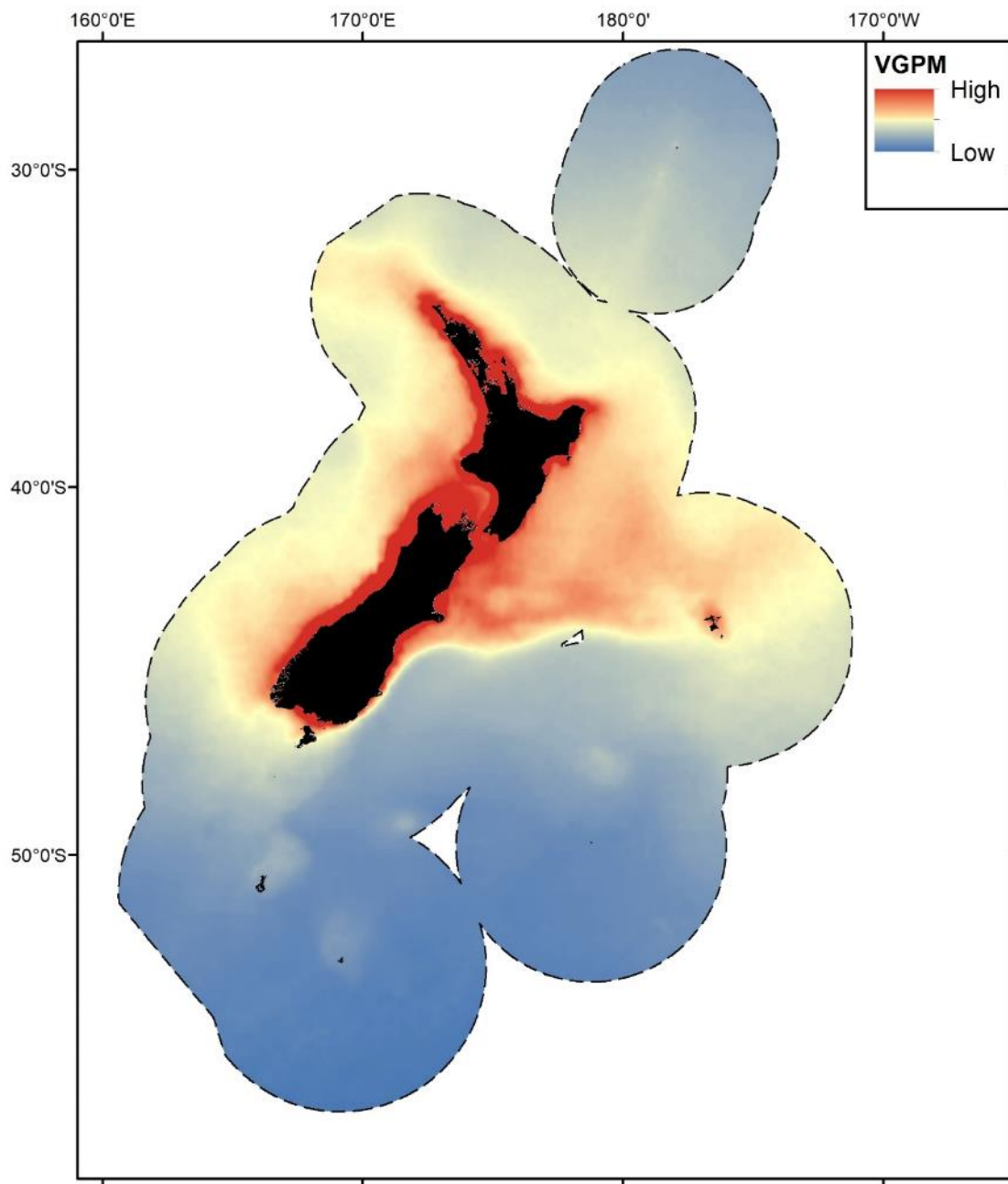


Figure 55: Annual mean productivity model (VGPM) within the study region (New Zealand Exclusive Economic Zone (EEZ), black dashed line).

APPENDIX 2: RES MODEL OUTPUTS

Table 26: RES values for variable Sea Surface Temperature for cetacean species with low numbers of observation records.

Common name	Species name	Sea Surface Temperature (degrees C)				Description
		Min	Preferred min	Preferred max	Max	
Andrew's beaked whale	<i>Mesoplodon bowdoini</i>	5	10	20	25	Cold temperate - subtropical
Arnoux's beaked whale	<i>Berardius arnuxii</i>	-2	0	20	25	Polar-subtropical
Blainville's beaked whale	<i>Mesoplodon densirostris</i>	10	15	30	35	Warm temperate - full tropical
Cuvier's beaked whale	<i>Ziphius cavirostris</i>	5	10	30	35	Cold temperate - full tropical
Dwarf minke whale	<i>Balaenoptera acutorostrata</i>	-2	0	25	30	Polar - tropical
False killer whale	<i>Pseudorca crassidens</i>	10	15	30	35	Warm temperate - full tropical
Gray's beaked whale	<i>Mesoplodon grayi</i>	0	5	20	25	Subpolar - subtropical
Hourglass dolphin	<i>Lagenorhynchus cruciger</i>	-2	0	15	20	Polar-warm temperate
Pygmy sperm whale	<i>Kogia breviceps</i>	10	15	30	35	Warm temperate - full tropical
Risso's dolphin	<i>Grampus griseus</i>	5	10	30	35	Cold temperate - full tropical
Shepherd's beaked whale	<i>Tasmacetus shepherdi</i>	0	5	15	20	Subpolar - warm temperate
Short finned pilot whale	<i>Globicephala macrorhynchus</i>	15	20	25	30	Subtropical - tropical
Short-beaked common dolphin	<i>Delphinus delphis</i>	5	10	30	35	Cold temperate - full tropical
Southern bottlenose whale	<i>Hyperoodon planifrons</i>	-2	0	20	25	Polar - subtropical
Southern right whale dolphin	<i>Lissodelphis peronii</i>	-2	0	20	25	Polar - subtropical
Spectacled porpoise	<i>Phocoena dioptrica</i>	-2	0	10	15	Polar - cold temperate
Striped dolphin	<i>Stenella coeruleoalba</i>	5	10	25	30	Cold temperate - tropical

Table 27: RES values for variable Bathymetry for cetacean species with low numbers of observation records.

Common name	Species name	Bathymetry (m)				Description
		Min	Preferred min	Preferred max	Max	
Andrew's beaked whale	<i>Mesoplodon bowdoini</i>	0	-200	-2000	-8000	Mainly continental slope to very deep water
Arnoux's beaked whale	<i>Berardius arnuxii</i>	0	-200	-2000	-8000	Mainly continental slope to very deep water
Blainville's beaked whale	<i>Mesoplodon densirostris</i>	0	-200	-2000	-8000	Mainly continental slope to very deep water
Cuvier's beaked whale	<i>Ziphius cavirostris</i>	0	-200	-2000	-8000	Mainly continental slope to very deep water
Dwarf minke whale	<i>Balaenoptera acutorostrata</i>	0	-10	-2000	-8000	Mainly coast. Continental slope to v deep water
False killer whale	<i>Pseudorca crassidens</i>	0	-1000	-2000	-8000	Mainly low. Continental slope to v deep water
Gray's beaked whale	<i>Mesoplodon grayi</i>	0	-200	-2000	-8000	Mainly continental slope to very deep water
Hourglass dolphin	<i>Lagenorhynchus cruciger</i>	0	-1000	-4000	-8000	Mainly low continental slope - abyss plains to v deep
Pygmy sperm whale	<i>Kogia breviceps</i>	0	-200	-2000	-8000	Mainly continental slope to very deep water
Risso's dolphin	<i>Grampus griseus</i>	0	-200	-1000	-6000	Mainly up continental slope to deep water
Shepherd's beaked whale	<i>Tasmacetus shepherdi</i>	0	-1000	-2000	-8000	Mainly low. Continental slope to v deep water
Short finned pilot whale	<i>Globicephala macrorhynchus</i>	0	-200	-2000	-8000	Mainly continental slope to very deep water
Short-beaked common dolphin	<i>Delphinus delphis</i>	0	-200	-2000	-8000	Mainly continental slope to very deep water
Southern bottlenose whale	<i>Hyperoodon planifrons</i>	0	-1000	-4000	-8000	Mainly low continental slope - abyss plains to v deep
Southern right whale dolphin	<i>Lissodelphis peronii</i>	0	-200	-2000	-8000	Mainly continental slope to very deep water
Spectacled porpoise	<i>Phocoena dioptrica</i>	0	-10	-2000	-8000	Mainly coast. Continental slope to v deep water
Striped dolphin	<i>Stenella coeruleoalba</i>	0	-200	-2000	-8000	Mainly continental slope to very deep water

Table 28: RES values for variable Distance to Shore for cetacean species with low numbers of observation records.

Common name	Species name	Distance to Shore (km)				Description
		Min	Preferred min	Preferred max	Max	
Andrew's beaked whale	<i>Mesoplodon bowdoini</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Arnoux's beaked whale	<i>Berardius arnuxii</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Blainville's beaked whale	<i>Mesoplodon densirostris</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Cuvier's beaked whale	<i>Ziphius cavirostris</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Dwarf minke whale	<i>Balaenoptera acutorostrata</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
False killer whale	<i>Pseudorca crassidens</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Gray's beaked whale	<i>Mesoplodon grayi</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Hourglass dolphin	<i>Lagenorhynchus cruciger</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Pygmy sperm whale	<i>Kogia breviceps</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Risso's dolphin	<i>Grampus griseus</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Shepherd's beaked whale	<i>Tasmacetus shepherdi</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Short finned pilot whale	<i>Globicephala macrorhynchus</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Short-beaked common dolphin	<i>Delphinus delphis</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Southern bottlenose whale	<i>Hyperoodon planifrons</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Southern right whale dolphin	<i>Lissodelphis peronii</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Spectacled porpoise	<i>Phocoena dioptrica</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species
Striped dolphin	<i>Stenella coeruleoalba</i>	10	100	1000	4000	Oceanic - continental shelf/pelagic species

Andrew's beaked whale (*Mesoplodon bowdoini*)

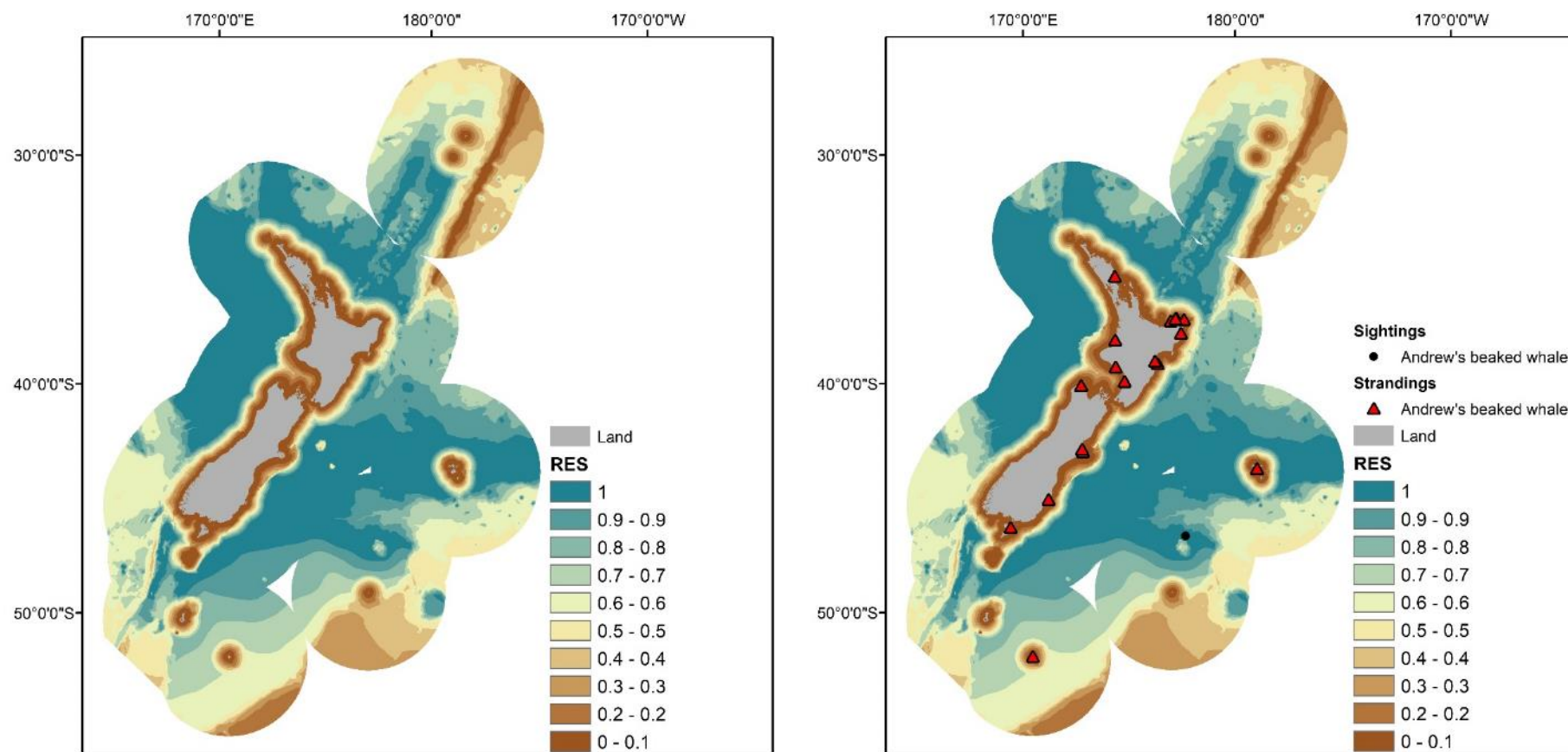


Figure 59: Predicted RES scores for Andrew's beaked whale (*Mesoplodon bowdoini*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Arnoux's beaked whale (*Berardius arnuxii*)

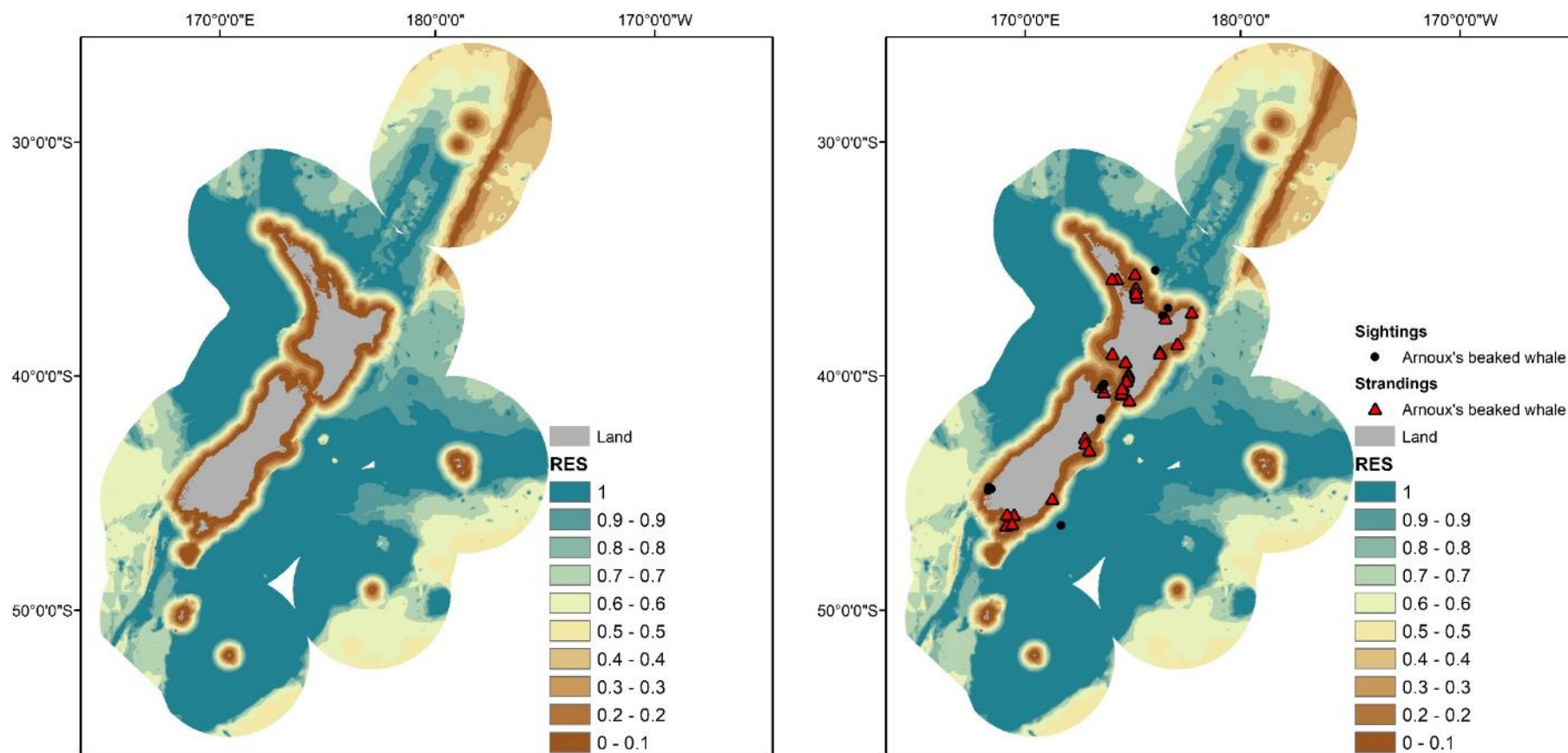


Figure 60: Predicted RES scores for Arnoux's beaked whale (*Berardius arnuxii*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Blainville's beaked whale (*Mesoplodon densirostris*)

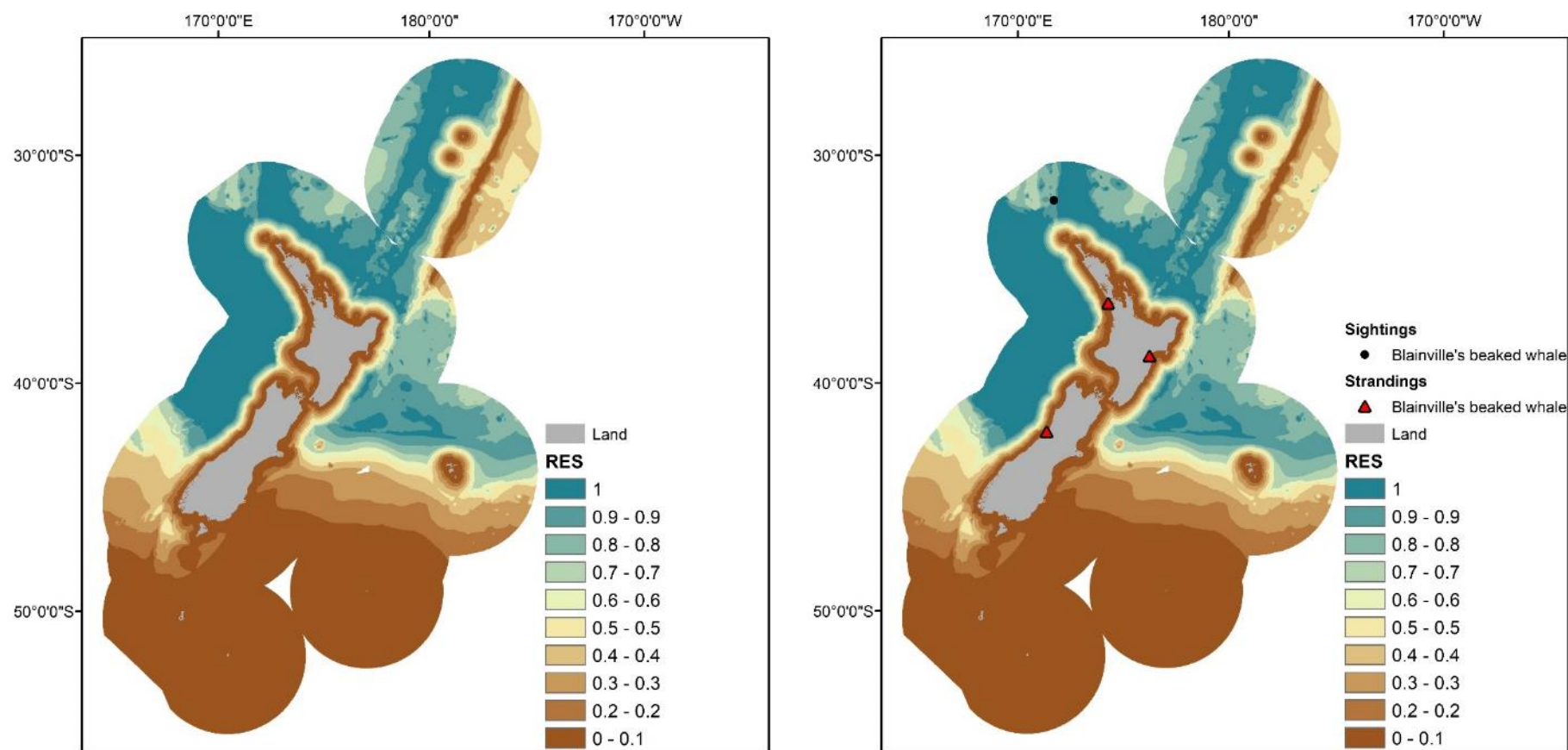


Figure 61: Predicted RES scores for Blainville's beaked whale (*Mesoplodon densirostris*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Cuvier's beaked whale (*Ziphius cavirostris*)

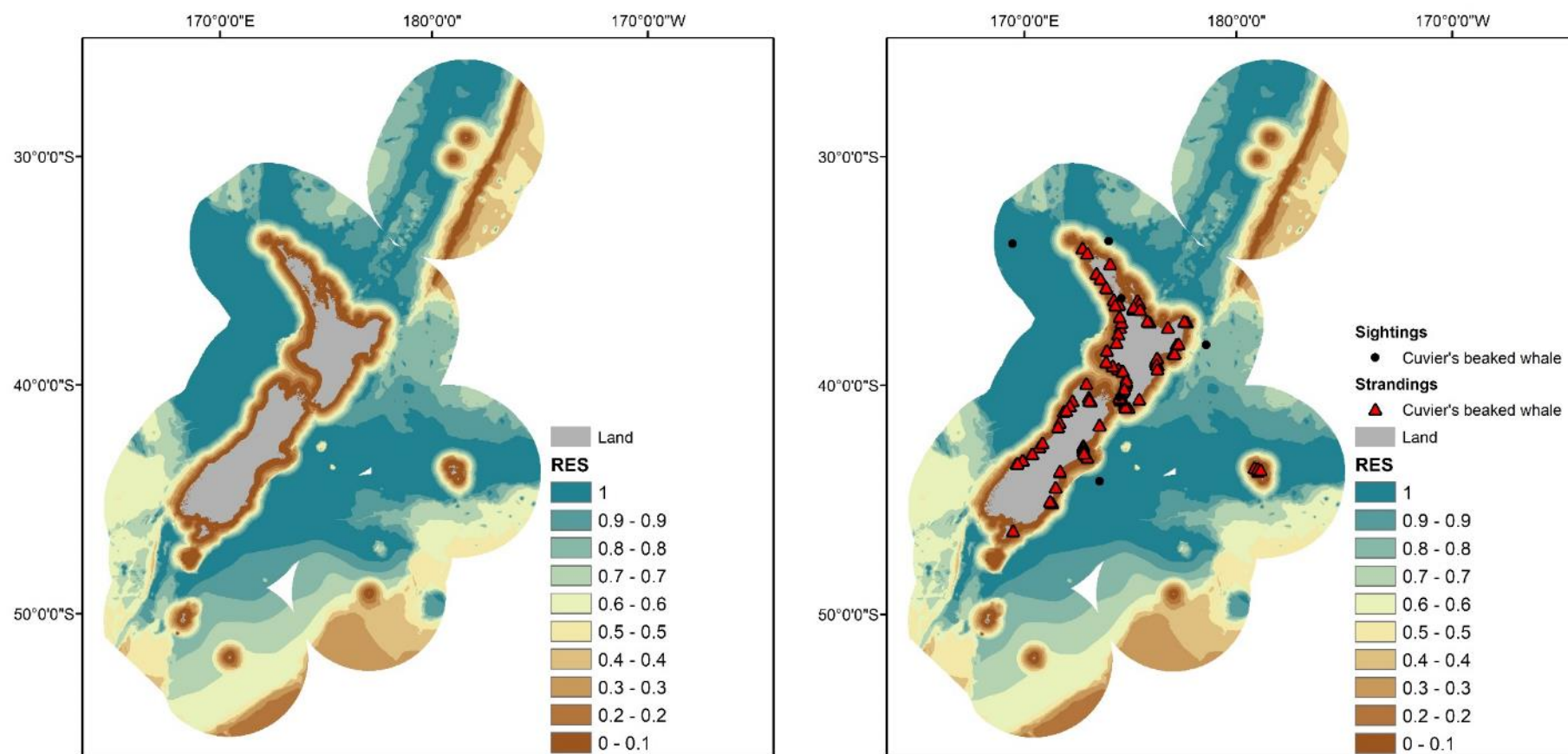


Figure 62: Predicted RES scores for Cuvier's beaked whale (*Ziphius cavirostris*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Dwarf minke whale (*Balaenoptera acutorostrata*)

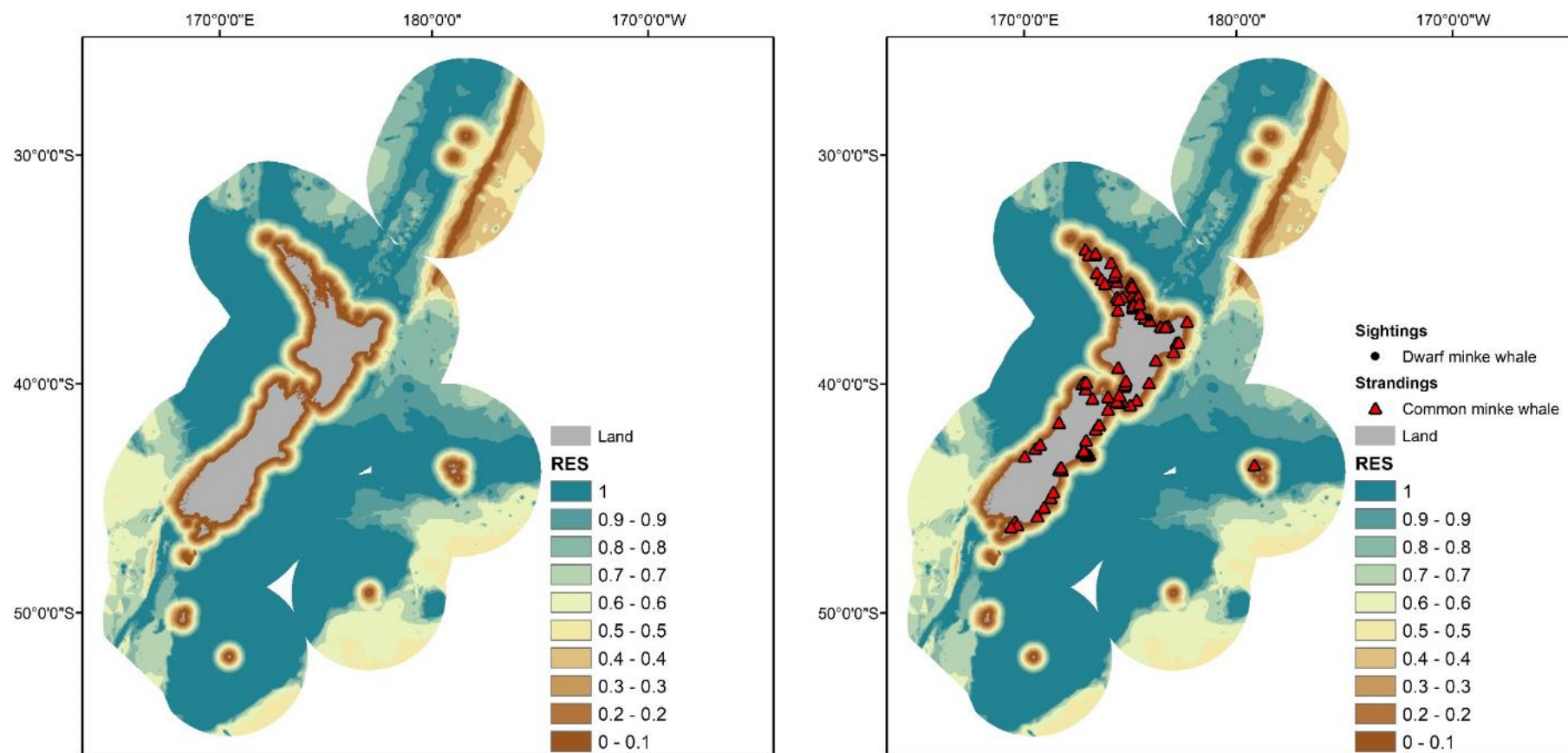


Figure 63: Predicted RES scores for dwarf minke whale (*Balaenoptera acutorostrata*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Arnoux's beaked Whale (*Berardius arnuxii*)

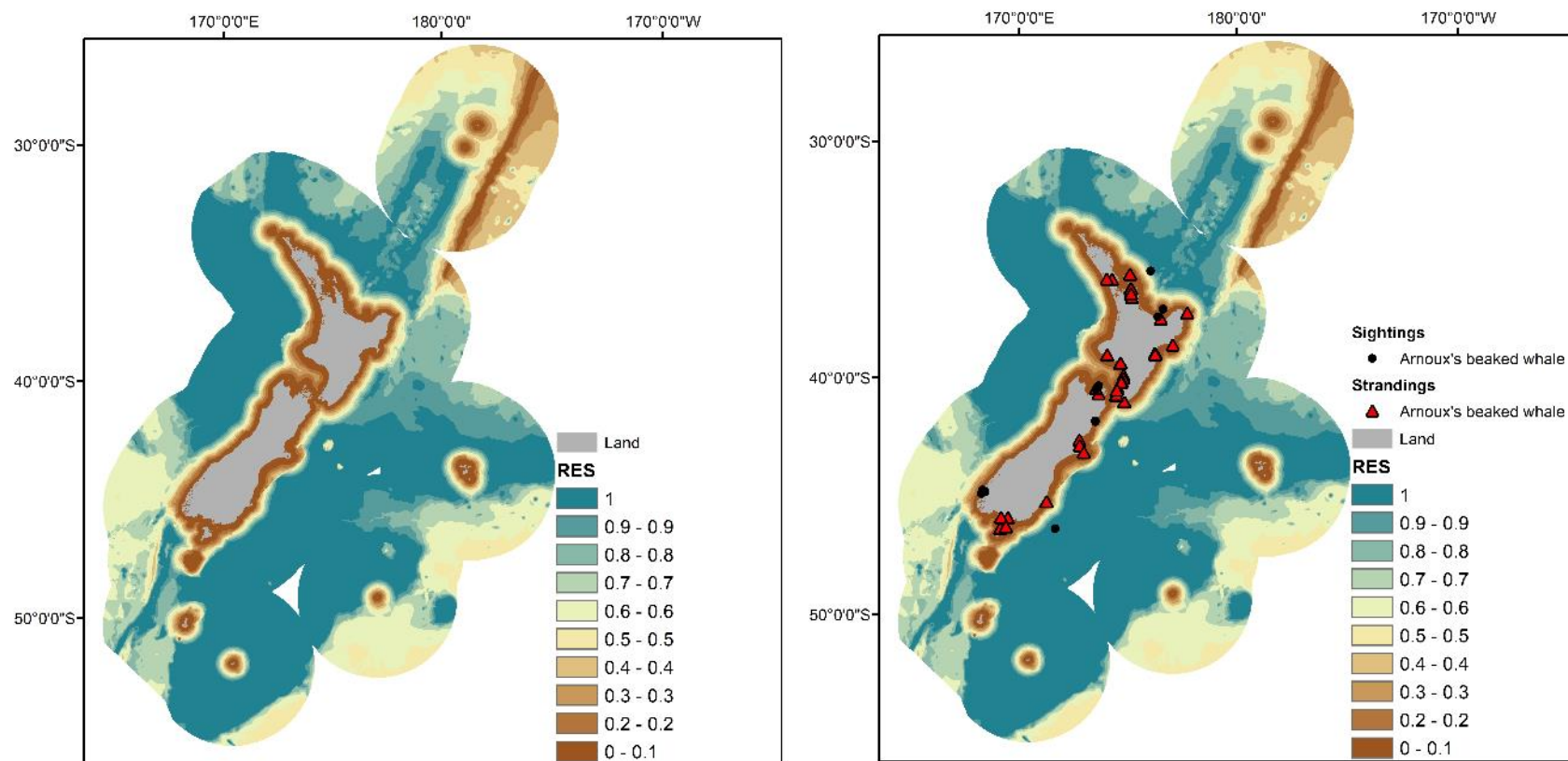


Figure 64: Predicted RES scores for Arnoux's beaked Whale (*Berardius arnuxii*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Gray's beaked whale (*Mesoplodon grayi*)

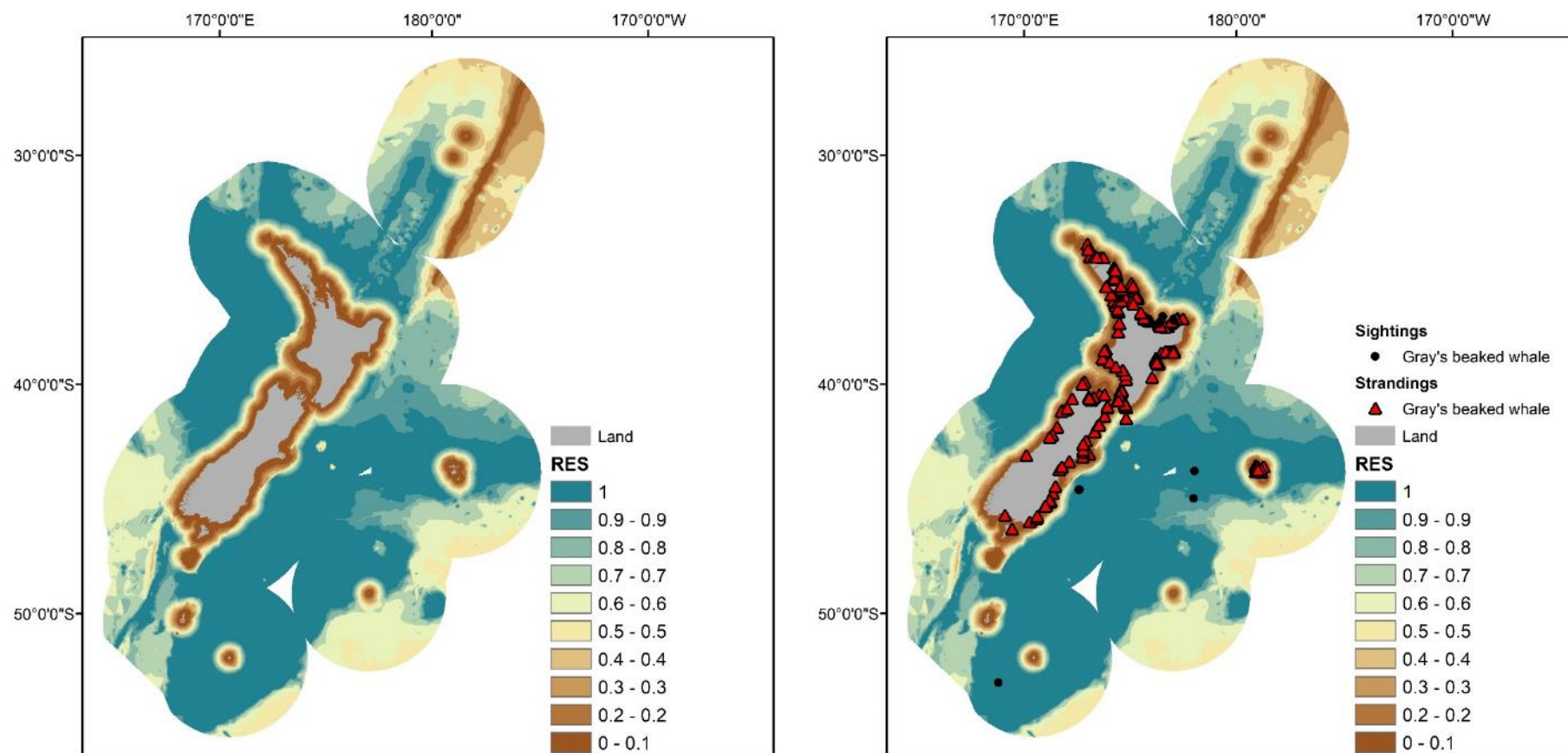


Figure 65: Predicted RES scores for gray's beaked whale (*Mesoplodon grayi*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Hourglass dolphin (*Lagenorhynchus cruciger*)

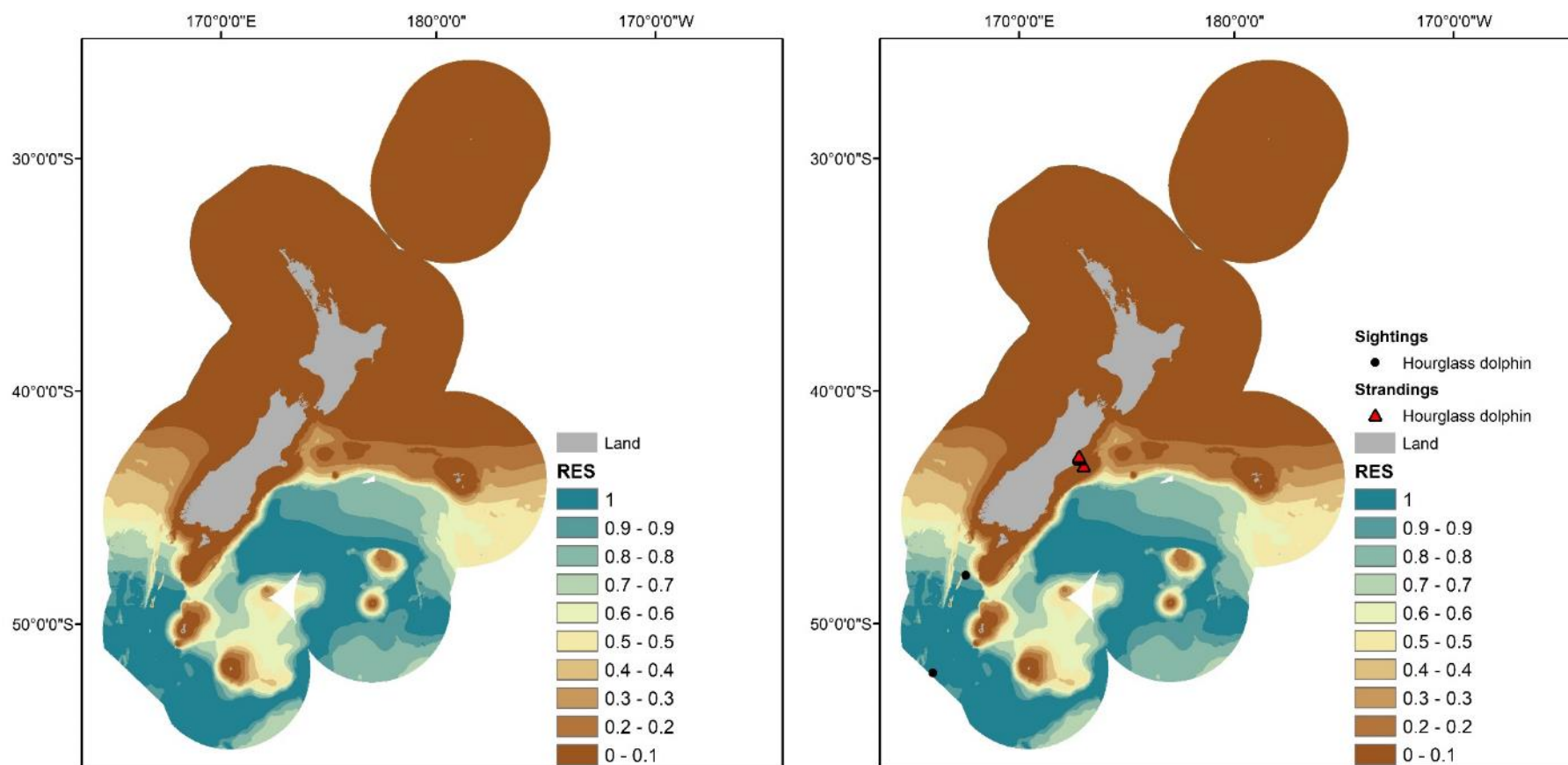


Figure 66: Predicted RES scores for hourglass dolphin (*Lagenorhynchus cruciger*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Pygmy sperm whale (*Kogia breviceps*)

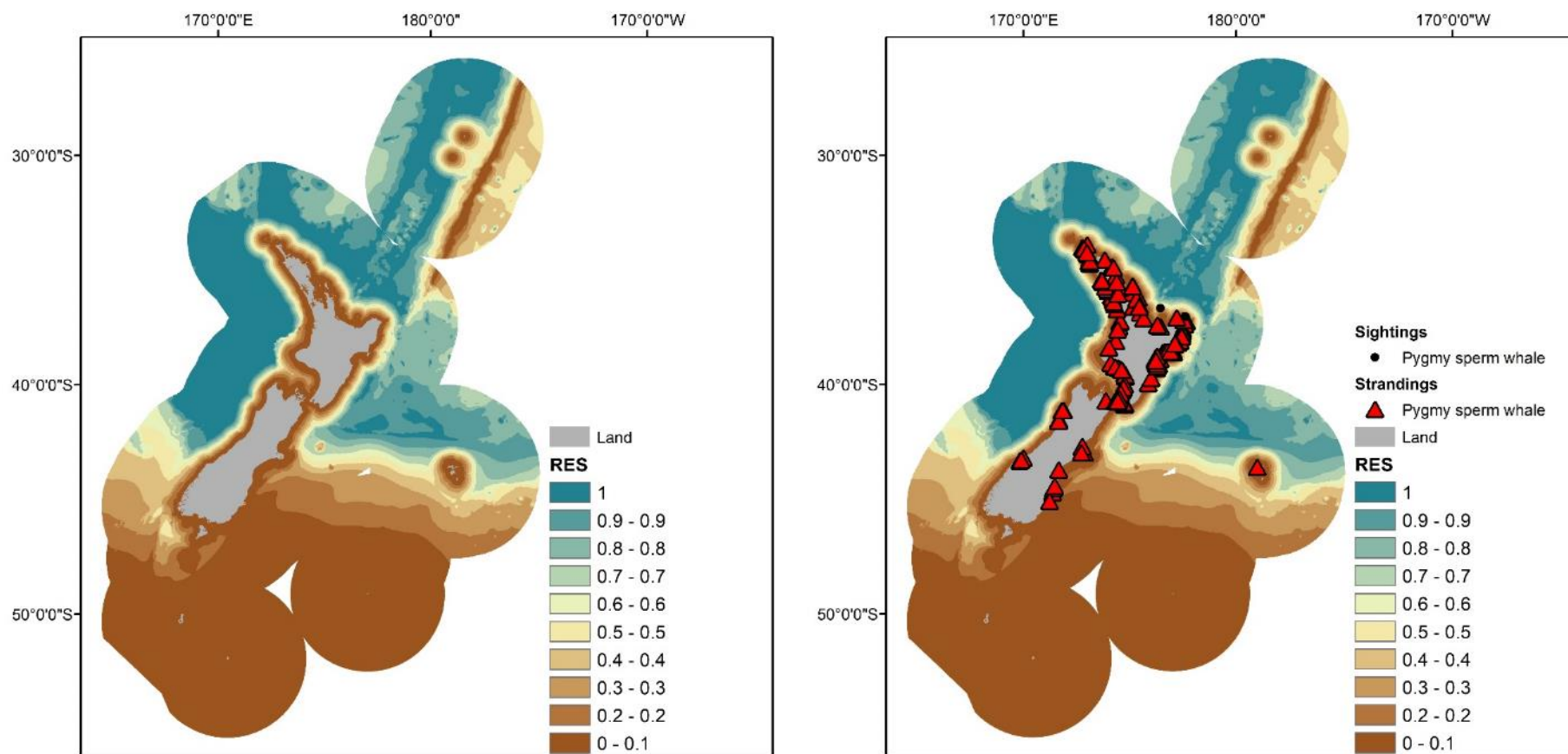


Figure 67: Predicted RES scores for pygmy sperm whale (*Kogia breviceps*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Risso's dolphin (*Grampus griseus*)

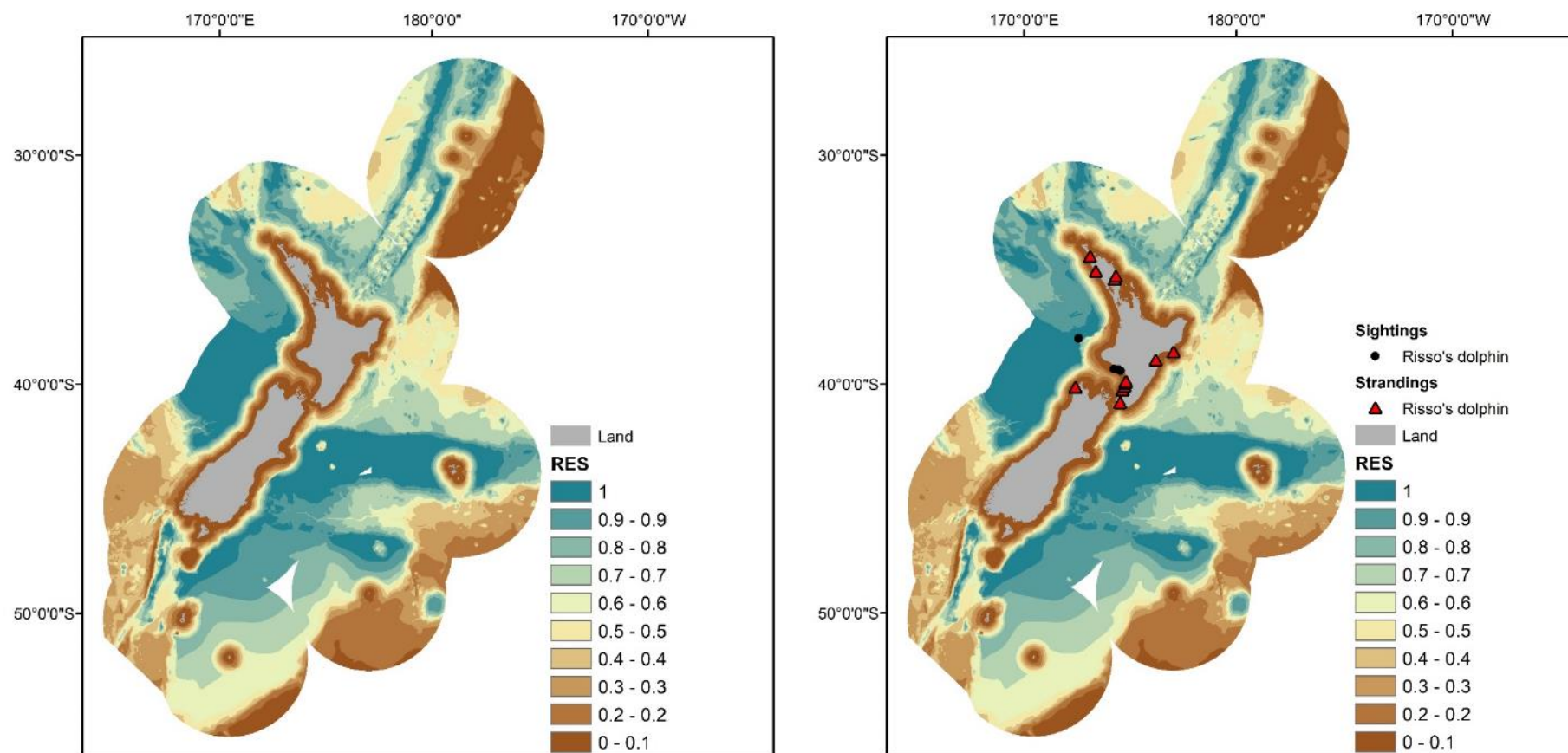


Figure 68: Predicted RES scores for Risso's dolphin (*Grampus griseus*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Shepherd's beaked whale (*Tasmacetus shepherdi*)

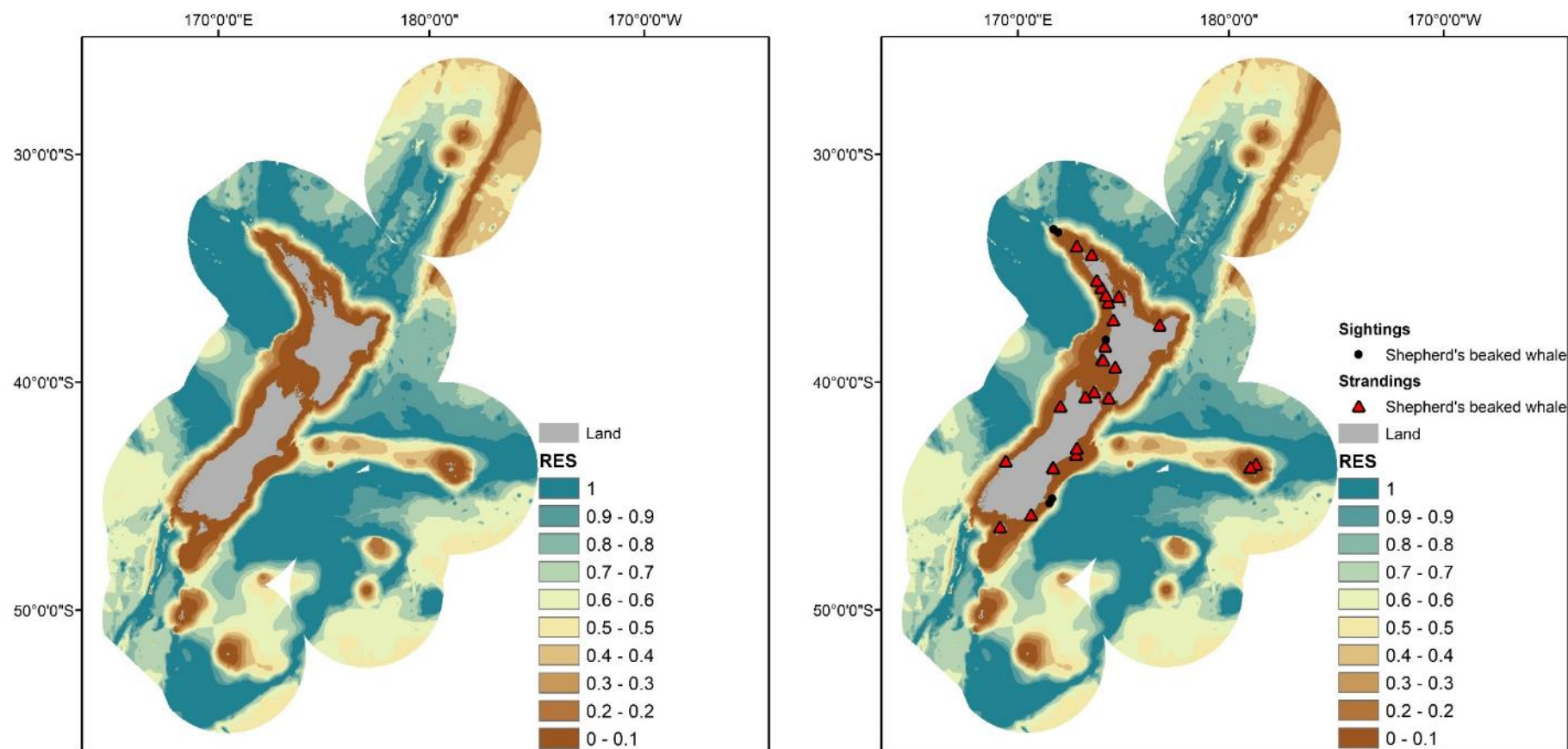


Figure 69: Predicted RES scores for Shepherd's beaked whale (*Tasmacetus shepherdi*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Short finned pilot whale (*Globicephala macrorhynchus*)

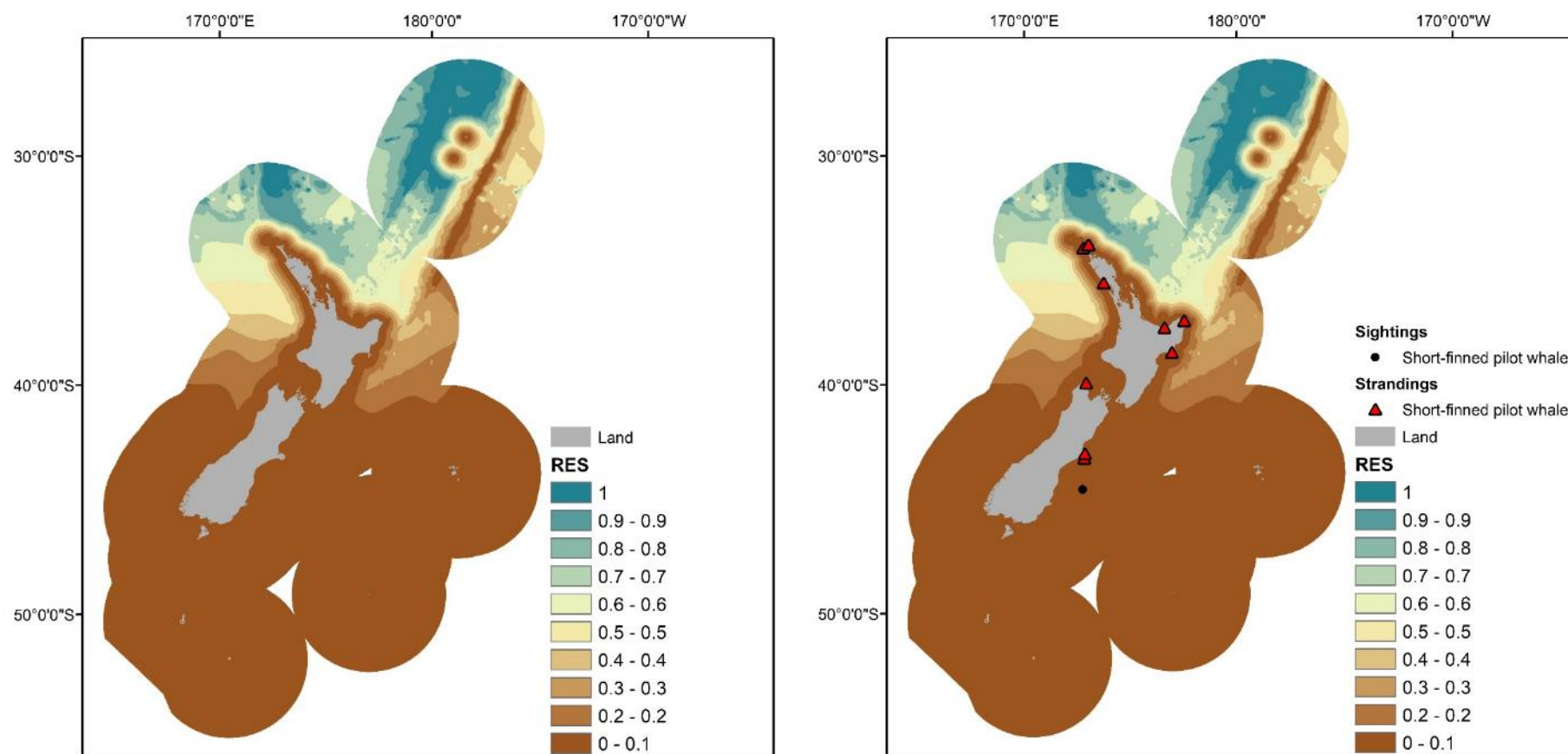


Figure 70: Predicted RES scores for short finned pilot whale (*Globicephala macrorhynchus*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Southern bottlenose whale (*Hyperoodon planifrons*)

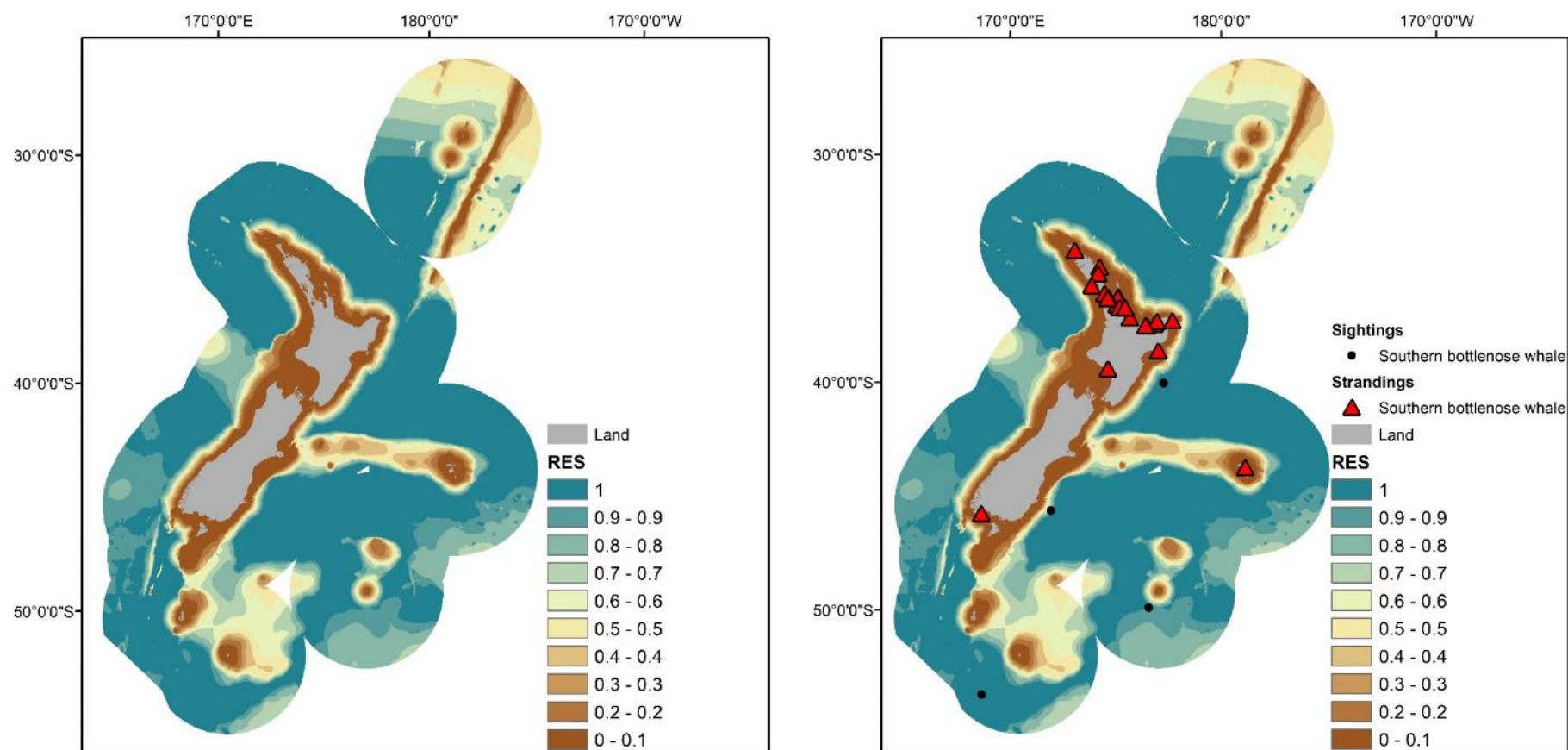


Figure 71: Predicted RES scores for southern bottlenose whale (*Hyperoodon planifrons*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Southern right whale dolphin (*Lissodelphis peronii*)

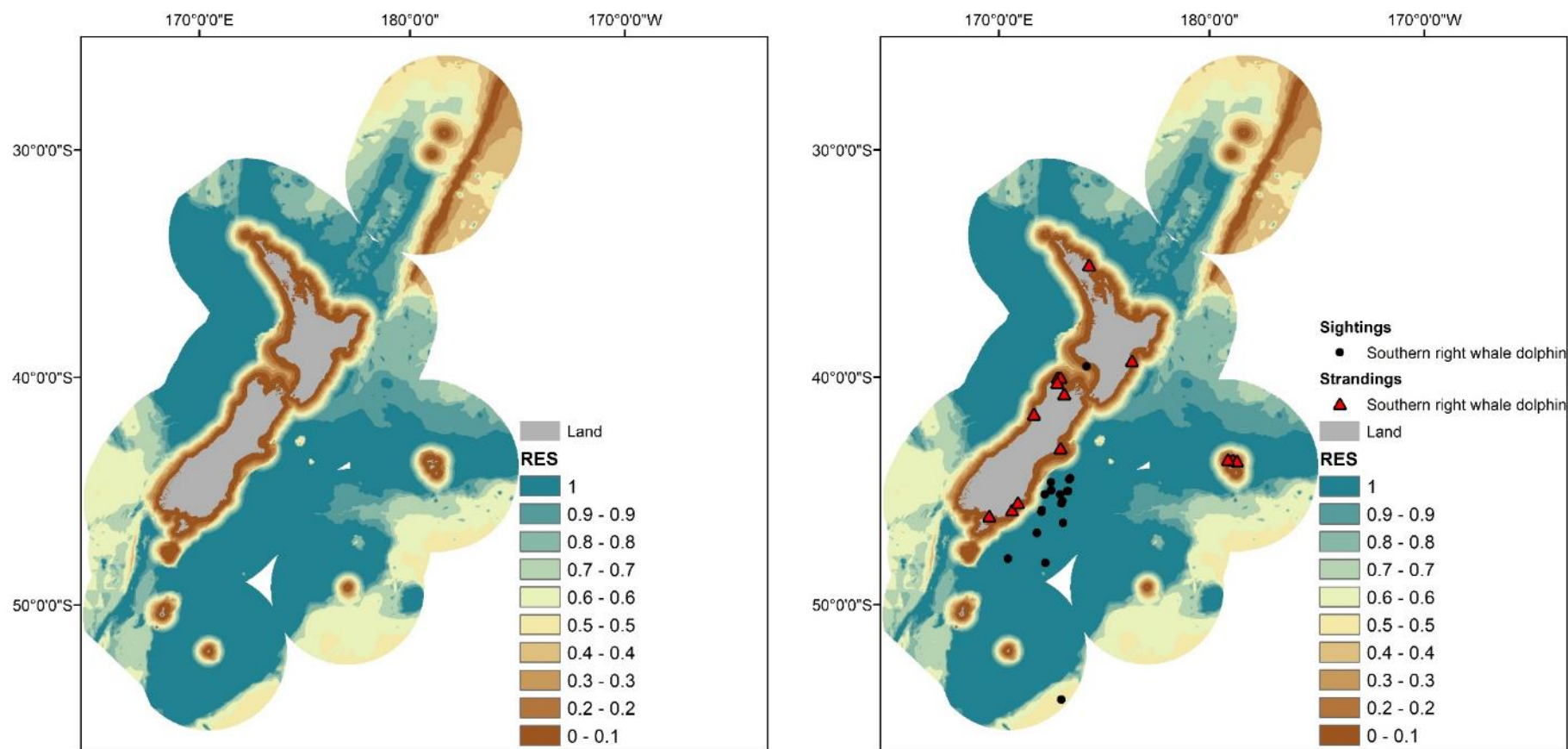


Figure 72: Predicted RES scores for southern right whale dolphin (*Lissodelphis peronii*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Spectacled porpoise (*Phocoena dioptrica*)

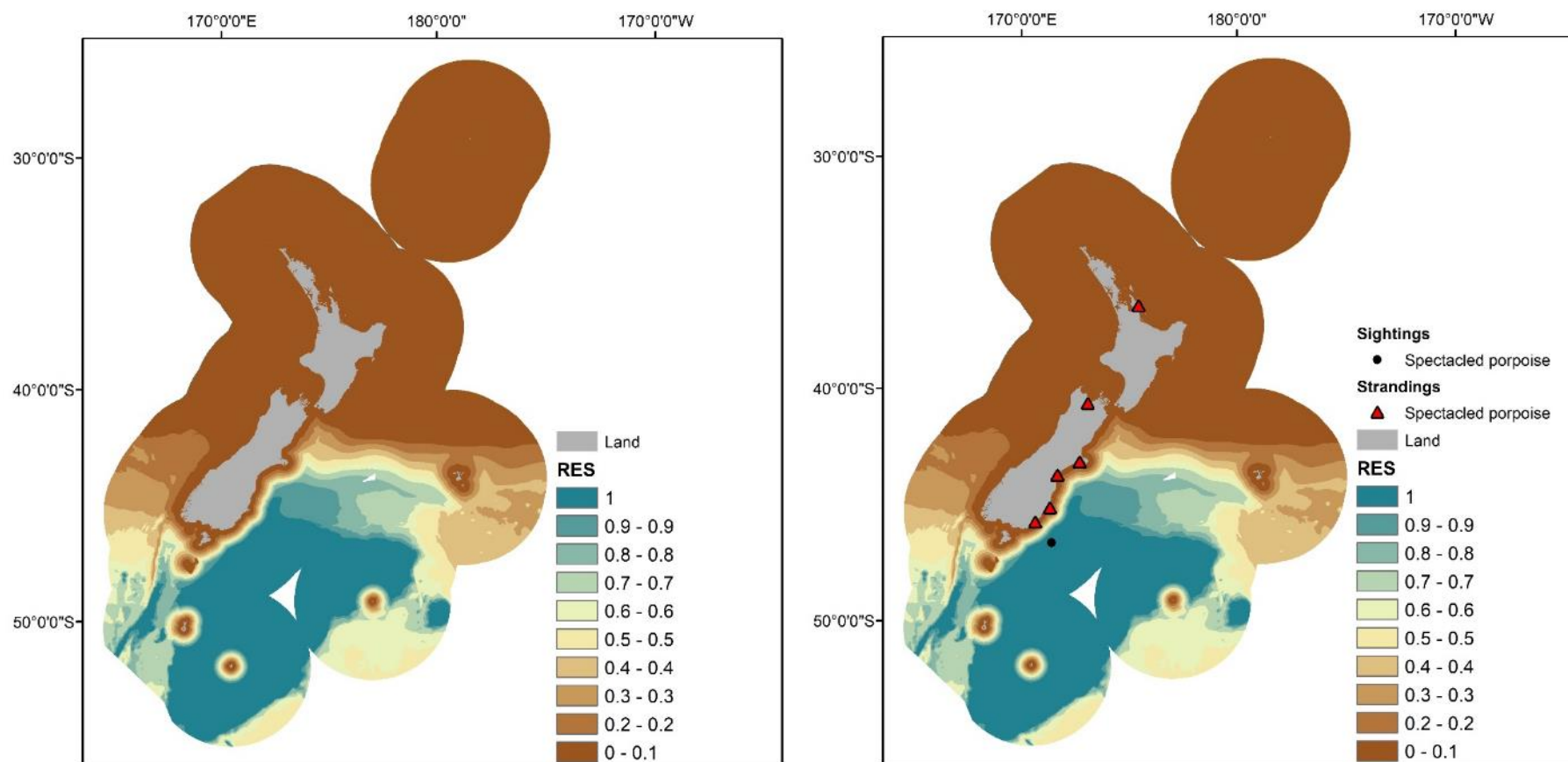


Figure 73: Predicted RES scores for spectacled porpoise (*Phocoena dioptrica*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

Striped dolphin (*Stenella coeruleoalba*)

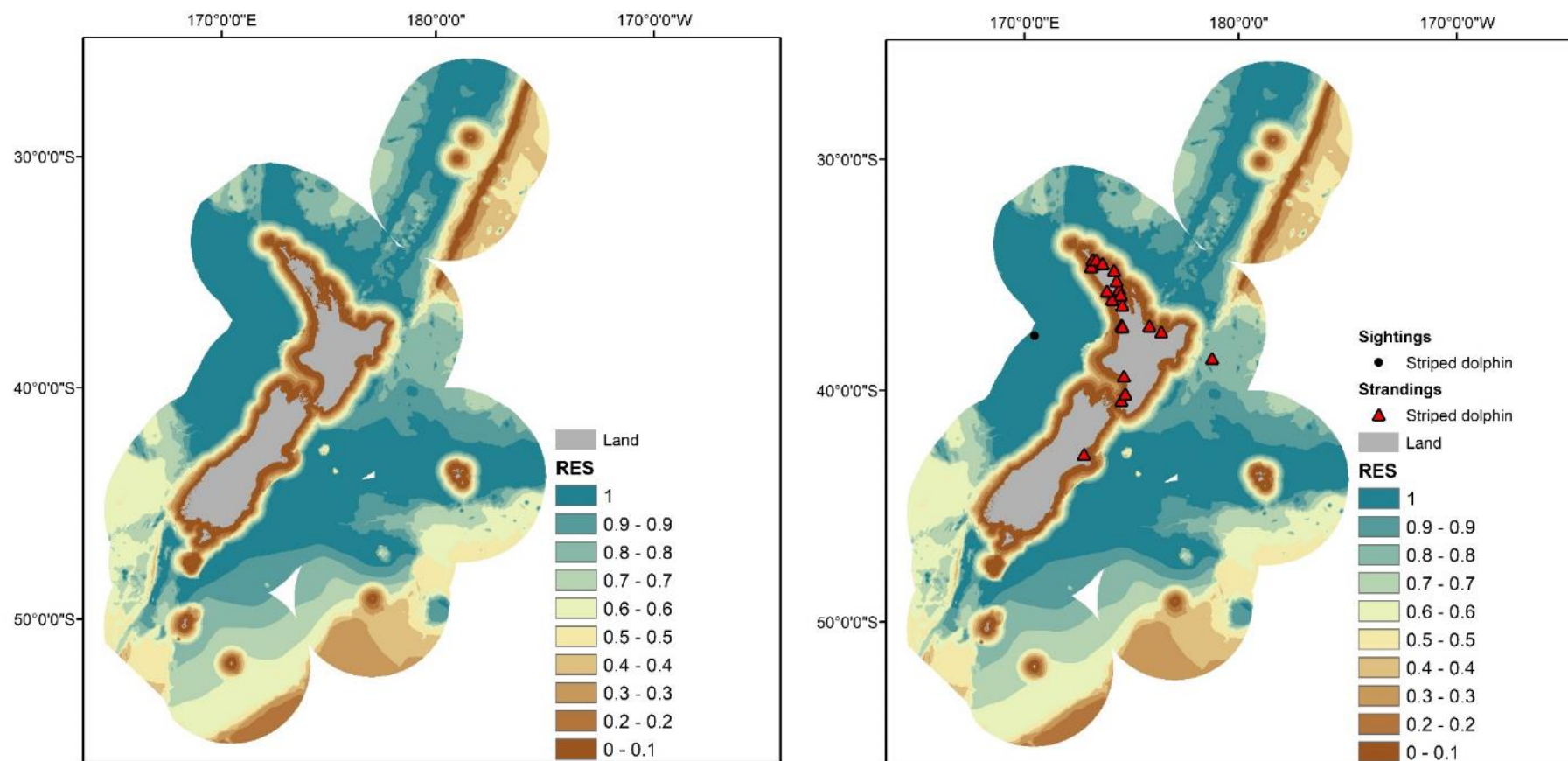


Figure 74: Predicted RES scores for striped dolphin (*Stenella coeruleoalba*) ranging from less suitable (brown) to very suitable (blue) (left); Predicted RES scores are shown with sightings at sea and location of recorded strandings (right).

APPENDIX 3: BRT MODEL OUTPUTS

Presence/relative absence BRT models

Bottlenose dolphin (*Tursiops truncatus*) – see section 3.2.3.1 in main body of the report

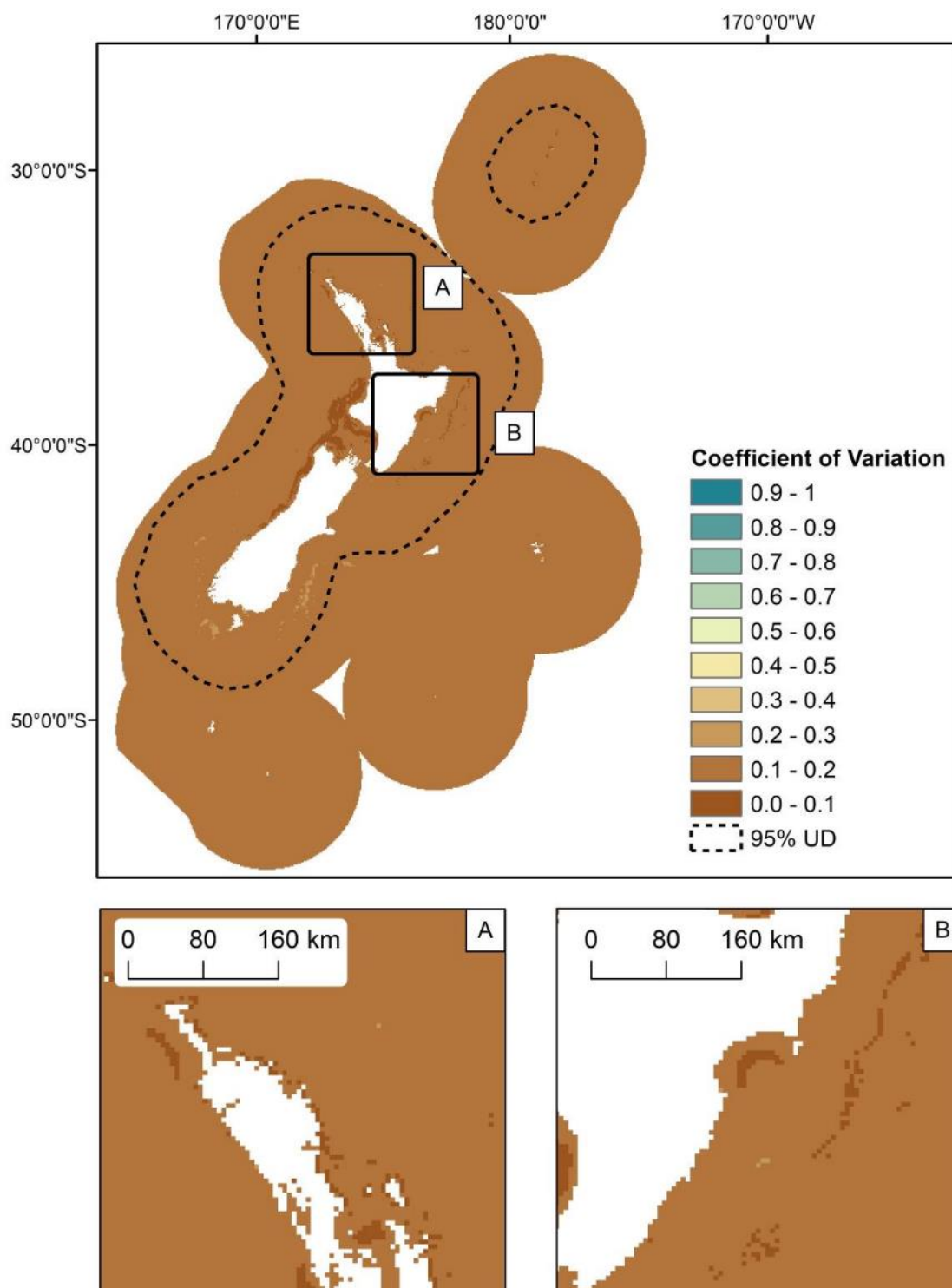


Figure 72: Uncertainty estimates (coefficient of variation, CV) of bottlenose dolphin (*Tursiops truncatus*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island and Hauraki Gulf; B) south east coast of North Island.

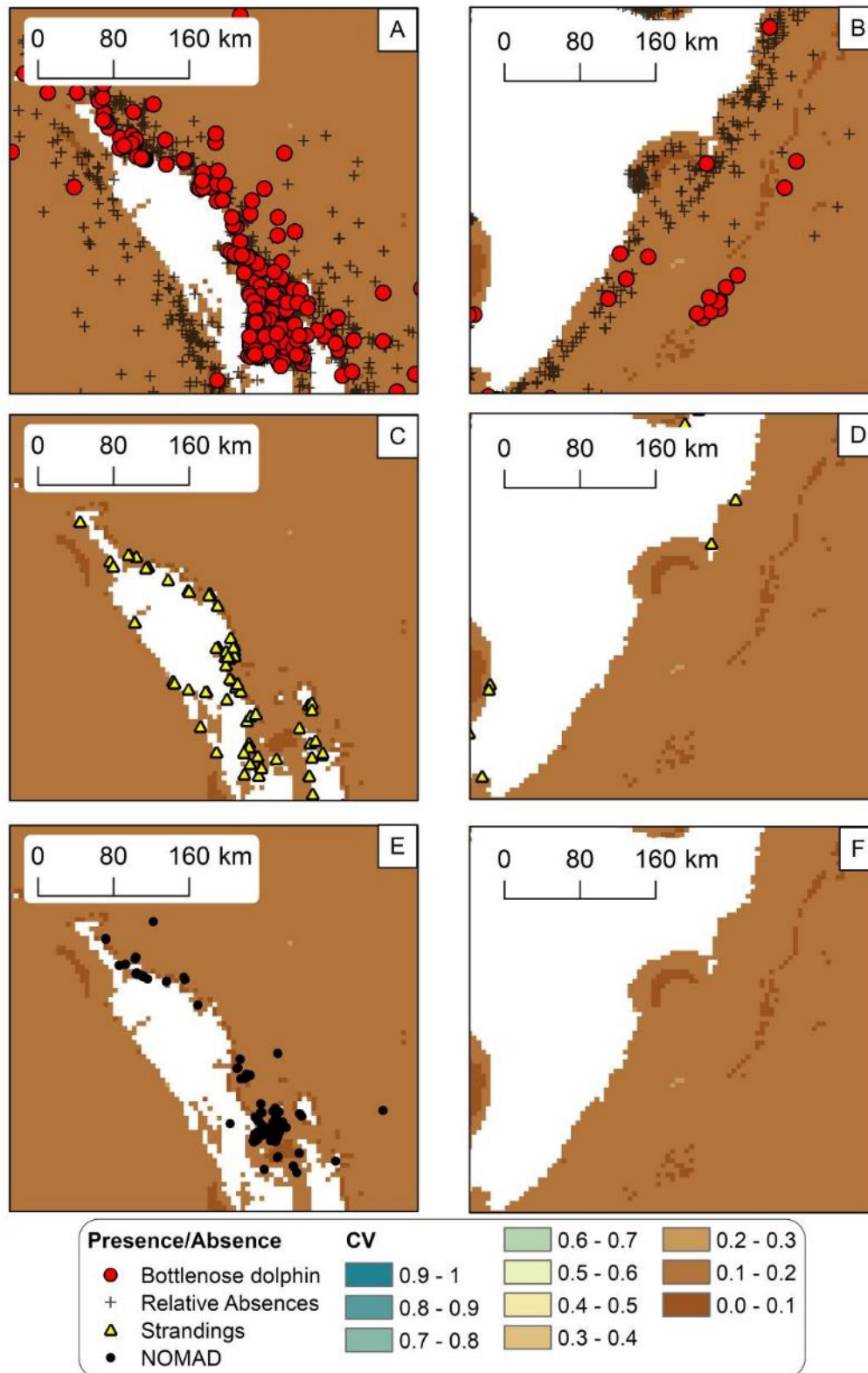


Figure 73: Uncertainty estimates (coefficient of variation, CV) of bottlenose dolphin (*Tursiops truncatus*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Bottlenose dolphin probability presence models in the north of North Island and Hauraki Gulf are shown with presence / relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of Bottlenose dolphin probability presence models on the south east coast of North Island are shown with presence / relative absences (B), stranding locations (D) and NOMAD sightings (F).

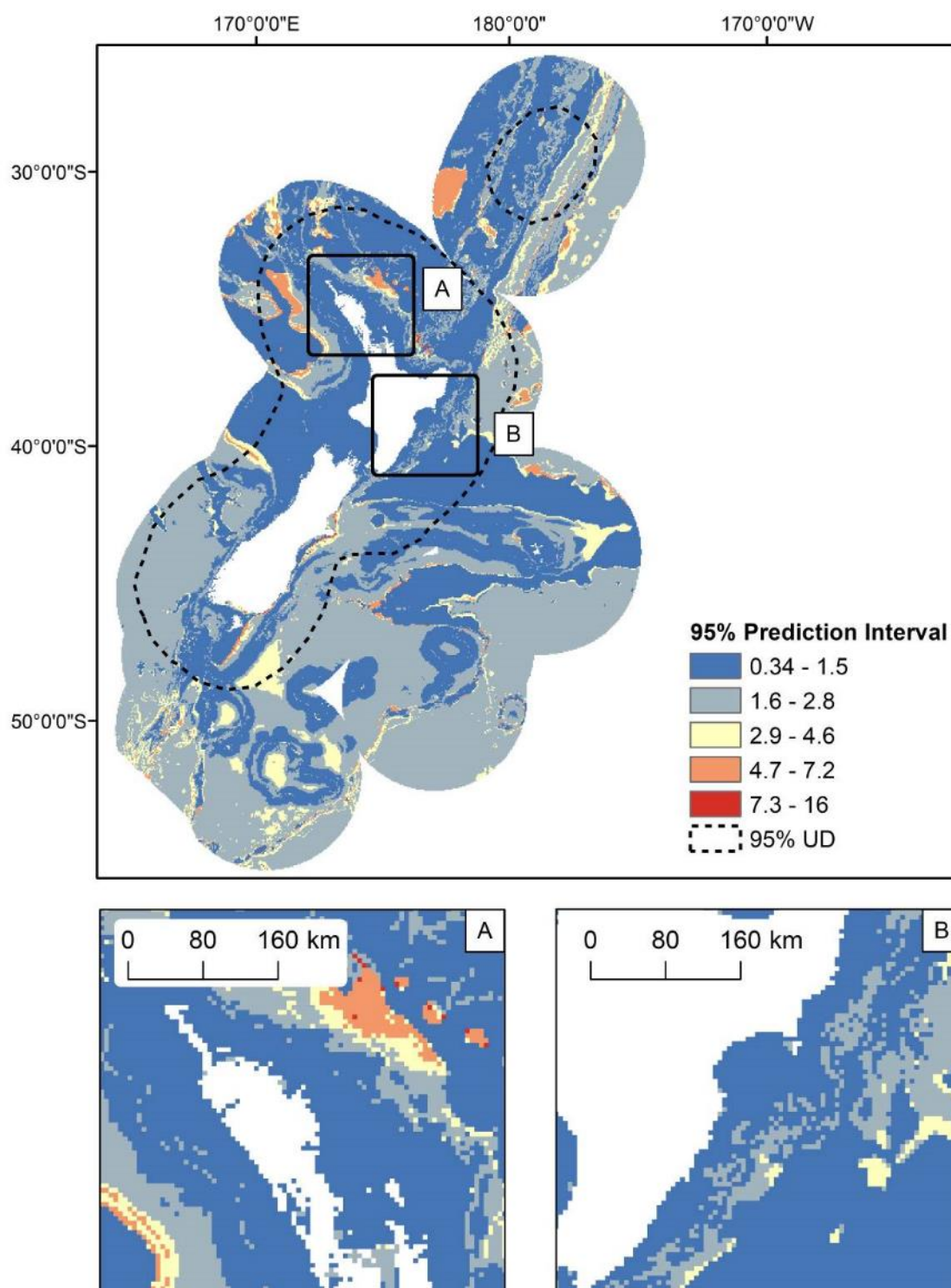


Figure 74: Uncertainty estimates (95% prediction interval) of bottlenose dolphin (*Tursiops truncatus*) predicted relative densities in the New Zealand EEZ modelled using bootstrapped BRTs. 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island and Hauraki Gulf; B) south east coast of North Island.

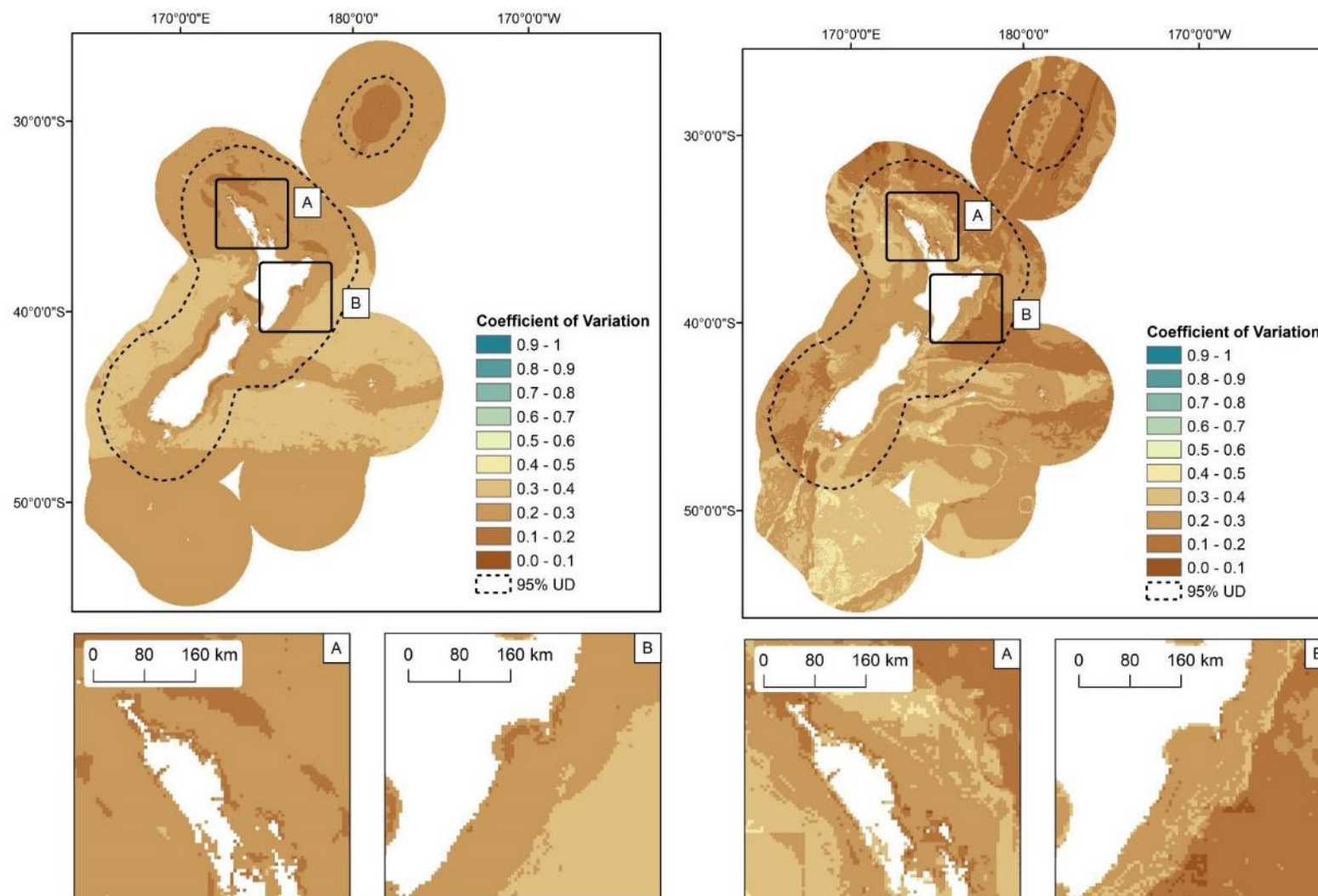


Figure 75: Seasonal uncertainty estimates (coefficient of variation, CV) of bottlenose dolphin (*Tursiops truncatus*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs fitted with Winter (May - Oct) presence/relative absence sightings records (left) and Summer (Nov - Apr) sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island and Hauraki Gulf; B) south east coast of North Island.

Common dolphin (*Delphinus delphis*) – see section 3.2.3.2 in main body of report

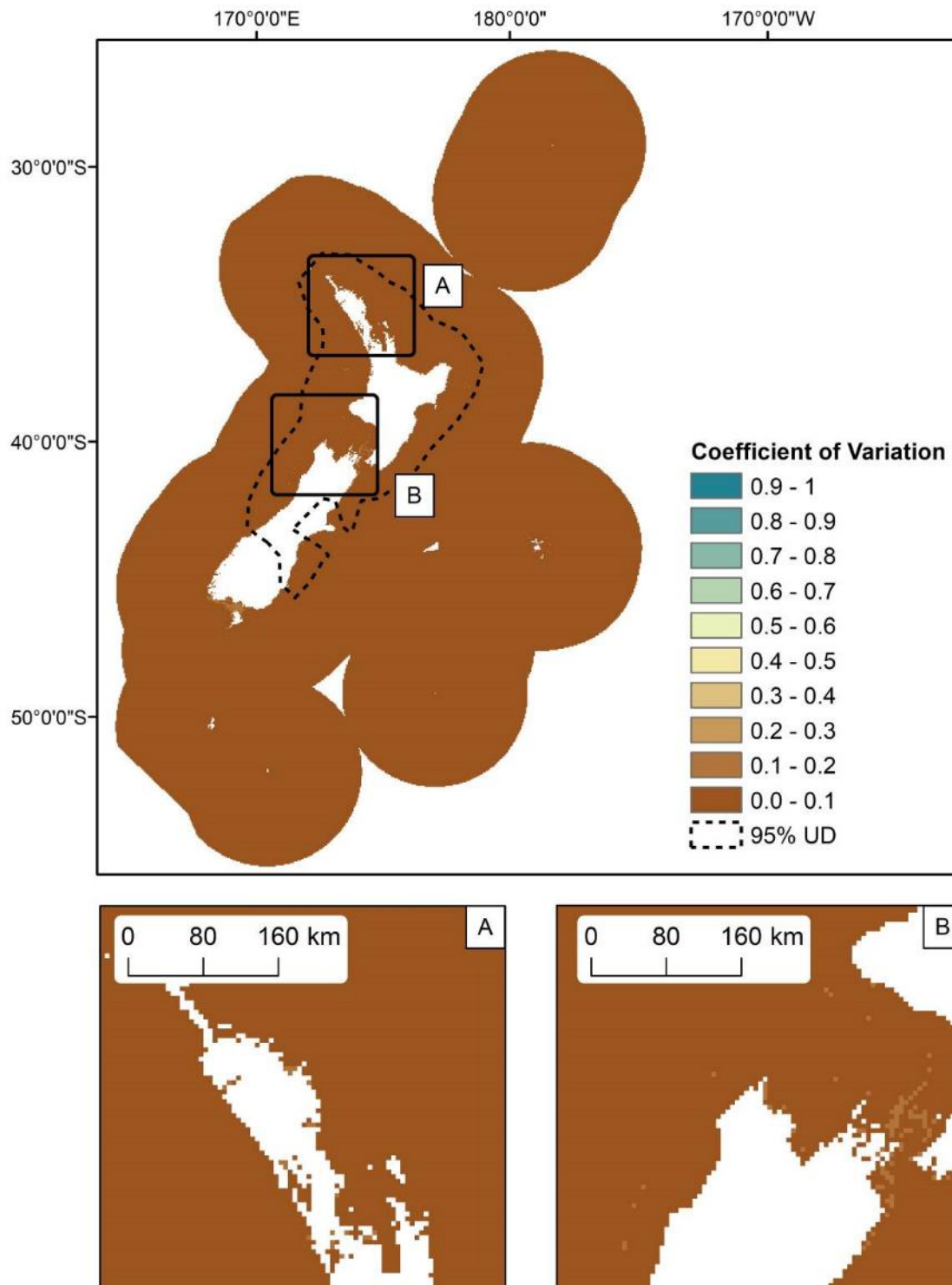


Figure 76: Uncertainty estimates (coefficient of variation, CV) of common dolphin (*Delphinus delphis*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island; B) Taranaki Bight.

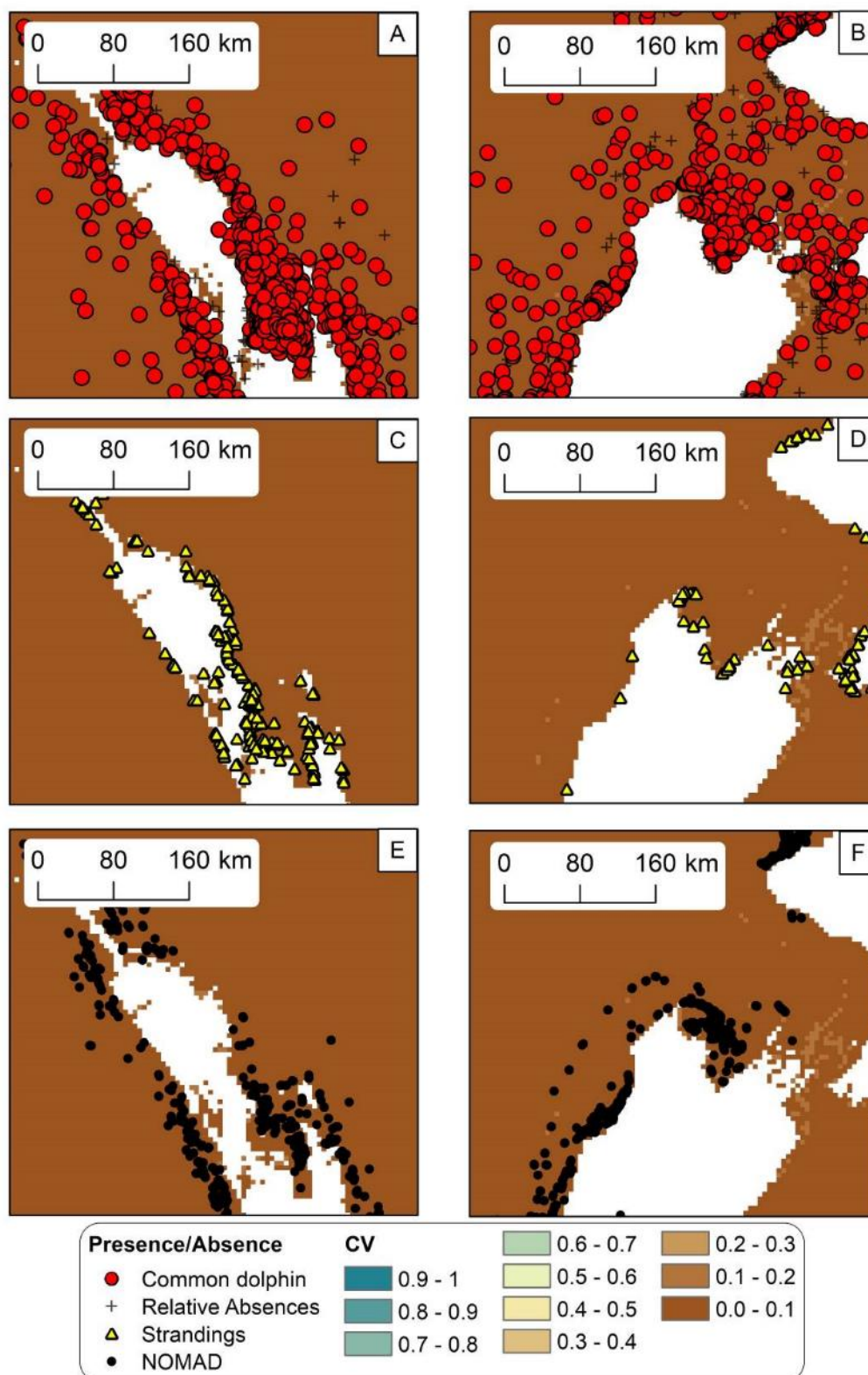


Figure 77: Uncertainty estimates (coefficient of variation, CV) of Common dolphin (*Delphinus delphis*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Common dolphin probability presence models in the north of North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of Common dolphin probability presence models in the Taranaki Bight are shown with Presence/relative Absences (B), stranding locations (D) and NOMAD sightings (F).

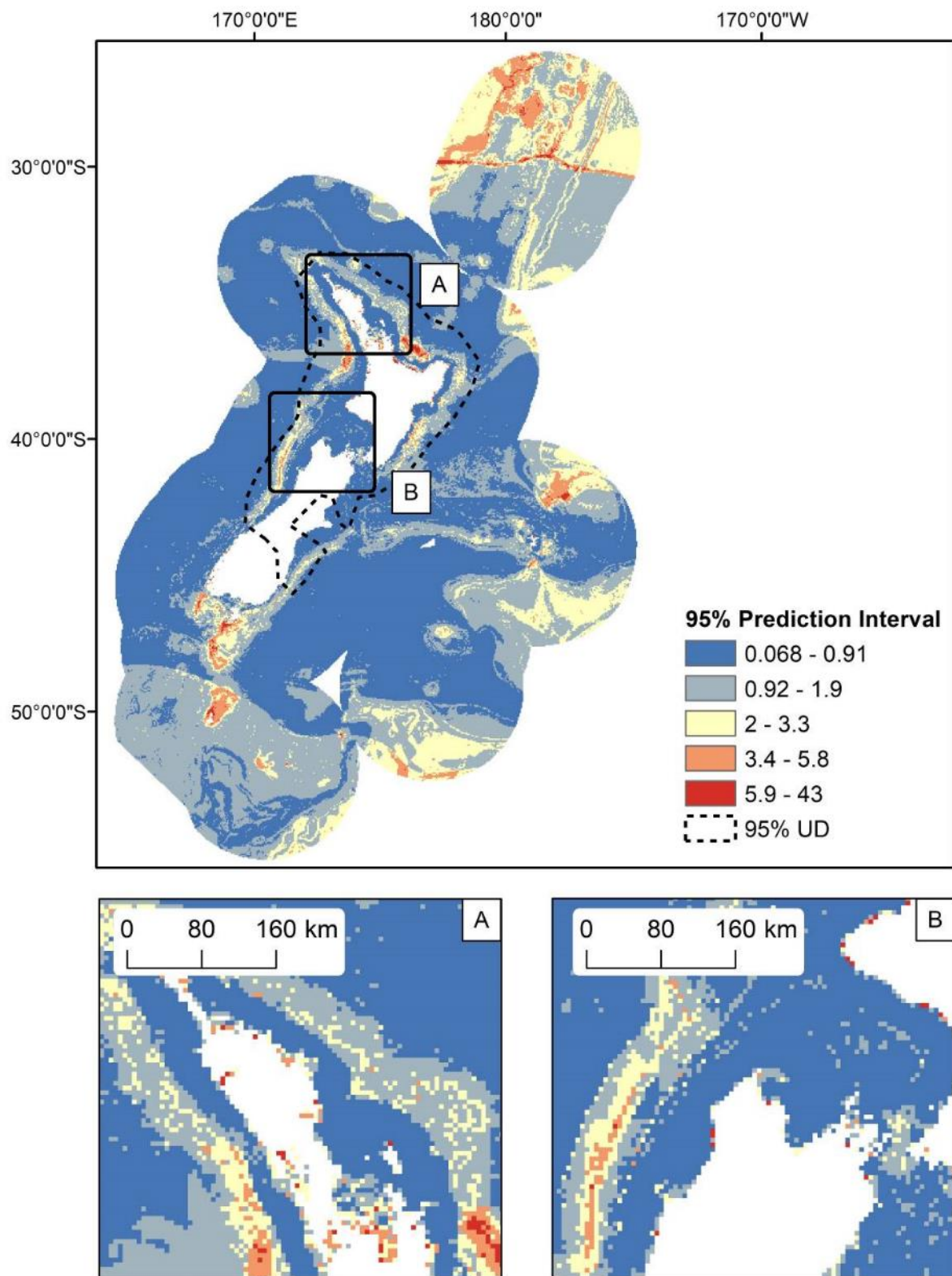


Figure 78: Uncertainty estimates (95% prediction interval) of common dolphin (*Delphinus delphis*) predicted relative densities in the New Zealand EEZ modelled using bootstrapped BRTs. 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island; B) Taranaki Bight.

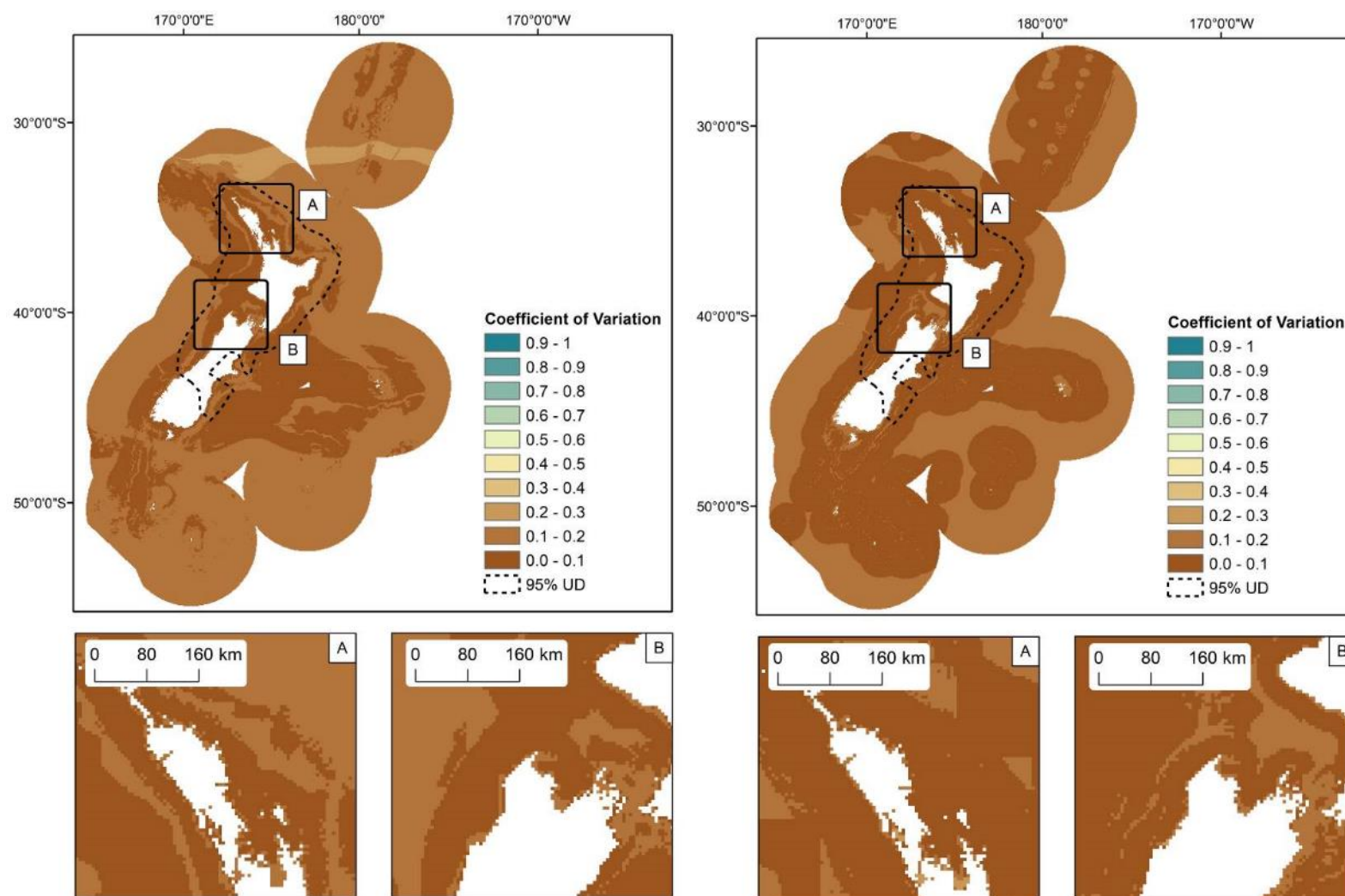


Figure 79: Seasonal uncertainty estimates (coefficient of variation, CV) of common dolphin (*Delphinus delphis*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs fitted with Winter (May - Oct) presence/relative absence sightings records (left) and Summer (Nov - Apr) sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island; B) Taranaki Bight.

Hector's dolphin (*Cephalorhynchus hectori hectori*) – see section 3.2.3.3 in main body of report

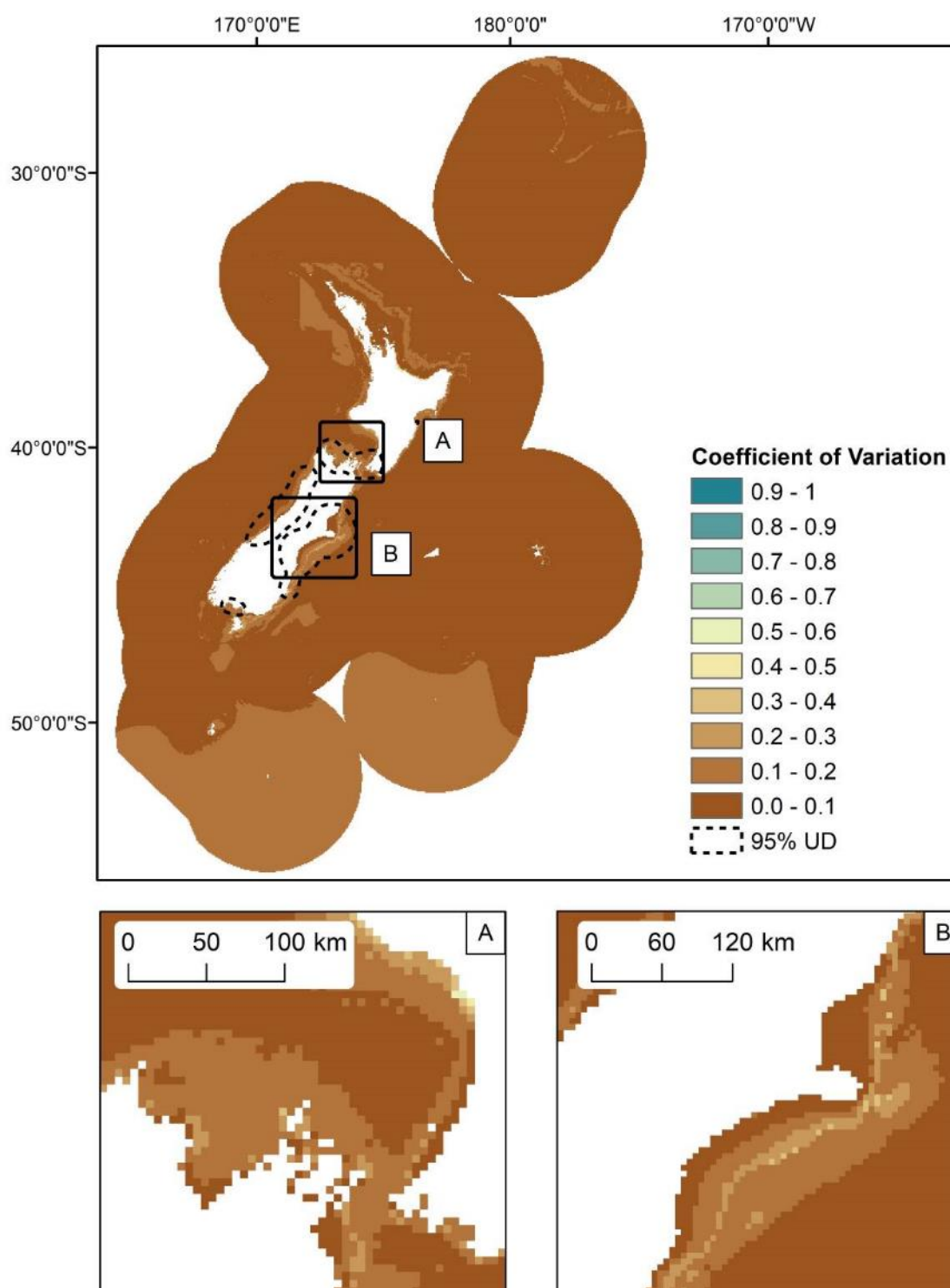


Figure 80: Uncertainty estimates (coefficient of variation, CV) of Hector's dolphin (*Cephalorhynchus hectori hectori*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) Taranaki Bight; B) east coast of South Island.

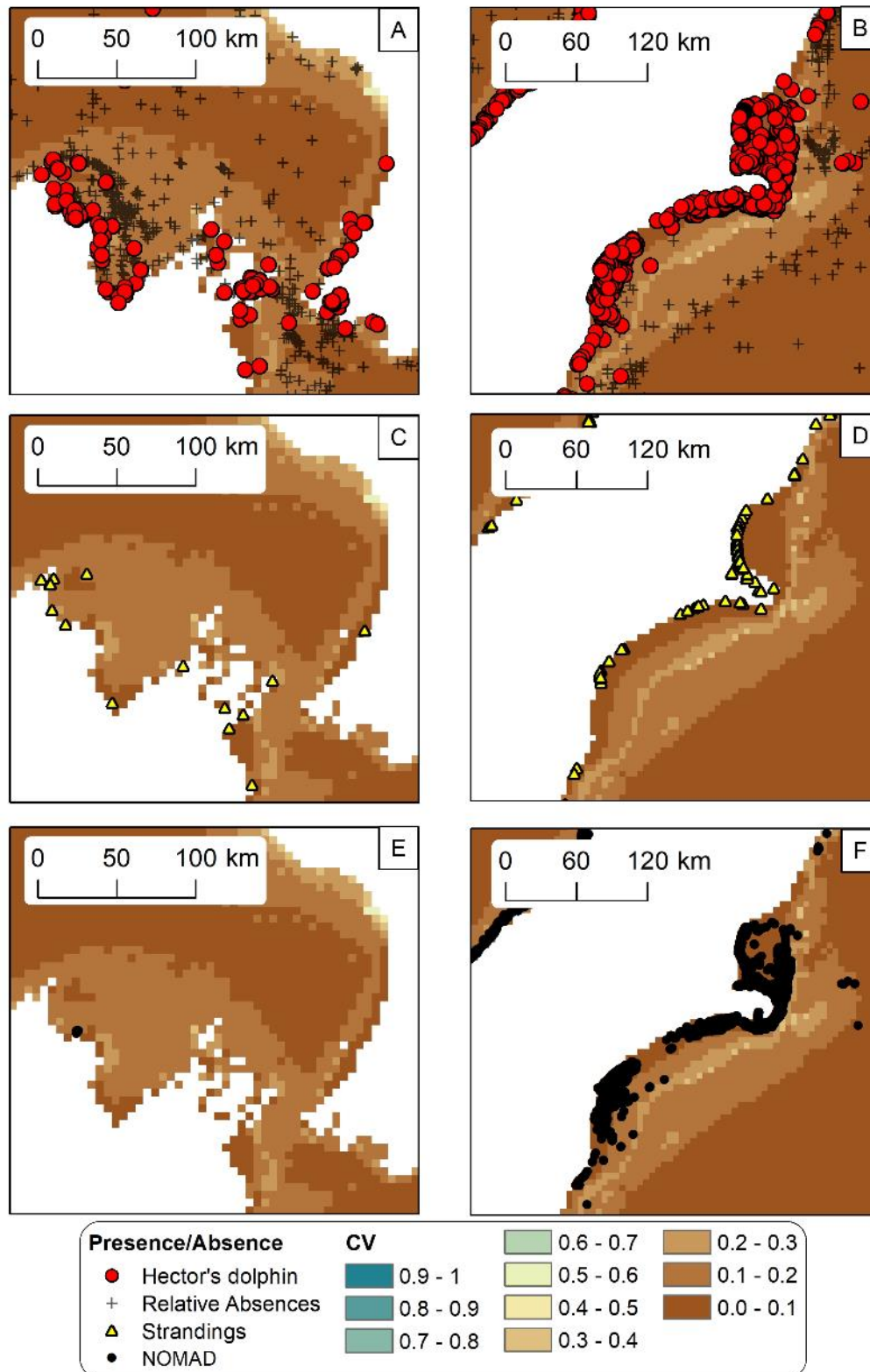


Figure 81: Uncertainty estimates (coefficient of variation, CV) of Hector's dolphin (*Cephalorhynchus hectori hectori*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Hector's dolphin probability of presence models in the Taranaki Bight are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of Hector's dolphin probability of presence models in the east coast of South Island are shown with Presence/relative Absences (B), stranding locations (D) and NOMAD sightings (F).

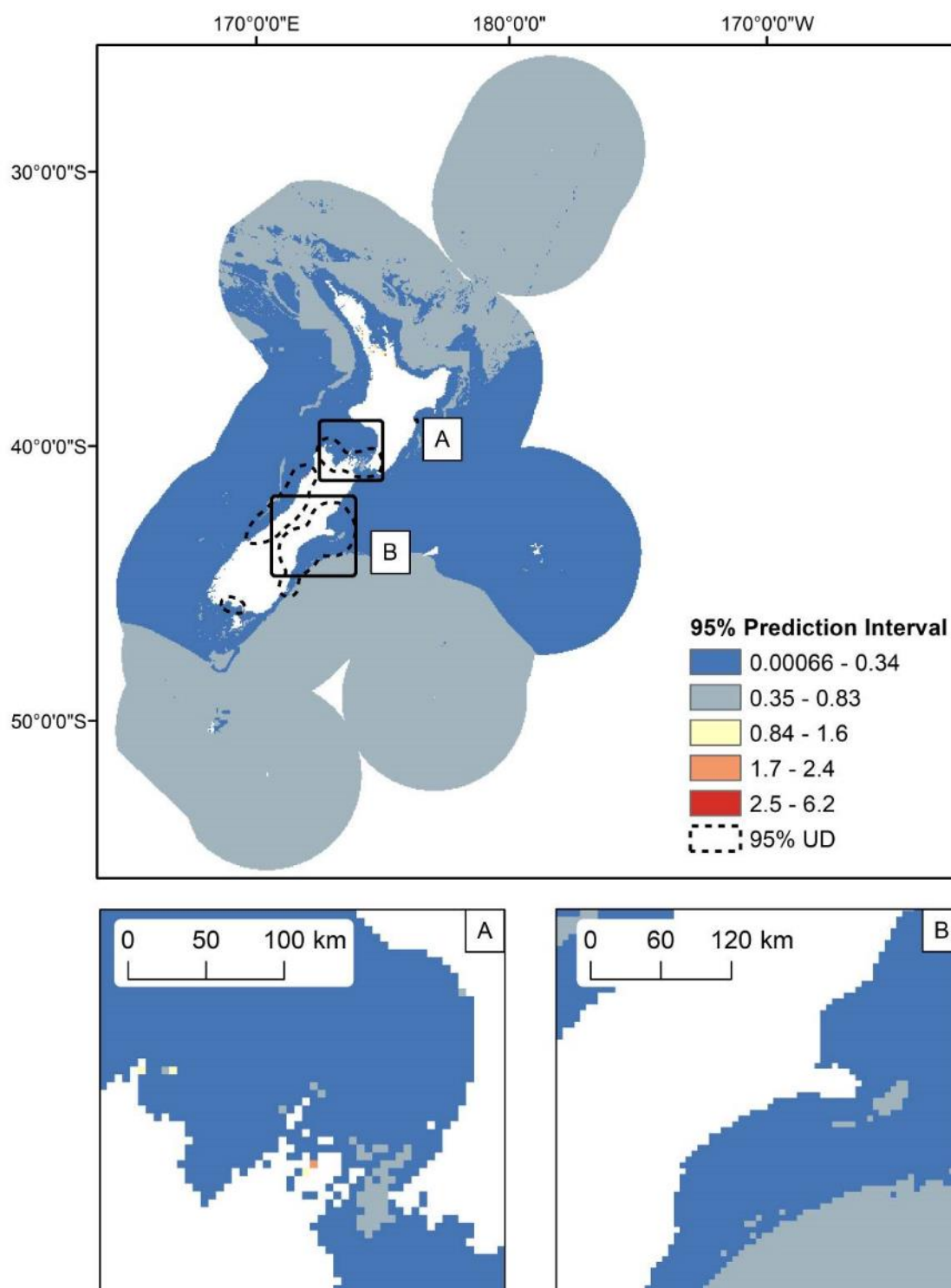


Figure 82: Uncertainty estimates (95% prediction interval) of Hector's dolphin (*Cephalorhynchus hectori*) predicted relative densities in the New Zealand EEZ modelled using bootstrapped BRTs. 95% utilisation distribution is shown as dashed line. Inset maps: A) Taranaki Bight; B) east coast of South Island.

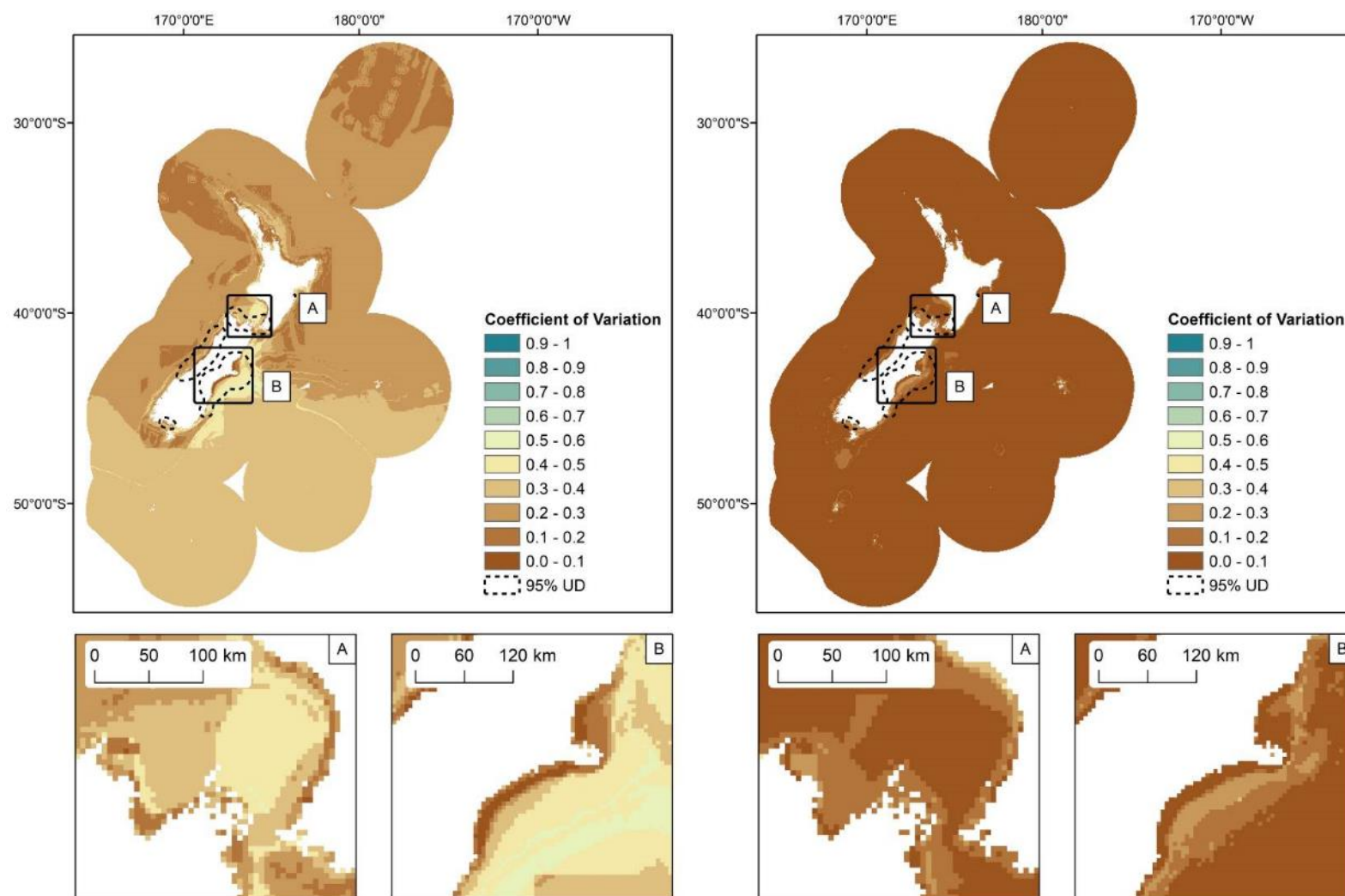


Figure 83: Seasonal uncertainty estimates (coefficient of variation, CV) of Hector's dolphin (*Cephalorhynchus hectori hectori*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs fitted with Winter (May - Oct) (left) and Summer (Nov - Apr) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) Taranaki Bight; B) east coast of South Island.

Dusky dolphin (*Lagenorhynchus obscurus*) – see section 3.2.3.3 in main body of the report

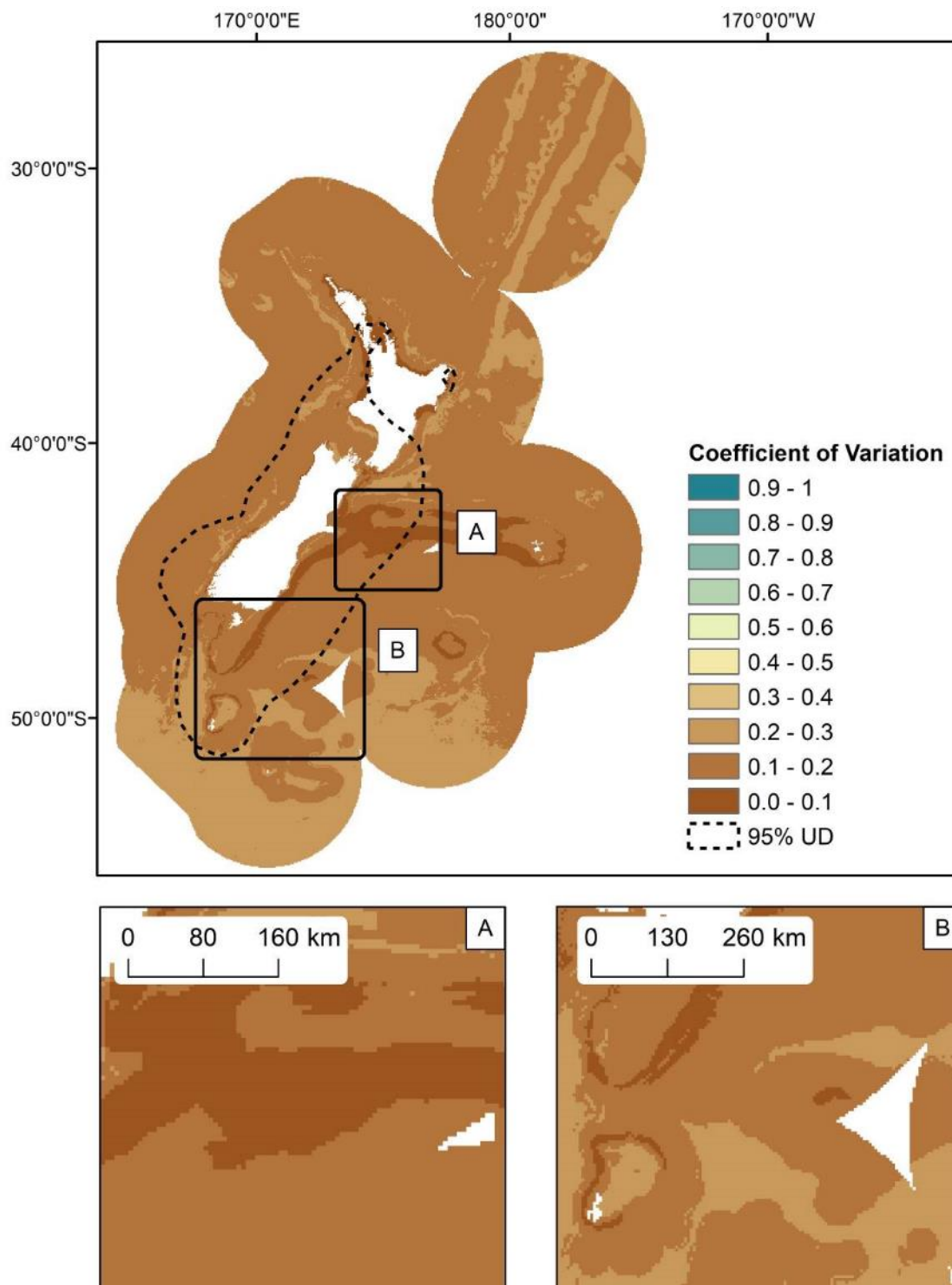


Figure 84: Uncertainty estimates (coefficient of variation, CV) of dusky dolphin (*Lagenorhynchus obscurus*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) west Chatham Rise; B) Campbell Plateau.

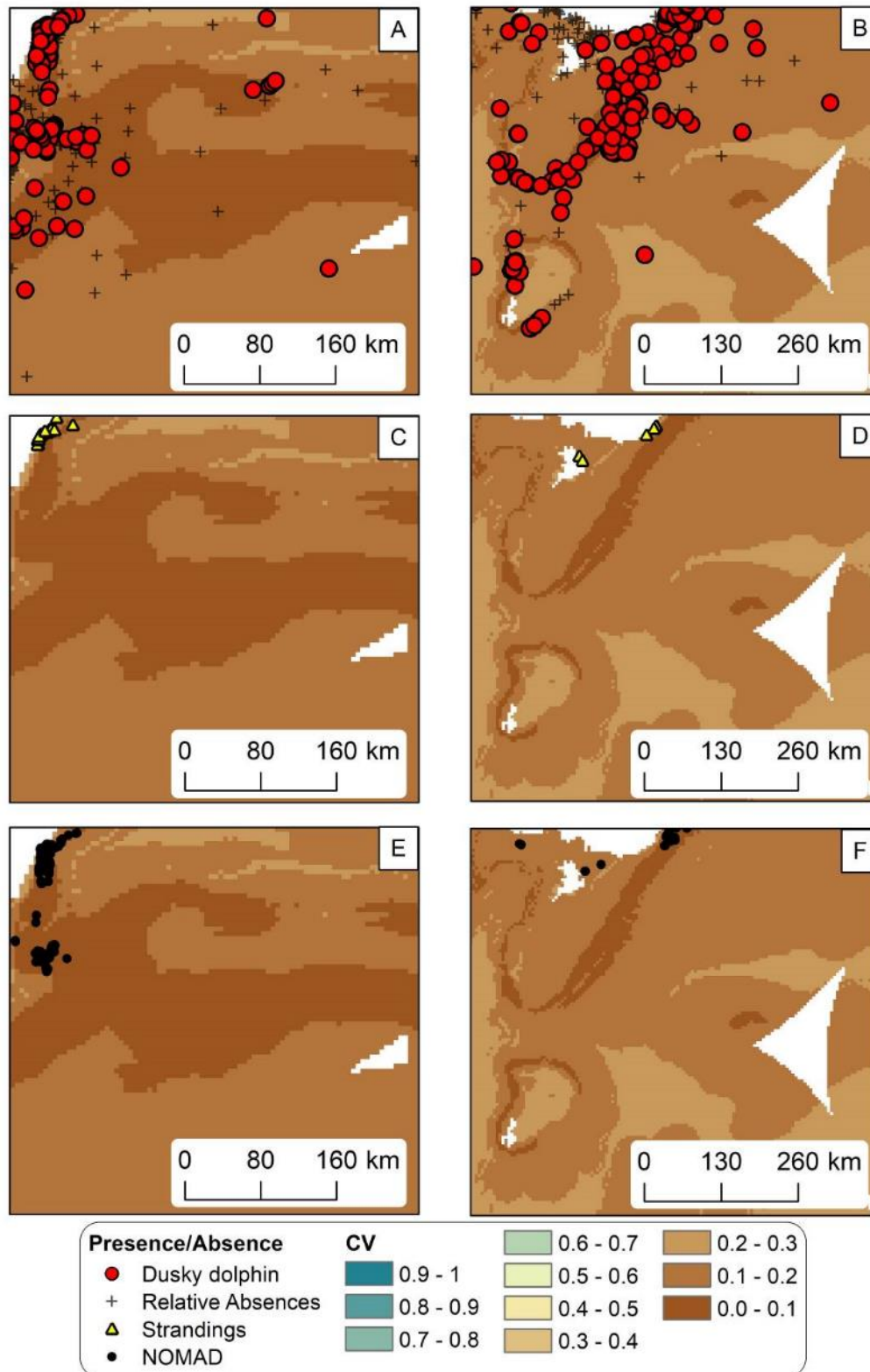


Figure 85: Uncertainty estimates (coefficient of variation, CV) of dusky dolphin (*Lagenorhynchus obscurus*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Dusky dolphin probability presence models on the west Chatham Rise are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of Dusky dolphin probability presence models on the Campbell Plateau are shown with Presence/relative Absences (B), stranding locations (D) and NOMAD sightings (F).

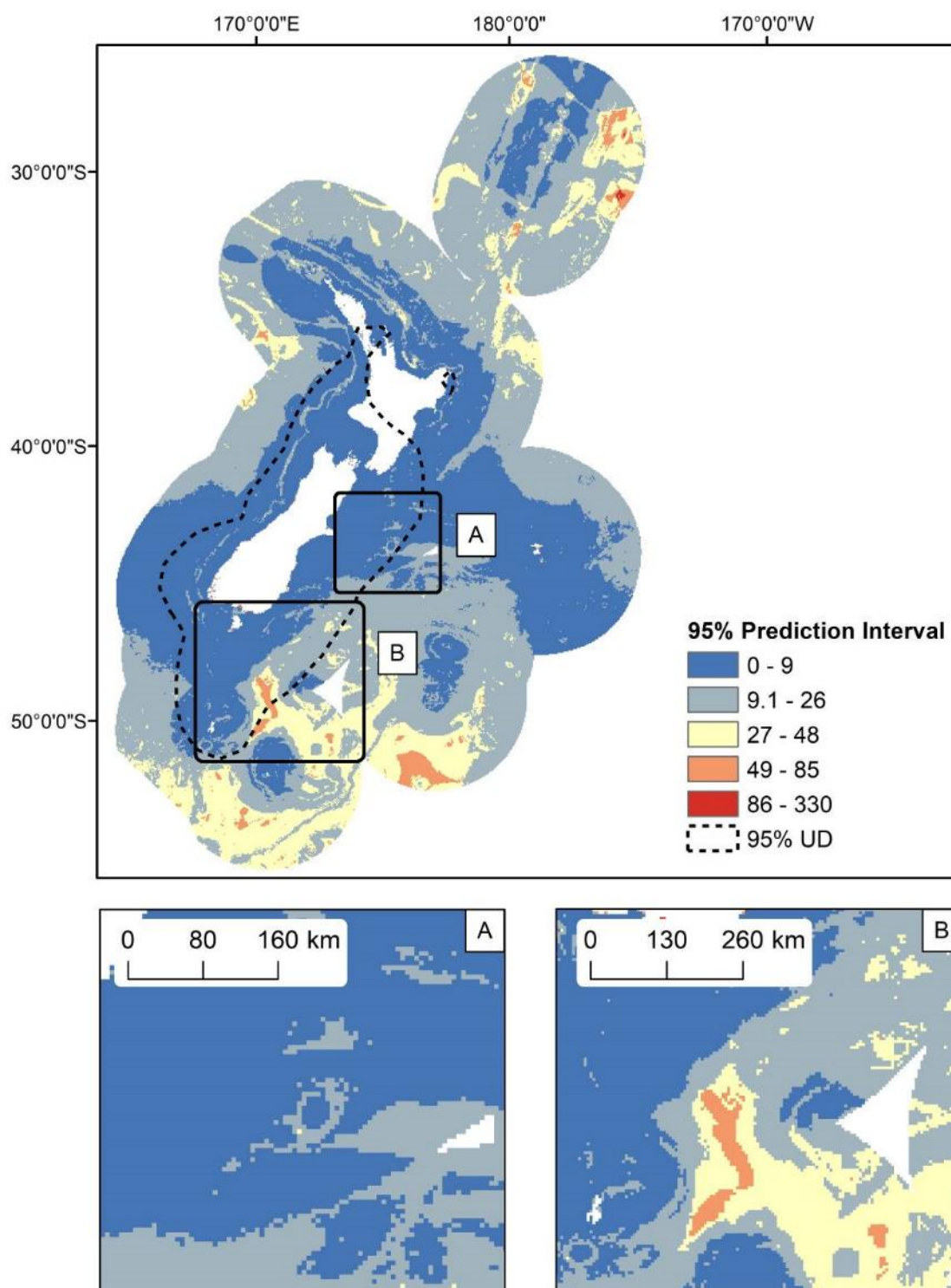


Figure 86: Uncertainty estimates (95% prediction interval) of dusky dolphin (*Lagenorhynchus obscurus*) predicted relative densities in the New Zealand EEZ modelled using bootstrapped BRTs. 95% utilisation distribution is shown as dashed line. Inset maps: A) west Chatham Rise; B) Campbell Plateau.

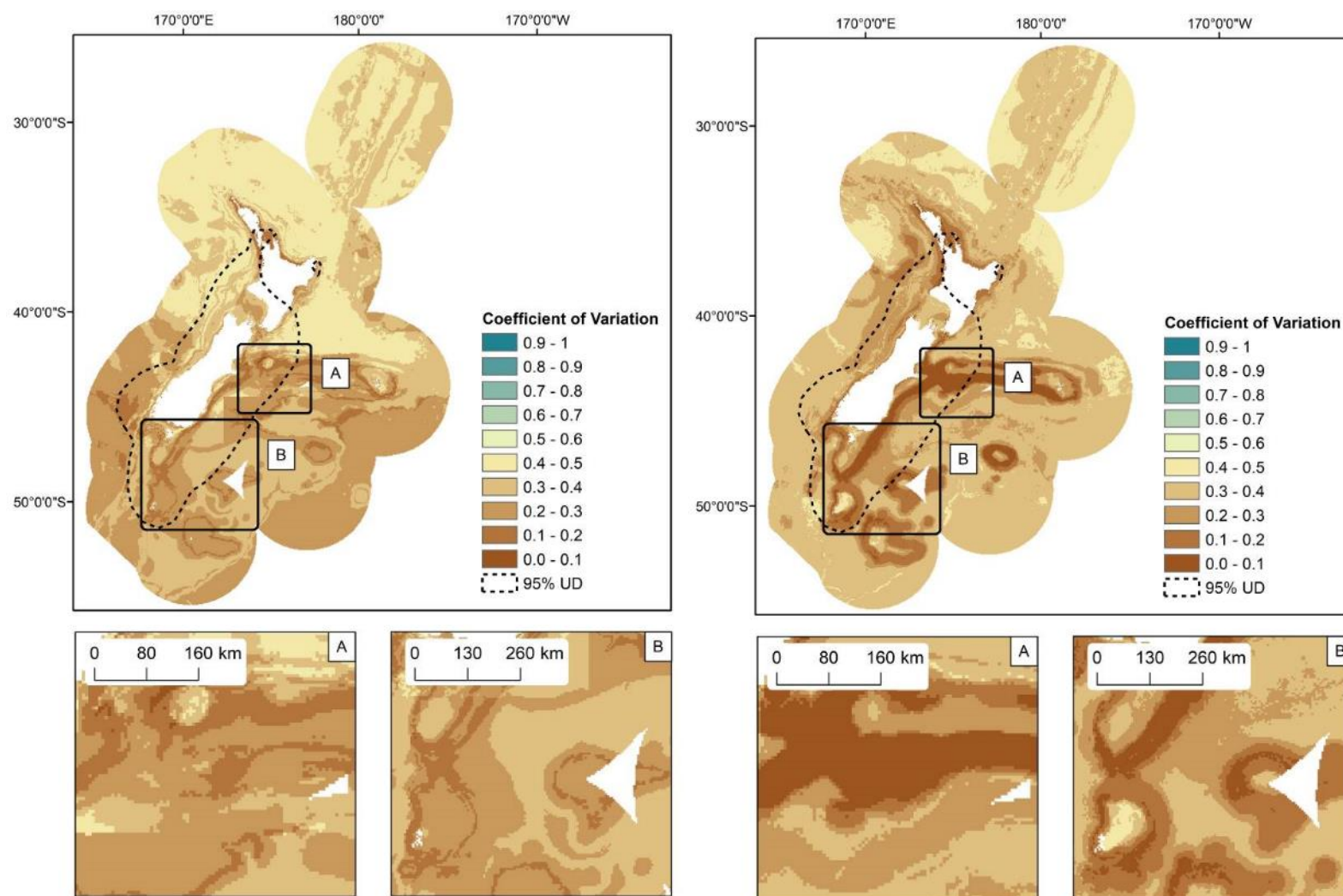


Figure 87 Seasonal uncertainty estimates (coefficient of variation, CV) of dusky dolphin (*Lagenorhynchus obscurus*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs fitted with Winter (May - Oct) (left) and Summer (Nov - Apr) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) west Chatham Rise; B) Campbell Plateau.

Bryde's whale (*Balaenoptera brydei*) – see section 3.2.3.5 in main body of the report

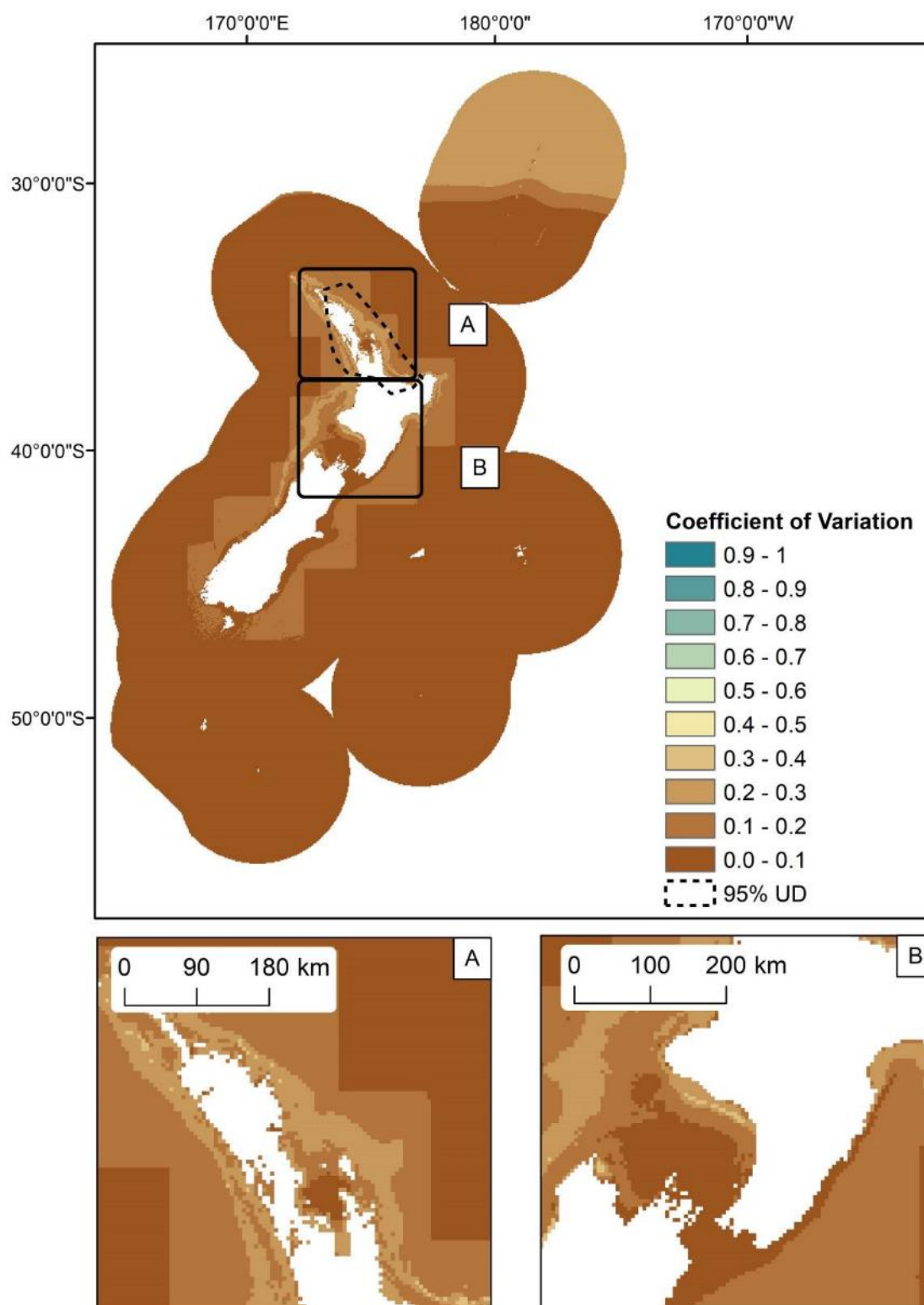


Figure 88: Uncertainty estimates (coefficient of variation, CV) of Bryde's whale (*Balaenoptera brydei*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed. Inset maps: A) north of North Island; B) Taranaki Bight and southeast of North Island.

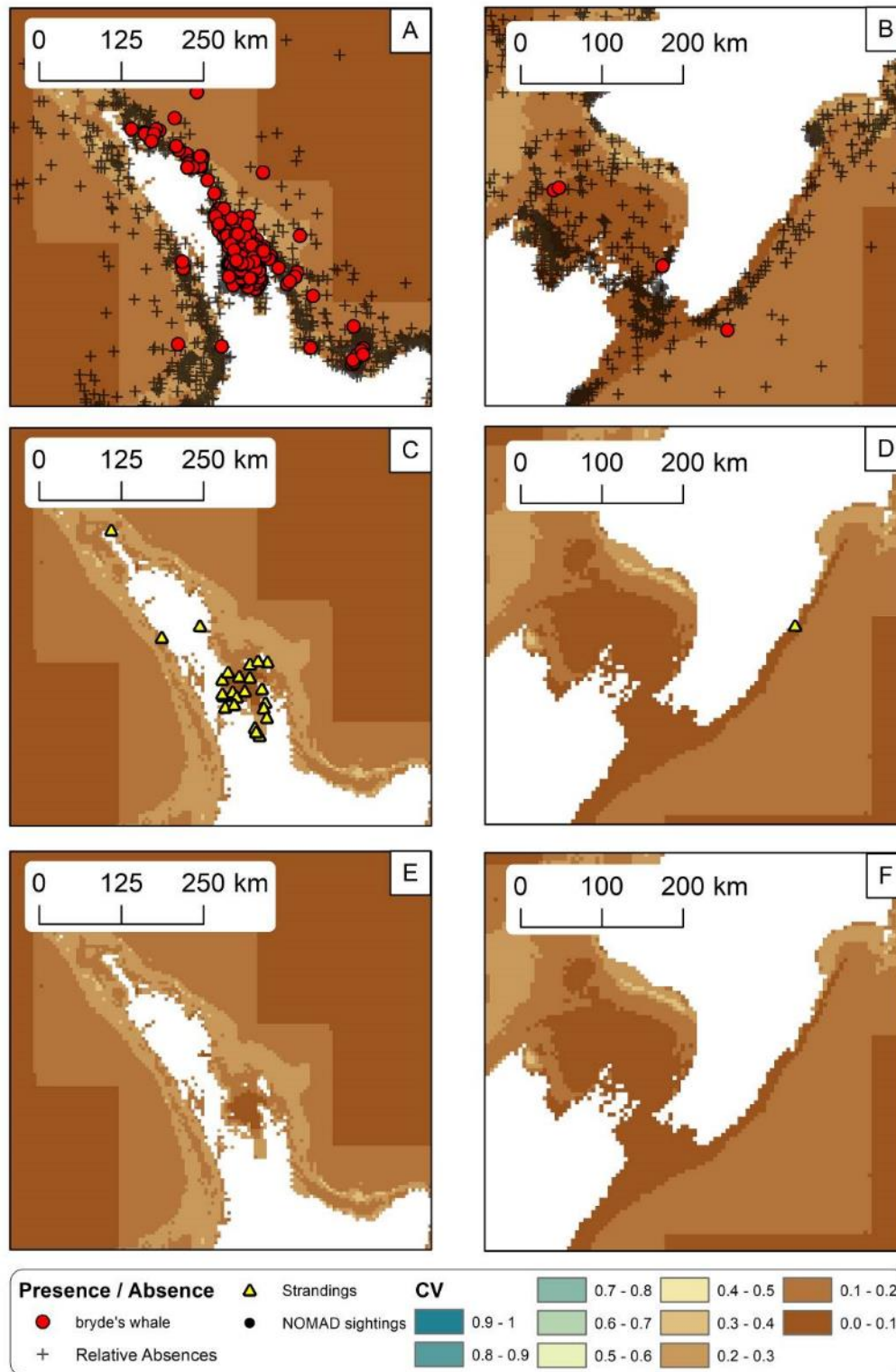


Figure 89: Uncertainty estimates (coefficient of variation, CV) of Bryde's whale (*Balaenoptera brydei*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Bryde's whale probability of presence models on the north of North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of Bryde's whale probability of presence models in the Taranaki Bight and southeast of North Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

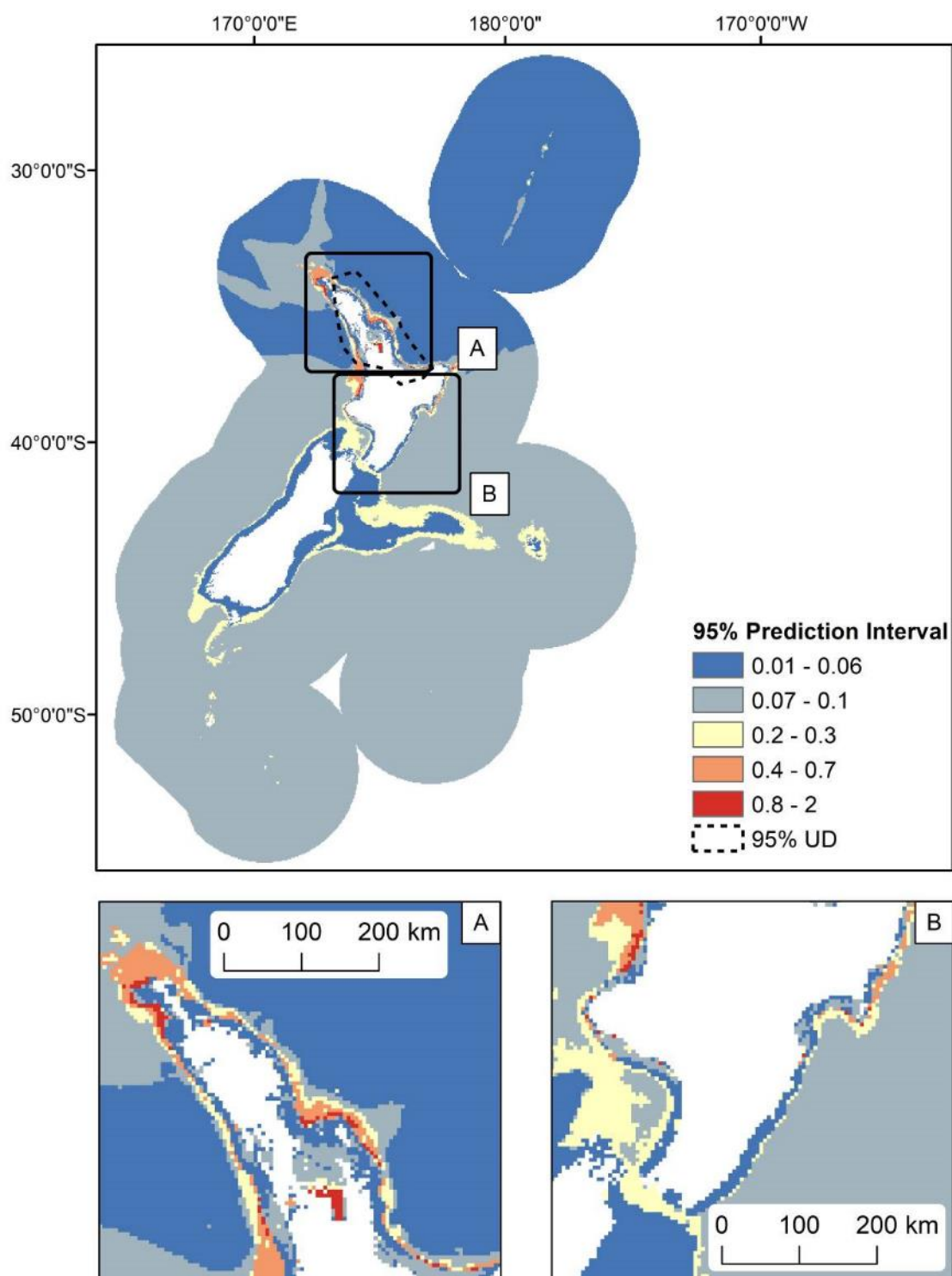


Figure 90: Uncertainty estimates (95% prediction interval) of Bryde's whale (*Balaenoptera brydei*) predicted relative densities in the New Zealand EEZ modelled using bootstrapped BRTs. 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island; B) Taranaki Bight and southeast of North Island.

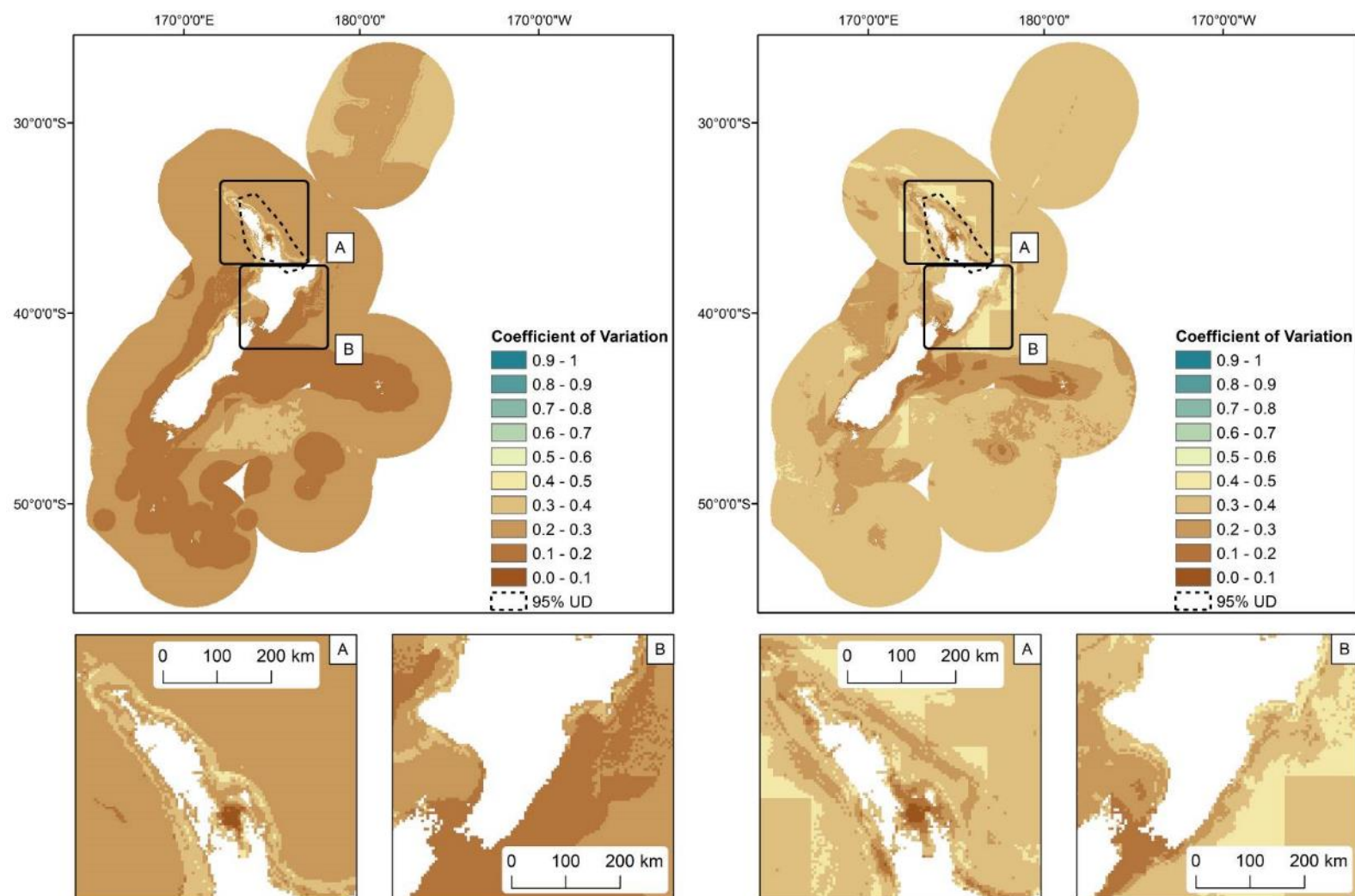


Figure 91: Seasonal uncertainty estimates (coefficient of variation, CV) of Bryde's whale (*Balaenoptera brydei*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs fitted with Winter (May - Oct) (left) and Summer (Nov - Apr) presence relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) north of North Island; B) Taranaki Bight and southeast of North Island.

Fin whale (*Balaenoptera physalus*)

Table 29: Fin whale (*Balaenoptera physalus*): Highlights of model fits and geographic prediction. Group size estimates and seasonal distributions were not modelled due to low number of records.

Sample number	Distribution	P/RA		Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
Low	Cosmopolitan	Good	Fair	NA	NA	NA

Table 30: Model performance measures (deviance explained and AUC) for BRT models of fin whale (*Balaenoptera physalus*) presence/relative absence. Measures of error are not available because the low number of presence records did not allow bootstrapping.

	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Mean	0.20	0.19	0.81	0.90
Standard Deviation	NA	NA	NA	NA

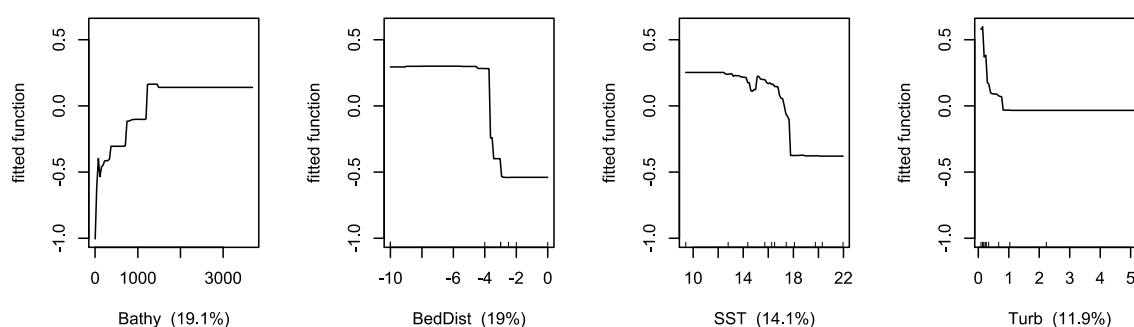


Figure 92: Partial dependence plots showing the relationships between predictor variables and probability presence of fin whale (*Balaenoptera physalus*) modelled using BRTs. The four most influential environmental predictors in the model are shown. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

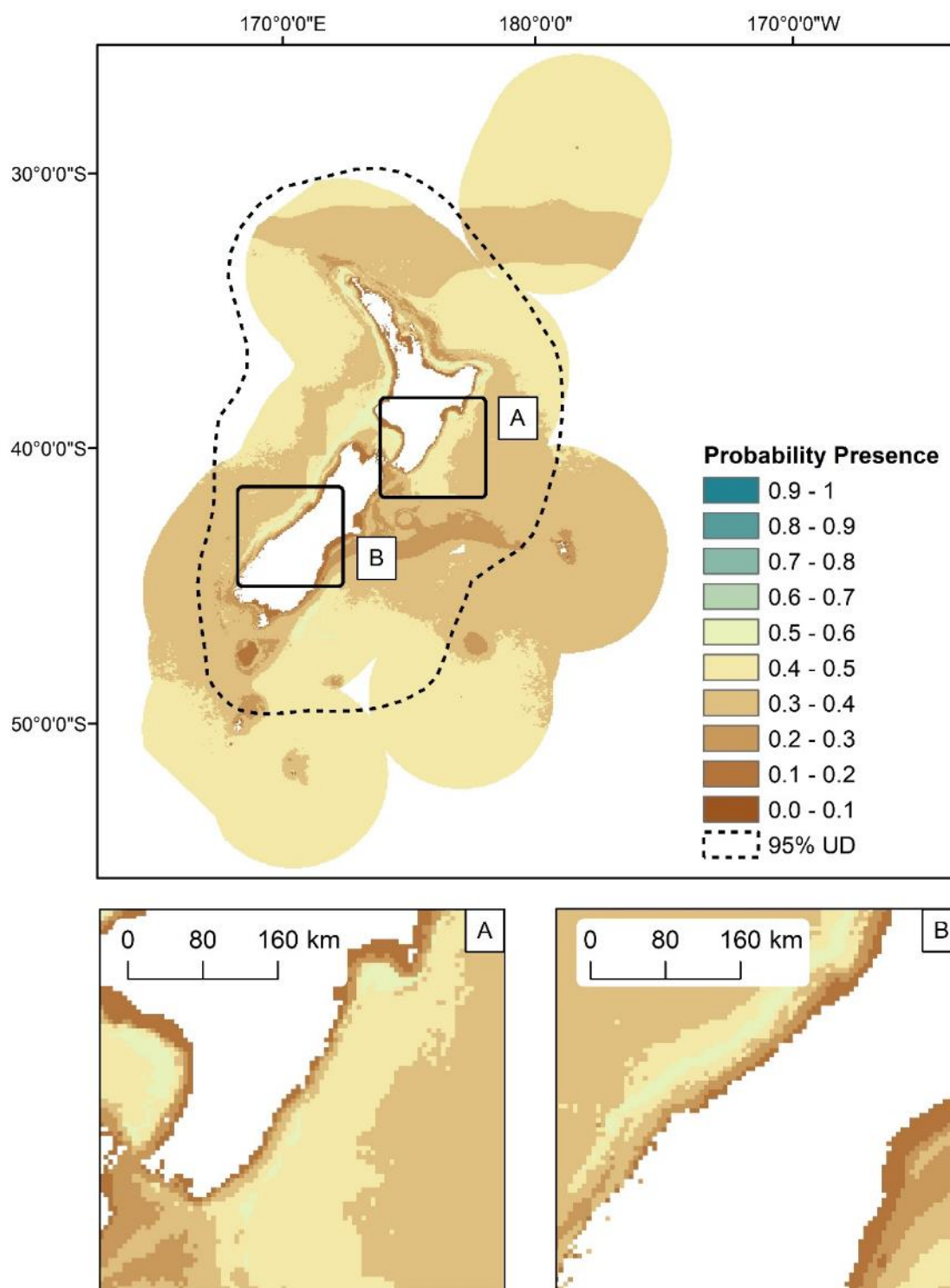


Figure 93: The predicted probability presence of fin whale (*Balaenoptera physalus*) in the New Zealand EEZ modelled using BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) South North Island; B) West Coast and south Canterbury Coast, South Island.

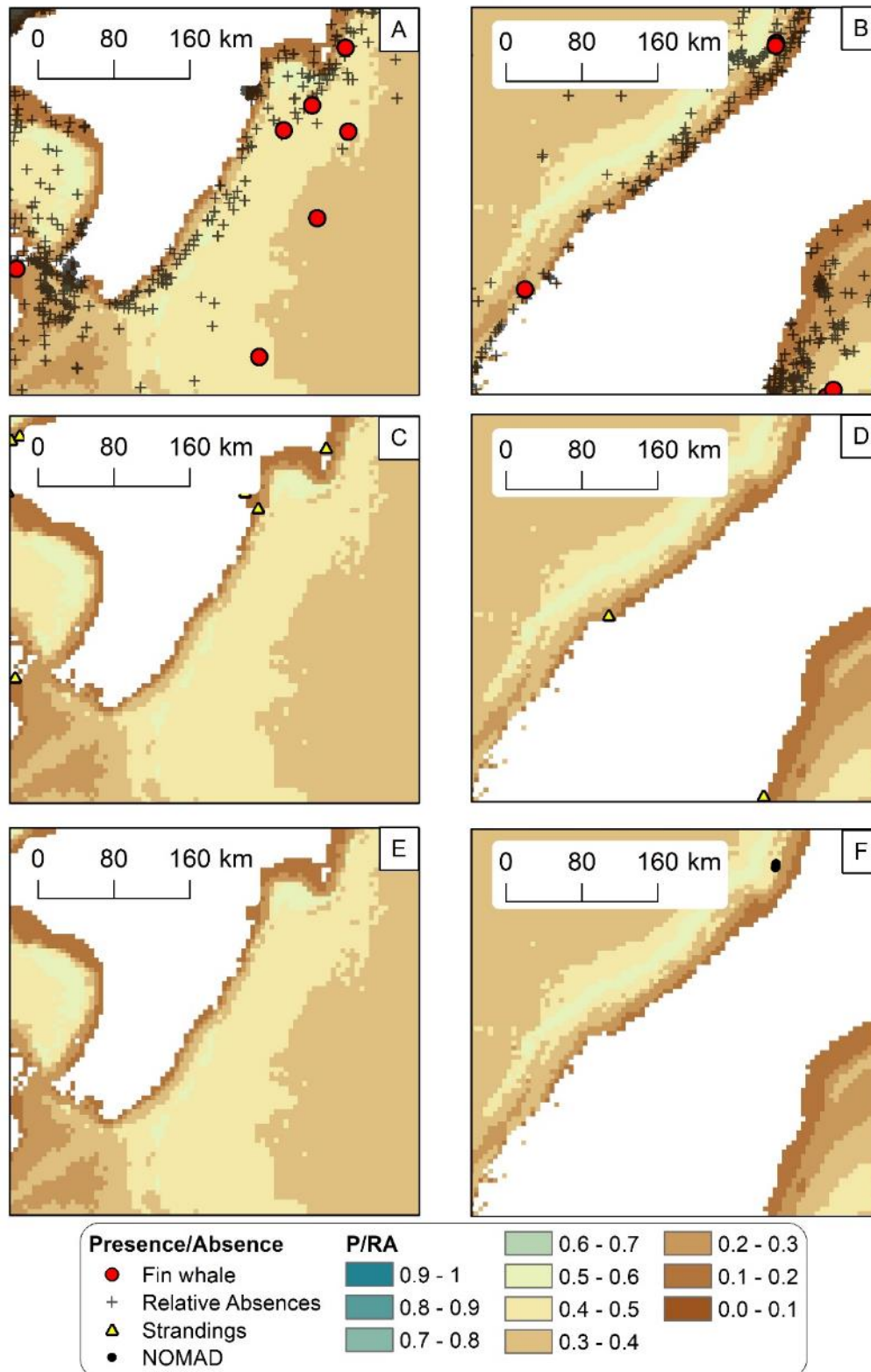


Figure 94: The predicted probability presence of fin whale (*Balaenoptera physalus*) modelled using BRTs. Predicted probability presence of fin whale in the south North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted probability presence of Fin whale on the West Coast and south Canterbury Coast, South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

Minke whale (*Balaenoptera acutorostrata*)

Table 31: Minke whale (*Balaenoptera acutorostrata*): Highlights of model fits and geographic prediction. Group size estimates and seasonal distributions were not modelled due to low number of records.

Sample number	Distribution	P/RA		Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
Low	Cosmopolitan	Fair	Fair	NA	NA	NA

Table 32: Model performance measures (deviance explained and AUC) for BRT models of minke whale (*Balaenoptera acutorostrata*) presence / relative absence. Measures of error are not available because the low number of presence records did not allow bootstrapping.

	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Mean	0.18	0.25	0.79	0.88
Standard Deviation	NA	NA	NA	NA

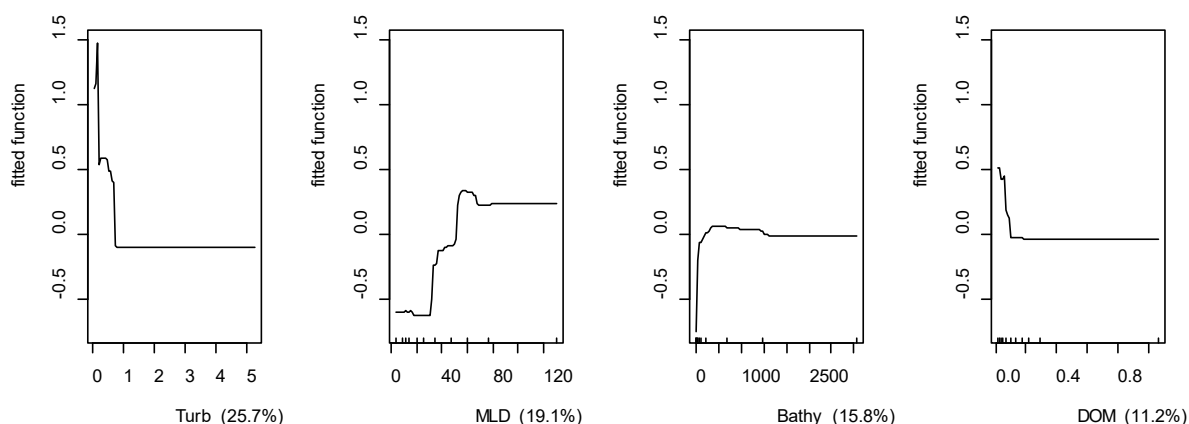


Figure 95: Partial dependence plots showing the relationships between predictor variables and probability presence of minke whale (*Balaenoptera acutorostrata*) modelled using BRTs. The four most influential environmental predictors in the model are shown. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

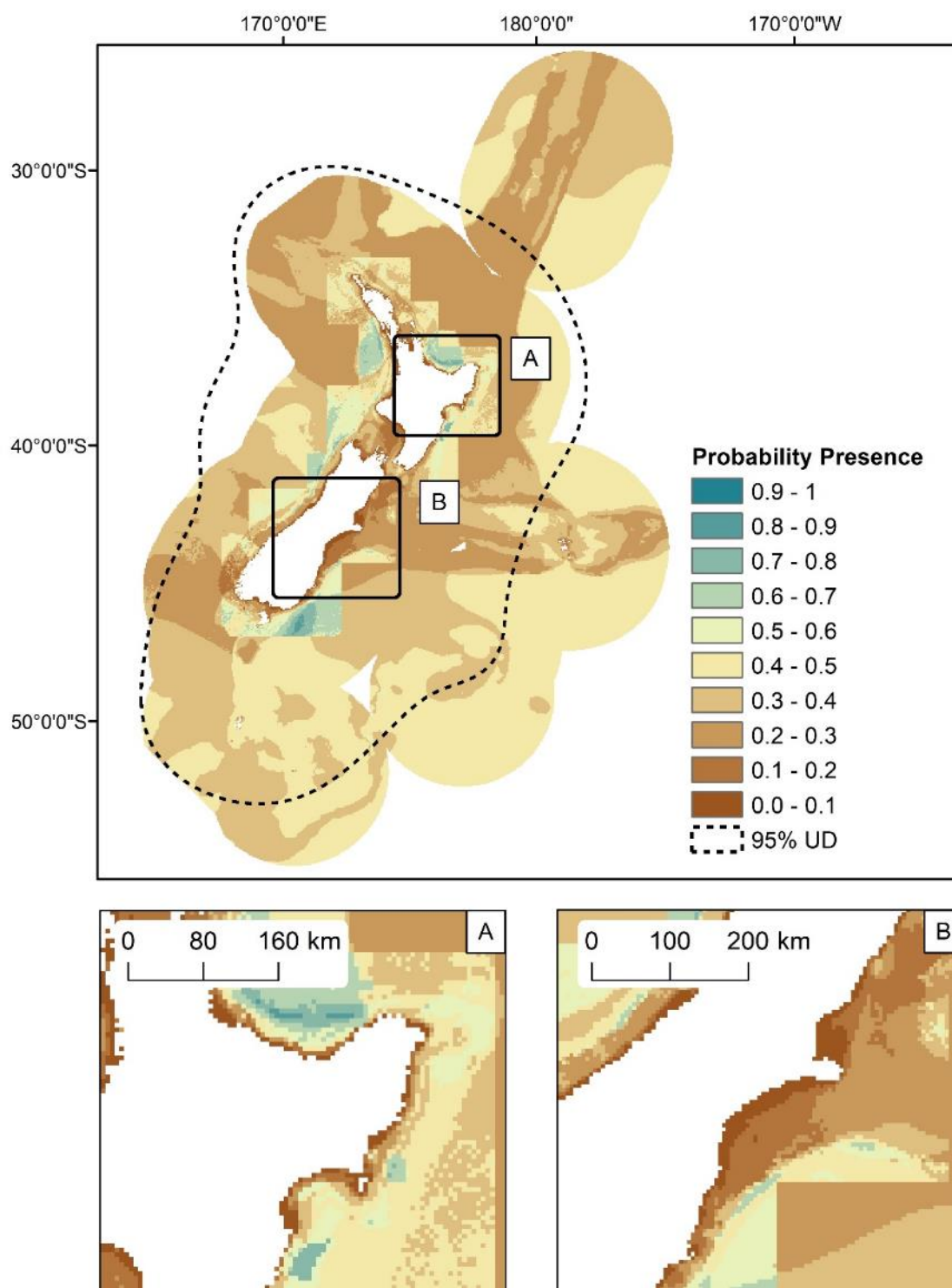


Figure 96: The predicted probability of minke whale (*Balaenoptera acutorostrata*) presence in the New Zealand EEZ modelled using BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) East coast of the North Island; B) East and west coasts of the South Island.

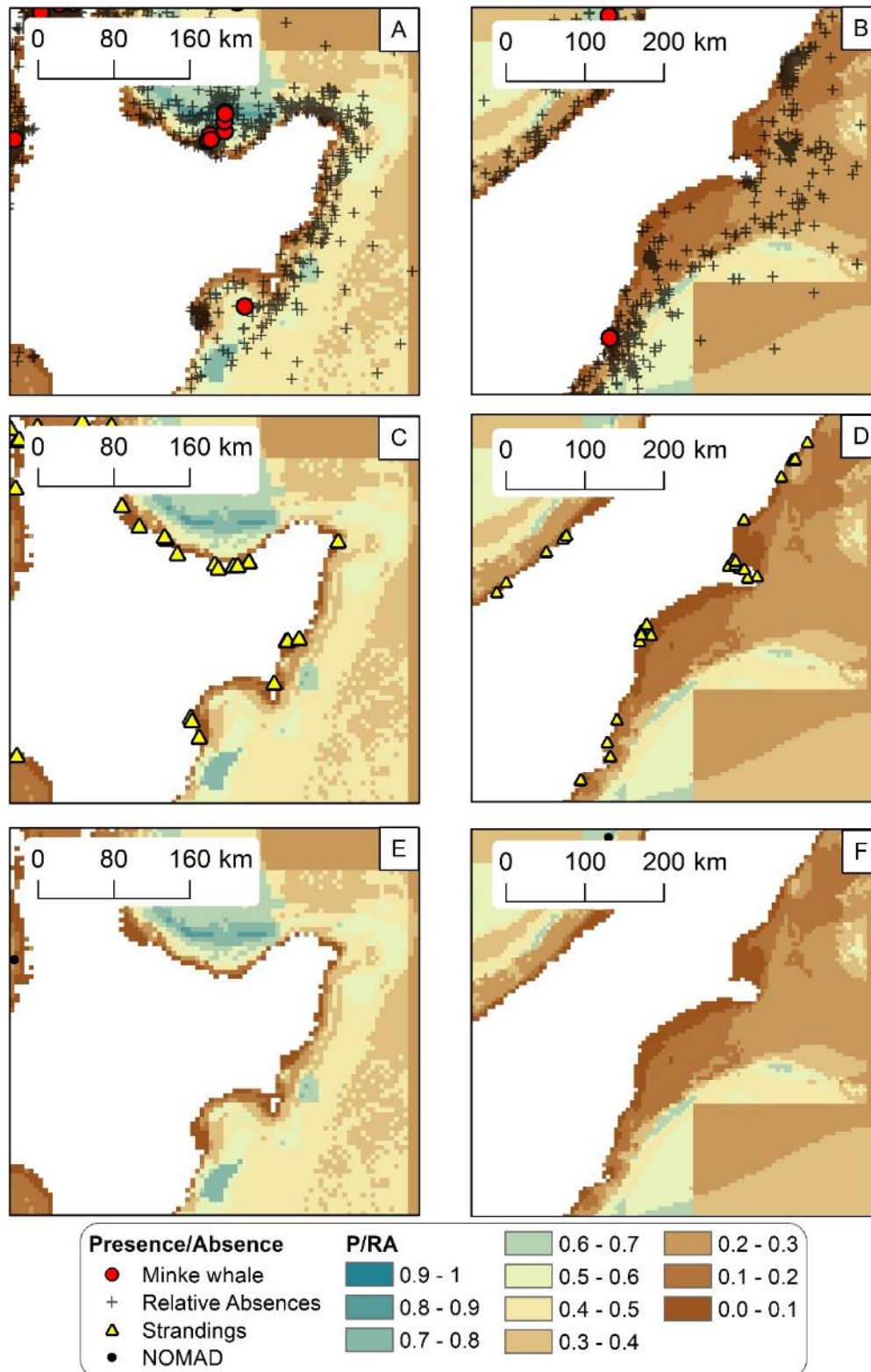


Figure 97: The predicted probability of minke whale (*Balaenoptera acutorostrata*) presence modelled using BRTs. Predicted probability presence of Minke whale on the east coast of the North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted probability presence of Minke whale on west and east coasts of the South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

Sei whale (*Balaenoptera borealis*)

Table 33: Sei whale (*Balaenoptera borealis*): Highlights of model fits and geographic prediction. Group size estimates and seasonal distributions were not modelled due to low number of records.

Sample number	Distribution	P/RA		Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
Low	Cosmopolitan	Good	Good	NA	NA	NA

Table 34: Mean model performance measures (deviance explained and AUC) for BRT models of sei whale (*Balaenoptera borealis*) presence/relative absence. Measures of error are not available because the low number of presence records did not allow bootstrapping.

	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Mean	0.24	0.24	0.81	0.88
Standard Deviation	NA	NA	NA	NA

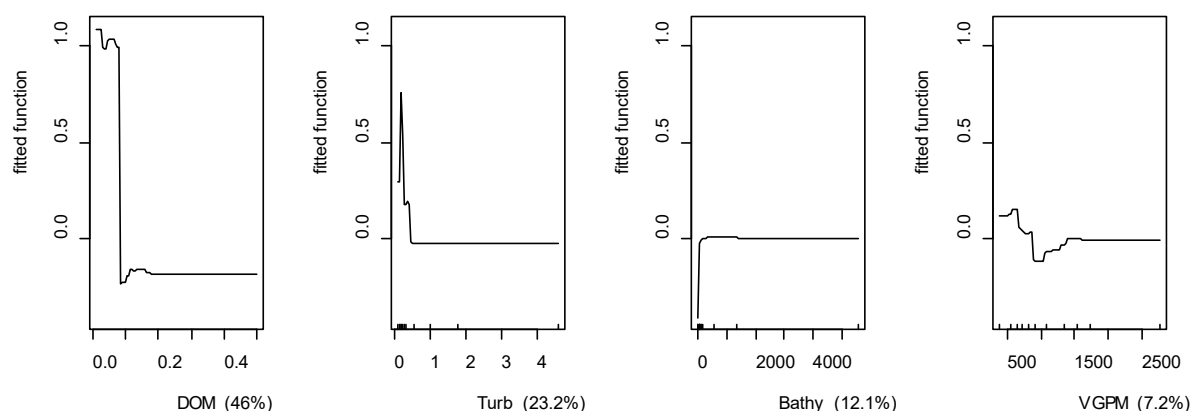


Figure 98: Partial dependence plots showing the relationships between predictor variables and probability presence of sei whale (*Balaenoptera borealis*) modelled using BRTs. The four most influential environmental predictors in the model are shown. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

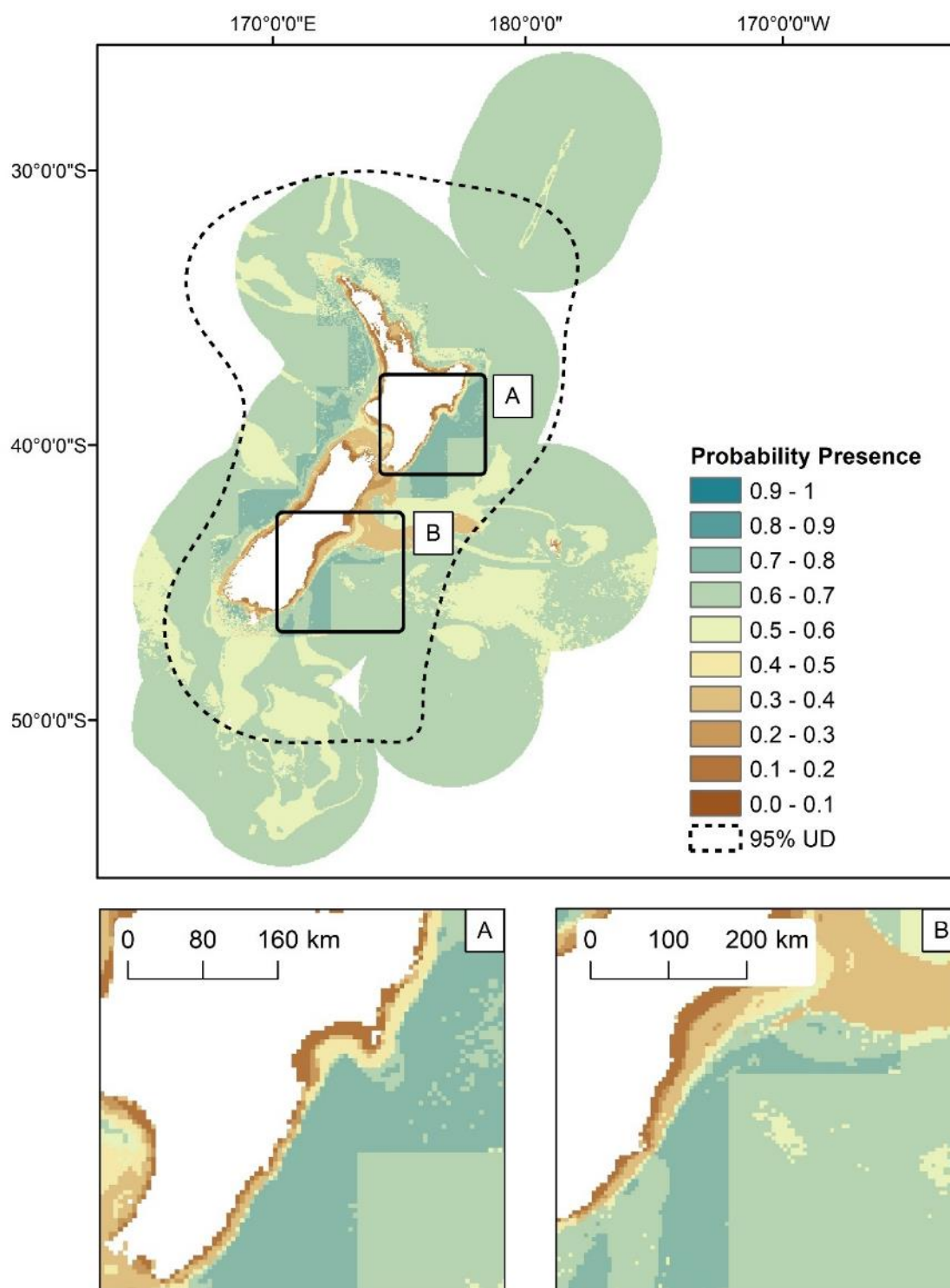


Figure 99: The predicted probability of sei whale (*Balaenoptera borealis*) presence in the New Zealand EEZ modelled using BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) South east coast of the North Island; B) South East coast of the South Island.

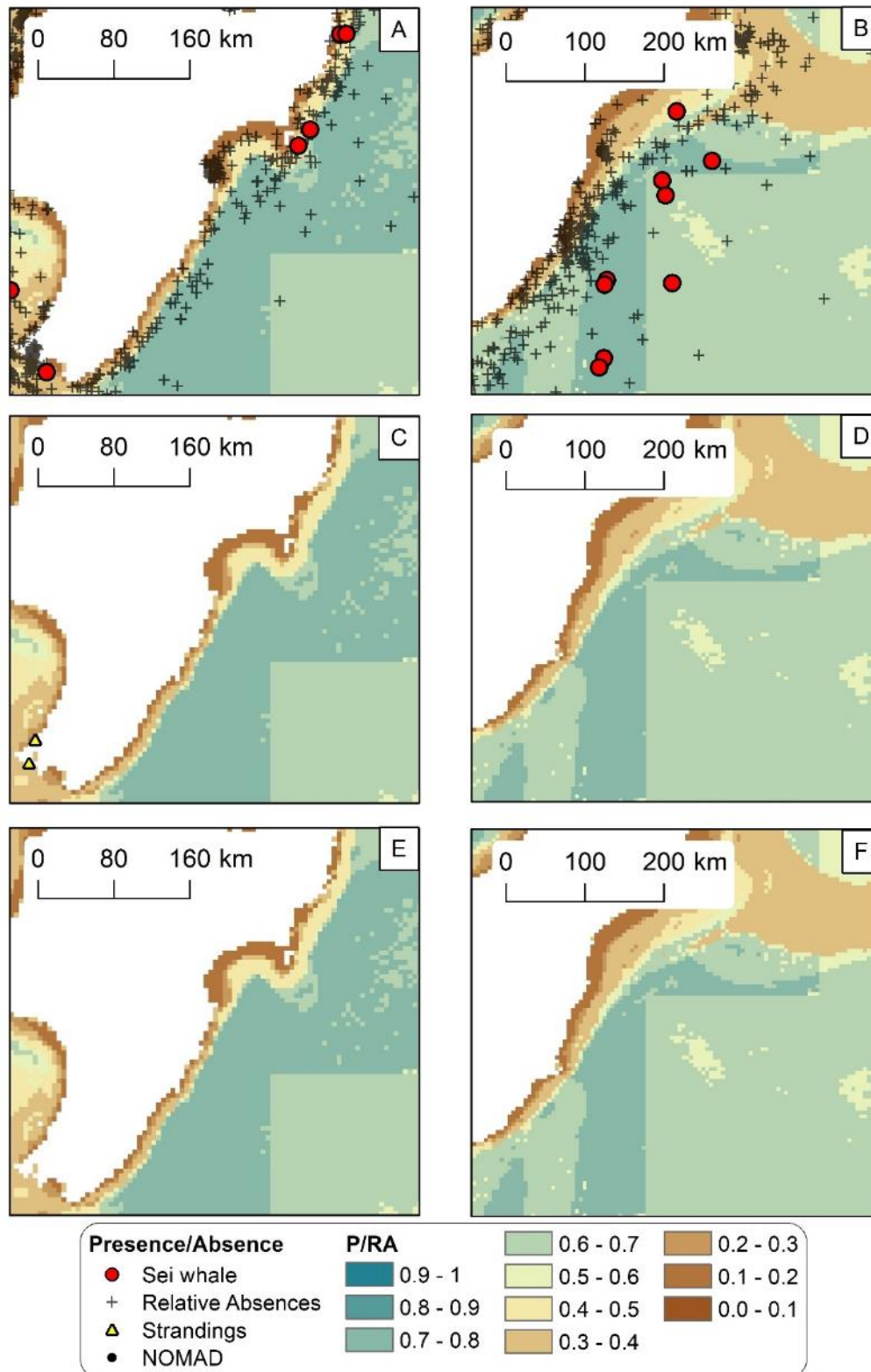


Figure 100: The predicted probability of sei whale (*Balaenoptera borealis*) presence modelled using BRTs. Predicted probability presence of Sei whale on the south east coast of the North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted probability presence of Sei whale on the south east coast of the South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

Blue whale (*Balaenoptera musculus*)

Table 35: Blue whale (*Balaenoptera musculus*): Highlights of model fits and geographic prediction. Group size estimates models did not converge and there were too few records in winter to model seasonal differences.

Sample number	Distribution	P/RA		Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
High	Localised – High probability presence in Taranaki Bight	Excellent	Excellent	NA	NA	NA

Table 36 Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of blue whale (*Balaenoptera musculus*).

	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Mean	0.58	0.58	0.95	0.95
Standard Deviation	0.02	0.05	0.01	0.01

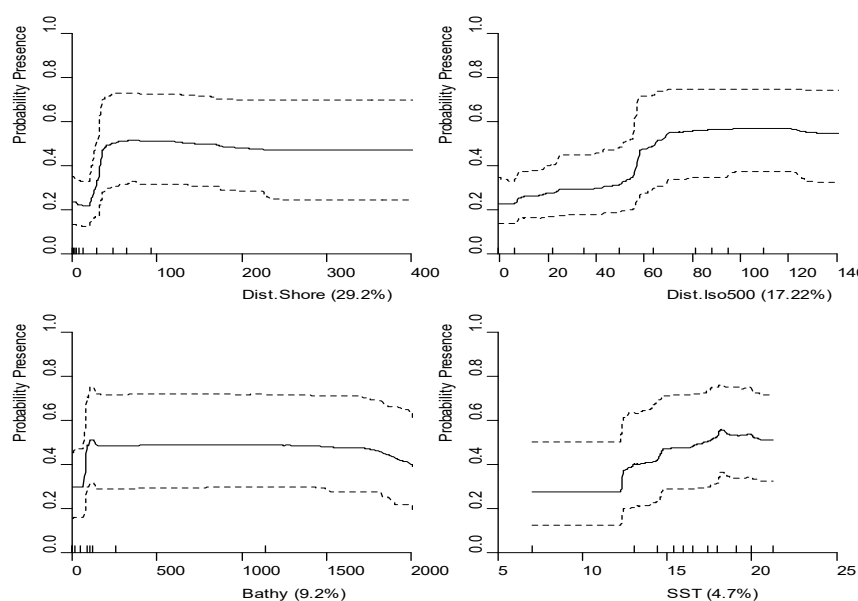


Figure 101: Partial dependence plots showing the relationships between predictor variables and probability presence of blue whale (*Balaenoptera musculus*) modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

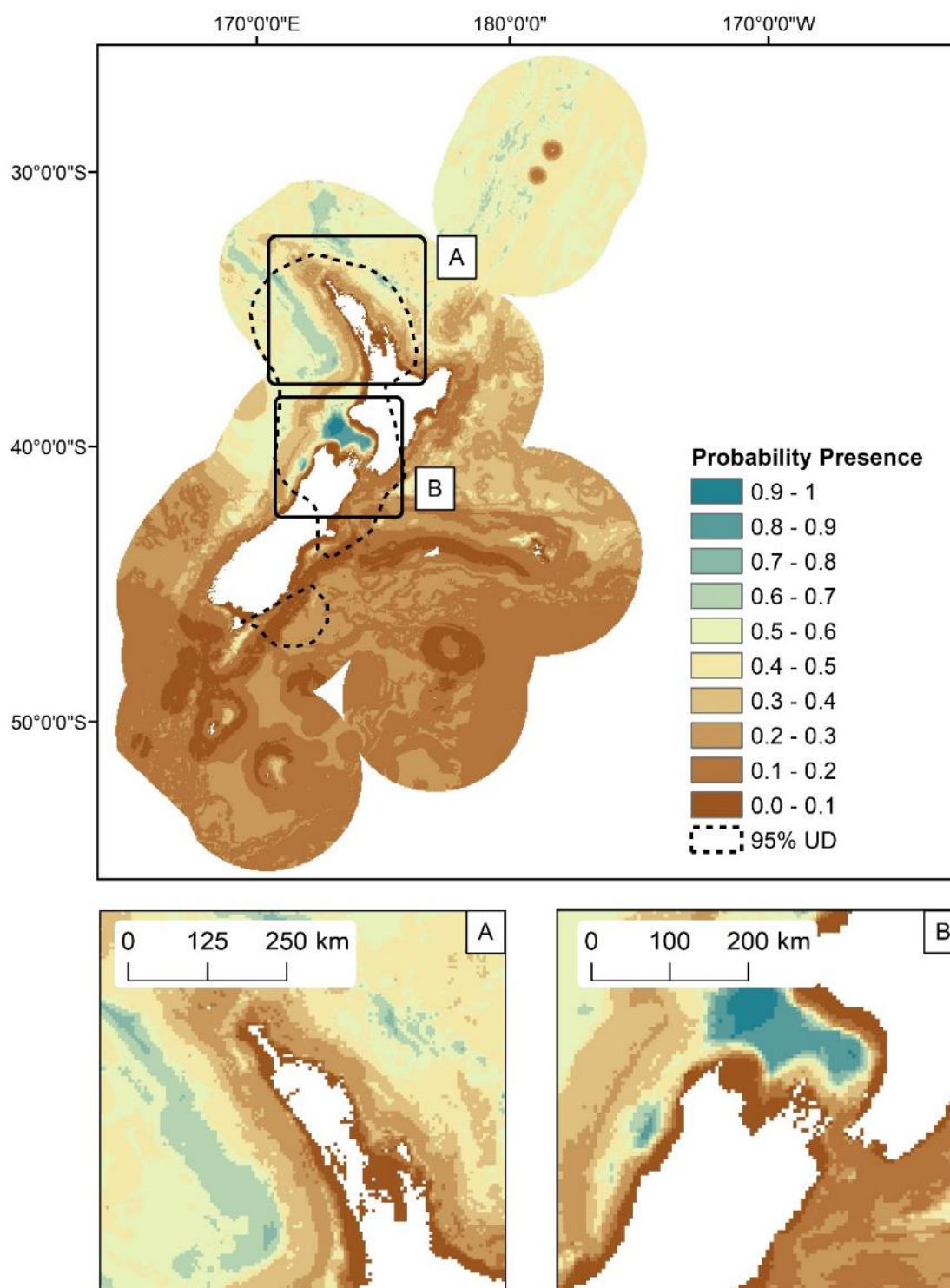


Figure 102: The predicted probability of blue whale (*Balaenoptera musculus*) presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island and Hauraki Gulf; B) Southern North Island and northern South Island including the Cook Strait.

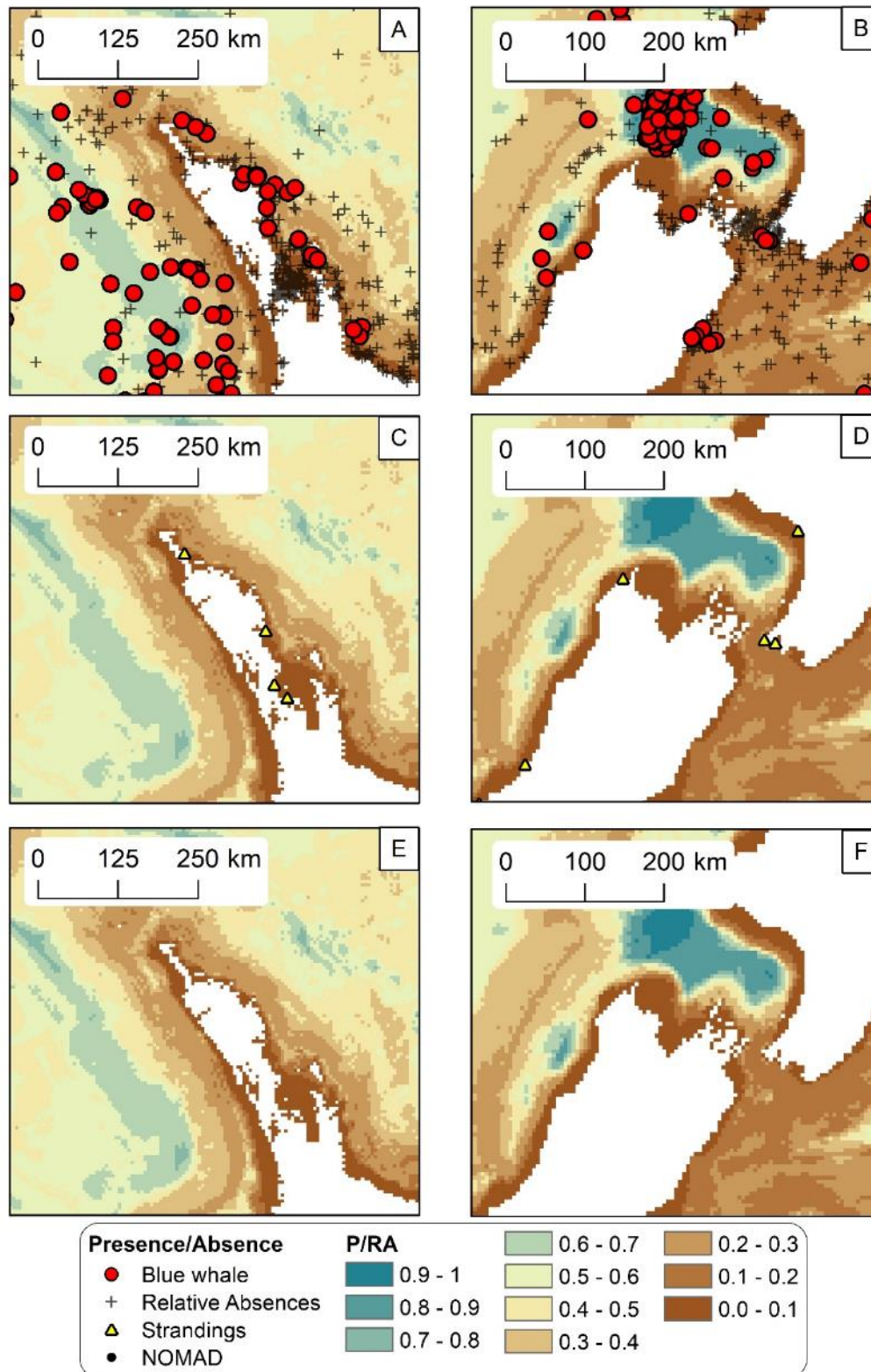


Figure 103: The predicted probability of blue whale (*Balaenoptera musculus*) presence modelled using bootstrapped BRTs. Predicted probability presence of Blue whale in the north of North Island and Hauraki Gulf are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted probability presence of Blue whale on southern North Island and northern South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

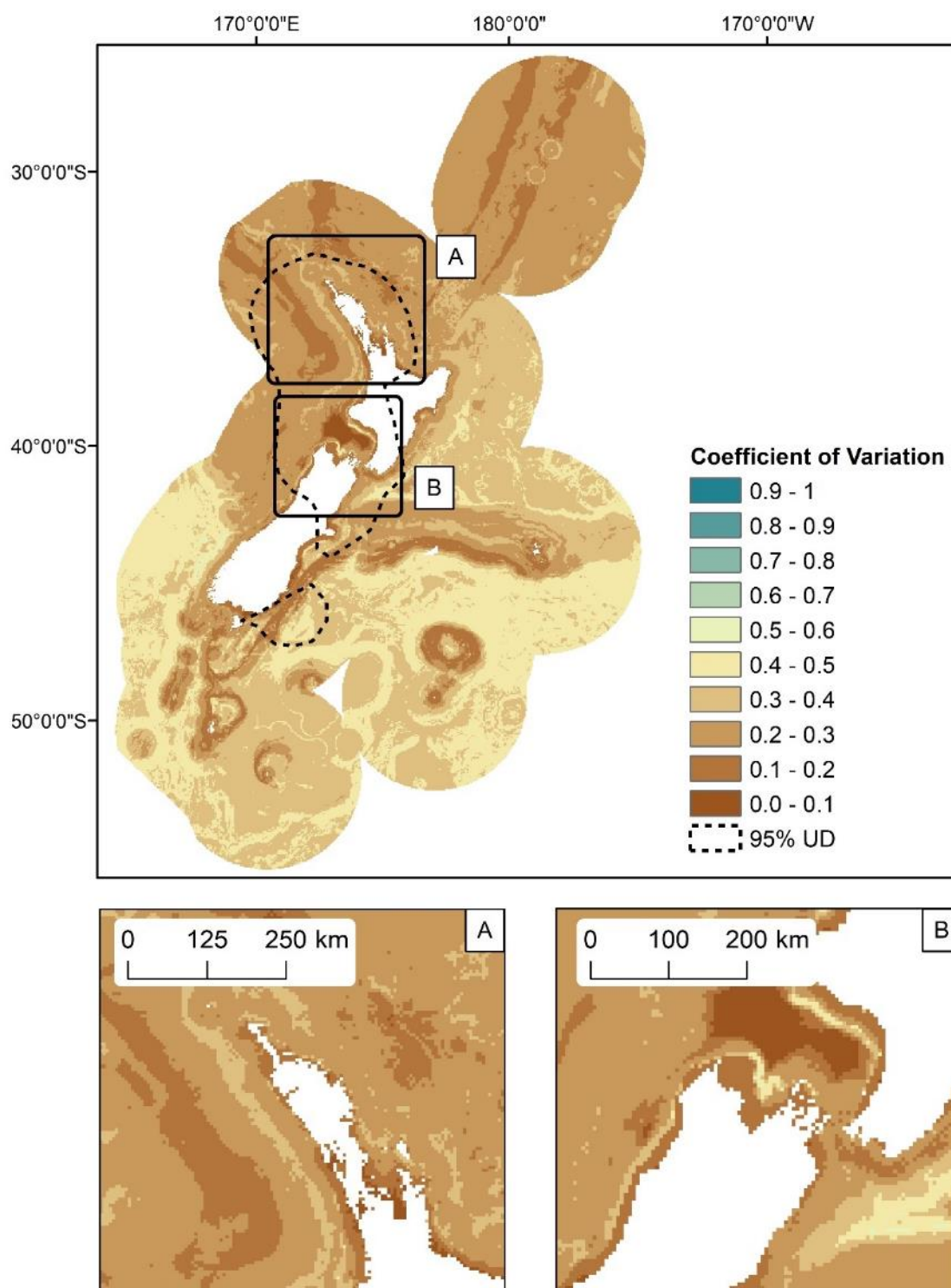


Figure 104: Uncertainty estimates (coefficient of variation, CV) of blue whale (*Balaenoptera musculus*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) north North Island and Hauraki Gulf; B) Southern North Island and northern South Island including the Cook Strait.

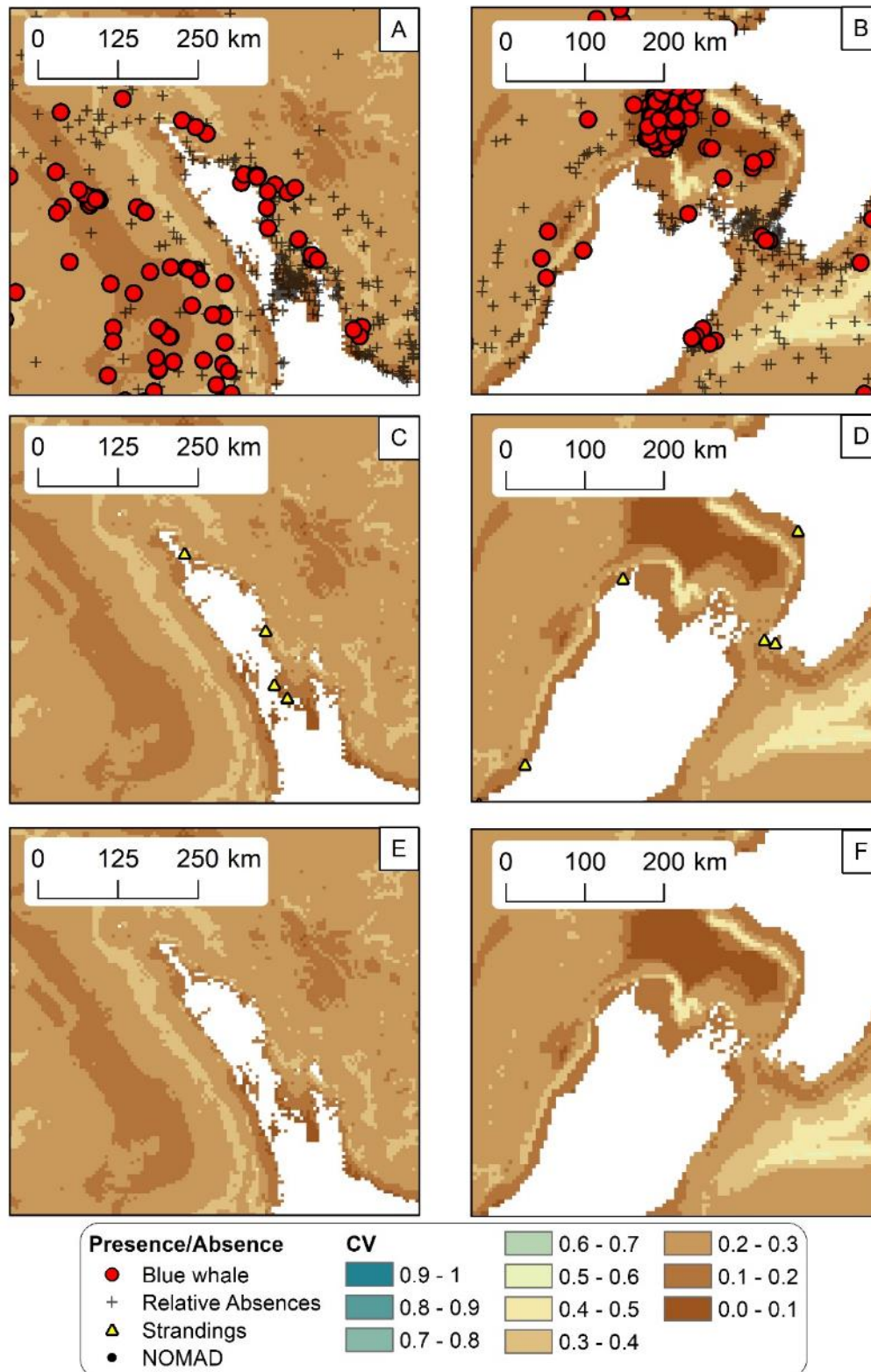


Figure 105: Uncertainty estimates (coefficient of variation, CV) of blue whale (*Balaenoptera musculus*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Blue whale probability presence models in the north of North Island and Hauraki Gulf are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of Blue whale probability presence models on the southern North Island and northern South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

Right whale (*Eubalaena australis*)

Table 37: Right whale (*Eubalaena australis*): Highlights of model fits and geographic prediction. There were too few records in summer to model seasonal differences

Sample number	Distribution	P/RA		Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
High	Localised – Southern areas in EEZ	Excellent	Excellent	Fair	Good	NA

Table 38: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of right whale (*Eubalaena australis*).

	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Mean	0.53	0.53	0.94	0.94
Standard Deviation	0.02	0.04	0.01	0.01

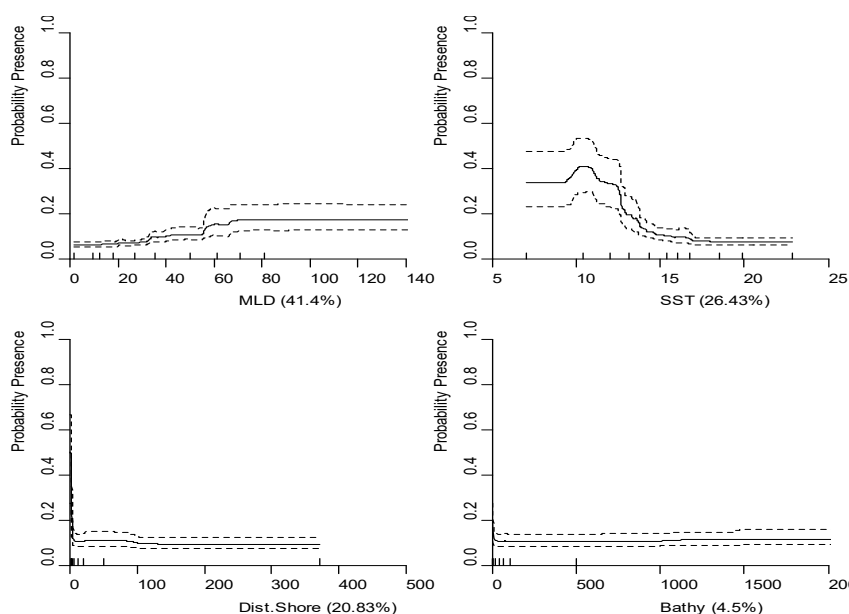


Figure 106: Partial dependence plots showing the relationships between predictor variables and probability presence of right whale (*Eubalaena australis*) modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

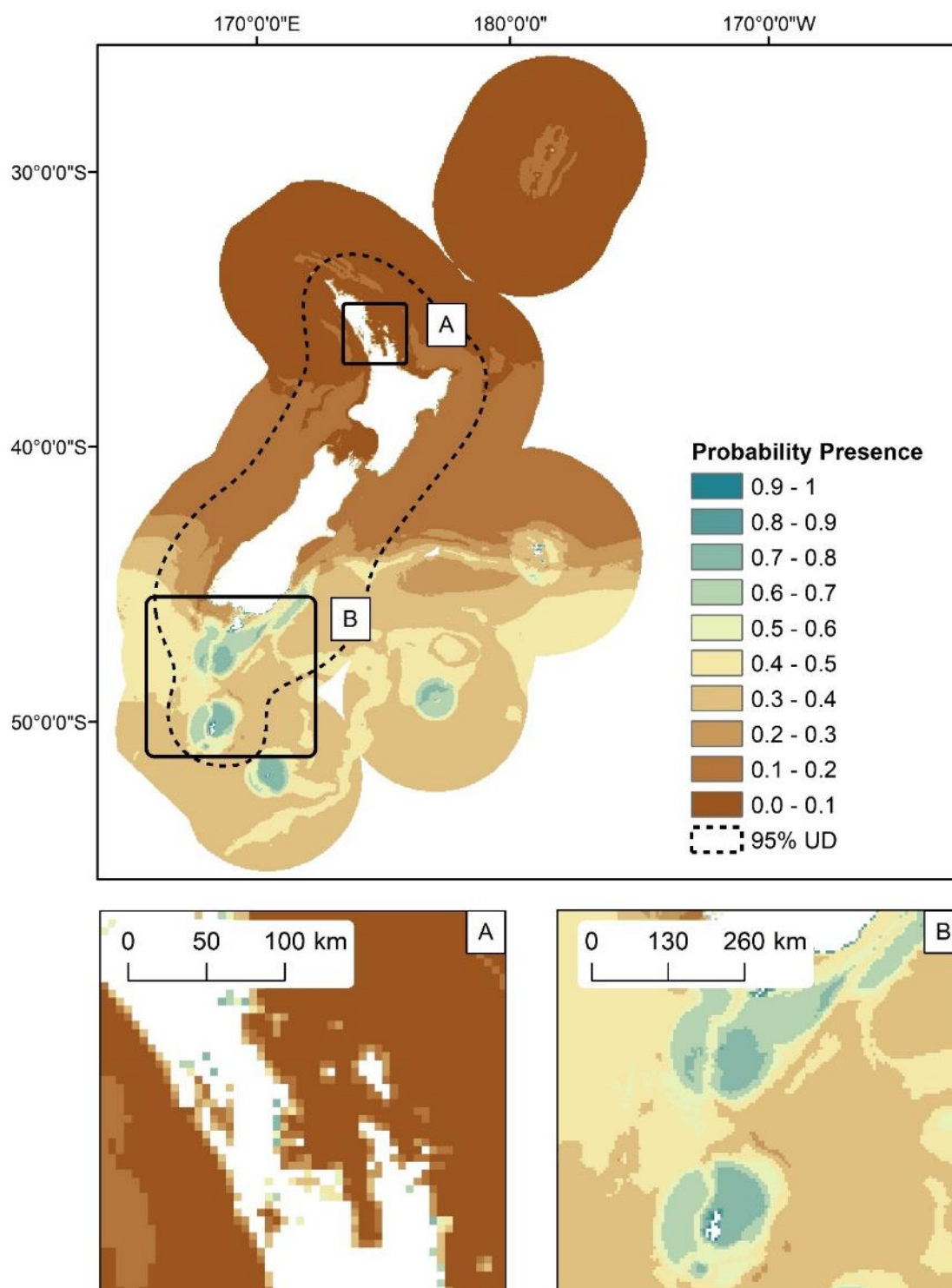


Figure 107: The predicted probability of right whale (*Eubalaena australis*) presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) Hauraki Gulf, North Island; B) South of the South Island.

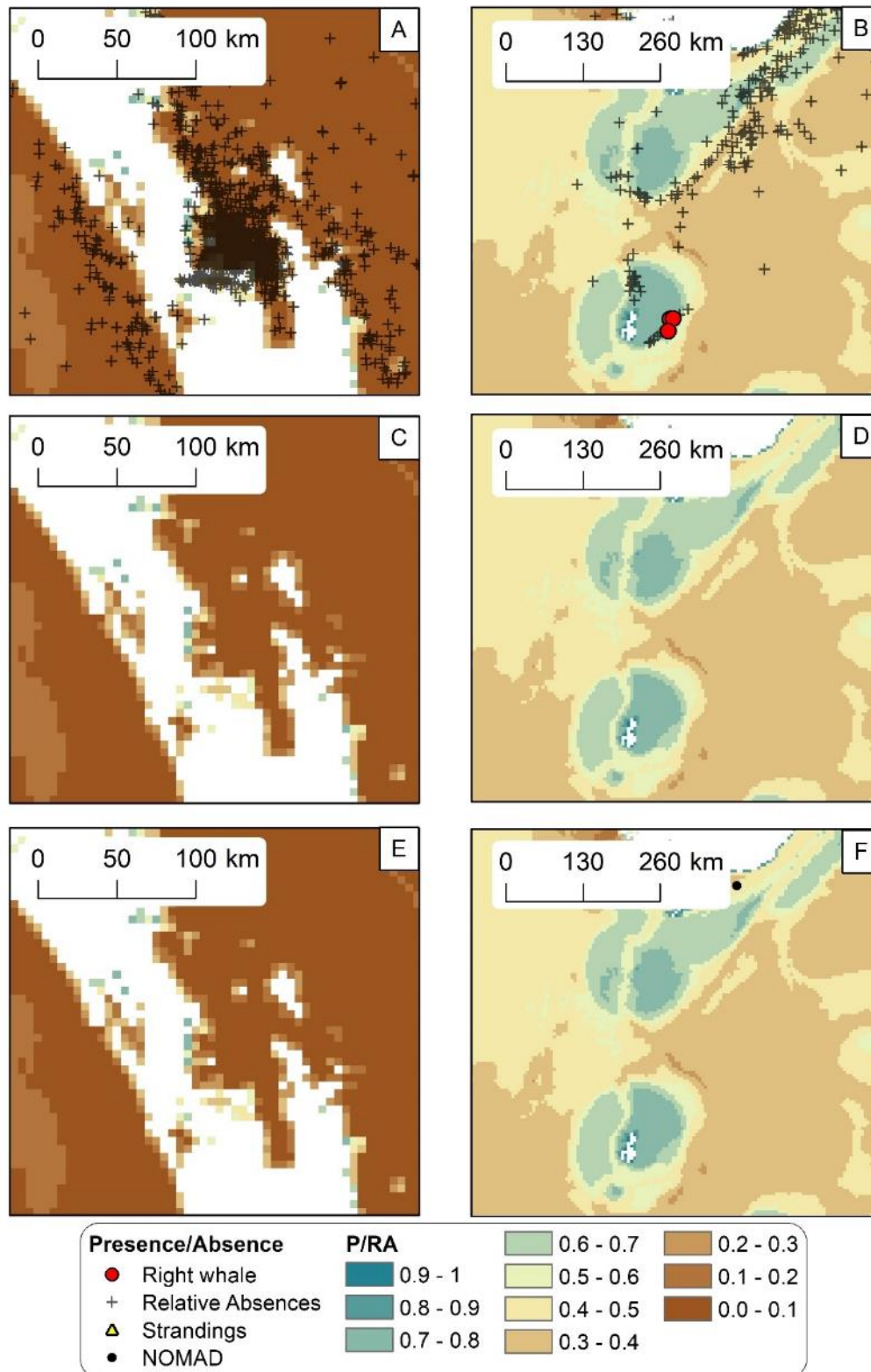


Figure 108: The predicted probability of right whale (*Eubalaena australis*) presence modelled using bootstrapped BRTs. Predicted probability presence of Right whale in the Hauraki Gulf, North Island, are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted probability presence of Right whale south of the South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

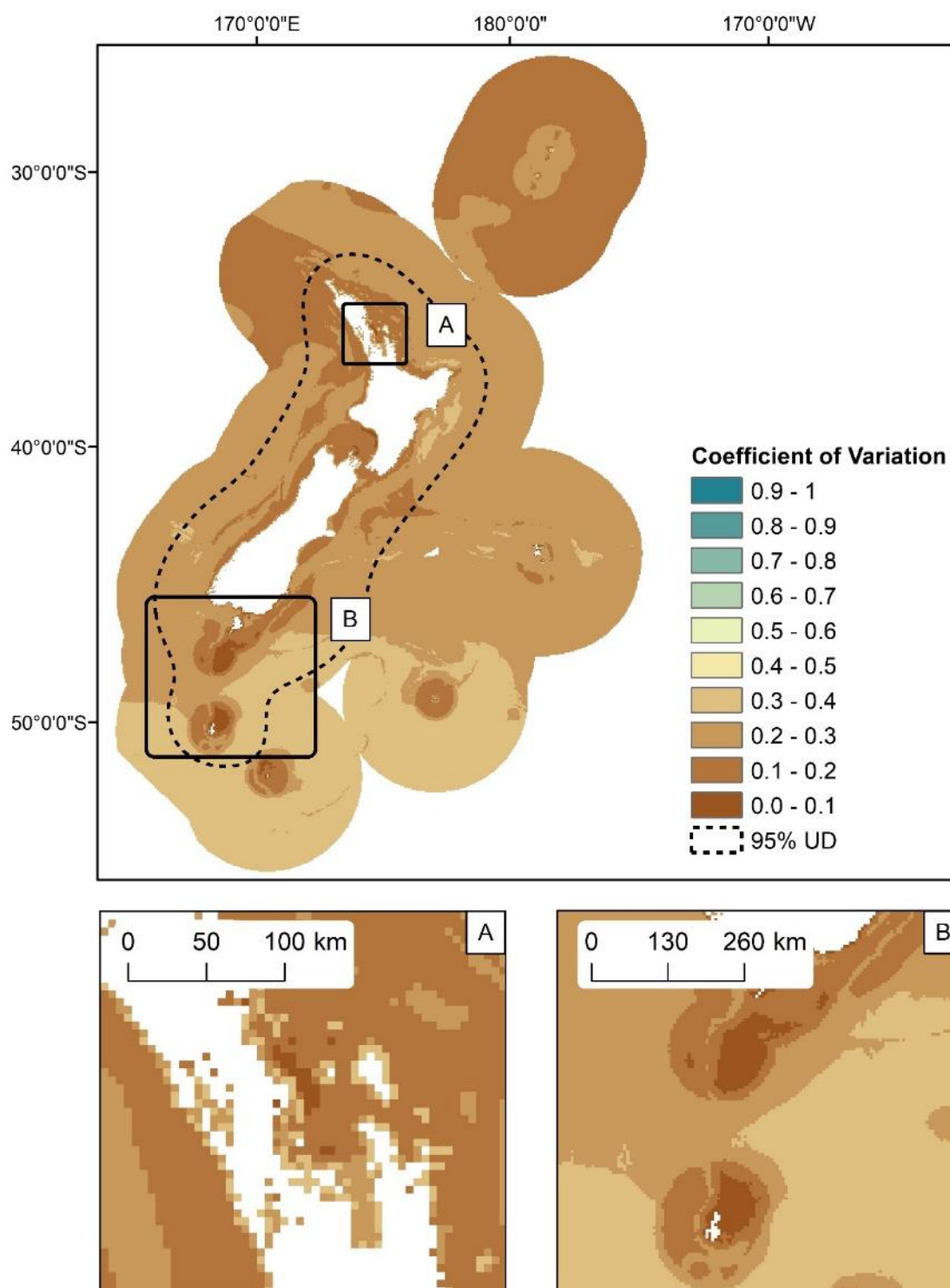


Figure 109: Uncertainty estimates (coefficient of variation, CV) of right whale (*Eubalaena australis*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) Hauraki Gulf, North Island; B) South of the South Island.

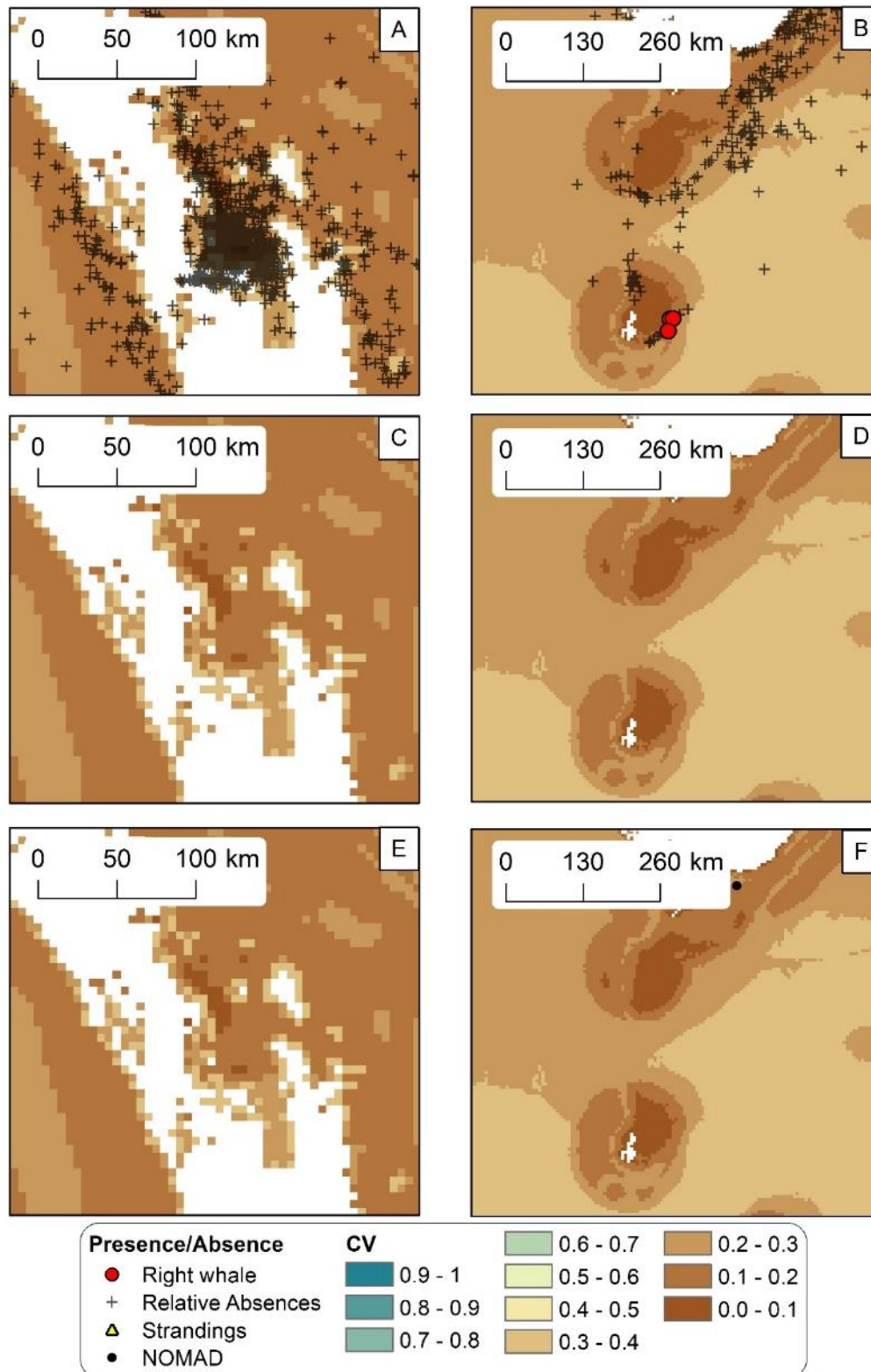


Figure 110: Uncertainty estimates (coefficient of variation, CV) of right whale (*Eubalaena australis*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Right whale probability presence models in the Hauraki Gulf, North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of Right whale probability presence models in the south of the South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

Table 39: Mean model performance measures (deviance explained and Pearson's correlation of predicted vs observed group sizes (R^2)) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of right whale (*Eubalaena australis*).

	Deviance explained (training data)	Deviance explained (validation data)	Pearson's correlation of predicted vs observed species group sizes (R^2)
Mean	0.14	0.08	0.27
Standard Deviation	0.06	0.12	0.14

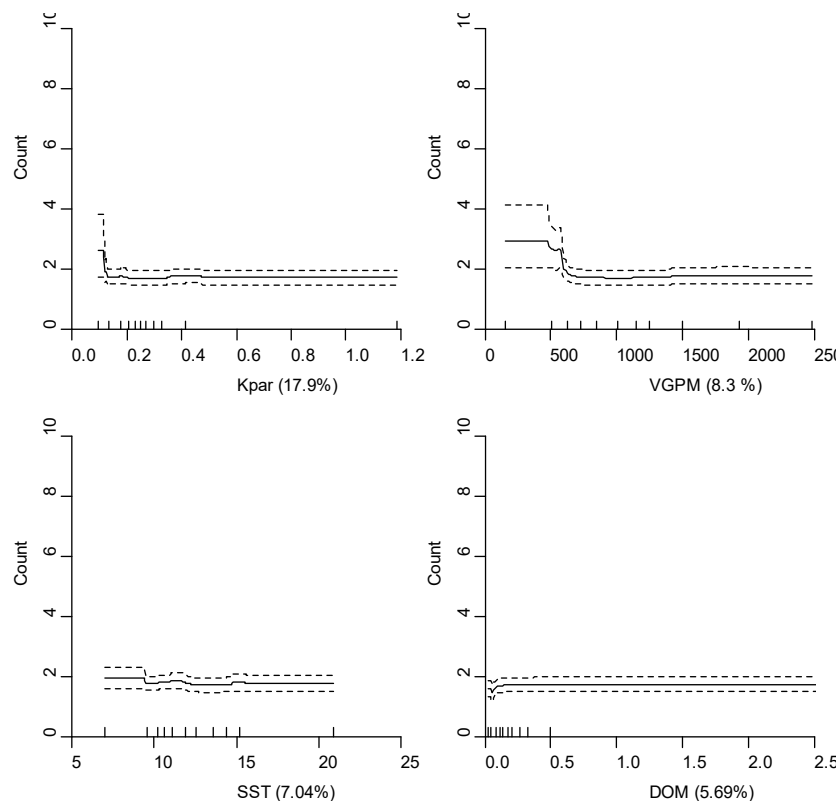


Figure 111: Partial dependence plots showing the relationships between predictor variables and right whale (*Eubalaena australis*) group sizes modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x axes).

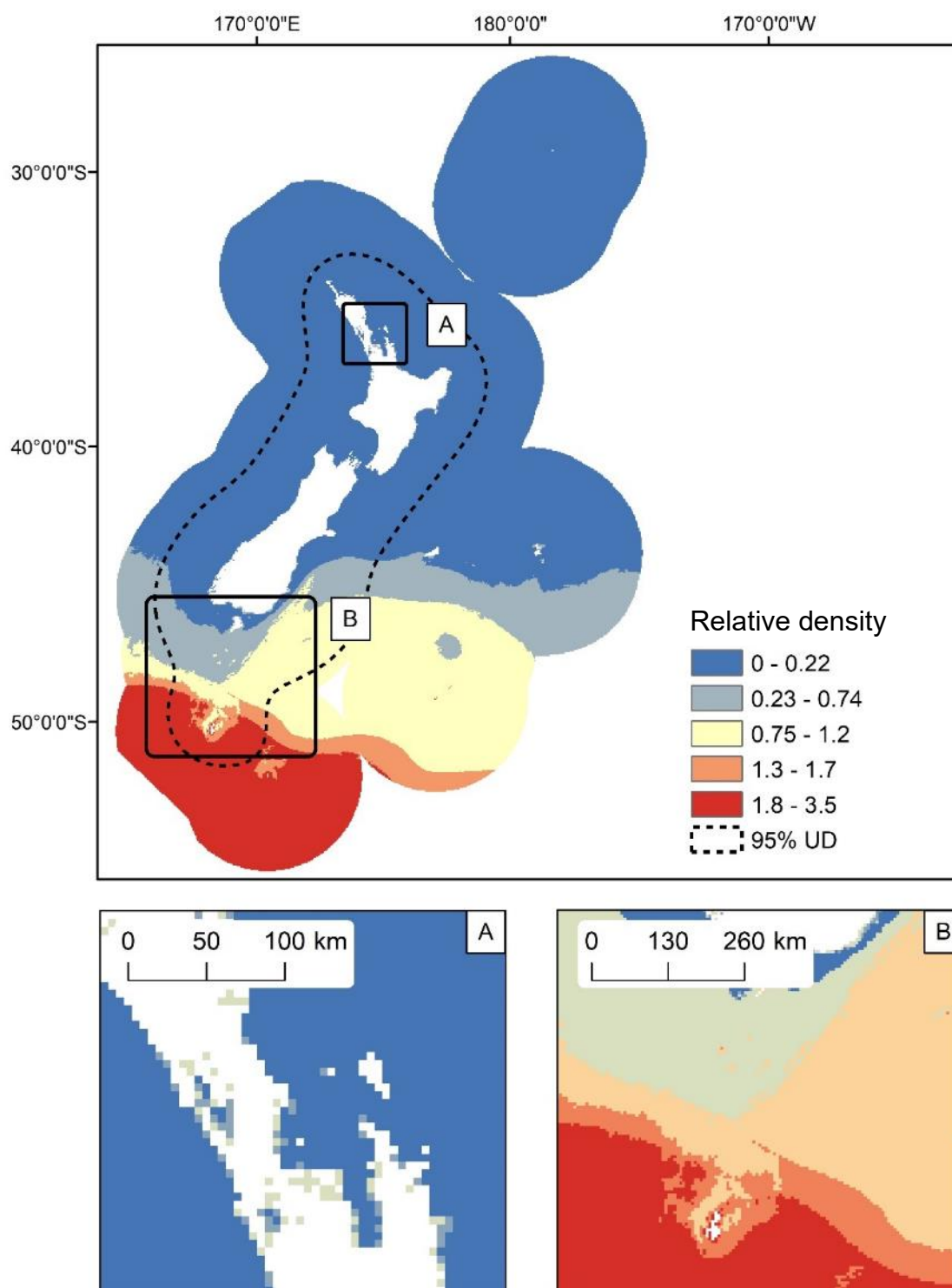


Figure 112: Predicted relative density of right whale (*Eubalaena australis*) in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) Hauraki Gulf, North Island; B) South of the South Island.

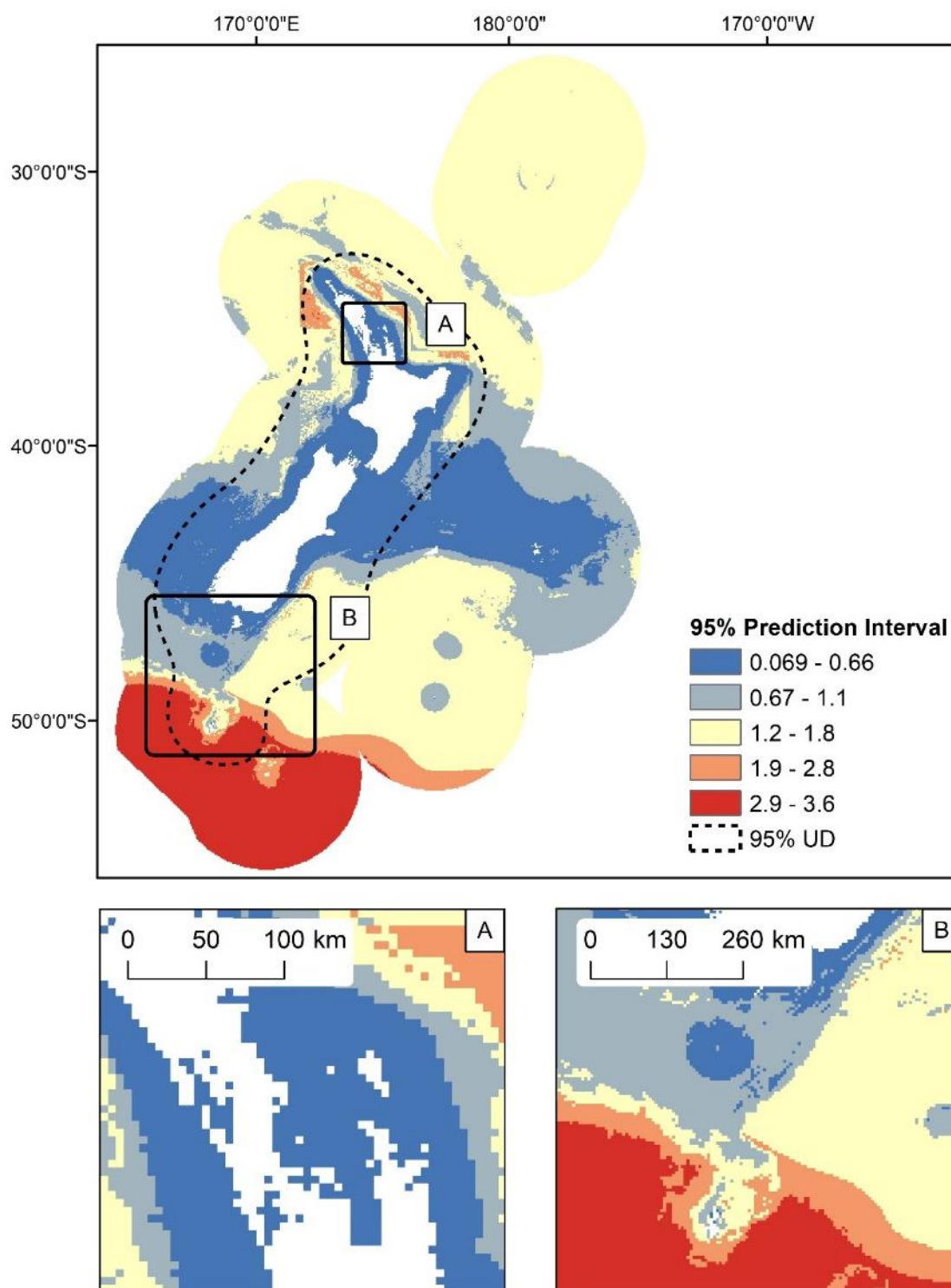


Figure 113: Uncertainty estimates (95% prediction interval) of right whale (*Eubalaena australis*) predicted relative density in the New Zealand EEZ modelled using bootstrapped BRTs. 95% utilisation distribution is shown as dashed line. Inset maps: A) Hauraki Gulf, North Island; B) South of the South Island.

Humpback whale (*Megaptera novaeangliae*)

Table 40: Humpback whale (*Megaptera novaeangliae*): Highlights of model fits and geographic prediction. Group size estimates models did not converge.

Sample number	Distribution	P/RA		Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
High	Cosmopolitan – highest probability presence close to shore	Good	Good	NA	NA	Inshore – offshore shift in habitat use from winter to summer

Table 41: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of humpback whale (*Megaptera novaeangliae*).

	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Mean	0.29	0.28	0.85	0.85
Standard Deviation	0.02	0.03	0.01	0.02

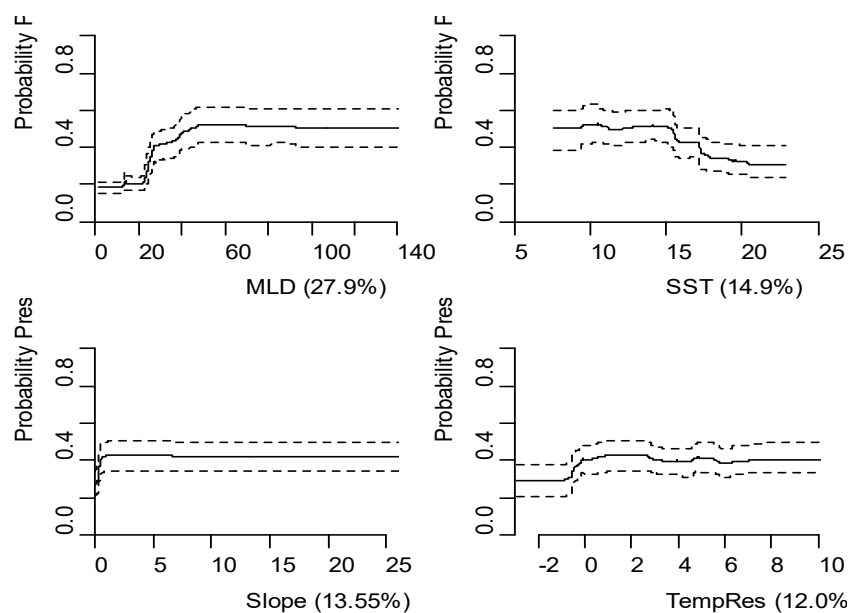


Figure 114: Partial dependence plots showing the relationships between predictor variables and probability of humpback whale (*Megaptera novaeangliae*) presence modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

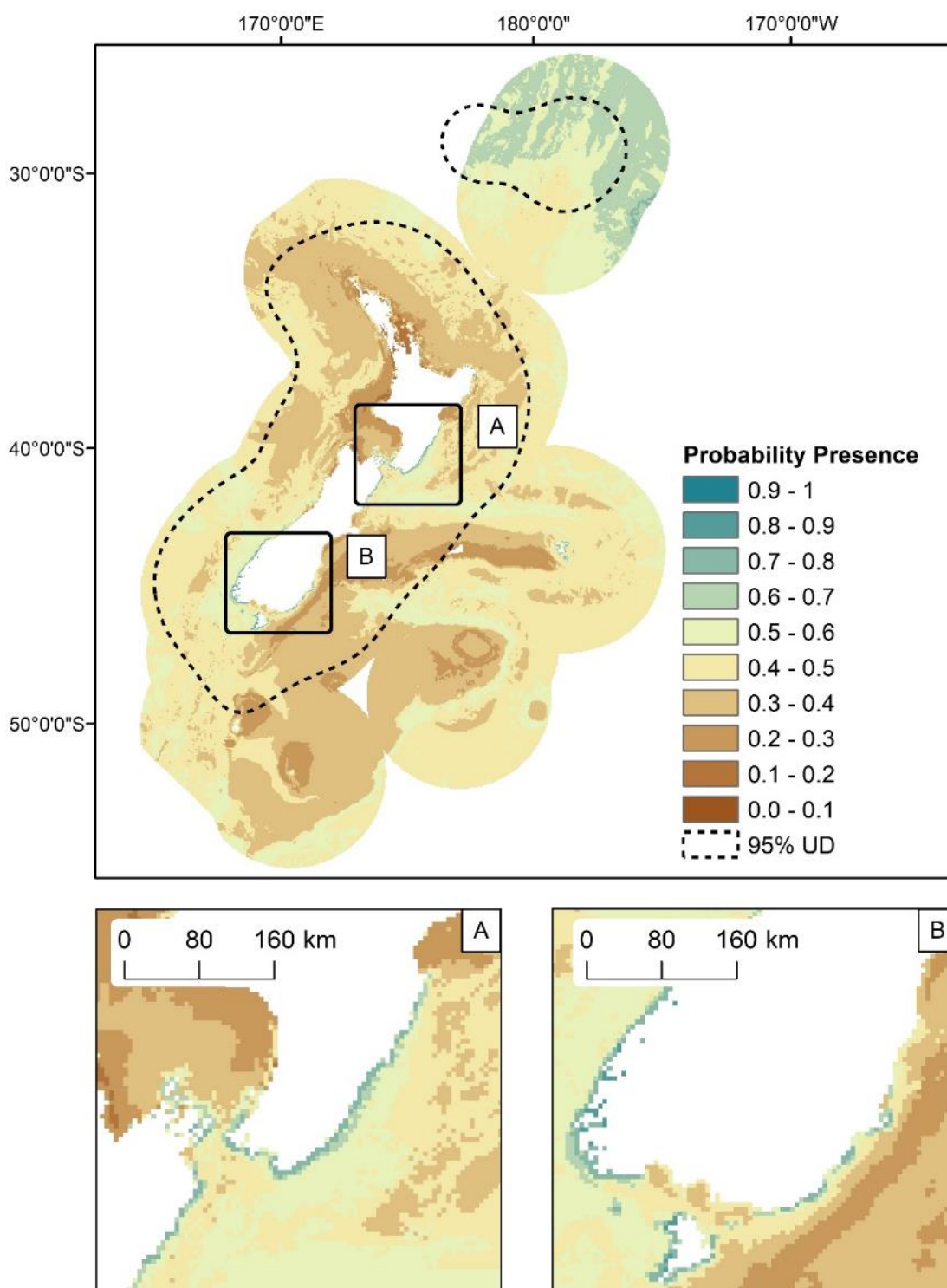


Figure 115: The predicted probability of humpback whale (*Megaptera novaeangliae*) presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) South North Island and north South Island; B) South South Island, including Stewart Island.

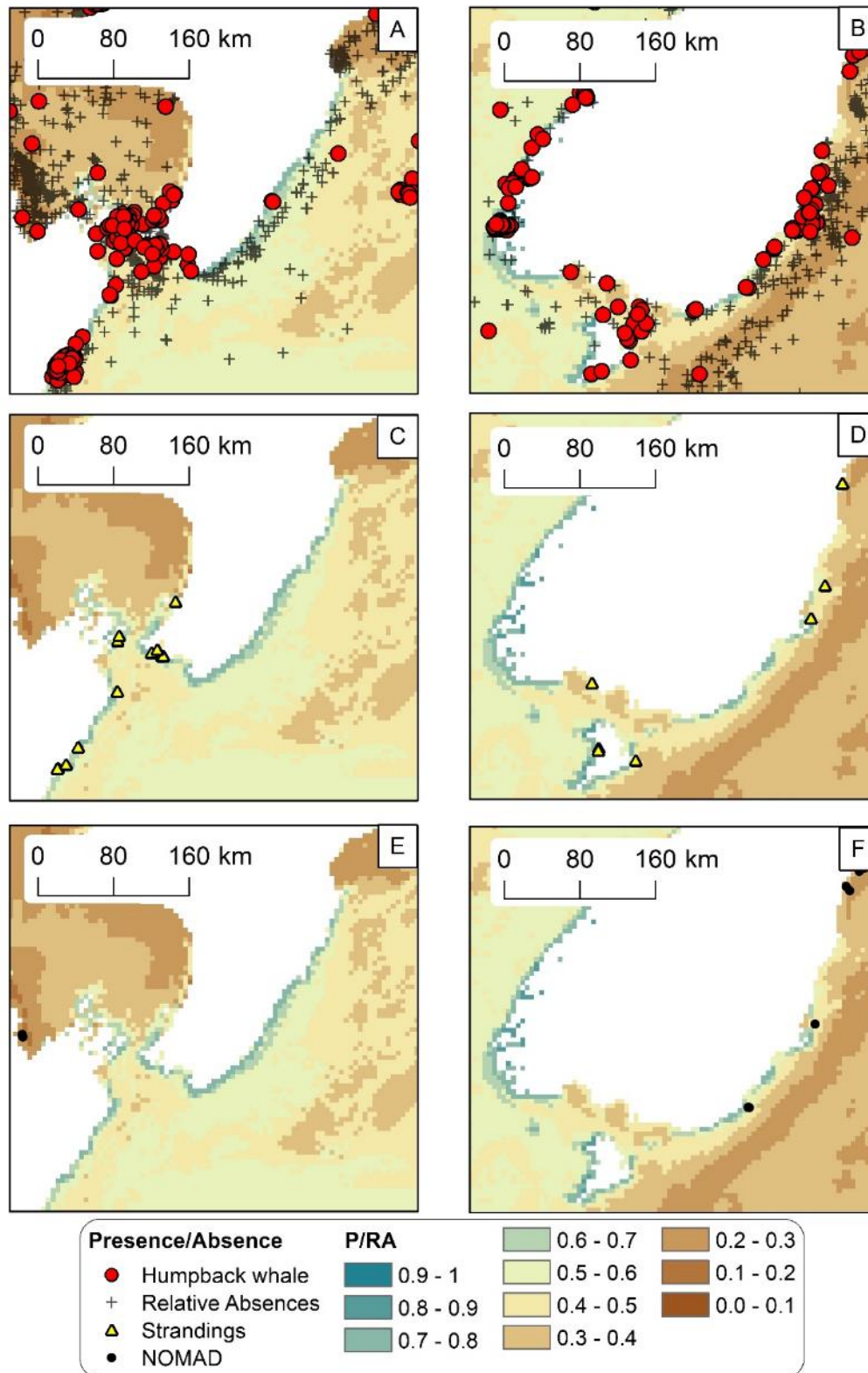


Figure 116: The predicted probability of Humpback whale (*Megaptera novaeangliae*) presence modelled using bootstrapped BRTs. Predicted probability presence of Humpback whale in the south North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted probability presence of Humpback whale on south South Island, including Stewart Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

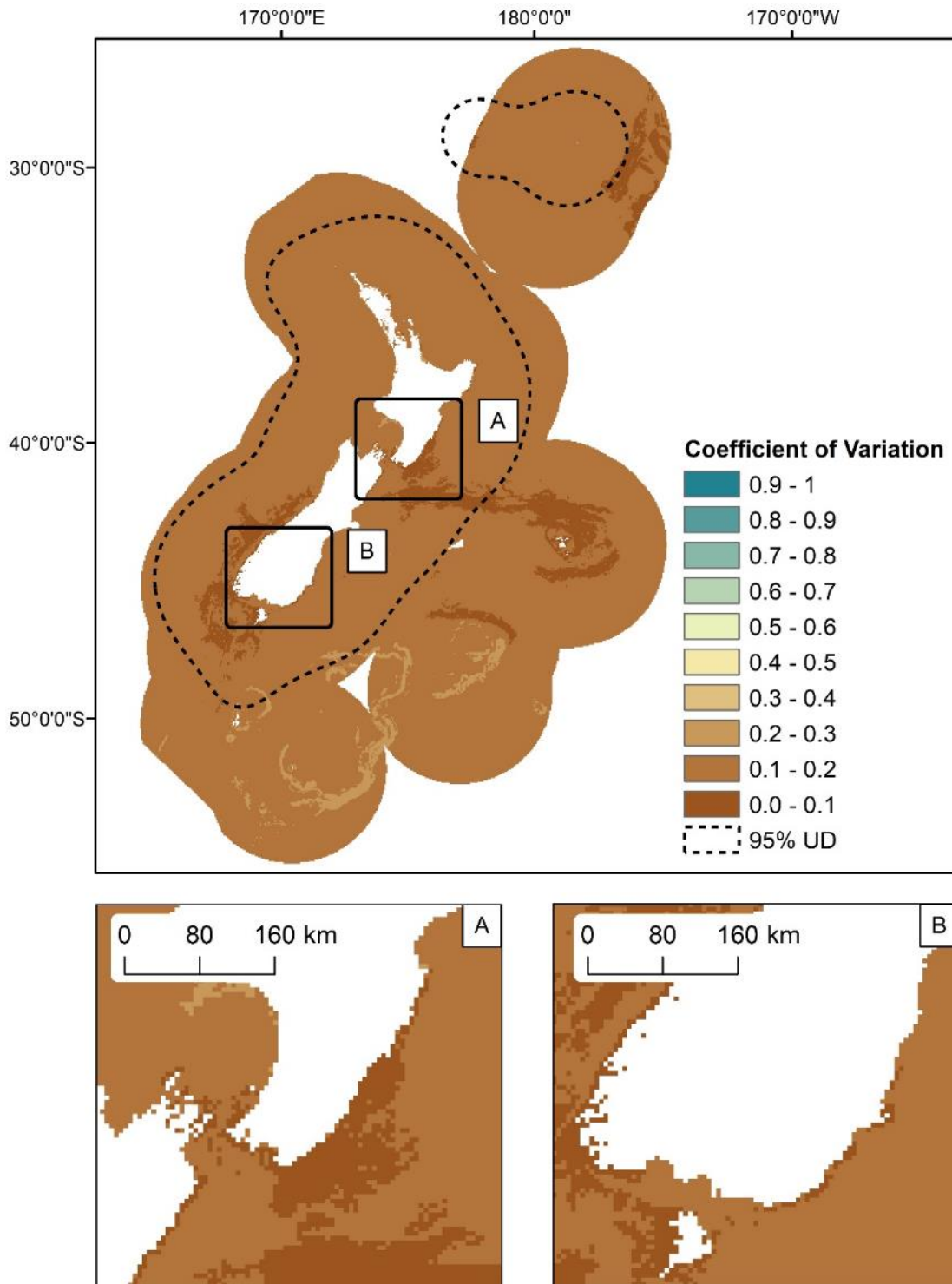


Figure 117: Uncertainty estimates (coefficient of variation, CV) of humpback whale (*Megaptera novaeangliae*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) South North Island and north South Island; B) South South Island, including Stewart Island.

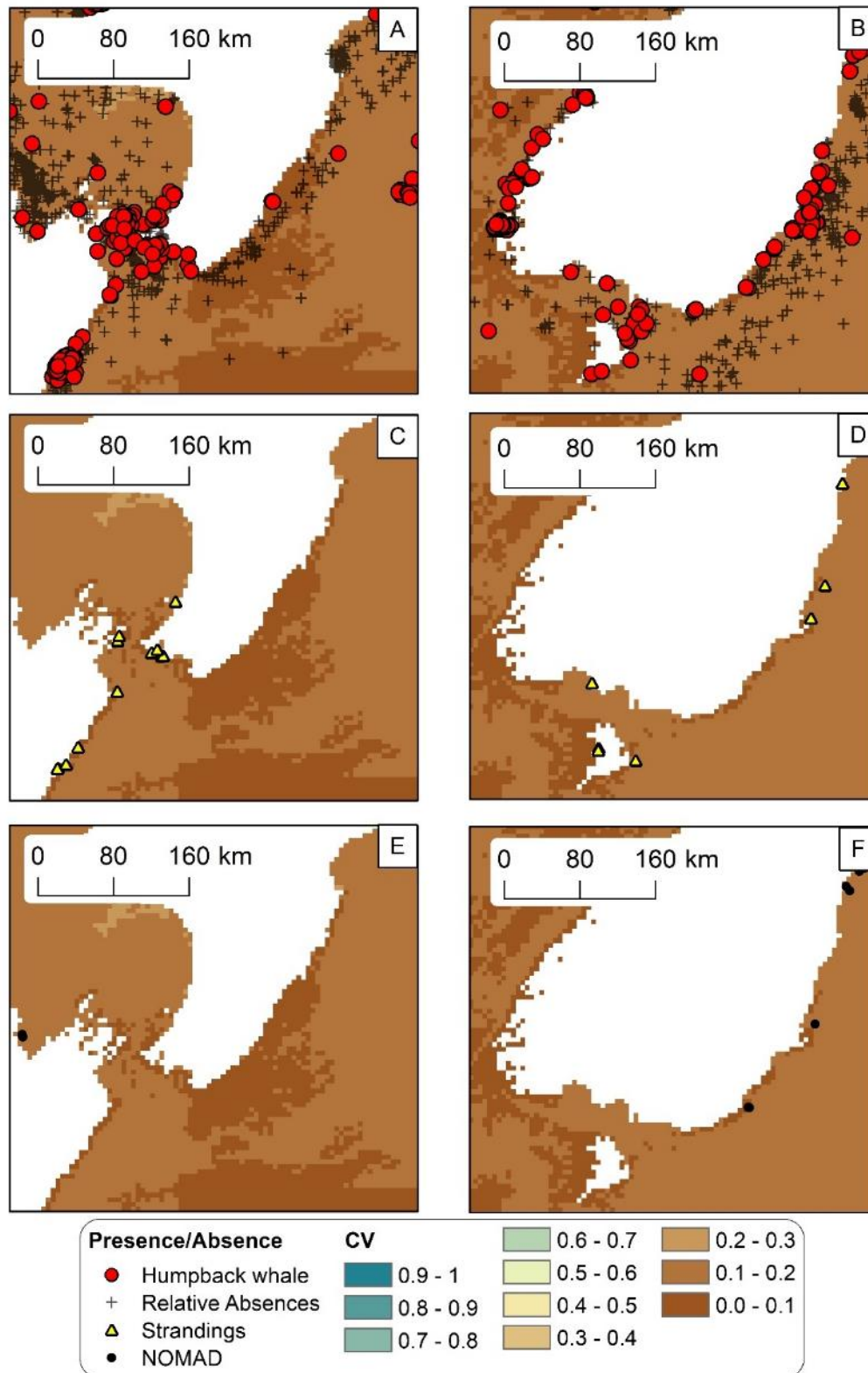


Figure 118: Uncertainty estimates (coefficient of variation, CV) of humpback whale (*Megaptera novaeangliae*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Humpback whale probability presence models in the south North Island and north South Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of Humpback whale probability presence models on the south South Island, including Stewart Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

Table 42: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) from Winter (May - Oct) and Summer (Nov - Apr) sightings of humpback whale (*Megaptera novaeangliae*) records.

Season	Metric	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Winter	Mean	0.30	0.29	0.85	0.85
	Standard Deviation	0.06	0.13	0.01	0.03
Summer	Mean	0.27	0.27	0.84	0.84
	Standard Deviation	0.11	0.19	0.02	0.03

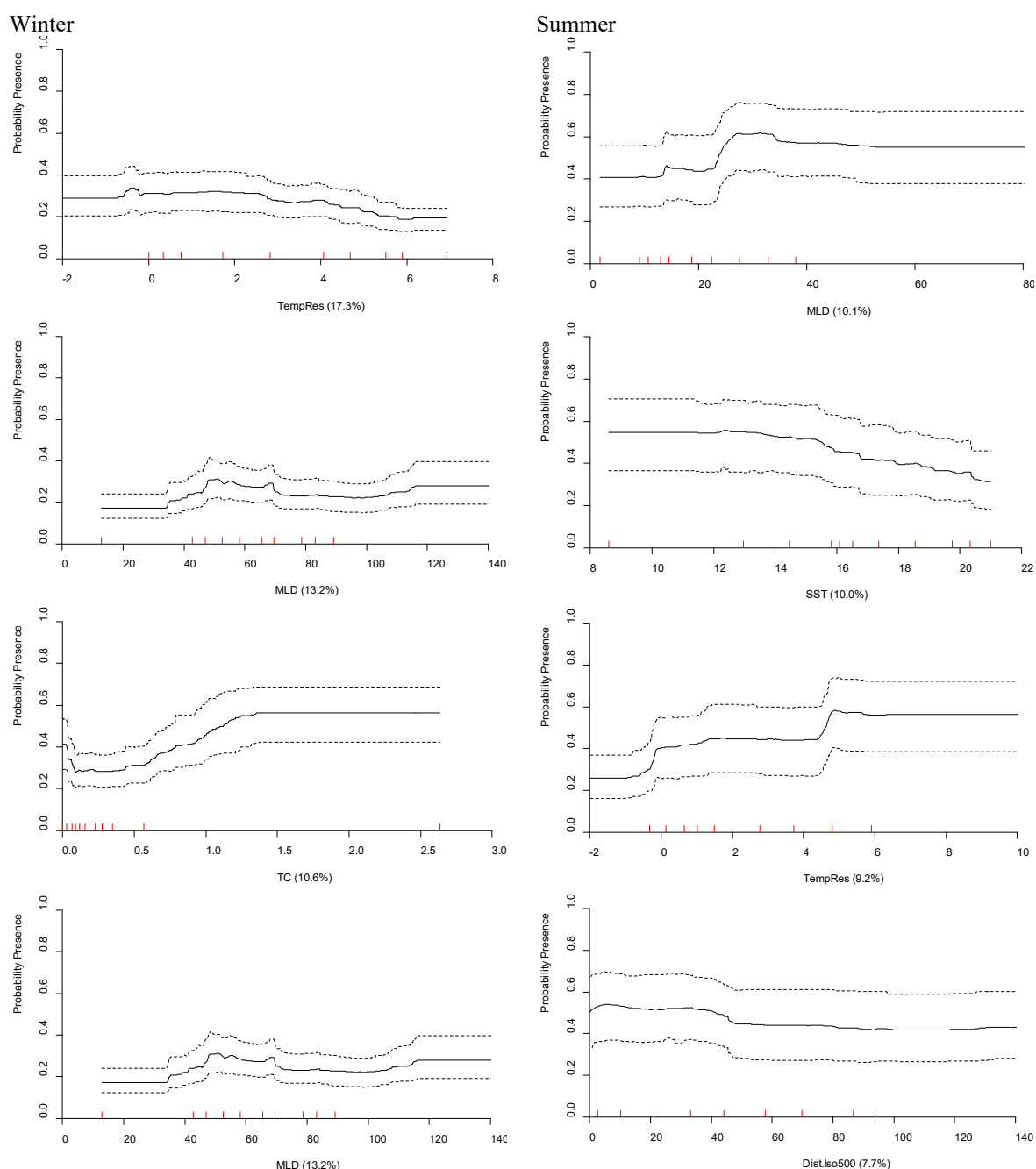


Figure 119: Partial dependence plots showing the relationships between predictor variables and predicted presence of humpback whale (*Megaptera novaeangliae*) modelled using bootstrapped BRTs for Winter (left) and Summer (right). The four most influential environmental predictors in the model are shown for Winter (left) and Summer (right). Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x axes).

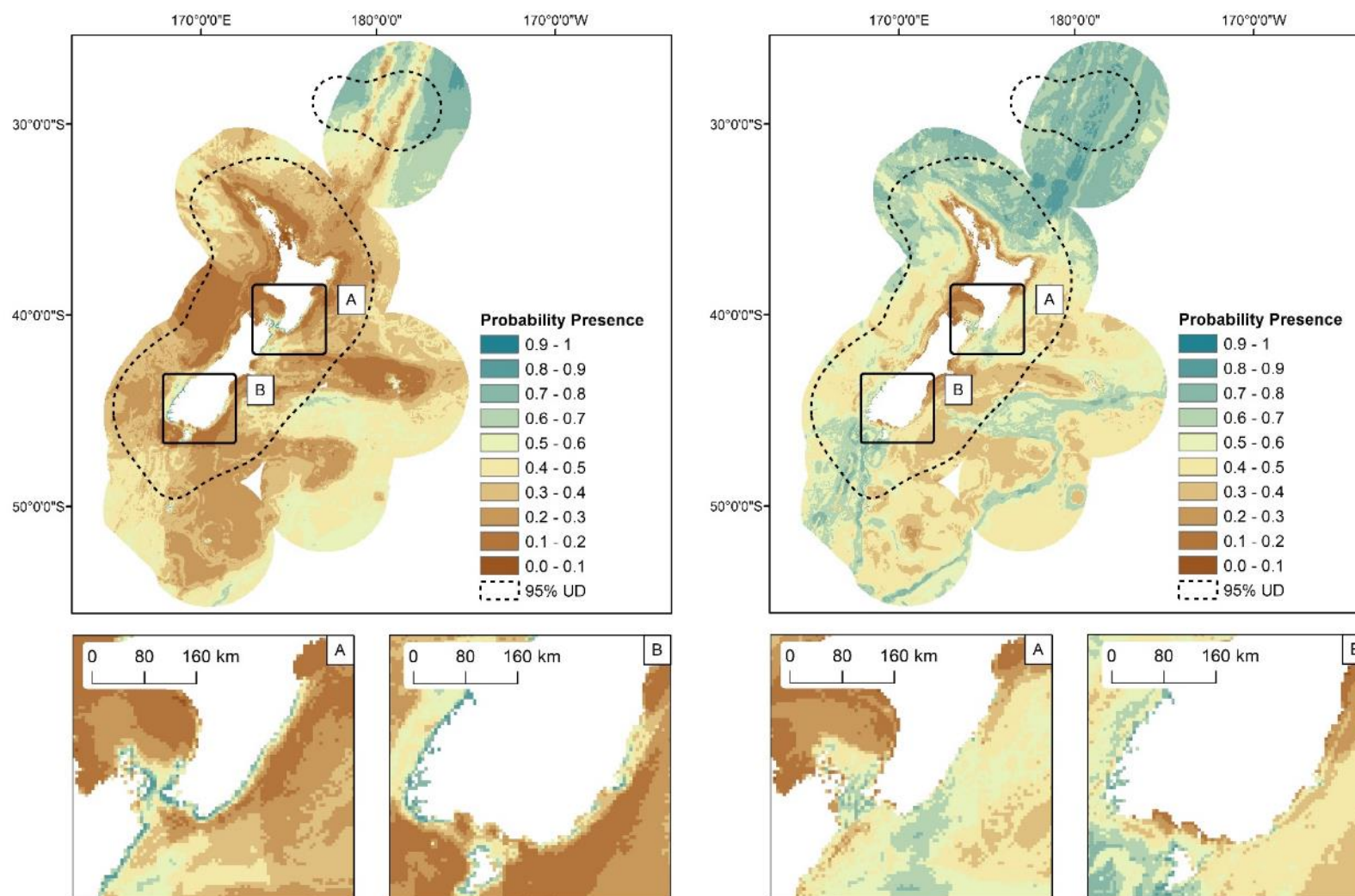


Figure 123: Seasonal predicted probability presence of humpback whale (*Megaptera novaeangliae*) modelled using bootstrapped BRTs fitted with Winter (May - Oct, n = 169) (left) and Summer (Nov - Apr, n = 654) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) South North Island and north South Island; B) South South Island, including Stewart Island.

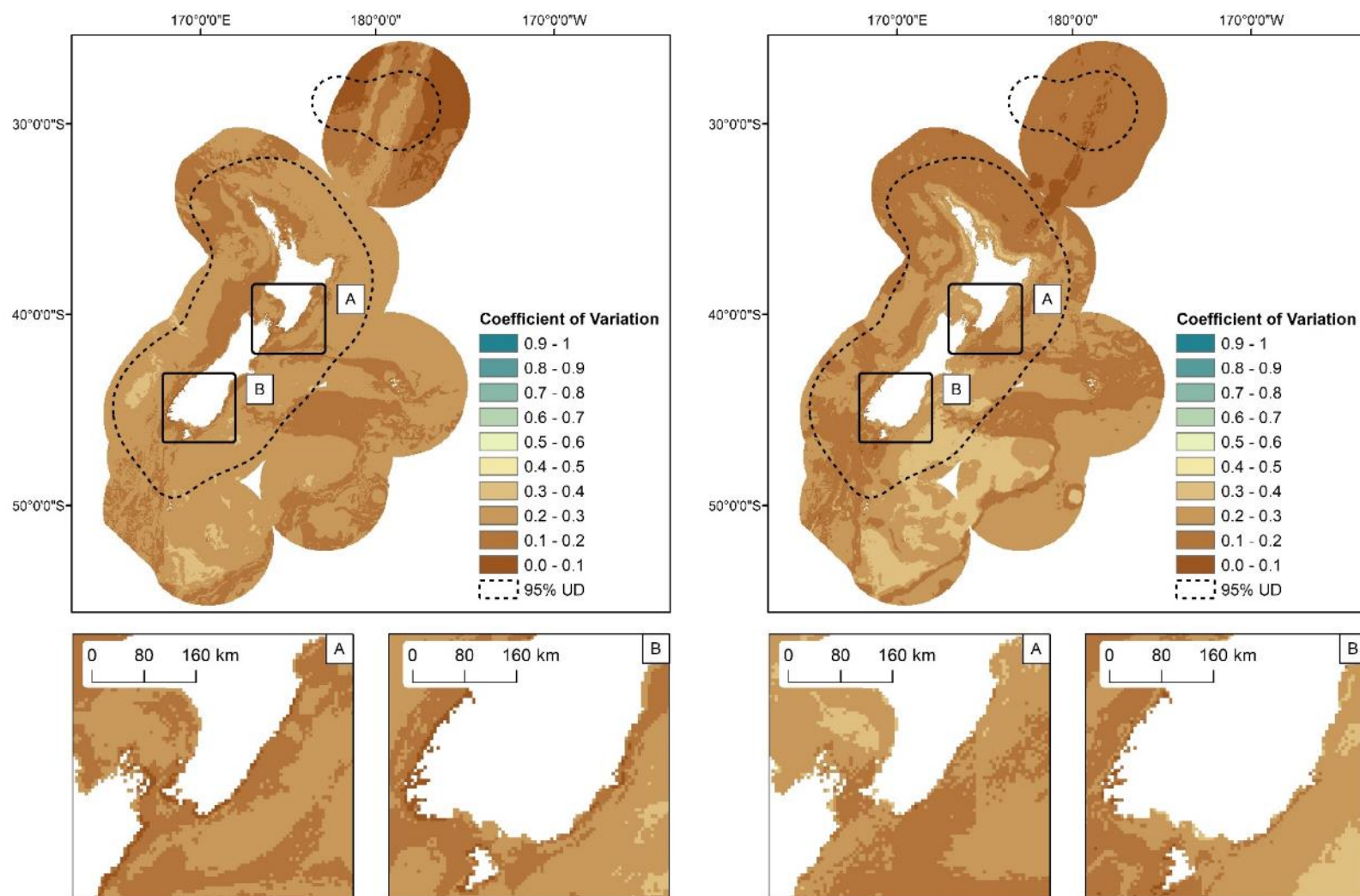


Figure 124: Seasonal uncertainty estimates (coefficient of variation, CV) of humpback whale (*Megaptera novaeangliae*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs fitted with Winter (May - Oct) (left) and Summer (Nov - Apr) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) South North Island and north South Island; B) South South Island, including Stewart Island.

Sperm whale (*Physeter microcephalus*)

Table 43: Sperm whale (*Physeter microcephalus*): Highlights of model fits and geographic prediction.

Sample number	Distribution	P/RA		Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
High	Cosmopolitan – offshore	Excellent	Very good	Fair	Fair	Little change in pattern between seasons, lower probability occurrence in winter

Table 44: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of sperm whale (*Physeter microcephalus*).

	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Mean	0.46	0.46	0.92	0.92
Standard Deviation	0.02	0.04	0.01	0.01

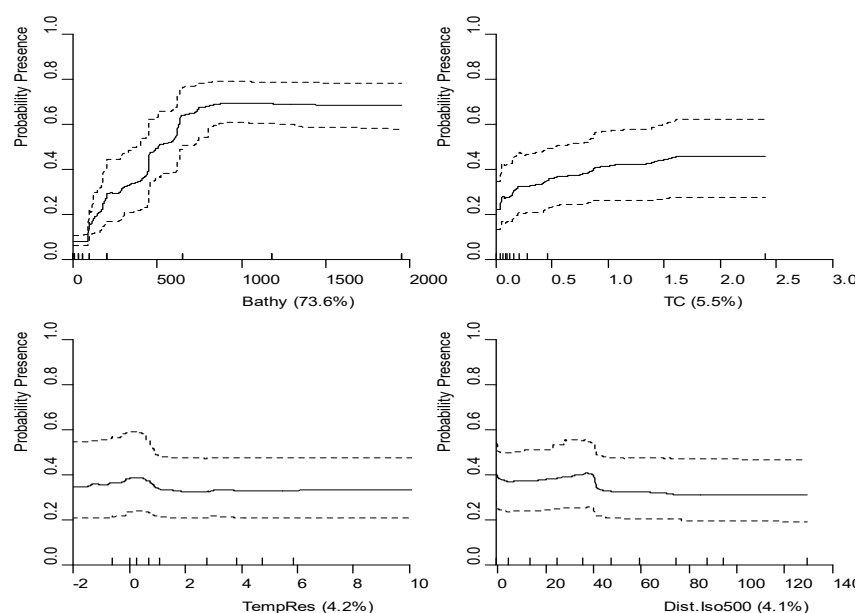


Figure 122: Partial dependence plots showing the relationships between predictor variables and probability presence of sperm whale (*Physeter microcephalus*) modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

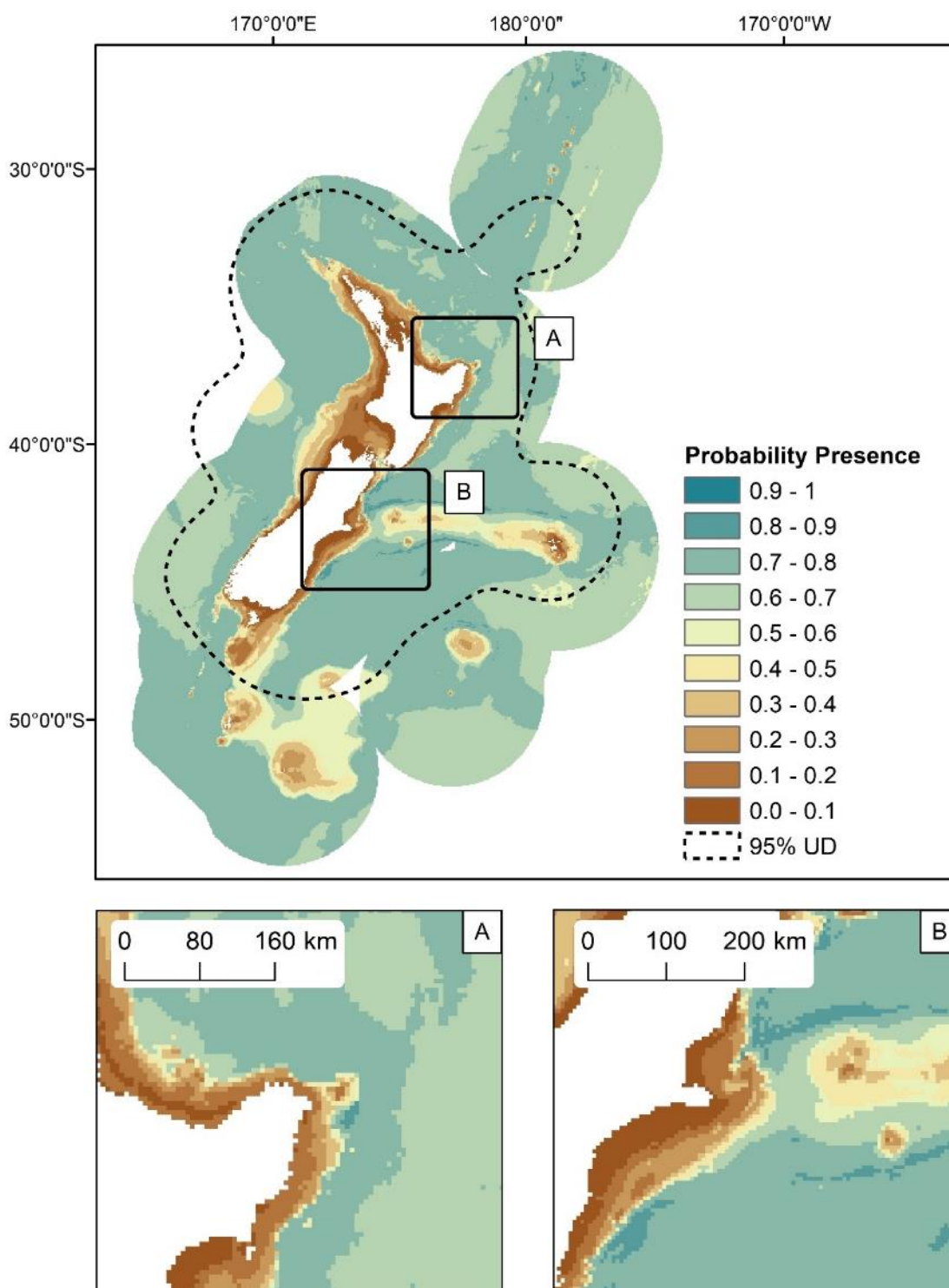


Figure 123: The predicted probability of sperm whale (*Physeter microcephalus*) presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) East Cape, North Island; B) East coast, South Island.

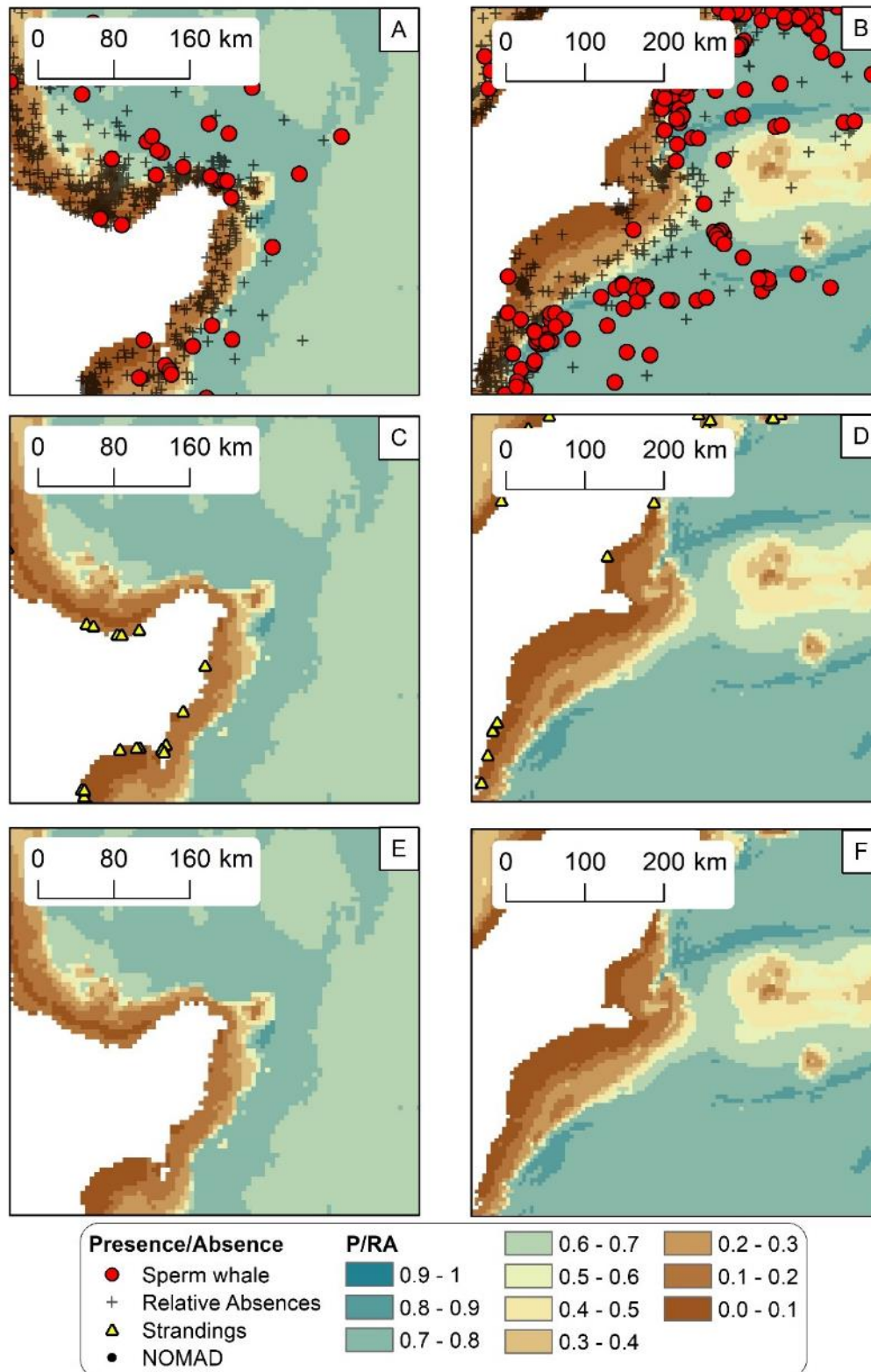


Figure 124: The predicted probability of sperm whale (*Physeter microcephalus*) presence modelled using bootstrapped BRTs. Predicted probability presence of Sperm whale on East Cape, North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted probability presence of Sperm whale on the east coast of the South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

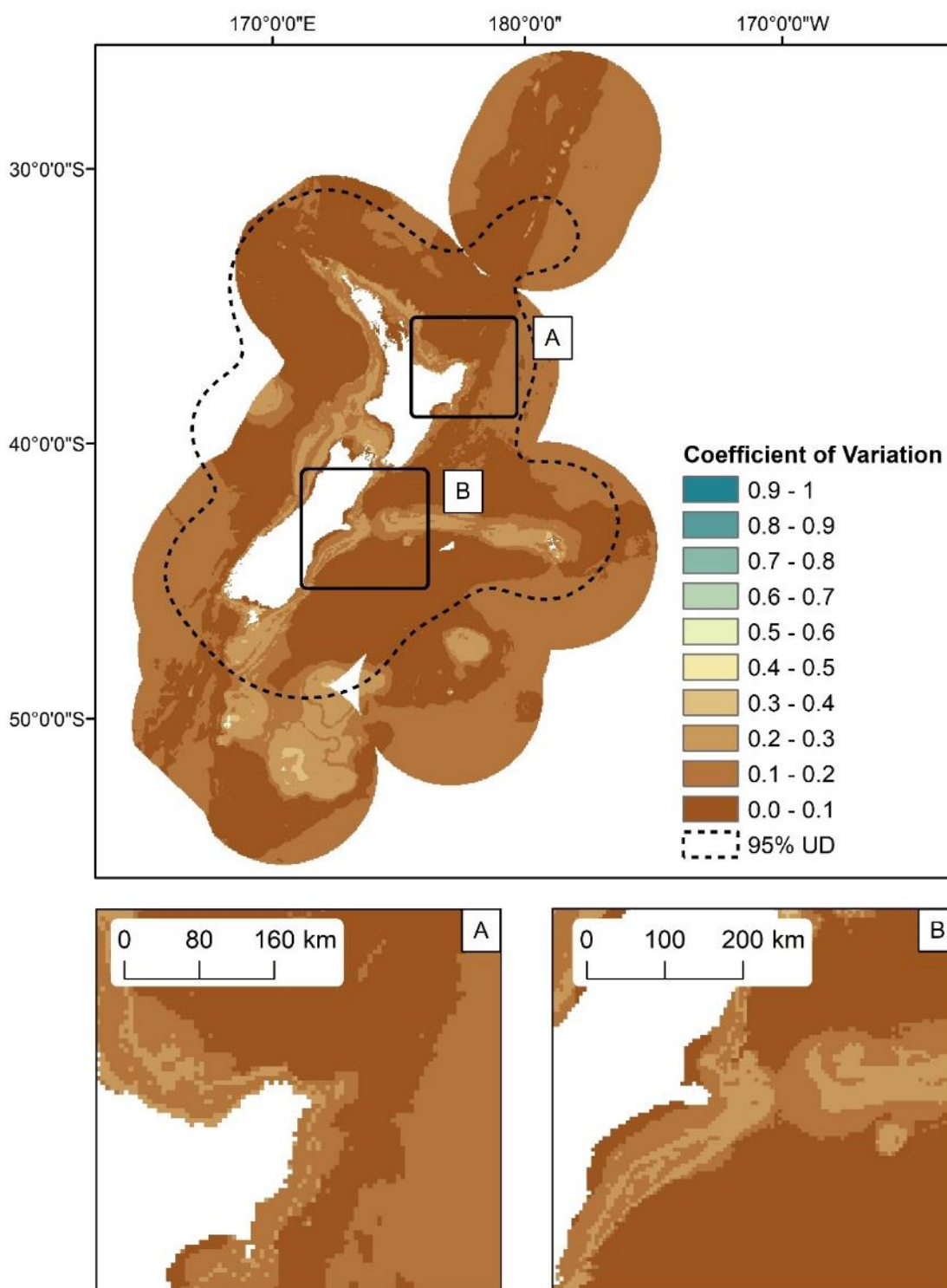


Figure 125: Uncertainty estimates (coefficient of variation, CV) of sperm whale (*Physeter microcephalus*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) East Cape, North Island; B) East coast, South Island.

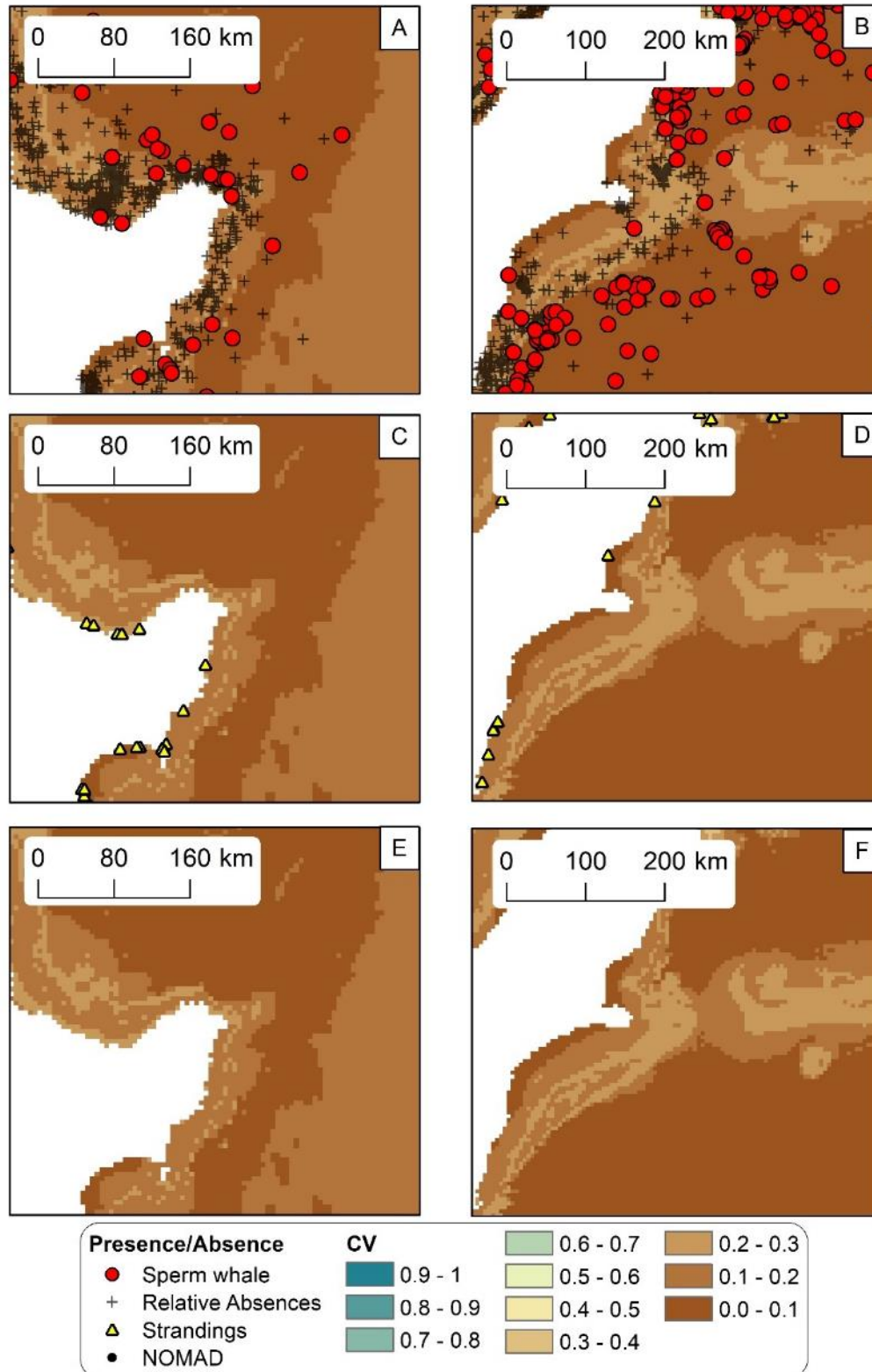


Figure 126: Uncertainty estimates (coefficient of variation, CV) of sperm whale (*Physeter microcephalus*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Sperm whale probability presence models in on East Cape, North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of Sperm whale probability presence models on the east coast of the South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

Table 45: Mean model performance measures (deviance explained and Pearson's correlation of predicted vs observed group sizes (R^2)) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of sperm whale (*Physeter microcephalus*).

	Deviance explained (training data)	Deviance explained (validation data)	Pearson's correlation of predicted vs observed species group sizes (R^2)
Mean	0.35	0.10	0.26
Standard Deviation	0.28	0.30	0.15

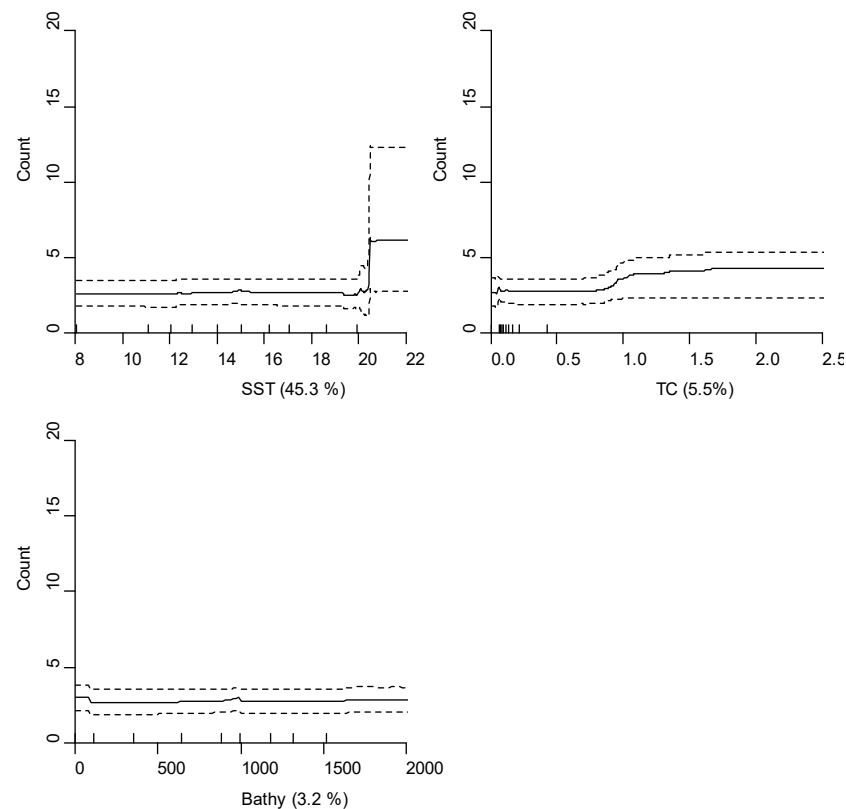


Figure 127: Partial dependence plots showing the relationships between predictor variables and sperm whale (*Physeter microcephalus*) group sizes modelled using bootstrapped BRTs. The three most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x axes).

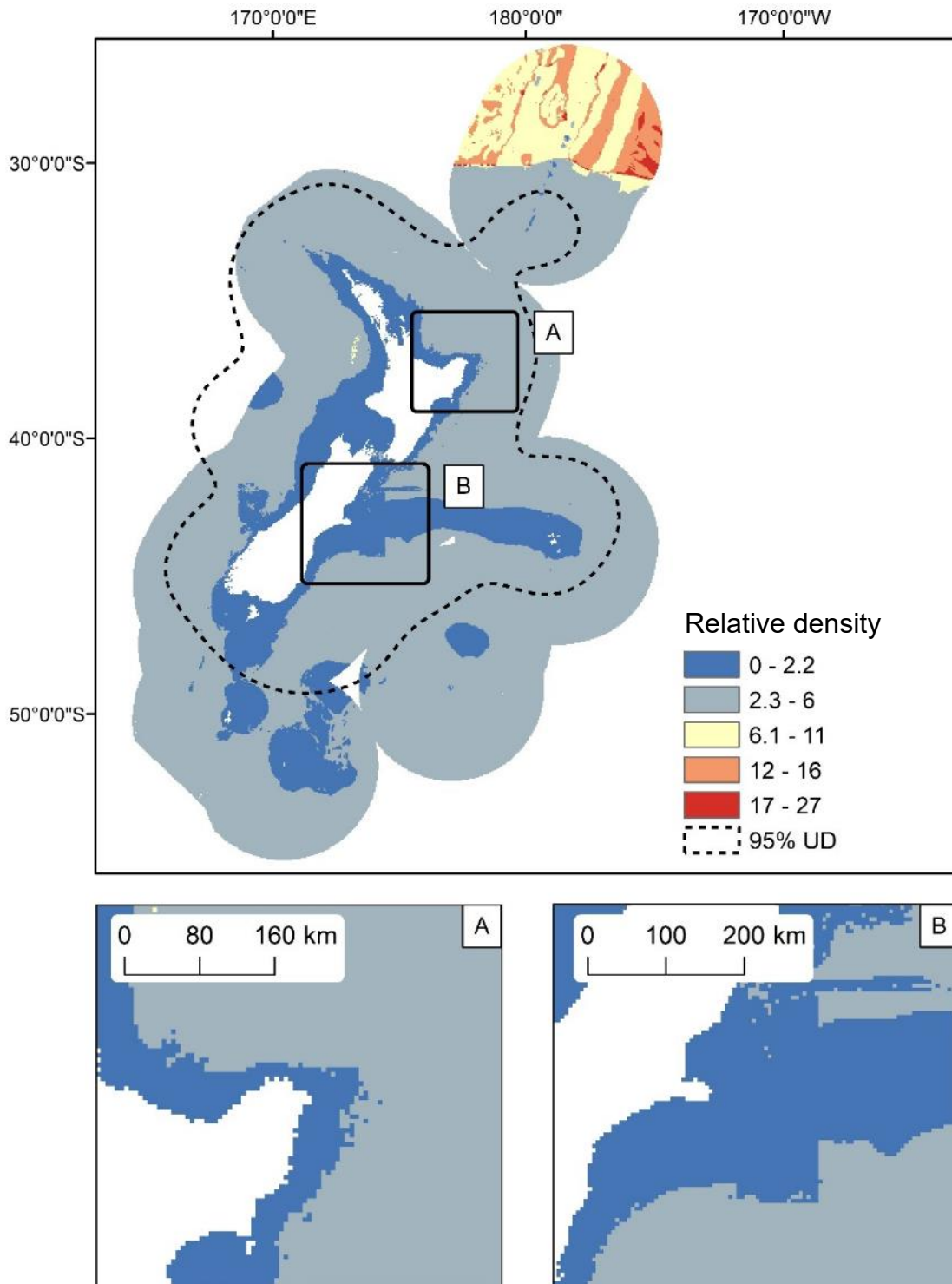


Figure 128: Predicted relative density of sperm whale (*Physeter microcephalus*) in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) East Cape, North Island; B) East coast, South Island.

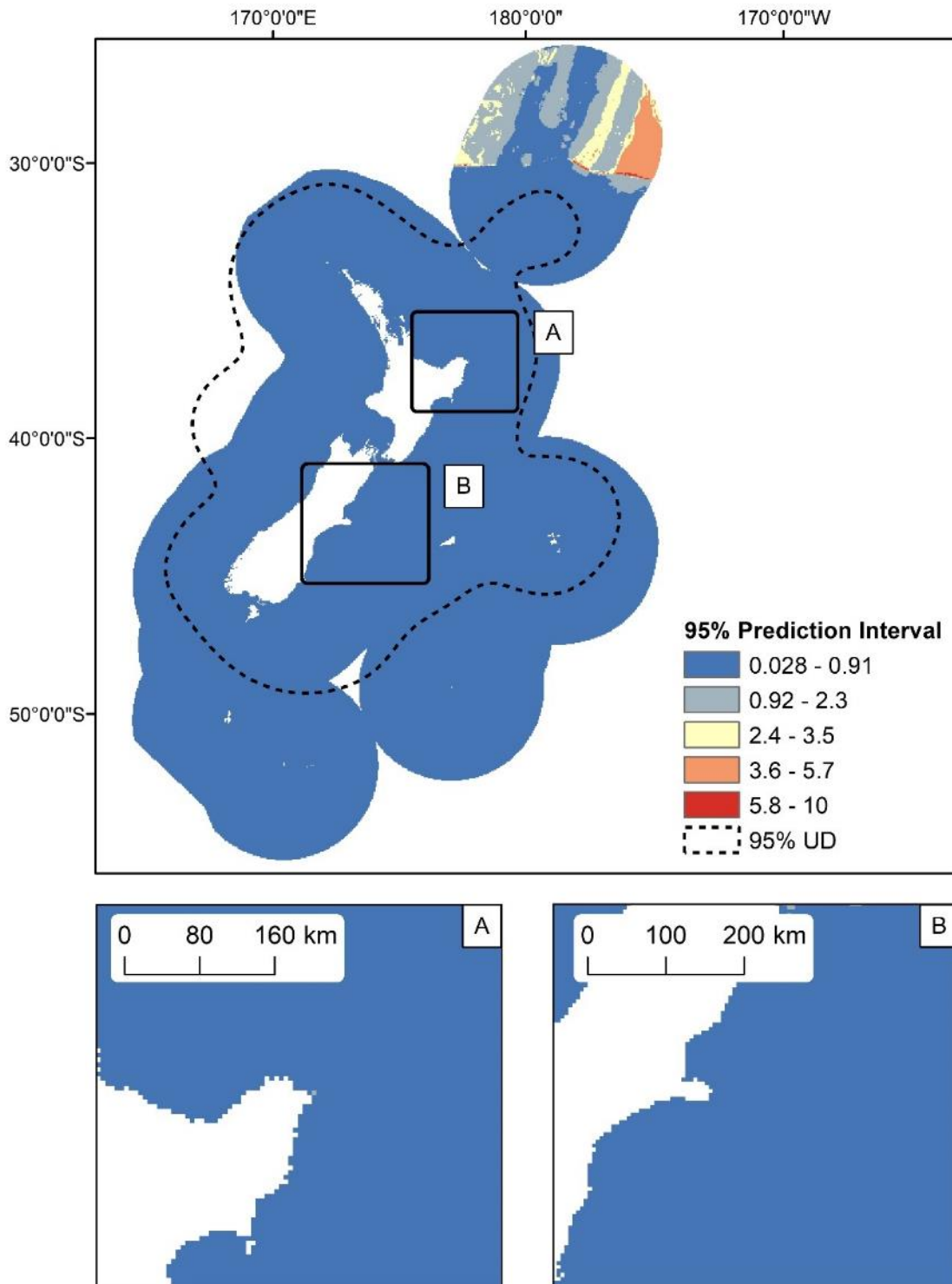


Figure 129: Uncertainty estimates (95% prediction interval) of sperm whale (*Physeter microcephalus*) predicted relative density in the New Zealand EEZ modelled using bootstrapped BRTs. 95% utilisation distribution is shown as dashed line. Inset maps: A) East Cape, North Island; B) East coast, South Island.

Table 46: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) from Winter (May - Oct) and Summer (Nov - Apr) sightings of sperm whale (*Physeter microcephalus*) records.

Season	Metric	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Winter	Mean	0.47	0.48	0.93	0.93
	Standard Deviation	0.08	0.16	0.01	0.03
Summer	Mean	0.48	0.48	0.92	0.92
	Standard Deviation	0.06	0.10	0.01	0.02

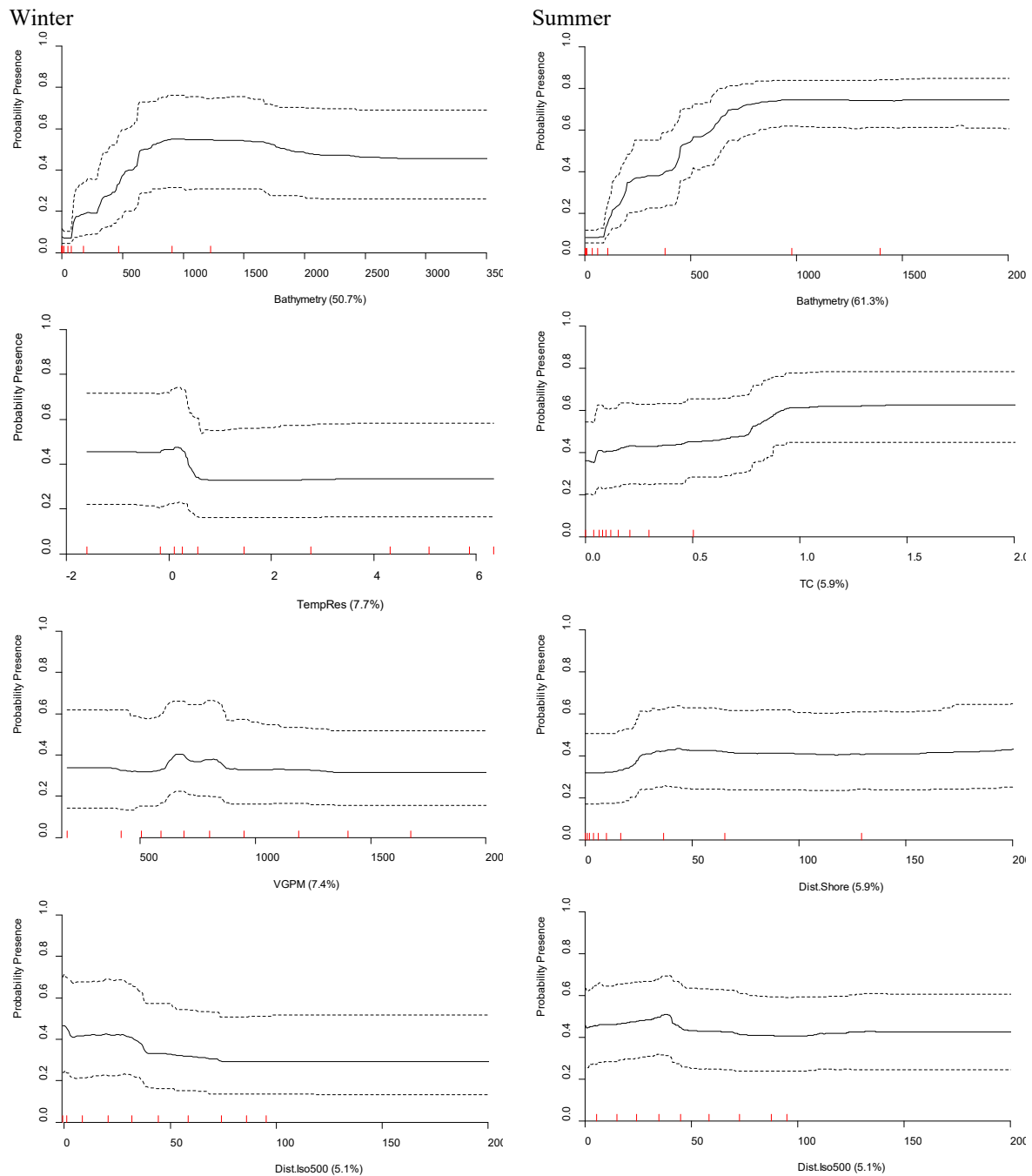


Figure 130: Partial dependence plots showing the relationships between predictor variables and predicted of sperm whale (*Physeter microcephalus*) presence modelled using bootstrapped BRTs for Winter (left) and Summer (right). The four most influential environmental predictors in the model are shown for Winter (left) and Summer (right). Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x axes).

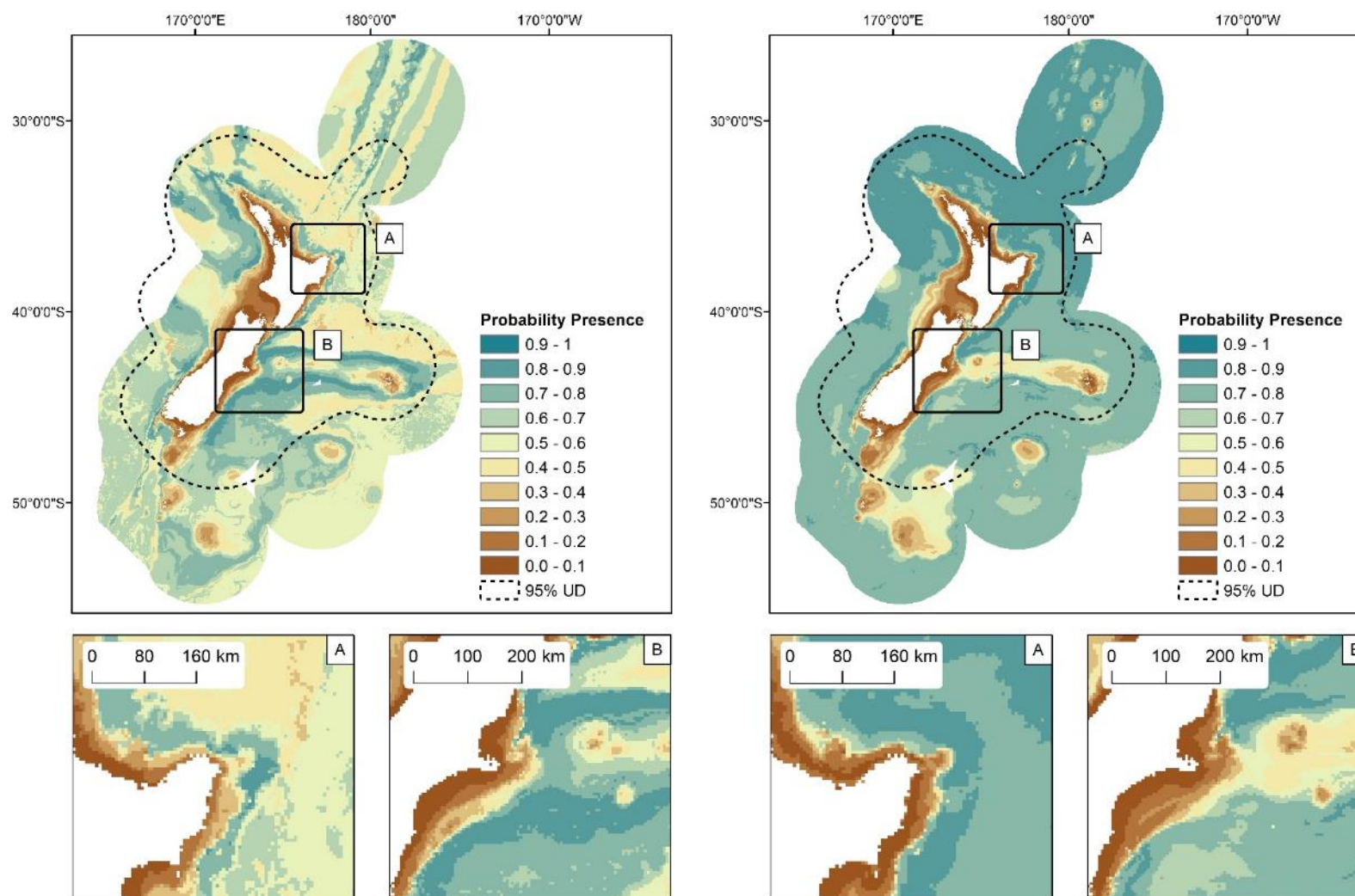


Figure 134: Seasonal predicted probability of sperm whale (*Physeter microcephalus*) presence modelled using bootstrapped BRTs fitted with Winter (May - Oct, n = 169) (left) and Summer (Nov – Apr, n = 654) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) East Cape, North Island; B) East coast, South Island.

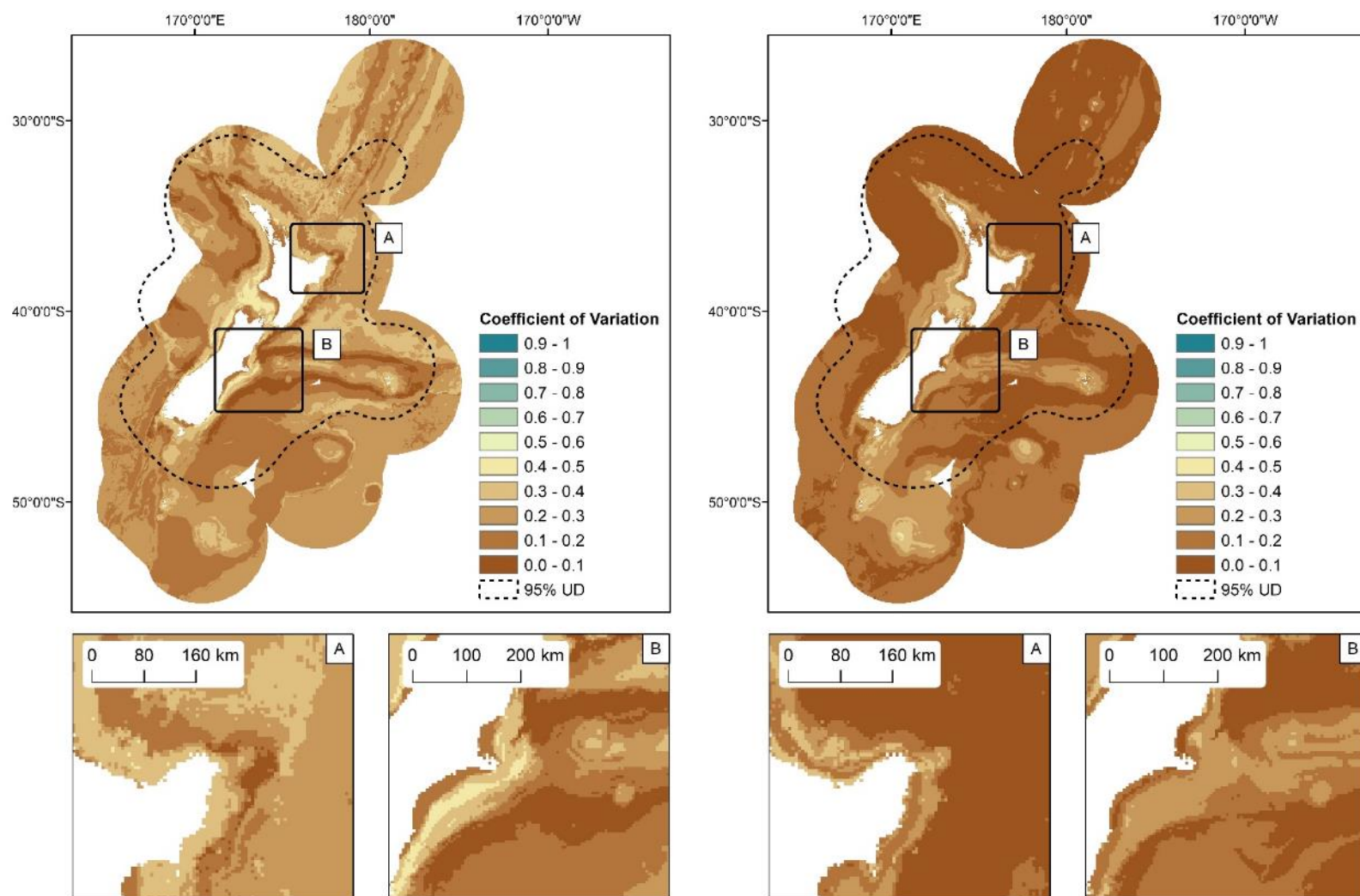


Figure 135: Seasonal uncertainty estimates (coefficient of variation, CV) of sperm whale (*Physeter microcephalus*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs fitted with Winter (May - Oct) (left) and Summer (Nov - Apr) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) East Cape, North Island; B) East coast, South Island.

Pilot whales (*Globicephala melas* and *Globicephala macrorhynchus*)

Table 47: Pilot whales (*Globicephala melas* and *Globicephala macrorhynchus*): Highlights of model fits and geographic prediction. Group size estimates models did not converge.

Sample number	Distribution	P/RA		Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R^2)	Model fit (dev. Exp)	
High	Cosmopolitan – offshore	Excellent	Very good	NA	NA	Southerly move in probability presence further south in summer

Table 48: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of pilot whales (*Globicephala melas* and *Globicephala macrorhynchus*).

	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Mean	0.43	0.43	0.91	0.91
Standard Deviation	0.02	0.03	0.01	0.01

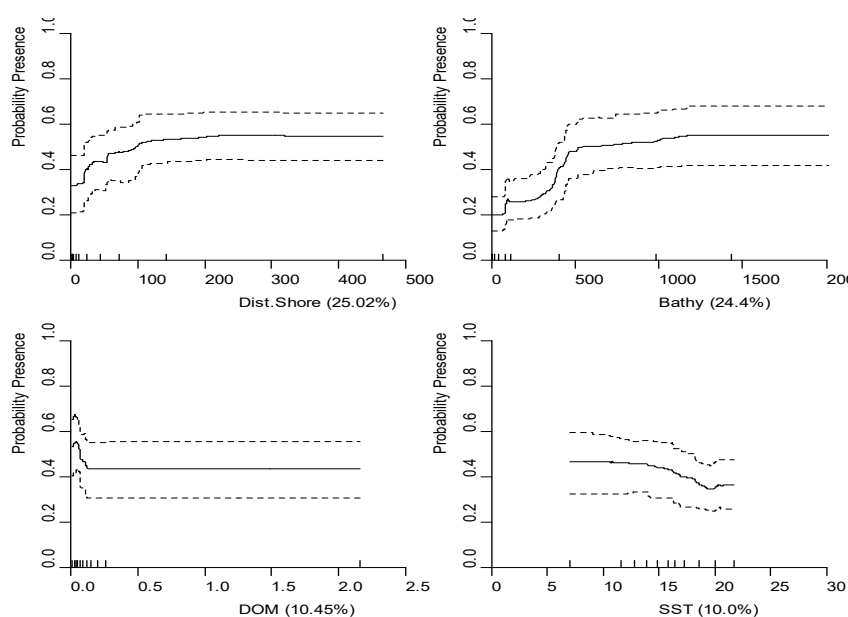


Figure 133: Partial dependence plots showing the relationships between predictor variables and probability presence of pilot whales (*Globicephala melas* and *Globicephala macrorhynchus*) modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

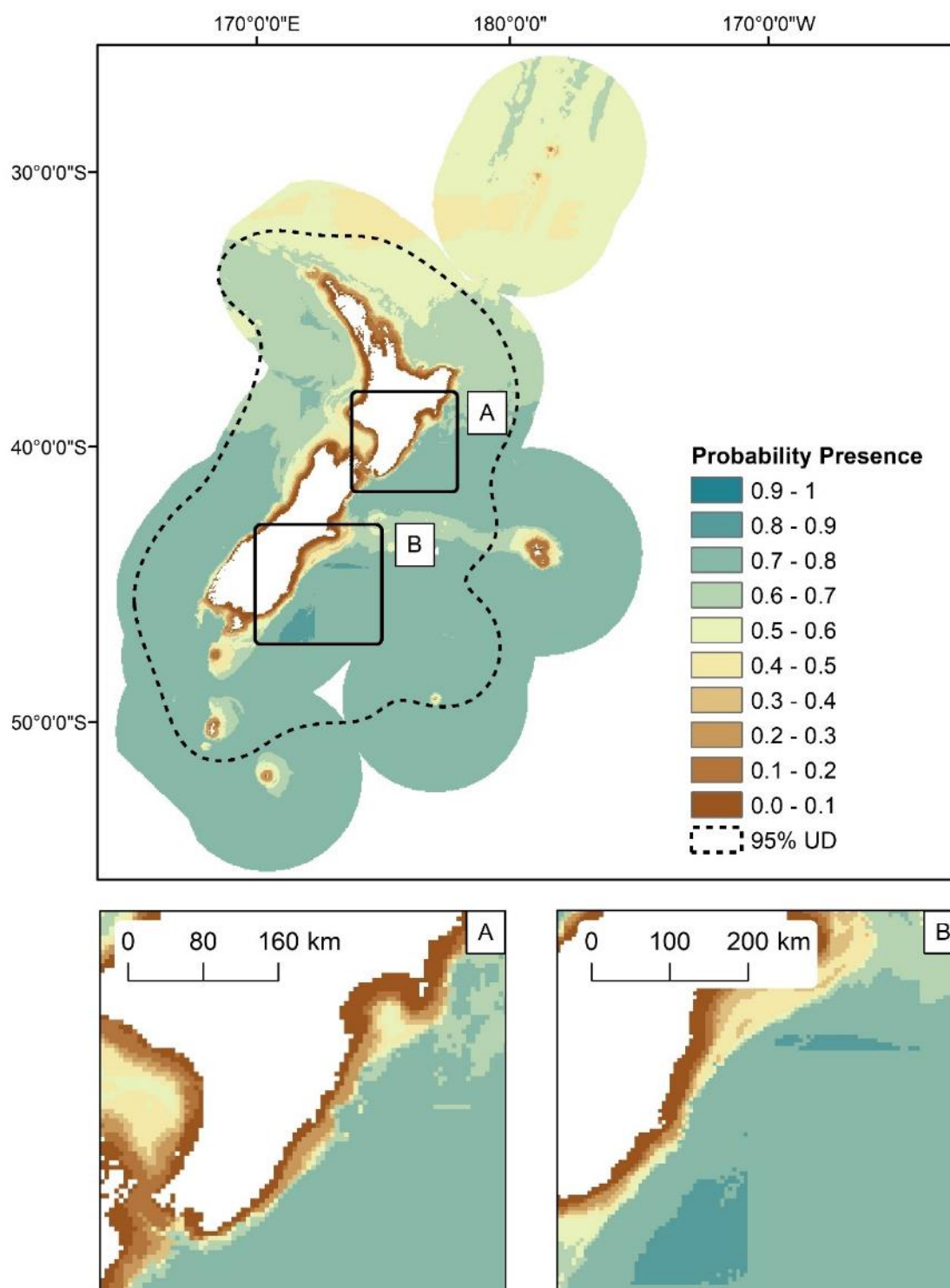


Figure 134: The predicted probability of pilot whales (*Globicephala melas* and *Globicephala macrorhynchus*) presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) South east coast of the North Island; B) South east coast of the South Island.

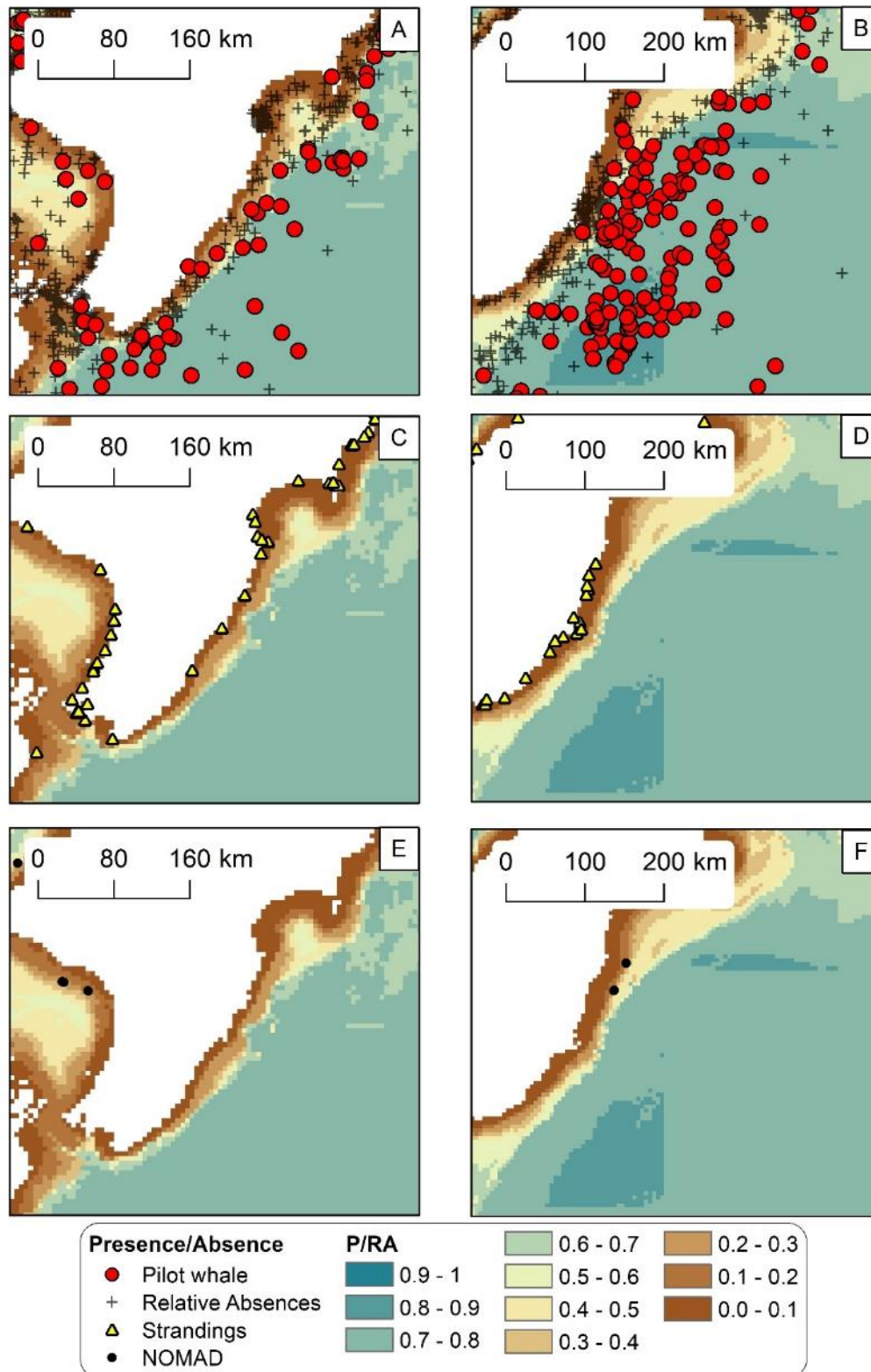


Figure 135: The predicted probability presence of pilot whales (*Globicephala melas* and *Globicephala macrorhynchus*) modelled using bootstrapped BRTs. Predicted probability presence of Pilot whales on the south east of the North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted probability presence of Pilot whales on the south east coast of the South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

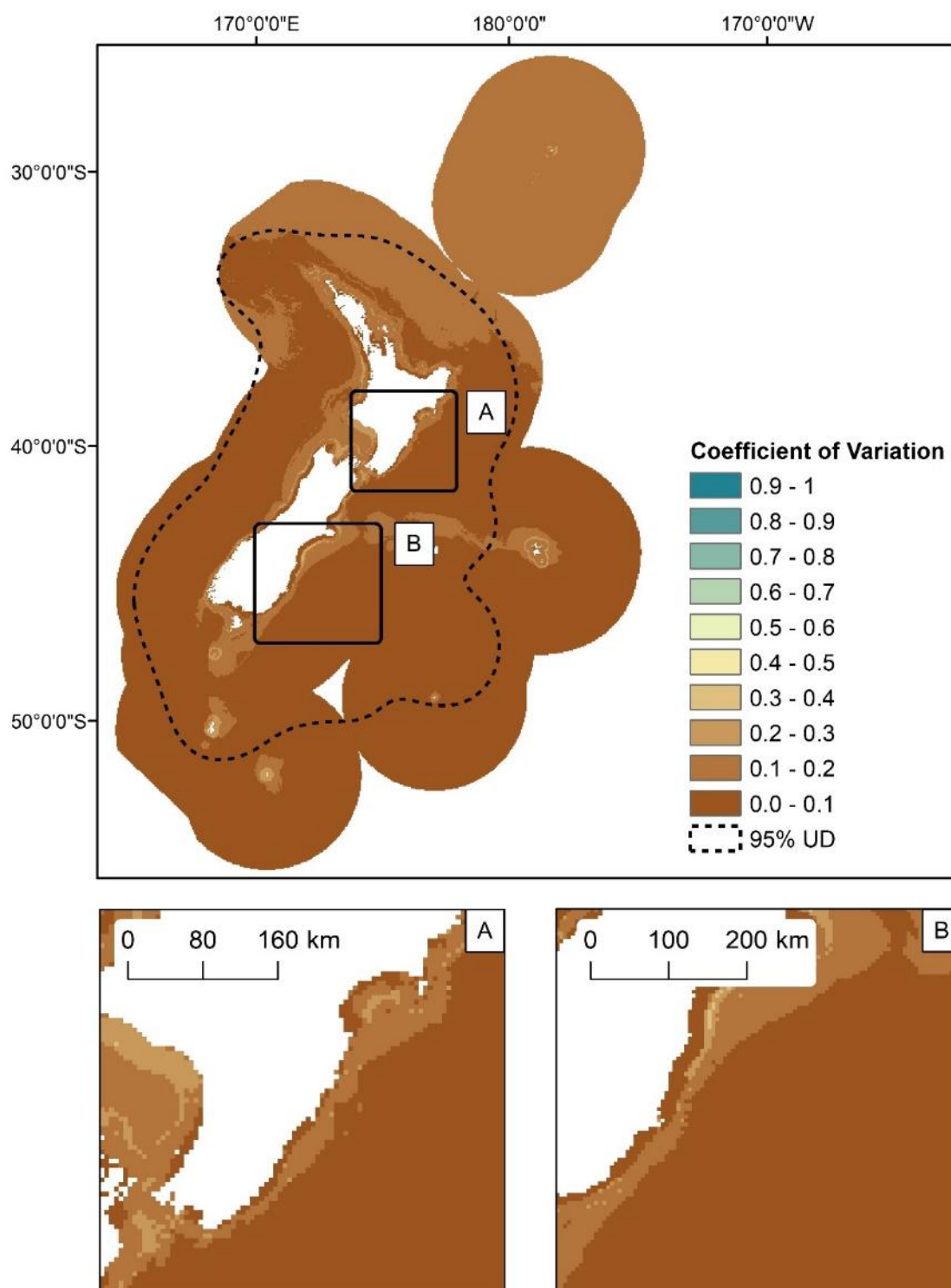


Figure 136: Uncertainty estimates (coefficient of variation, CV) of pilot whales (*Globicephala melas* and *Globicephala macrorhynchus*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) South east coast of the North Island; B) South east coast of the South Island.

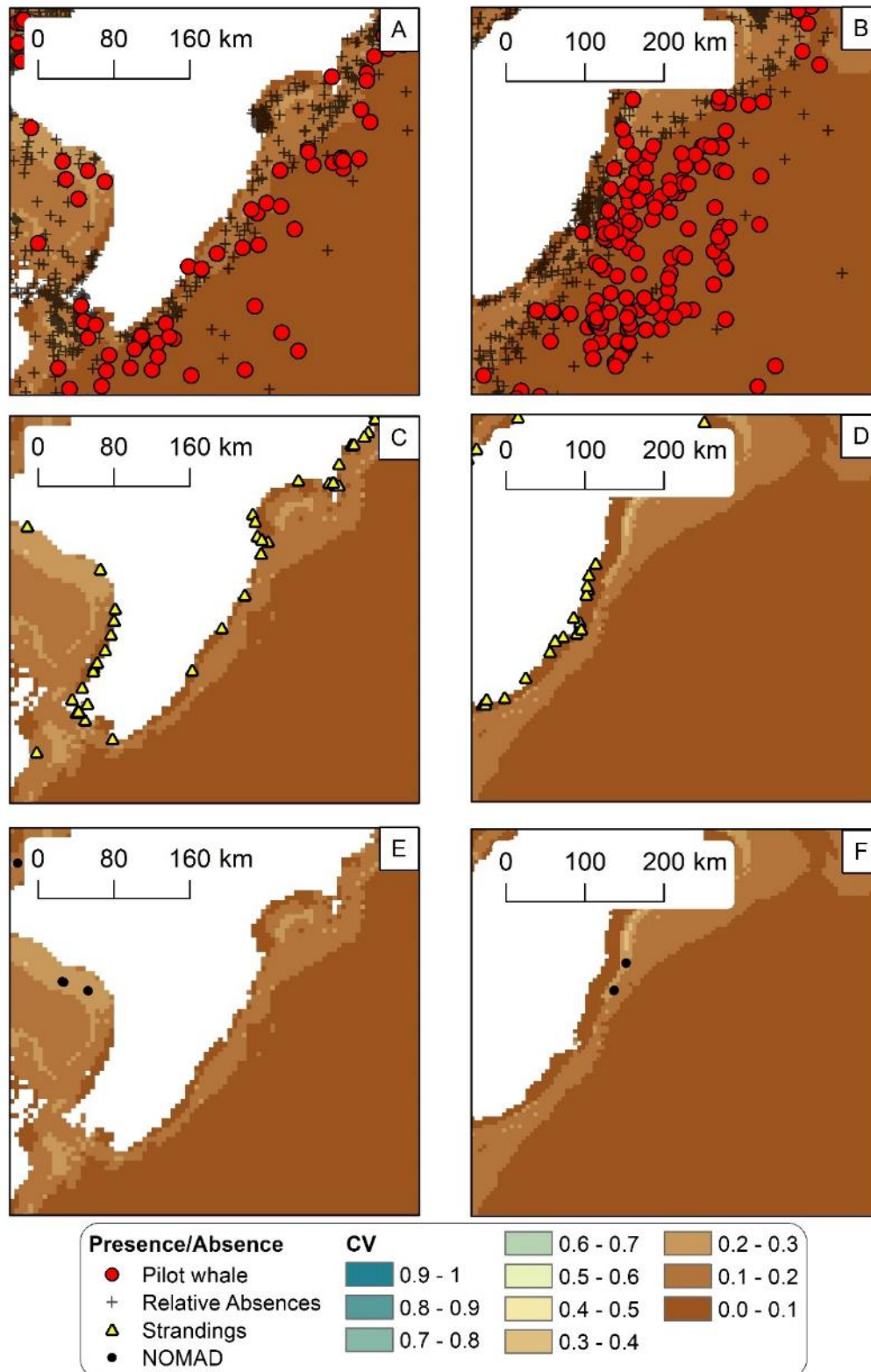


Figure 137: Uncertainty estimates (coefficient of variation, CV) of pilot whales (*Globicephala melas* and *Globicephala macrorhynchus*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Pilot whales probability presence models in the south east of the North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of pilot whales probability presence models on the south east coast of the South Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

Table 49: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) from Winter (May - Oct) and Summer (Nov - Apr) sightings of pilot whales (*Globicephala melas* and *Globicephala macrorhynchus*) records.

Season	Metric	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Winter	Mean	0.47	0.49	0.91	0.92
	Standard Deviation	0.09	0.14	0.02	0.03
Summer	Mean	0.44	0.45	0.91	0.92
	Standard Deviation	0.05	0.09	0.01	0.02

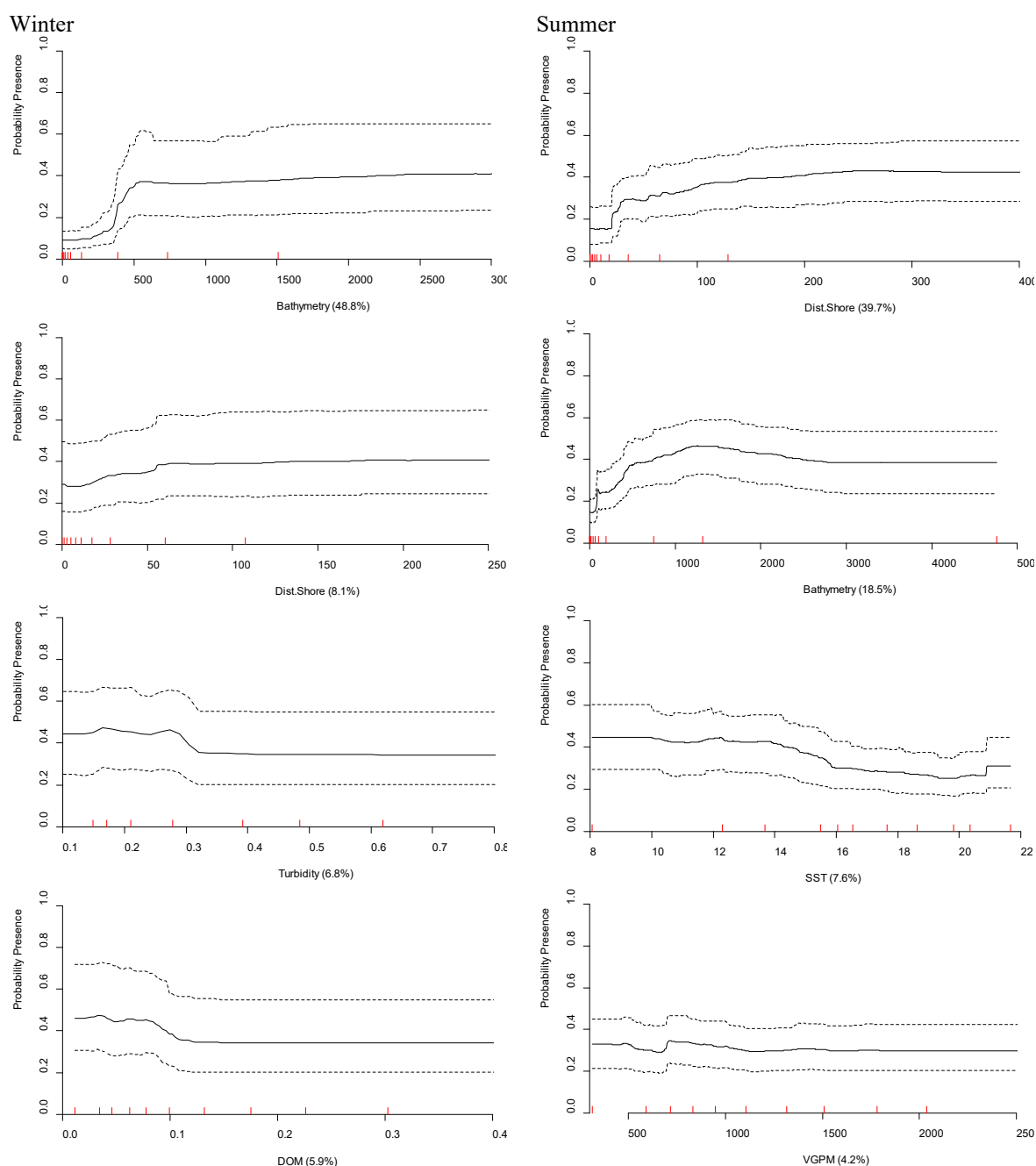


Figure 138: Partial dependence plots showing the relationships between predictor variables and predicted presence of pilot whales (*Globicephala melas* and *Globicephala macrorhynchus*) modelled using bootstrapped BRTs for Winter (left) and Summer (right). The four most influential environmental predictors in the model are shown for Winter (left) and Summer (right). Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x axes).

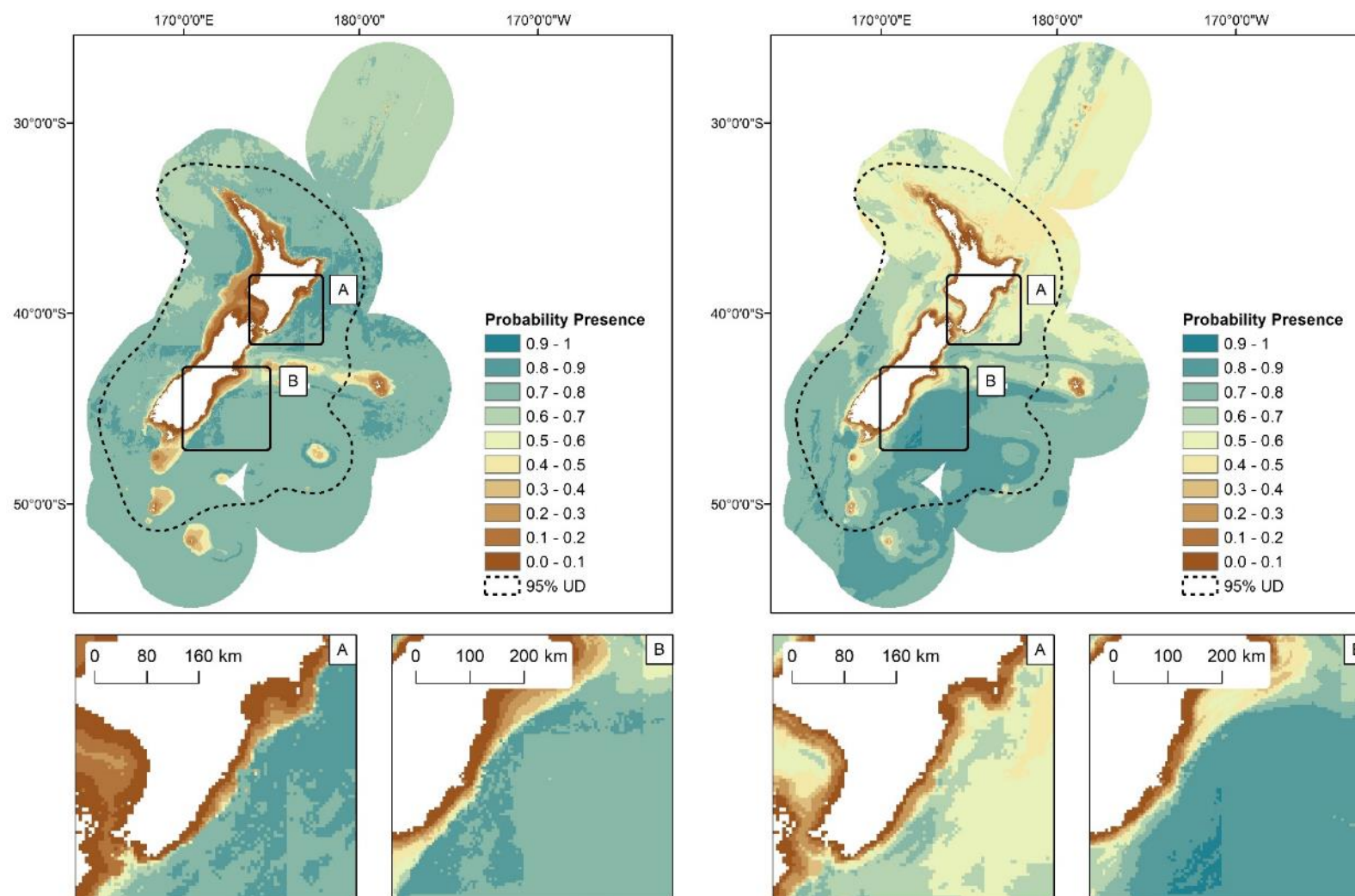


Figure 142: Seasonal predicted probability presence of pilot whale (*Globicephala melas* and *Globicephala macrorhynchus*) modelled using bootstrapped BRTs fitted with Winter (May - Oct, n = 169) (left) and Summer (Nov - Apr, n = 654) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line (see section x for further details on methods and interpretation). Inset maps: A) South east coast of the North Island; B) South east coast of the South Island.

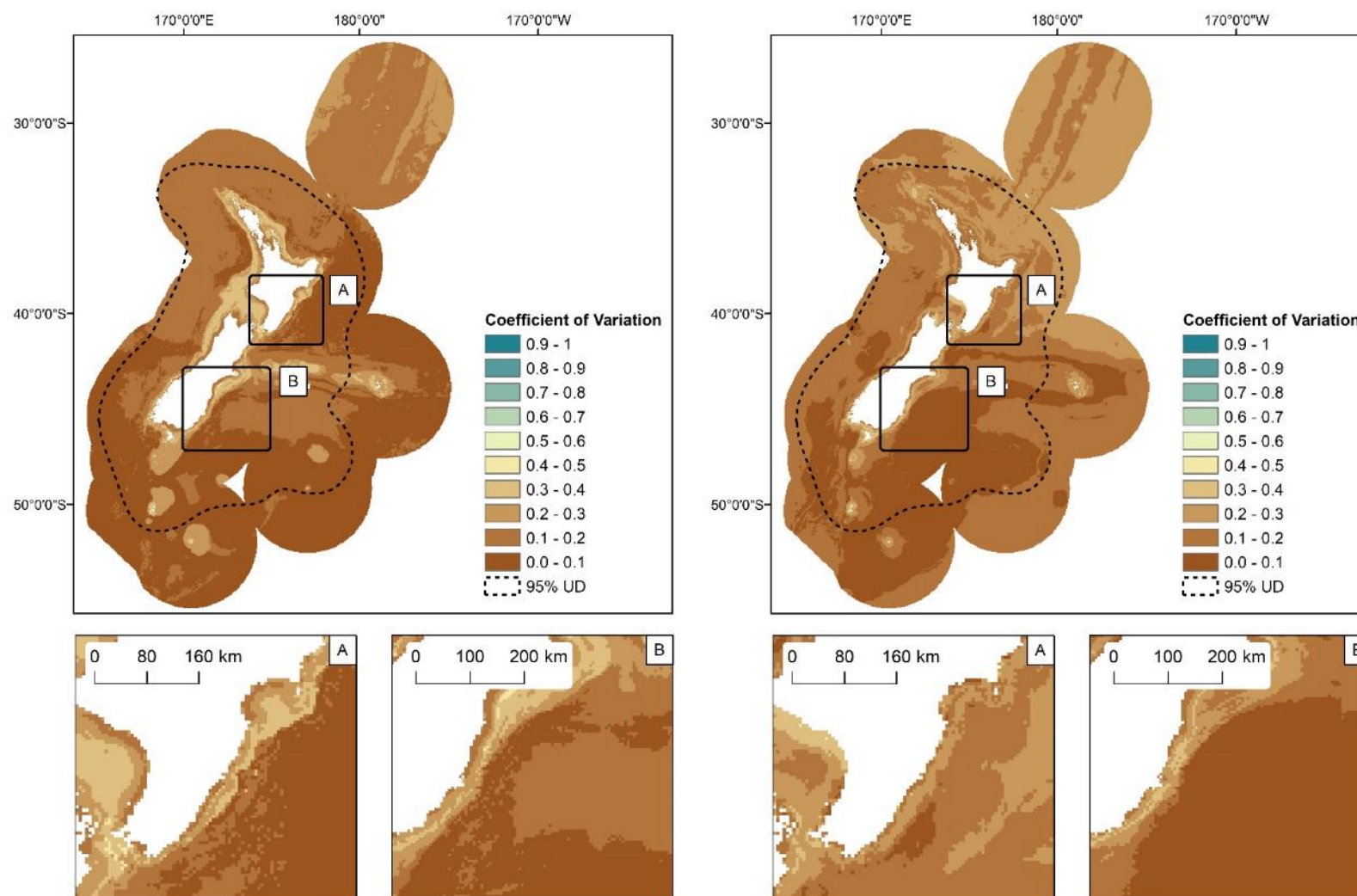


Figure 143: Seasonal uncertainty estimates (coefficient of variation, CV) of pilot whale (*Globicephala melas* and *Globicephala macrorhynchus*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs fitted with Winter (May - Oct) (left) and Summer (Nov - Apr) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line (see section x for further details on methods and interpretation). Inset maps: A) South east coast of the North Island; B) South east coast of the South Island.

Killer whale (*Orcinus orca*)

Table 50: Killer whale (*Orcinus orca*): Highlights of model fits and geographic prediction.

Sample number	Distribution	P/RA		Group size model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
High	Cosmopolitan – highest probability presence close to shore	Fair	Fair	Fair	Poor	Increasing inshore use in summer compared to winter

Table 51: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of killer whale (*Orcinus orca*).

	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Mean	0.18	0.16	0.79	0.79
Standard Deviation	0.02	0.02	0.01	0.02

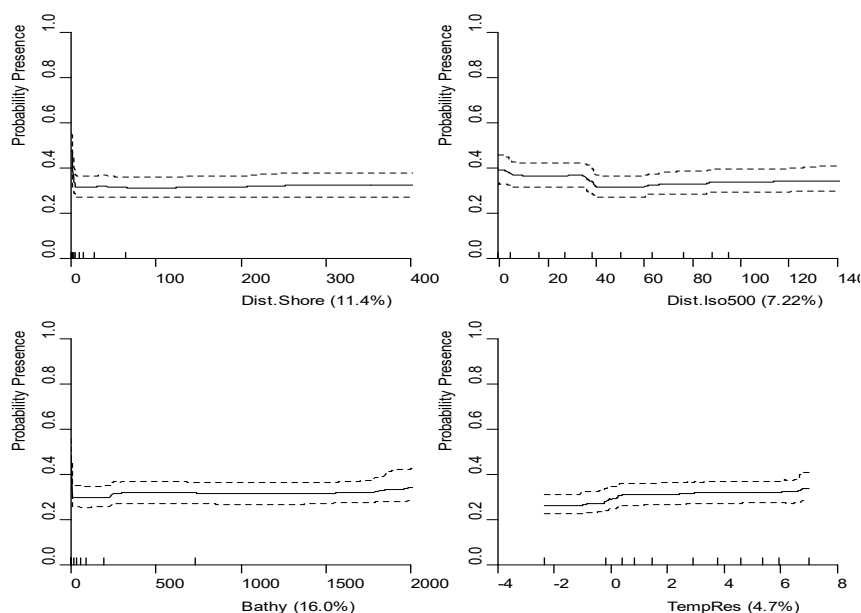


Figure 141: Partial dependence plots showing the relationships between predictor variables and probability presence of killer whale (*Orcinus orca*) modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

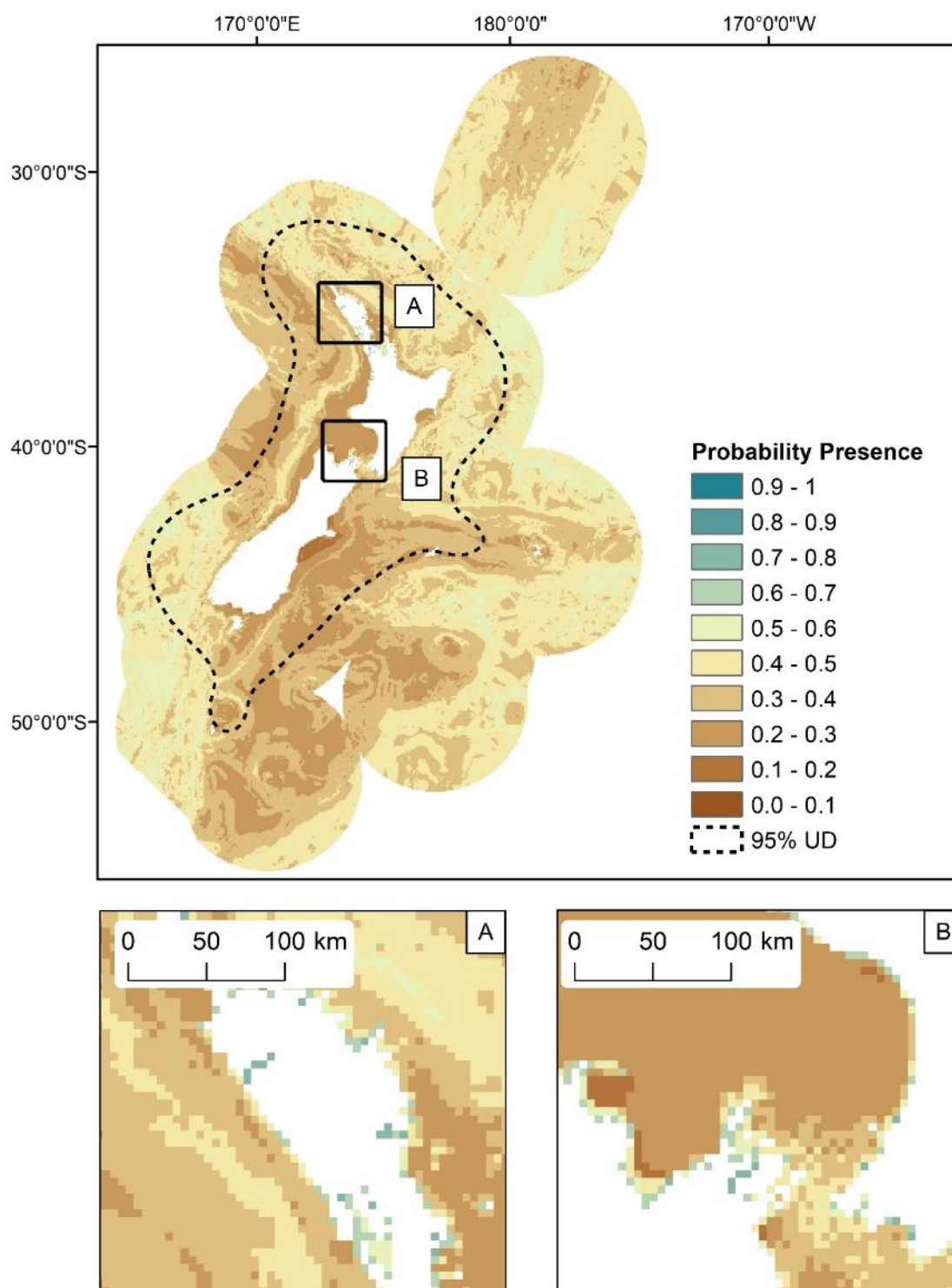


Figure 142: The predicted probability of killer whale (*Orcinus orca*) presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island; B) Cook Strait, including south North Island and north South Island.

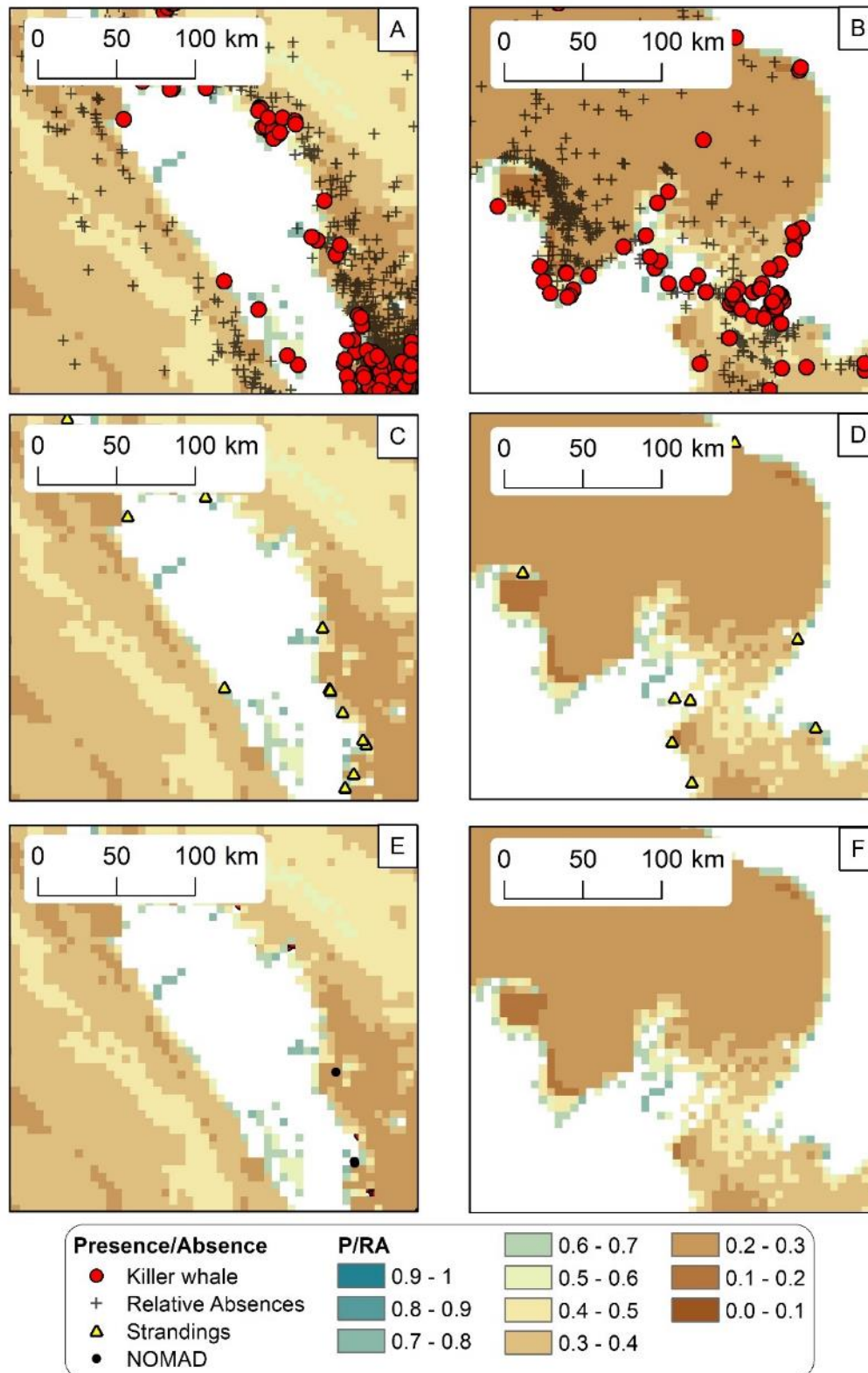


Figure 143: The predicted probability of killer whale (*Orcinus orca*) presence modelled using bootstrapped BRTs. Predicted probability presence of killer whale in the north North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted probability presence of Humpback whale in the Cook Strait are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

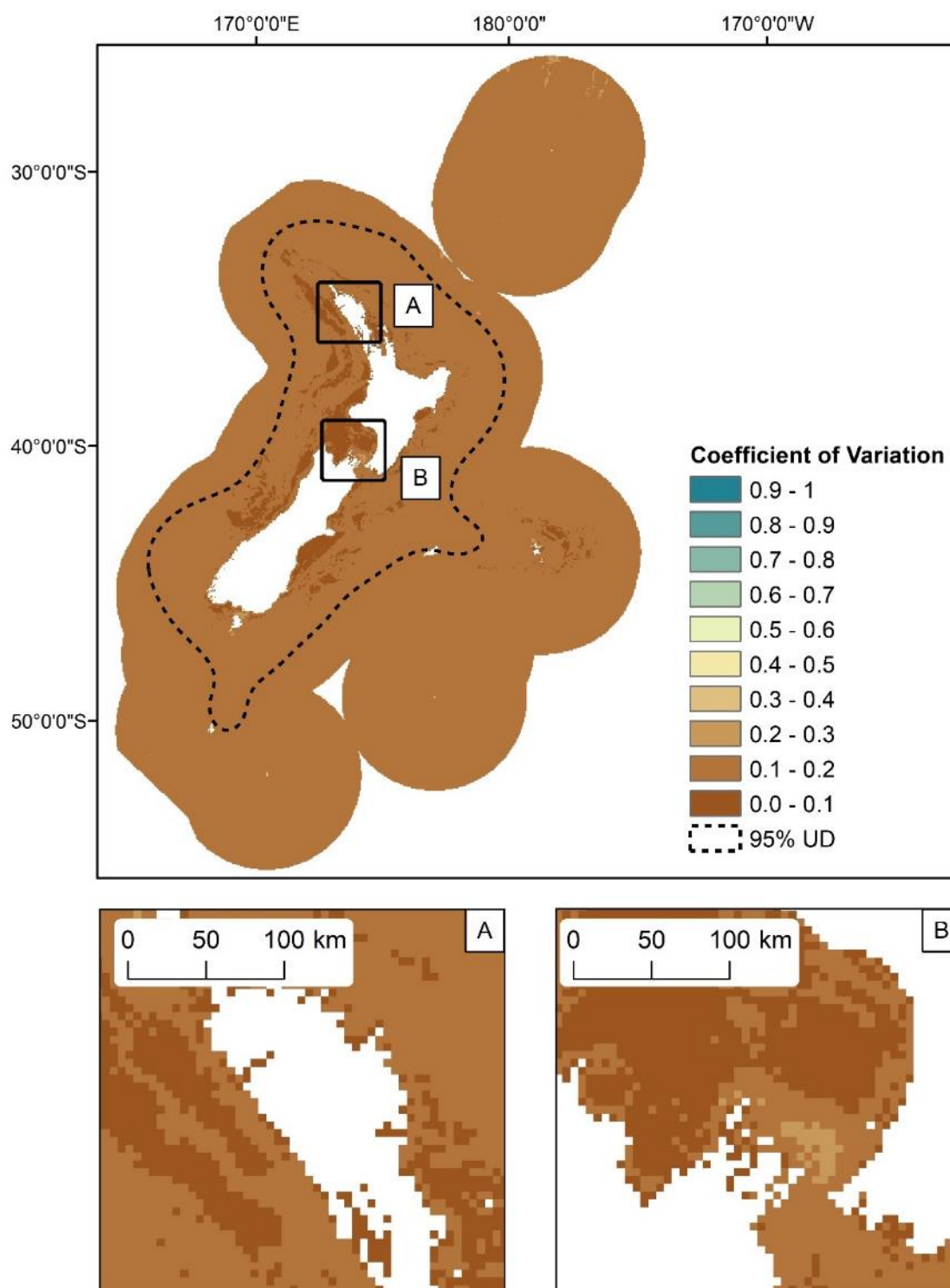


Figure 144: Uncertainty estimates (coefficient of variation, CV) of killer whale (*Orcinus orca*) probability of presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island; B) Cook Strait, including south North Island and north South Island.

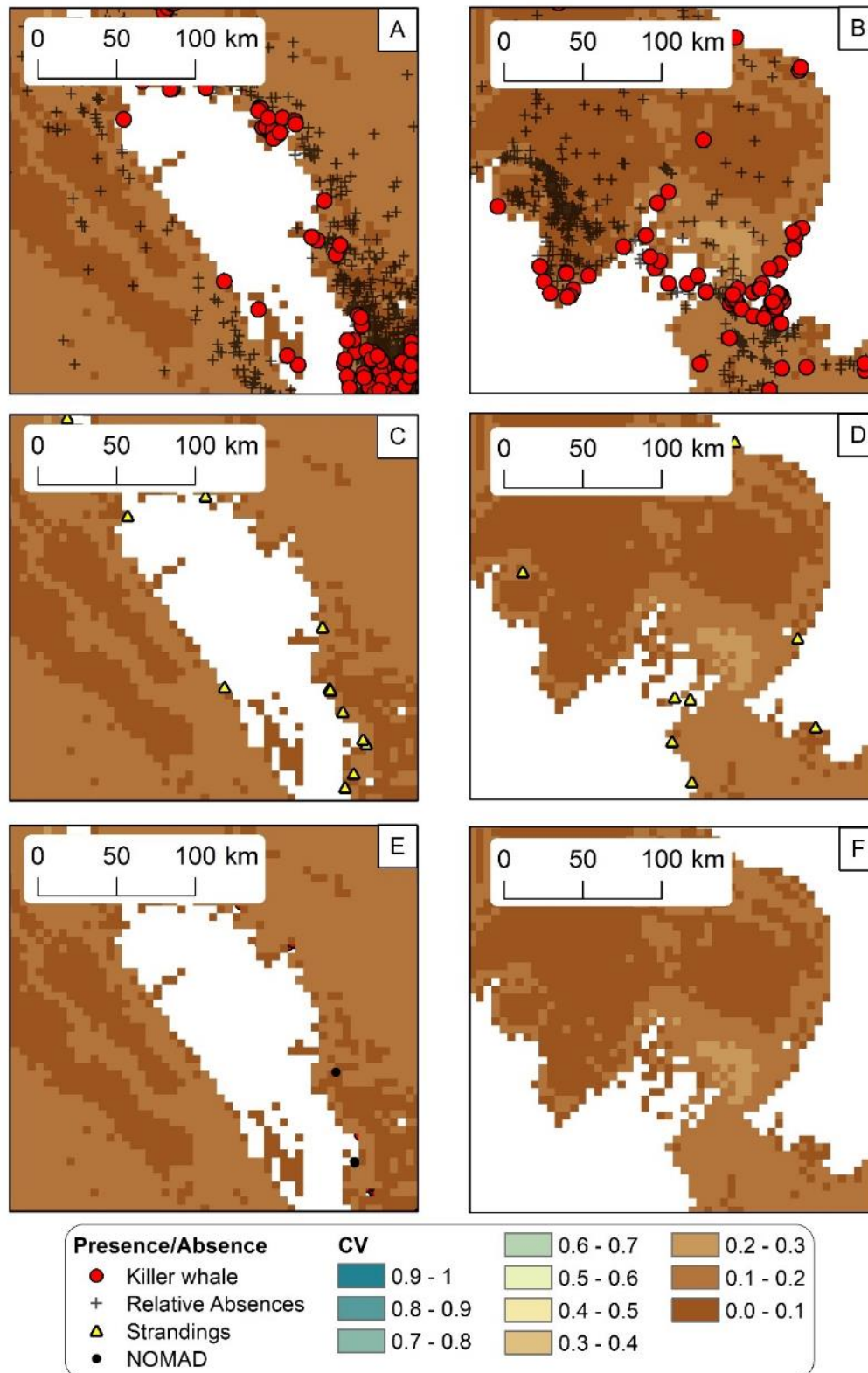


Figure 145: Uncertainty estimates (coefficient of variation, CV) of killer whale (*Orcinus orca*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Killer whale probability presence models in the north of North Island are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of killer whale probability presence models in the Cook Strait are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

Table 52: Mean model performance measures (deviance explained and Pearson's correlation of predicted vs observed group sizes (R^2)) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of killer whale (*Orcinus orca*).

	Deviance explained (training data)	Deviance explained (validation data)	Pearson's correlation of predicted vs observed species group sizes (R^2)
Mean	0.03	0.00	0.22
Standard Deviation	0.03	0.05	0.11

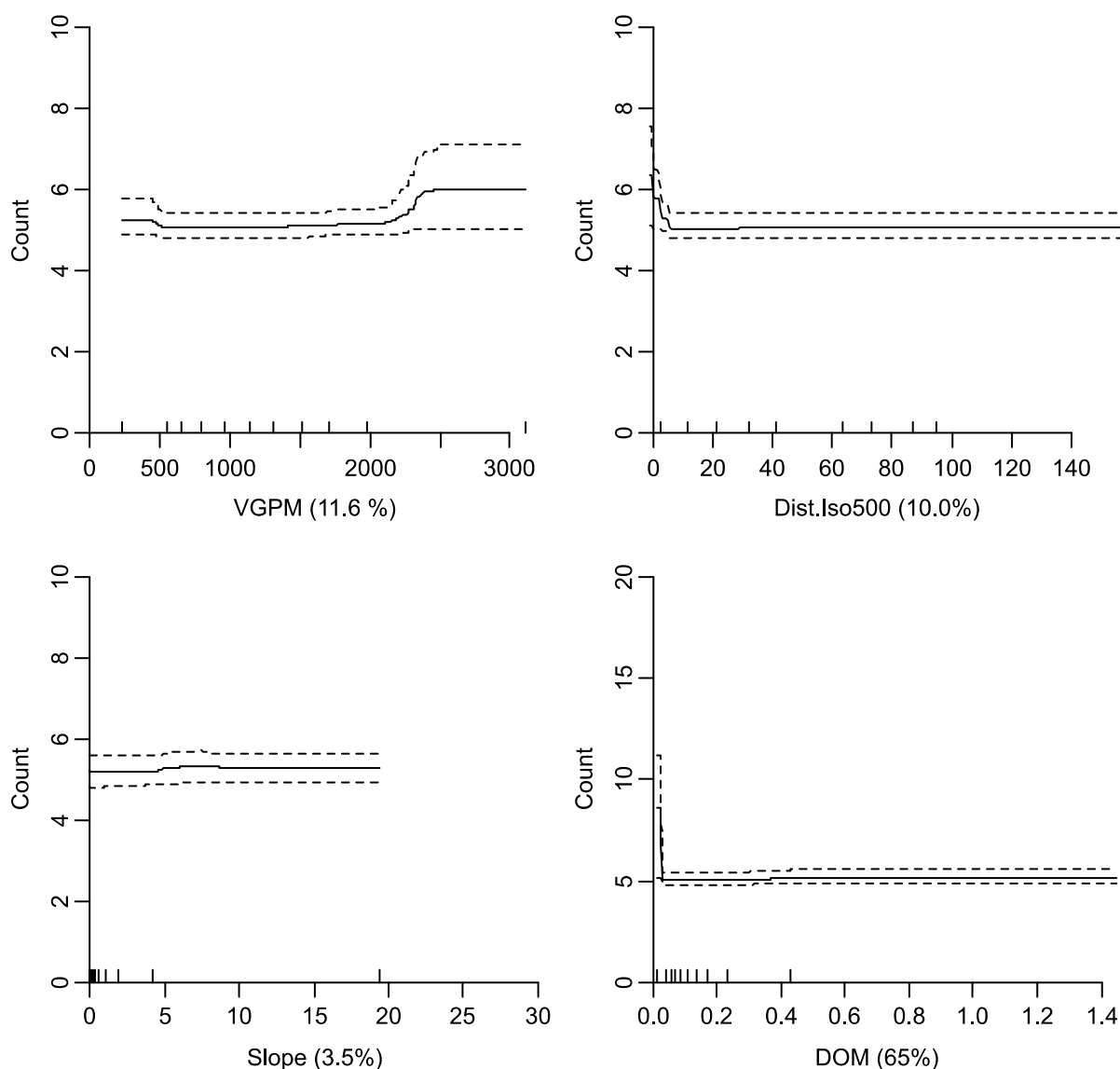


Figure 146: Partial dependence plot showing the relationship between the predictor variable (VGPM) and killer whale (*Orcinus orca*) group sizes modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x axes).

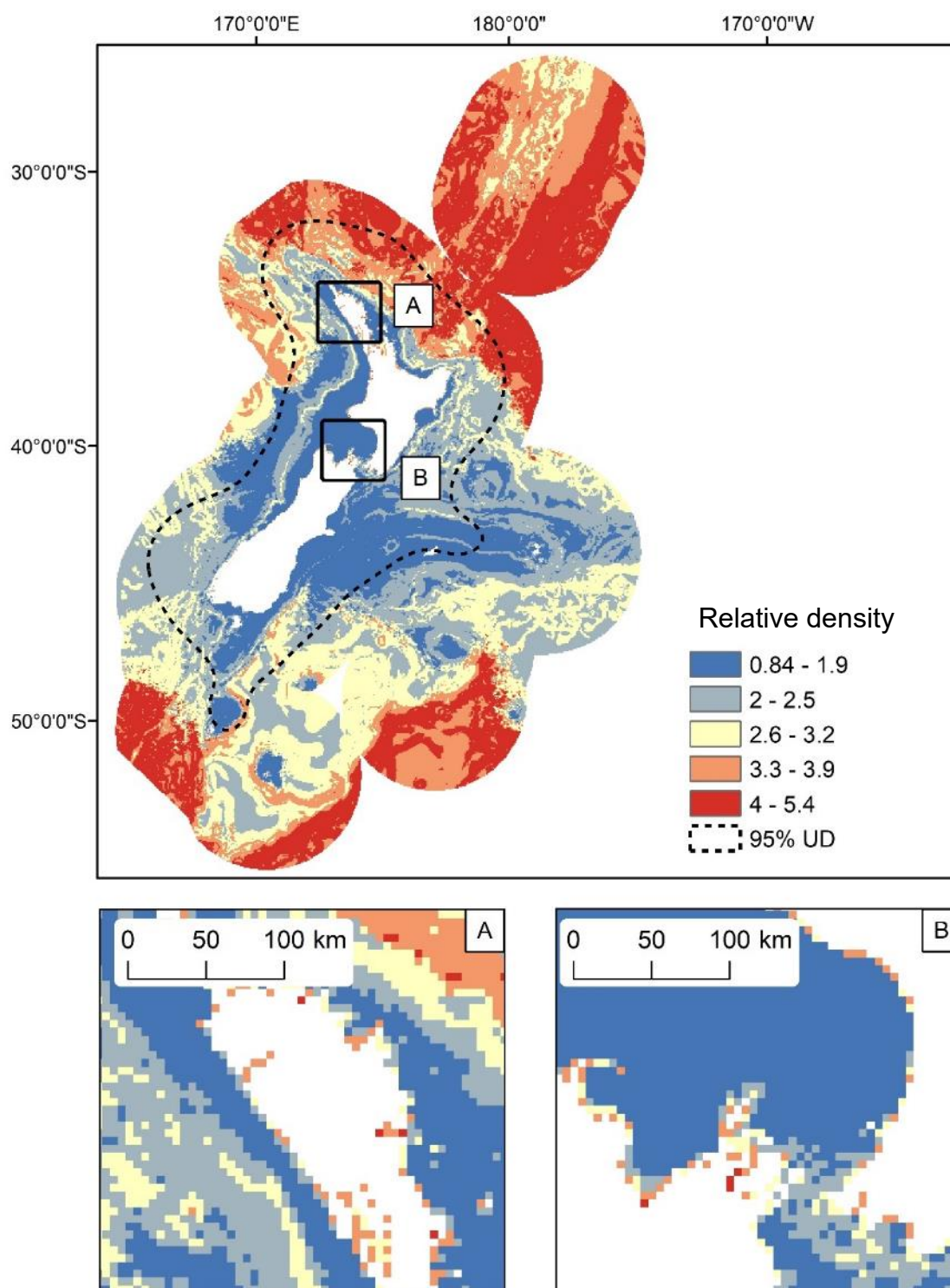


Figure 147: Predicted relative density of killer whale (*Orcinus orca*) in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island; B) Cook Strait, including south North Island and north South Island.

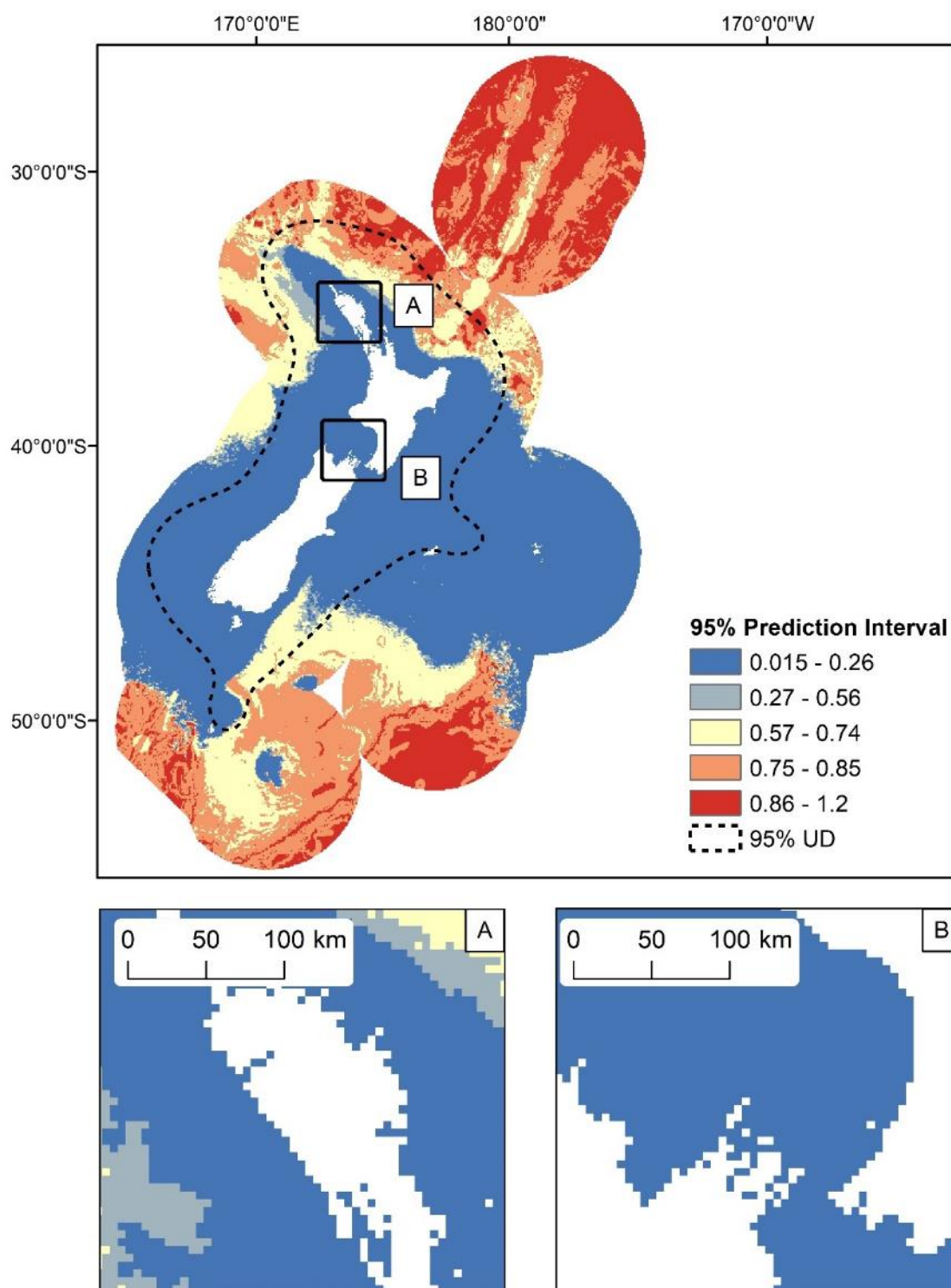


Figure 148: Uncertainty estimates (95% prediction interval) of killer whale (*Orcinus orca*) predicted counts in the New Zealand EEZ modelled using bootstrapped BRTs. 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island; B) Cook Strait, including south North Island and north South Island.

Table 53: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) from Winter (May - Oct) and Summer (Nov - Apr) sightings of killer whale (*Orcinus orca*) records.

Season	Metric	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Winter	Mean	0.09	0.08	0.71	0.73
	Standard Deviation	0.24	0.36	0.03	0.06
Summer	Mean	0.27	0.26	0.84	0.84
	Standard Deviation	0.09	0.16	0.02	0.03

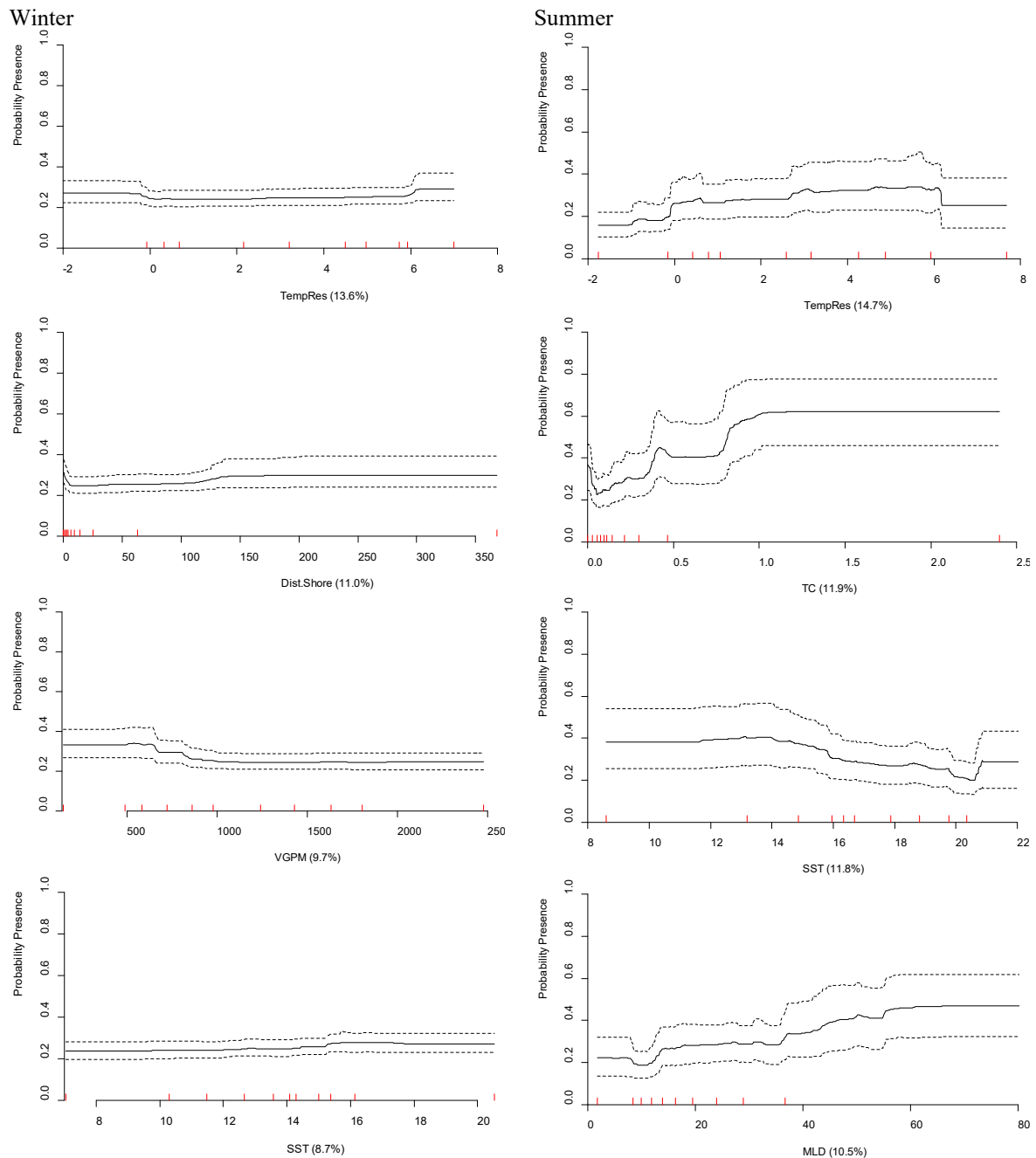


Figure 149: Partial dependence plots showing the relationships between predictor variables and predicted presence of killer whale (*Orcinus orca*) modelled using bootstrapped BRTs for Winter (left) and Summer (right). The four most influential environmental predictors in the model are shown for Winter (left) and Summer (right). Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x axes).

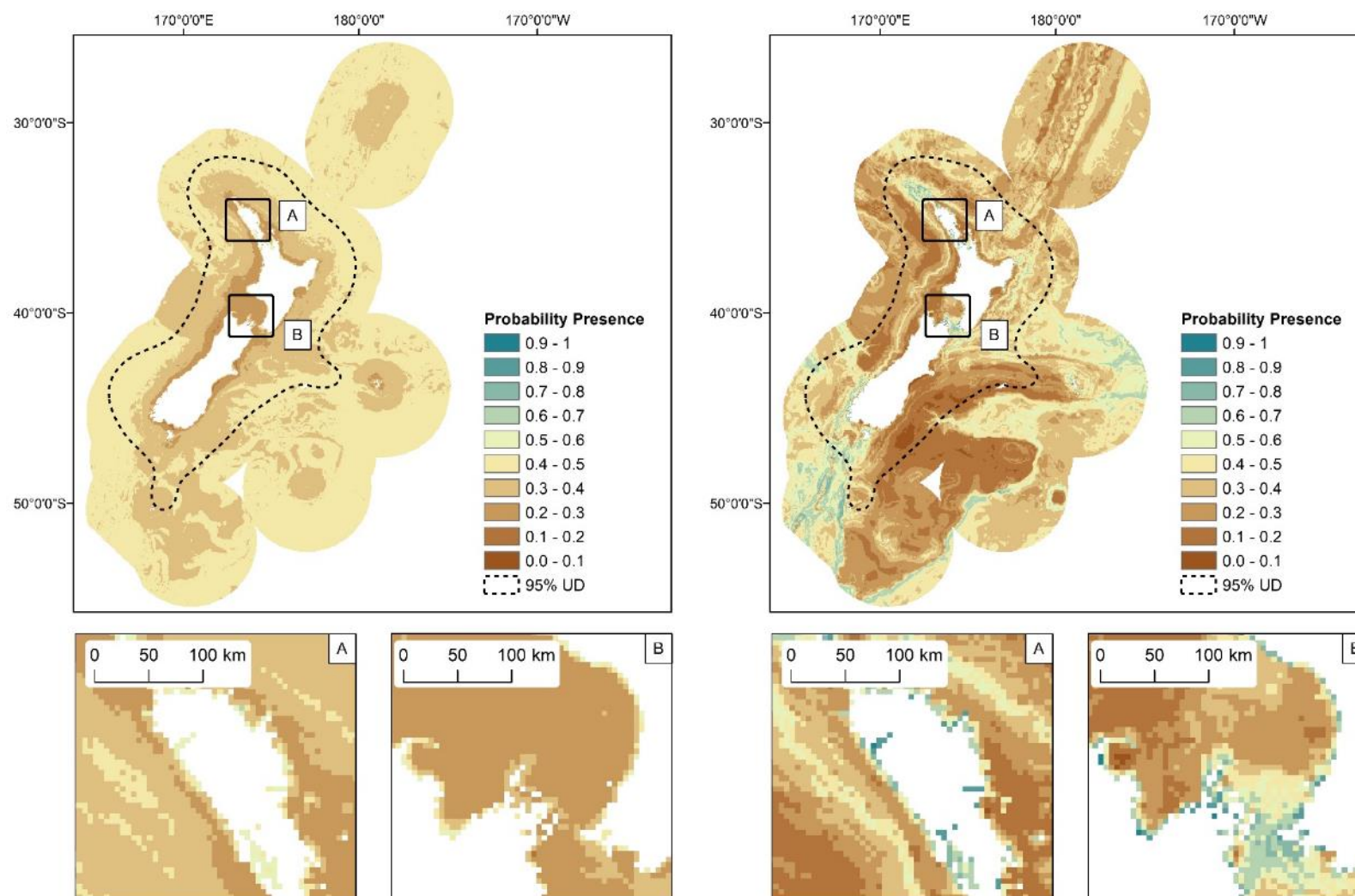


Figure 153: Seasonal predicted probability of killer whale (*Orcinus orca*) presence modelled using bootstrapped BRTs fitted with Winter (May - Oct, n = 169) (left) and Summer (Nov - Apr, n = 654) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line (see section x for further details on methods and interpretation). Inset maps: A) North North Island; B) Cook Strait, including south North Island and north South Island.

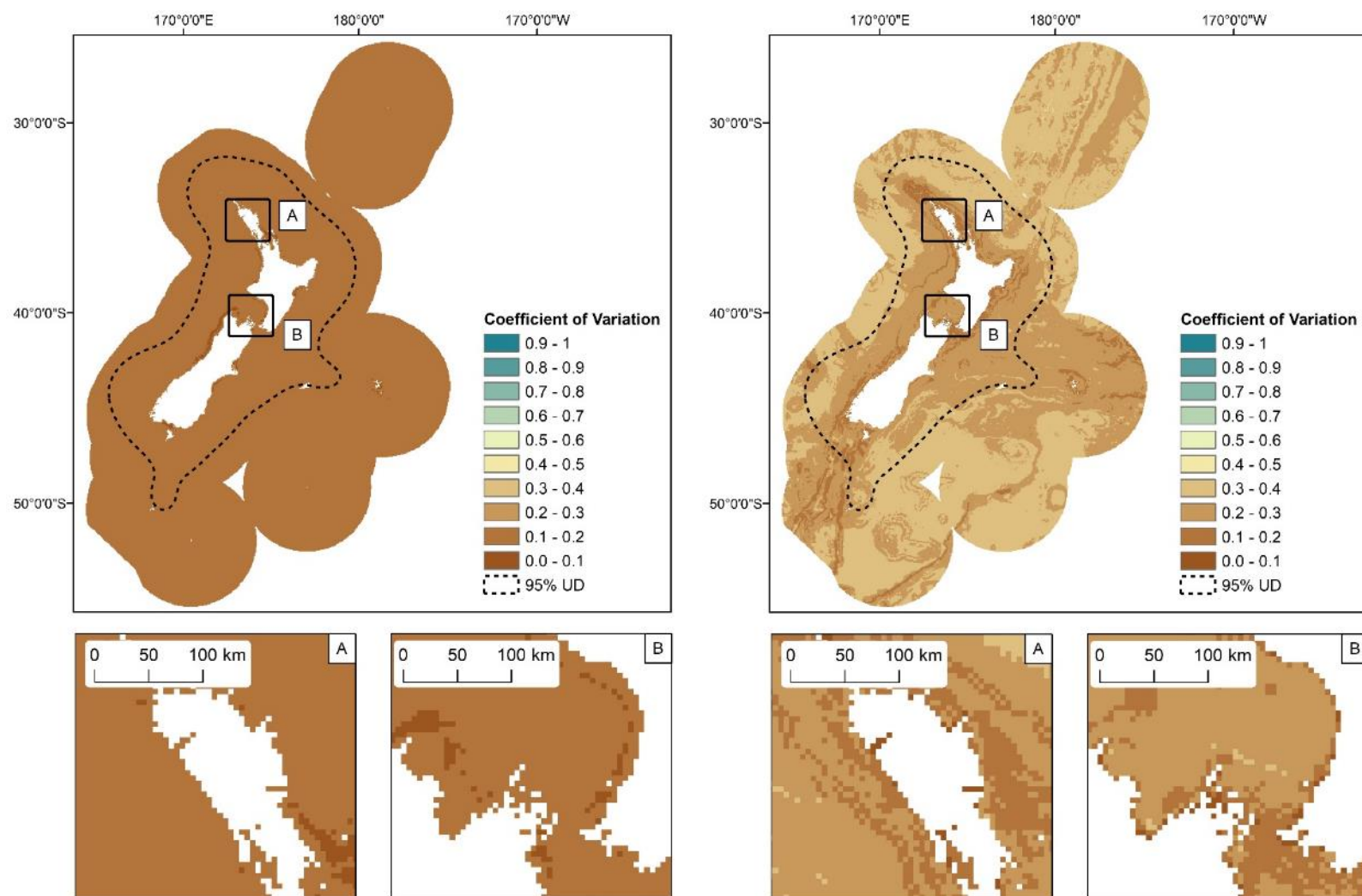


Figure 154: Seasonal uncertainty estimates (coefficient of variation, CV) of killer whale (*Orcinus orca*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs fitted with Winter (May - Oct) (left) and Summer (Nov - Apr) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island; B) Cook Strait, including south North Island and north South Island.

Māui dolphin (*Cephalorhynchus hectori māui*)

Table 54: Māui dolphin (*Cephalorhynchus hectori māui*): Highlights of model fits and geographic prediction.

Sample number	Distribution	P/RA		Count model		Changes in seasonal distribution
		Model fit (AUC)	Model fit (dev. Exp)	Model fit (R ²)	Model fit (dev. Exp)	
Very high	Highly localised to close to shore in north of North Island	Excellent	Excellent	Good	Poor	Little to no change seasonally

Table 55: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of Māui dolphin (*Cephalorhynchus hectori māui*).

	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Mean	0.87	0.88	0.99	0.99
Standard Deviation	0.01	0.02	0.00	0.00

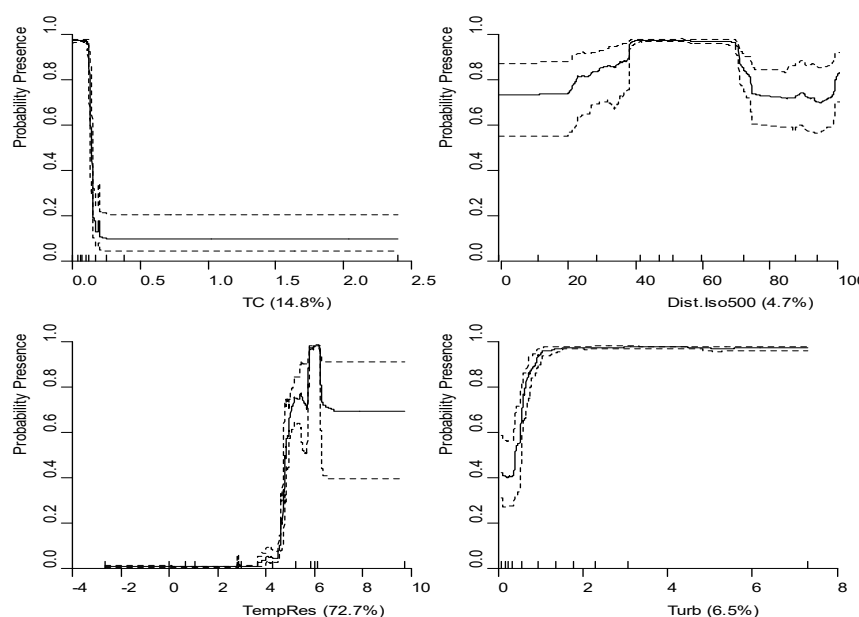


Figure 152: Partial dependence plots showing the relationships between predictor variables and probability presence of Māui dolphin (*Cephalorhynchus hectori māui*) modelled using bootstrapped BRTs. The four most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

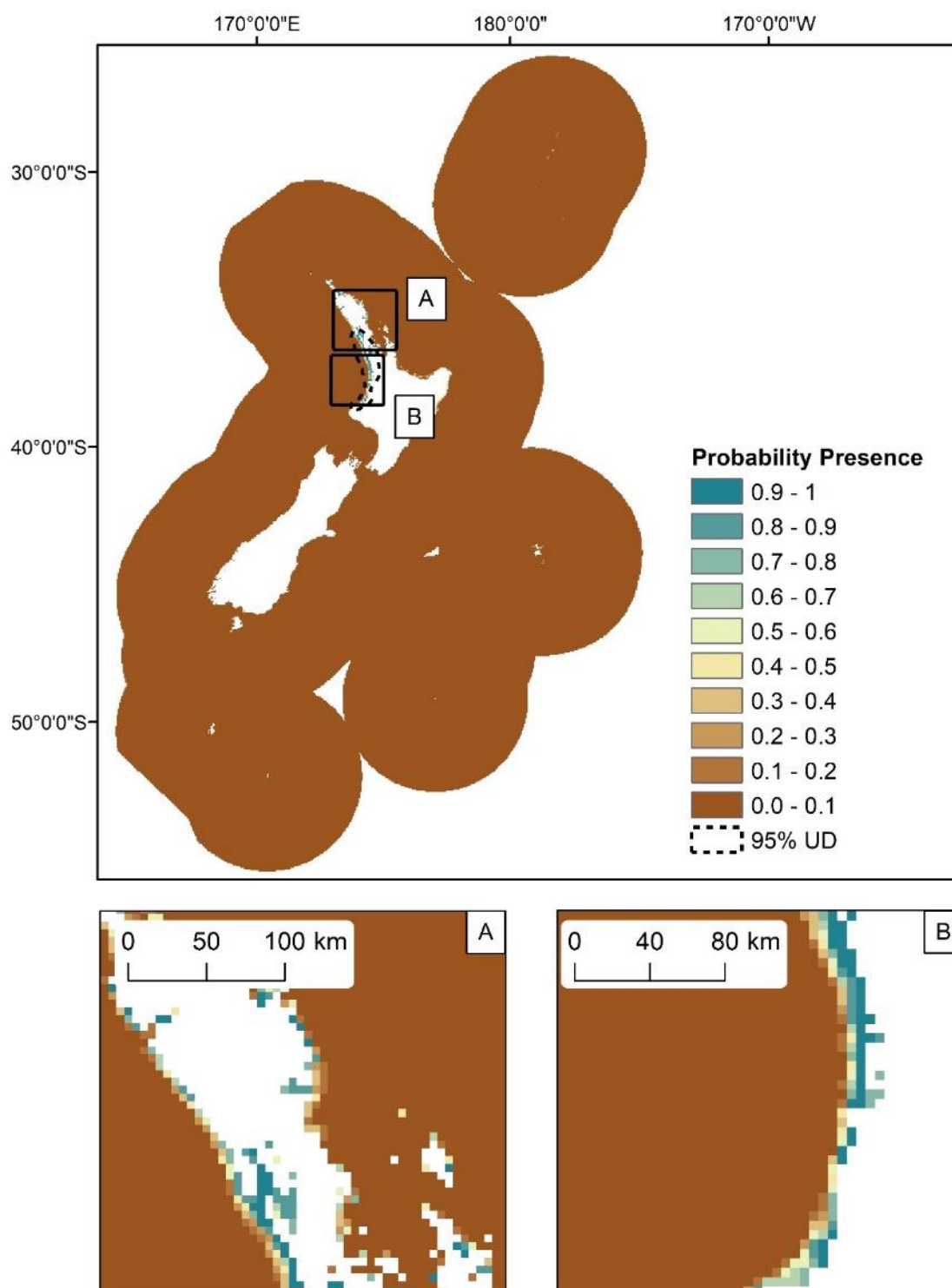


Figure 153: The predicted probability of Maui dolphin (*Cephalorhynchus hectori māui*) presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island and Hauraki Gulf; B) West coast of North Island.

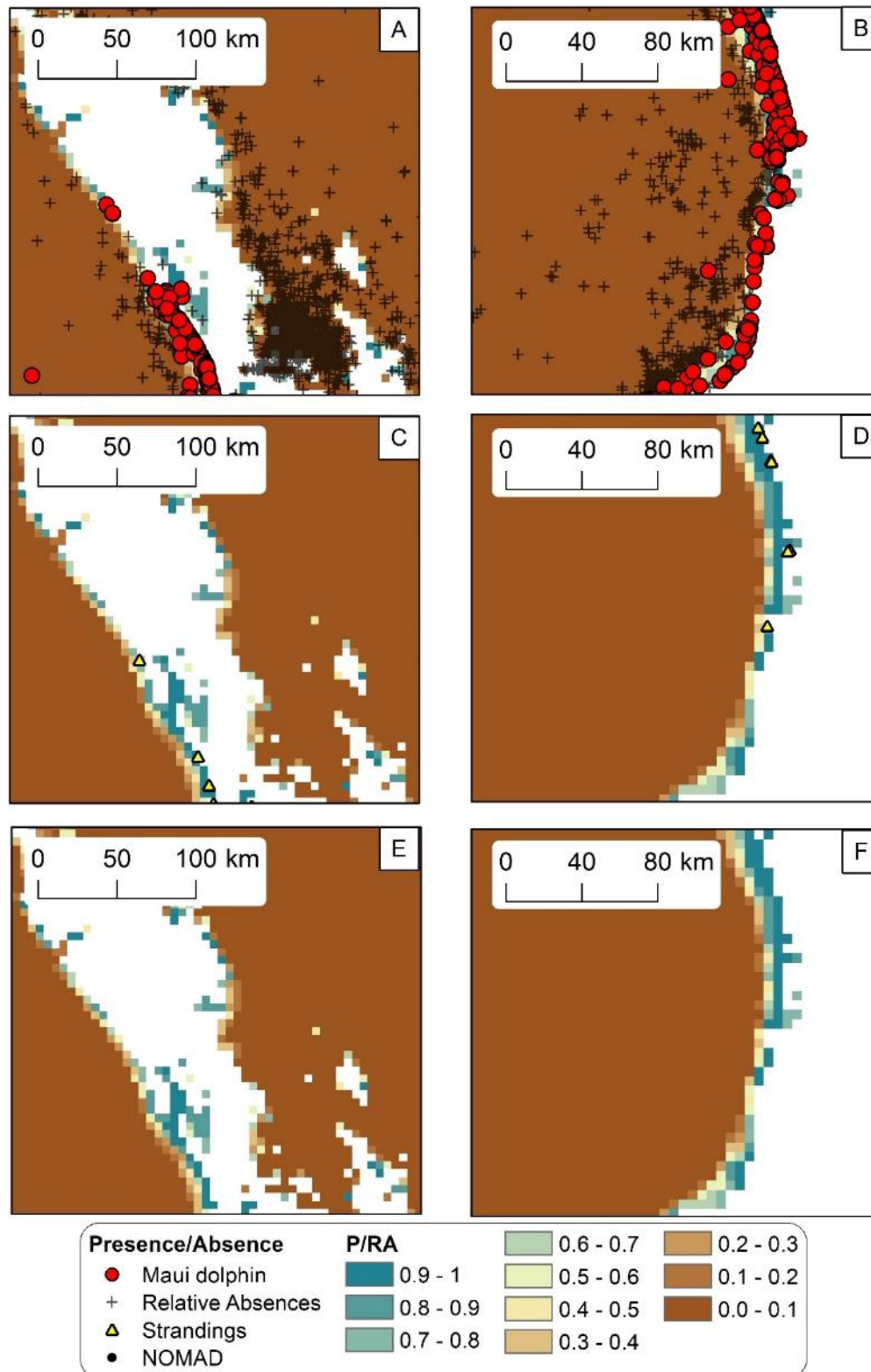


Figure 154: The predicted probability presence of Māui dolphin (*Cephalorhynchus hectori māui*) modelled using bootstrapped BRTs. Predicted probability presence of Māui dolphin in the north of North Island and Hauraki Gulf are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted probability presence of Māui dolphin on the west coast of the North Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

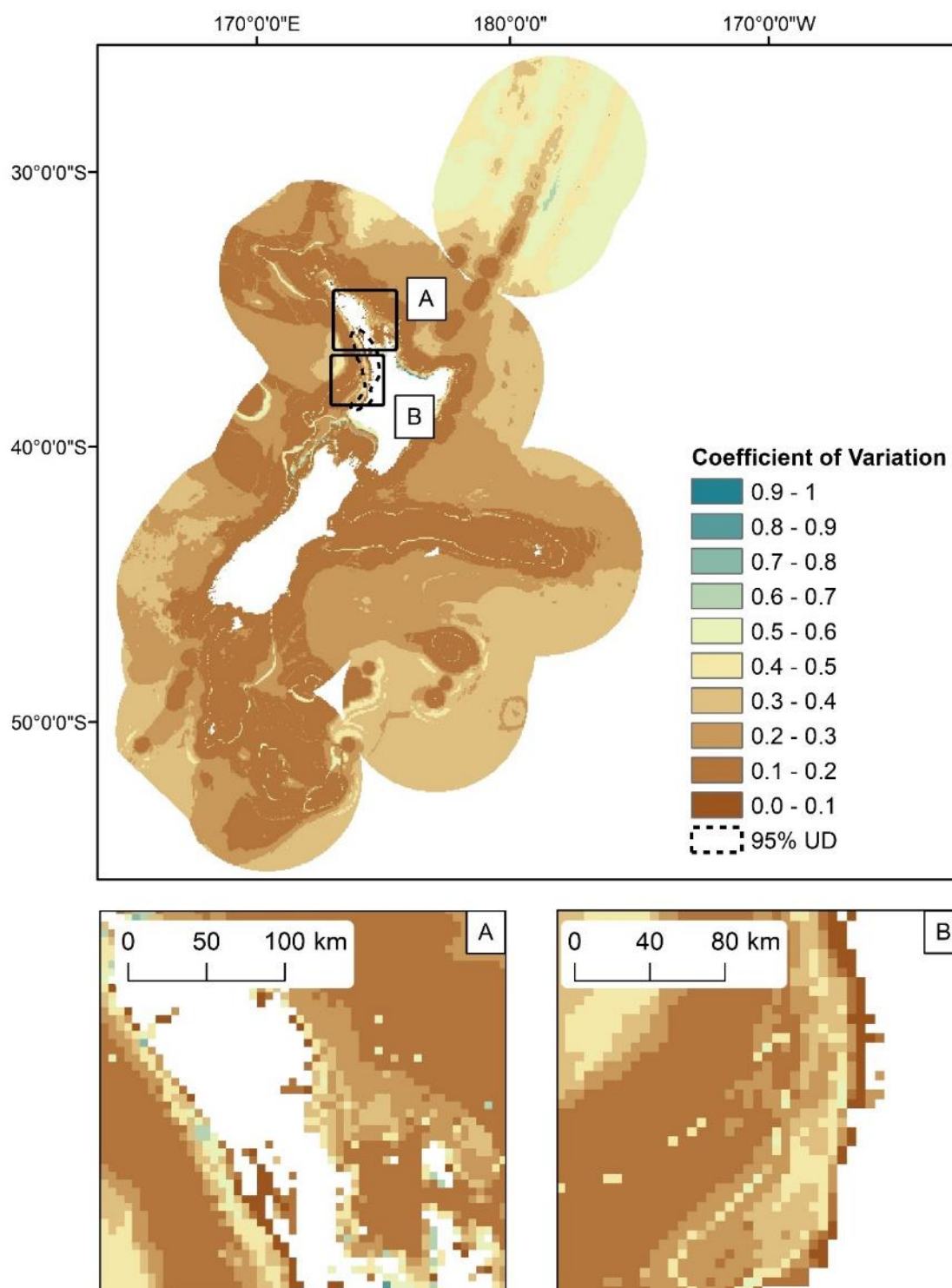


Figure 155: Uncertainty estimates (coefficient of variation, CV) of Māui dolphin (*Cephalorhynchus hectori māui*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island and Hauraki Gulf; B) West coast of North Island.

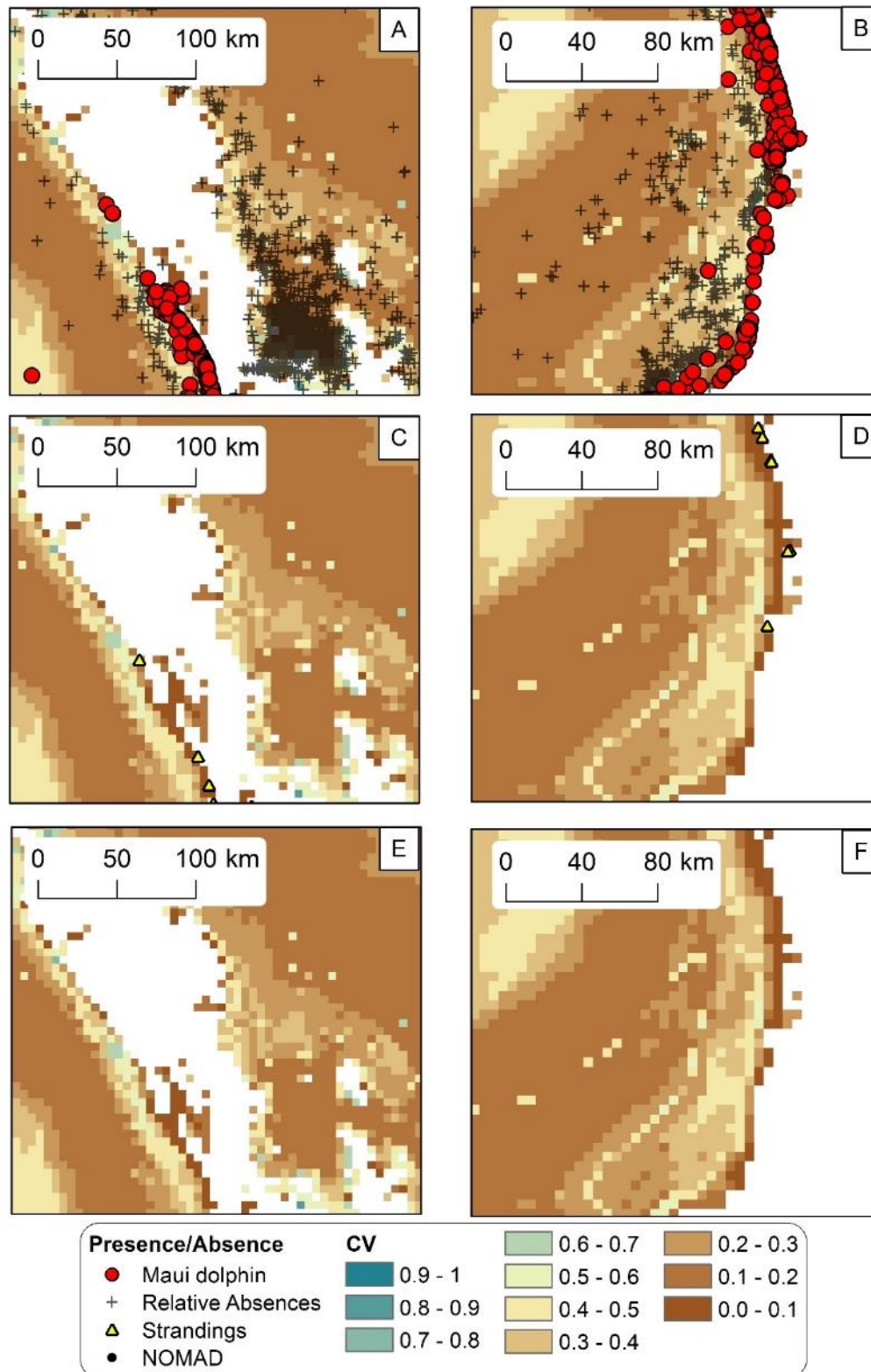


Figure 156: Uncertainty estimates (coefficient of variation, CV) of Māui dolphin (*Cephalorhynchus hectori māui*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs. Predicted CV of Māui dolphin probability presence models in the north of the North Island and Hauraki Gulf are shown with presence/relative absences (red circles and black crosses respectively) (A), DOC stranding locations (C) and NOMAD sightings (E). Predicted CV of Māui dolphin probability presence models on the west coast of the North Island are shown with presence/relative absences (B), stranding locations (D) and NOMAD sightings (F).

Table 56: Mean model performance measures (deviance explained and Pearson's correlation of predicted vs observed counts (R^2)) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) of Māui dolphin (*Cephalorhynchus hectori māui*).

	Deviance explained (training data)	Deviance explained (validation data)	Pearson's correlation of predicted vs observed species counts (R^2)
Mean	0.09	0.09	0.27
Standard Deviation	0.02	0.03	0.05

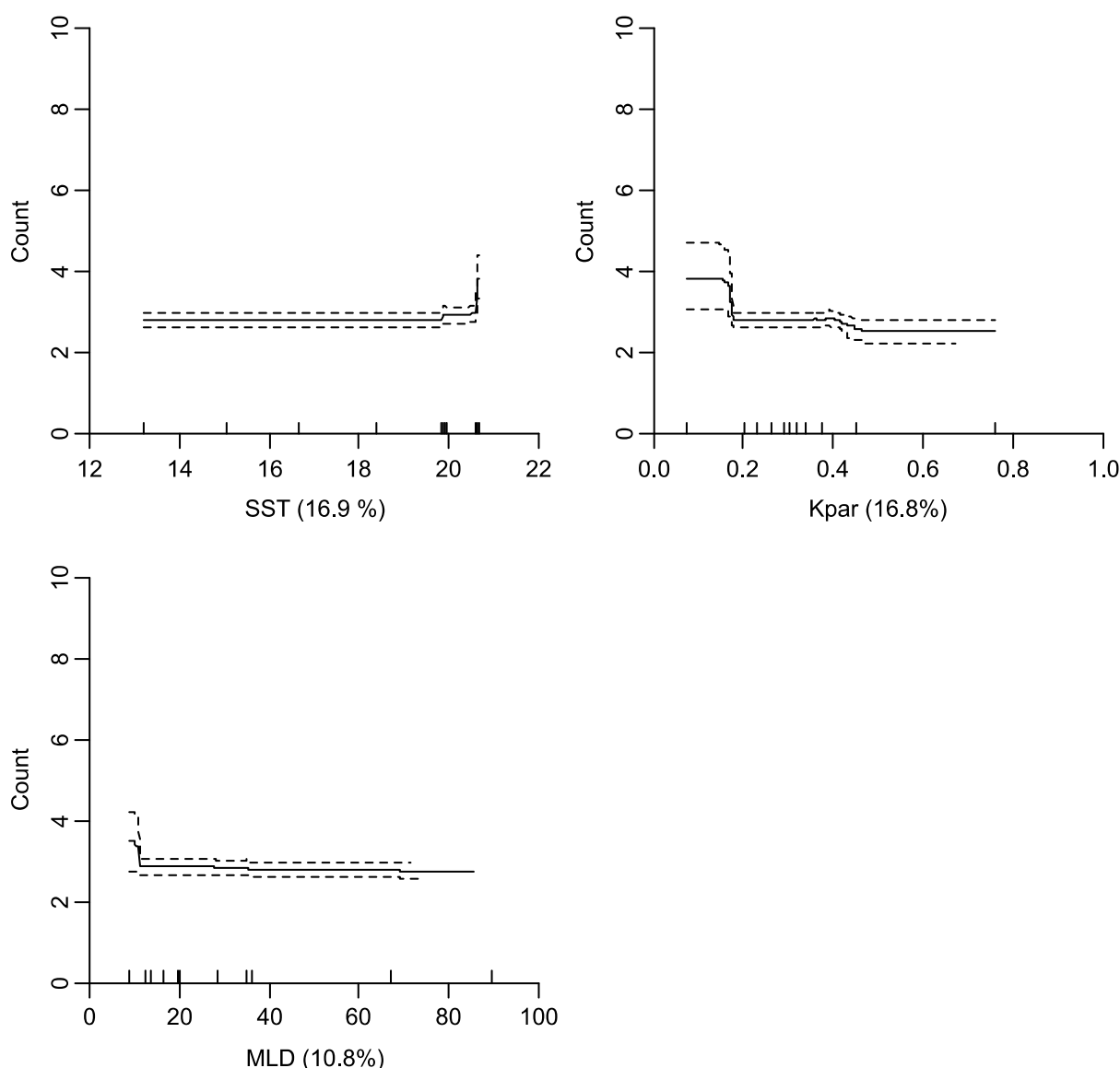


Figure 157: Partial dependence plot showing the relationship between the predictor variable (SST) and Māui dolphin (*Cephalorhynchus hectori māui*) counts modelled using bootstrapped BRTs. The three most influential environmental predictors in the model are shown. Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x-axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x-axes).

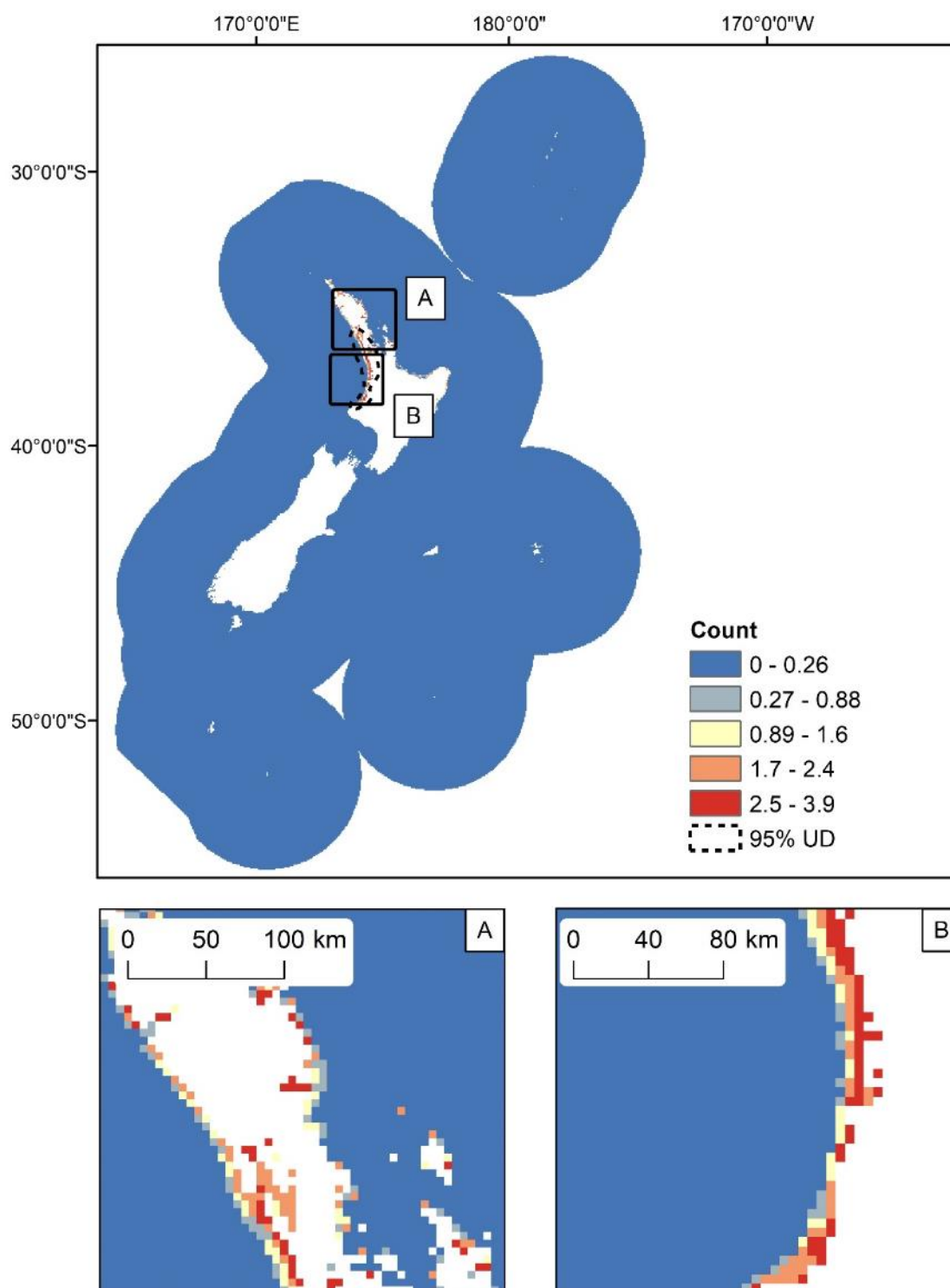


Figure 158: Predicted counts of Maui dolphin (*Cephalorhynchus hectori māui*) in the New Zealand EEZ modelled using bootstrapped BRTs. The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island and Hauraki Gulf; B) West coast of North Island.

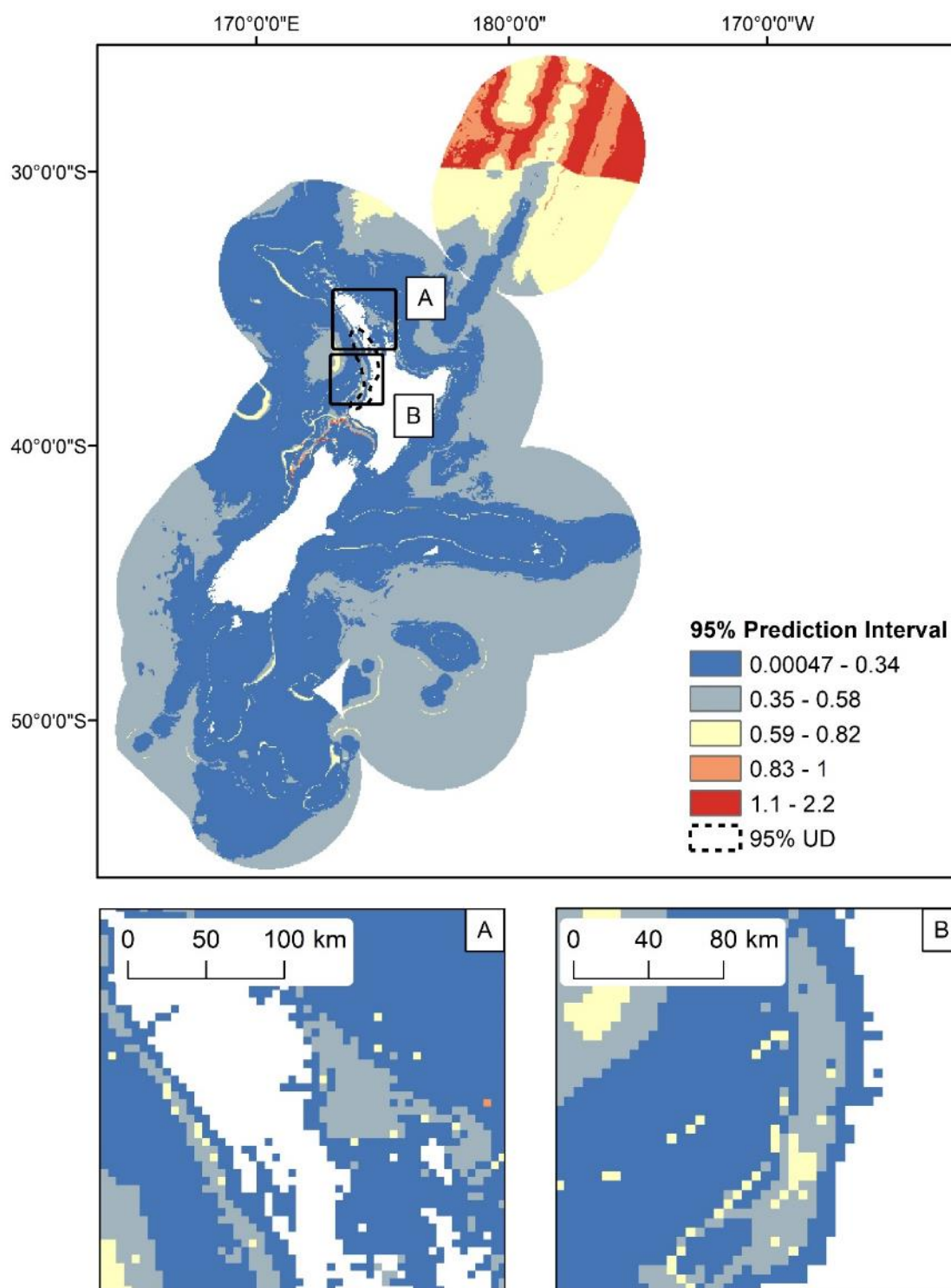


Figure 159: Uncertainty estimates (95% prediction interval) of Maui dolphin (*Cephalorhynchus hectori māui*) predicted counts in the New Zealand EEZ modelled using bootstrapped BRTs. 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island and Hauraki Gulf; B) West coast of North Island.

Table 57: Mean model performance measures (deviance explained and AUC) for bootstrapped BRT models fitted with training records (75%) and evaluation records (25%) from Winter (May - Oct) and Summer (Nov - Apr) sightings of Māui dolphin (*Cephalorhynchus hectori māui*) records.

Season	Metric	Deviance explained (training data)	Deviance explained (validation data)	AUC (training data)	AUC (validation data)
Winter	Mean	0.83	0.85	0.99	0.99
	Standard Deviation	0.03	0.08	0.00	0.01
Summer	Mean	0.89	0.89	1.00	1.00
	Standard Deviation	0.01	0.03	0.00	0.00

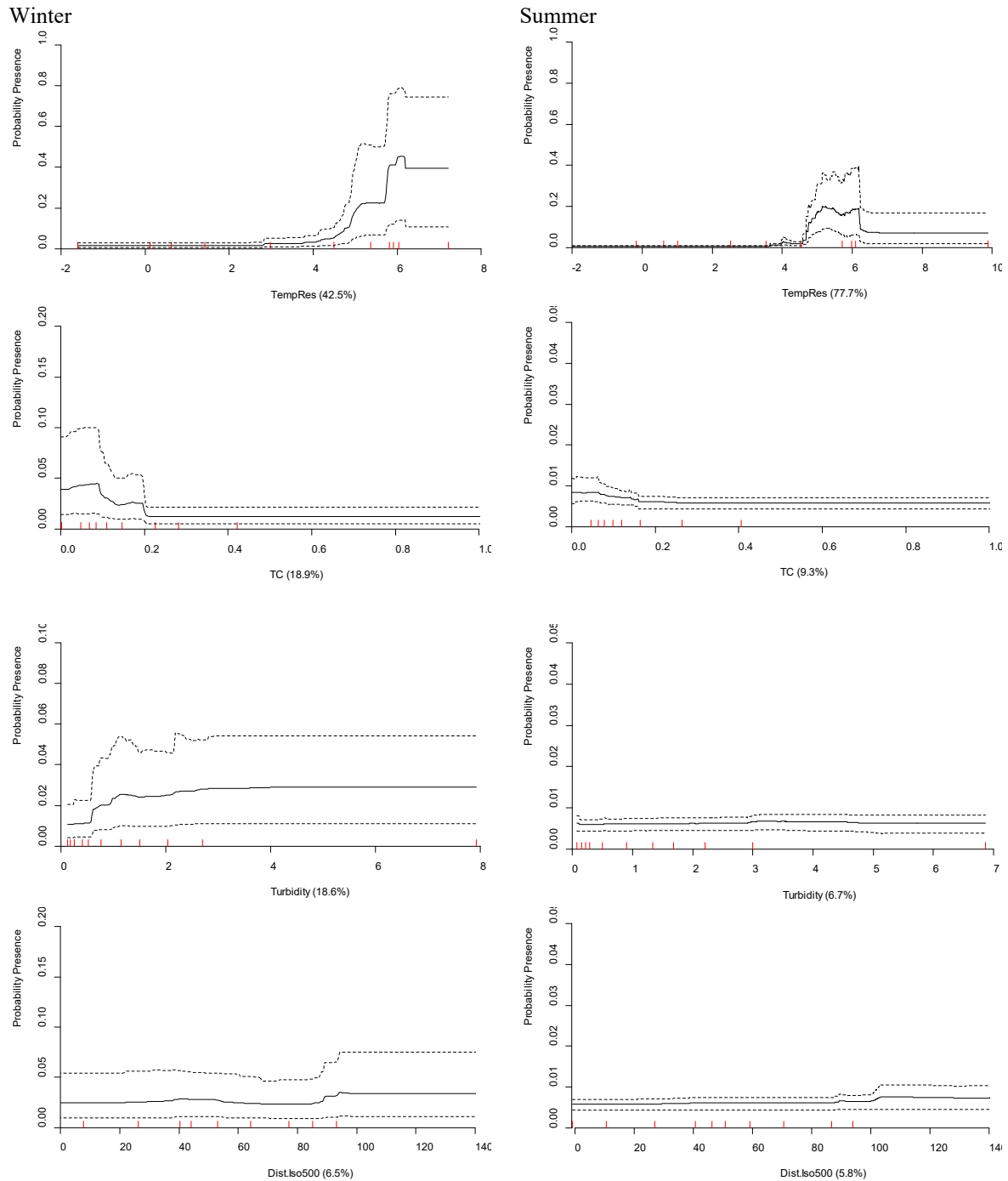


Figure 160: Partial dependence plots showing the relationships between predictor variables and predicted presence of Māui dolphin (*Cephalorhynchus hectori māui*) modelled using bootstrapped BRTs for Winter (left) and Summer (right). The four most influential environmental predictors in the model are shown for Winter (left) and Summer (right). Solid lines represent the mean of 100 bootstrap predictions and dashed lines the 95% prediction interval. Quantiles of each environmental predictor are shown on the x axes. Each plot represents a predictor variable (labels and relative percentage contribution in parentheses are shown on the x axes).

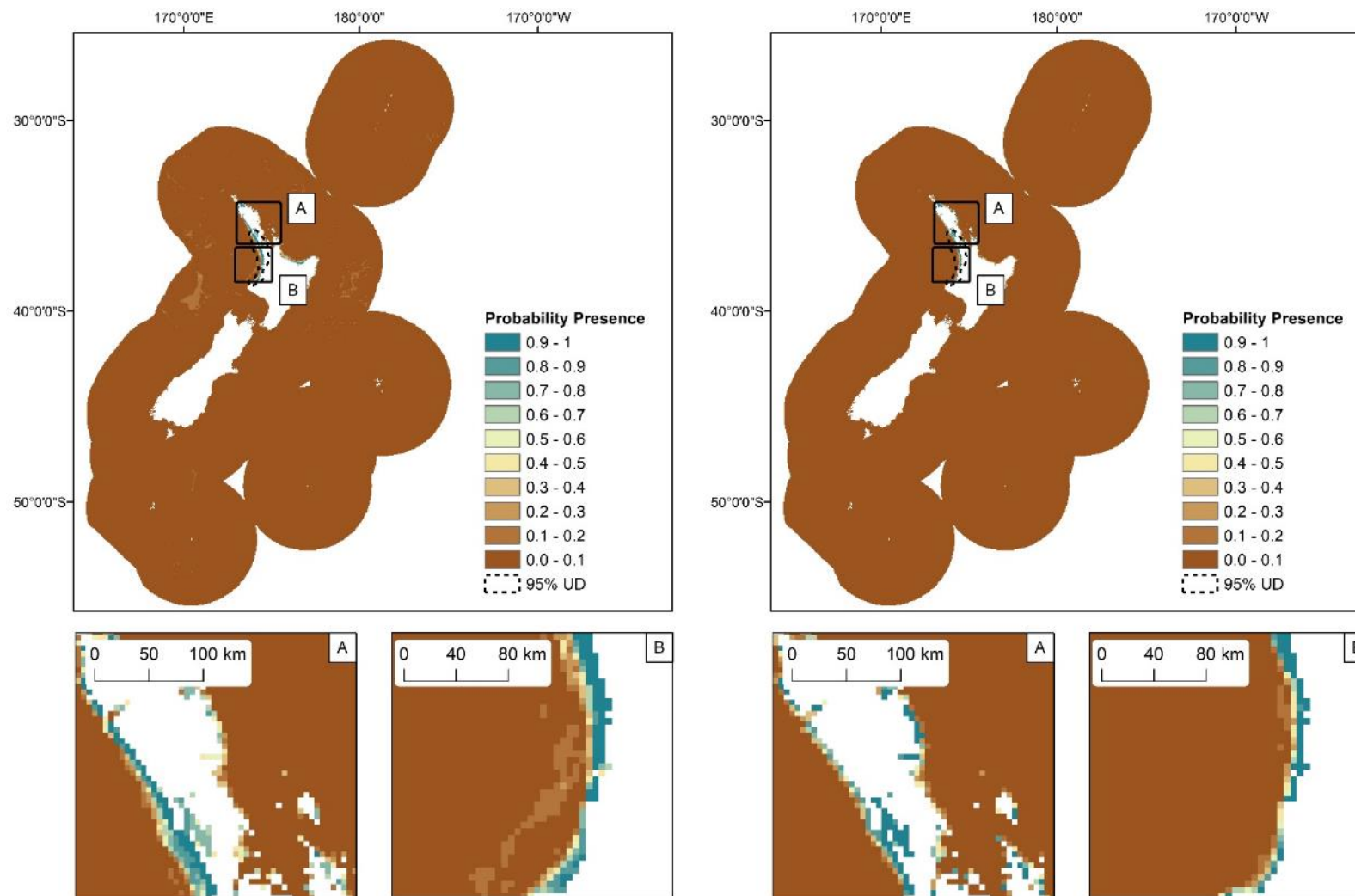


Figure 164: Seasonal predicted probability presence of Māui dolphin (*Cephalorhynchus hectori māui*) modelled using bootstrapped BRTs fitted with Winter (May - Oct, n = 169) (left) and Summer (Nov - Apr, n = 654) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island and Hauraki Gulf; B) West coast of North Island.

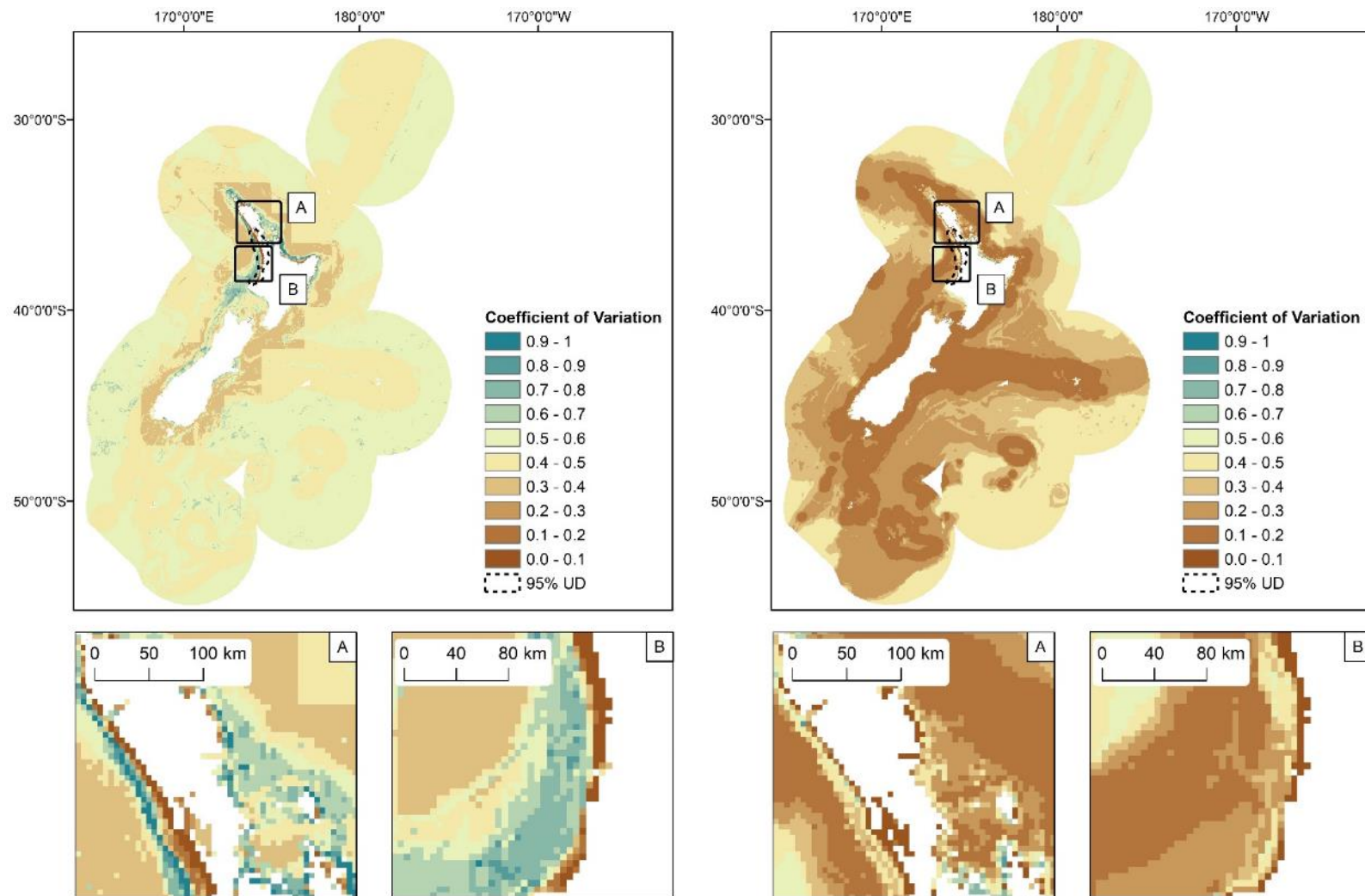


Figure 165: Seasonal uncertainty estimates (coefficient of variation, CV) of Māui dolphin (*Cephalorhynchus hectori māui*) probability presence in the New Zealand EEZ modelled using bootstrapped BRTs fitted with Winter (May - Oct) (left) and Summer (Nov - Apr) presence/relative absence sightings records (right). The predicted 95% utilisation distribution is shown as dashed line. Inset maps: A) North North Island and Hauraki Gulf; B) West coast of North Island.