



Predicting the distribution and relative abundance of fishes on shallow subtidal reefs around New Zealand

**NIWA Client Report: WLG2008-9
April 2008**

NIWA Project: DOC08304

Predicting the distribution and relative abundance of fishes on shallow subtidal reefs around New Zealand

Author
Adam N. H. Smith

Prepared for
Department of Conservation

NIWA Client Report: WLG2008-9
April 2008

NIWA Project: DOC08304

National Institute of Water & Atmospheric Research Ltd
301 Evans Bay Parade, Greta Point, Wellington
Private Bag 14901, Kilbirnie, Wellington, New Zealand
Phone +64-4-386 0300, Fax +64-4-386 0574
www.niwa.co.nz

Contents

Executive Summary	iv
1. Introduction	1
2. Methods	2
2.1 Datasets	2
2.1.1 Diver survey of the abundance of reef fishes	2
2.1.2 Predictor variables	4
2.1.2.1. Environmental predictors	4
2.1.2.2. Geographic predictors	4
2.1.2.3. Dive-specific predictors	6
2.2 Statistical analysis	6
2.2.1 Fitting boosted regression tree models	6
2.2.2 Grooming models	7
2.2.3 Weighting of sites	7
2.2.4 Predictions	8
2.2.5 Scaling the predictions	9
2.3 Uncertainty	10
2.3.1 Bootstrapped confidence intervals	10
2.3.2 Coverage of the environmental space by samples	10
3. Results	10
3.1 Model output	10
3.2 Influences of the predictor variables	11
3.3 Uncertainty	11
4. Discussion	12
4.1 Predicted distributions of abundance	12
4.2 Measures of uncertainty	17
4.3 Limitations and assumptions	17
4.4 Potential uses of this work	18
5. Acknowledgements	19
6. References	19
7. Appendix I – List of data outputs provided	22
8. Appendix II – List of species modelled	23
9. Appendix III – Maps of predicted abundance of reef fishes	25

Reviewed by:

Approved for release by:

Dr Alison Macdiarmid

Dr Don Robertson

Executive Summary

The New Zealand government requires nationally consistent knowledge of the patterns of our marine flora and fauna to inform the design of a network of Marine Protected Areas (MPAs) around the country. To provide such knowledge for a culturally and ecologically important part of our marine fauna, this study used relatively recent statistical methods to model the distribution and abundance of 72 species of rocky reef fish, using a suite of environmental variables. These models were then used to predict and map the reef fish fauna across all known coastal reefs around the country. The predicted distributions have a range of potential applications. First, they provide an important resource for the management of coastal biodiversity and, in particular, for assessing areas of high priority for protection. In addition, they contain a wealth of ecological information and are a useful platform on which to base future research into the environmental drivers of the distributions of these species.

1. Introduction

The New Zealand government has committed to putting in place a network of Marine Protected Areas (MPAs) that cover 10% of the marine environment by the year 2010 (Anonymous 2000). Part of the requirements of this network is that the full range of biodiversity and habitats is protected. Accordingly, the Department of Conservation is currently gathering data on the distributions of marine species to inform this process.

Shallow, coastal rocky reefs are some of the most familiar marine ecosystems to New Zealanders, largely due to their proximity to human settlements and the variety of activities that they offer. Likewise, reef fishes are some of the most conspicuous and recognisable marine species, and have undeniable ecological, social and economic significance. Coastal reefs and reef fishes must therefore be taken into consideration in the planning of networks of MPAs. Indeed, the majority of our marine reserves at present are centred upon shallow coastal reefs (Enderby & Enderby 2006). To date, research into the distributional patterns of reef fishes has focused on broad spatial scales, with detailed accounts published comparing assemblages of each bioregion (Francis 1996; Paulin & Roberts 1992, 1993). However, MPA planning requires biological information at much finer resolution than this, but nationally consistent information of this sort has not yet been obtainable for reef fish.

In fact, biological surveys that have (a) a fine enough spatial resolution and (b) broad enough spatial extent (i.e. national) for use in planning of MPAs are virtually impossible to effectively carry out. To ameliorate this, existing data from biological surveys can be interpolated to areas where knowledge is lacking by using environmental information which is available at fine resolution and broad scales. The relationships between species distributions and environment can be examined using statistical models, and then predictions can be made for areas where observational data are lacking. A variety of modelling techniques are available for this purpose. For example, Leathwick, Francis et al. (2006b) recently modelled the distributions of demersal fishes across New Zealand's EEZ using Boosted Regression Trees (BRTs, De'ath 2007; Friedman 2001; Hastie et al. 2001). The outputs of these models were then analysed using reserve selection software, and areas that would provide the most protection to the biodiversity of demersal fishes were identified (Leathwick et al. 2006c).

In this study, the BRT modelling approach was applied to a dataset of relative abundance of rocky reef fishes from diver surveys around shallow, subtidal reefs of New Zealand. The aim here was to provide fine-scale, nationally consistent predictions of the distribution and abundance of reef fishes, so that robust spatial

information about these culturally and ecologically important species can be integrated into the management of coastal biodiversity and, in particular, the planning of a network of MPAs.

2. Methods

2.1 Datasets

2.1.1 Diver survey of the abundance of reef fishes

The biological data are observations from SCUBA dives around the coast of New Zealand over a period of 18 years from November 1986 to December 2004. The majority of the data used in this study were collected by C. A. J. Duffy of the Department of Conservation, and a small number were collected by the author. The total number of sites was 467. The median time spent per dive was 46 min, and the median maximum depth was 17 m. The locations of the sites are shown in Figure 1.

At each site, a thorough search for fish was undertaken. The relative abundance of each species of fish observed was recorded on a scale of zero to four (Table 1). This scale roughly represents orders of magnitude of abundance. Although developed independently, researchers in the Caribbean have been using this same technique, including the abundance scale, and they called it the Roving Diver Technique (Schmitt et al. 2002; Schmitt & Sullivan 1996; Semmens et al. 2004).

Table 1: Ordinal scale of relative abundance of fish recorded.

Value	Name	No. fish observed
0	absent	0
1	single	1
2	few	2-10
3	many	11-100
4	abundant	> 100

The original dataset contained 212 species. Many species were excluded from the analysis if they are considered to be pelagic, highly cryptic or more associated with soft sediment than reefs. Other species were excluded because they were too rare to be effectively modelled using the techniques used herein – it is difficult to detect

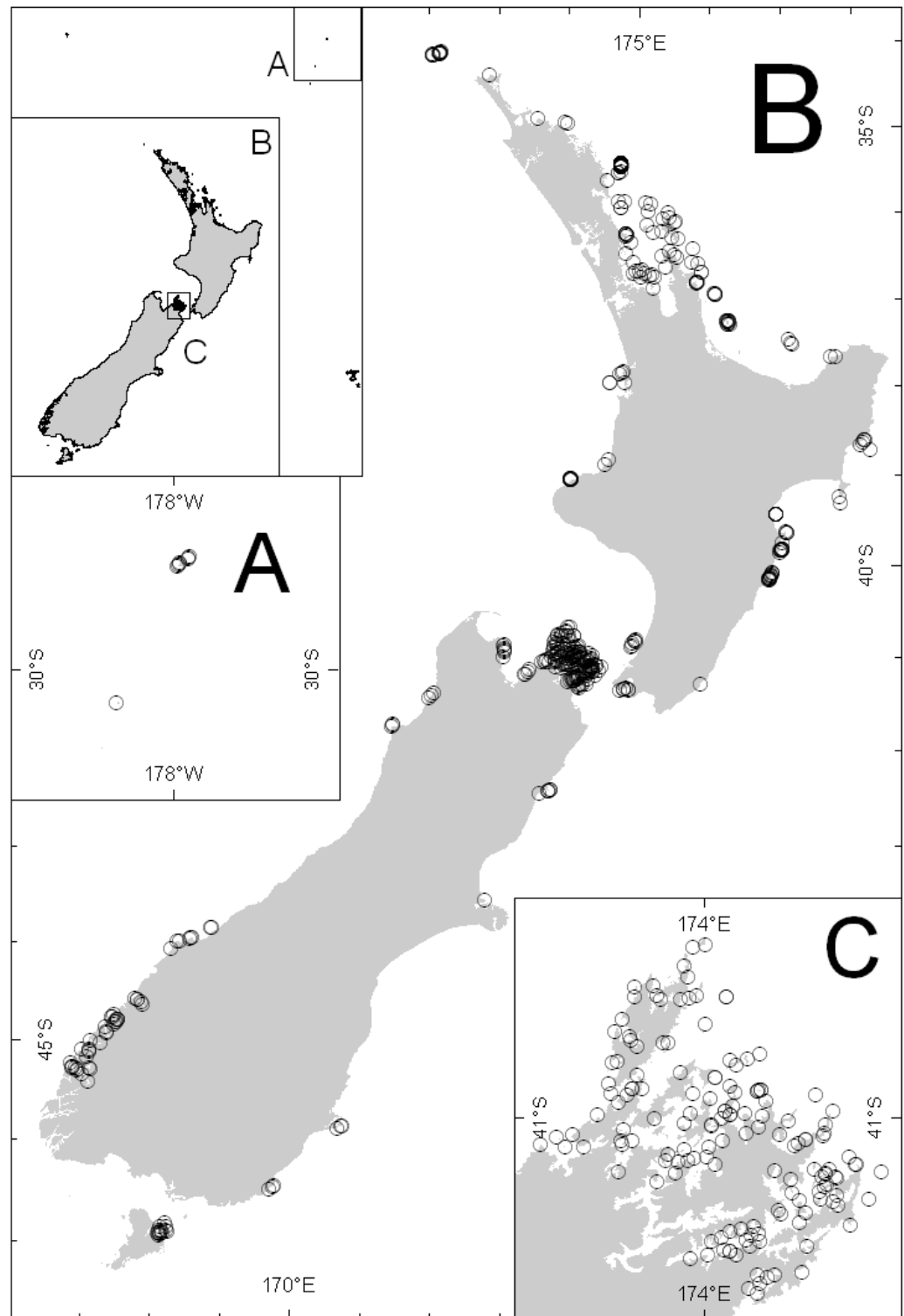


Figure 1 **Locations of 467 sample sites.**

consistent patterns in response to environment from only few sightings. Species that were observed in fewer than 20 sites were not included unless they were deemed of particular interest by the client. The final list contained 72 species.

2.1.2 Predictor variables

A total of 15 variables were available to the models, each falling into one of three categories: environmental, geographic and dive related (Table 2). Some predictor variables were log-transformed, though this was purely to enable easier visualisation of the results, as monotonic transformations have no effect on BRT models or their predictions (Friedman & Meulman 2003).

2.1.2.1. Environmental predictors

The environmental variables were provided as Geographic Information Systems (GIS) raster layers, and were developed for previous projects by NIWA. Most of the layers were developed as part of the Marine Environment Classification (MEC, Snelder et al. 2004; Snelder et al. 2007). Note that the tidal current layer used was not the same as the one developed for the MEC, but a more recent one that has higher resolution (200 m) and accuracy.

Other variables that were available from the MEC were not used, as they were considered too inaccurate in shallow coastal areas (Snelder et al. 2005). Bathymetry was not used because it was only available at the 1 km scale, which is not resolute nor accurate enough for coastal areas (Smith 2006). Freshwater fraction was excluded because of concerns for its accuracy (as it was for the MEC). An orbital velocity layer which predicts the velocity of water at the sea bed, as induced by swell waves, was available from the MEC work, but this layer was not reliable in coastal areas as it did not take into account sheltering or refraction by land. Instead, a geographically derived proxy was used for wave exposure.

2.1.2.2. Geographic predictors

Two geographic variables were the shortest distance to land and average fetch. The shortest distance to land was calculated using ArcGIS. Average fetch is essentially the average distance to land, and was used here as a surrogate for wave exposure. It was calculated following the method developed by Fletcher et al. (2005), where the average distance to land of 36 radial lines radiating from every point, separated by 10 degrees, is taken. In the cases where land is not encountered, the lines were cropped at 10 km. The rationale for this variable is that the more enclosed a site is by land, the level of wave action to which it is likely to be exposed is less. This sort of approach to approximating wave exposure has recently been extended and validated in a study by Burrows et al. (2008).

Table 2 **The variables used in the modelling, and the number of models in which they were used (out of 72 models).**

Type	Name	Explanation	No. models	Average percent contribution
environmental	sstwint ¹	wintertime sea surface temperature	70	33.6
	seabedsal	salinity at the sea bed	43	18.7
	sstanamp ¹	annual amplitude of sea surface temperature	32	18.1
	logdisorgm ^{1,2}	log of dissolved organic matter	32	14.9
	logtidalspeed	log of tidal speed	31	12.0
	logsuspartmat ^{1,2}	log of suspended particulate matter	29	12.2
	sstanom ¹	sea surface temperature anomaly	29	18.8
	logsstgrad ¹	log of sea surface temperature gradient	25	11.7
geographic	chla2 ¹	concentration of chlorophyll a	22	16.2
	avefetch ³	average fetch	48	12.6
dive	dcoast ⁴	distance from coast	7	11.8
	dmax	maximum depth of dive	17	12.9
	dur	duration of dive	12	7.6
	dmin	minimum depth of dive	10	5.9
	vis	visibility of dive	1	5.1

¹ Developed for the MEC (Hadfield et al. 2002; Snelder et al. 2004; Snelder et al. 2007)

² Pinkerton & Richardson (2005)

³ Produced by the author using a program developed by Fletcher et al. (2005)

⁴ Calculated using ArcGIS

2.1.2.3. Dive-specific predictors

The level of effort and the depths of the dive surveys were not standardised throughout the dataset, creating a potential source of bias in the abundance estimates. To control for variability attributable to these factors, some dive-specific variables were included as predictors in the models and given fixed values for the predictions. These variables were the dive duration, visibility and the minimum and maximum depths that were surveyed.

The response functions fitted to these variables were forced to be monotonically positive for duration, visibility and maximum depth, and monotonically negative for minimum depth. The logic behind this forcing is that poor visibility should never increase observed abundance, only decrease it. In contrast, surveying a reef from a minimum depth at 5 m, as opposed to all the way to the surface, should not increase the observed abundance of a species of fish, but may decrease the probability of observing those species that are confined to very shallow water. Allowing these functions to fluctuate freely may allow the model to falsely attribute noise to these variables or, worse, variation that may be better attributed to the environmental variables.

2.2 Statistical analysis

2.2.1 Fitting boosted regression tree models

Independent models were used to model the abundance of each of the 72 species of reef fish. All statistical analyses were undertaken in R (R Development Core Team 2007) using the GBM library (Ridgeway 2006) and code developed by Leathwick et al. (2006a; 2006b). The models were built using the boosted regression tree (BRT) method. This approach uses several (sometimes thousands) of individual regression trees which are combined to form a single ensemble model. The 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. A more complete description of the BRT method can be found elsewhere (De'ath 2007; Friedman 2001; Hastie et al. 2001; Leathwick et al. 2006a; Ridgeway 2006). The specifications used in this study were to fit five trees at a time with a learning rate of 0.002 and a tree depth of 5. A Gaussian error distribution was used, as it produced a better overall fit and residual pattern than the alternatives. Because the scale of abundance is roughly logarithmic,

this is analogous to using a log-normal approach to modelling abundance, which is relatively common in ecology (Legendre & Legendre 1998).

A stepwise, 10-fold, cross-validation procedure was employed to objectively determine the number of trees to be fit in each model, thus reducing the risk of over-fitting. This approach divides the dataset into 10 subsets, which are each withheld in turn while models are fitted to each group of 90% of remaining sites. The holdout deviance is then calculated from the average of the prediction errors of the models to the respective withheld subsets. The final number of trees is given by that which minimises the holdout deviance. Goodness of fit statistics were calculated from the cross-validation routine, by taking the mean and standard error of the correlation between the observed and predicted values for the holdout sites. See Hastie et al. (2001) for a more detailed description of the cross-validation method.

2.2.2 Grooming models

At the first stage, the models were fitted using all available predictor variables. Although the cross-validation process goes some way to ensure that the models are parsimonious in terms of the number of trees fitted to the data, it is acknowledged that over-fitting can also occur by including more predictor variables than are necessary. To ameliorate this, the global models (those with all predictor variables included) for each species in turn were subjected to a simplification process wherein variables were removed from the models, and then the final models were created by refitting with the reduced variable set. Although the simplification process was essentially subjective, in that it was not done automatically, it was informed by some objective criteria. First, the relative contributions of each variable, in terms of deviance explained, was examined. Second, a procedure was used where the lowest contributing variables were sequentially removed from the model, then the model was refit and the change in deviance explained that resulted from removing a variable was examined.

2.2.3 Weighting of sites

The geographic placement of the dive surveys was neither random nor representative. In fact, their placement was highly skewed, with many more sites occurring in areas where the principal collector had done intensive surveys (e.g. the Marlborough Sounds and the Poor Knights Islands) (Figure 1). To avoid these areas having a disproportionately high influence on the models, the sites were given a weighting that reflected the prevalence of sites with its environmental characteristics. This meant that sites with environment types that were poorly represented in the samples were

weighted higher, and those with environment types that are over-represented were down-weighted.

To achieve this, a BRT model was used to calculate the probability of a sampling site being present in parts of the environmental space into which predictions are made. For input to this model, the predictive environmental space of interest was represented by 1000 points that were randomly generated from within the predictive domain (produced using Hawth's Tools for ArcGIS, Beyer 2004). Each of the random points was assigned a zero and each sampling site was assigned a one, and this variable was used to produce a binary response variable that was modelled using the environmental variables. The fitted values were transformed (by the log of the inverse, i.e. $\log(1/x)$, where x is the fitted value) and used as the weighting in the predictive models of the abundances of reef fishes. Only the environmental predictors were used in this model, not the geographic or dive-specific predictors.

The majority of the randomly allocated "zero" points were located immediately adjacent to the coast. This was done by selecting them according to a spatially-explicit probability distribution that decreased with distance from the coast. Although, for the final models, predictions were made beyond coastal areas in the BRT models, the coast is where the most accuracy is required because this is where the majority of sampling sites were collected. As a result of this, it can be said that although the predictive domain includes offshore areas, the models are more strongly weighted towards, and probably more accurate for, coastal areas.

2.2.4 Predictions

An area for which predictions were made was delineated, herein called the predictive domain. The latitudinal extent of the predictions was from the Kermadec and Norfolk Islands in the north to the Snares Islands in the south. Although no surveys were conducted at the Chatham Islands, predictions were made for this remote island group. However, these predictions should be treated with caution as it is known that some mainland species of reef fish have not been recorded at this island group.

The predictive domain was first produced as a 1 km grid, in which a pixel was included if it satisfied at least one of the following conditions: (a) it was within the 50 m depth contour, (b) within 1 km from the shore or (c) within 1 km of a sample site. This grid was then overlain with a shape file that showed the positions of known subtidal reefs around New Zealand, and those grid cells that contained no reef were removed. This predictive domain was then converted into a 1 km grid of points. Values for the environmental and geographic variables were extracted for each point,

and then the BRT models were used to make predictions for each point according to the environmental conditions.

For the dive variables, fixed values were assigned to the predictive points. For duration and visibility, the median values from the surveys were assigned, namely 46 min and 7 m, respectively. For depth, no bathymetric information was available that was sufficiently accurate, extensive or finely scaled to form the basis for predictions. Instead, the maximum depth was fixed to 30 m and the minimum depth was fixed to 0 m for points adjacent to the coast (i.e. within 1 km) and 10 m for offshore points.

Eight of the predicted species distributions in earlier modelling showed presences far southwards of their known ranges (as given by Francis 2001). These species tended to have a very low number of presences in the dataset. This problem was managed by restricting both the data that were used to produce the model, and the predictions, to northwards of specific latitudes. These species, with the latitudes in parentheses, are *Anampses elegans* (38 degrees south), *Atypichthys latus* (38), *Bodianus unimaculatus* (40), *Girella tricuspidata* (40), *Nemadactylus douglasii* (42), *Notolabrus inscriptus* (39), *Optivus elongatus* (43) and *Trachypoma macracanthus* (40).

2.2.5 Scaling the predictions

Because a Gaussian error distribution was used and the fact that predictions from the BRT method are the average taken from many models, the output from these models was on a continuous scale, rather than the ordinal one in the input data. Further, the results for many species were poorly scaled, so that the predictions at the lower end of the scale were overestimated and those at the upper end were underestimated.

To correct these issues, a second step was used to rescale the predictions to the original abundance scale of the input data. This was done by fitting a single classification tree (using the *rpart* library in R, see Breiman et al. 1984) to the observed values using the predicted values. The predictions from the entire domain were then rescaled using this model and rounded to one decimal place. This returned an ordinal response that contained values on the same scale as the original abundance data.

2.3 Uncertainty

2.3.1 Bootstrapped confidence intervals

To obtain estimates of certainty for each prediction, bootstrapped confidence intervals were produced (Manly 1997). This was done by refitting the model to each of 500 bootstrap samples of the original data and making predictions (including scaling, section 2.2.5) for each species. The 0.025 and 0.975 quantiles were taken for every predicted point to provide 95% confidence intervals.

2.3.2 Coverage of the environmental space by samples

The degree of confidence in the predictions given by the model in different areas should be guided by whether the environmental conditions found there were adequately sampled. To quantify this, a similar approach to that of the site weights was taken. A sample of 50000 random values were taken from the environmental space and assigned a zero, and these were combined with the true samples to which a one was assigned. A BRT model was then used to model the probability of a site occurring in the environmental space, using the Bernoulli error distribution. Five-hundred trees with an interaction depth of two were used, so that only pair-wise combinations of the environmental variables were regarded. Predictions were then made for the predictive domain using this model, generating values between zero and one, according to how well each predictive site was represented by the samples.

3. Results

3.1 Model output

Predictions were initially made for 52110 points, but this was reduced to 9605 when only the areas with rocky reef were selected. The predicted distribution and abundance for the 72 species on known reefs are depicted in Appendix III. The predictions for reef-only areas and full domain were both provided as geographic grid files. These, and other relevant data layers and maps, are provided on a data CD (see Appendix I).

The stepwise routine fitted between 675 and 9110 trees to the models, and took between 48 seconds and 46 minutes of computation time (medians 1698 and 4 min, respectively). As assessed from the cross-validation routine, the models were able to

explain between 28 (*Notoclinops caerulepunctus*) and 93 (*Odax cyanoallix*) percent of the deviance in species abundances, with a median of 64 percent (Table 3).

The most widespread species were predicted to be *Forsterygion malcolmi* (98% of predicted reef sites), *Caesioperca lepidoptera* (95%), *Forsterygion varium* (93%), *Notolabrus fucicola* (91%), *Lotella rhacinus* (90%), *Hypoplectrodes huntii* (88%), *Pseudolabrus miles* (87%), *Parika scaber* (85%), *Parapercis colias* (84%) and *Ruanoho whero* (82%). These proportions can be compared with the number of sample sites at which these species were observed (Appendix II).

3.2 Influences of the predictor variables

Appendix IV shows the fitted functions for each predictor variable for each species. The selection process produced models with a median of 5 variables. The most consistently important variable for explaining variation in abundance was sea surface temperature, which was selected in 70 models. It was the highest contributing variable in 30 of these models, with an average of 33% contribution in those models in which it was selected (Table 2). For many species with exclusively northern distributions, the shape of the response to sea surface temperature forms a threshold, from which minimum temperature requirements for these species can be deduced. The only two species for which sea surface temperature was not selected were *Forsterygion lapillum* and *Trachurus novaezelandiae*. The next most important variables were average fetch and salinity of the seabed.

The dive-specific variables were seldom included in the models, mostly due to their poor predictive power. However, some species did show responses to the dive variables, and maximum depth of dive in particular. Two species, namely *Caesioperca lepidoptera* and *Forsterygion flavonigrum*, had maximum depth as their best predictor, and both are often seen at deeper depths.

3.3 Uncertainty

The bootstrapped 95% confidence intervals for the predictions for each species were produced and are available as GIS grid files. They are not presented here because of the additional space required to include them in this already voluminous report.

The modelling of the coverage of the environmental space by the samples produced a spatially explicit layer indicating the areas where we are extrapolating beyond the environmental characteristics of the input data (Figure 2). Poorly covered areas

included the inner Hauraki Gulf, the east and south coasts of the South Island, most of Stewart Island (with the exception of Paterson Inlet) and most of the Chatham Islands. Areas where coverage was good included the rest of north-eastern North Island, East Cape, Taranaki, southern North Island, the outer Marlborough Sounds and Fiordland.

Chlorophyll *a* was the highest contributing factor to distinguishing the sampled and non-sampled parts of the environmental space, with most surveyed sites having values of between one and two (Figure 3). For other variables, surveyed sites appear to be in areas with high gradients of sea surface temperature, moderate levels of sea surface temperature, tidal speed, suspended particulate matter, dissolved organic matter, salinity, anomaly and annual amplitude of sea surface temperature, and low fetch.

4. Discussion

4.1 Predicted distributions of abundance

This study has predicted patterns of distribution and abundance for 72 reef fishes around New Zealand (Appendix III). The spatial resolution of the predictions is a 1 km grid, the extent of which includes all known shallow (< 50 m) reefs from the Kermadec Islands in the north to the Chatham Islands in the east and the Snares Islands in the south. The only shallow areas in New Zealand's exclusive economic zone (EEZ) that were not included were those around the subantarctic islands, which are known to be relatively depauperate in their fauna of reef fishes (Kingsford et al. 1989). The fact that, on average, 64% of the variation in reef fish abundance was explained by the models (according to cross-validation statistics) is very encouraging, and implies that we may place a reasonable degree of confidence in these estimates (Table 3).

To date, only broad scale information on the distributions of these species had been published (Francis 1996; Paulin & Roberts 1992, 1993). This work presents the first attempt at modelling reef fishes at such fine spatial resolution and broad extent. This work follows that of Leathwick et al. and (2006b) who first used this approach with marine species in New Zealand, predicting distributions and abundance of demersal fishes across the entire EEZ.

Table 3 **Model performance from boosted regression tree analysis of 72 rocky reef fishes.**

Species	No. predictor variables	No. trees fitted	Goodness of fit (from cross-validation)	
			Mean deviance explained	Standard error
<i>Aldrichetta forsteri</i>	8	820	0.39	0.08
<i>Amphichaetodon howensis</i>	3	905	0.79	0.05
<i>Anampses elegans</i>	2	9110	0.88	0.04
<i>Aplodactylus arctidens</i>	8	1825	0.49	0.03
<i>Aplodactylus etheridgii</i>	2	5165	0.77	0.11
<i>Atypichthys latus</i>	4	2845	0.55	0.12
<i>Bodianus unimaculatus</i>	5	2045	0.84	0.04
<i>Caesioperca lepidoptera</i>	11	6205	0.67	0.03
<i>Canthigaster callisterna</i>	3	1295	0.67	0.04
<i>Caprodon longimanus</i>	3	1965	0.87	0.03
<i>Centroberyx affinis</i>	4	5870	0.67	0.06
<i>Cheilodactylus spectabilis</i>	8	2310	0.66	0.02
<i>Chironemus marmoratus</i>	7	990	0.52	0.03
<i>Chromis dispilus</i>	5	2495	0.93	0.01
<i>Conger verreauxi</i>	10	1035	0.30	0.05
<i>Coris sandageri</i>	2	1545	0.87	0.02
<i>Decapterus koheru</i>	5	1215	0.65	0.06
<i>Epinephelus daemeli</i>	2	4505	0.90	0.02
<i>Forsterygion lapillum</i>	6	1835	0.71	0.04
<i>Forsterygion malcolmi</i>	6	2170	0.62	0.03
<i>Forsterygion varium</i>	12	1440	0.62	0.03
<i>Forsterygion flavonigrum</i>	9	2585	0.65	0.03
<i>Girella cyanea</i>	2	2705	0.65	0.09
<i>Girella tricuspidata</i>	6	1080	0.51	0.08
<i>Grahamina gymnota</i>	8	835	0.34	0.08
<i>Gymnothorax prasinus</i>	4	945	0.59	0.06
<i>Helicolenus percoides</i>	8	2640	0.62	0.07
<i>Hypoplectrodes huntii</i>	9	2330	0.56	0.05
<i>Hypoplectrodes sp.B</i>	3	1720	0.74	0.09
<i>Karalepis stewarti</i>	10	1090	0.41	0.05
<i>Kyphosus sydneyanus</i>	6	690	0.41	0.06
<i>Latridopsis ciliaris</i>	8	1870	0.61	0.03
<i>Latris lineata</i>	3	1720	0.67	0.11
<i>Lotella rhacinus</i>	10	1160	0.35	0.06
<i>Mendosoma lineatum</i>	6	2145	0.57	0.07
<i>Nemadactylus douglasii</i>	5	1550	0.79	0.03
<i>Nemadactylus macropterus</i>	6	1335	0.48	0.05
<i>Notoclinops caerulepunctus</i>	8	700	0.28	0.05
<i>Notolabrus celidotus</i>	10	3420	0.78	0.02
<i>Notolabrus cinctus</i>	5	1390	0.67	0.04
<i>Notolabrus fucicola</i>	11	4215	0.69	0.03

<i>Notolabrus inscriptus</i>	4	2945	0.57	0.11
<i>Notoclinops segmentatus</i>	9	1670	0.55	0.03
<i>Notoclinops yaldwyni</i>	5	1340	0.63	0.05
<i>Obliquichthys maryannae</i>	7	1755	0.62	0.03
<i>Odax cyanoallix</i>	3	5360	0.93	0.04
<i>Odax pullus</i>	12	2110	0.56	0.03
<i>Optivus elongatus</i>	5	1365	0.56	0.04
<i>Pagrus auratus</i>	5	1605	0.83	0.03
<i>Parma alboscaphularis</i>	4	1240	0.78	0.04
<i>Parapercis colias</i>	7	2575	0.77	0.02
<i>Parablennius laticlavus</i>	5	1770	0.78	0.03
<i>Parika scaber</i>	8	3565	0.64	0.03
<i>Paratrachichthys trailli</i>	7	1020	0.38	0.08
<i>Pempheris adspersa</i>	7	2975	0.80	0.02
<i>Plagiotremus tapeinosoma</i>	2	905	0.72	0.05
<i>Pseudophycis barbata</i>	5	1340	0.36	0.07
<i>Pseudocaranx dentex</i>	4	745	0.47	0.06
<i>Pseudolabrus luculentus</i>	2	1245	0.92	0.02
<i>Pseudolabrus miles</i>	7	2950	0.71	0.02
<i>Ruanoho whero</i>	7	1670	0.55	0.02
<i>Scorpaena cardinalis</i>	2	2260	0.85	0.03
<i>Scorpis lineolatus</i>	3	1355	0.58	0.04
<i>Scorpaena papillosus</i>	5	1370	0.57	0.05
<i>Scorpis violaceus</i>	4	2010	0.78	0.03
<i>Seriola lalandi</i>	6	990	0.56	0.05
<i>Suezichthys aylingi</i>	3	1560	0.66	0.05
<i>Trachypoma macracanthus</i>	2	3565	0.67	0.10
<i>Trachurus novaezelandiae</i>	4	840	0.32	0.05
<i>Upeneichthys lineatus</i>	5	2010	0.67	0.04
<i>Zanclistius elevatus</i>	2	720	0.41	0.10
<i>Zeus faber</i>	4	675	0.37	0.08

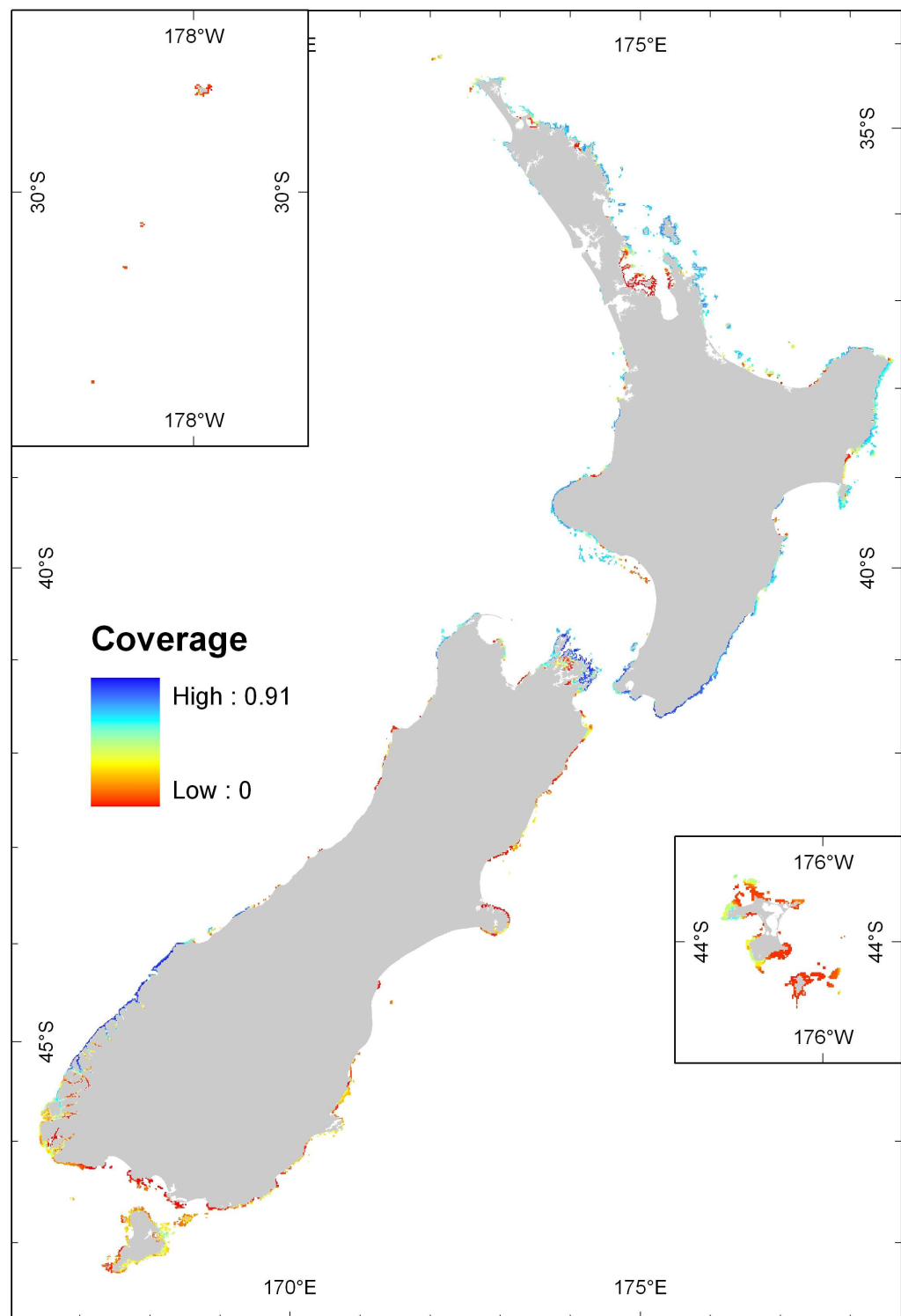


Figure 2 A map showing the predictive domain and how well the environmental conditions of each pixel was represented by the sampling sites. “Coverage” may be interpreted as the probability a sampling site occurring there given its environmental conditions, and provides an index of how much confidence can be placed in the predictions.

Coverage of environmental space

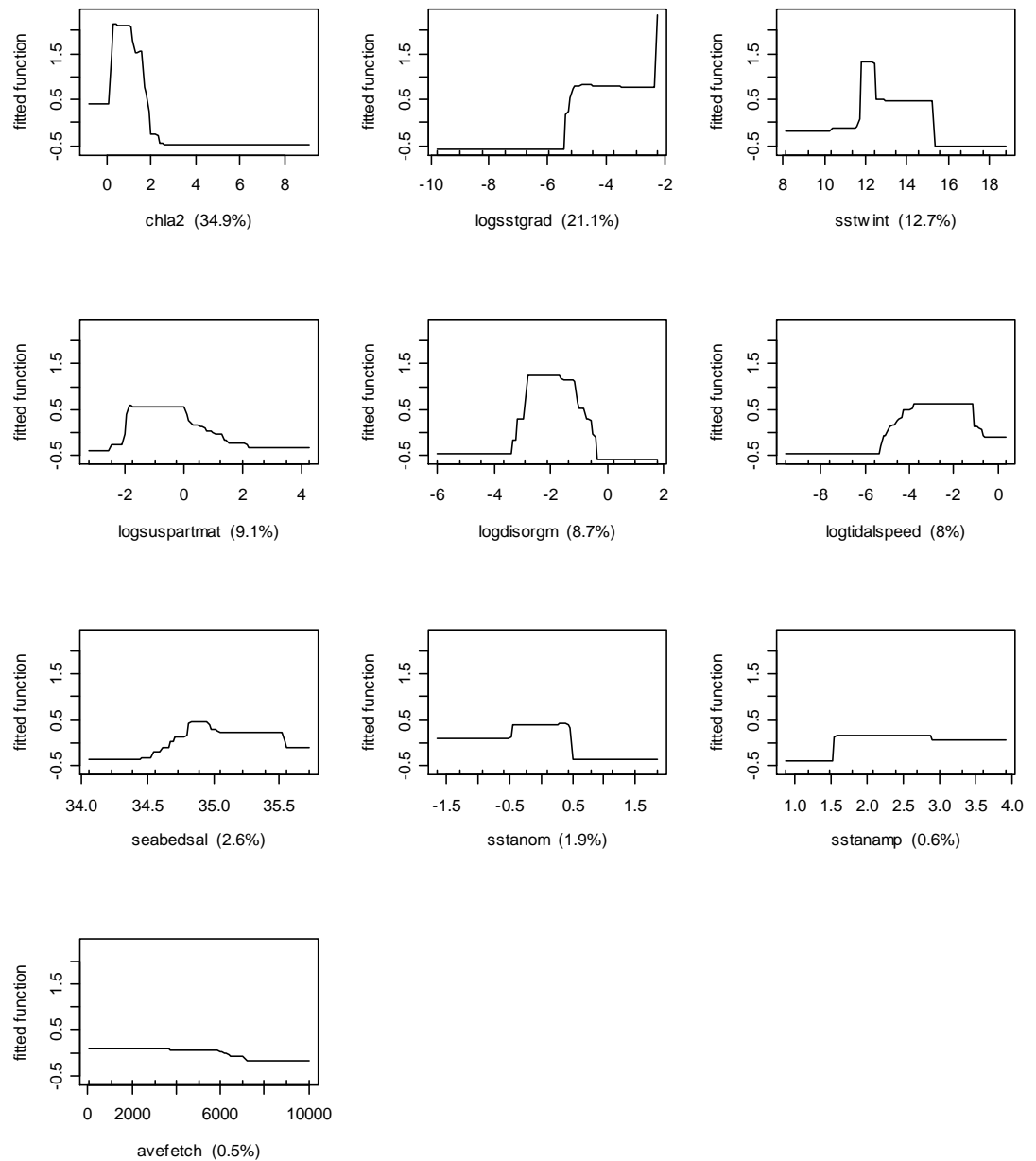


Figure 3 The contributions of and functions fitted to each environmental variable in the model determining the coverage of the environmental space by the survey sites.

4.2 Measures of uncertainty

It is useful to provide quantitative measures of the uncertainty of the predictions so that end users can know the degree of confidence that can be placed in them. Here, a number of measures were provided, pertaining to a variety of levels. First, the measure of deviance explained can be taken as an overall measure of the accuracy of each individual species map (Table 3). This is not spatially explicit, and is a single number that may be used as some sort of weighting for the species when combined in a single analysis, as more confidence can be placed in the models that have higher accuracy. These estimates of model fit are based on cross-validation statistics. Although this method is more robust than traditional measures (Hastie et al. 2001), it does not indicate how robust the models may be beyond the range of the biological dataset, chiefly because some areas into which we are predicting may have environmental characteristics that are not represented in the sites of the survey. However, a spatially explicit index of the coverage of the environmental character of each pixel by the survey was provided here (Figure 2). This has relevance across all 72 predicted distributions. Lastly, bootstrapped confidence intervals for each species at each predicted site were calculated to give within-species, spatially explicit measures of uncertainty. All of these measures of uncertainty can be incorporated into the use of the data in management applications, thanks to recent developments in reserve planning software (Moilanen et al. 2006).

4.3 Limitations and assumptions

It is important to note the limitations of these predictions imposed by the input data and the methods. First, they are not intended to be a definitive account of where each species can be found. The predictions are subject to the restrictions imposed by the survey method, and do not take into account factors such as diver positive or negative behaviour by fish (see Cole 1994), or variation in detection probabilities (see MacKenzie et al. 2003; MacKenzie et al. 2002). Rather, these layers represent predictions of the fish assemblages that might be seen on a typical dive at each of these locations, which we can fairly safely assume is correlated with true local abundance.

Also, the predictions were limited to diveable depths. The predictive domain here extended to the 50 m isobath but, because only 10% of the survey dives were deeper than 30 m, predictions beyond this depth should be used with caution.

The predictions are made primarily on the basis of the suitability of the environment for each species. The models contain no spatial information of any kind and disregards biogeographic factors such as dispersal distances and barriers. Therefore, caution should be taken when using the predictions on offshore island groups, such as the Snares and Chatham Islands. On the other hand, if a species does not in fact occur at a particular modelled location it highlights that factors other than the environmental variables used in the model are important in determining the species distribution.

A spatial resolution of 1 km applies to these predictions, because this was the resolution of most of the environmental variables. Although this is a far finer scale than that of any previous work on these species, it should be noted that variation in abundances will exist at smaller spatial scales, mainly due to variations in habitats between reefs within a pixel, or even within a single contiguous reef. The predictions are for known reef habitat, but they are based on assumption that suitable habitat for each species is available on each of these reefs.

Finally, it should be acknowledged that wave exposure is an important environmental factor that undoubtedly affects reef fish community structure around coasts. However, an appropriate GIS layer for this was not available for the entire coastline at the scale and accuracy that was required by this work, as the one used for the MEC is not considered accurate for coasts (Hadfield et al. 2002). In its place, a geographically derived variable, fetch, was used as a surrogate, but this does not take into account regional and directional variation in wave energy. Although this geographic approach has been shown to be a good proxy for wave exposure (Burrows et al. 2008), the inclusion of a more accurate and robust measure of this factor would probably improve the accuracy of the models.

4.4 Potential utility of these results

Improved knowledge of the spatial distributions of reef fishes will have a variety of potential uses. Its primary purpose is to inform the management of New Zealand's coastal biodiversity and, in particular, aid with the protection of biodiversity through the planning of MPAs and evaluating susceptibility to biosecurity threats. As detailed in another report, this information will be combined with predictions of coastal communities of demersal fishes and analysed using the reserve planning software, Zonation (Moilanen 2007; Moilanen et al. 2005).

In addition to management applications, these results contain a wealth of information about the environmental drivers of the distributions and abundance of reef fishes. This includes both the relative importance of different factors and the nature of the responses to these factors. For example, the sea surface temperature responses often show a clear step function, implying minimum thresholds of tolerance for many species. While conclusive statements about mechanistic processes cannot be made on the basis of this work, ecological hypotheses may be derived from the outputs of the models (Appendix IV) and then tested in future research.

5. Acknowledgements

This project was funded by Department of Conservation (DOC contract 3994; NIWA project DOC08304). NIWA thanks Clinton Duffy (DOC) for extensive advice, and providing the reef fish dataset and GIS layer of the locations of shallow reefs, John Leathwick for providing R scripts and statistical advice, David Plew and James Sturman for the GIS layer of tidal velocity, Ross Pickard and Eduardo Villouta for providing the fetch analysis program and assistance with it, Malcolm Francis and Clinton Duffy for comments on preliminary predictions, and James Sturman and Arne Pallentin for assistance with GIS. Thanks also to Clinton Duffy, Alison MacDiarmid and Don Robertson for providing comments on the report.

6. References

- Anonymous. 2000. The New Zealand Biodiversity Strategy - Our chance to turn the tide. Page 146. Department of Conservation and Ministry for the Environment, Wellington.
- Beyer, H. L. 2004. Hawth's Analysis Tools for ArcGIS, Available at <http://www.spatialecology.com/htools>.
- Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone 1984. Classification and Regression Trees. Wadsworth & Brooks, Belmont, California.
- Burrows, M. T., R. Harvey, and L. Robb. 2008. Wave exposure indices from digital coastlines and the prediction of rocky shore community structure. *Marine Ecology Progress Series* **353**:1-12.
- Cole, R. G. 1994. Abundance, size structure, and diver-oriented behaviour of three large benthic carnivorous fishes in a marine reserve in northeastern New Zealand. *Biological Conservation* **70**:93-99.

- De'ath, G. 2007. Boosted trees for ecological modeling and prediction. *Ecology* **88**:243-251.
- Enderby, J., and T. Enderby 2006. A guide to New Zealand's marine reserves. New Holland Publishers, Auckland.
- Fletcher, D., D. MacKenzie, and E. Villouta. 2005. Modelling skewed data with many zeros: a simple approach combining ordinary and logistic regression. *Environmental and Ecological Statistics* **12**:45-54.
- Francis, M. P. 1996. Geographic distribution of marine reef fishes in the New Zealand region. *New Zealand Journal of Marine and Freshwater Research* **30**:35-55.
- Francis, M. P. 2001. Coastal fishes of New Zealand: an identification guide. Reed Publishing, Auckland, New Zealand.
- Friedman, J. H. 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics* **29**:1189-1232.
- Friedman, J. H., and J. J. Meulman. 2003. Multiple additive regression trees with application in epidemiology. *Statistics in Medicine* **22**:1365-1381.
- Hadfield, M. G., M. J. Uddstrom, D. Goring, R. M. Gorman, M. Wild, S. Stephens, U. Shankar, K. Niven, and T. H. Snelder. 2002. Physical variables for the New Zealand Marine Environment Classification System: development and description of data layers. Unpublished report by the National Institute of Water & Atmospheric Research Ltd, Report No. CHC2002-043.
- Hastie, T., R. Tibshirani, and J. Friedman 2001. The elements of statistical learning: data mining, inference, and prediction. Springer, New York.
- Kingsford, M. J., D. R. Schiel, and C. N. Battershill. 1989. Distribution and abundance of fish in a rocky reef environment at the subantarctic Auckland Islands, New Zealand. *Polar Biology* **9**:179-186.
- Leathwick, J. R., J. Elith, M. P. Francis, T. Hastie, and P. Taylor. 2006a. 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. R., M. P. Francis, and K. Julian. 2006b. Development of a demersal fish community map for New Zealand's Exclusive Economic Zone. Unpublished report by the National Institute of Water and Atmospheric Research, No. HAM2006-062, Hamilton, New Zealand.
- Leathwick, J. R., K. Julian, and M. P. Francis. 2006c. Exploration of the use of reserve planning software to identify potential Marine Protected Areas in New Zealand's Exclusive Economic Zone. Unpublished report by the National Institute of Water and Atmospheric Research, No. HAM2006-064, Hamilton, New Zealand.
- Legendre, P., and L. Legendre 1998. Numerical Ecology. Elsevier Science, Amsterdam.
- MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology* **84**:2200-2207.

- MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* **83**:2248-2255.
- Manly, B. F. J. 1997. *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Chapman & Hall, London.
- Moilanen, A. 2007. Landscape Zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation* **134**:571-579.
- Moilanen, A., A. M. A. Franco, R. I. Early, R. Fox, B. Wintle, and C. D. Thomas. 2005. Prioritizing multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society B: Biological Sciences* **272**:1885-1891.
- Moilanen, A., B. A. Wintle, J. Elith, and M. Burgman. 2006. Uncertainty analysis for regional-scale reserve selection. *Conservation Biology* **20**:1688-1697.
- Paulin, C. D., and C. D. Roberts 1992. *The rockpool fishes of New Zealand*. Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand.
- Paulin, C. D., and C. D. Roberts. 1993. Biogeography of New Zealand rockpool fishes. Pages 191-199 in C. N. Battershill, D. R. Schiel, G. P. Jones, R. G. Creese, and A. B. MacDiarmid, editors. *Proceedings of the Second International Temperate Reef Symposium*, 7-10 January 1992, Auckland, New Zealand. National Institute of Water and Atmospheric Research, Wellington.
- Pinkerton, M. H., and K. R. Richardson. 2005. Case 2 Climatology of New Zealand: Final report. Page 16. Unpublished report by the National Institute of Water and Atmospheric Research, No. WLG2005-049, Wellington, New Zealand.
- R Development Core Team. 2007. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ridgeway, G. 2006. *Generalized Boosted Models: A guide to the gbm package*. gbm library for R.
- Schmitt, E., R. Sluka, and K. Sullivan-Sealey. 2002. Evaluating the use of roving diver and transect surveys to assess the coral reef fish assemblage off southeastern Hispaniola. *Coral Reefs* **21**:216-223.
- Schmitt, E. F., and K. M. Sullivan. 1996. Analysis of a volunteer method for collecting fish presence and abundance data in the Florida Keys. *Bulletin of Marine Science* **59**:404-416.
- Semmens, B. X., E. R. Buhle, A. K. Salomon, and C. V. Pattengill-Semmens. 2004. A hotspot of non-native marine fishes: evidence for the aquarium trade as an invasion pathway. *Marine Ecology Progress Series* **266**:239-244.
- Smith, A. N. H. 2006. *Evaluation of the New Zealand Marine Environment Classification for shallow coastal rocky reef fish communities*. Page 133+ix. Department of Statistics. University of Auckland, Auckland.
- Snelder, T., J. R. Leathwick, K. L. Dey, M. A. Weatherhead, G. D. Fenwick, M. P. Francis, R. Gorman, J. M. Grieve, M. G. Hadfield, J. E. Hewitt, T. M. Hume, K. M.

- Richardson, A. A. Rowden, M. J. Uddstrom, M. Wild, and J. R. Zeldis. 2005. The New Zealand Marine Environment Classification. Page 70. Ministry for the Environment.
- Snelder, T., J. R. Leathwick, K. Image, M. A. Weatherhead, and M. Wild. 2004. The New Zealand Marine Environment Classification. Unpublished report by the National Institute of Water & Atmospheric Research Ltd, Report No. CHC2004-071.
- Snelder, T. H., J. R. Leathwick, K. L. Dey, A. A. Rowden, M. A. Weatherhead, G. D. Fenwick, M. P. Francis, R. M. Gorman, J. M. Grieve, M. G. Hadfield, J. E. Hewitt, K. M. Richardson, M. J. Uddstrom, and J. R. Zeldis. 2007. Development of an ecological marine classification in the New Zealand region. *Environmental Management* **39**:12-29.

7. Appendix I – List of data outputs provided

The following is a list of data outputs provided to the client on a CD. They are provided in both ArcGIS and ASCII grid format, using the same projection as that of the MEC (Mercator, Clarke 1866). Note that producing confidence intervals for the entire predictive domain were not feasible with respect to the amount of time that was required to compute them for that many predictions. In addition to the GIS grids listed below, the CD also contains maps of the layers.

- Rocky reefs only
 - Abundance
 - 72 GIS grids of estimated abundance of reef fishes
 - 72 GIS grids of upper 95% confidence intervals
 - 72 GIS grids of lower 95% confidence intervals
 - Coverage of environmental space
 - 1 GIS grid of coverage of environmental space
 - Species richness
 - 1 GIS grid of estimated species richness of reef fishes
 - 1 GIS grid of upper 95% confidence intervals
 - 1 GIS grid of lower 95% confidence intervals
- Entire predictive domain
 - Abundance
 - 72 GIS grids of estimated abundance of reef fishes
 - Coverage of environmental space
 - 1 GIS grid of coverage of environmental space
 - Species richness
 - 1 GIS grid of estimated species richness of reef fishes

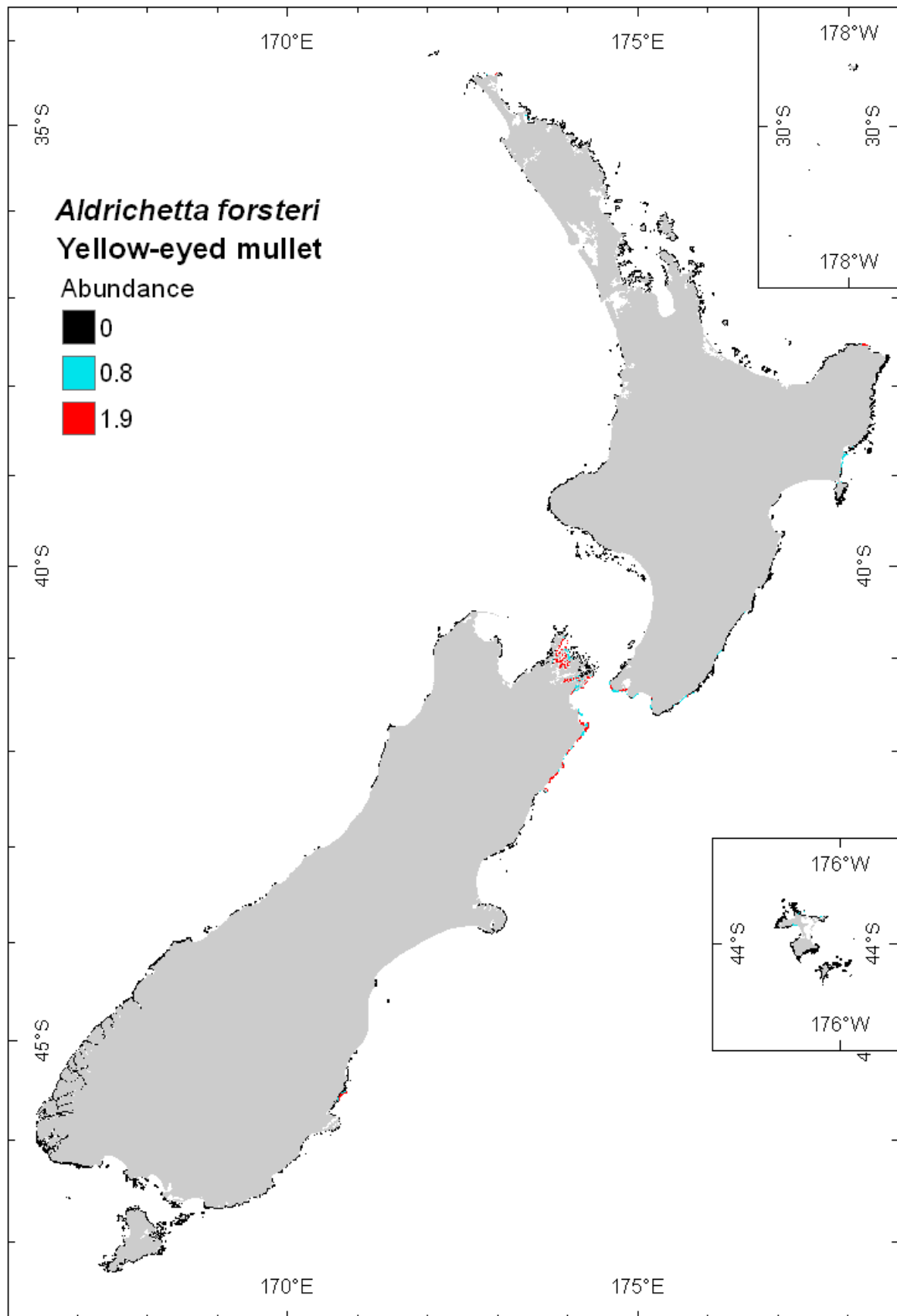
8. Appendix II – List of species modelled

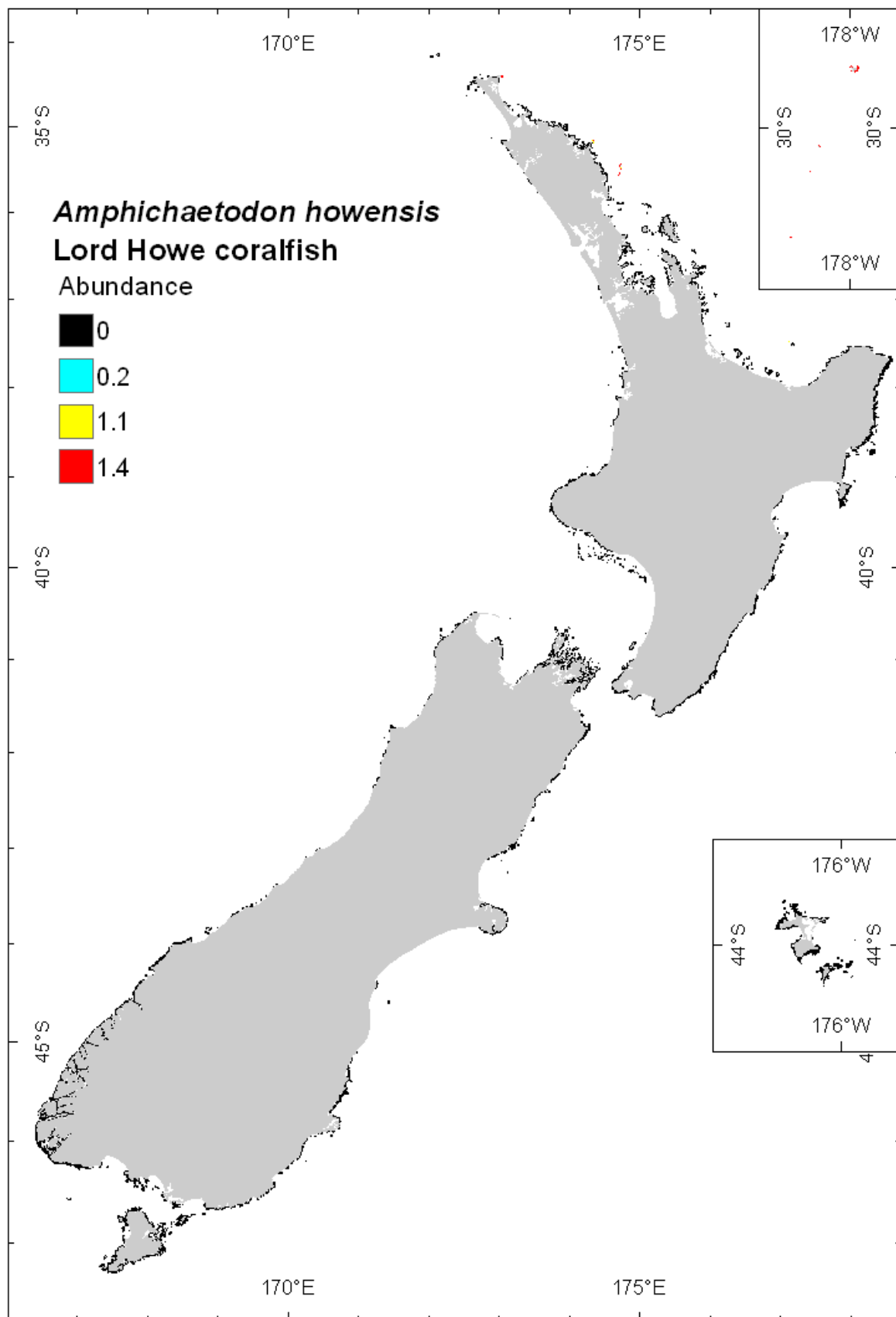
Code	Family	Genus species	Common name	No. sites observed (of 467)
Apl.arc	Aplodactylidae	<i>Aplodactylus arctidens</i>	Marblefish	191
Apl.eth	Aplodactylidae	<i>Aplodactylus etheridgii</i>	Notch-head marblefish	16
Cen.aff	Berycidae	<i>Centroberyx affinis</i>	Golden snapper	29
Par.lat	Blenniidae	<i>Parablennius laticlavus</i>	Crested blenny	97
Pla.tap	Blenniidae	<i>Plagiotremus tapeinosoma</i>	Mimic blenny	33
Dec.koh	Carangidae	<i>Decapterus koheru</i>	Koheru	65
Pse.den	Carangidae	<i>Pseudocaranx dentex</i>	Trevally	65
Ser.lal	Carangidae	<i>Seriola lalandi</i>	Kingfish	95
Tra.sp.	Carangidae	<i>Trachurus novaezelandiae</i>	Jack mackerel	47
Amp.how	Chaetodontidae	<i>Amphichaetodon howensis</i>	Lord Howe coralfish	27
Che.spe	Cheilodactylidae	<i>Cheilodactylus spectabilis</i>	Red moki	212
Nem.dou	Cheilodactylidae	<i>Nemadactylus douglasii</i>	Porae	84
Nem.mac	Cheilodactylidae	<i>Nemadactylus macropterus</i>	Tarakihi	132
Chi.mar	Chironemidae	<i>Chironemus marmoratus</i>	Hiwihwi	69
Con.ver	Congridae	<i>Conger verreauxi</i>	Common conger eel	85
Aty.lat	Kyphosidae	<i>Atypichthys latus</i>	Mado	22
Gir.cya	Kyphosidae	<i>Girella cyanea</i>	Bluefish	22
Gir.tri	Kyphosidae	<i>Girella tricuspidata</i>	Parore	26
Kyp.syd	Kyphosidae	<i>Kyphosus sydneyanus</i>	Silver drummer	23
Sco.lin	Kyphosidae	<i>Scorpis lineolatus</i>	Sweep	182
Sco.vio	Kyphosidae	<i>Scorpis violaceus</i>	Blue maomao	110
Ana.ele	Labridae	<i>Anampses elegans</i>	Elegant wrasse	12
Bod.vul	Labridae	<i>Bodianus unimaculatus</i>	Red pigfish	69
Cor.san	Labridae	<i>Coris sandageri</i>	Sandager's wrasse	74
Not.cel	Labridae	<i>Notolabrus celidotus</i>	Spotty	349
Not.cin	Labridae	<i>Notolabrus cinctus</i>	Girdled wrasse	45
Not.fuc	Labridae	<i>Notolabrus fucicola</i>	Banded wrasse	340
Not.ins	Labridae	<i>Notolabrus inscriptus</i>	Green wrasse	25
Pse.luc	Labridae	<i>Pseudolabrus luculentus</i>	Orange wrasse	61
Pse.mil	Labridae	<i>Pseudolabrus miles</i>	Scarlet wrasse	283
Sue.ayl	Labridae	<i>Suezichthys aylingi</i>	Crimson cleanerfish	46
Lat.cil	Latridae	<i>Latridopsis ciliaris</i>	Blue moki	148
Lat.for	Latridae	<i>Latridopsis forsteri</i>	Copper moki	15
Lat.lin	Latridae	<i>Latris lineata</i>	Trumpeter	21
Men.lin	Latridae	<i>Mendosoma lineatum</i>	Telescopefish	32

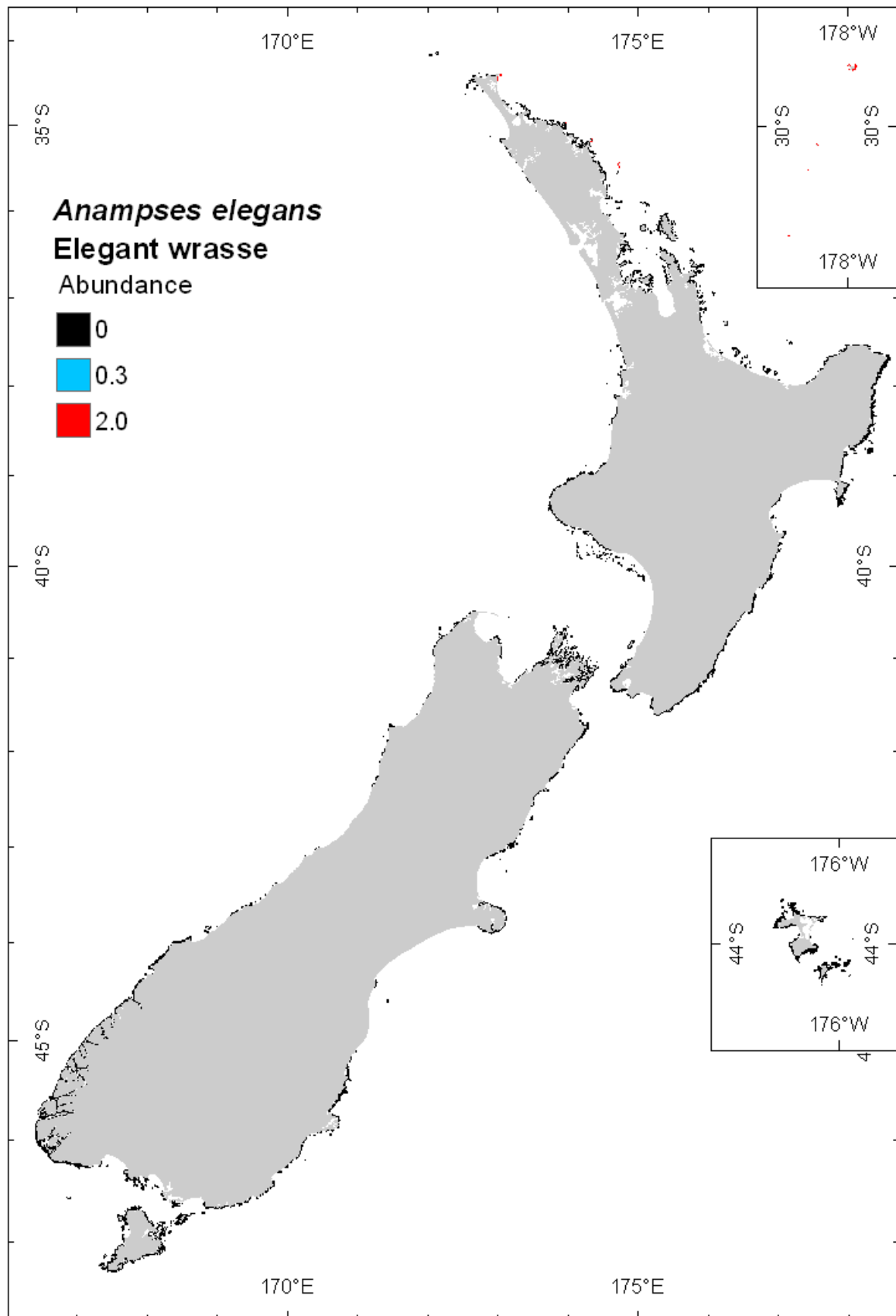
Par.sca	Monacanthidae	<i>Parika scaber</i>	Leatherjacket	244
Lot.rha	Moridae	<i>Lotella rhacinus</i>	Rock cod	144
Pse.bar	Moridae	<i>Pseudophycis barbata</i>	Southern bastard cod	28
Ald.for	Mugilidae	<i>Aldrichetta forsteri</i>	Yellow-eyed mullet	55
Upe.lin	Mullidae	<i>Upeneichthys lineatus</i>	Goatfish	154
Gym.pra	Muraenidae	<i>Gymnothorax prasinus</i>	Yellow moray	44
Oda.cya	Odacidae	<i>Odax cyanoallix</i>	Blue-finned butterfly	15
Oda.pul	Odacidae	<i>Odax pullus</i>	Butterfish	157
Pem.ads	Pempheridae	<i>Pempheris adspersa</i>	Bigeye	83
Zan.ele	Pentacerotidae	<i>Zanclistius elevatus</i>	Long-finned boarfish	13
Par.col	Pinguipedidae	<i>Parapercis colias</i>	Blue cod	275
Chr.dis	Pomacentridae	<i>Chromis dispilus</i>	Demoiselle	133
Par.alb	Pomacentridae	<i>Parma alboscapularis</i>	Black angelfish	69
Hel.per	Scorpaenidae	<i>Helicolenus percoides</i>	Sea perch	63
Sco.car	Scorpaenidae	<i>Scorpaena cardinalis</i>	Northern scorpionfish	44
Sco.pap	Scorpaenidae	<i>Scorpaena papillosus</i>	Dwarf scorpionfish	154
Cae.lep	Serranidae	<i>Caesioperca lepidoptera</i>	Butterfly perch	216
Cap.lon	Serranidae	<i>Caprodon longimanus</i>	Pink maomao	60
Epi.dae	Serranidae	<i>Epinephelus daemeli</i>	Spotted black grouper	19
Hyp.hun	Serranidae	<i>Hypoplectrodes huntii</i>	Red-banded perch	80
Hyp.spB	Serranidae	<i>Hypoplectrodes sp.B</i>	Half-banded perch B	57
Tra.mac	Serranidae	<i>Trachypoma macracanthus</i>	Toadstool grouper	14
Pag.aur	Sparidae	<i>Pagrus auratus</i>	Snapper	85
Can.cal	Tetraodontidae	<i>Canthigaster callisterna</i>	Clown toado	34
Opt.elo	Trachichthyidae	<i>Optivus elongatus</i>	Slender roughy	137
Par.tra	Trachichthyidae	<i>Paratrachichthys trailli</i>	Common roughy	46
For.lap	Tripterygiidae	<i>Forsterygion lapillum</i>	Common triplefin	203
For.mal	Tripterygiidae	<i>Forsterygion malcolmi</i>	Banded triplefin	284
For.var	Tripterygiidae	<i>Forsterygion varium</i>	Variable triplefin	343
Fst fla	Tripterygiidae	<i>Forsterygion flavonigrum</i>	Yellow-black triplefin	258
Gra.gym	Tripterygiidae	<i>Grahamina gymnota</i>	Robust triplefin	23
Kar.ste	Tripterygiidae	<i>Karalepis stewarti</i>	Scaly-headed triplefin	82
Not.cae	Tripterygiidae	<i>Notoclinops caerulepunctus</i>	Blue dot triplefin	67
Not.seg	Tripterygiidae	<i>Notoclinops segmentatus</i>	Blue-eyed triplefin	269
Not.yal	Tripterygiidae	<i>Notoclinops yaldwyni</i>	Yaldwyn's triplefin	86
Obl.mar	Tripterygiidae	<i>Obliquichthys maryannae</i>	Oblique-swimming triplefin	251
Rua.whe	Tripterygiidae	<i>Ruanoho whereo</i>	Spectacled triplefin	250
Zeu.fab	Zeidae	<i>Zeus faber</i>	John dory	27

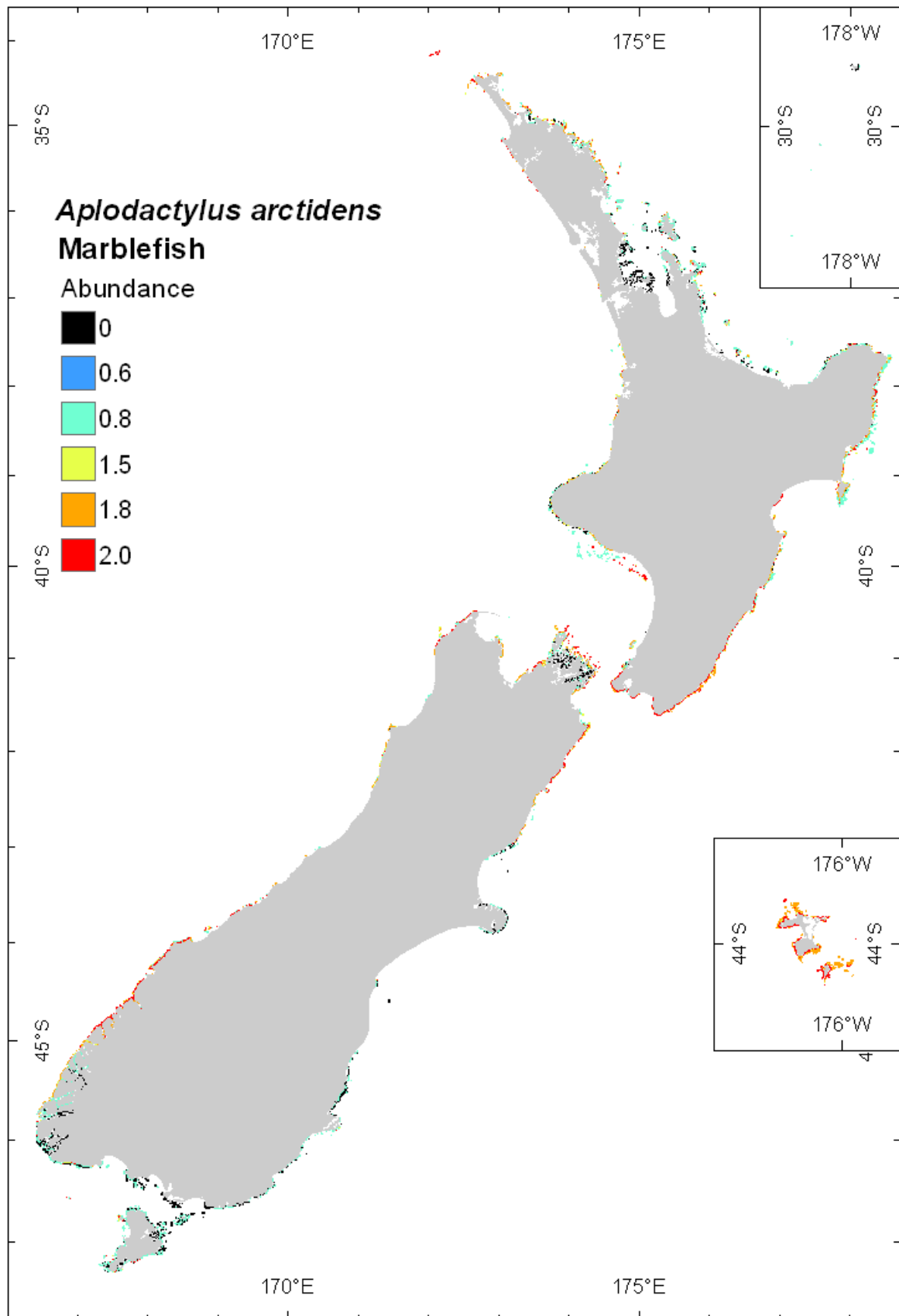
9. Appendix III – Maps of predicted abundance of reef fishes

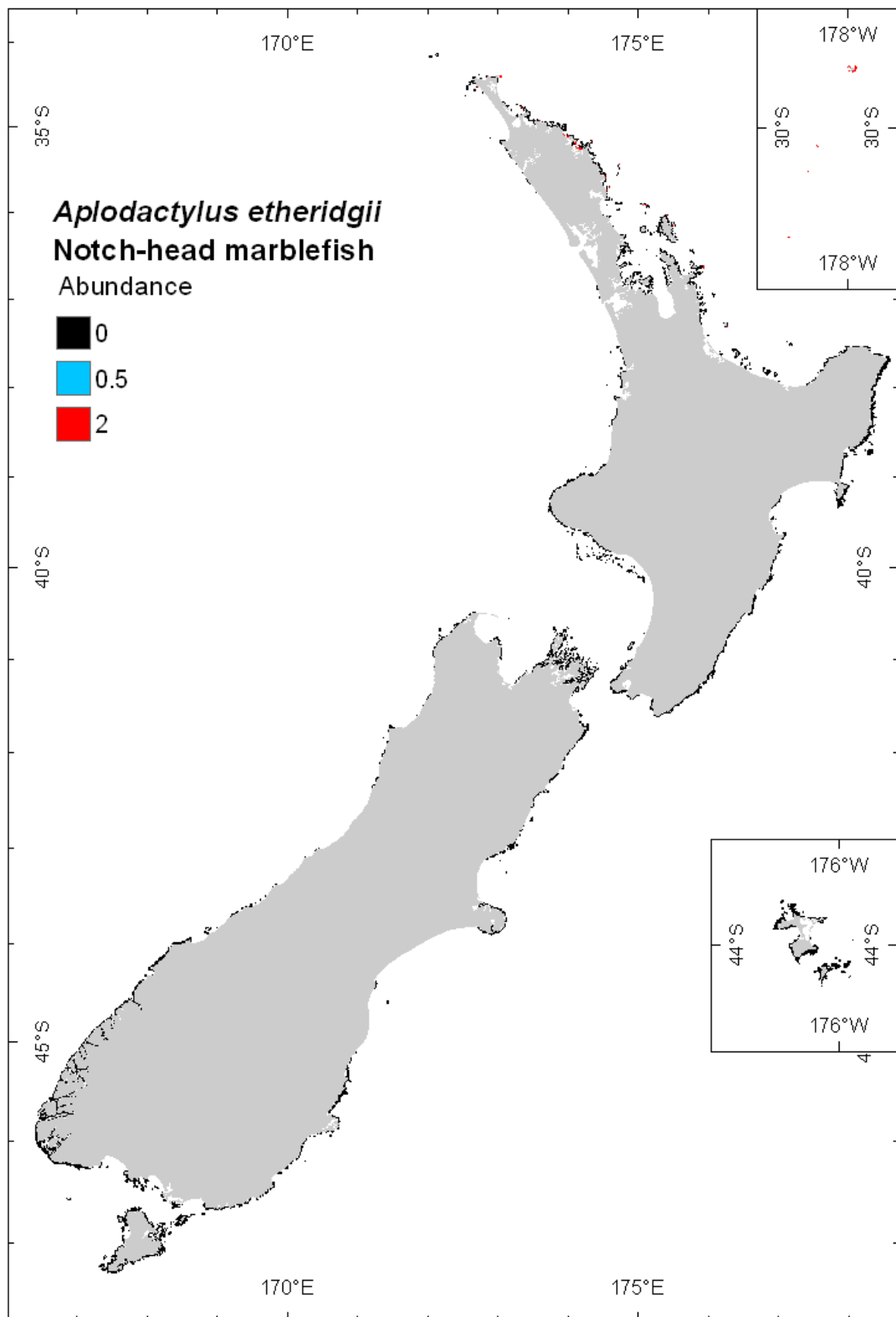
This appendix shows 72 maps of predicted distributions and relative abundance of fishes on shallow subtidal reefs around New Zealand. The predictions were produced by applying boosted regression tree regression to dive surveys of fishes, using environmental and geographic variables as predictors. The scales of relative abundance used on the maps are on the same scale as the original data, which is a scale that roughly corresponds to the log of the number of that species seen per dive. Specifically, 0 = absent, 1 = single (1 individual seen), 2 = few (2 – 10), 3 = many (11 – 100) and 4 = abundant (> 100). As the predictions shown here are for grid squares of 1 km² each, much of the finer detail is not clear on a national scale map. To view the predictions properly, the grid data should be loaded into a GIS program.

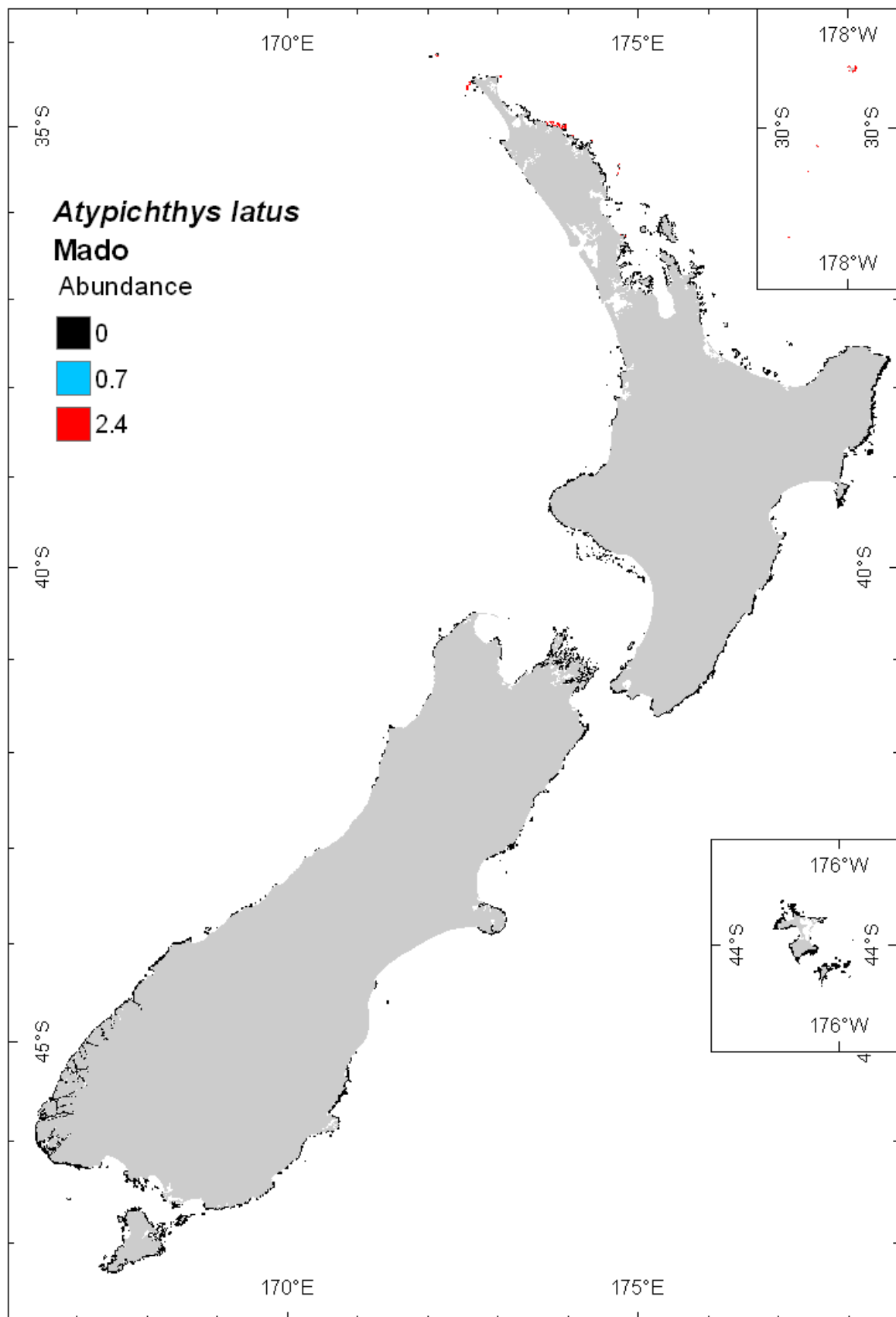


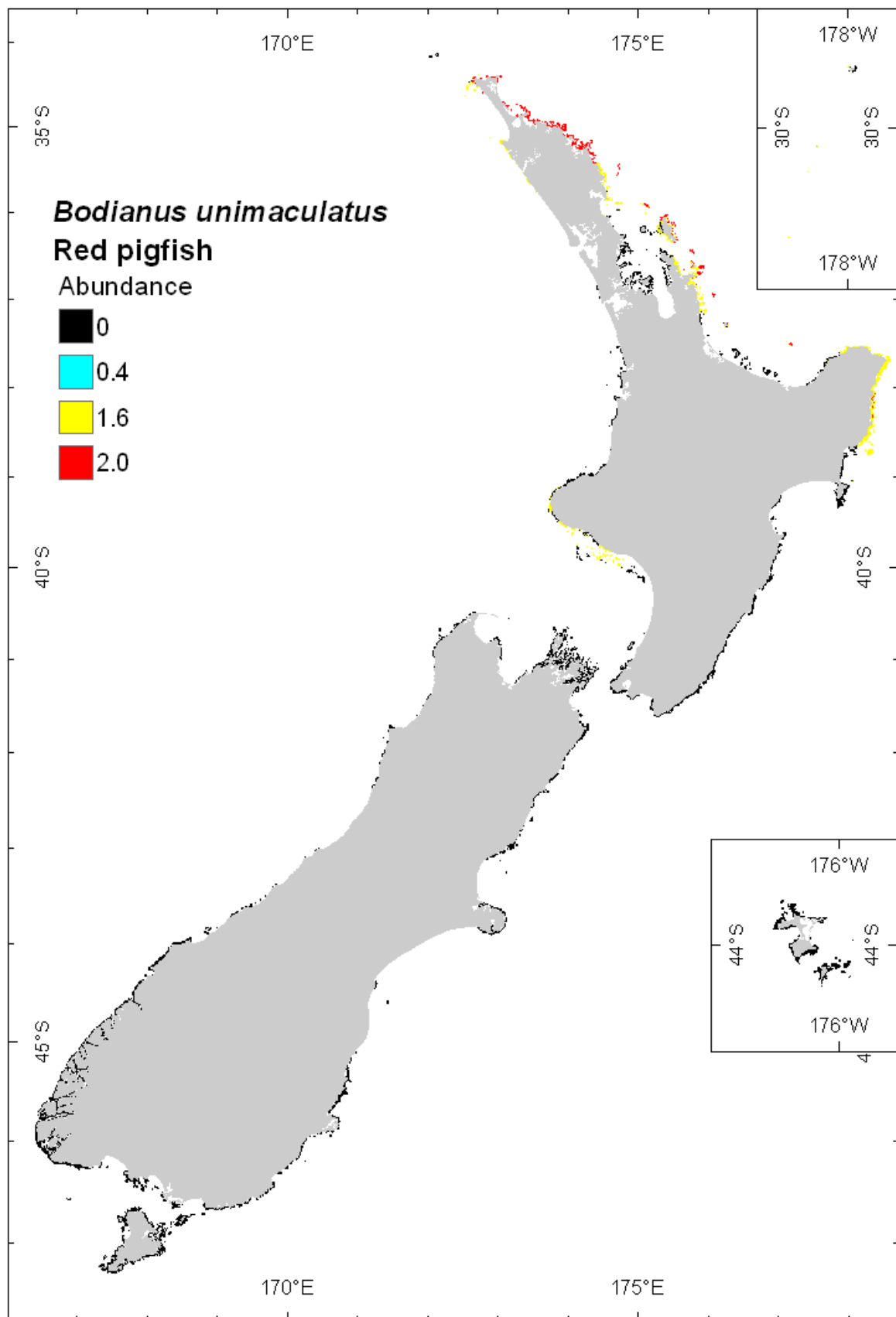


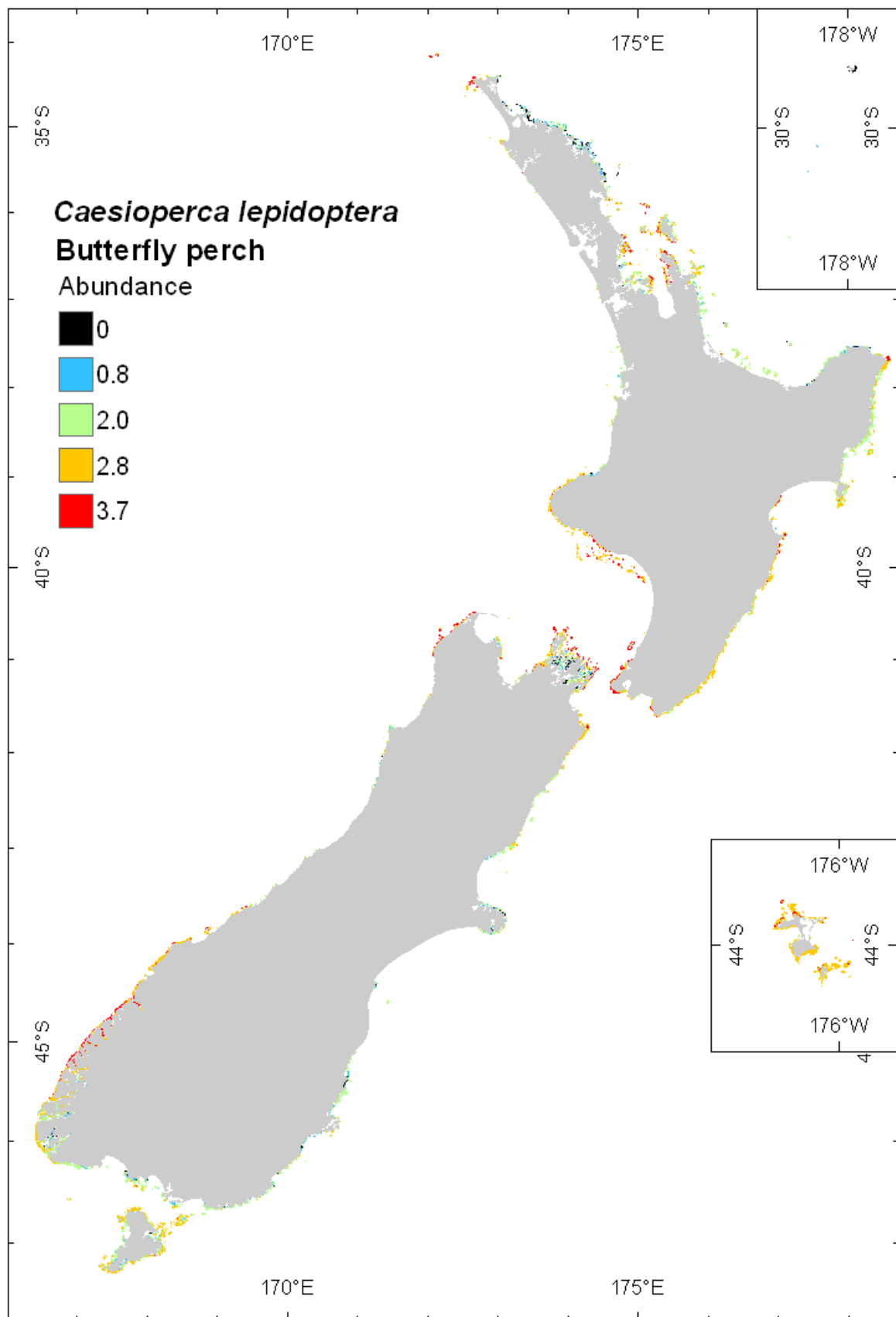


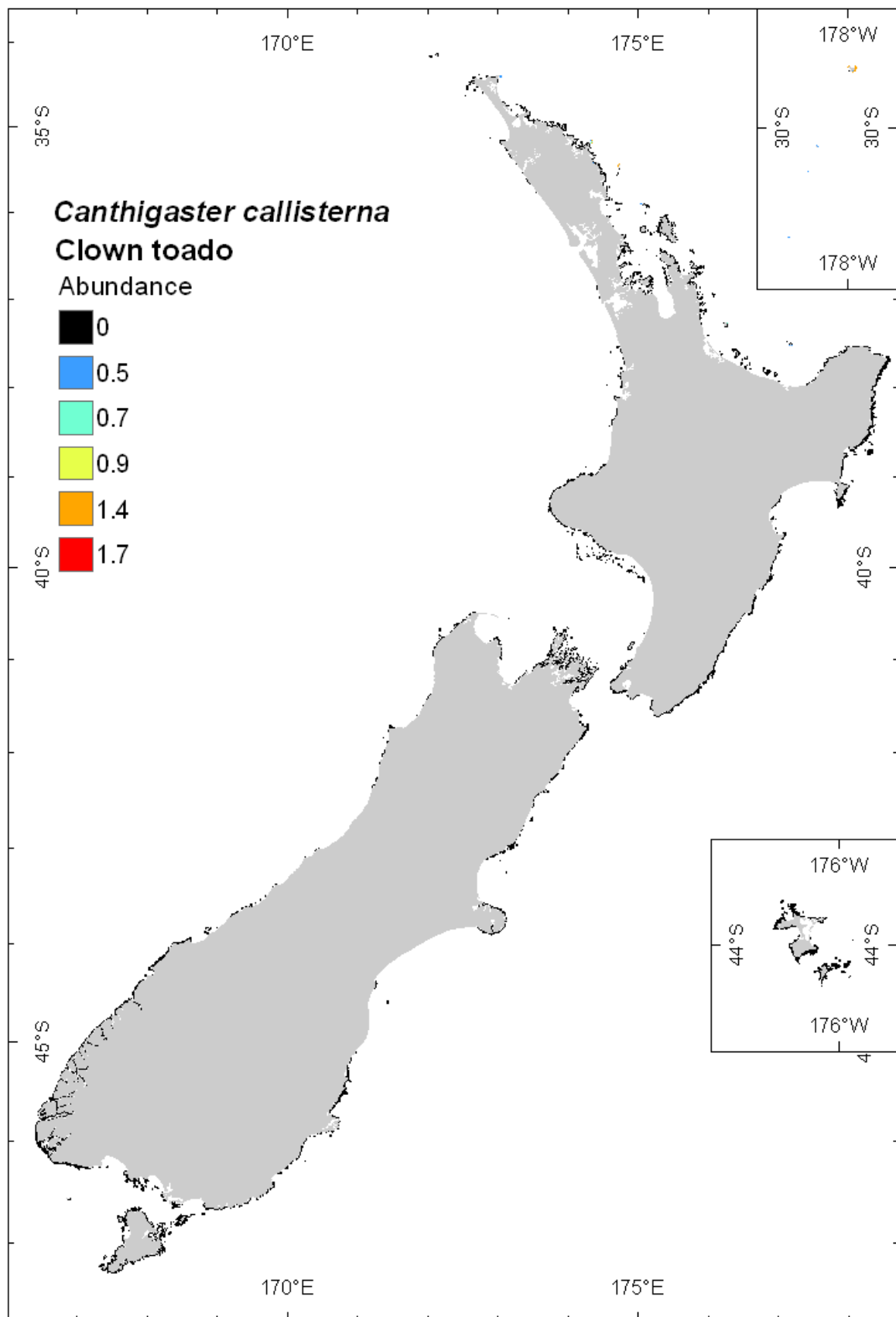


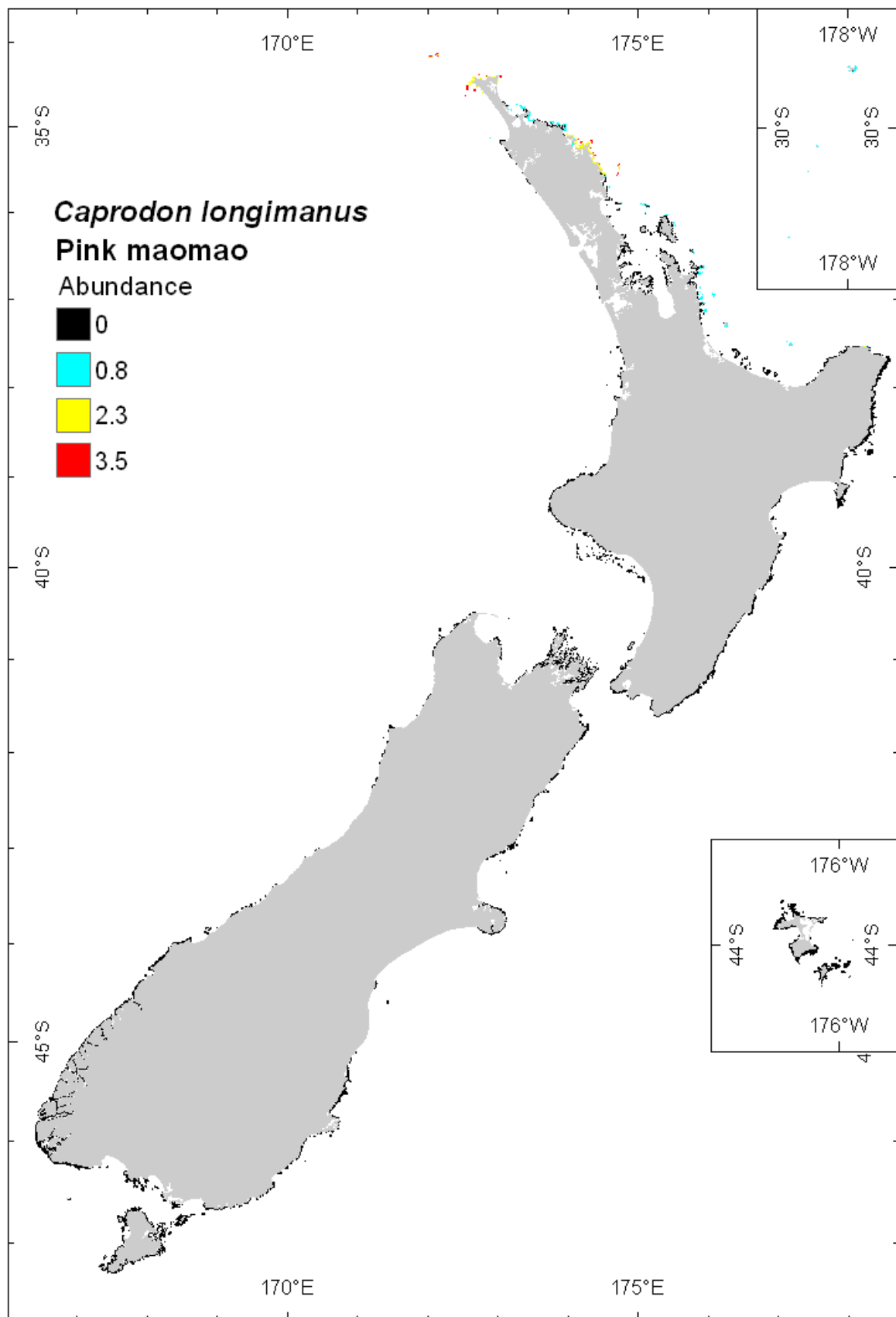


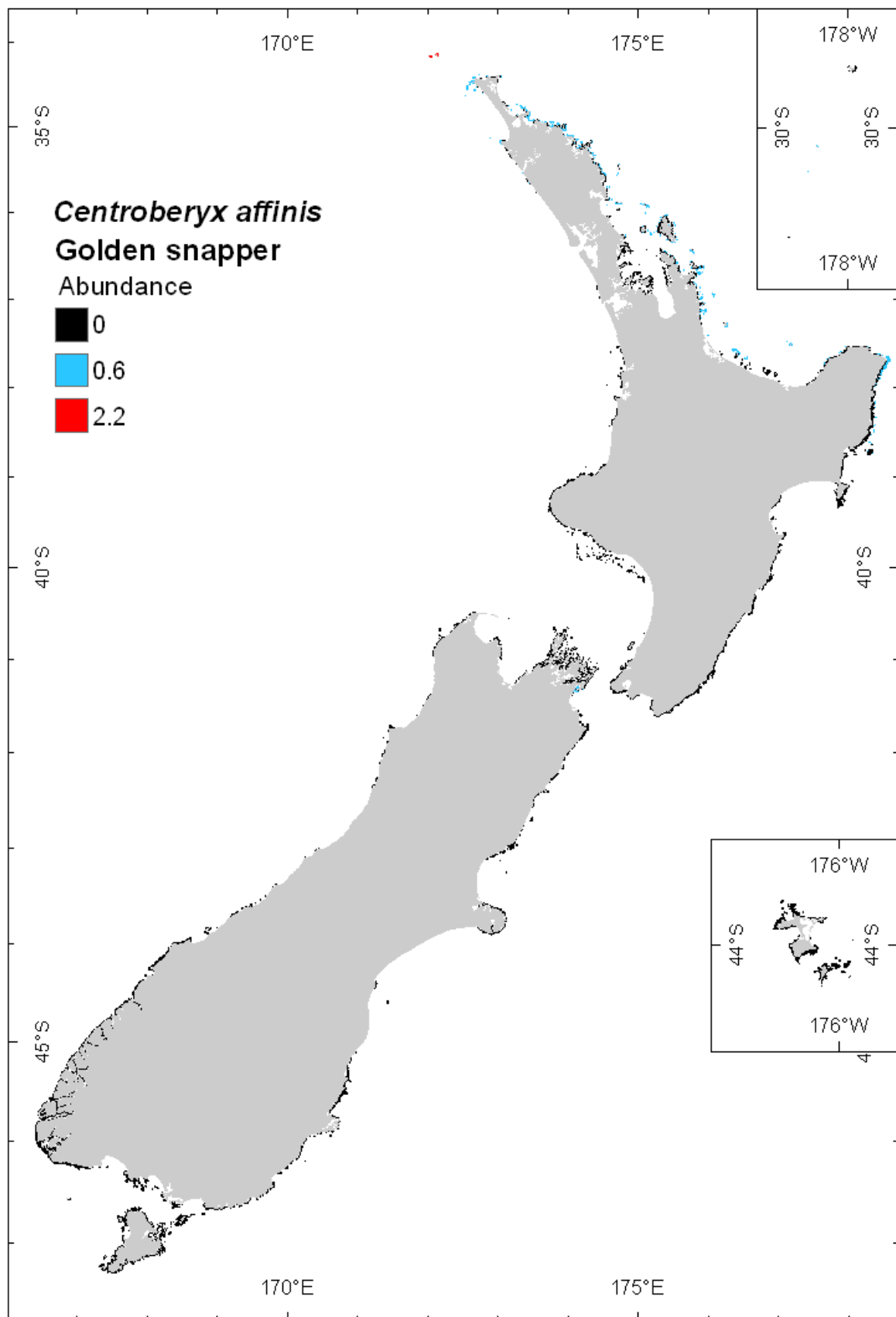


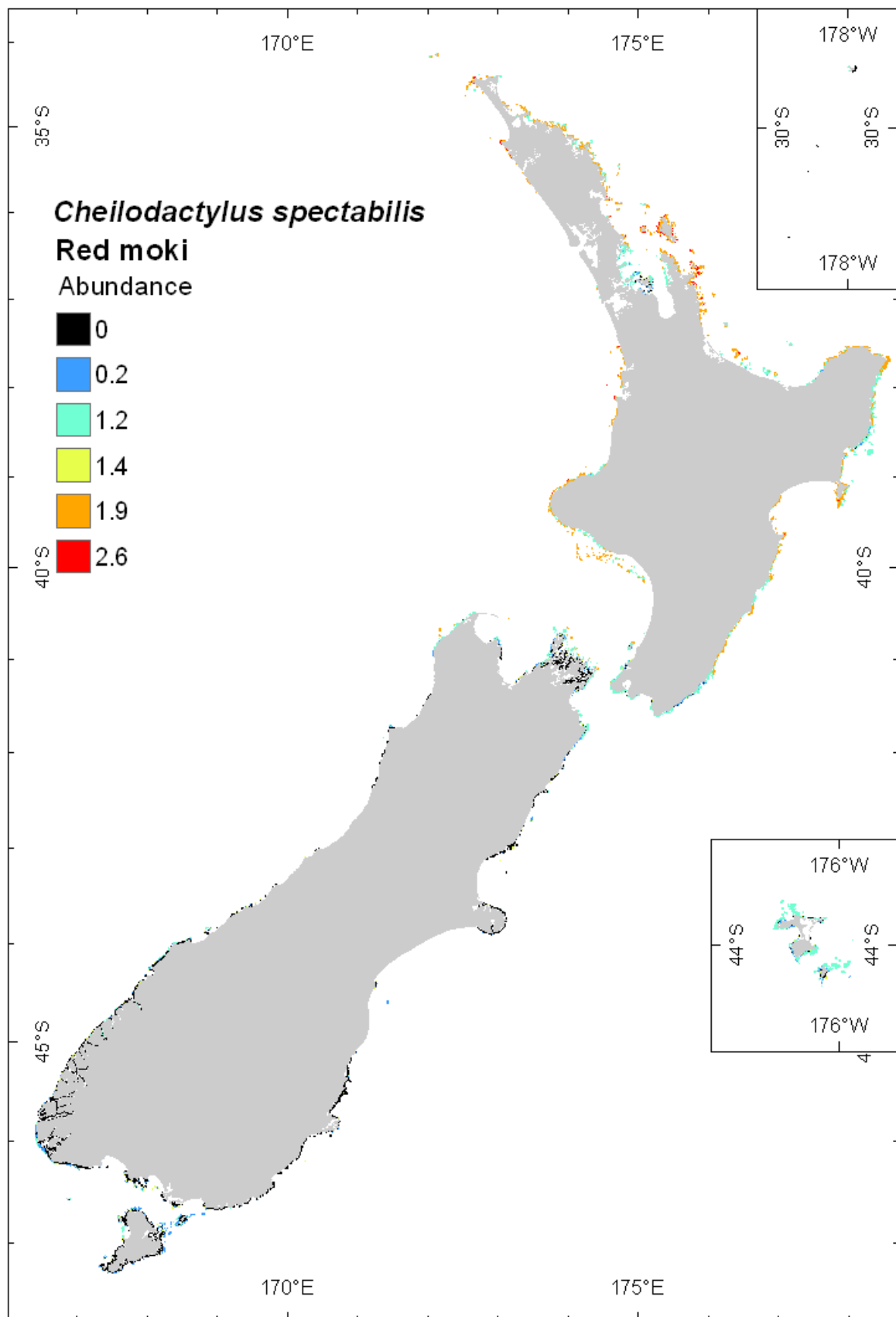


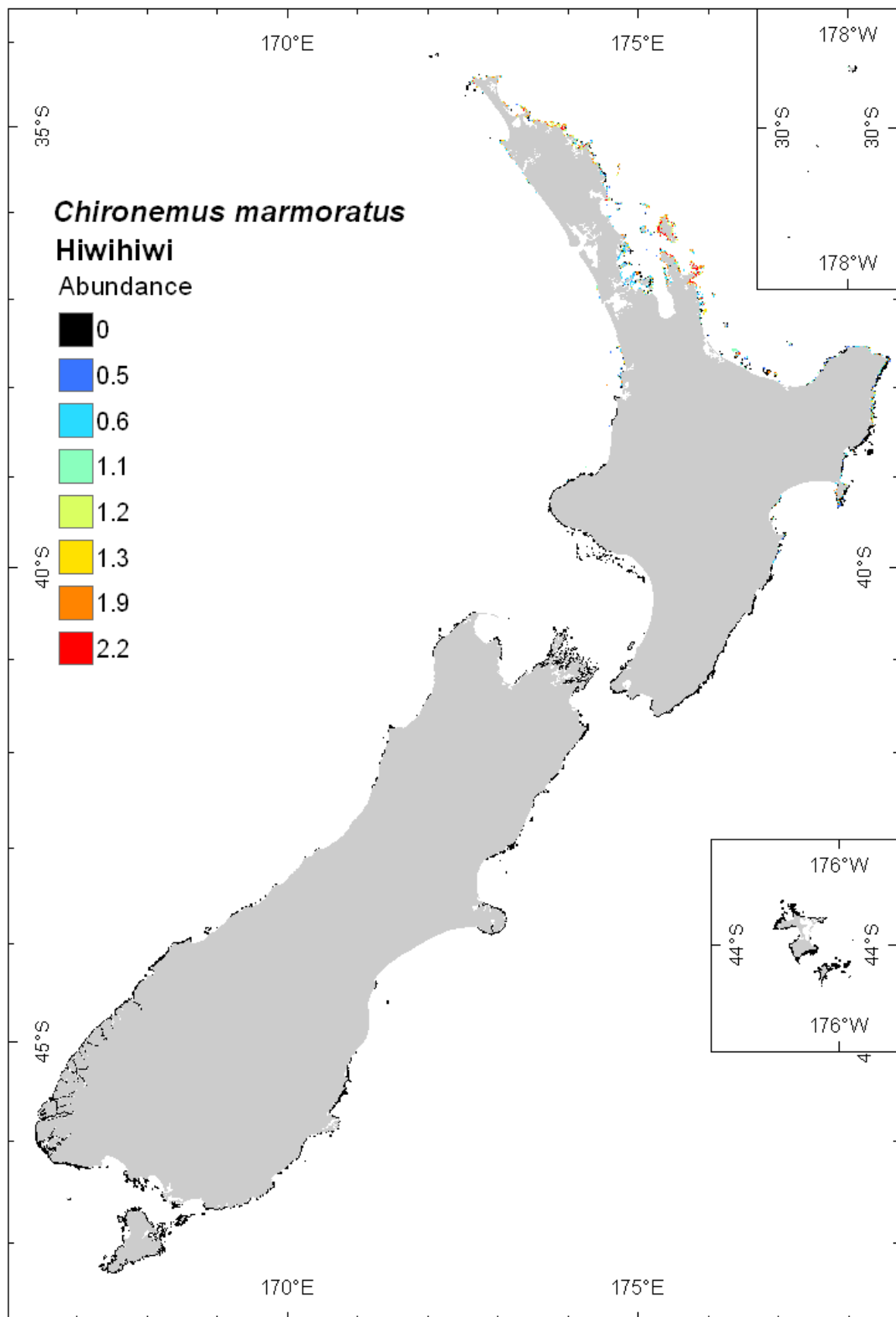


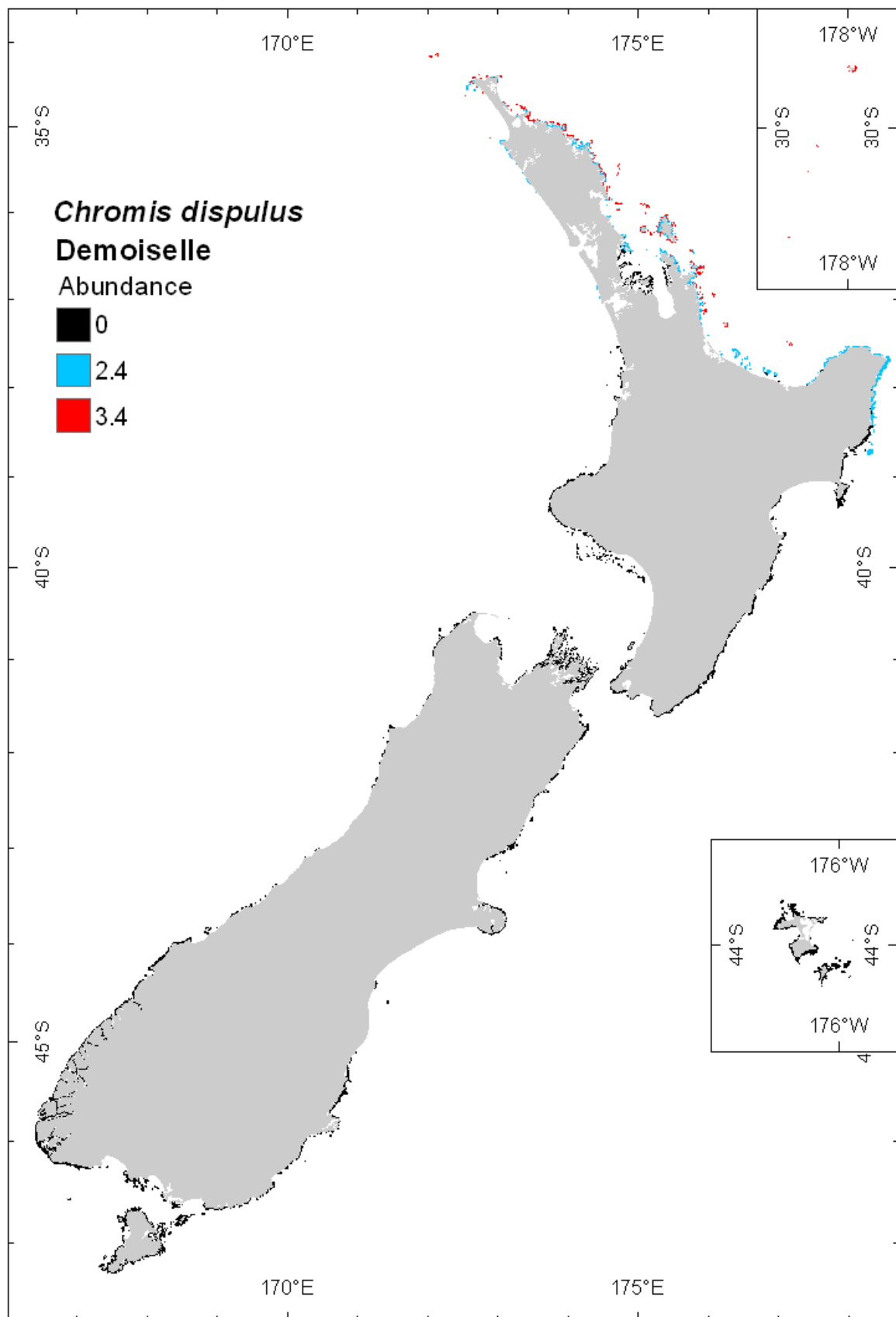


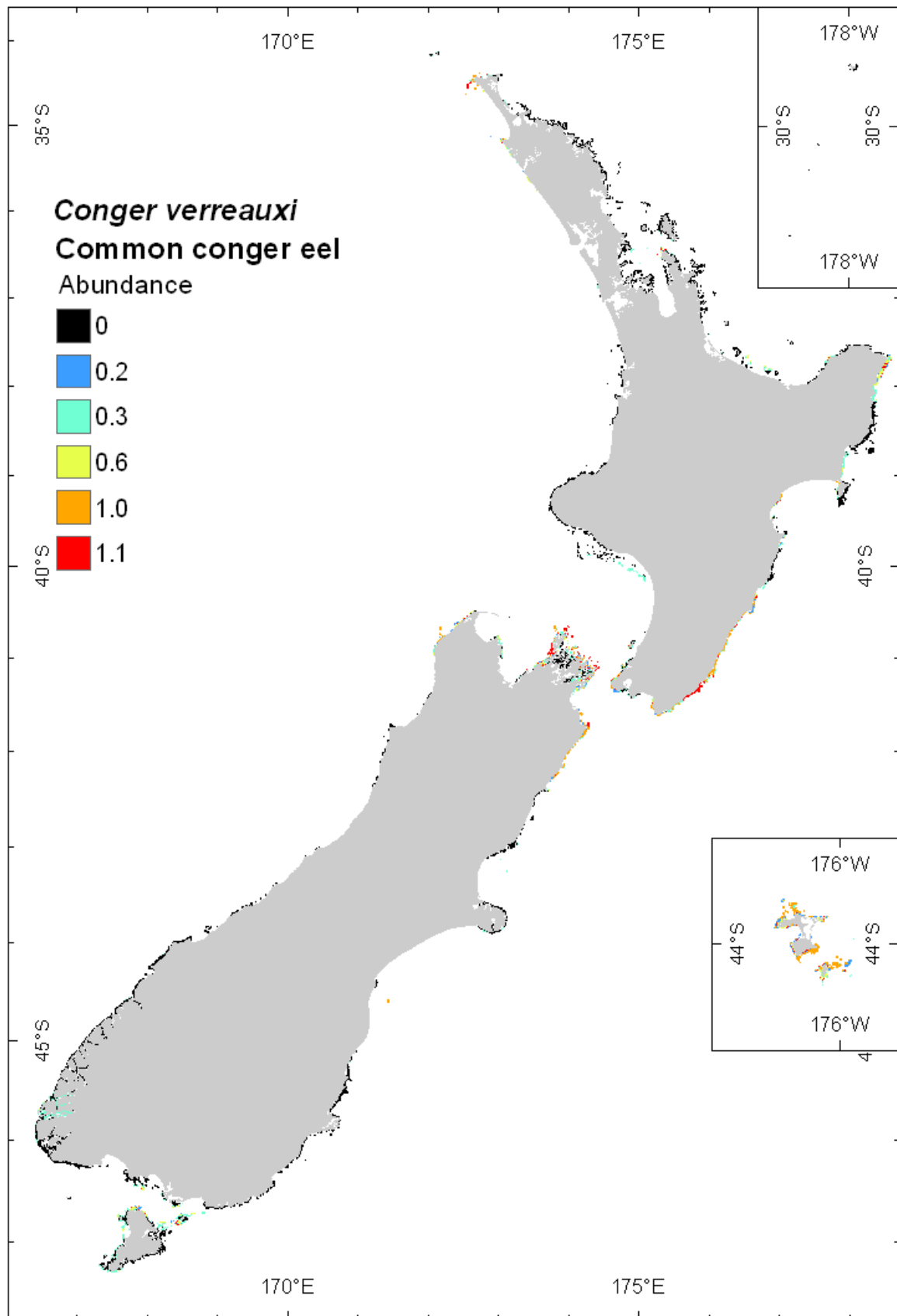


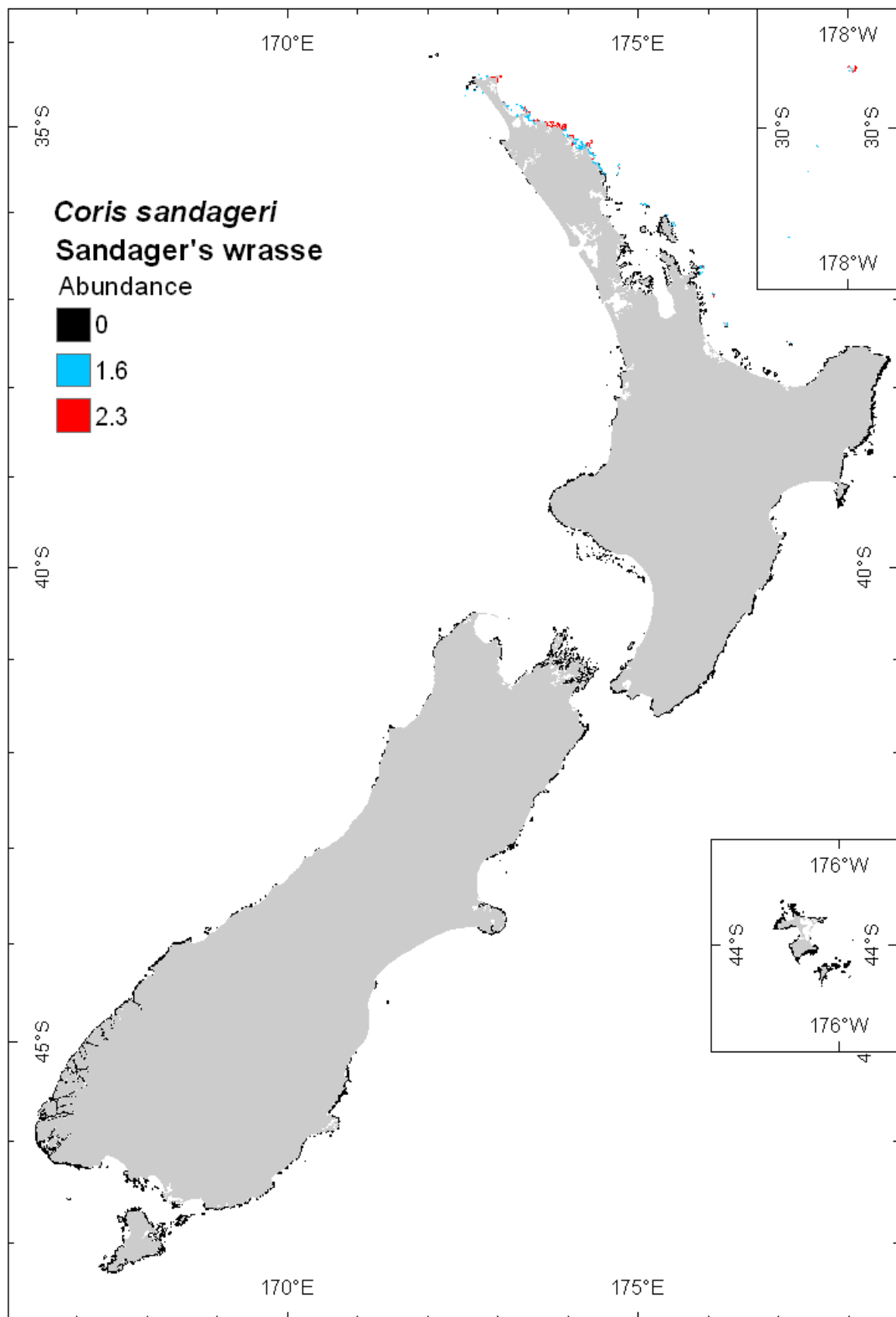


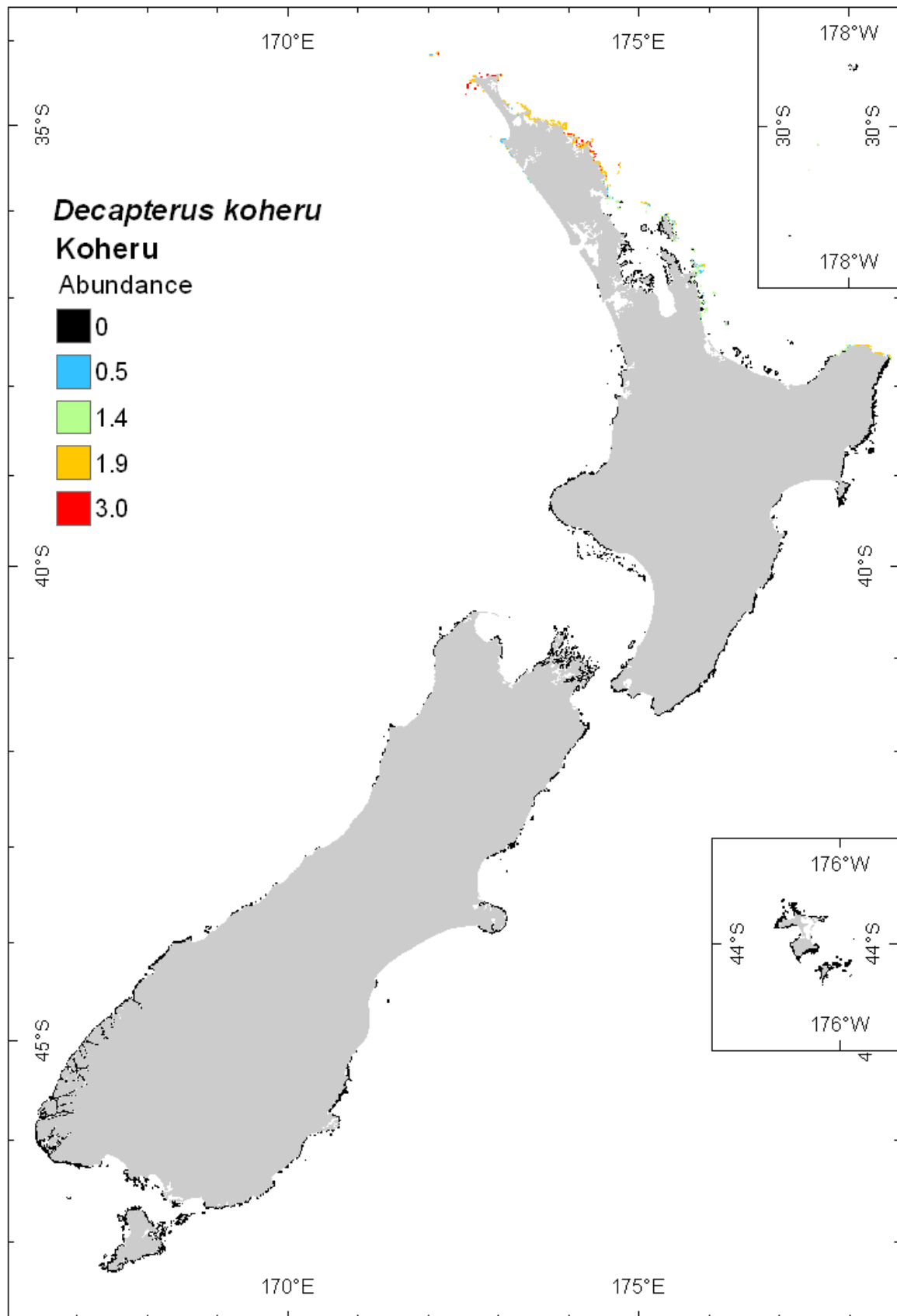


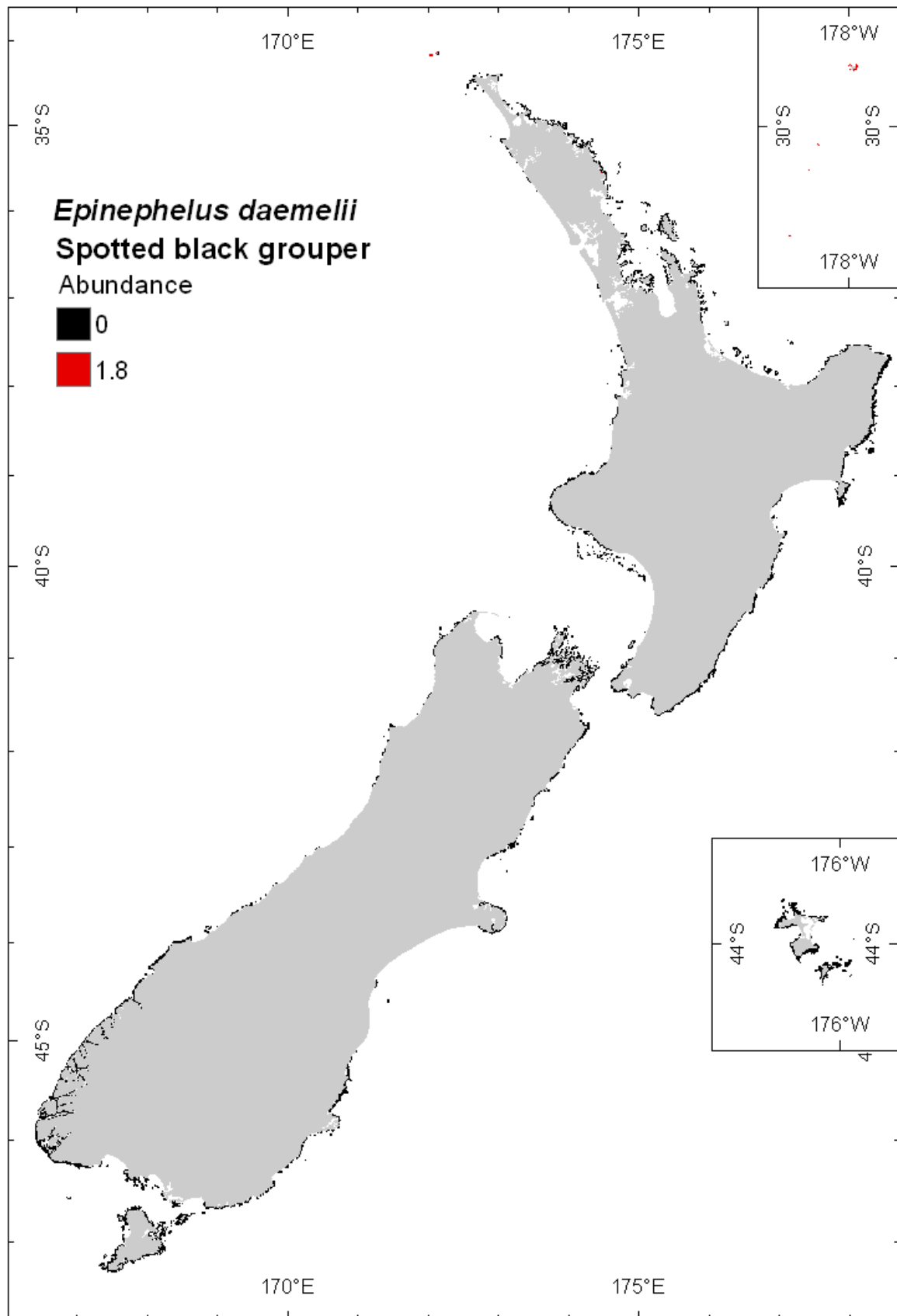


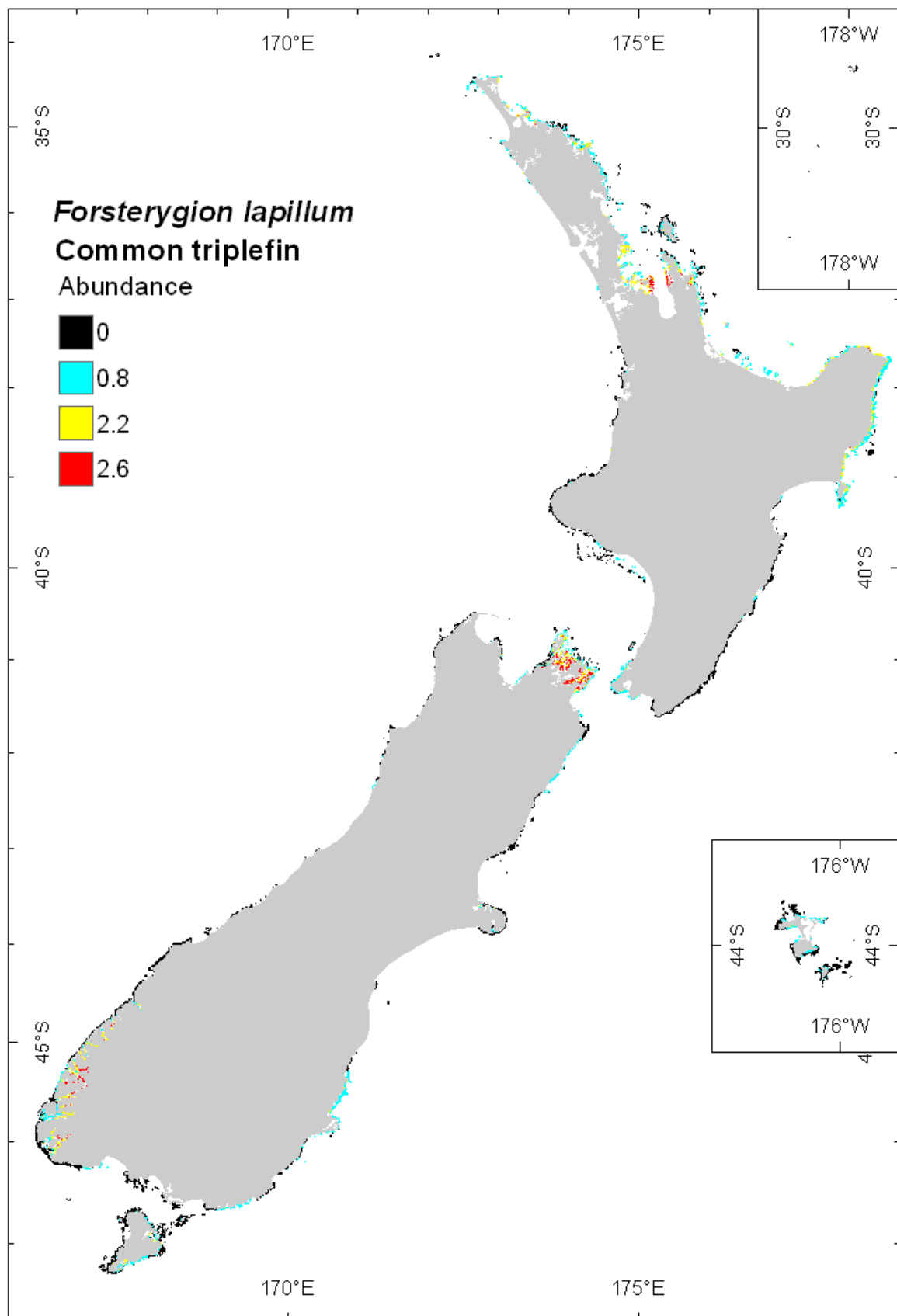


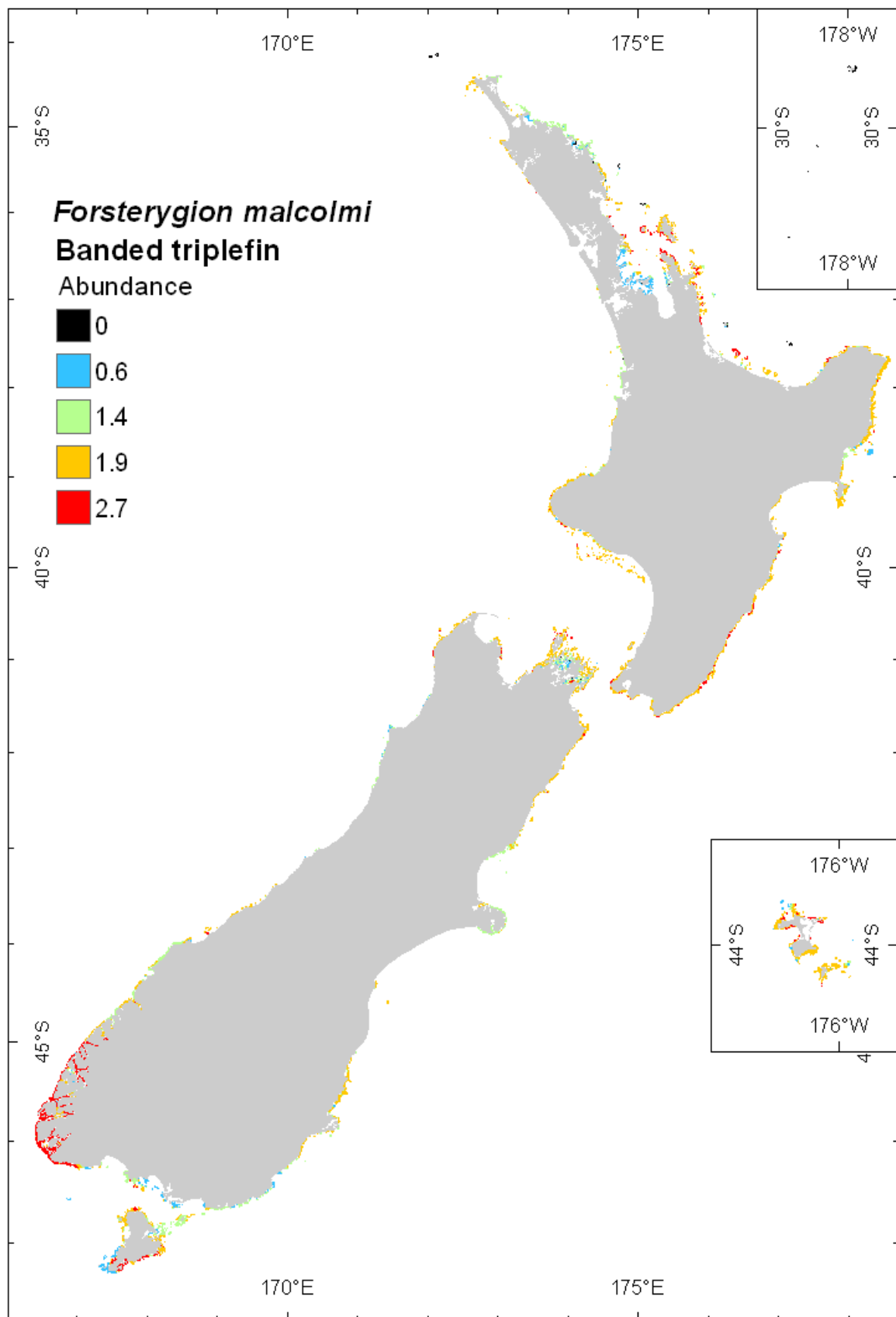


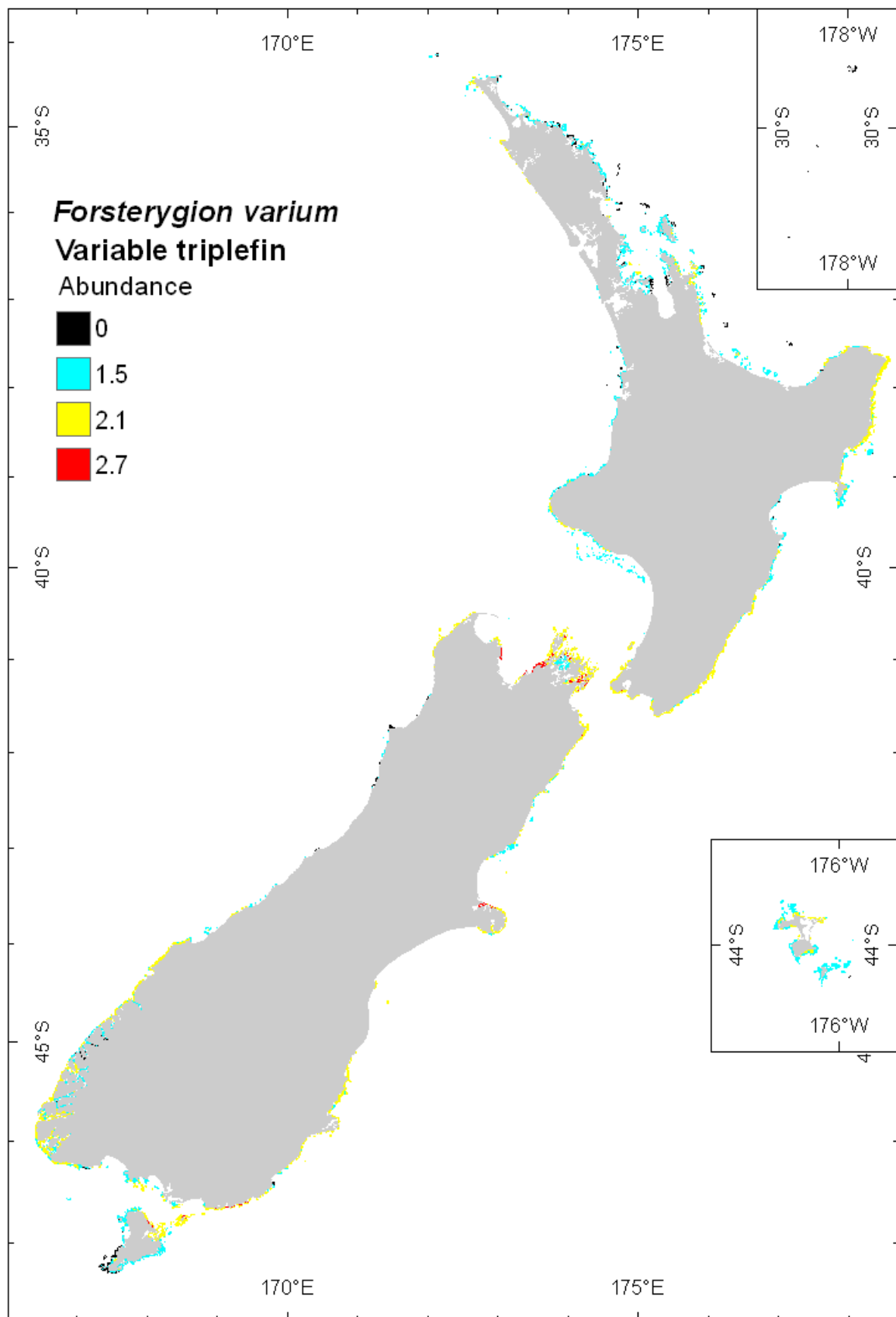


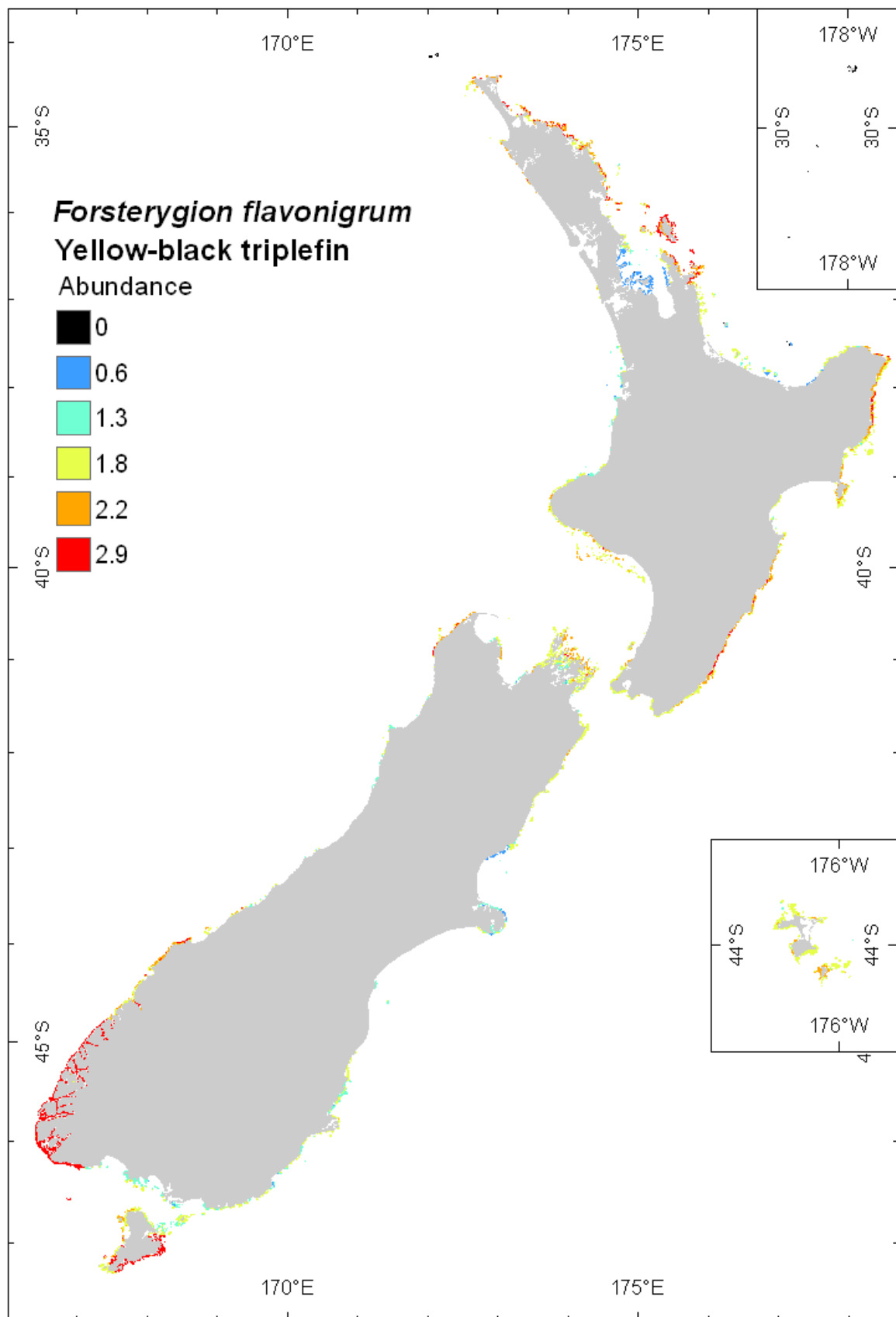


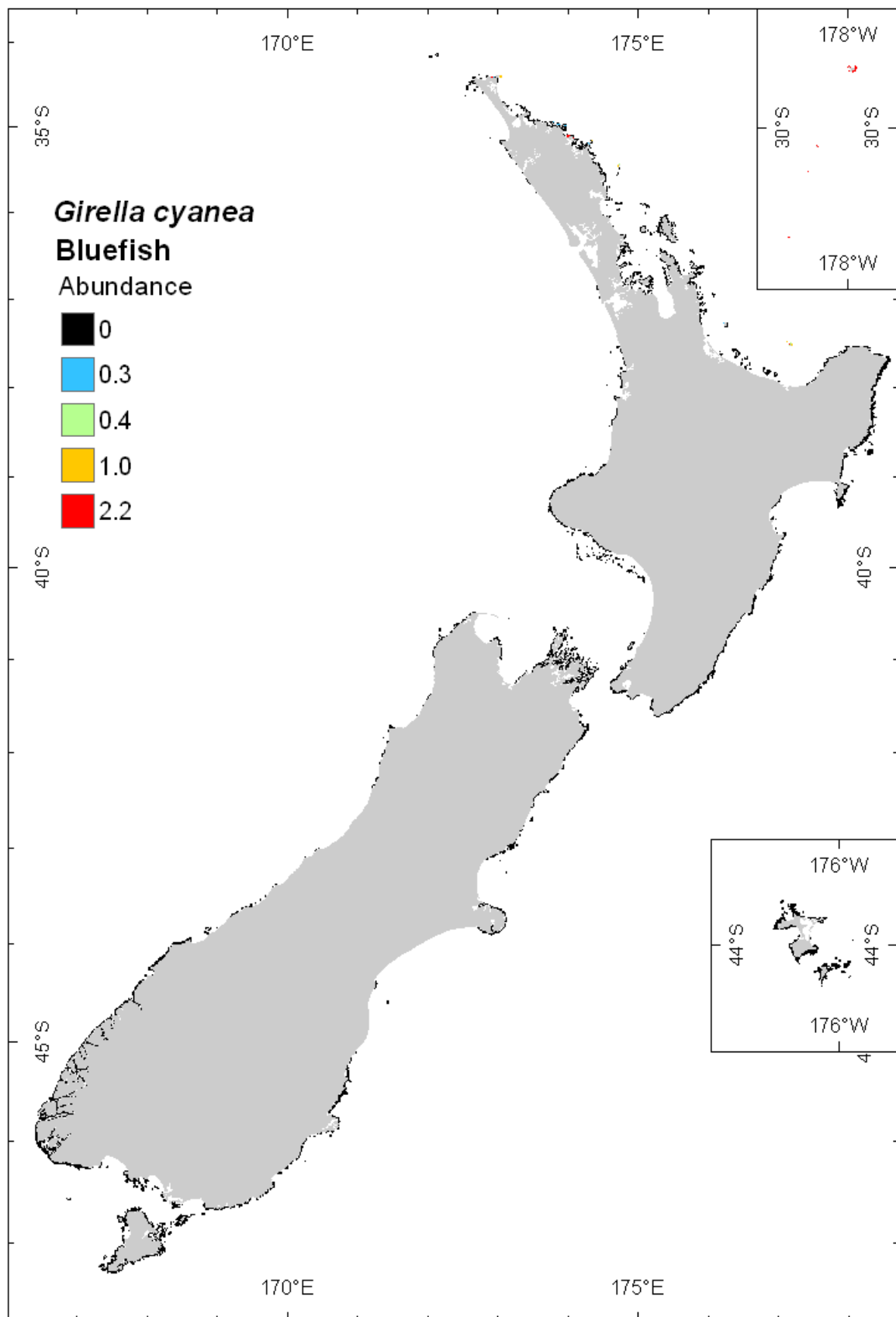


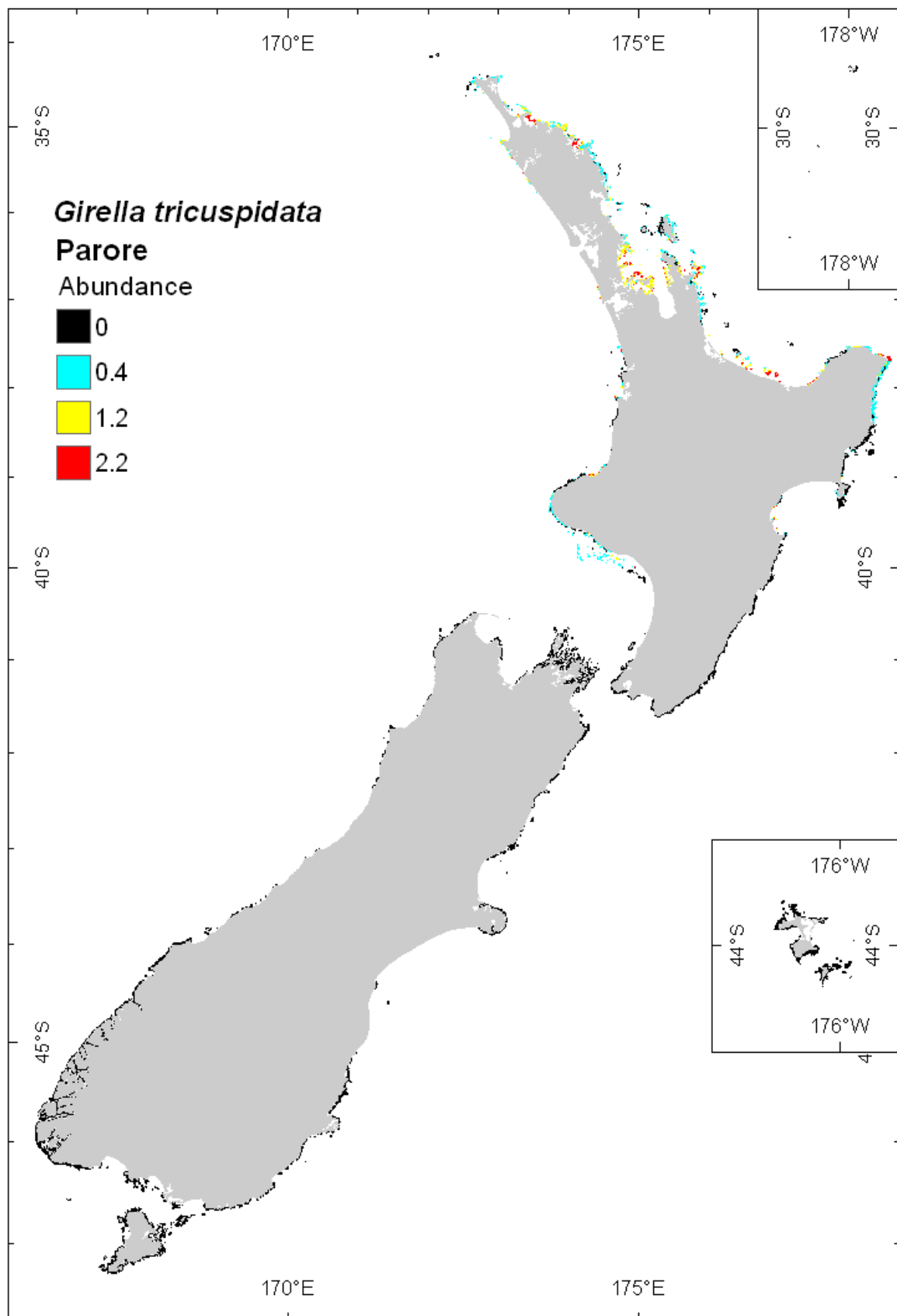


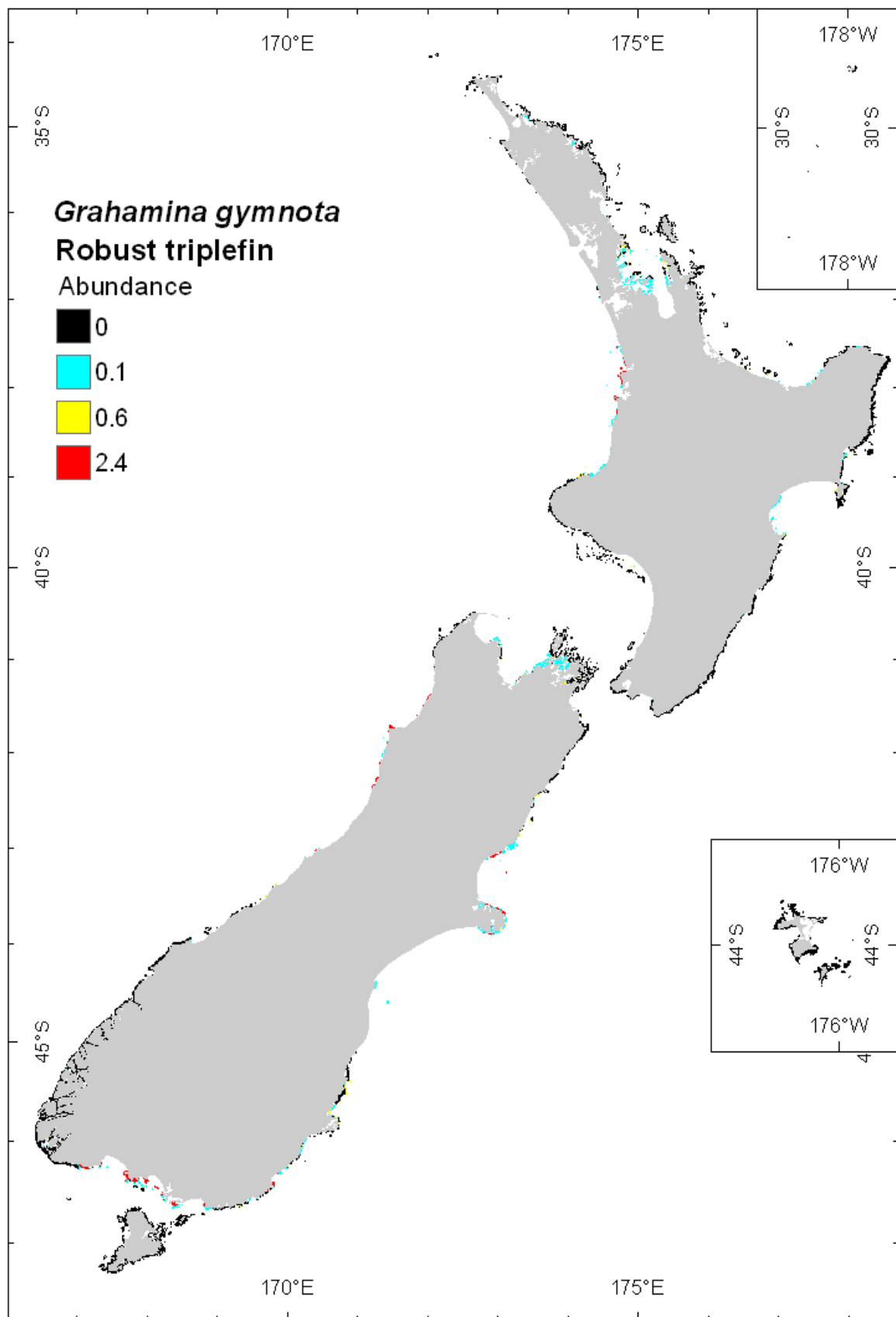


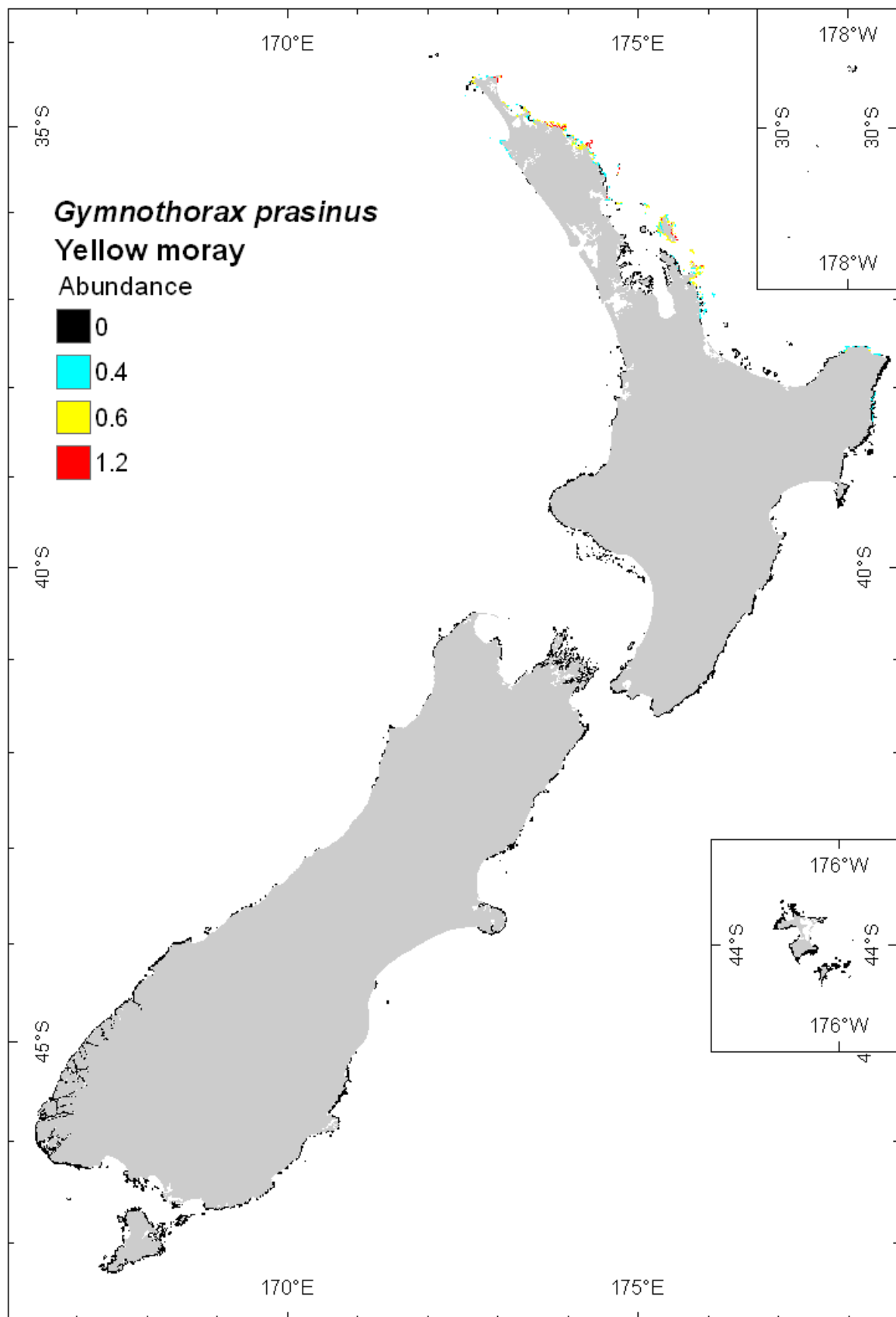


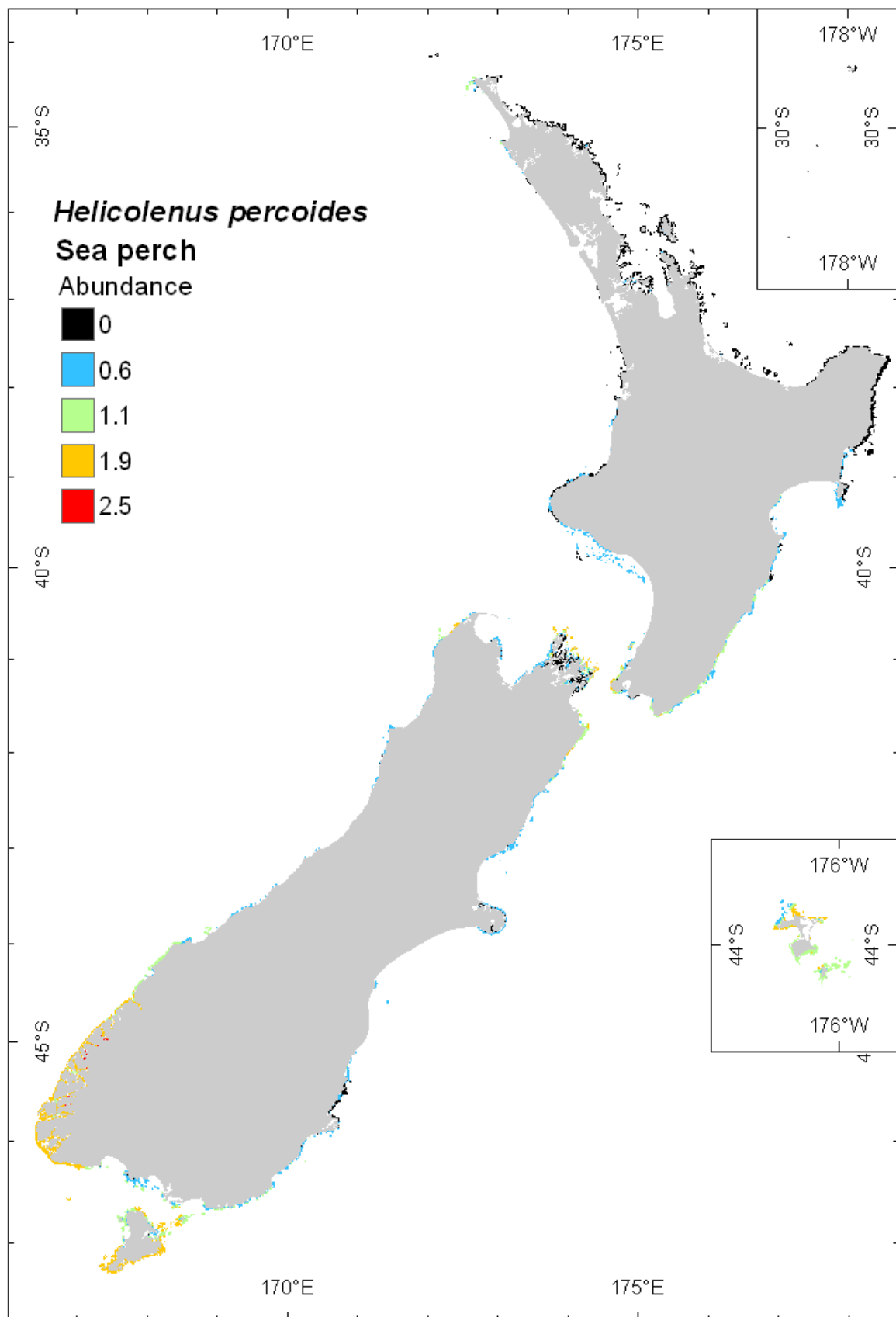


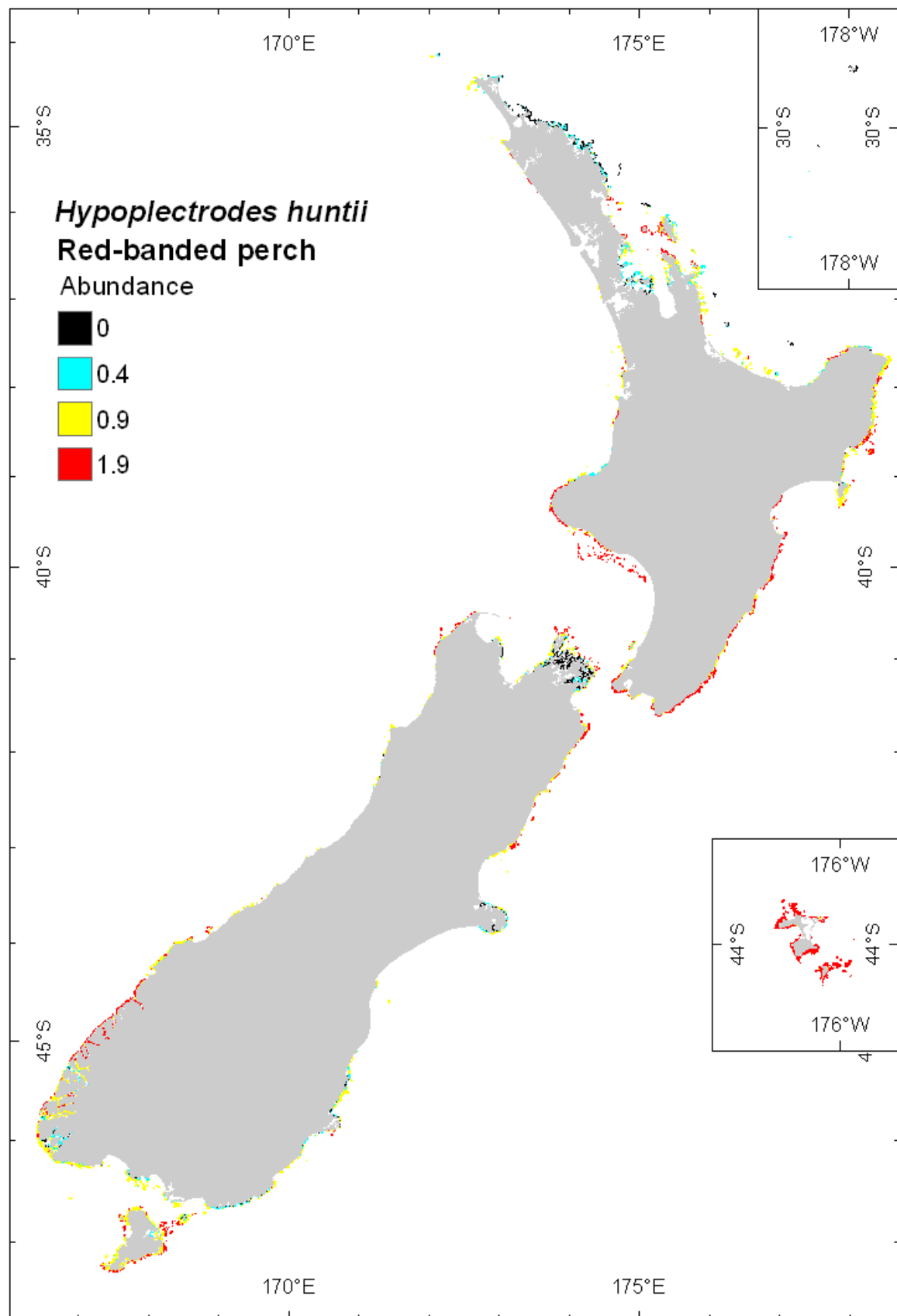


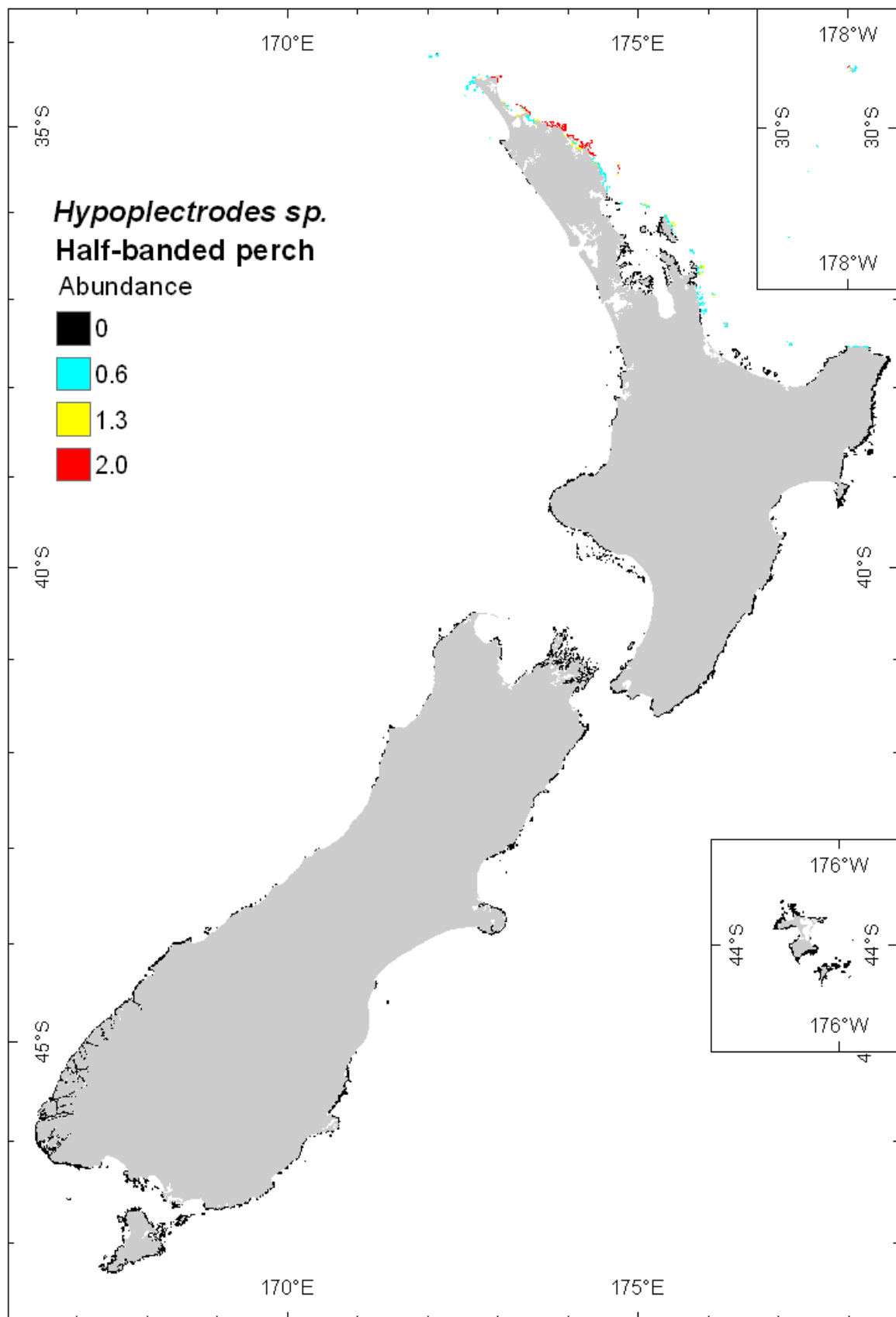


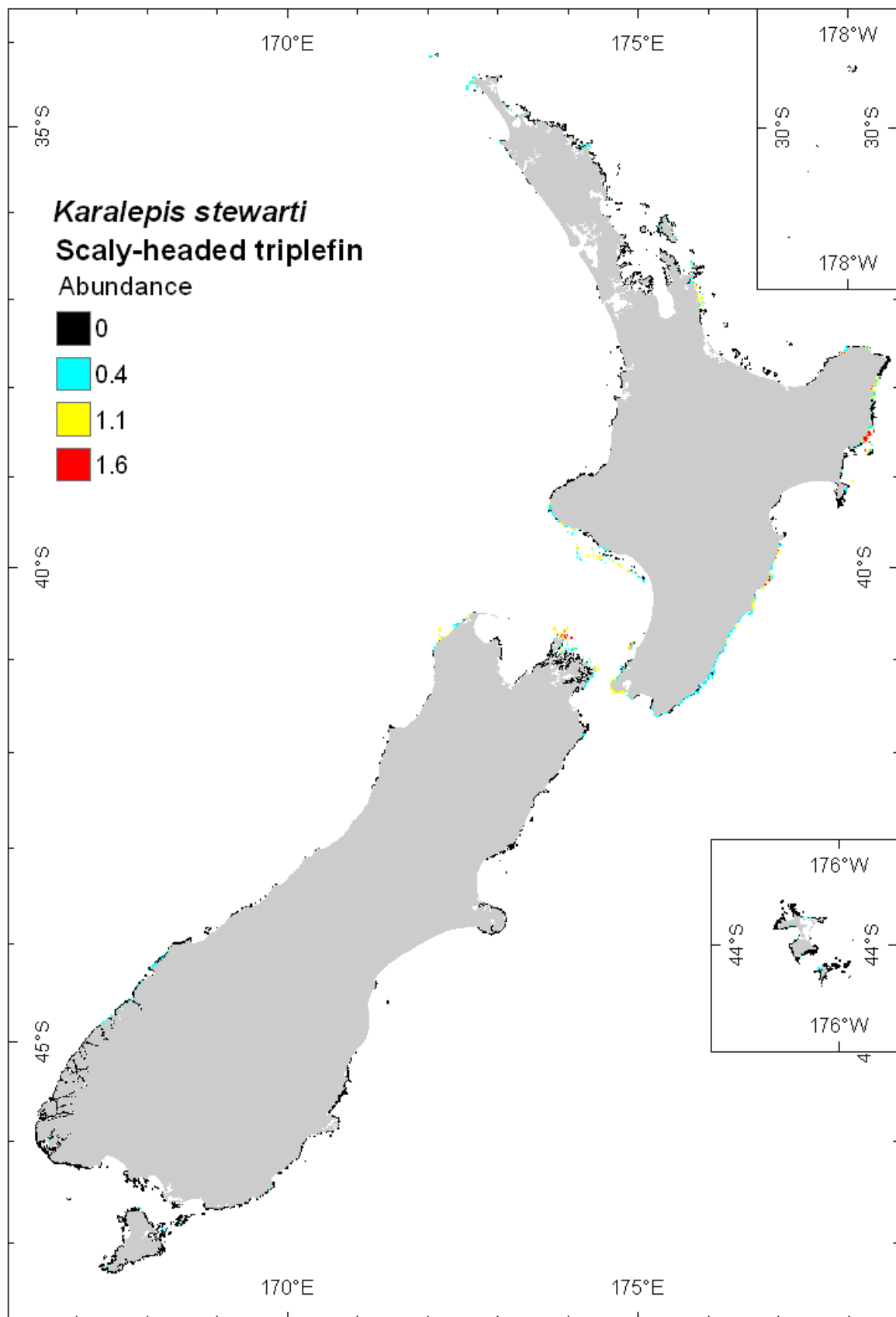


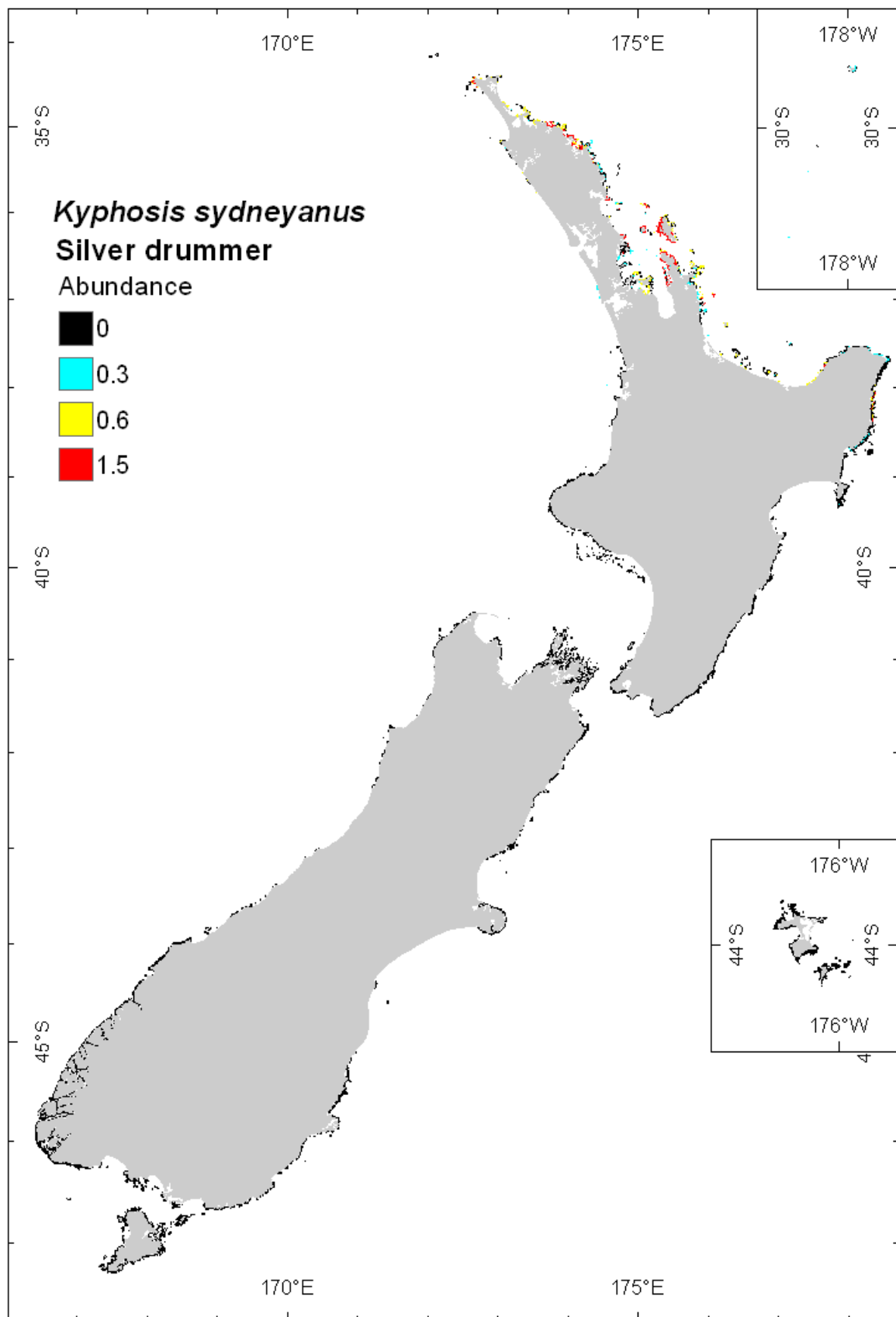


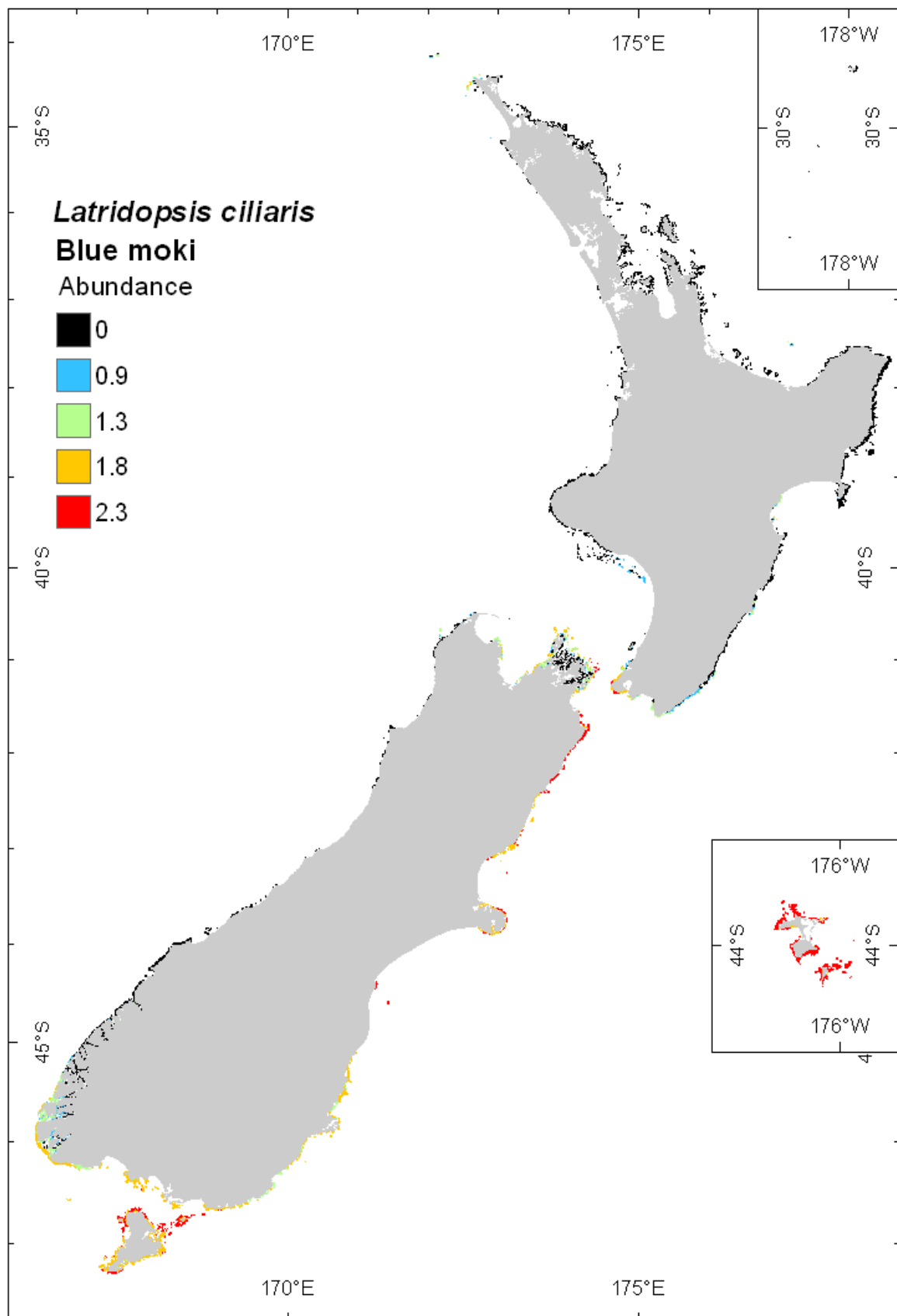


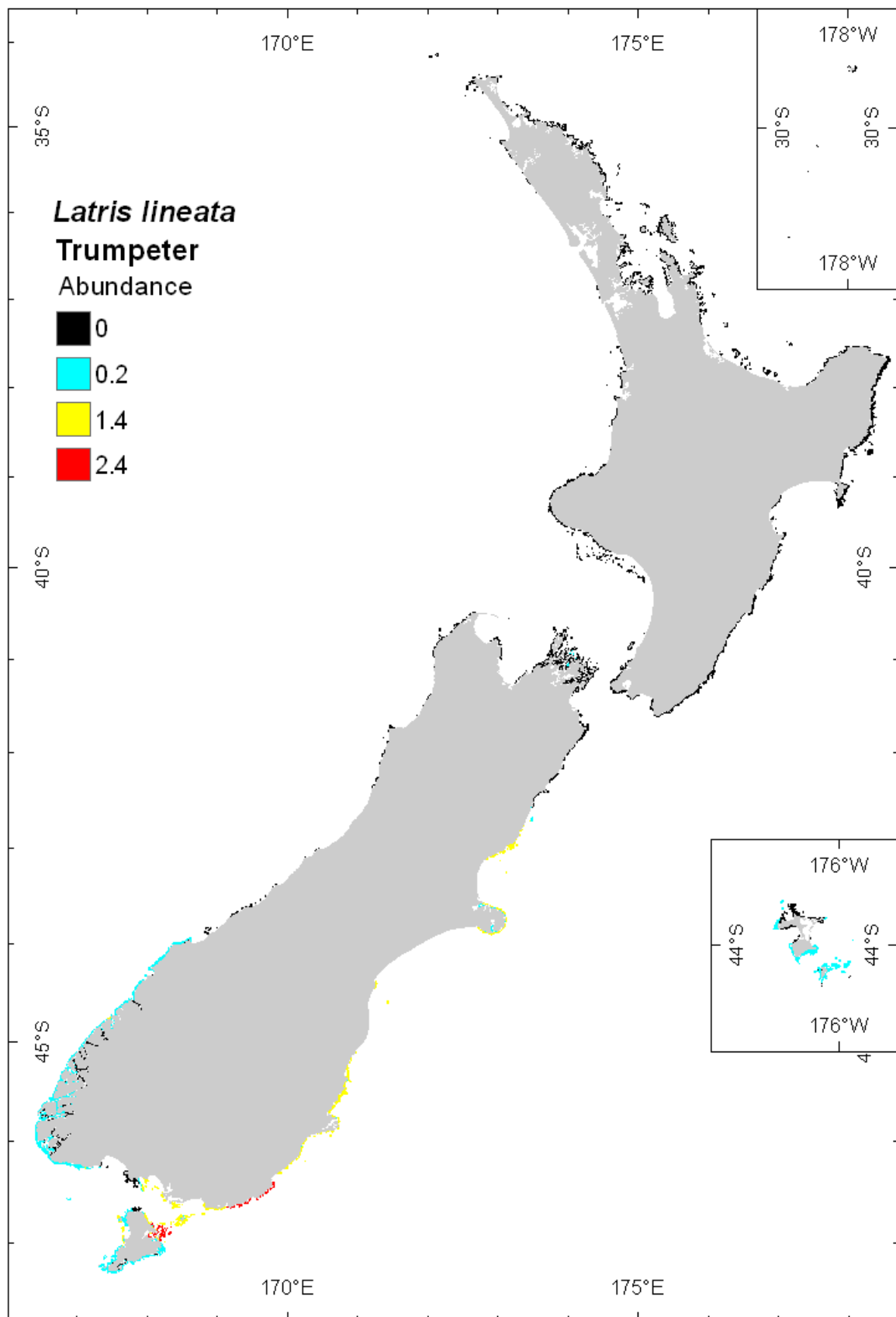


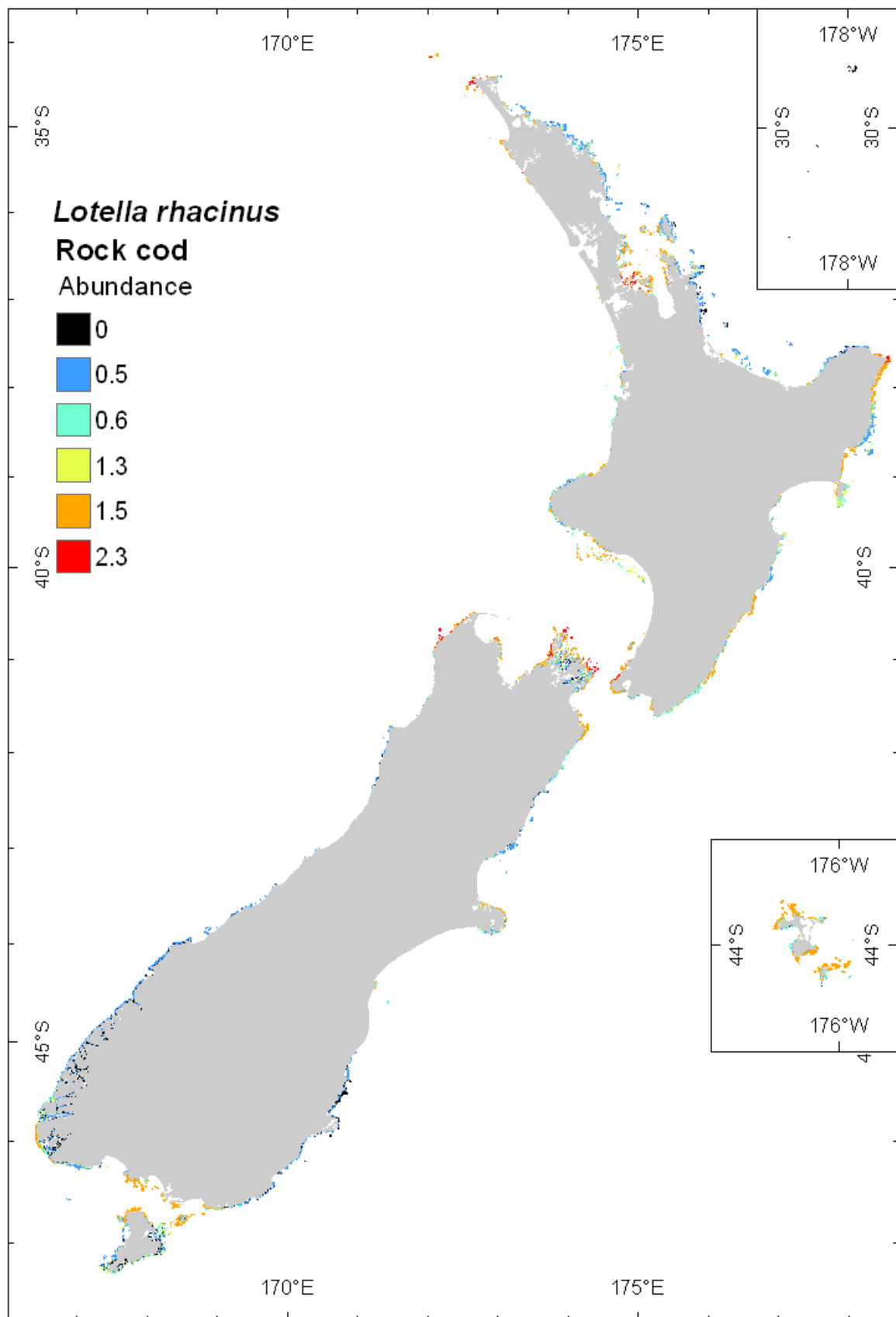


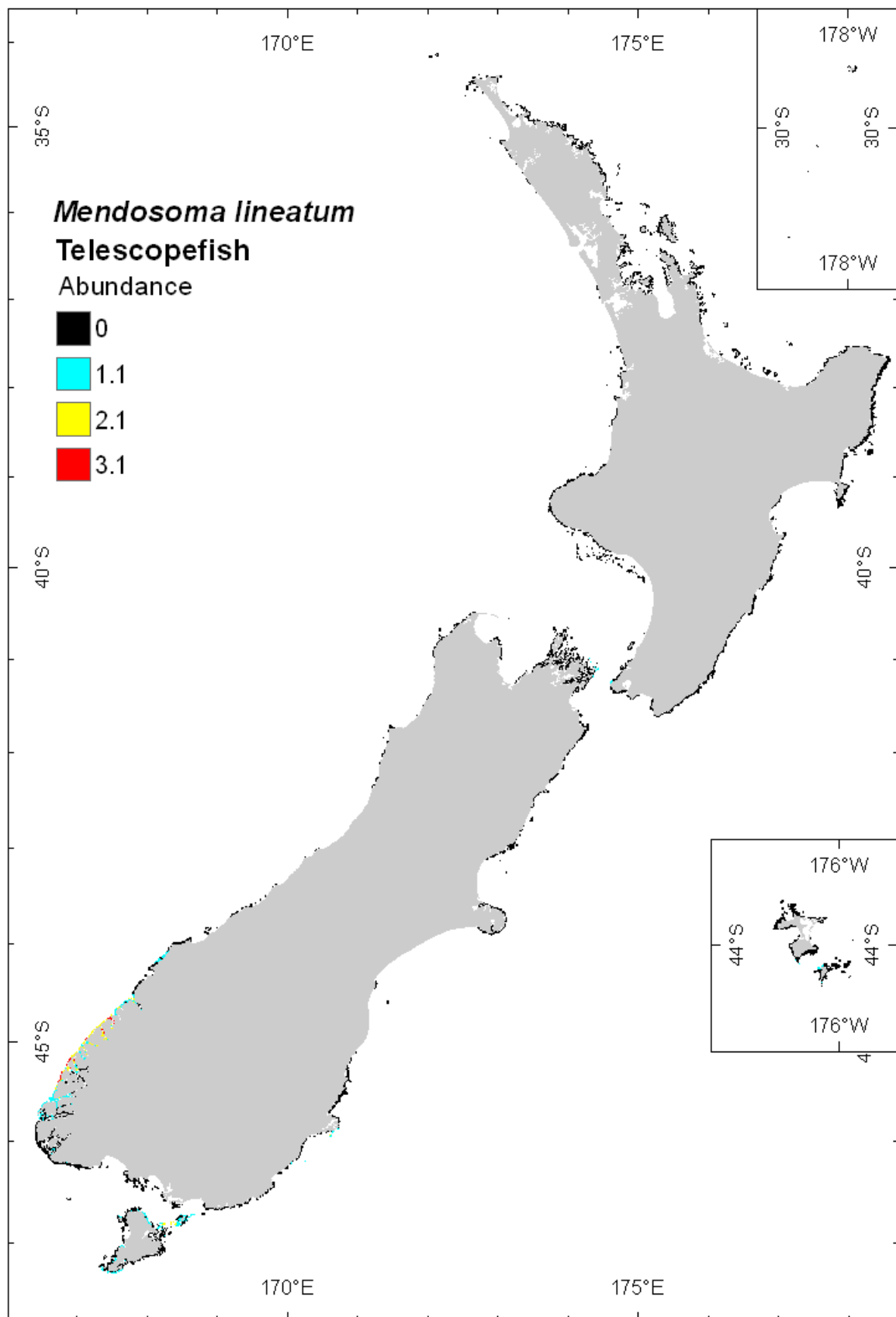


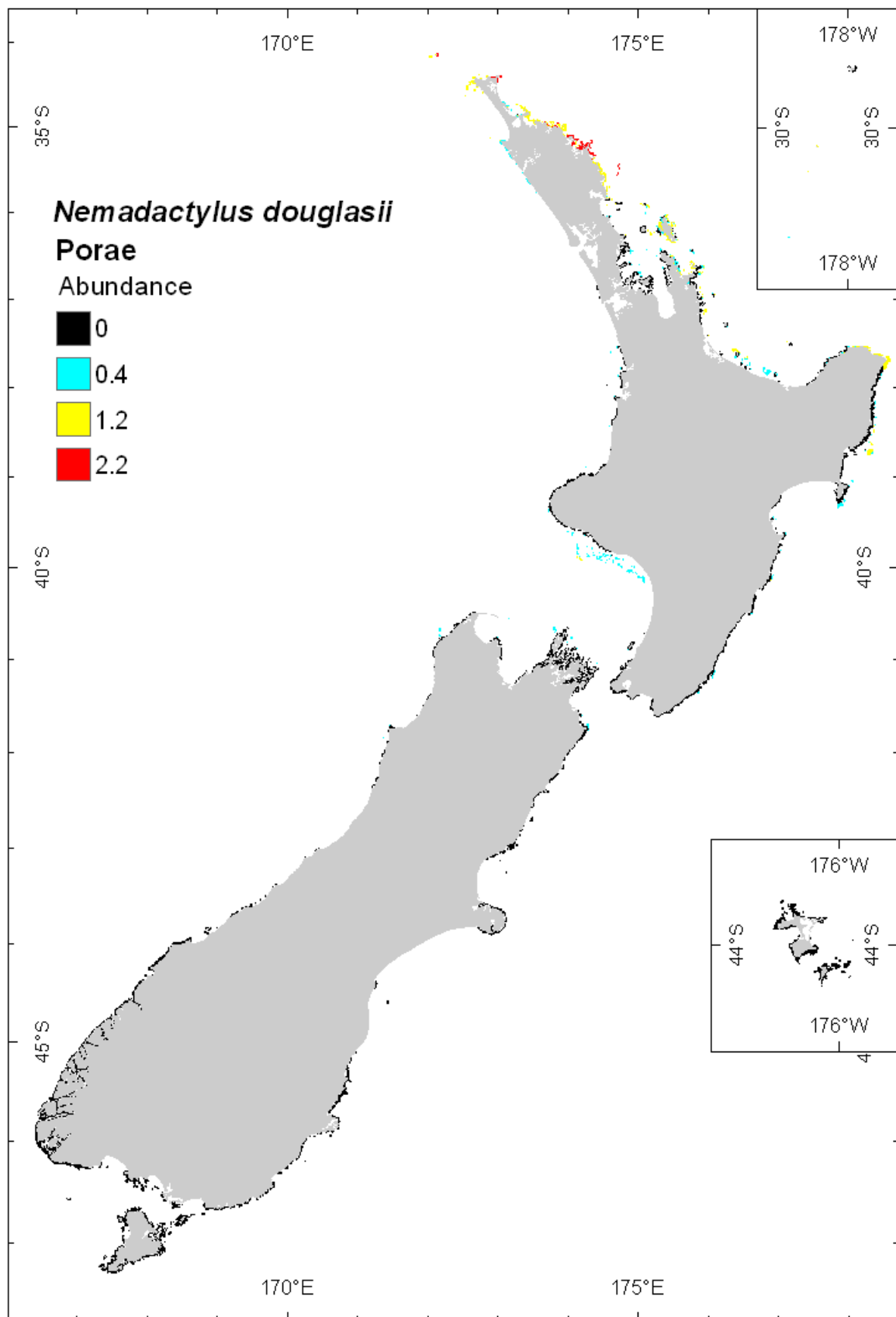


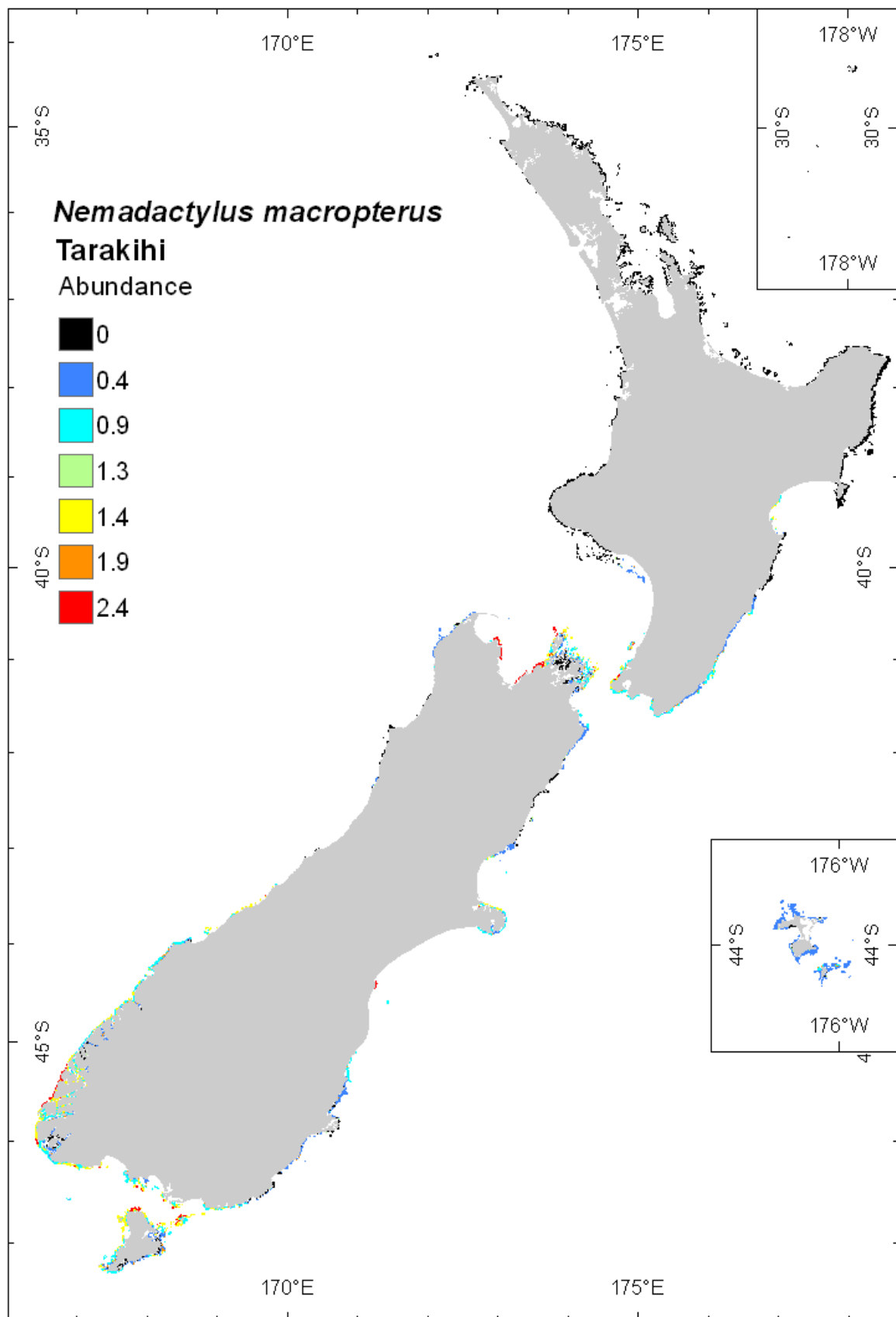


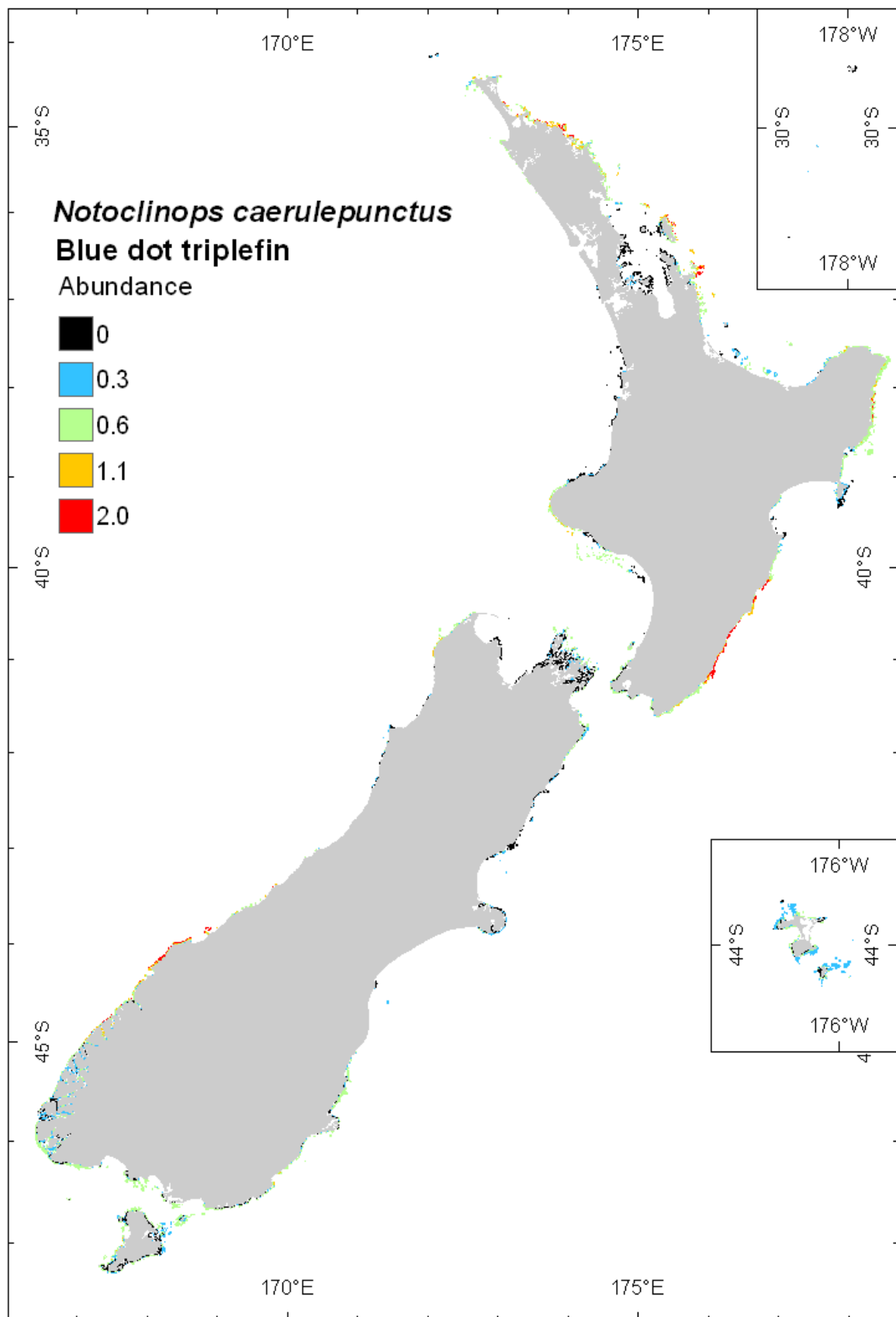


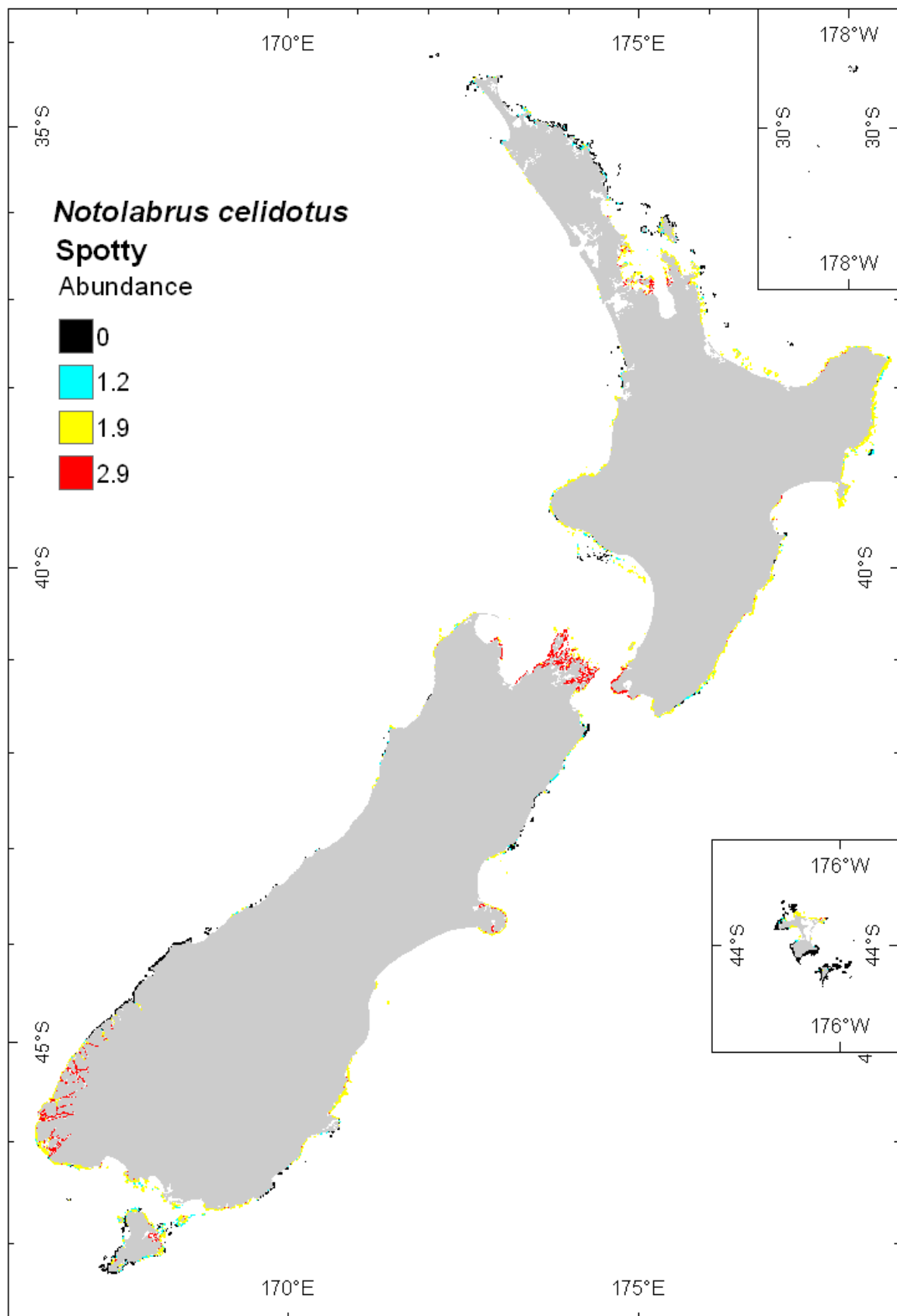


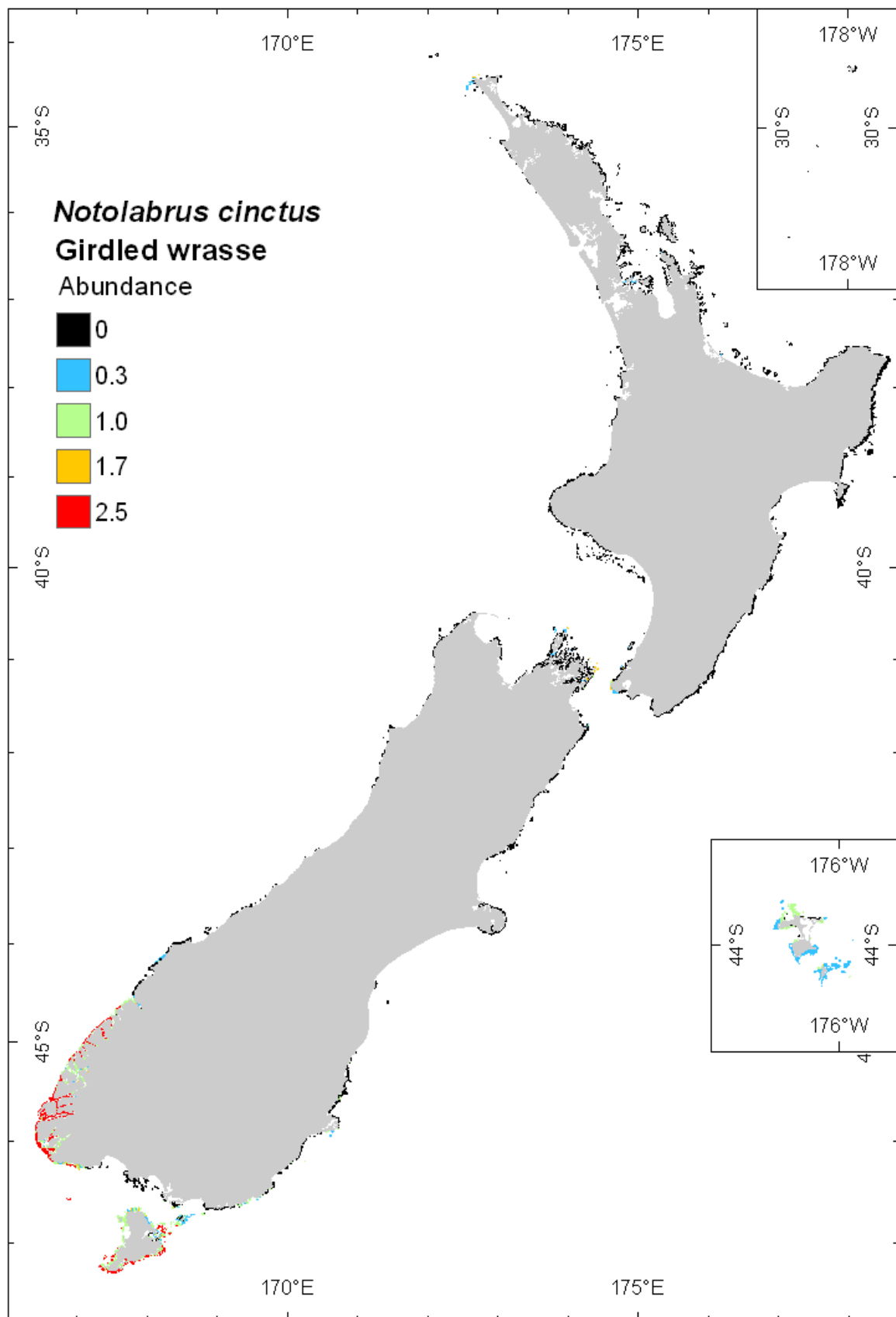


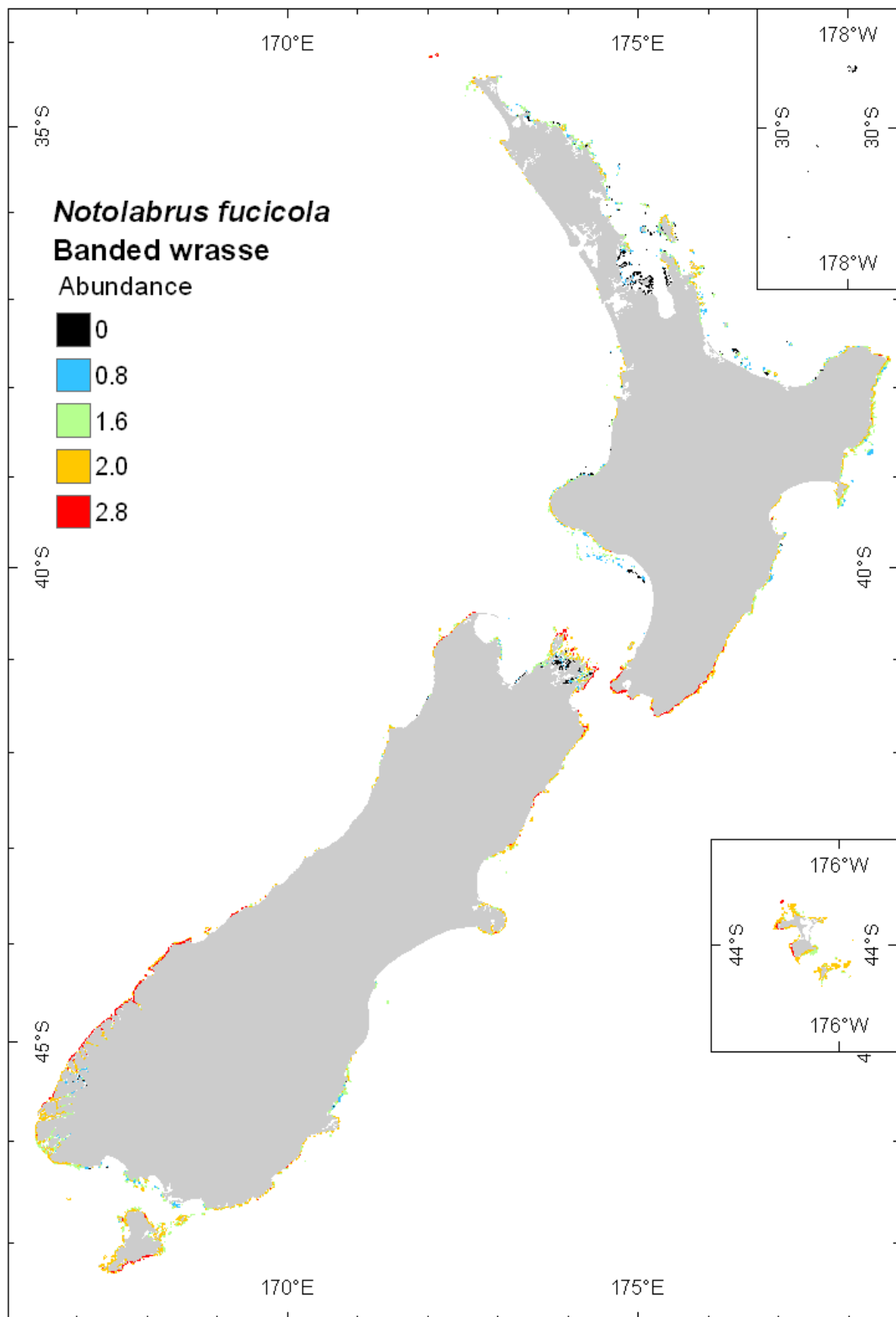


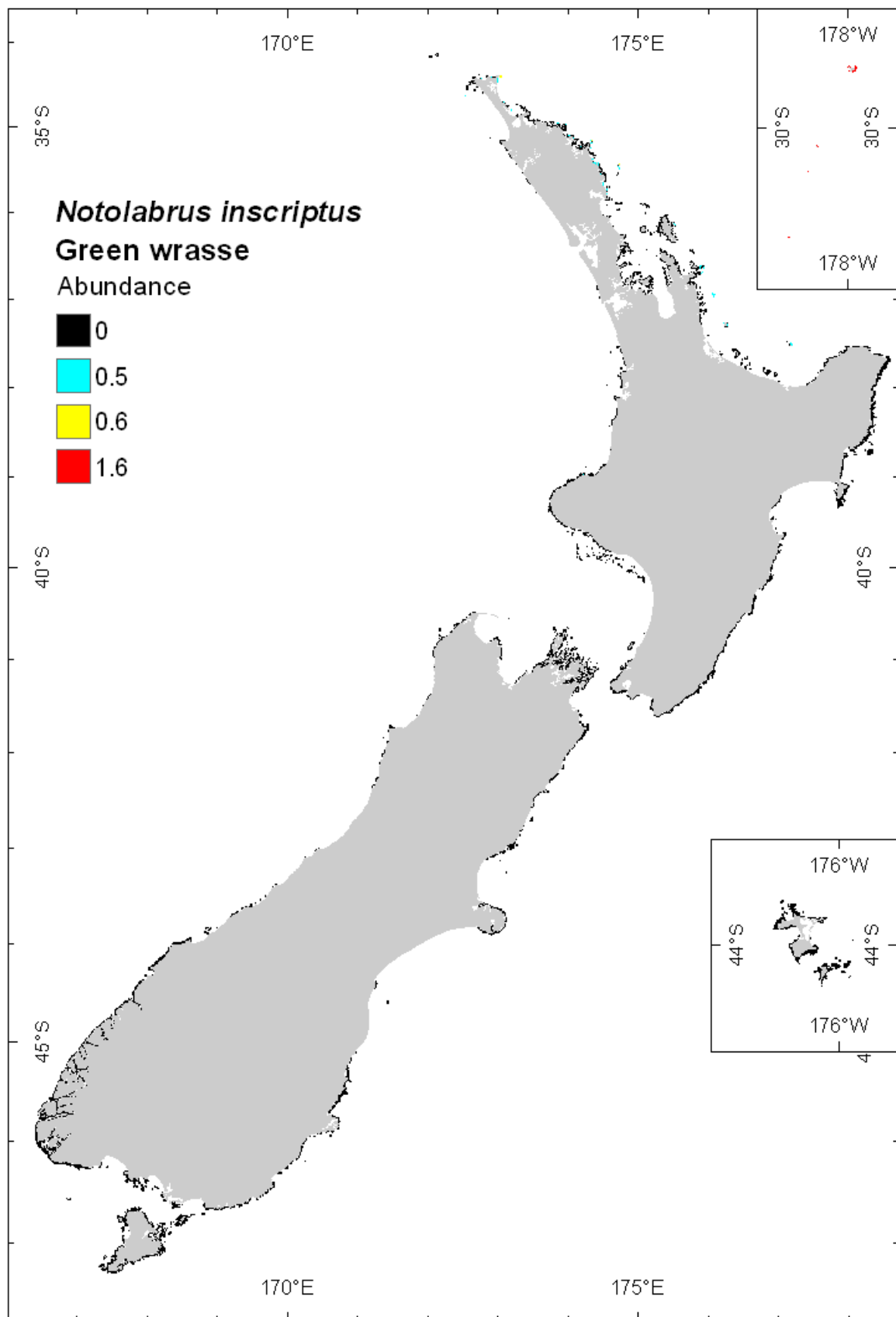


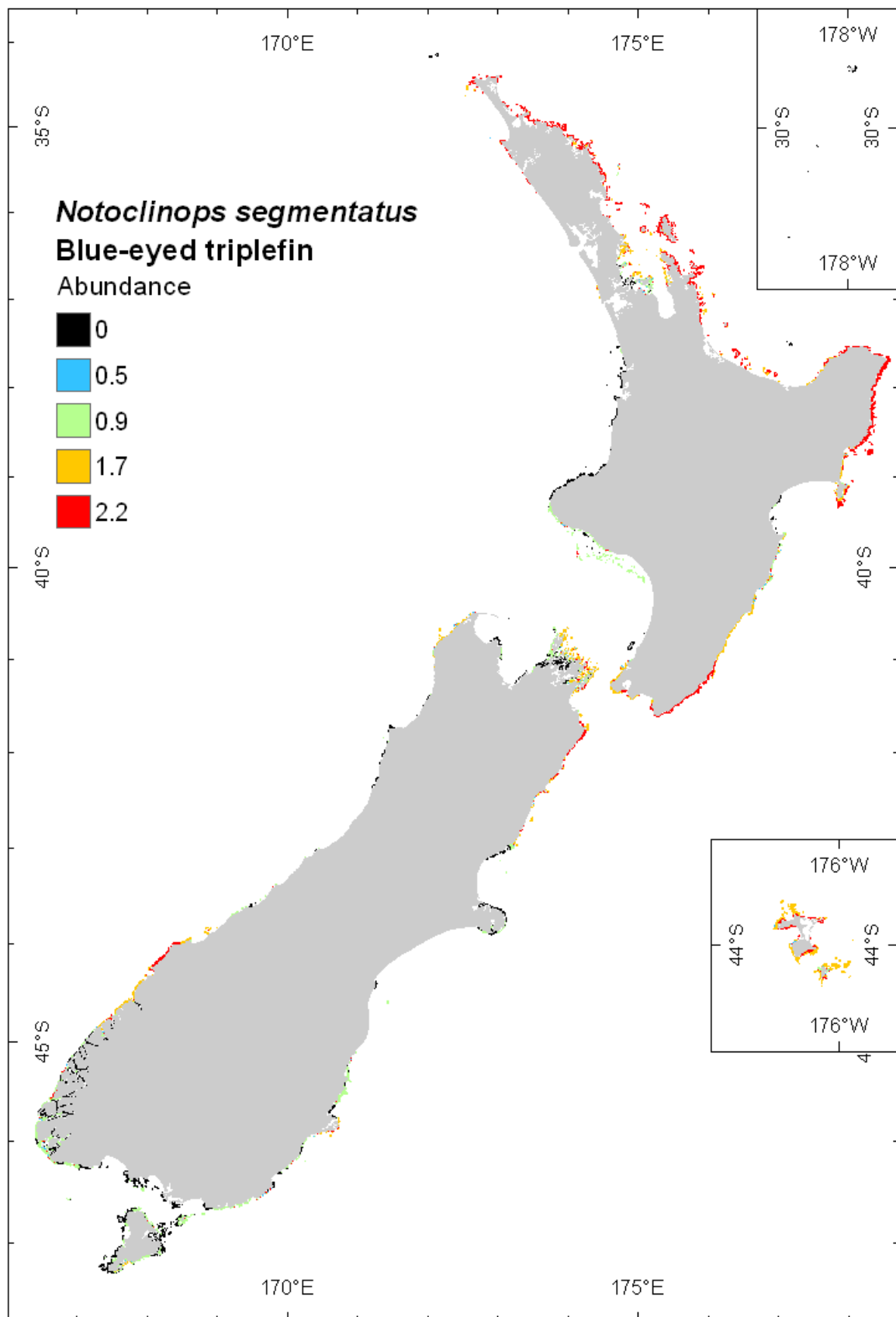


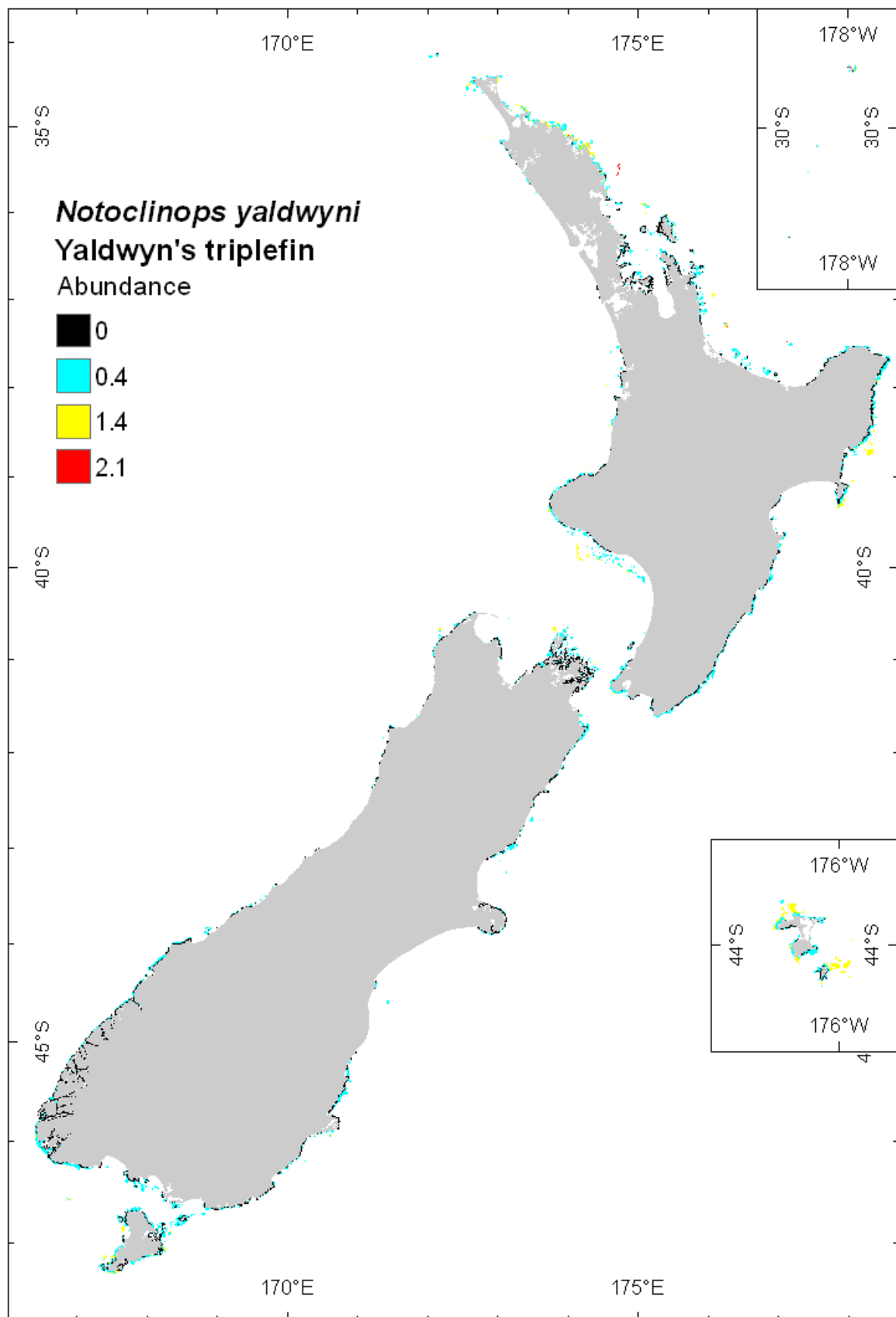


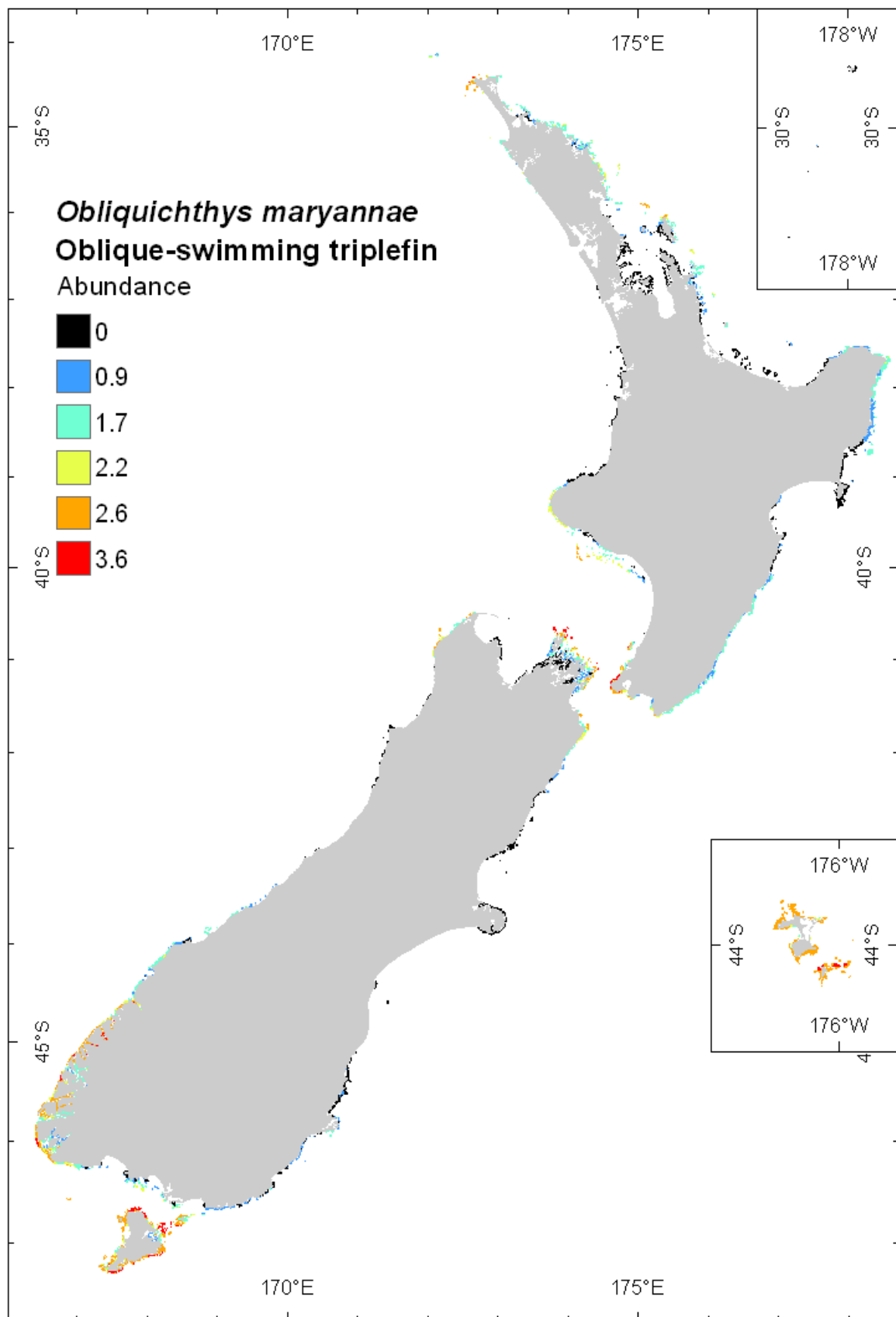


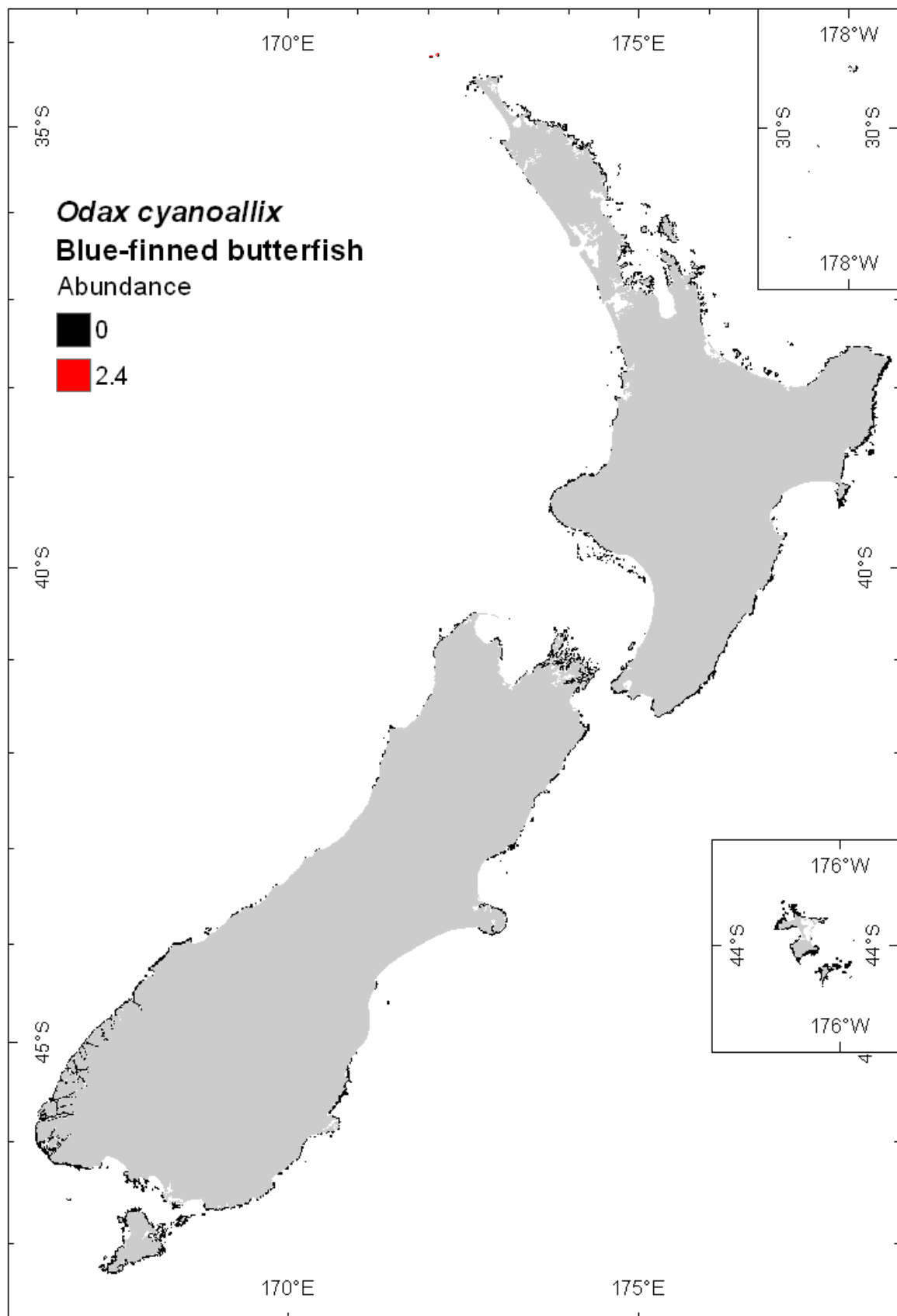


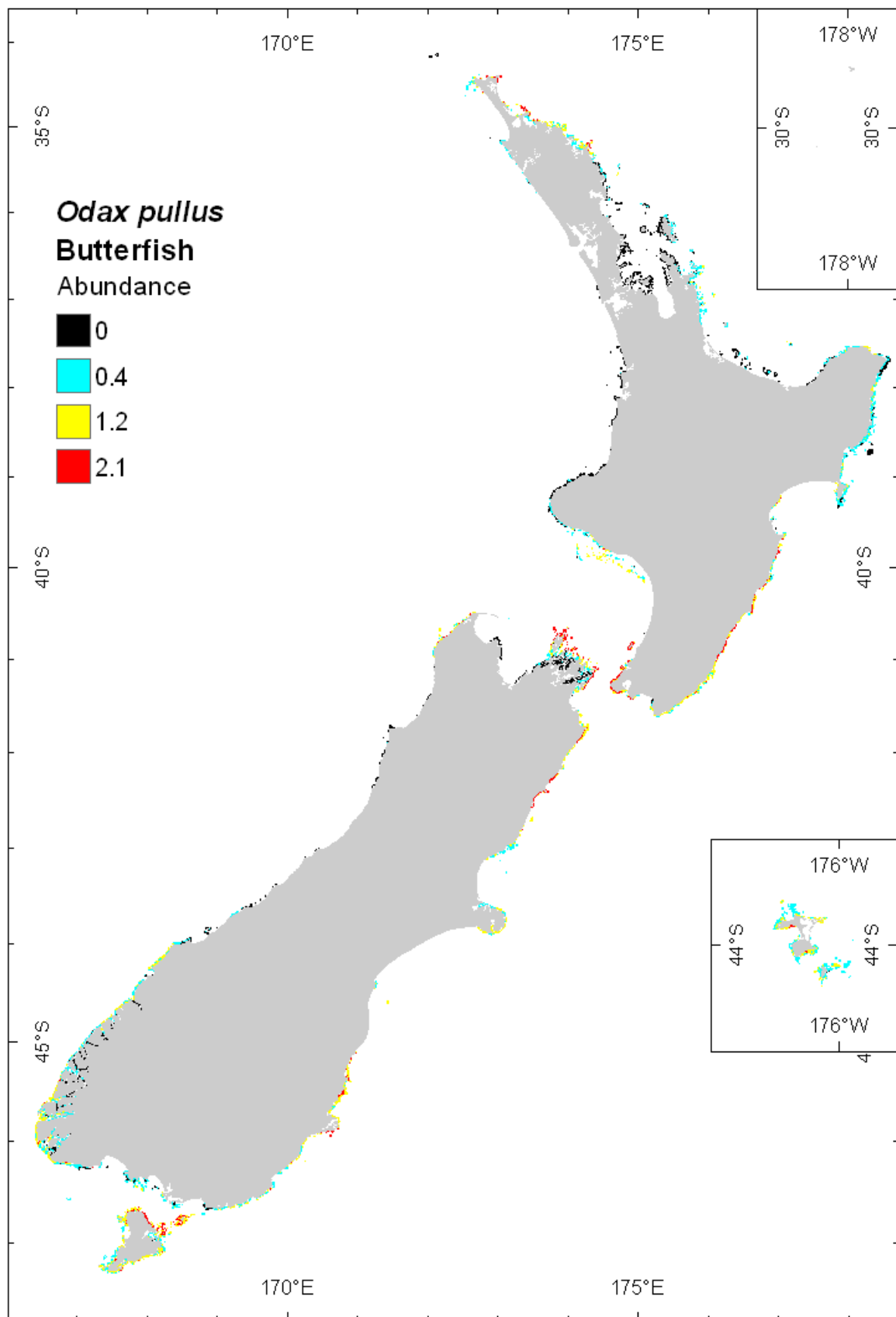


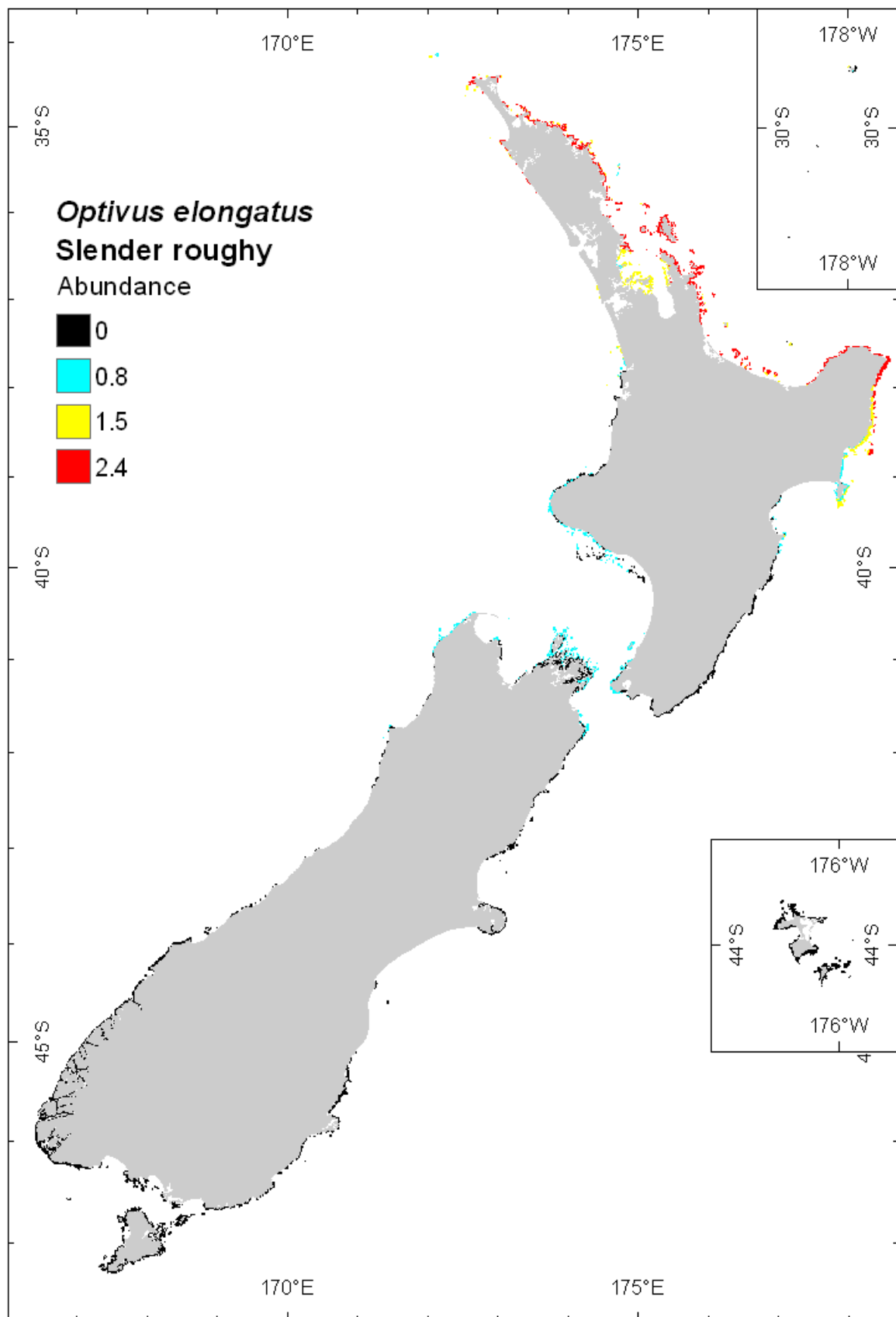


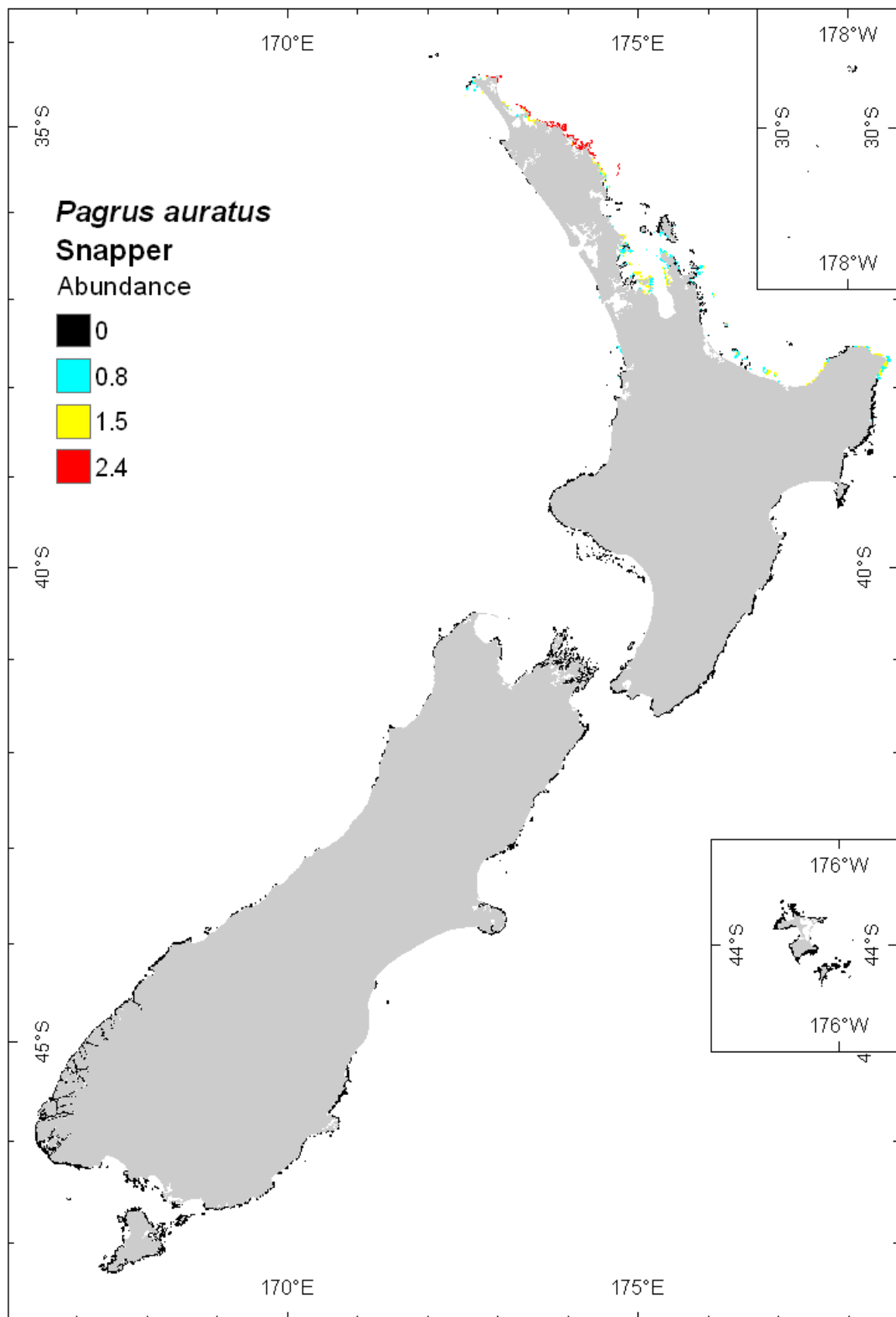


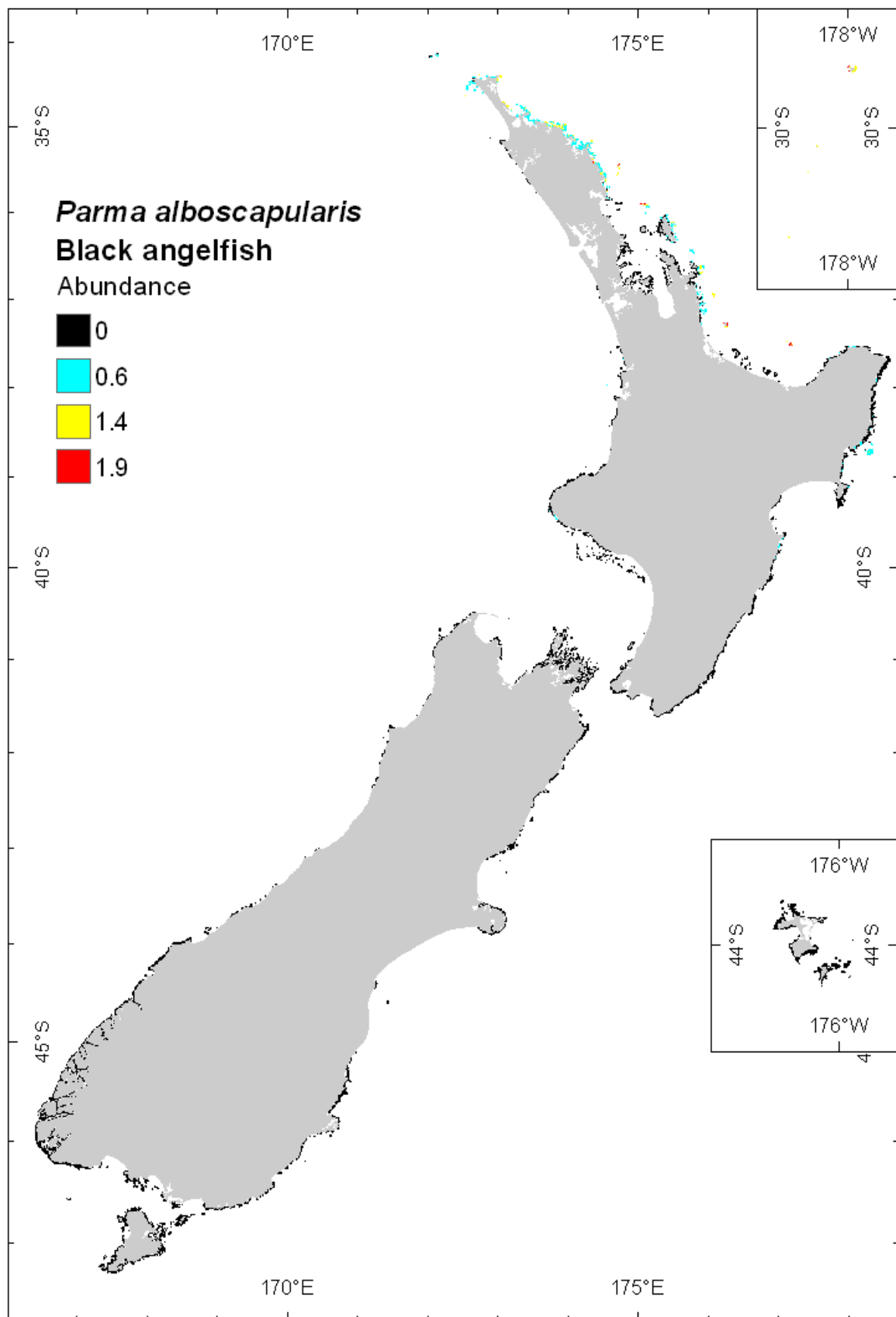


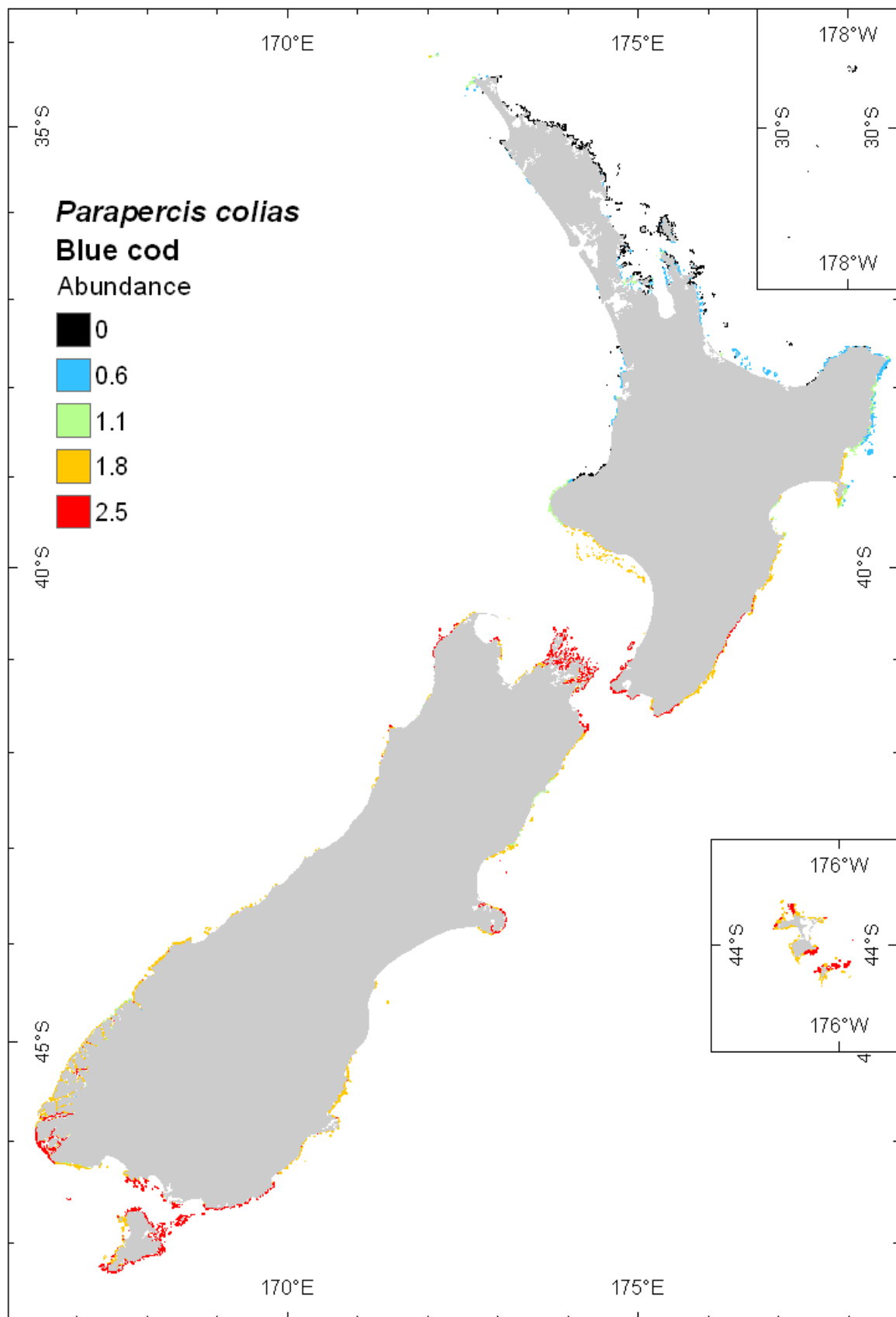


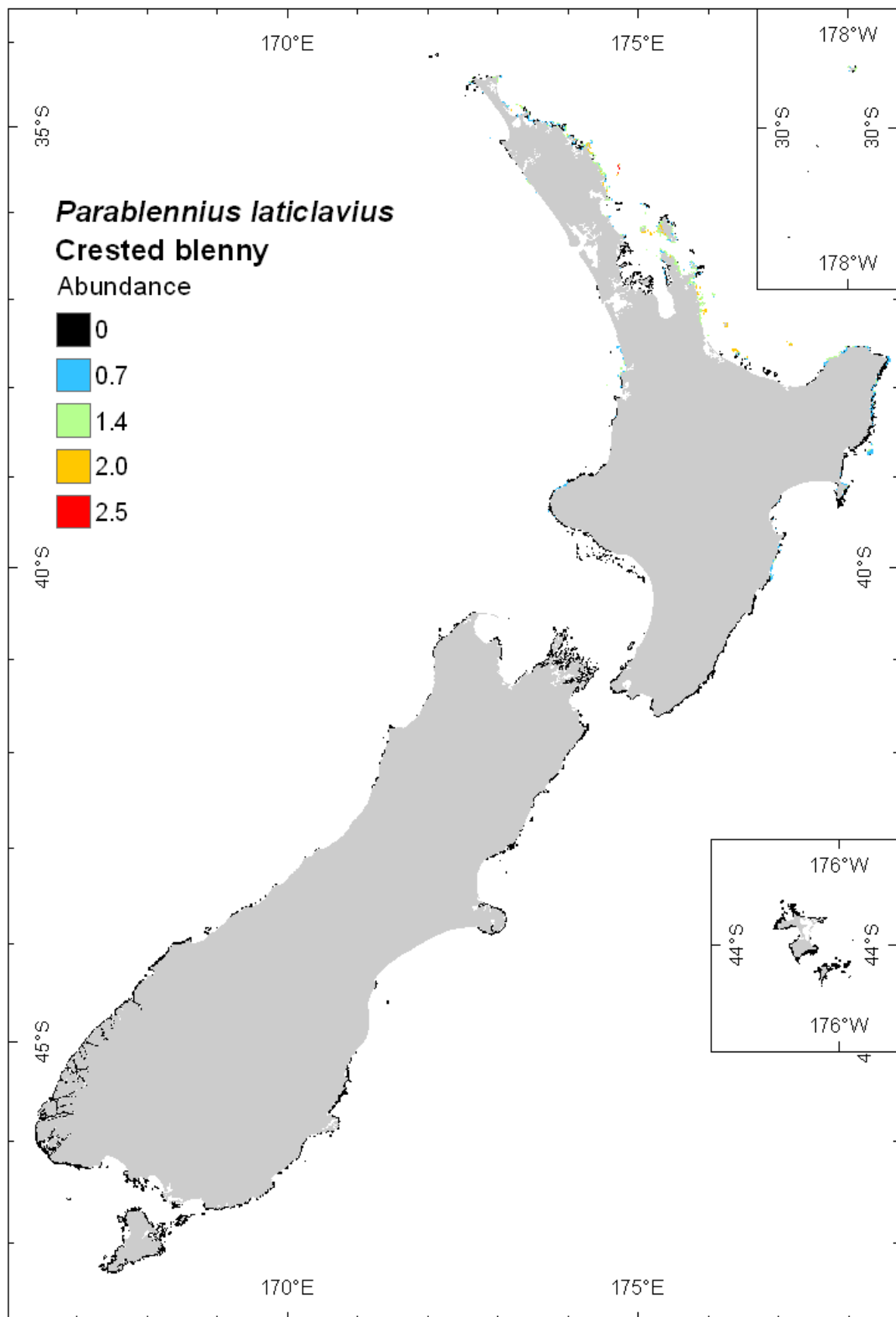


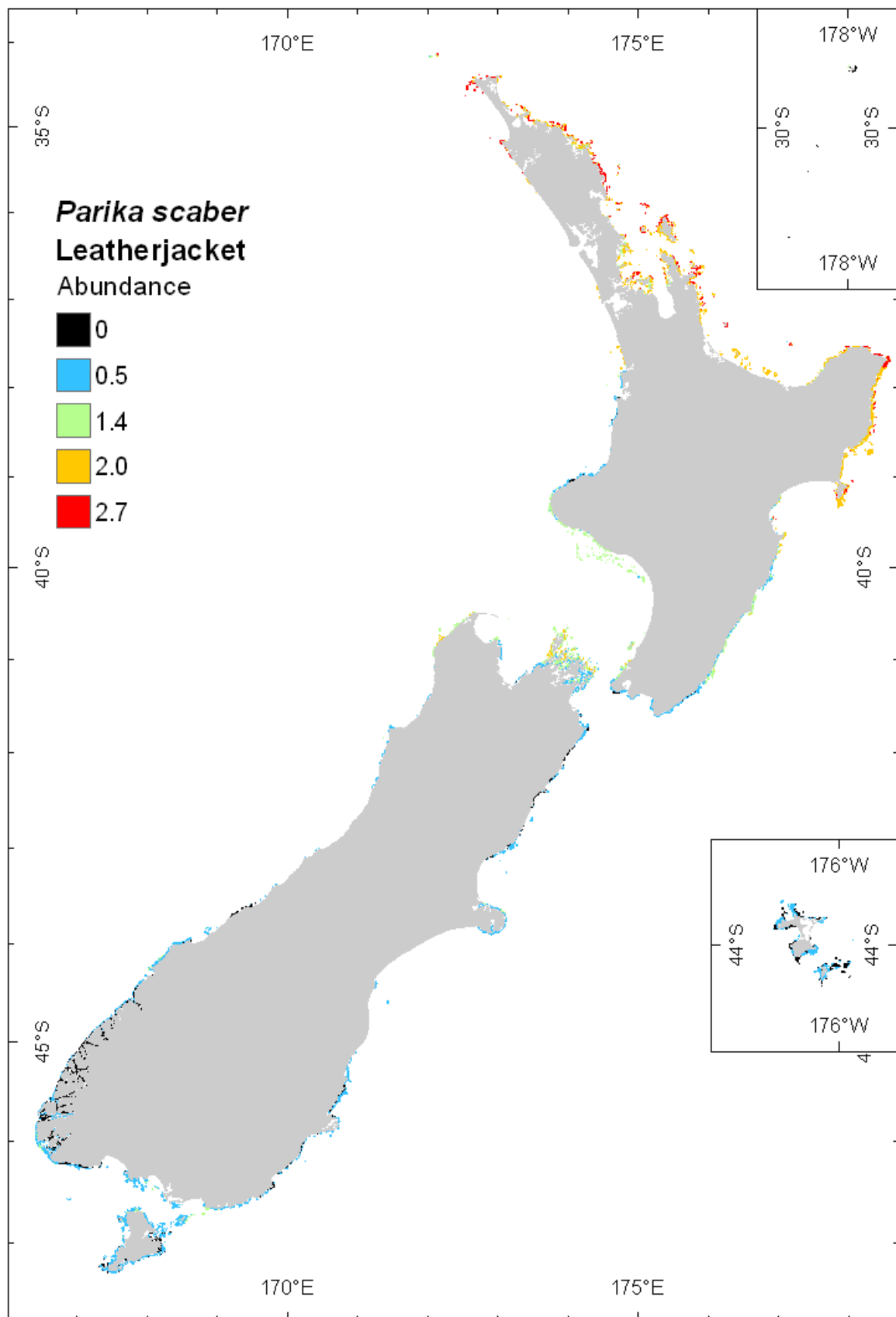


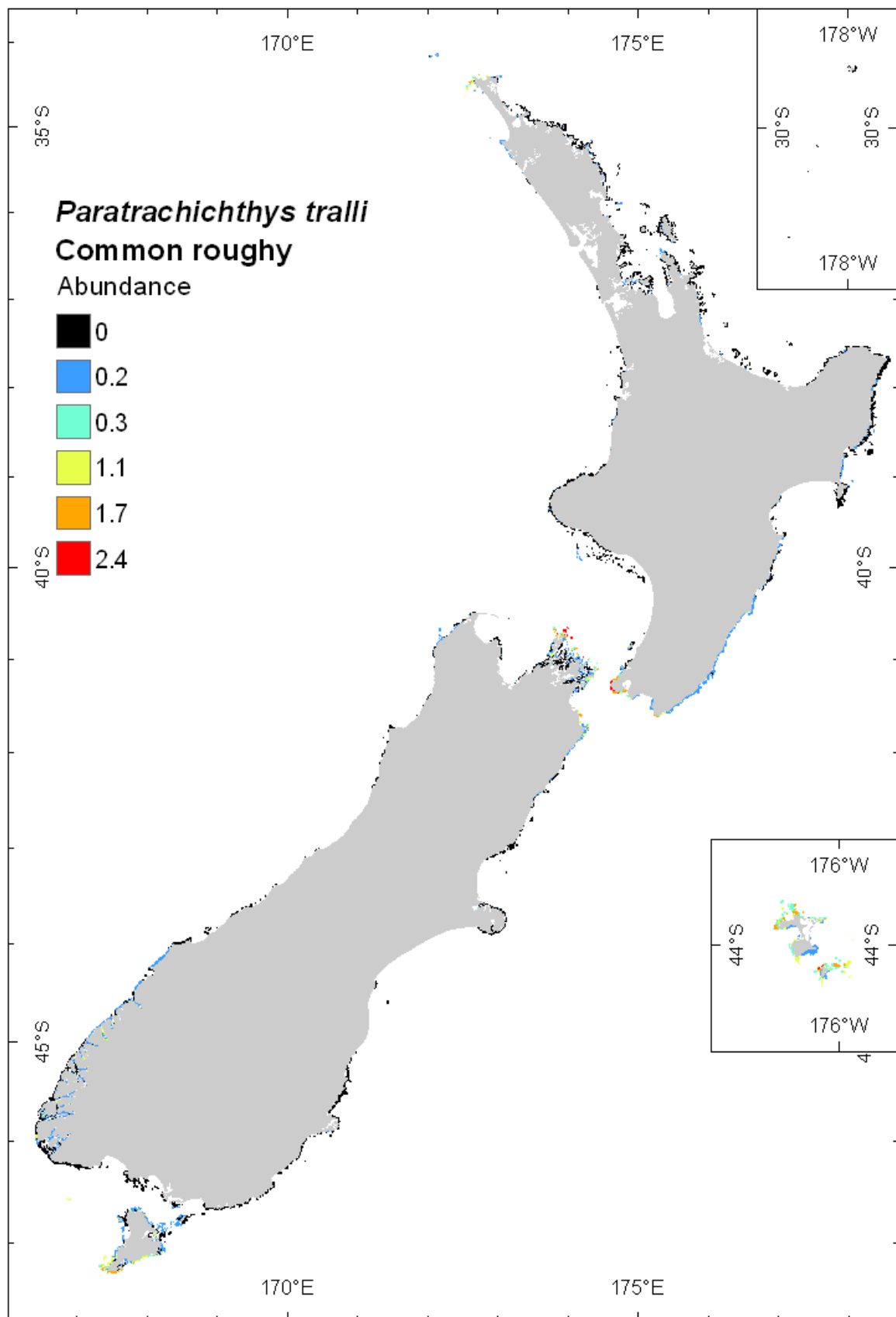


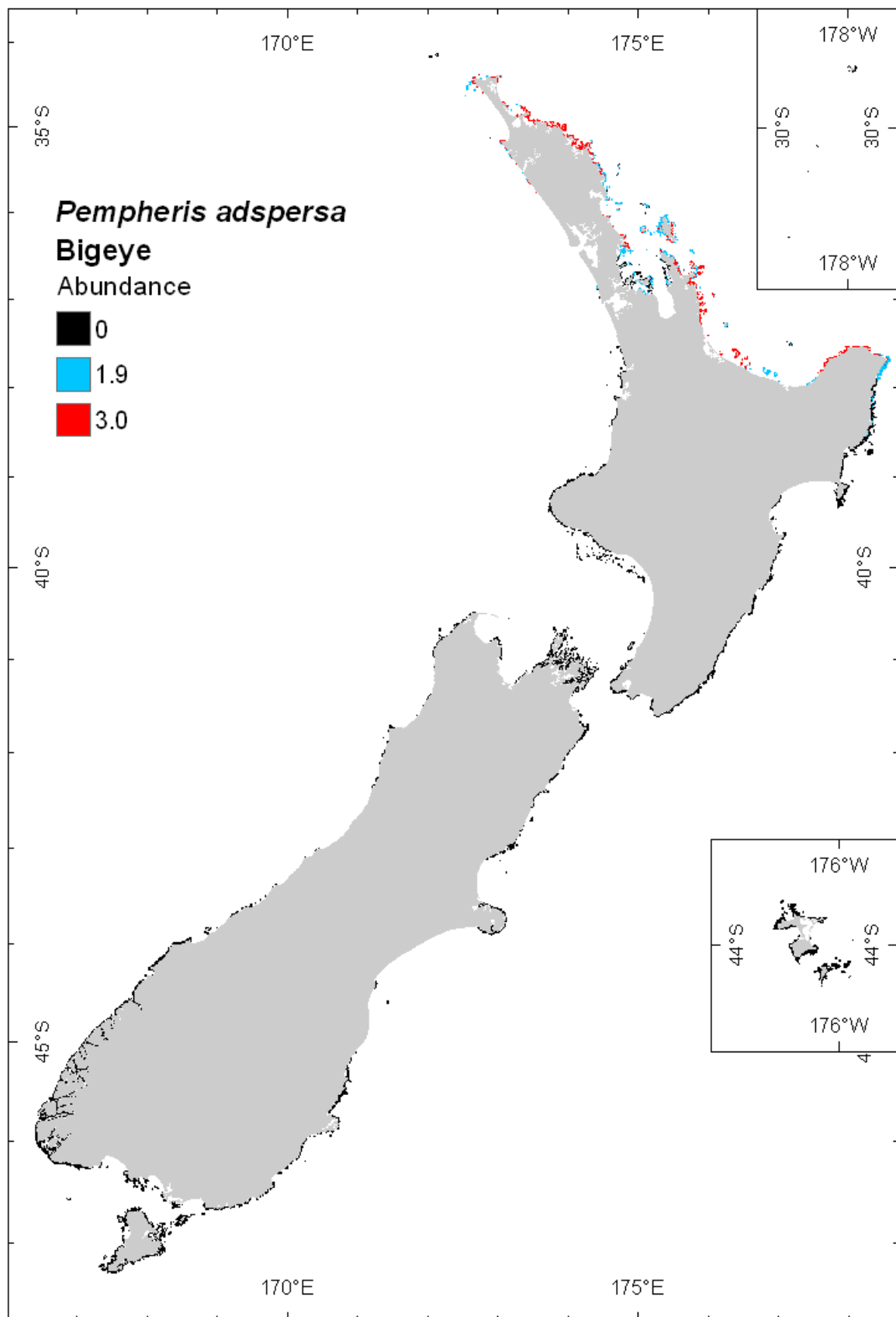


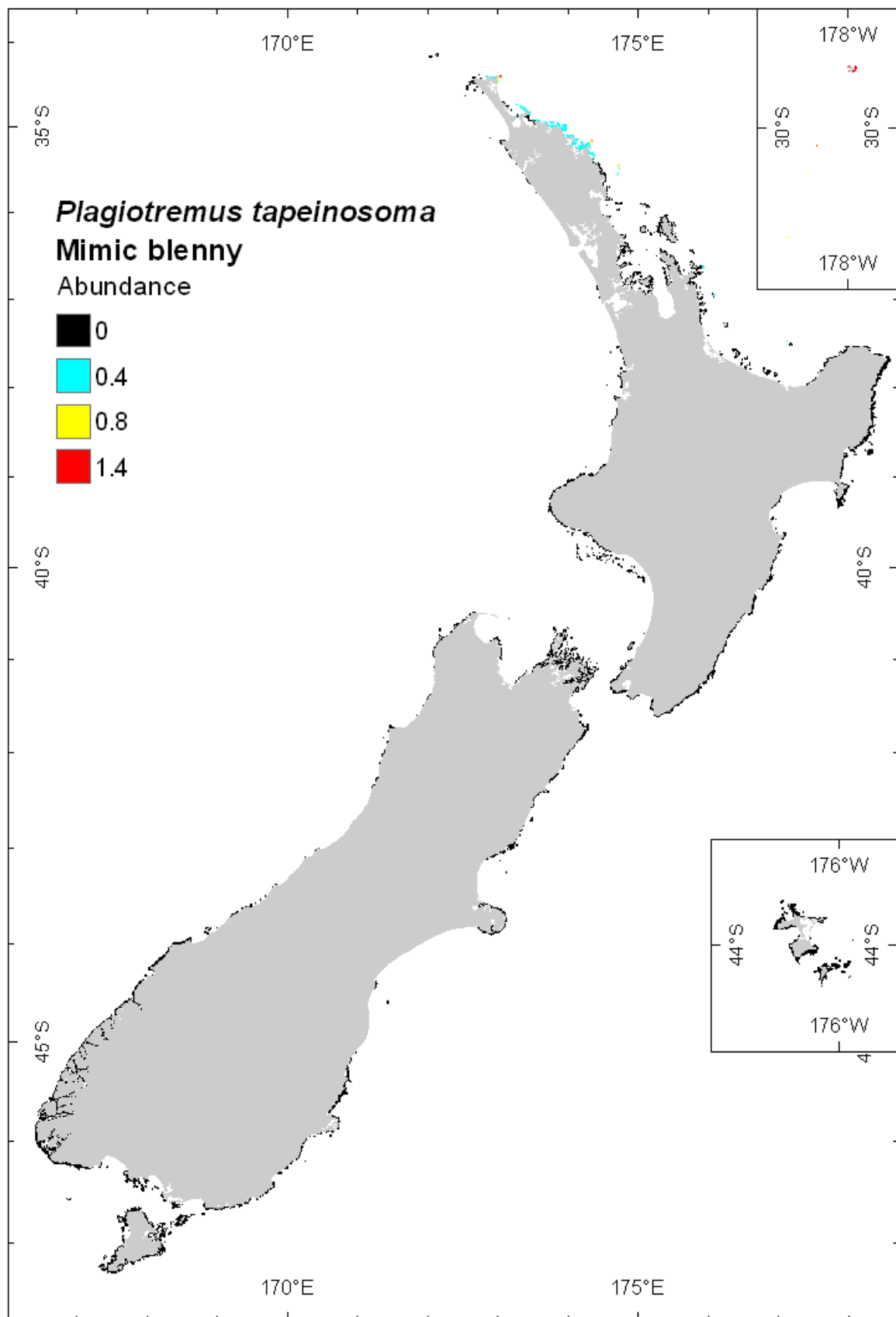


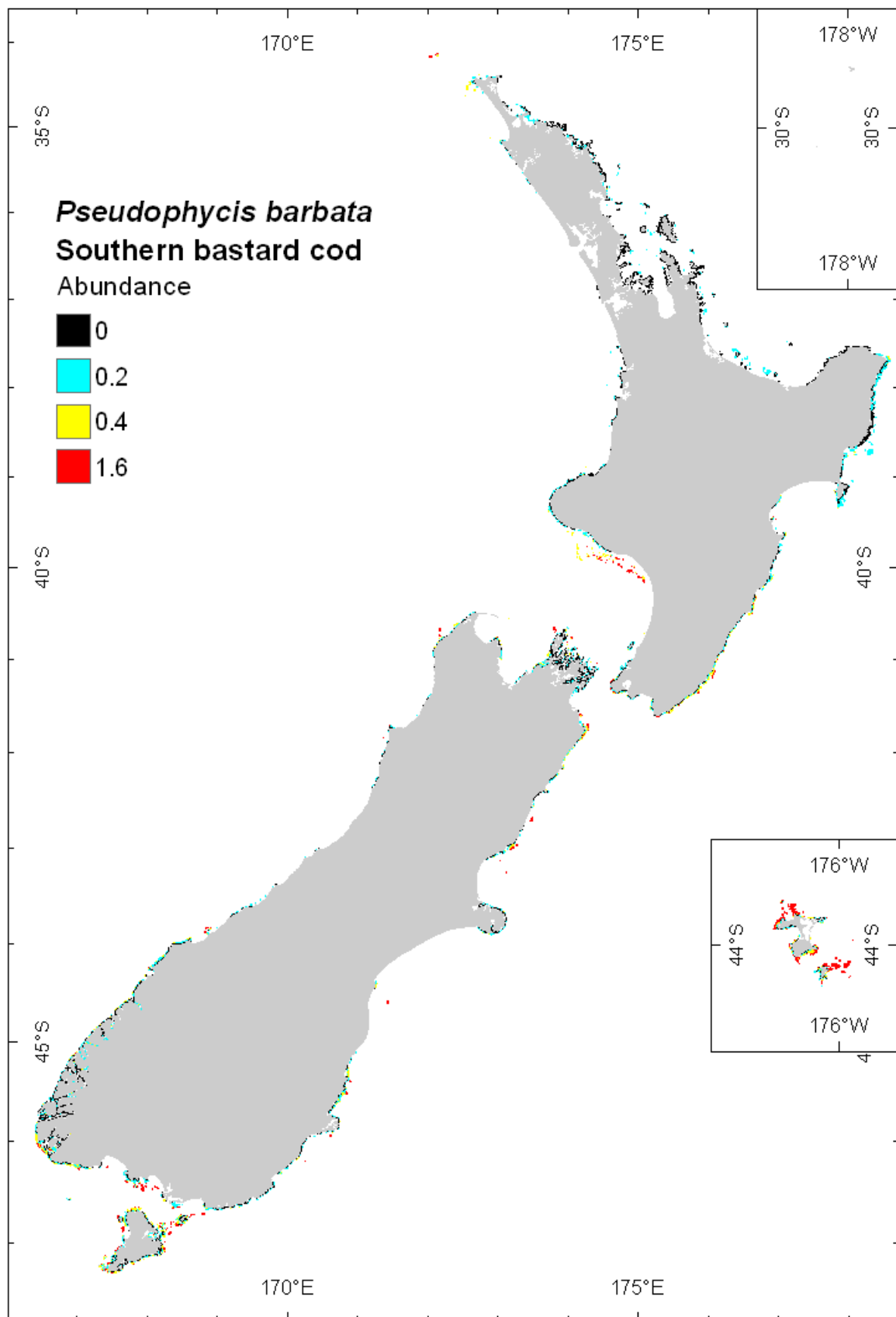


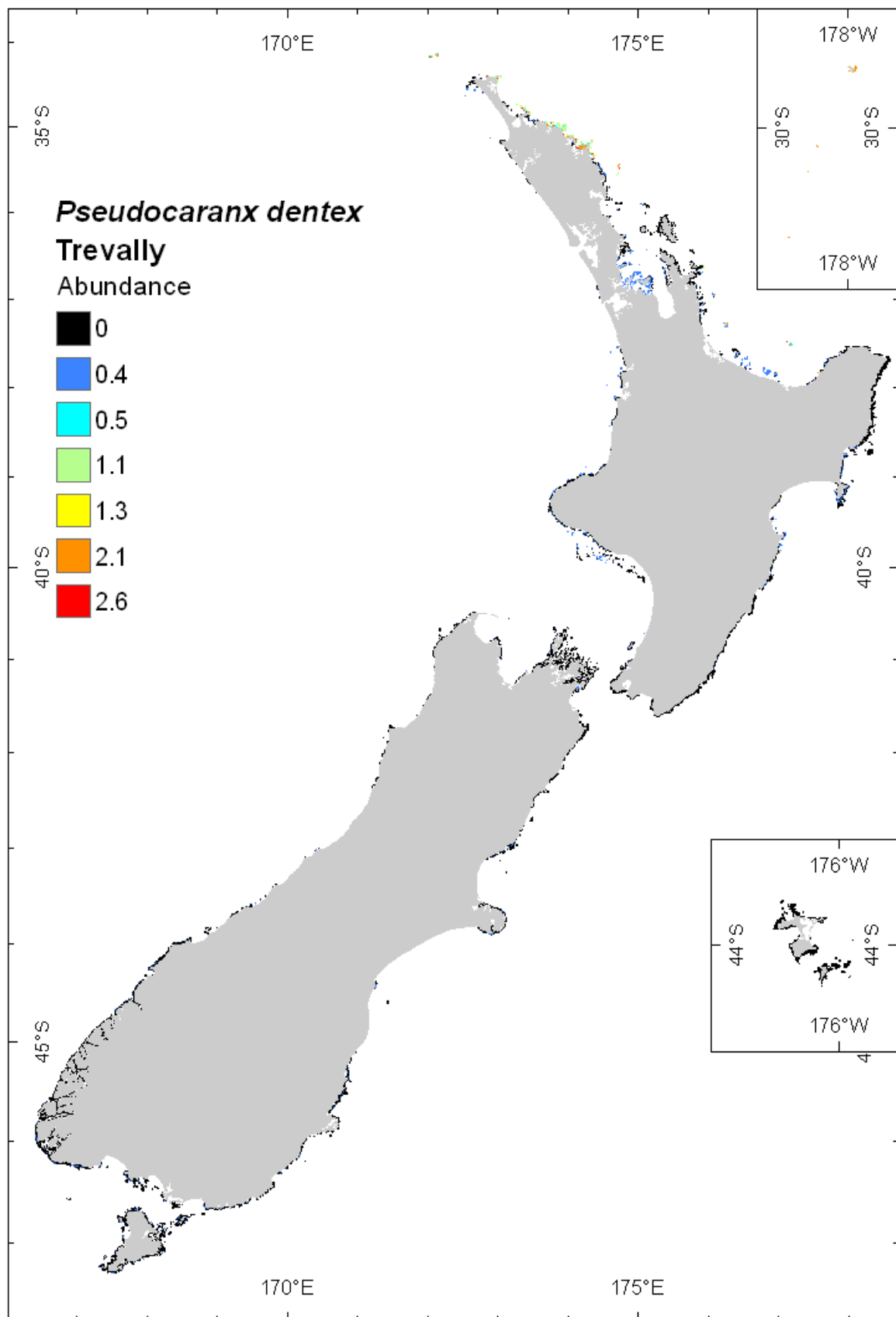


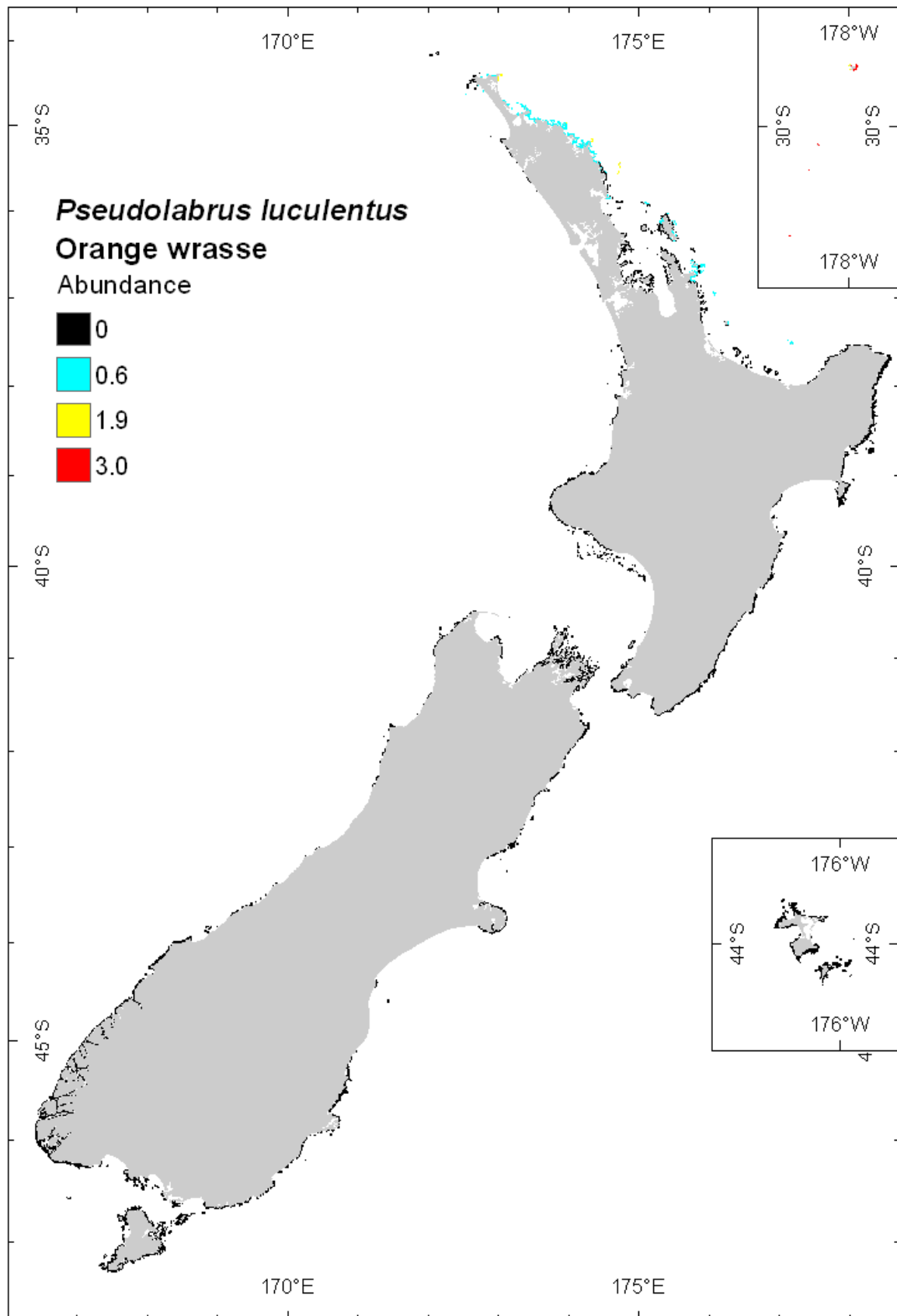


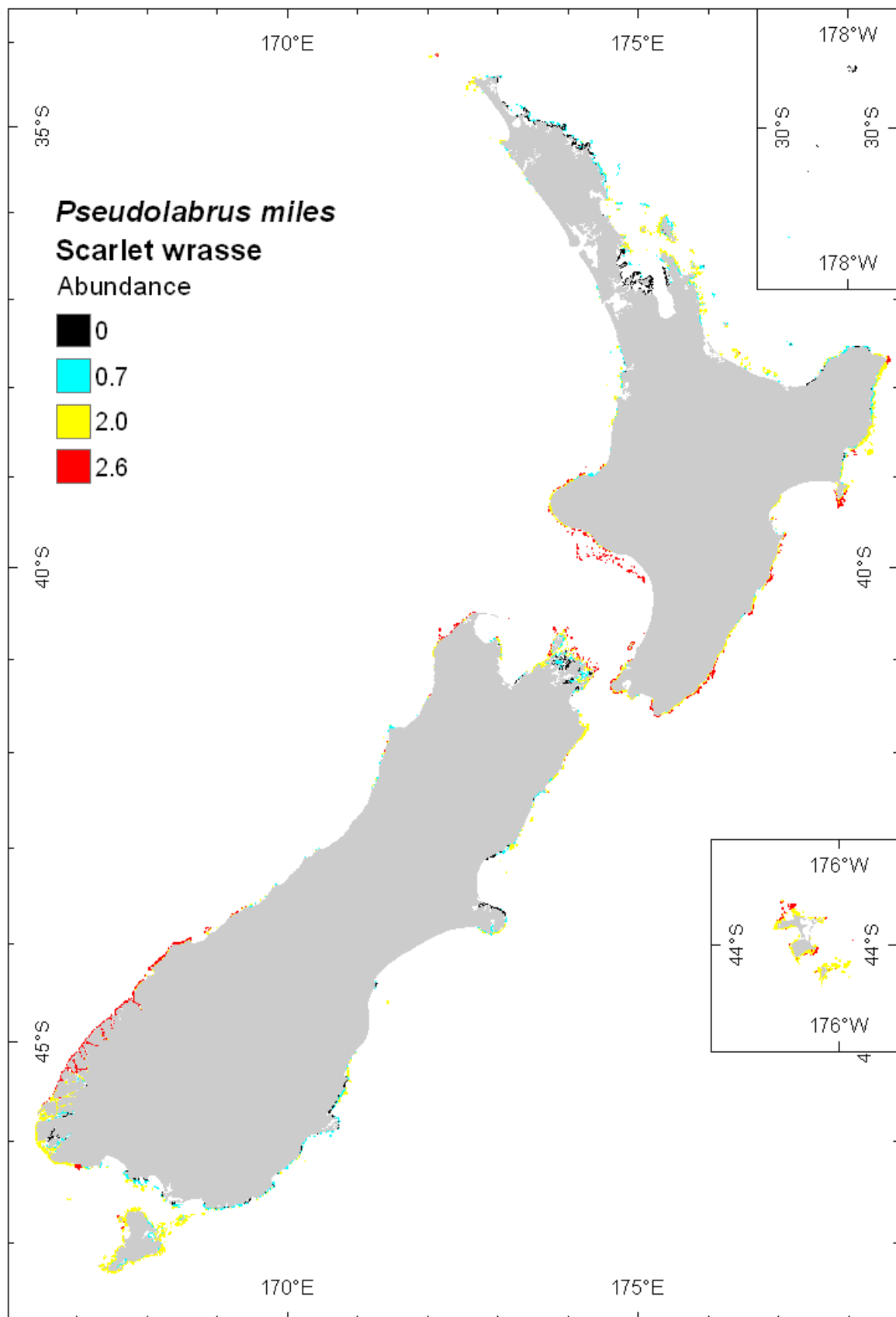


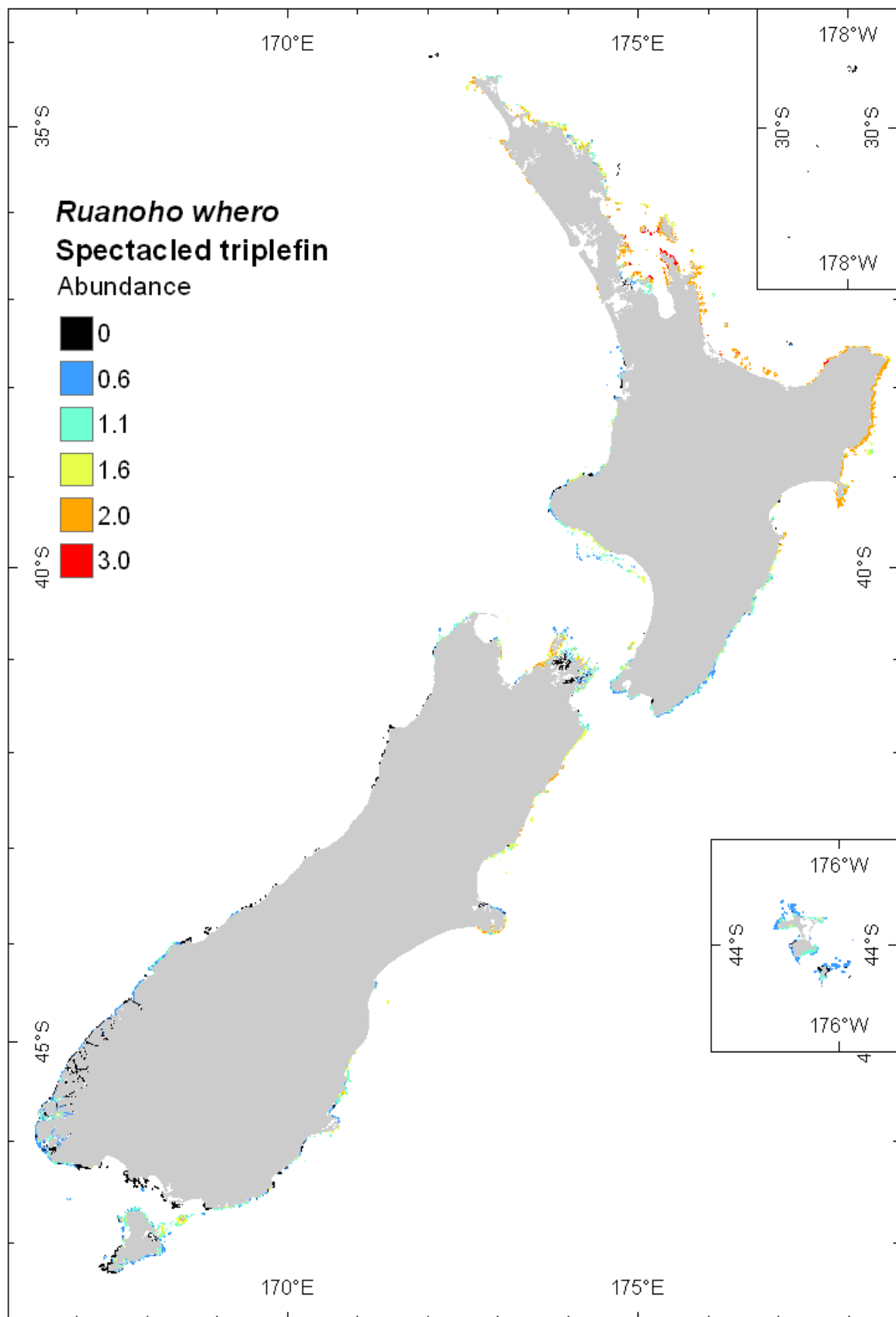


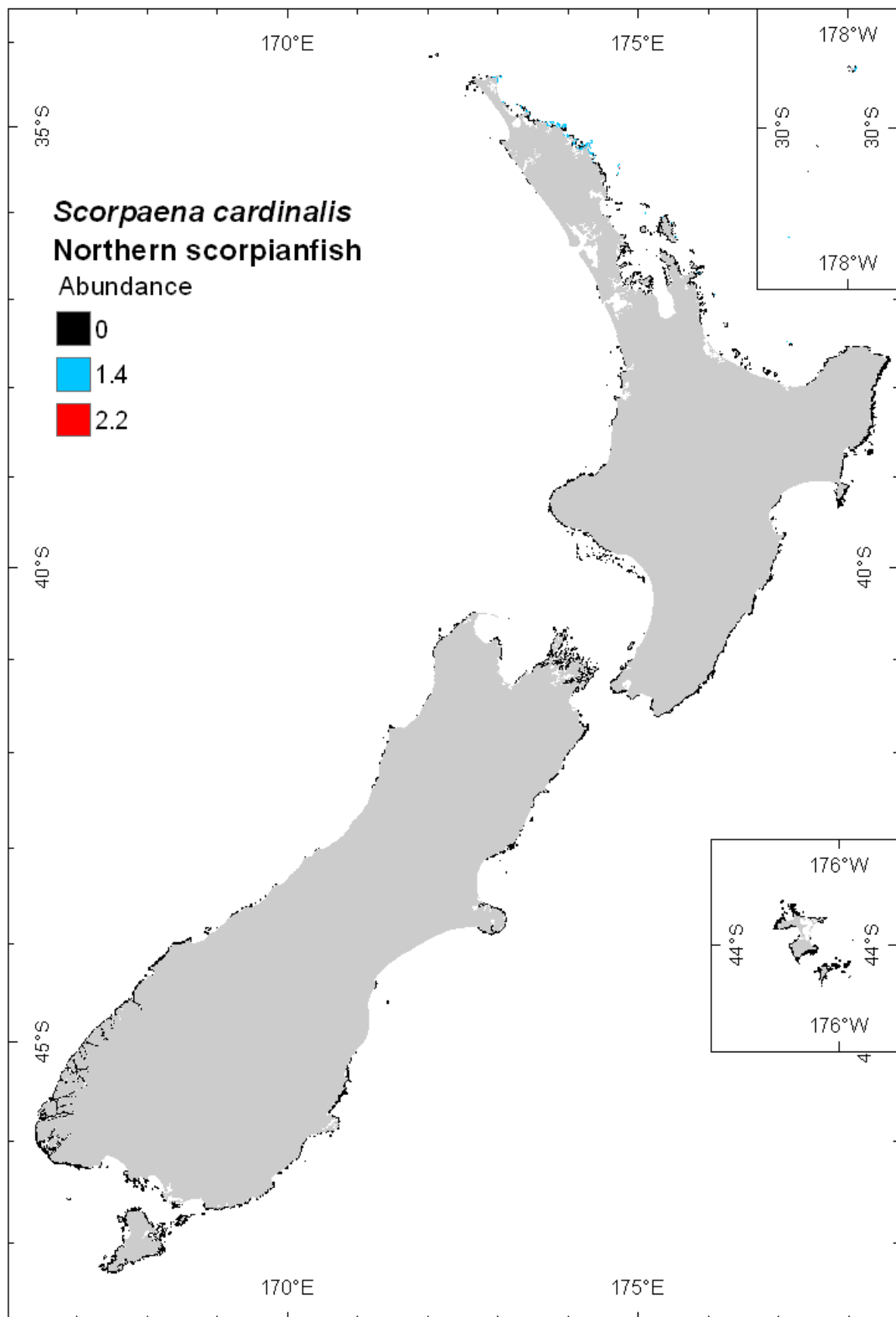


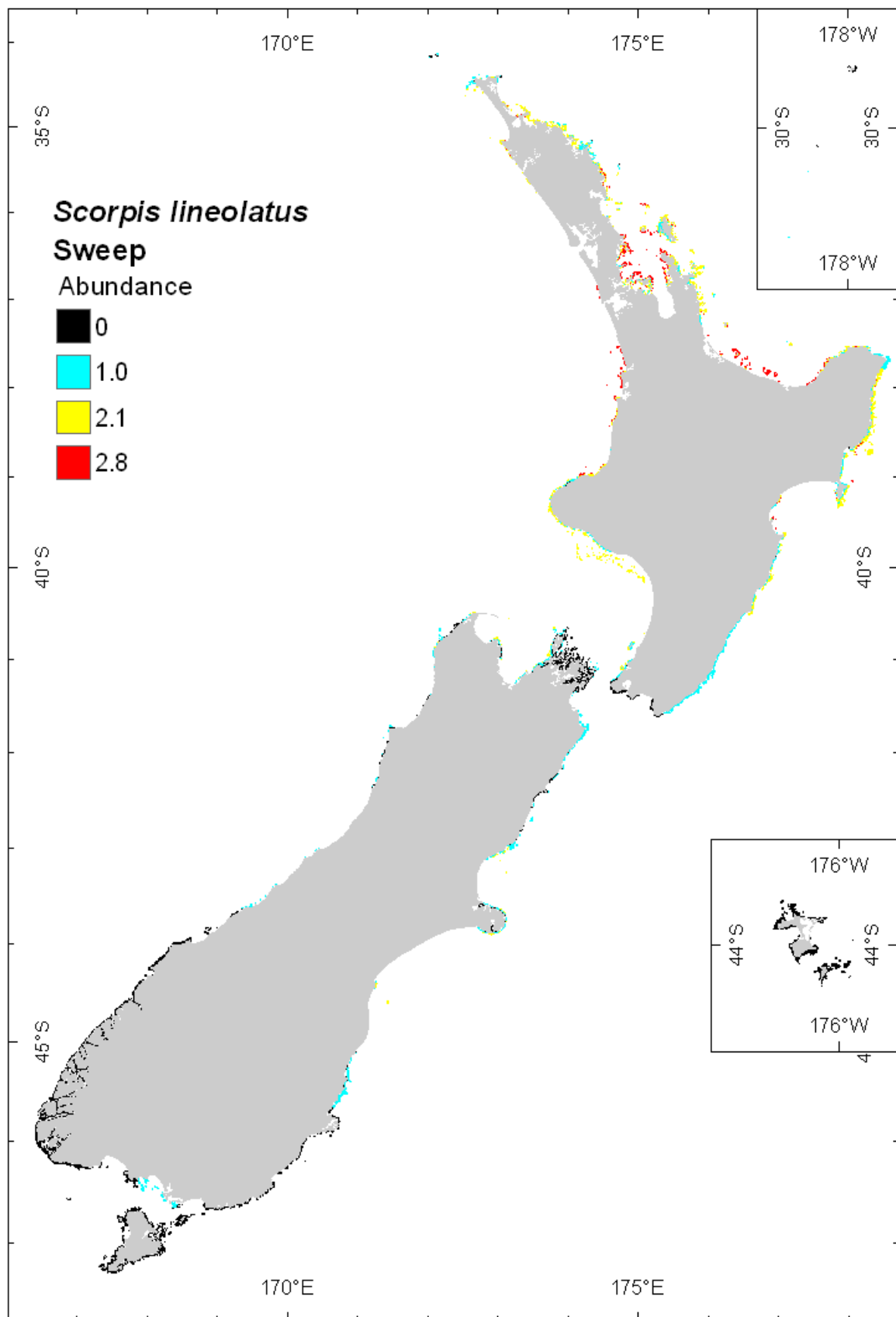


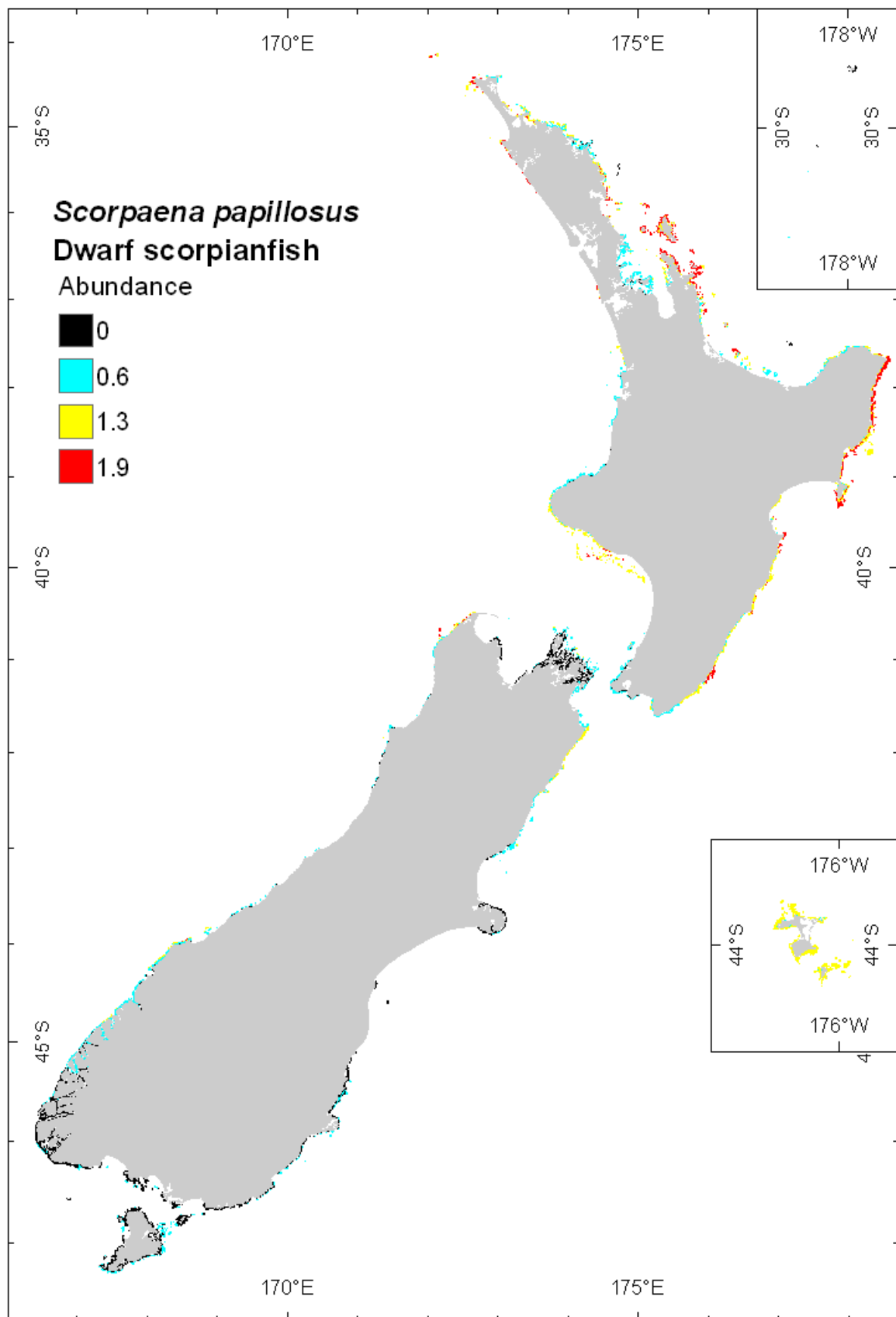


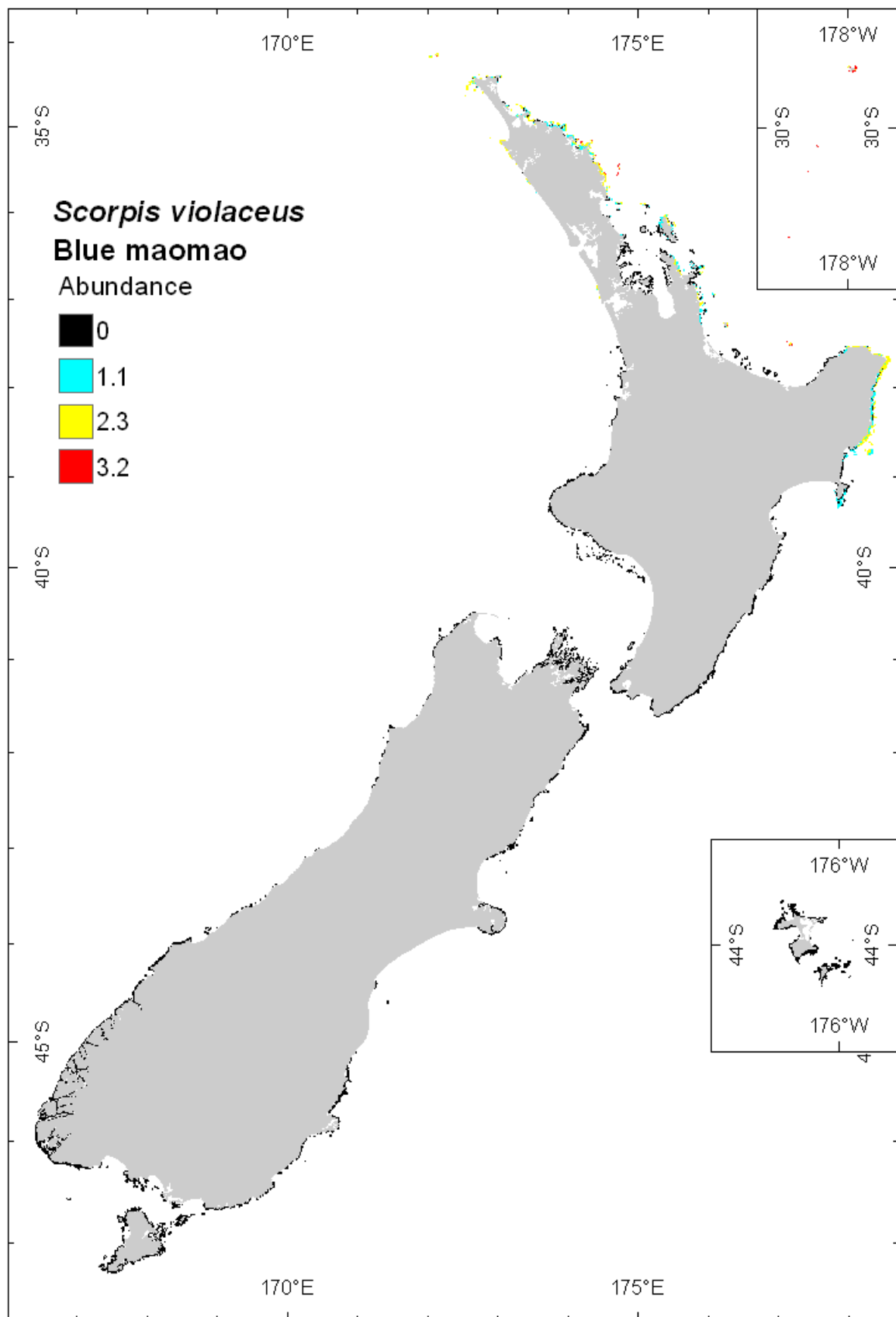


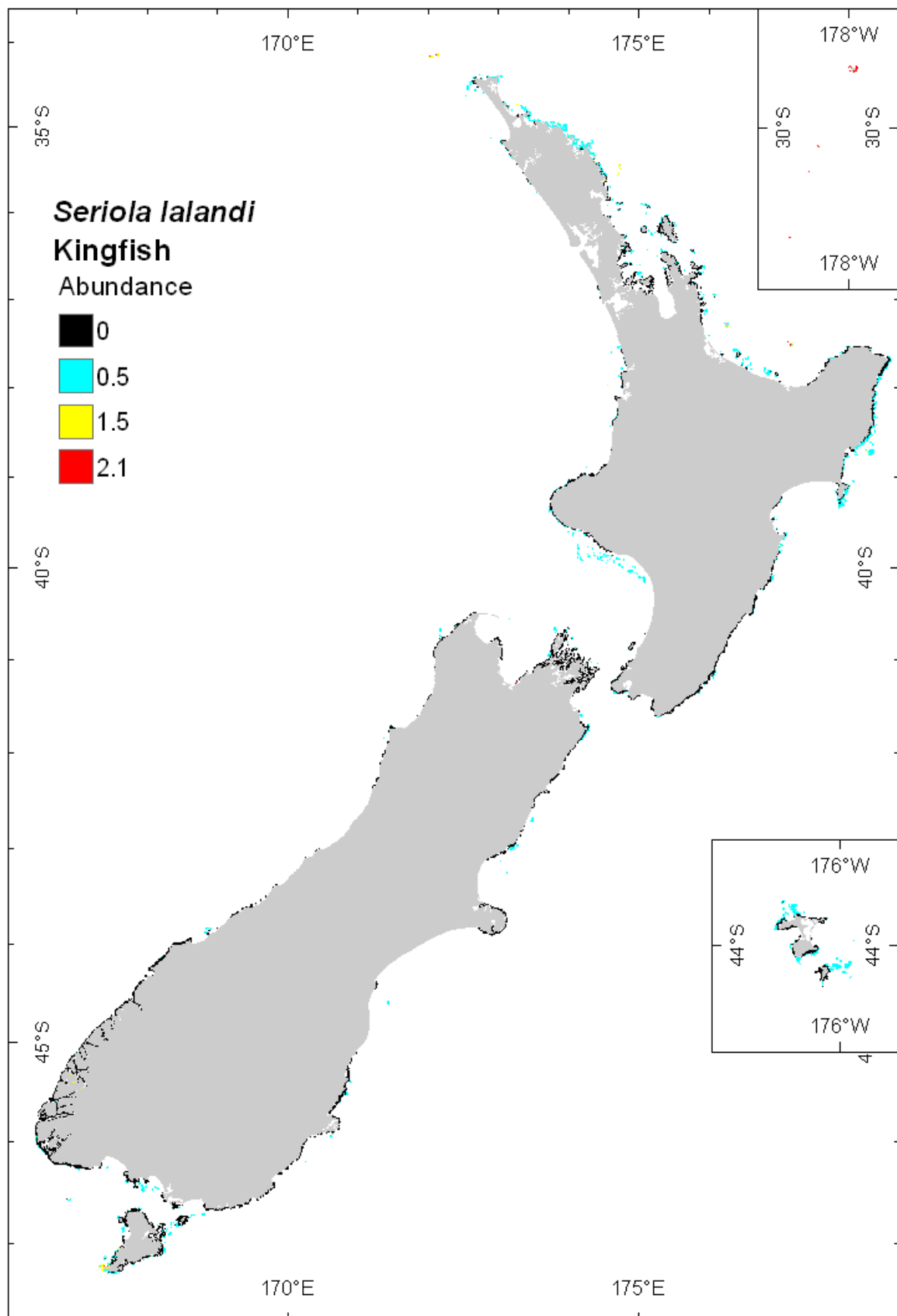


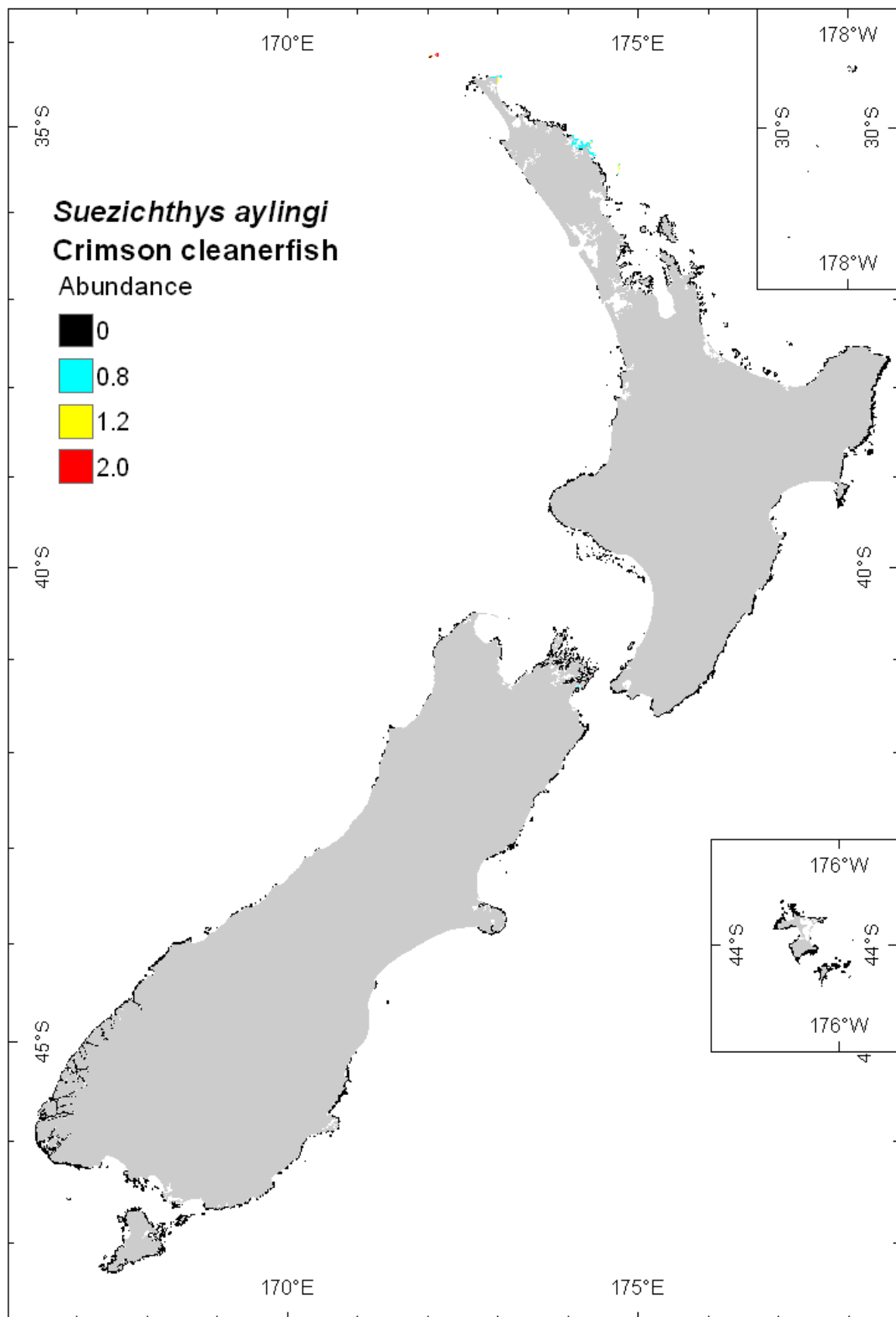


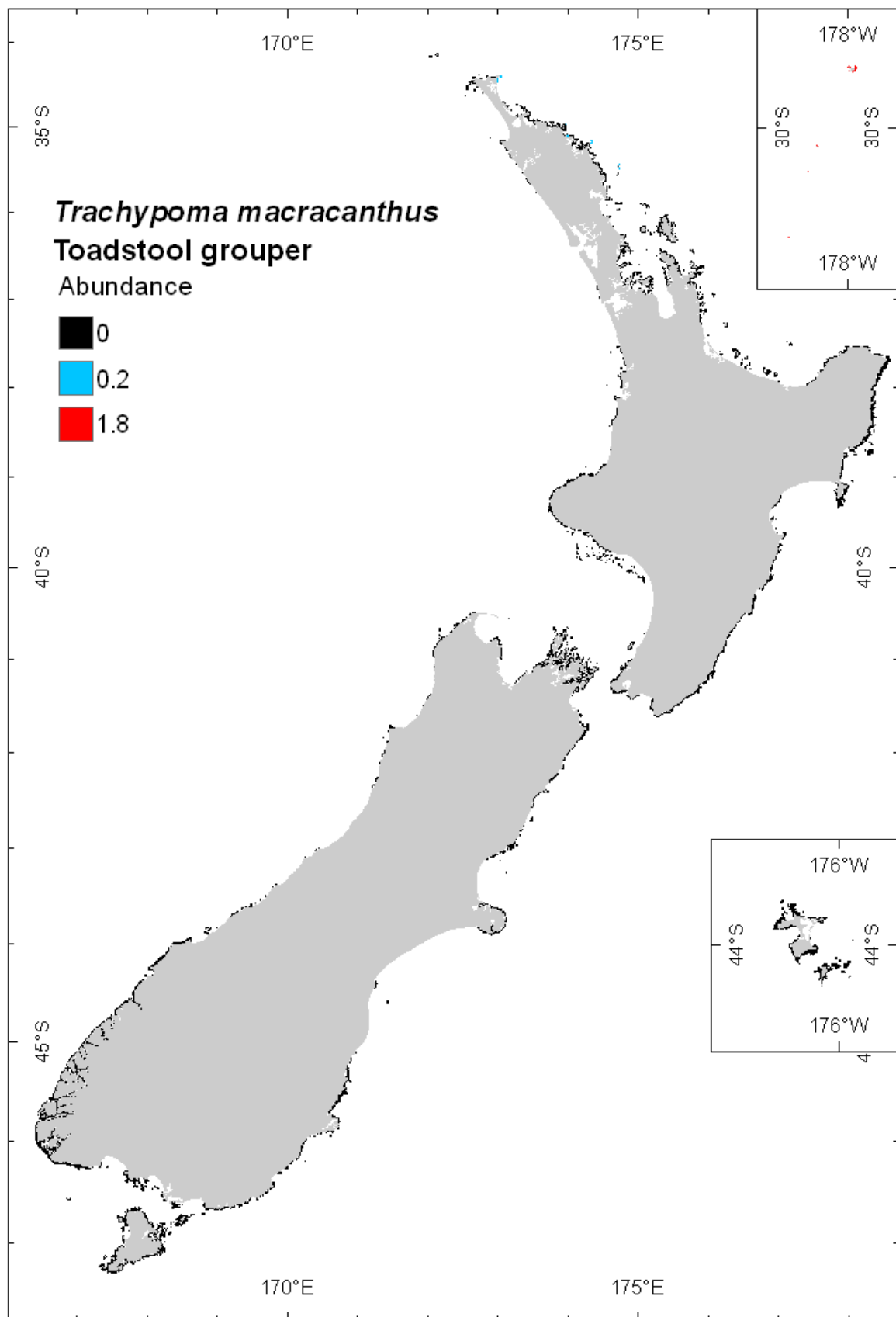


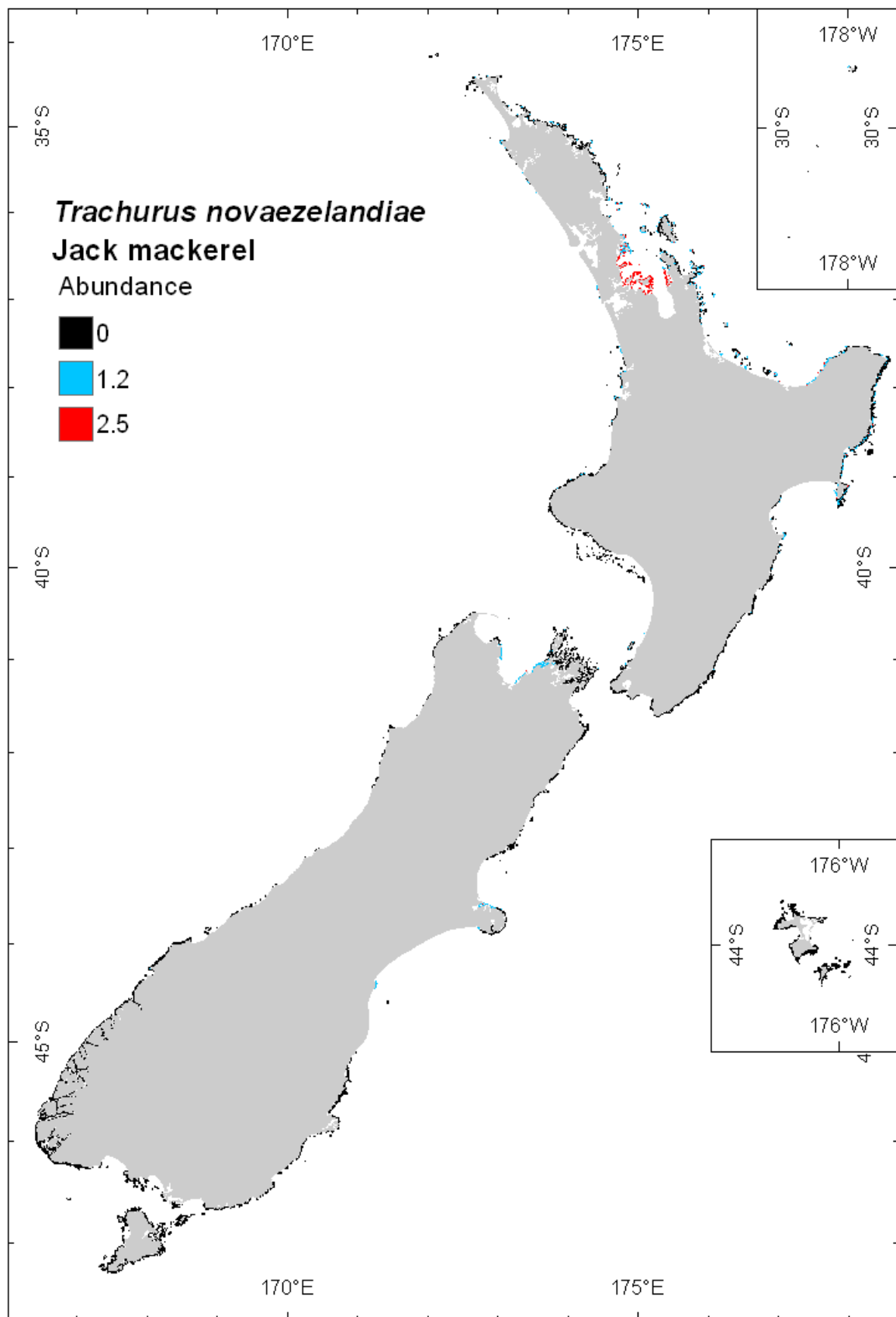


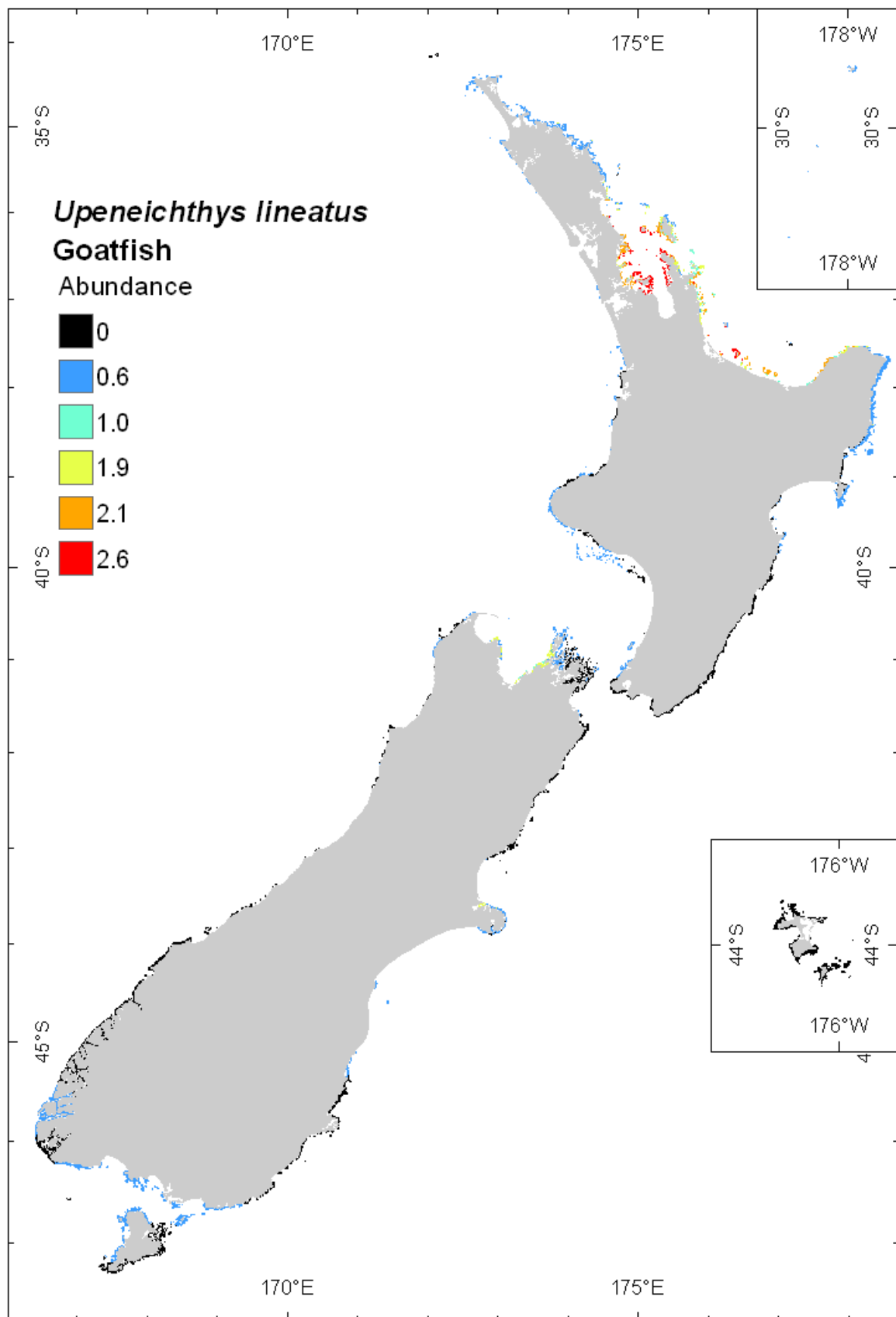


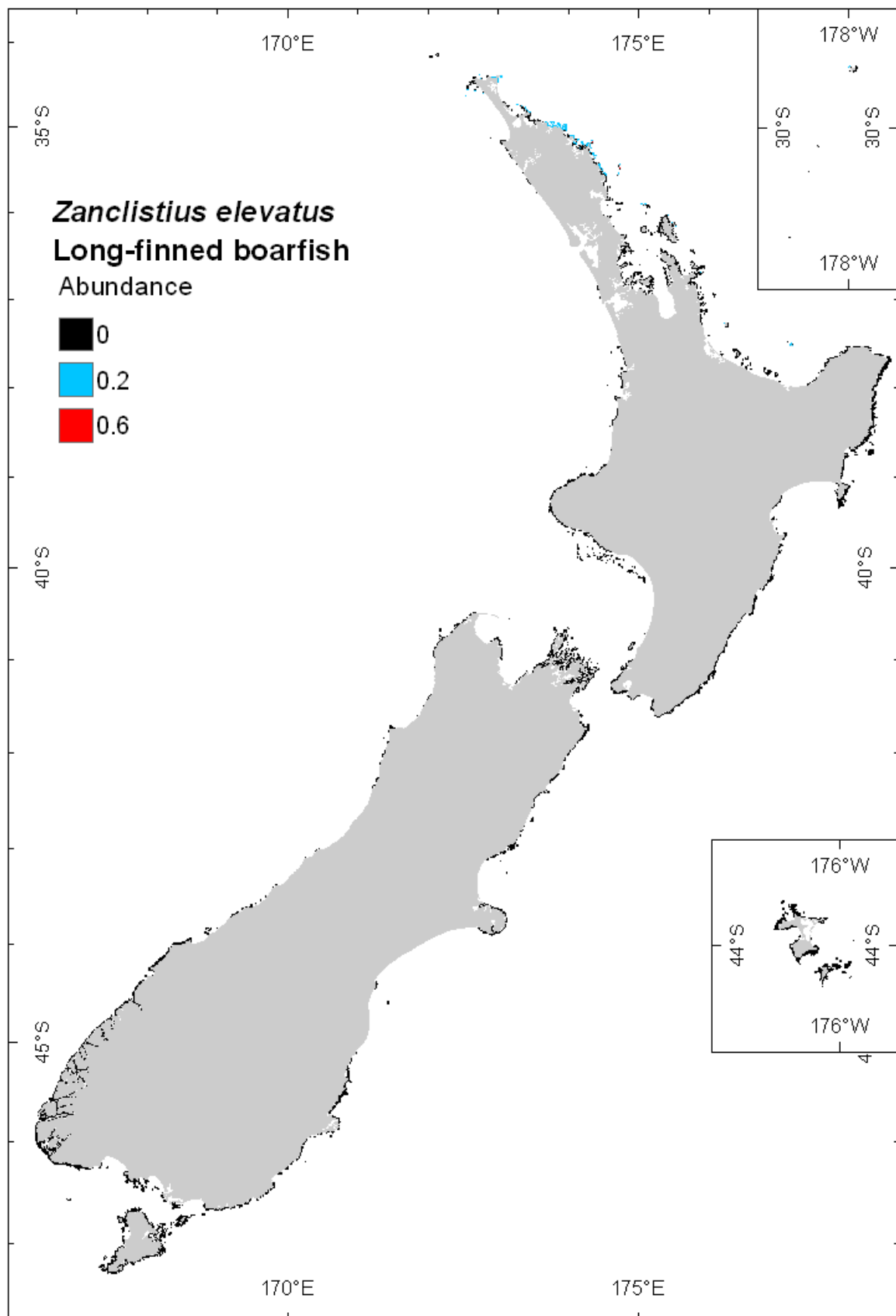


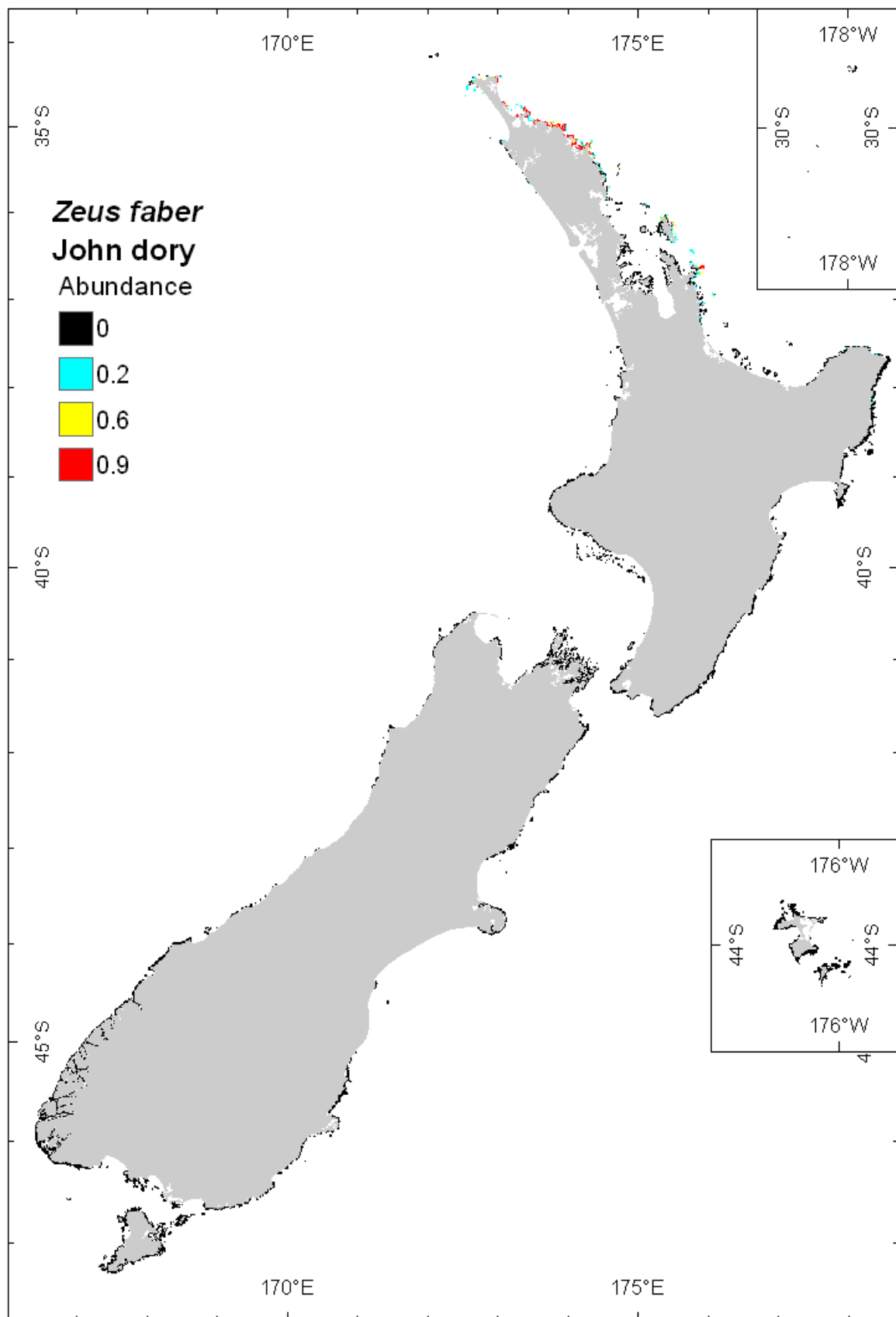










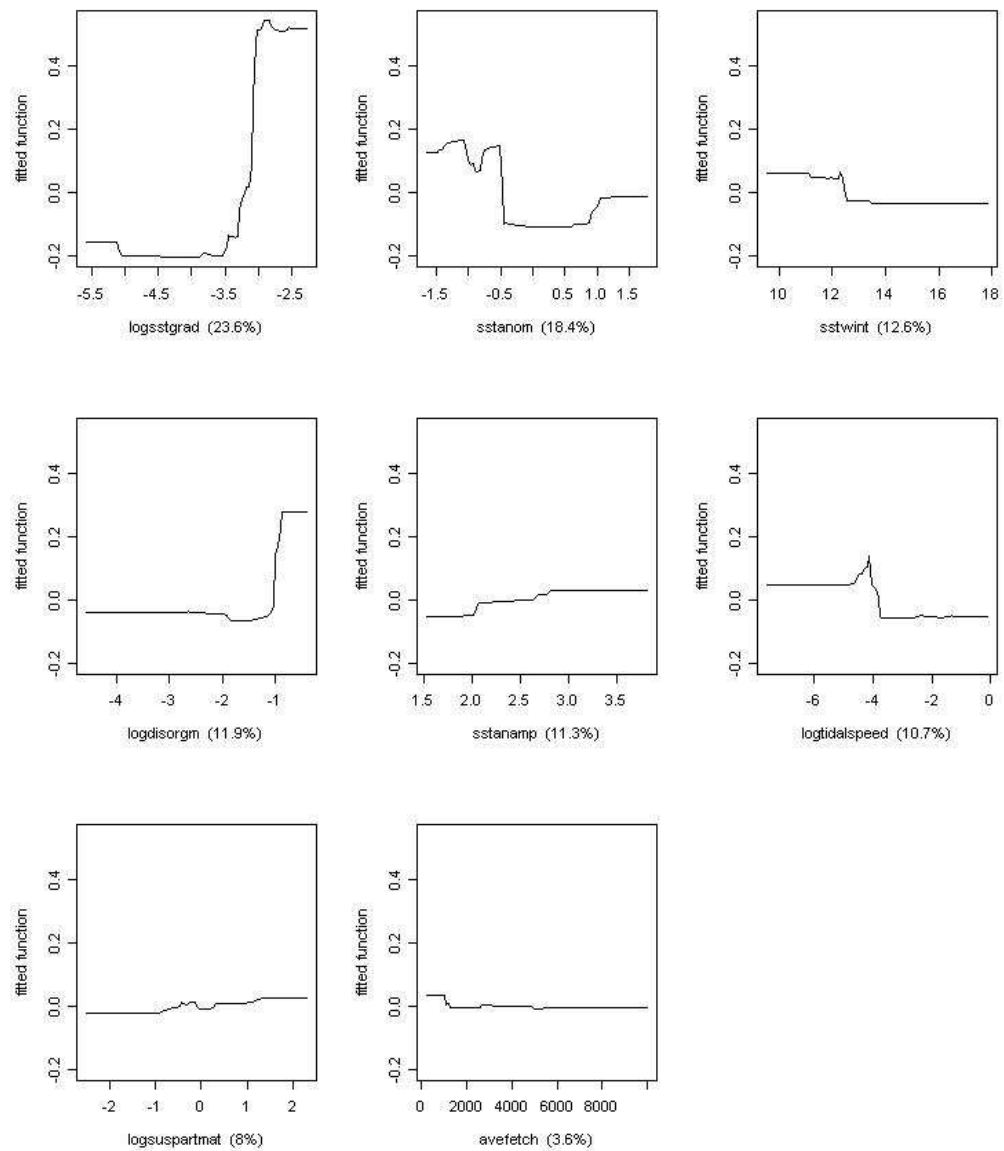


10. Appendix IV – Plots of responses to predictor variables

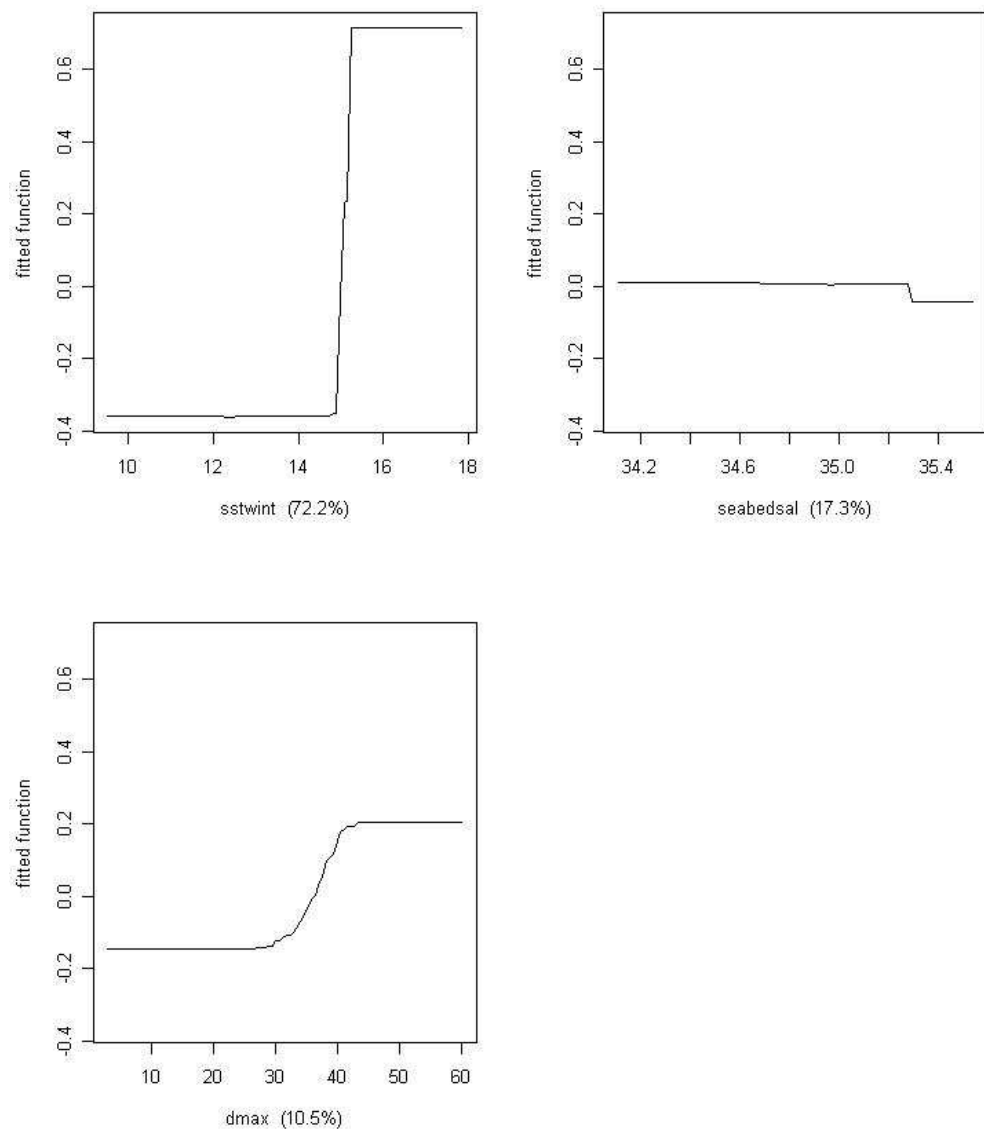
This appendix shows the influence of environmental, geographic and dive-specific predictor variables on the relative abundance of 72 fishes on shallow subtidal reefs around New Zealand. A set of graphs, containing one for each selected variable, is shown for each species. Below each graph are the variable names and, in parentheses, the relative influence on the model by each variable, expressed as a percentage. The graphs themselves show the overall effects of the variables on the abundance of the species, shown on a relative scale. For example, high values of the fitted function along part of the range of a variable suggest that this species will tend to have higher relative abundance in that range.

Variables offered to the model were *sstwint* (wintertime sea surface temperature), *seabedsal* (salinity at the sea bed), *sstanamp* (annual amplitude of sea surface temperature), *logdisorgm* (log of dissolved organic matter), *logtidalspeed* (log of tidal speed), *logsuspartmat* (log of suspended particulate matter), *sstanom* (sea surface temperature anomaly), *logsstgrad* (log of sea surface temperature gradient), *chla* (concentration of chlorophyll *a*) *avefetch* (average fetch), *dcoast* (distance from coast), *dmax* (maximum depth of dive), *dur* (duration of dive), *dmin* (minimum depth of dive) and *vis* (visibility of dive).

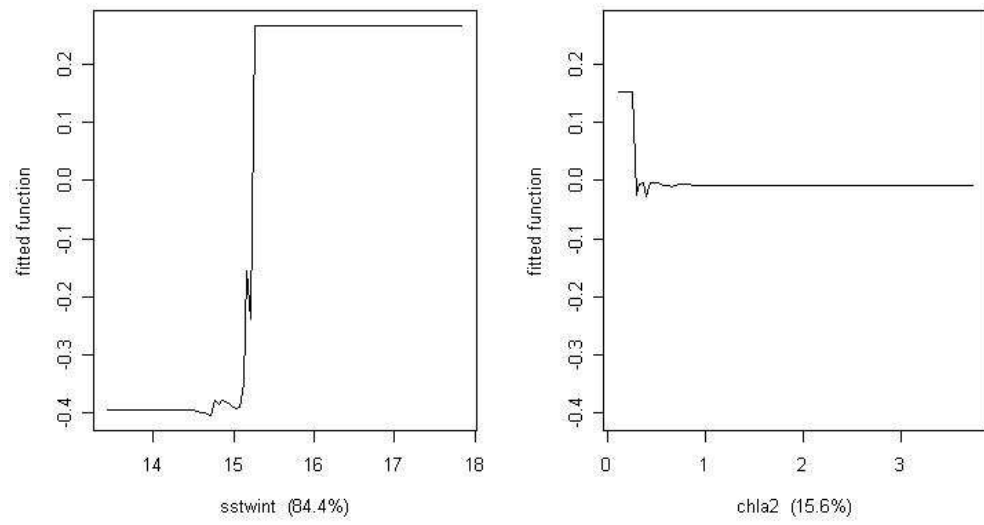
Aldrichetta forsteri : Yellow-eyed mullet



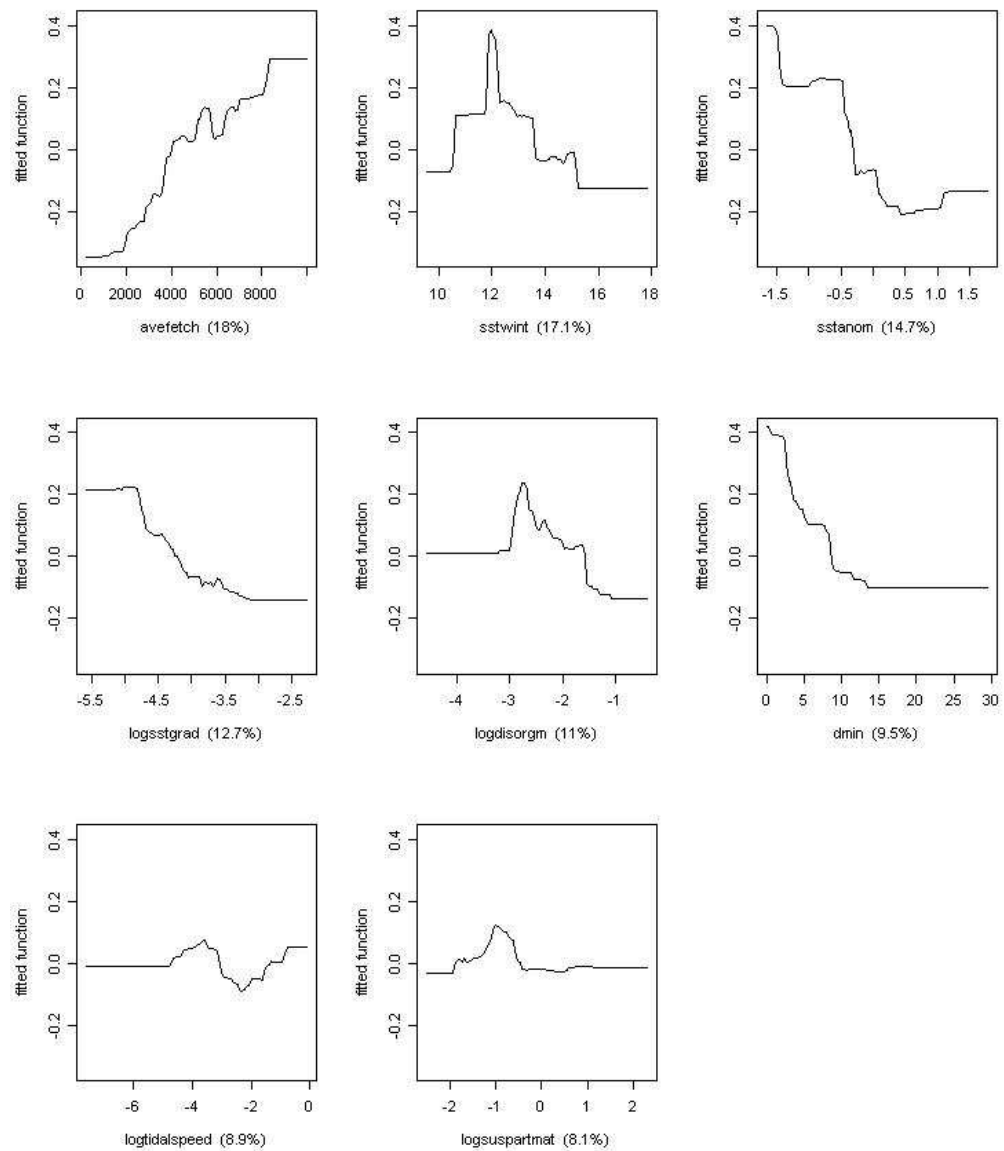
Amphichaetodon howensis : Lord Howe coralfish



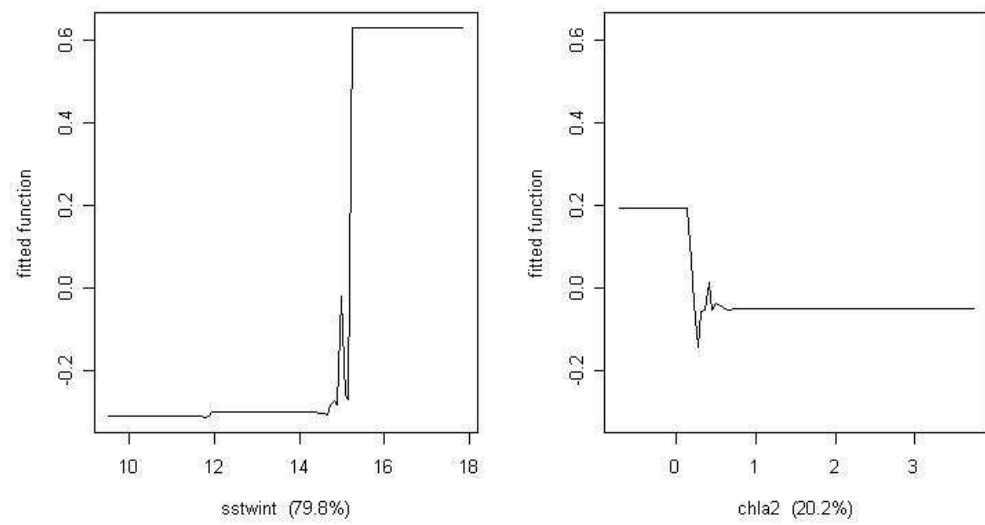
Anampses elegans : Elegant wrasse



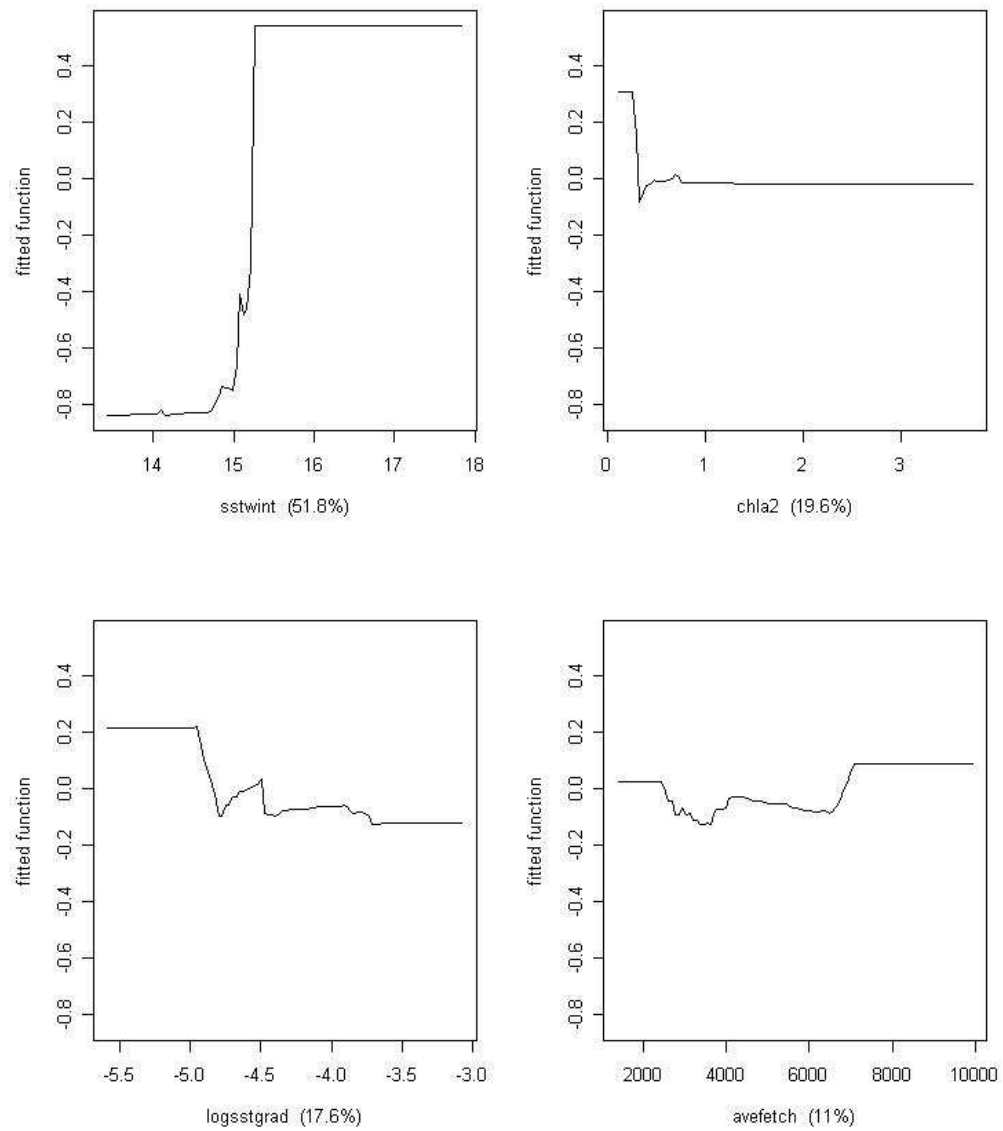
Aplodactylus arcidens : Marblefish



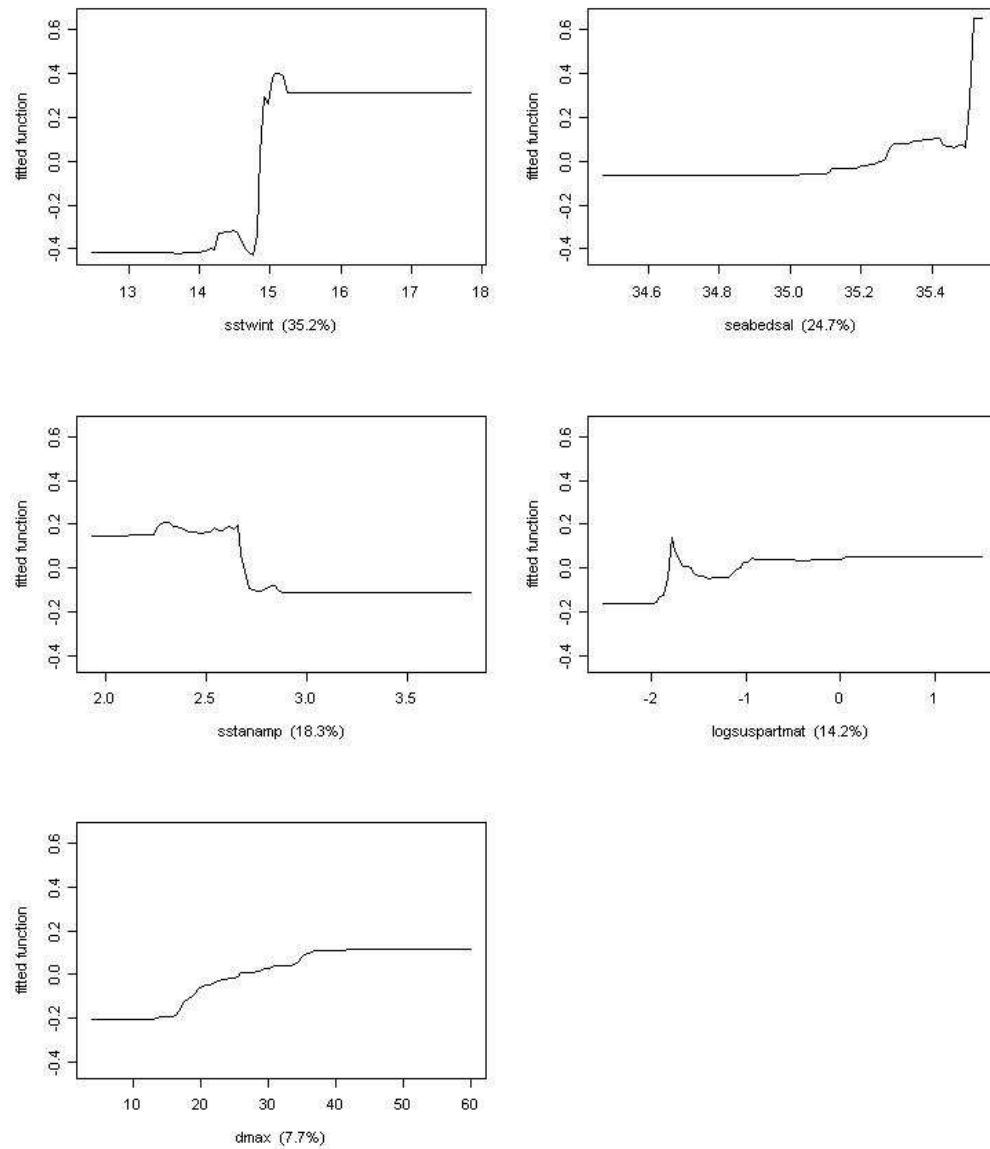
Aplodactylus etheridgii : Notch-head marblefish



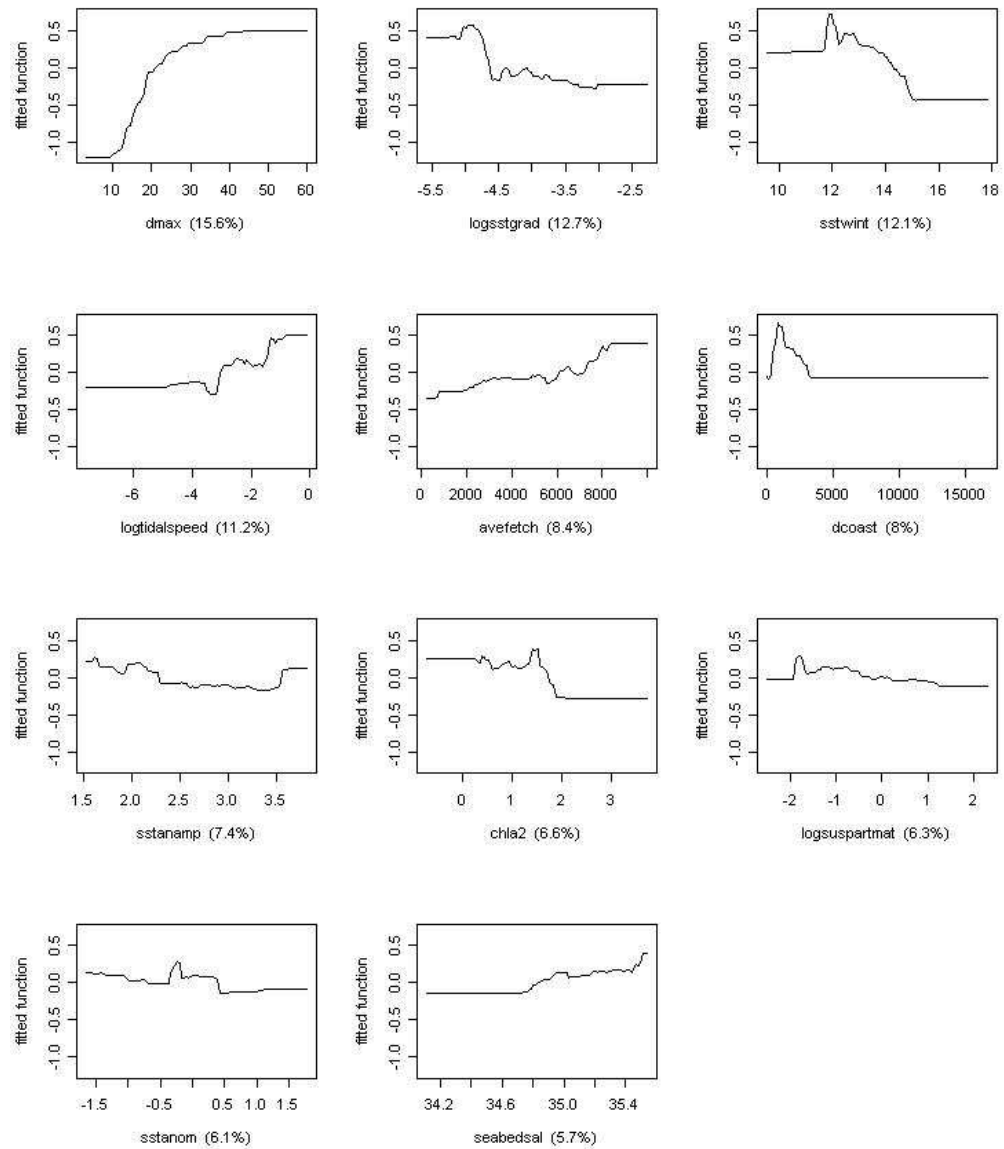
***Atypichthys latus* : Mado**



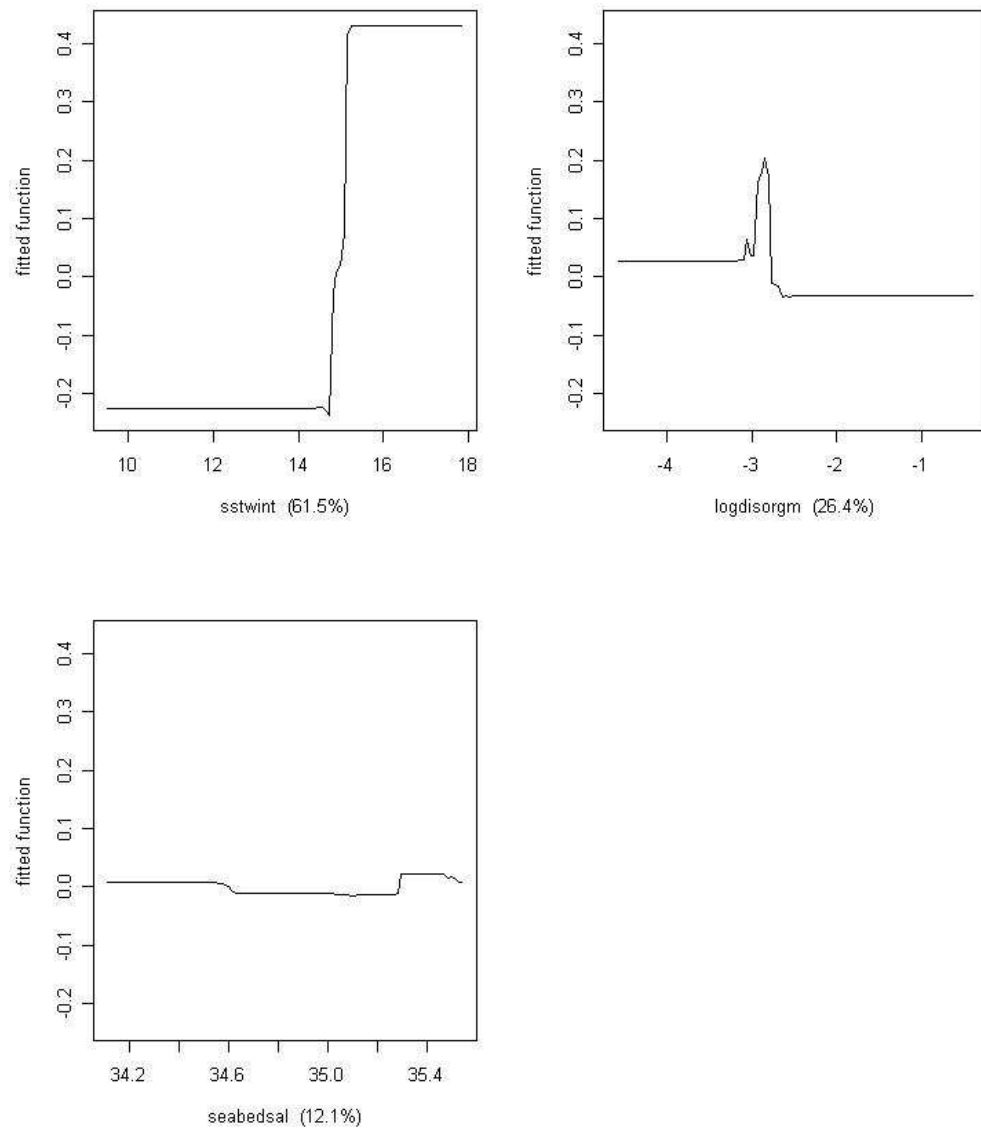
Bodianus unimaculatus : Red pigfish



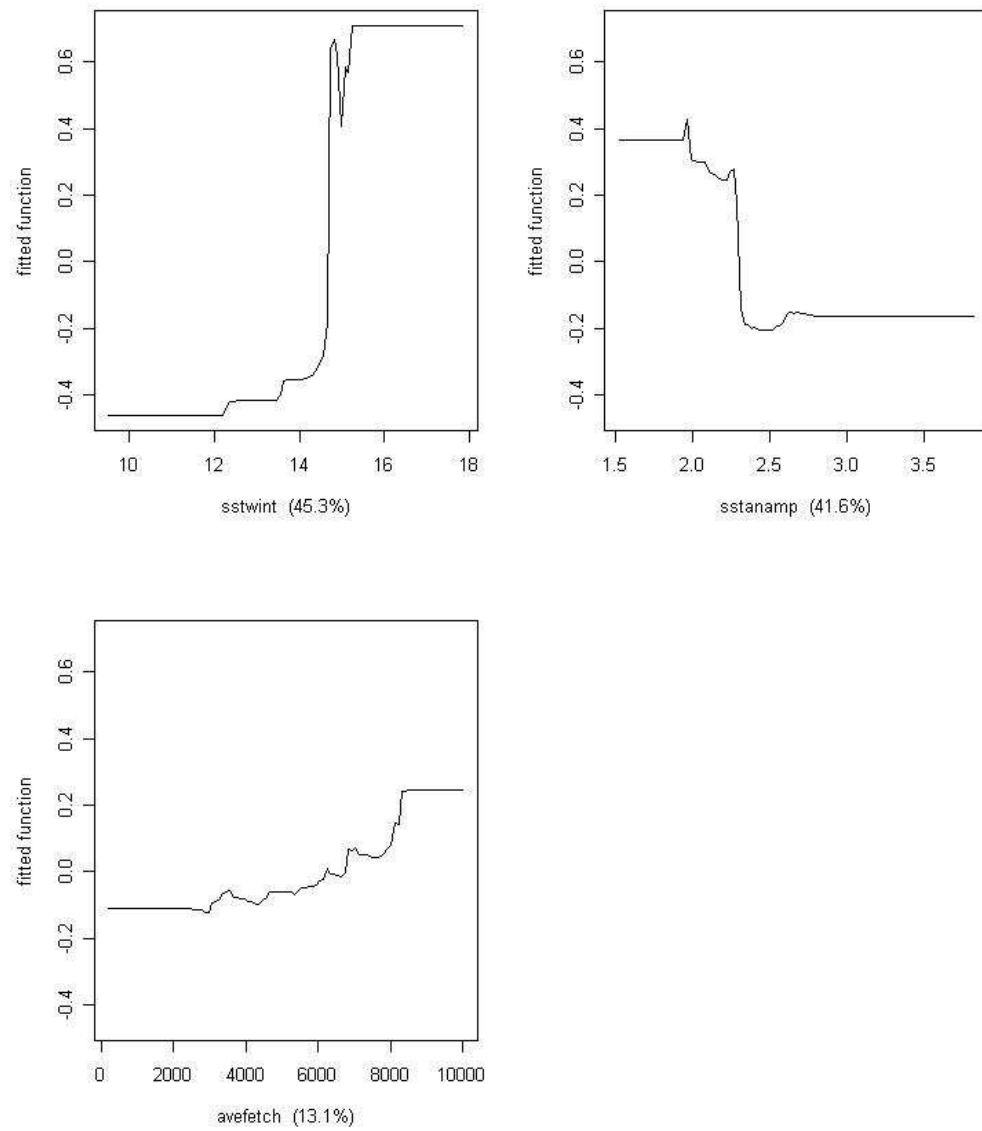
Caesioperca lepidoptera : Butterfly perch



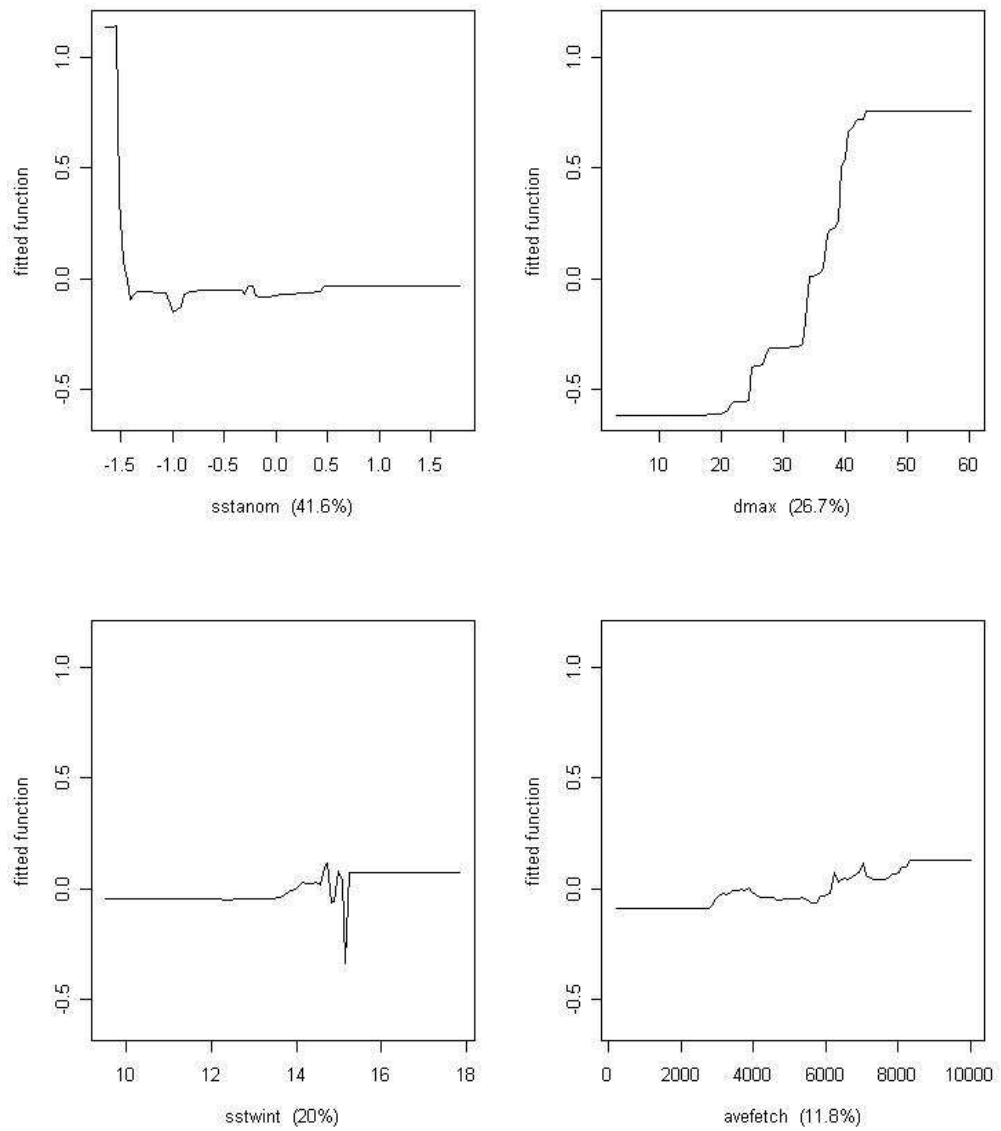
Canthigaster callisterna : Clown toado



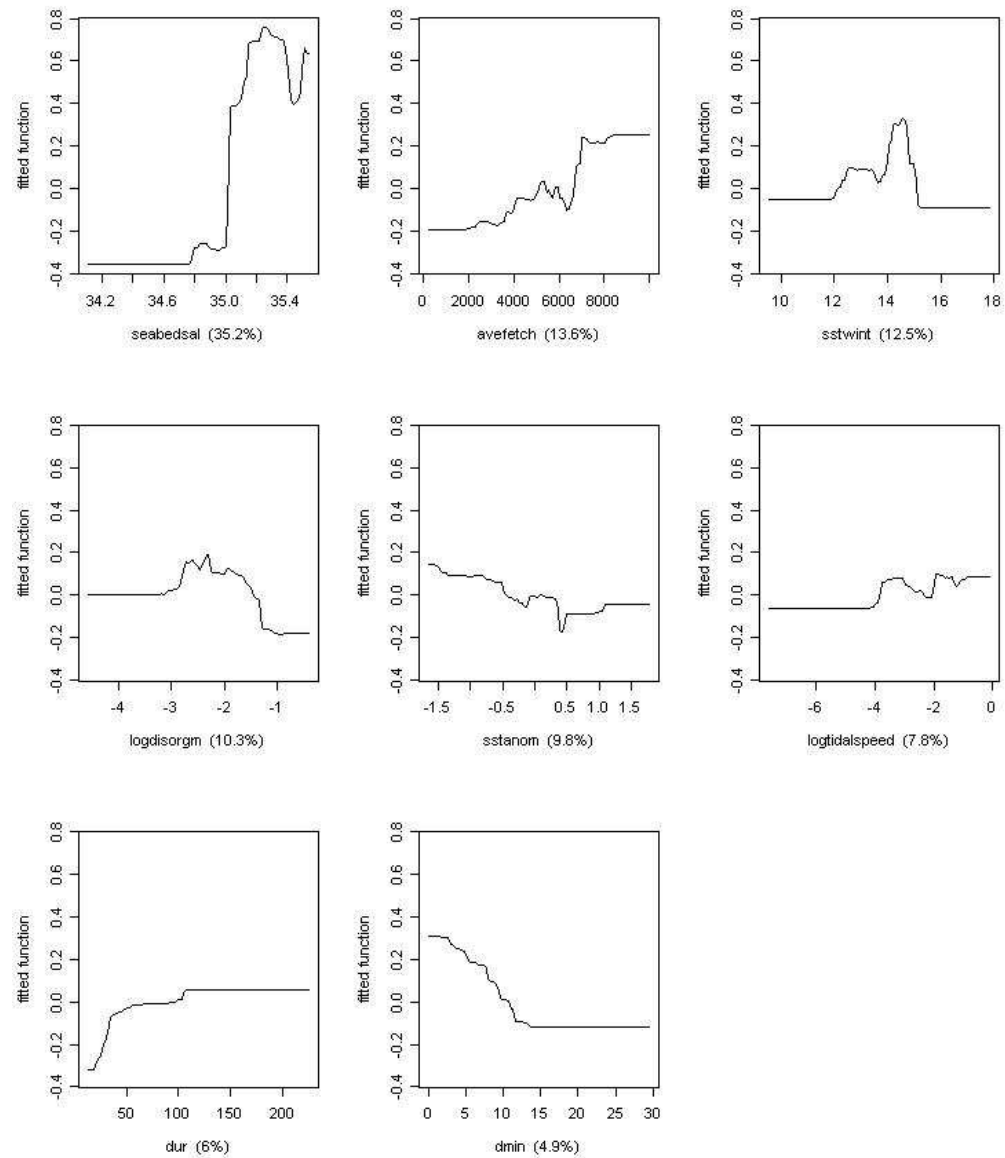
Caprodon longimanus : Pink maomao



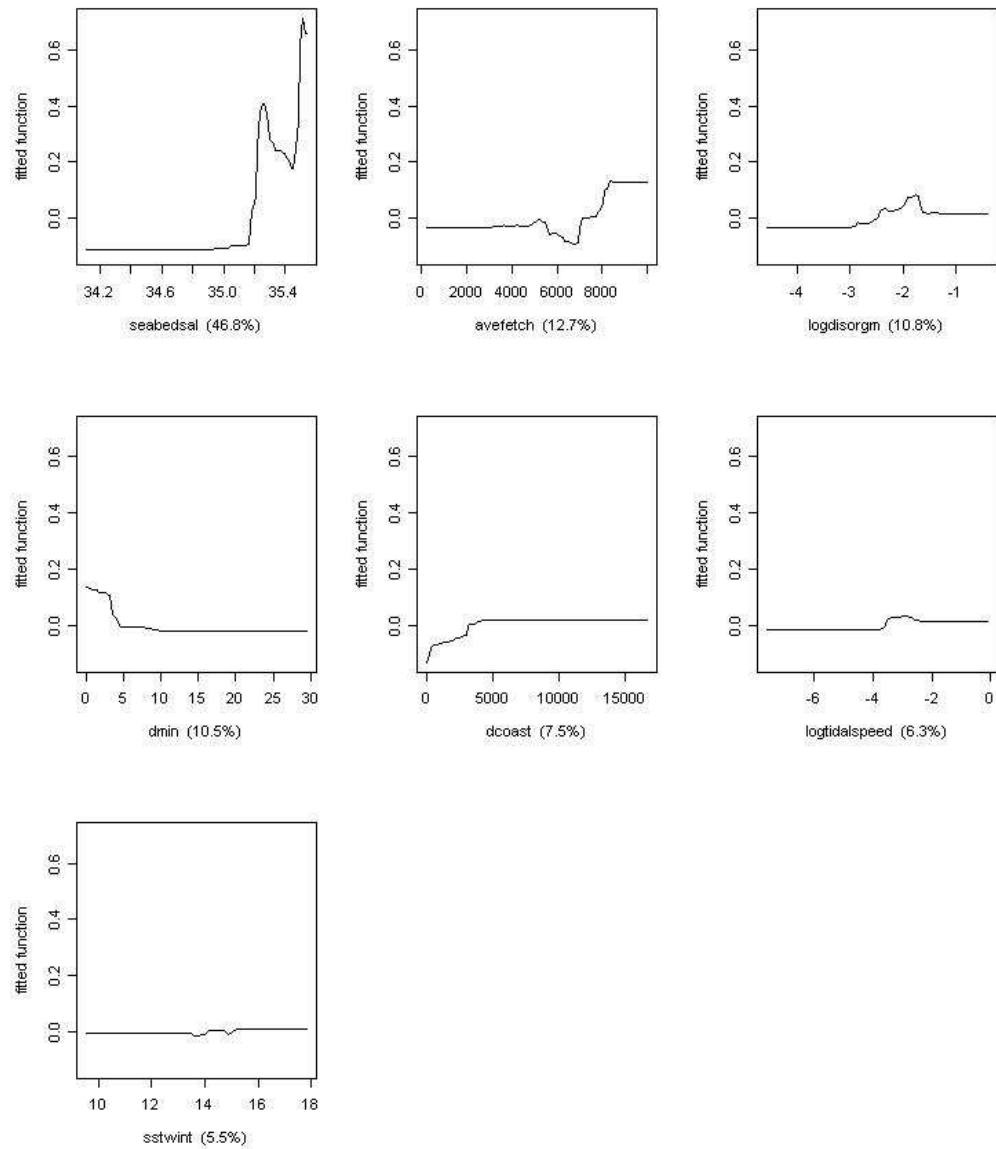
Centroberyx affinis : Golden snapper



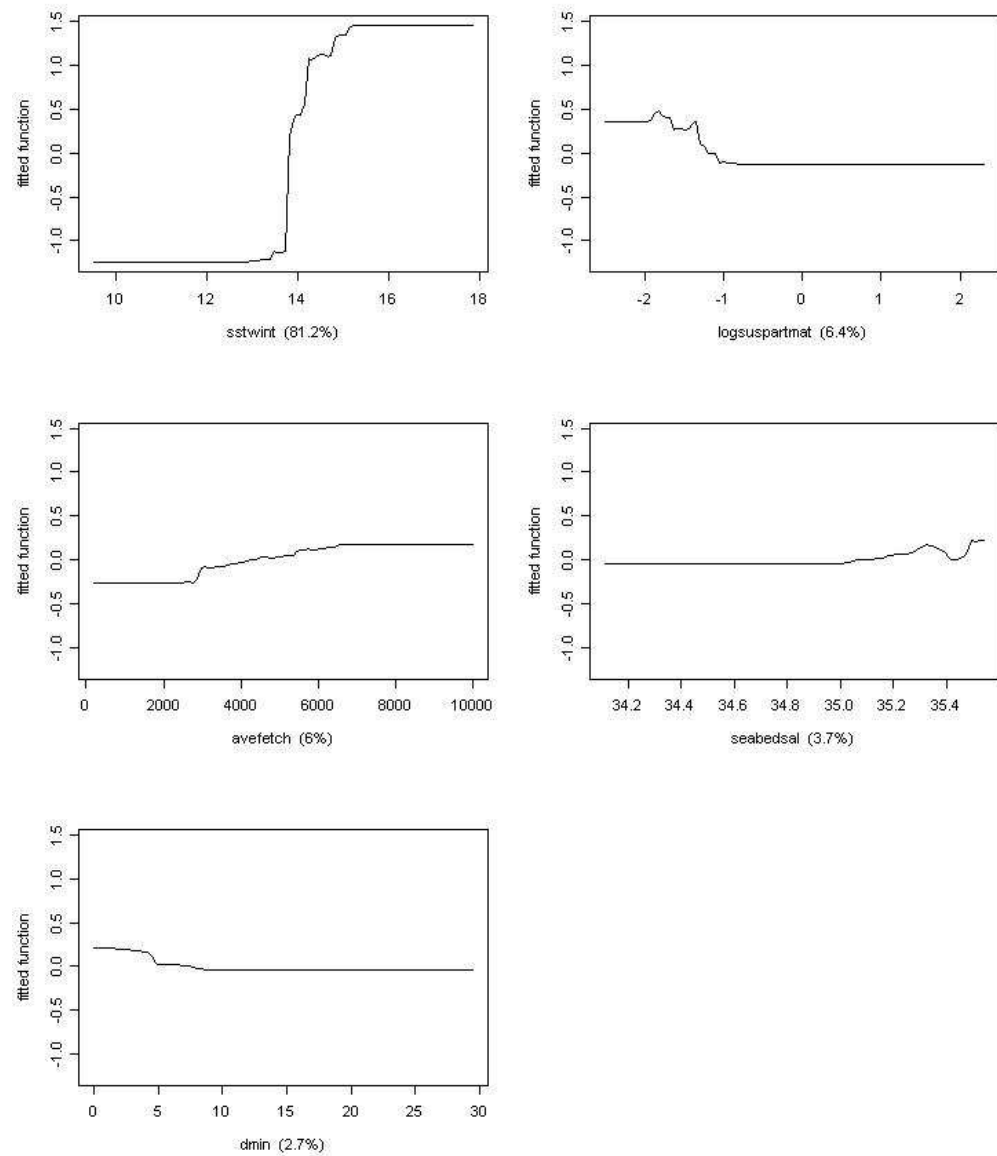
Cheilodactylus spectabilis : Red moki



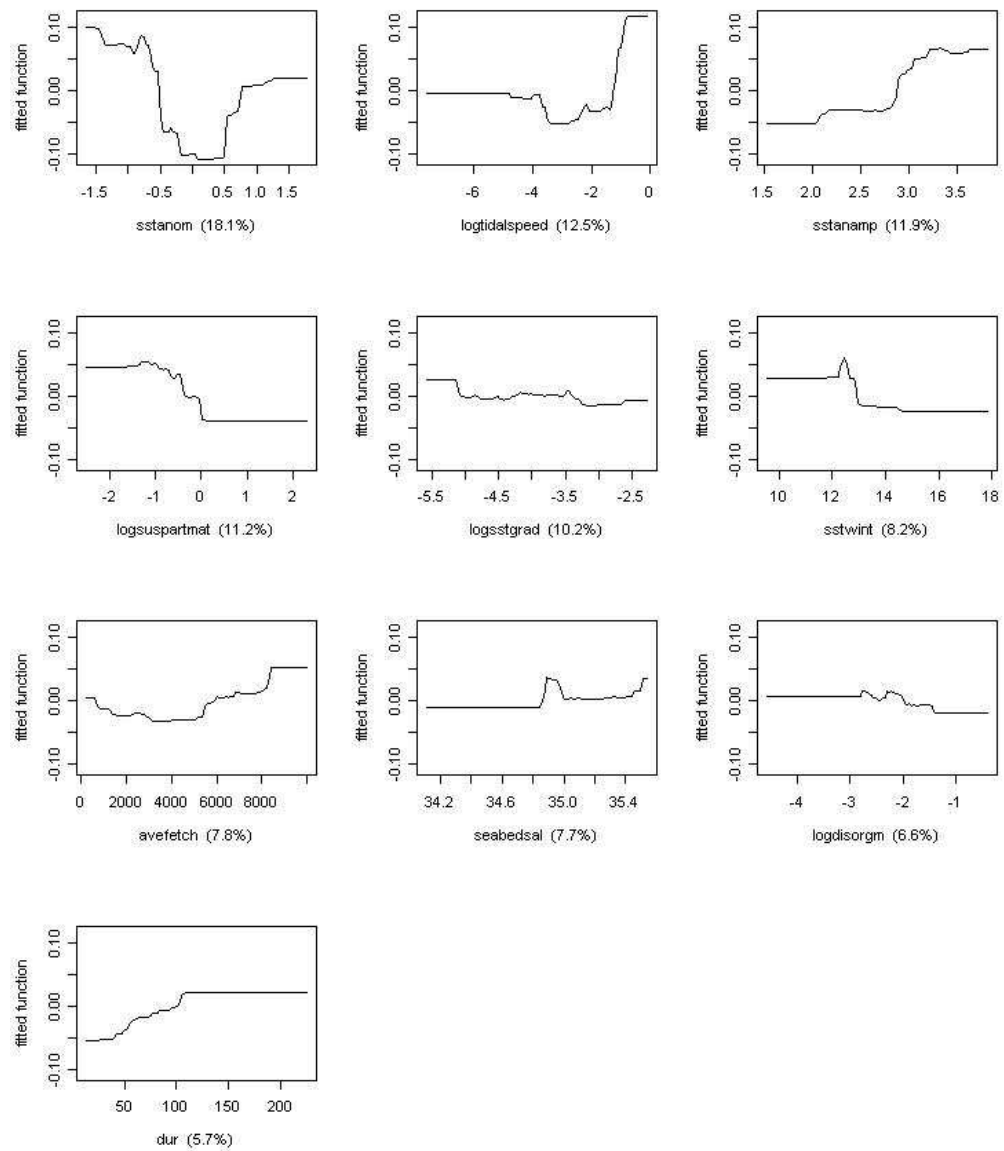
Chironemus marmoratus : Hiwihiwi



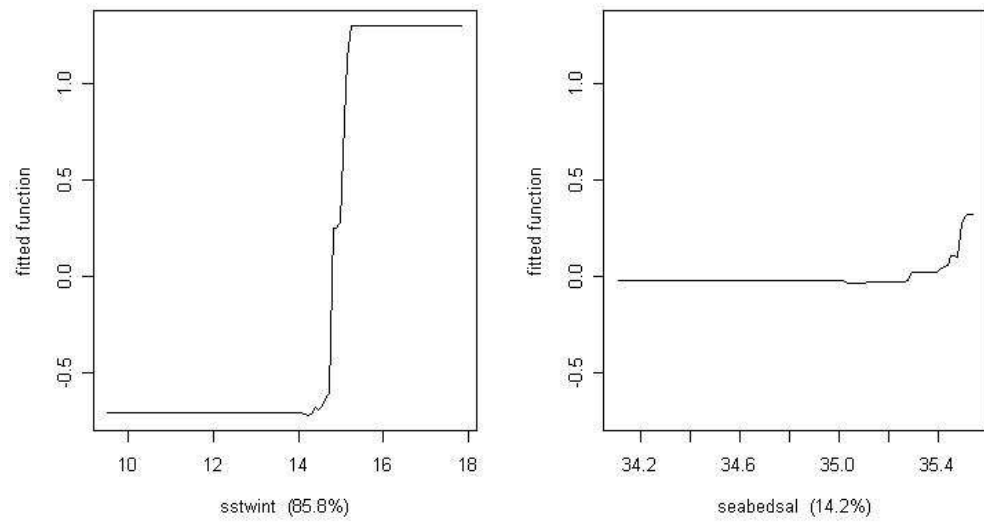
Chromis dispilus : Demoiselle



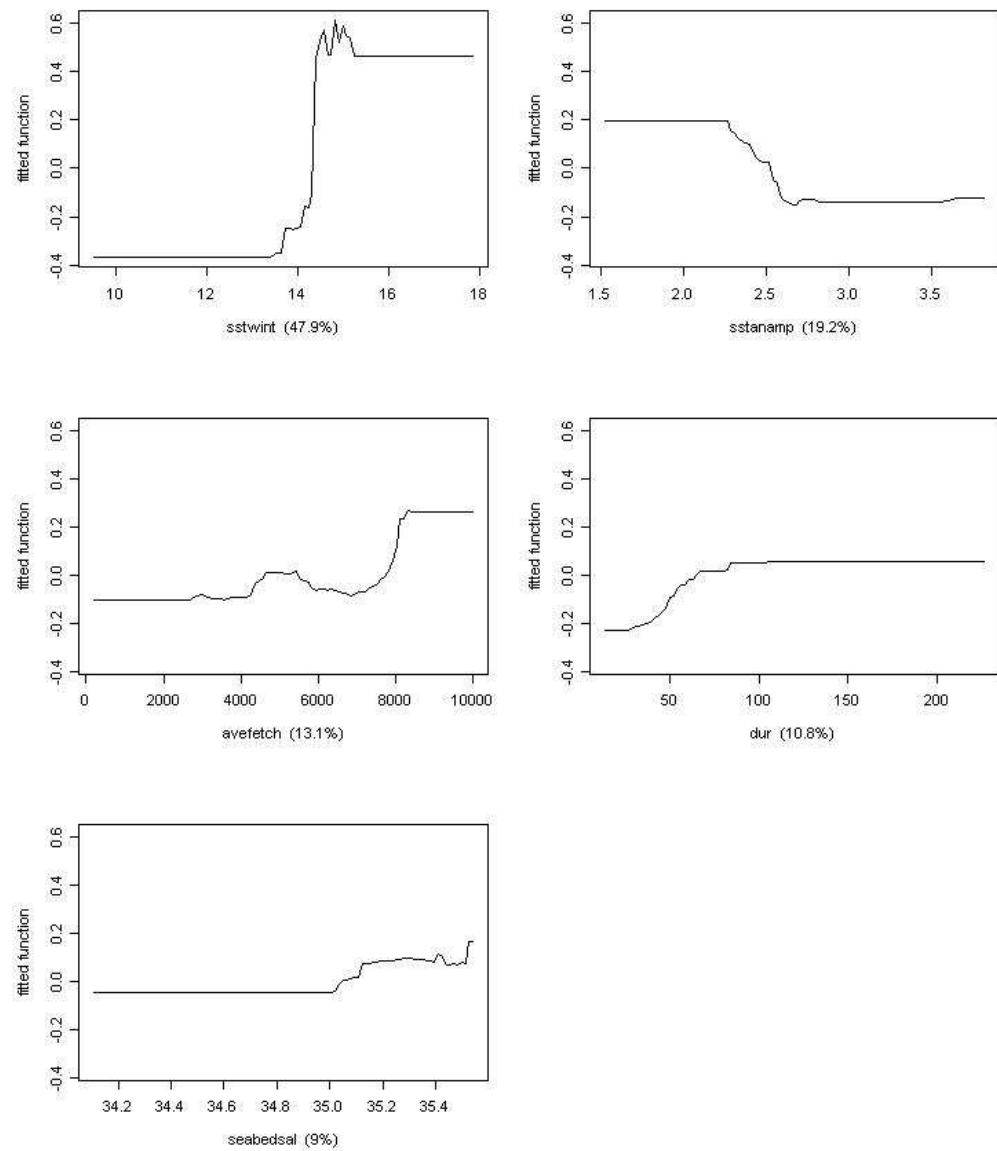
Conger verreauxi : Common conger eel



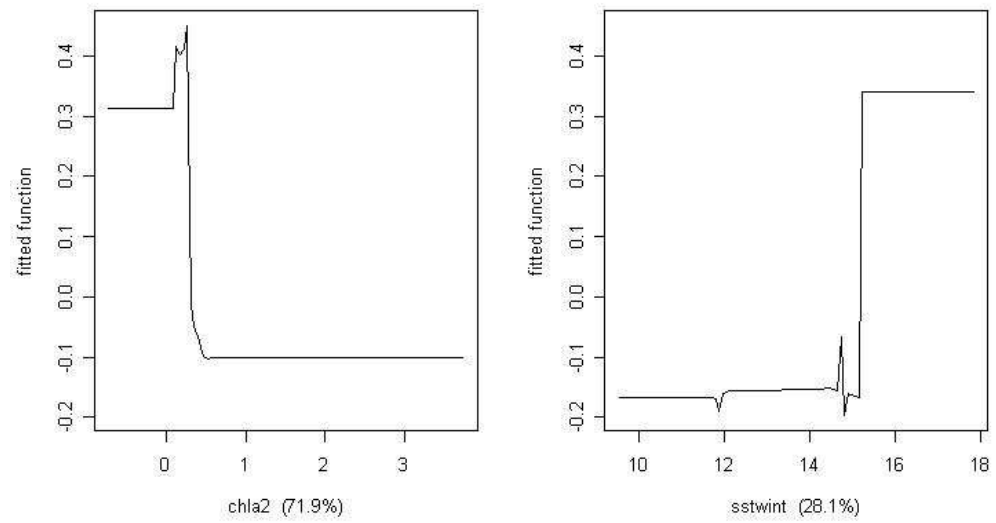
Coris sandageri : Sandager's wrasse



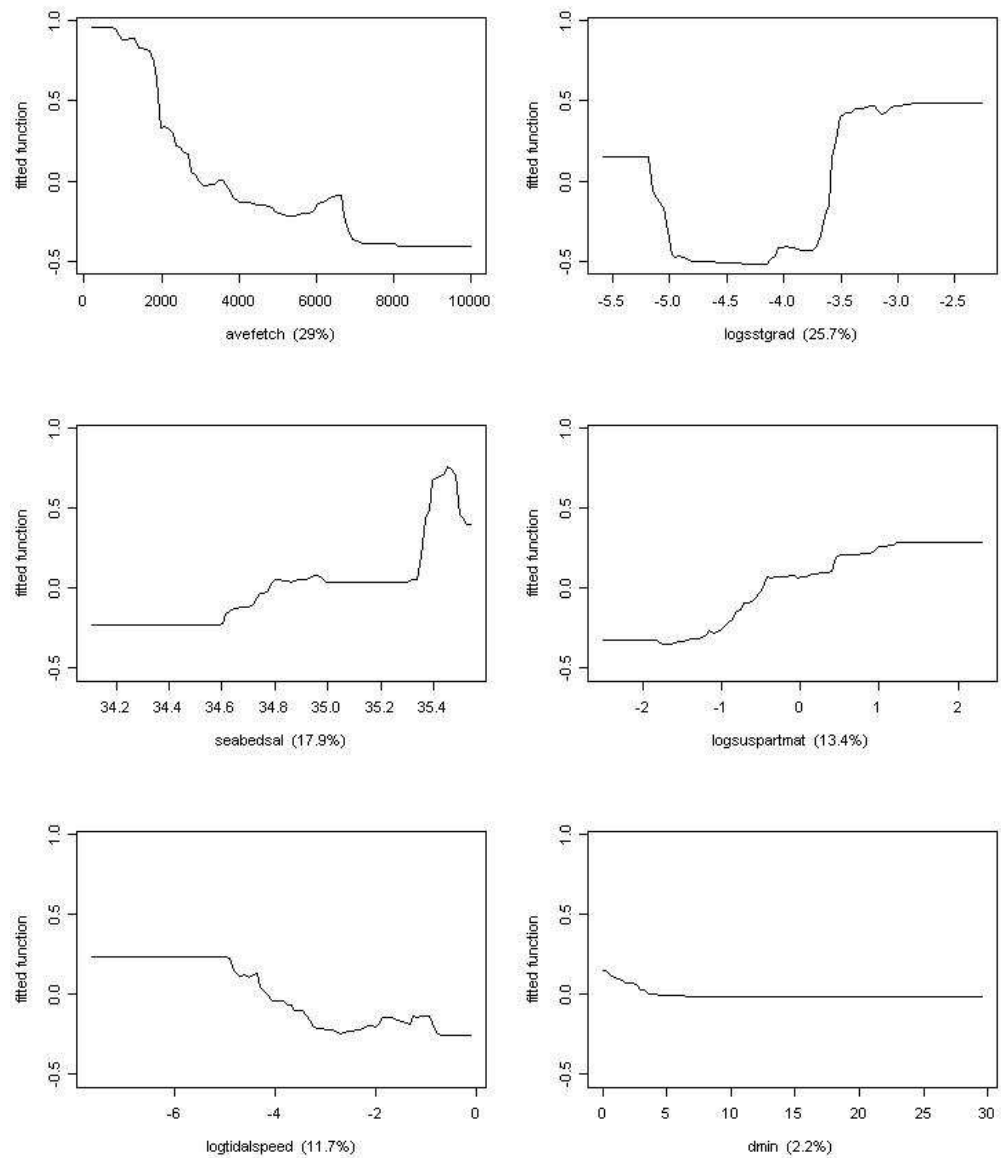
Decapterus koheru : Koheru



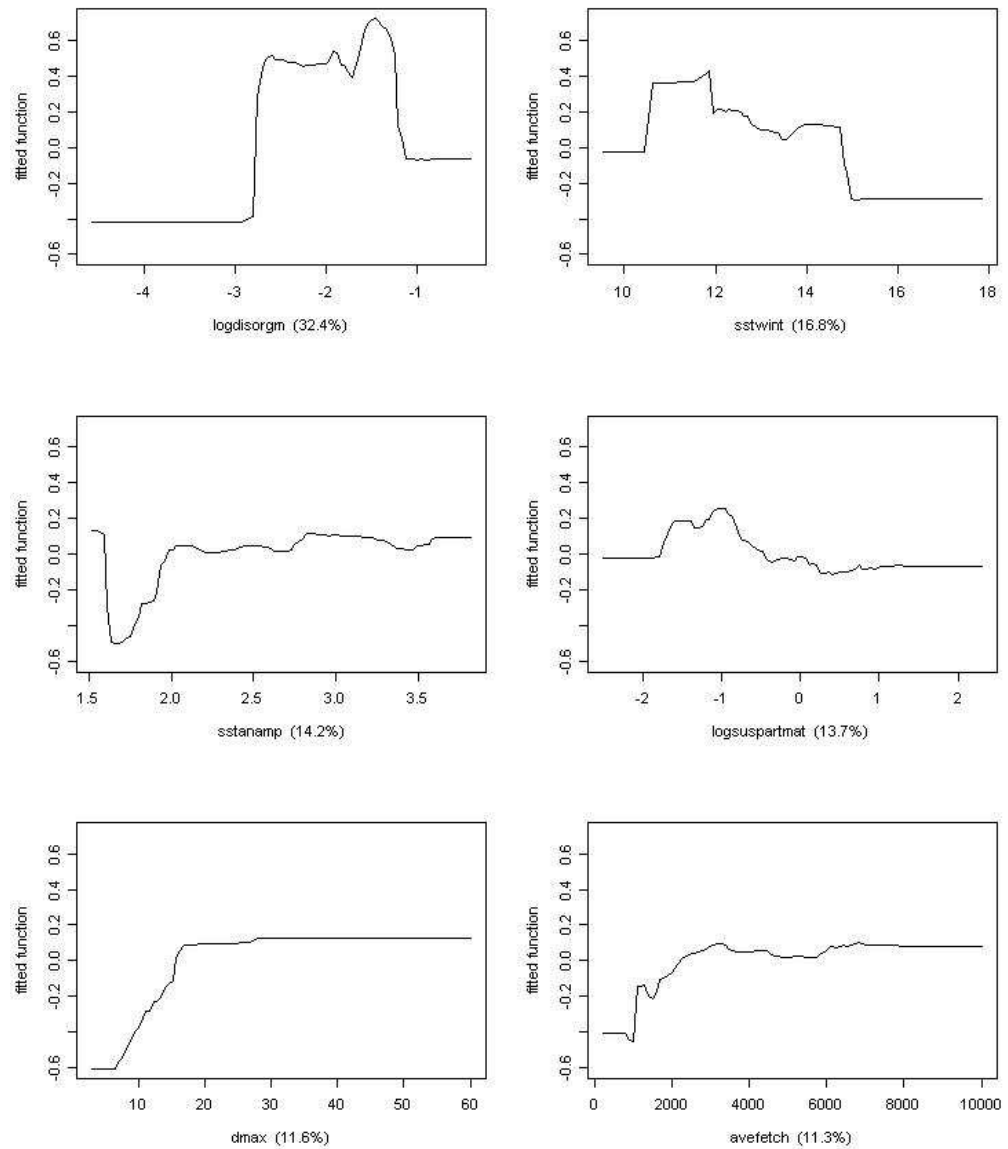
Epinephelus daemeli : Spotted black grouper



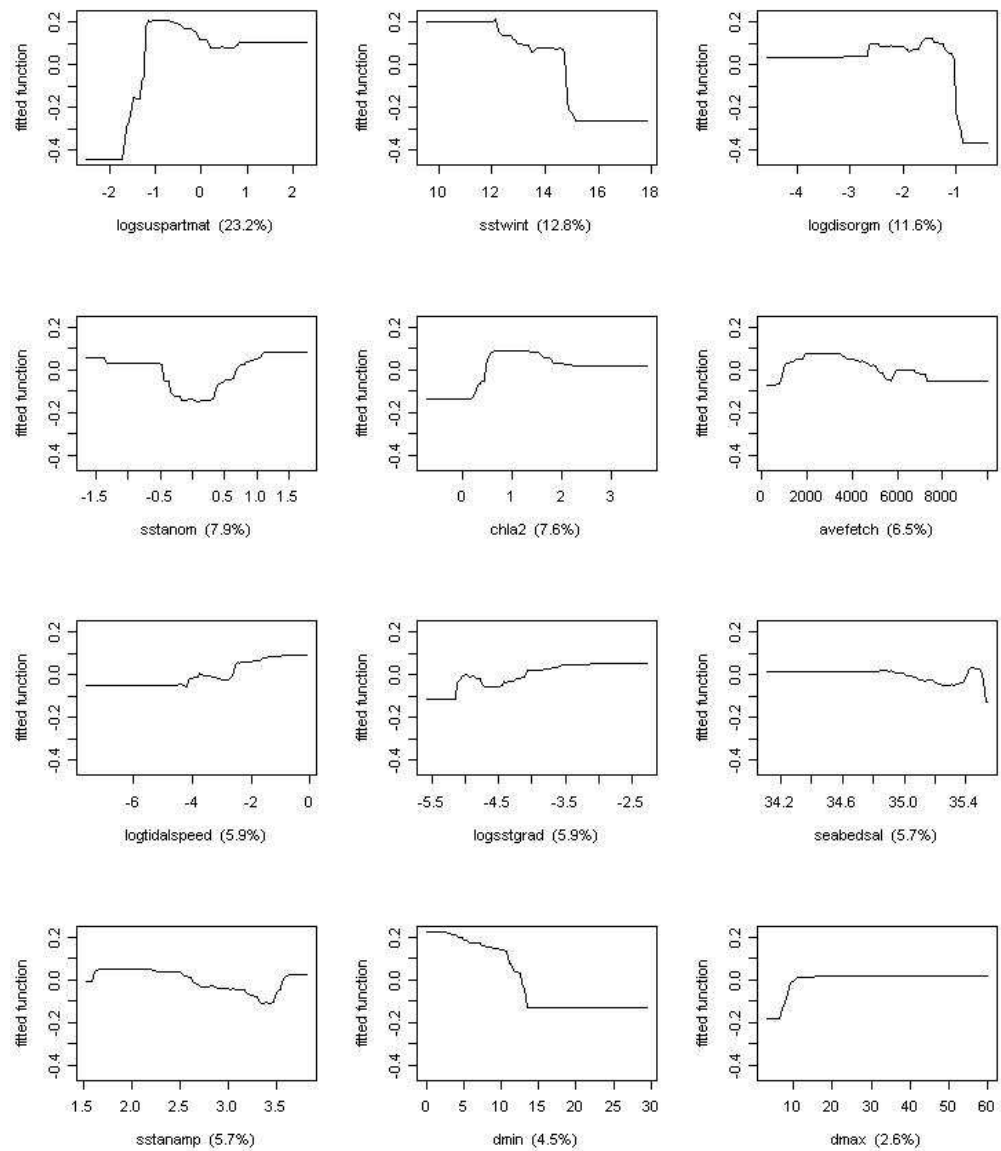
Forsterygion lapillum : Common triplefin



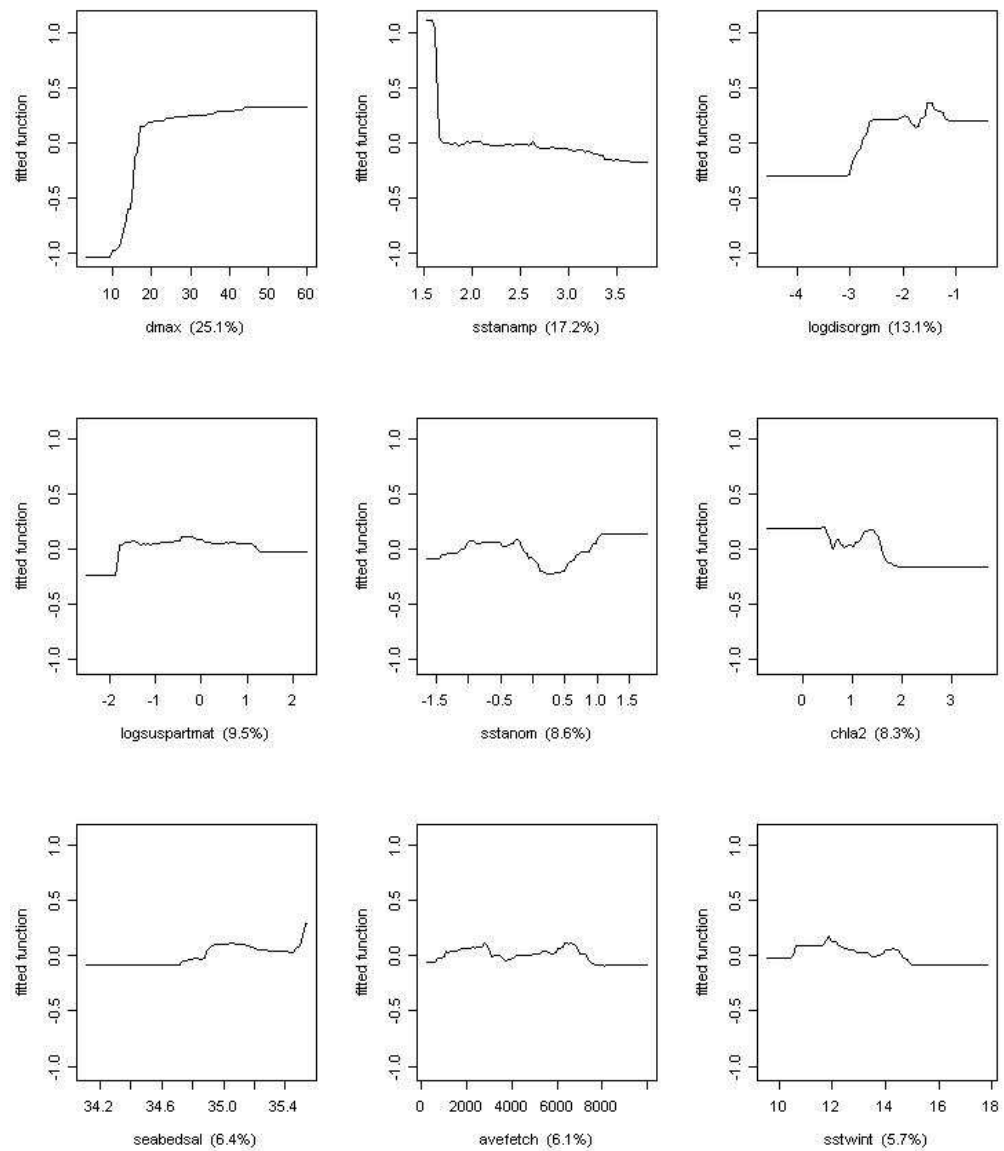
Forsterygion malcolmi : Banded triplefin



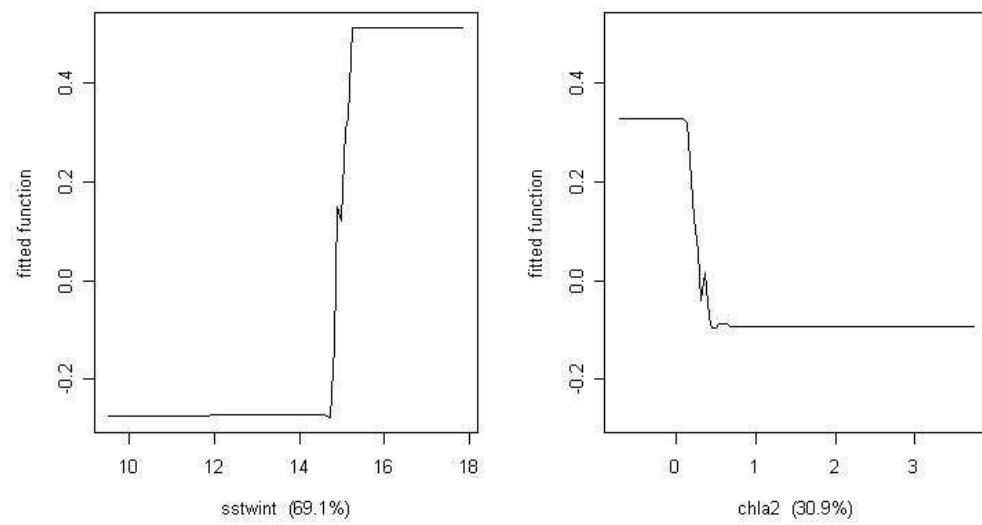
Forsterygion varium : Variable triplefin



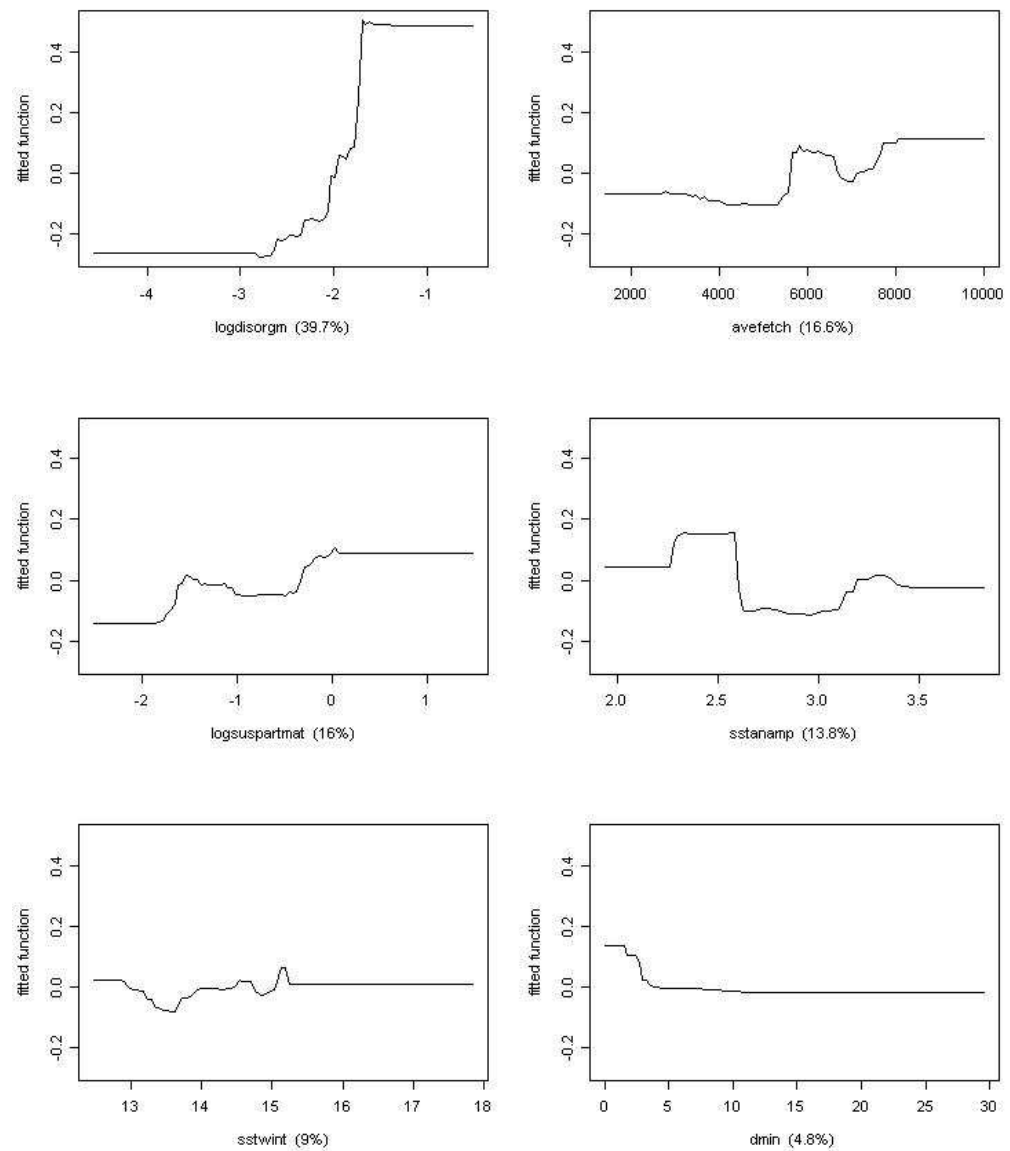
Forsterygion flavonigrum : Yellow-black triplefin



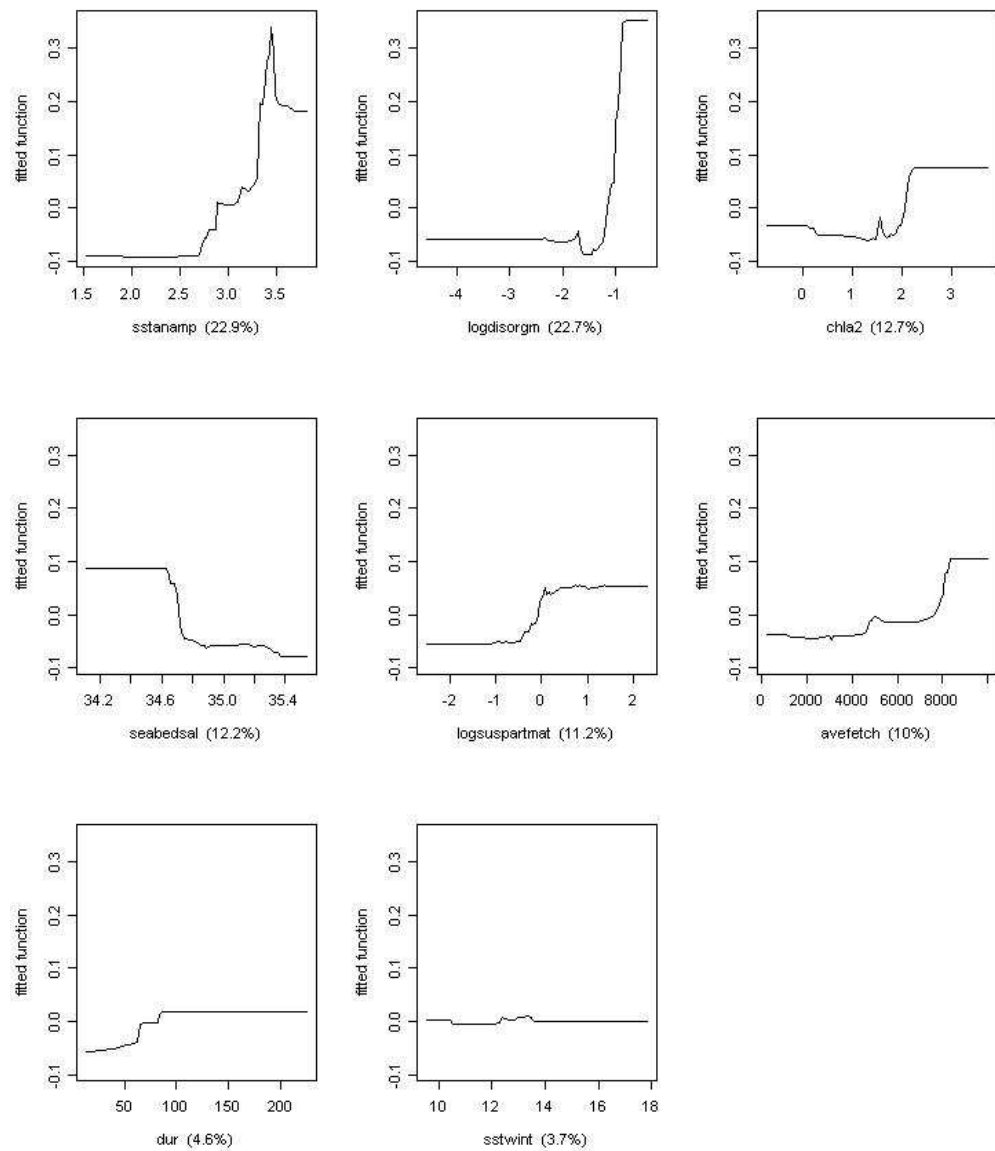
Girella cyanea : Bluefish



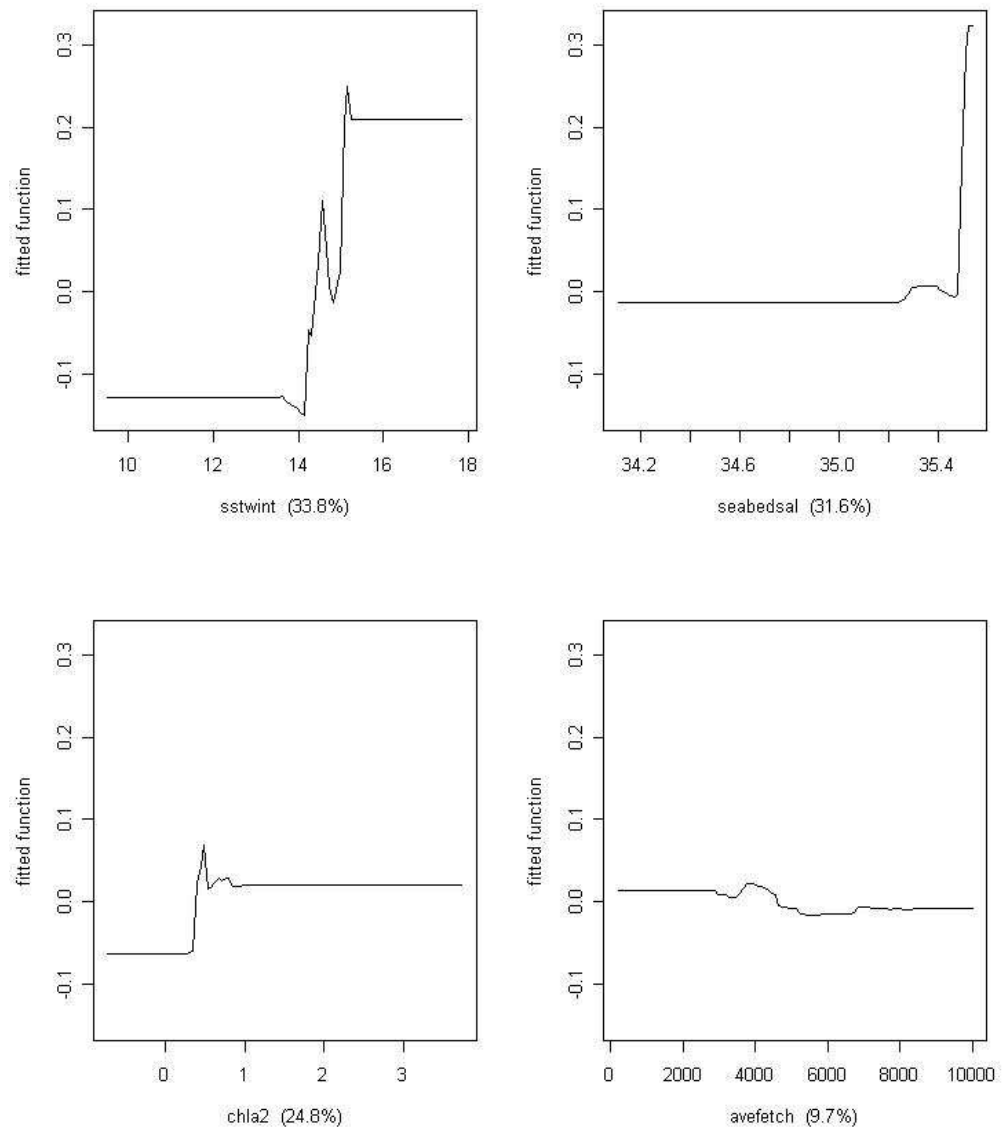
Girella tricuspidata : Parore



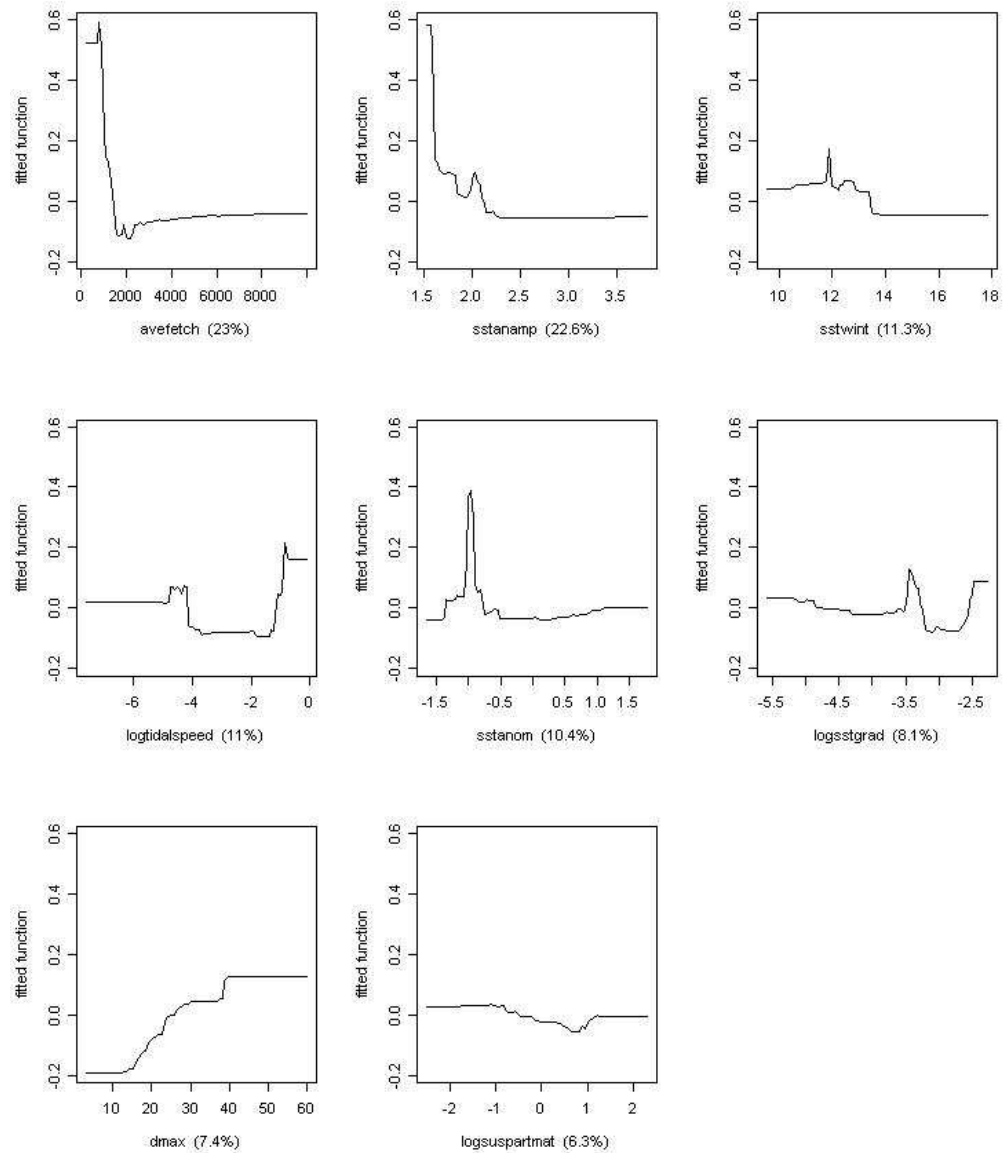
Grahamina gymnota : Robust triplefin



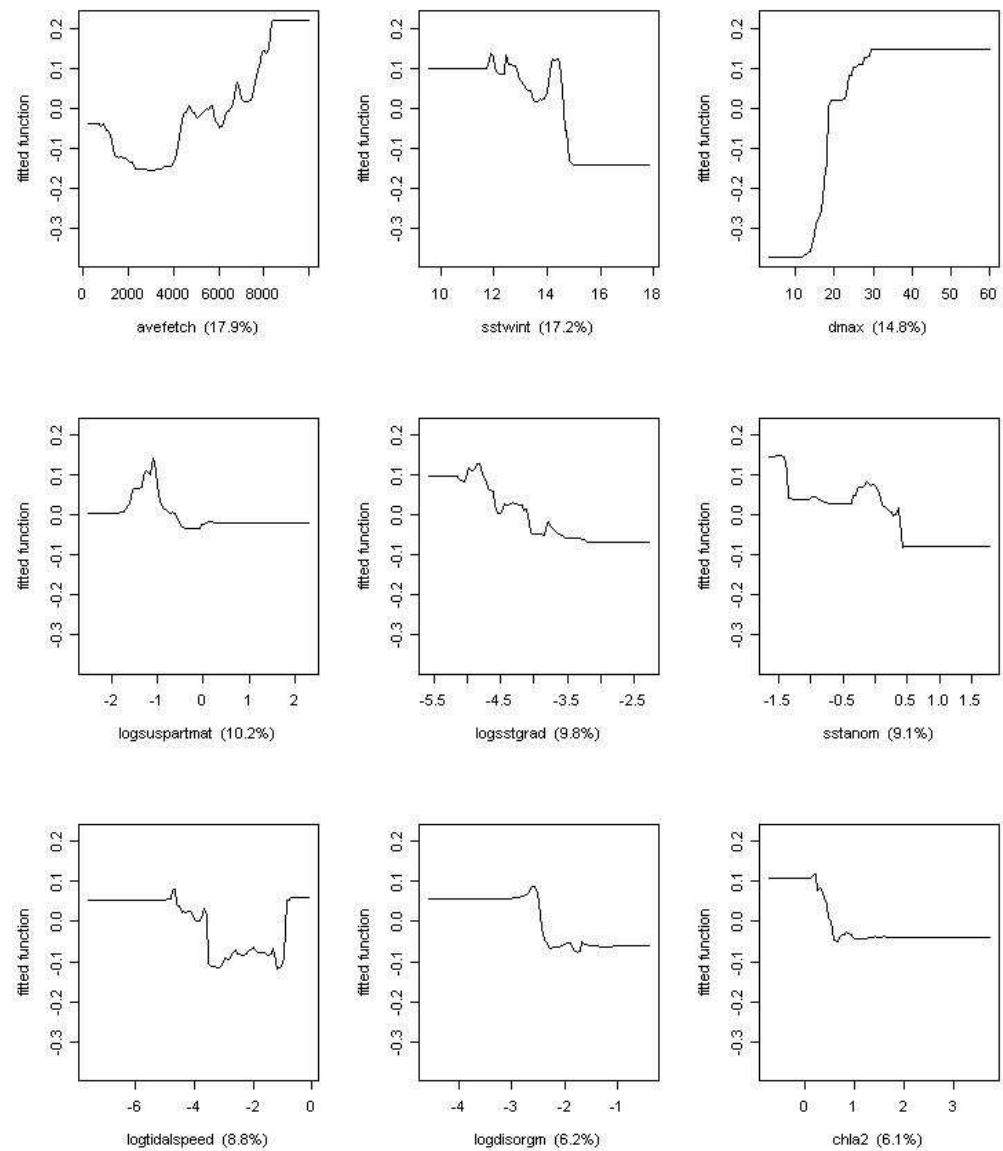
Gymnothorax prasinus : Yellow moray



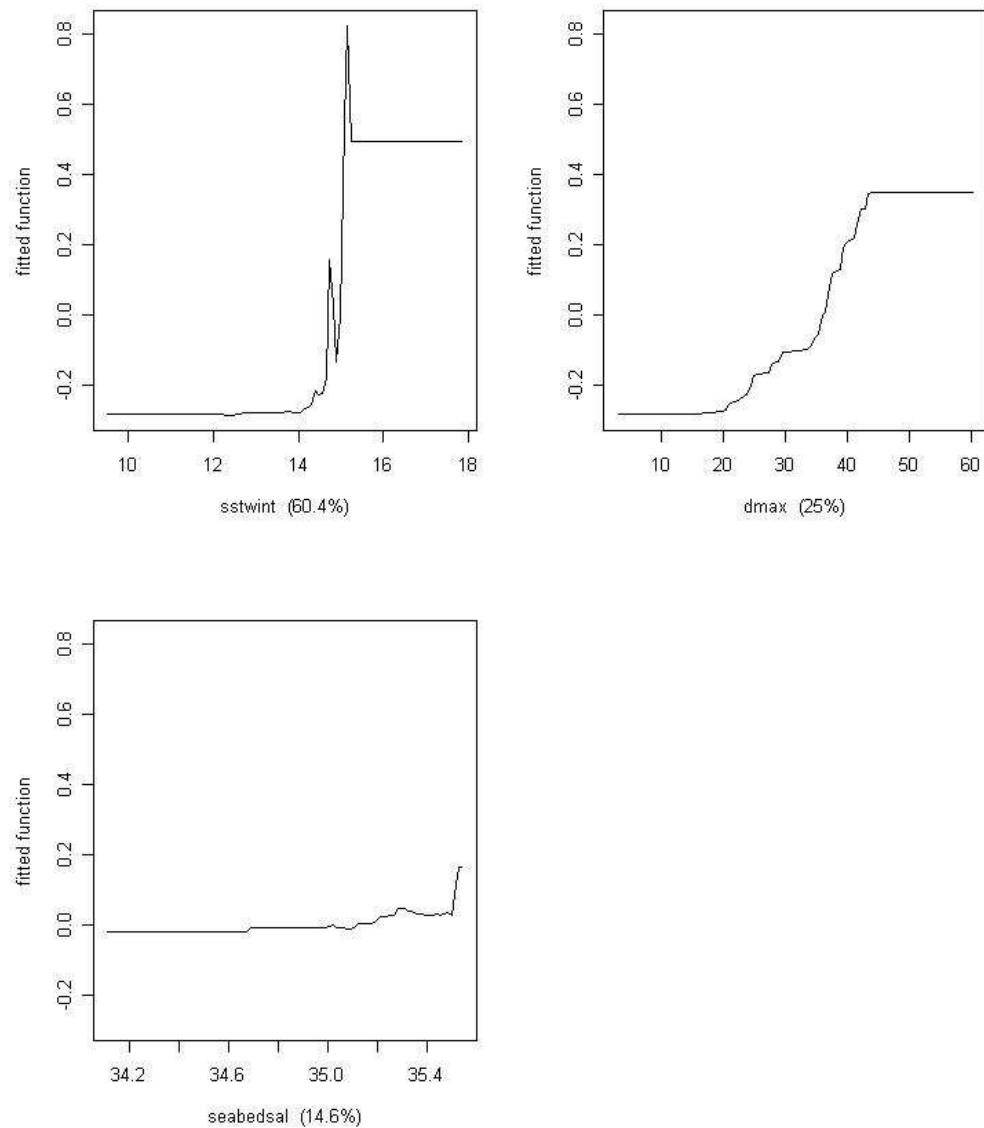
Helicolenus percoides : Sea perch



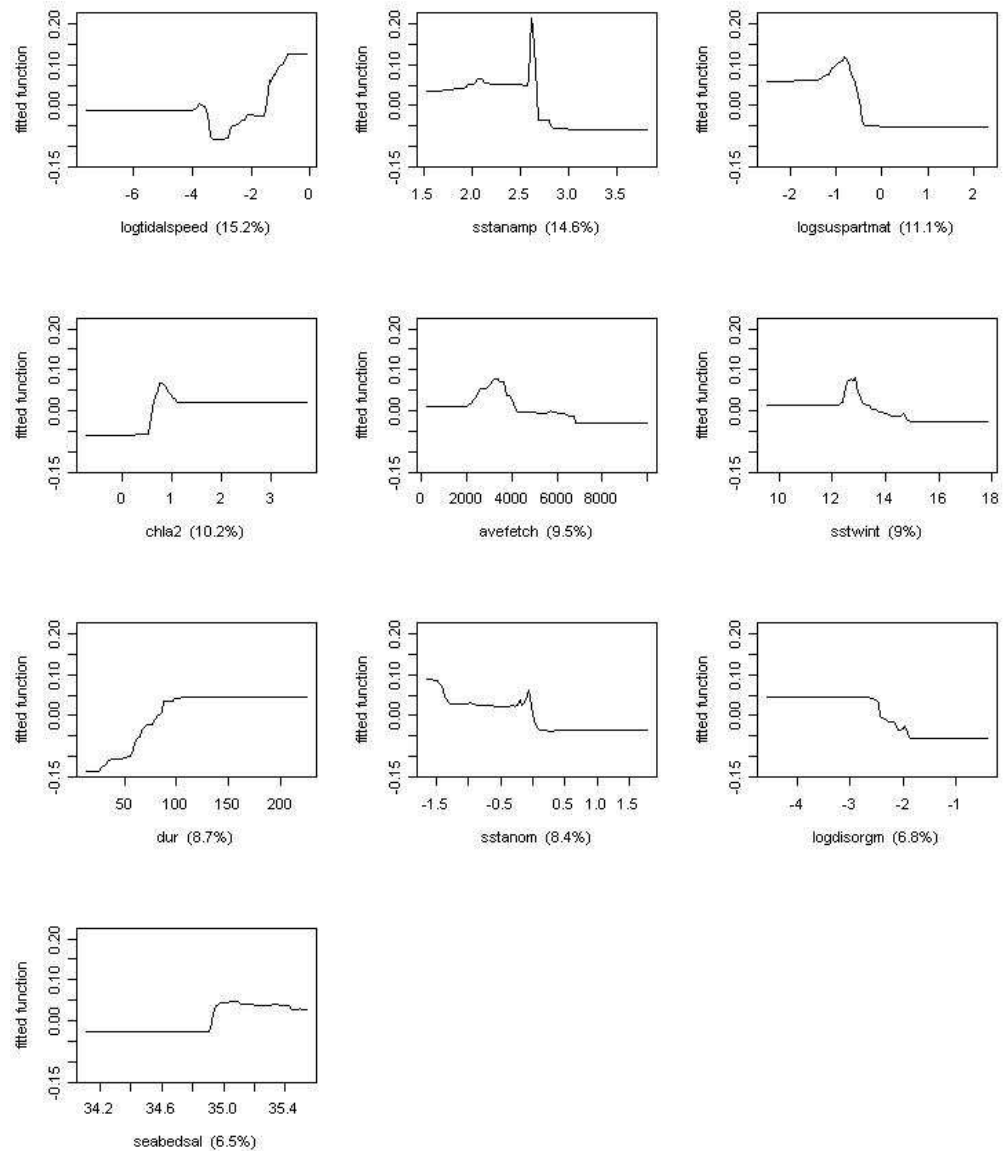
Hypoplectrodes huntii : Red-banded perch



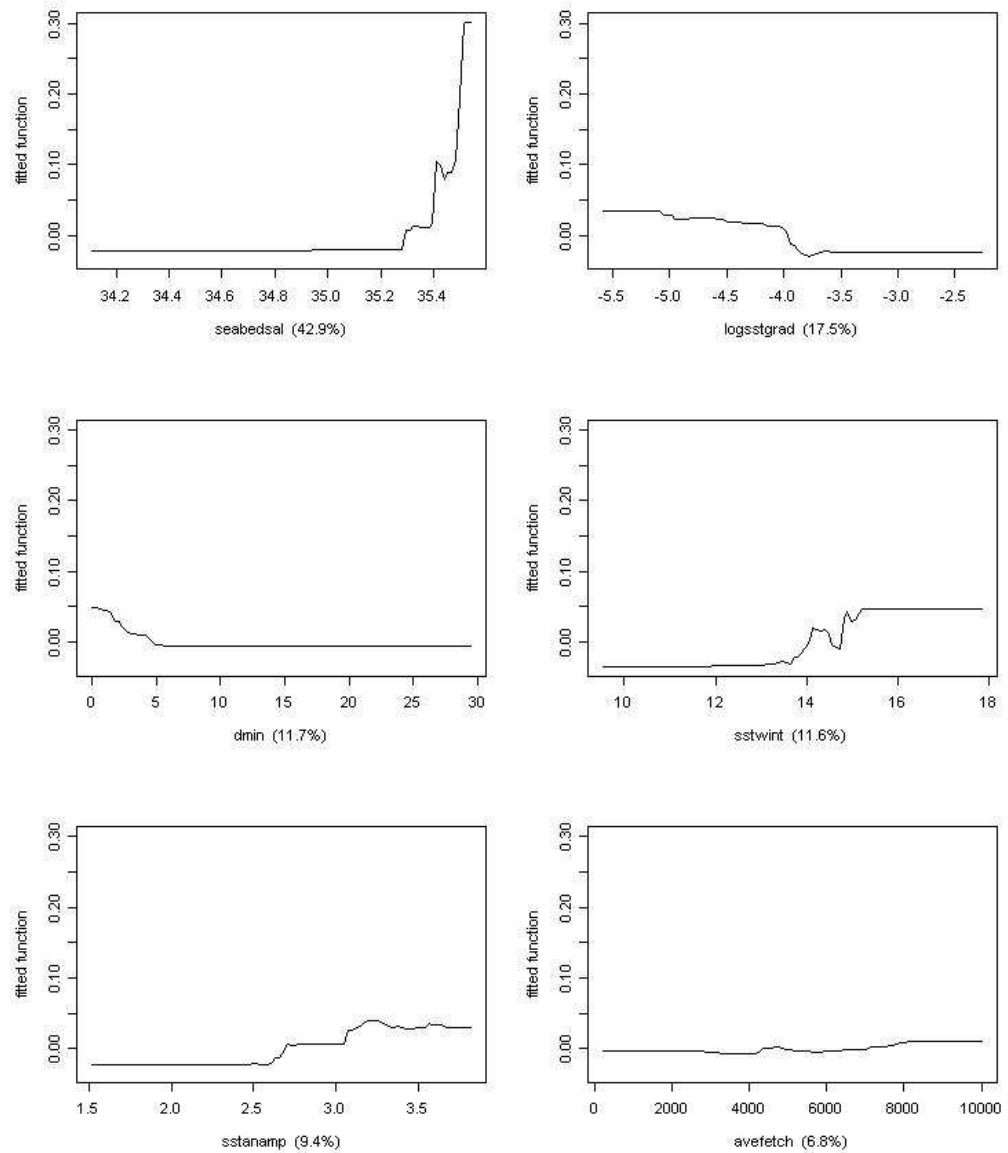
Hypoplectrodes sp.B : Half-banded perch B



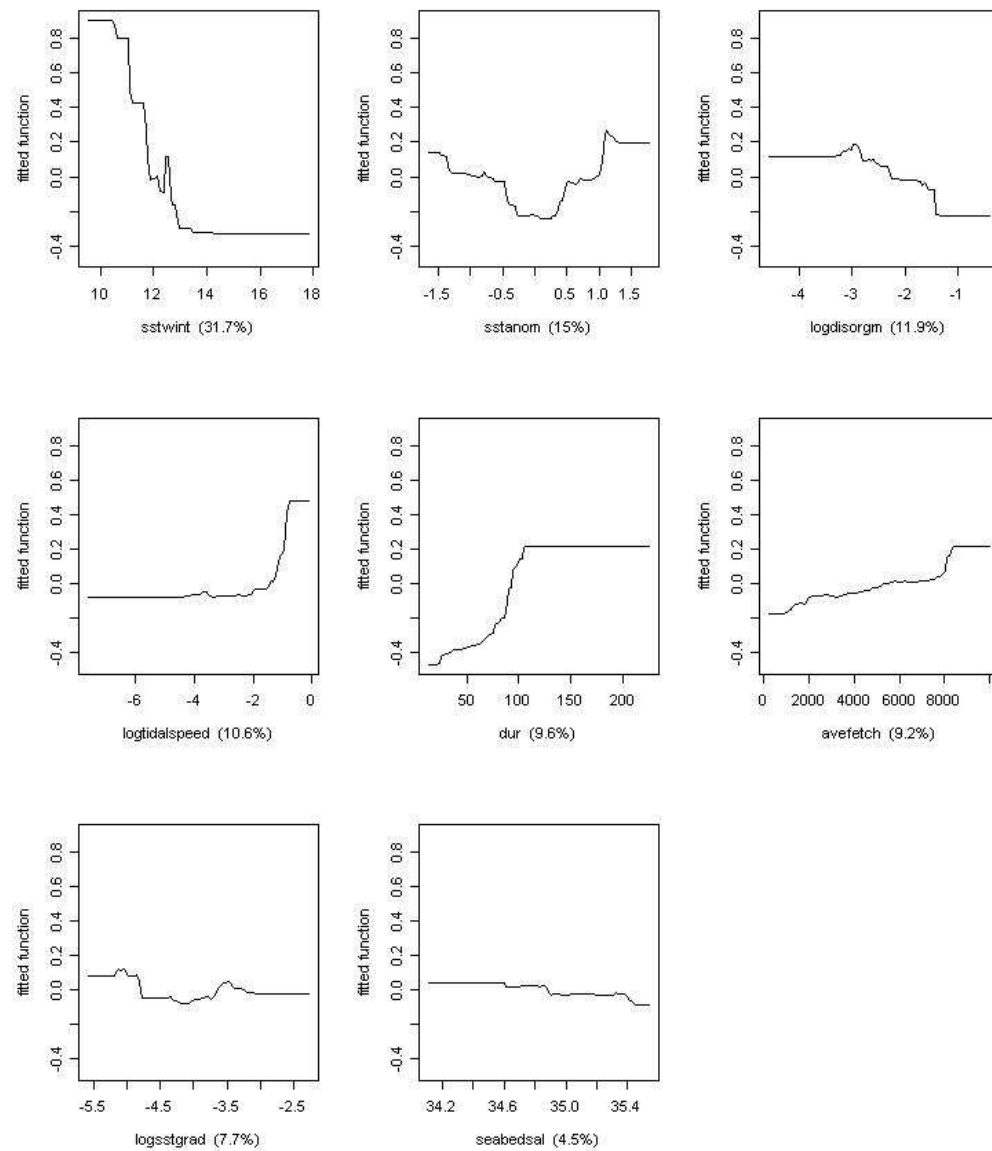
Karalepis stewarti : Scaly-headed triplefin



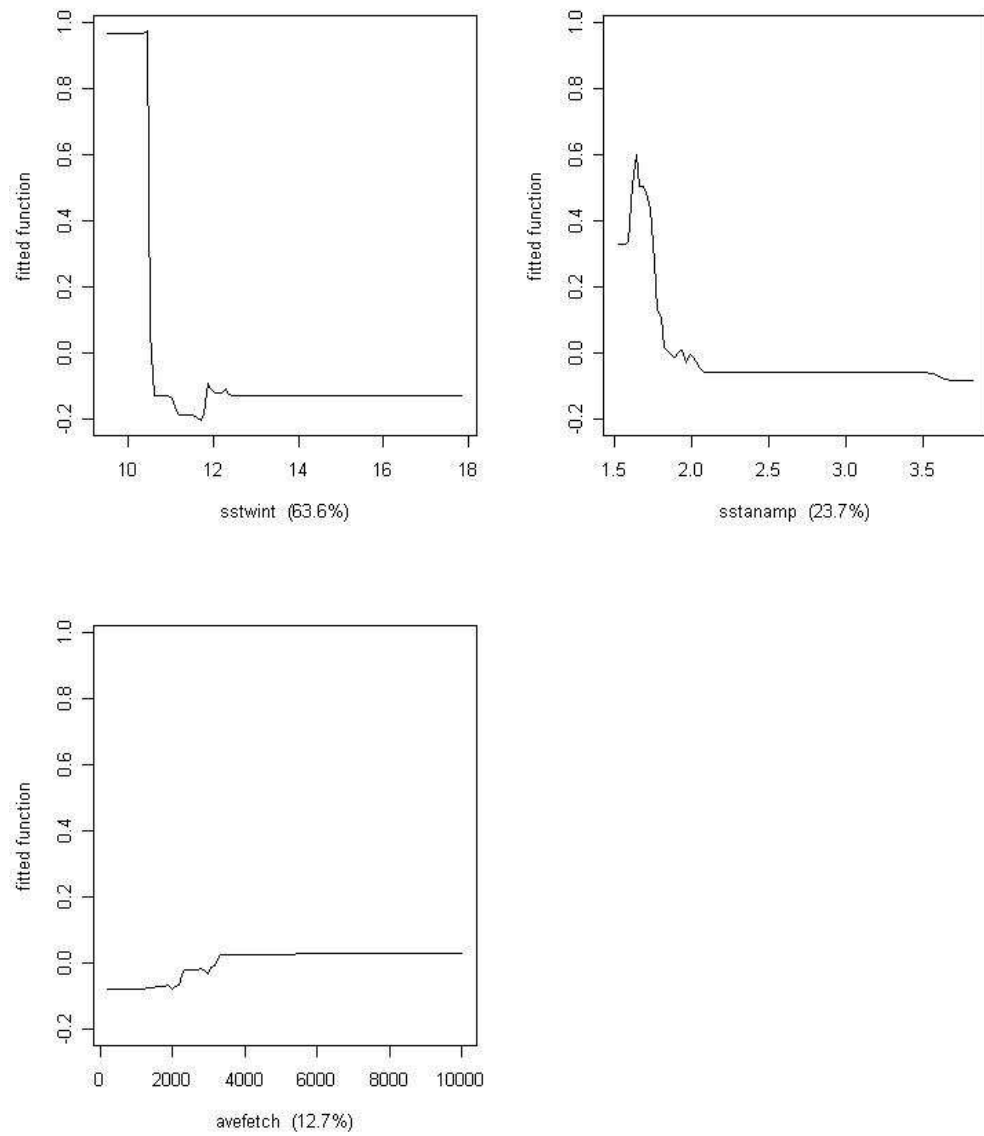
Kyphosus sydneyanus : Silver drummer



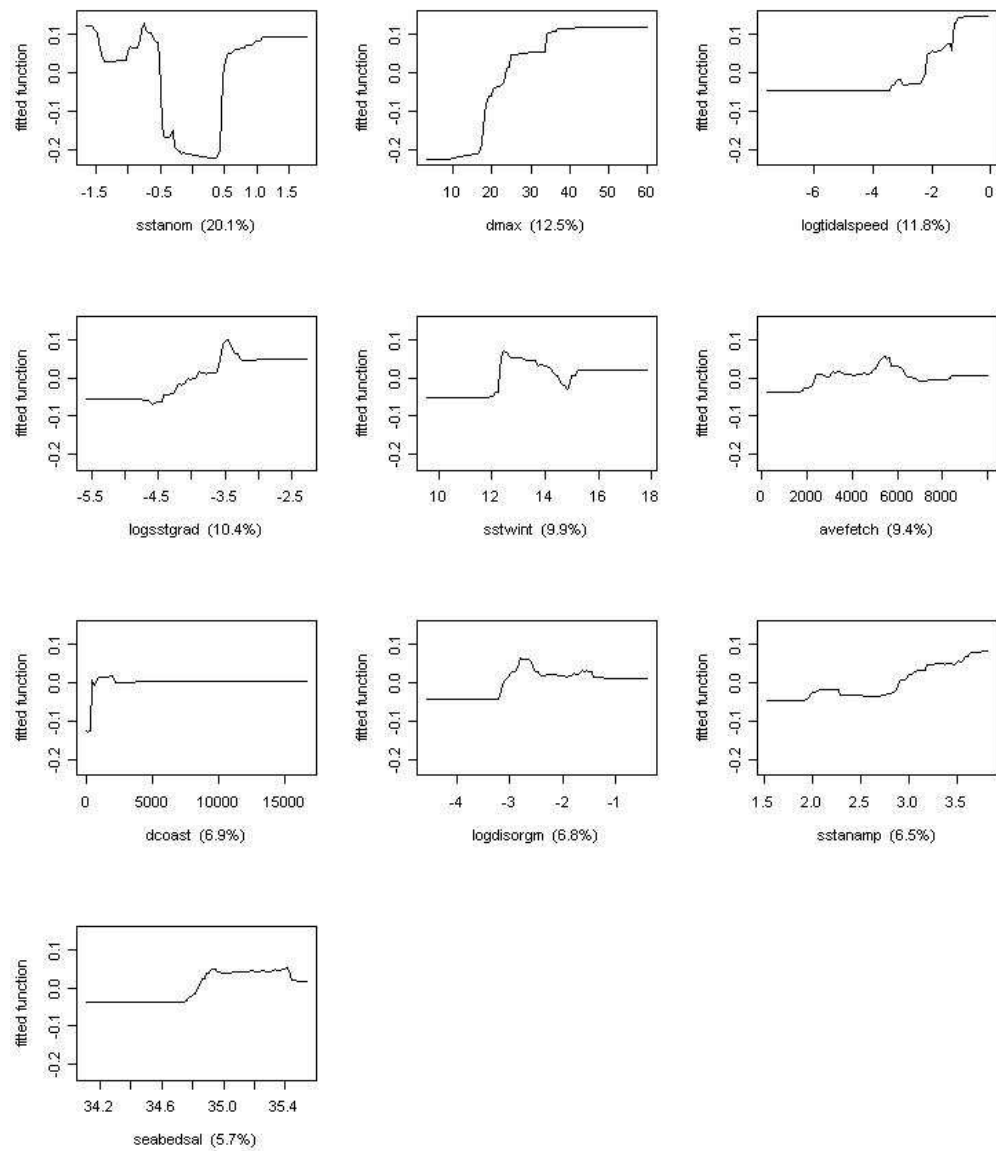
Latridopsis ciliaris : Blue moki



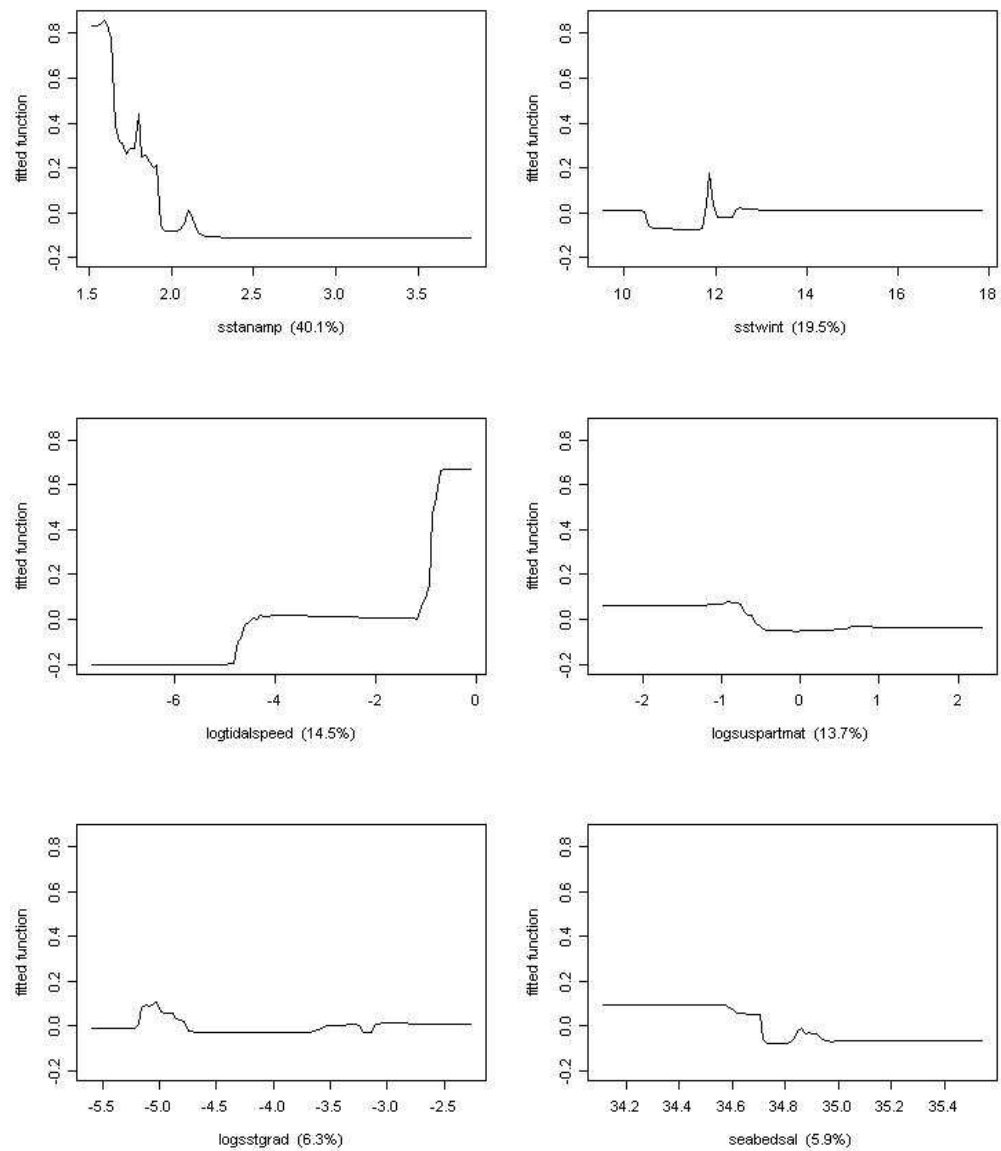
Latris lineata : Trumpeter



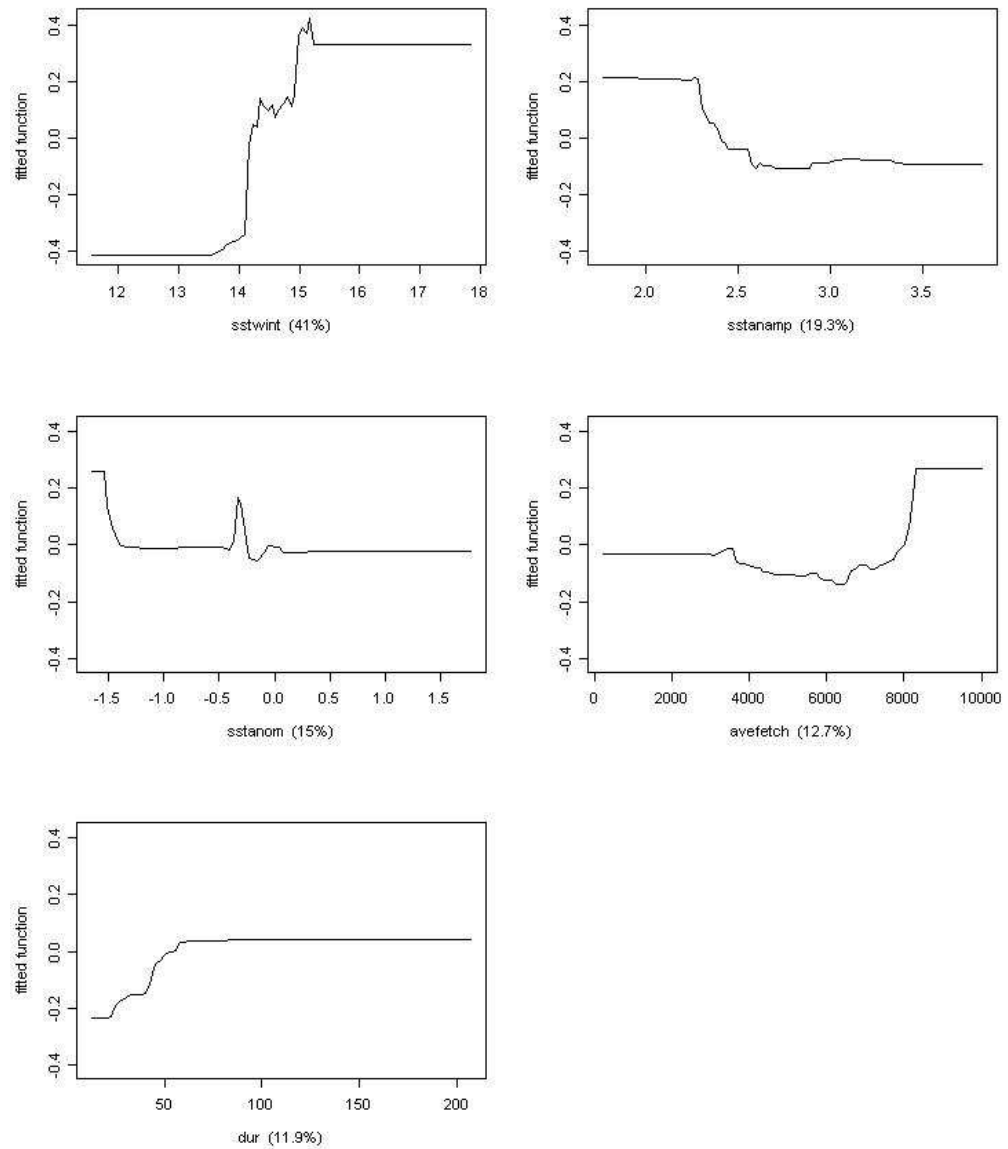
Lotella rhacinus : Rock cod



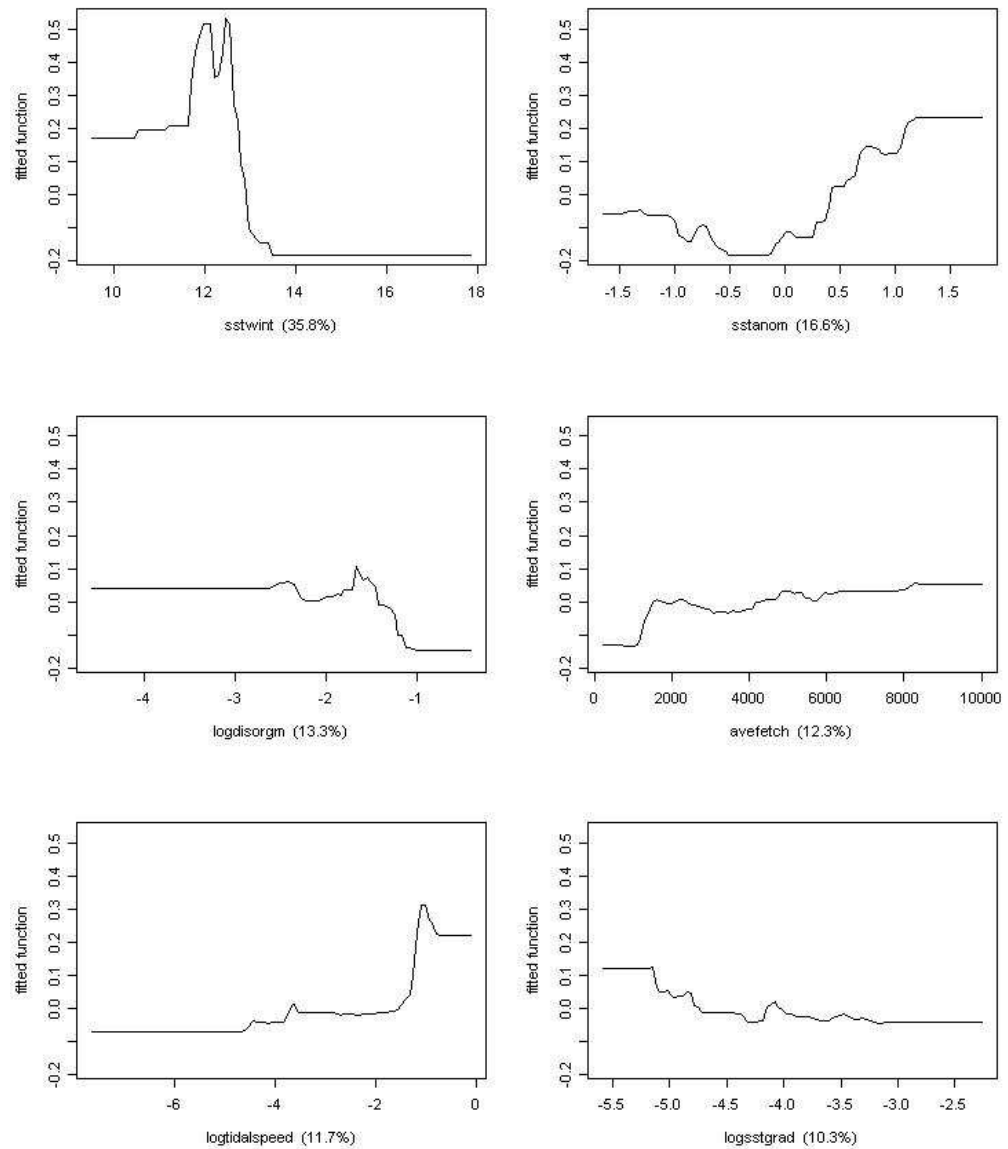
Mendosoma lineatum : Telescopefish



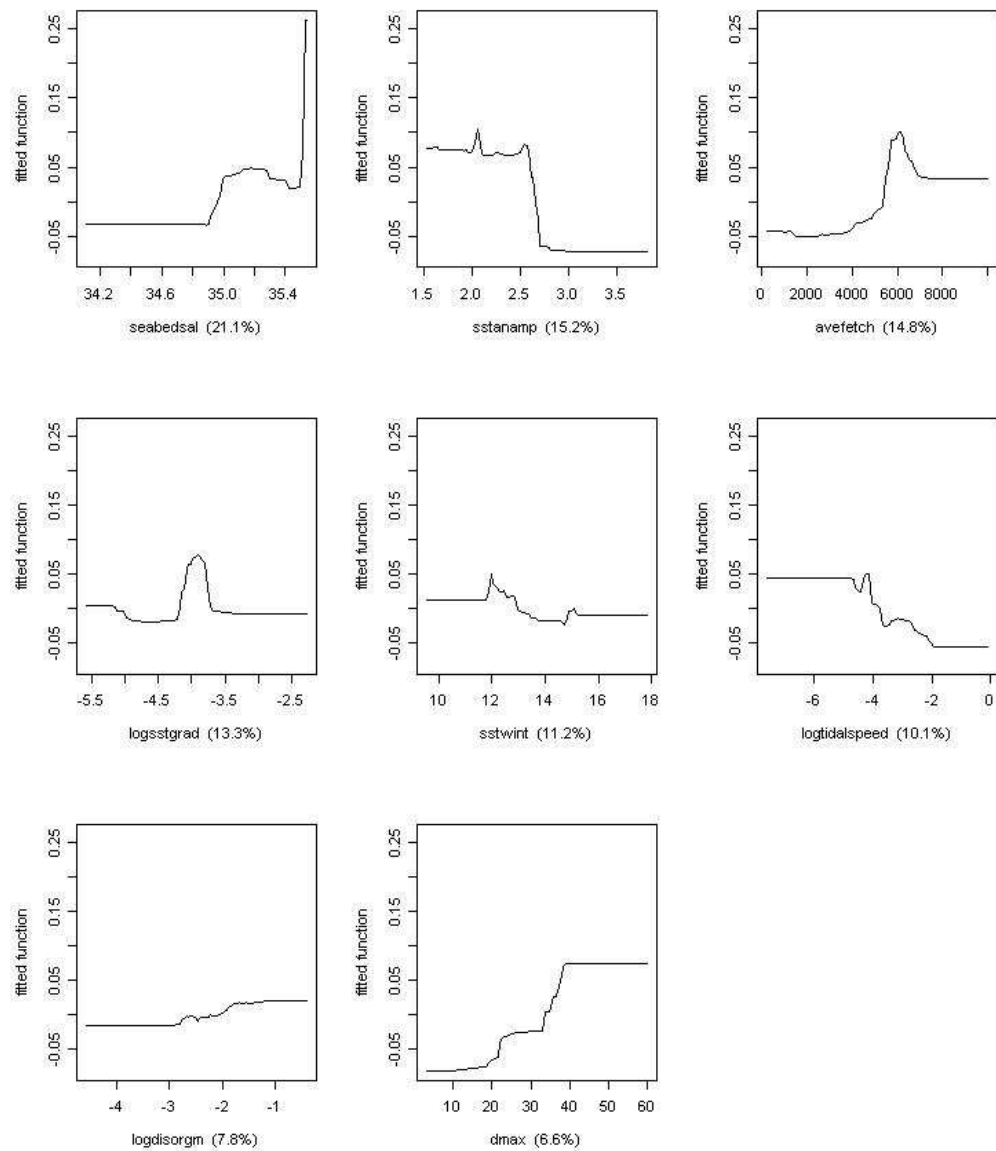
Nemadactylus douglasii : Porae



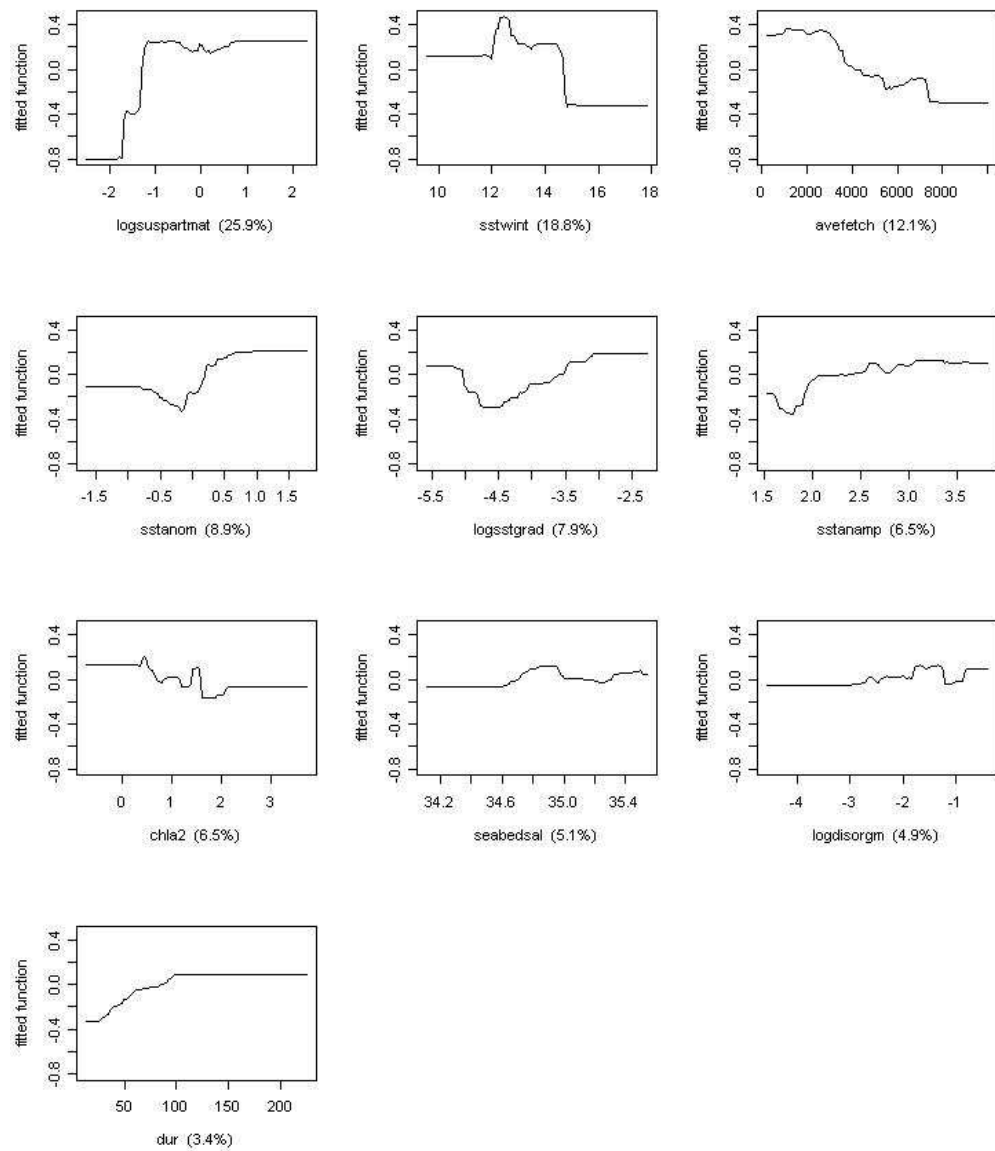
Nemadactylus macropterus : Tarakihi



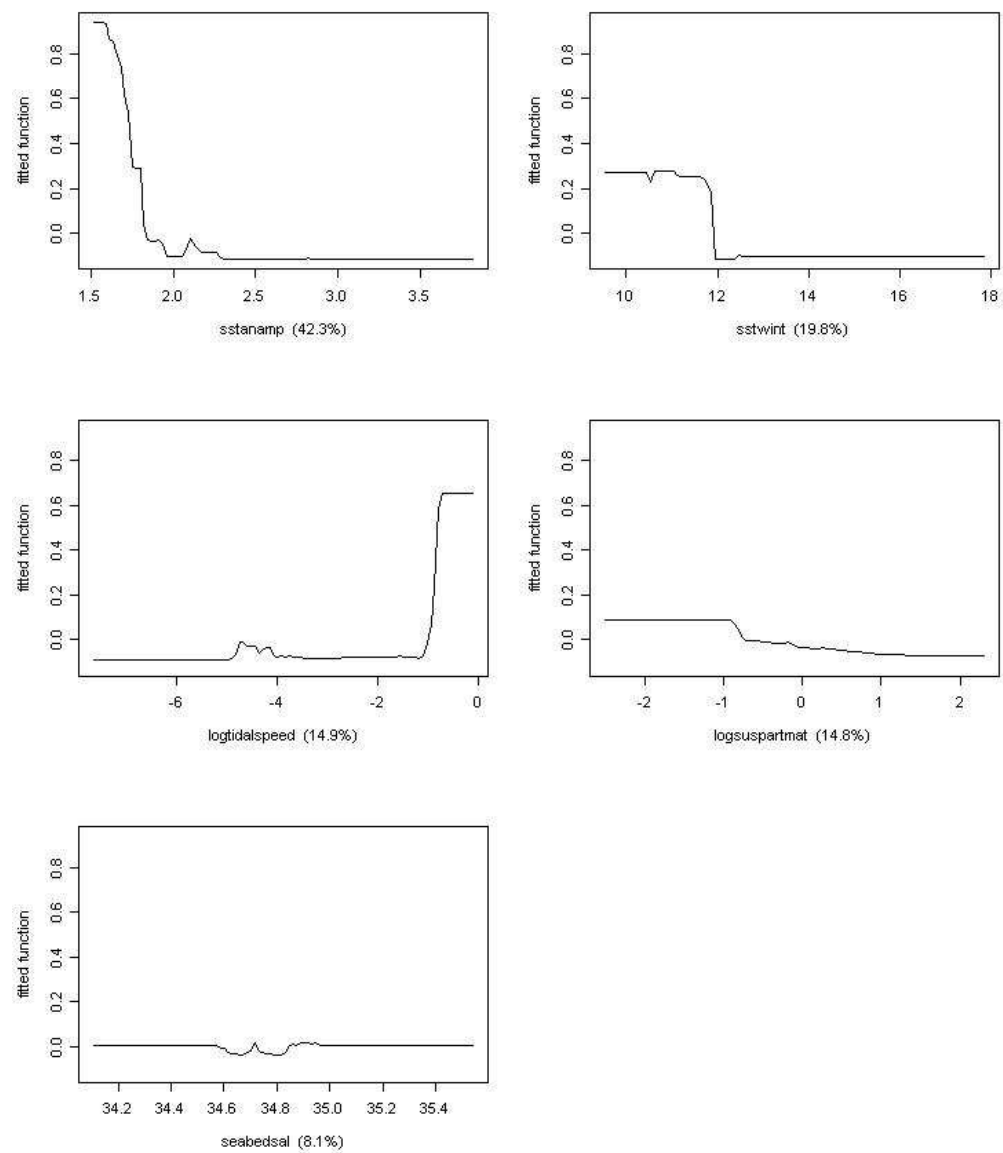
Notoclinops caerulepunctus : Blue dot triplefin



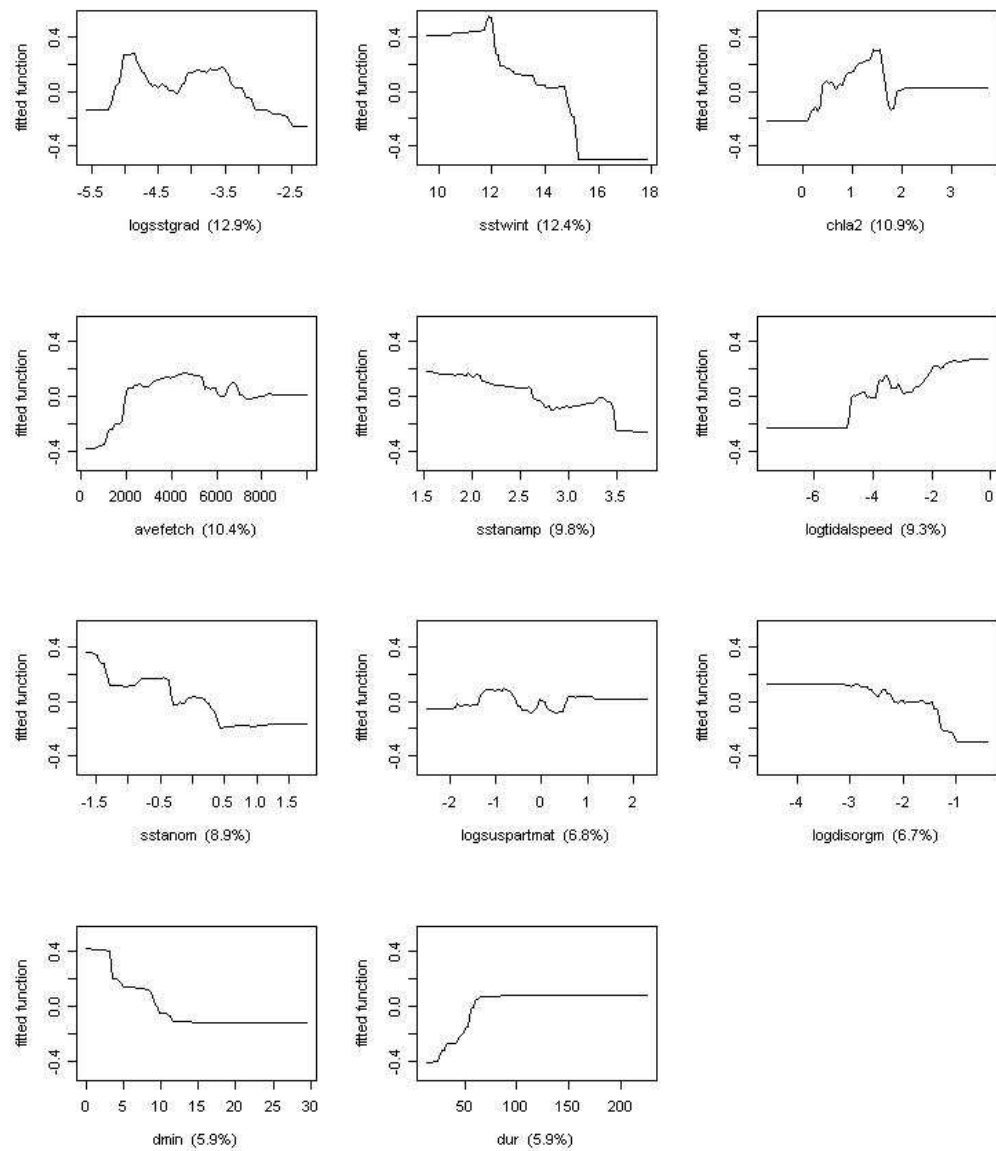
Notolabrus celidotus : Spotty



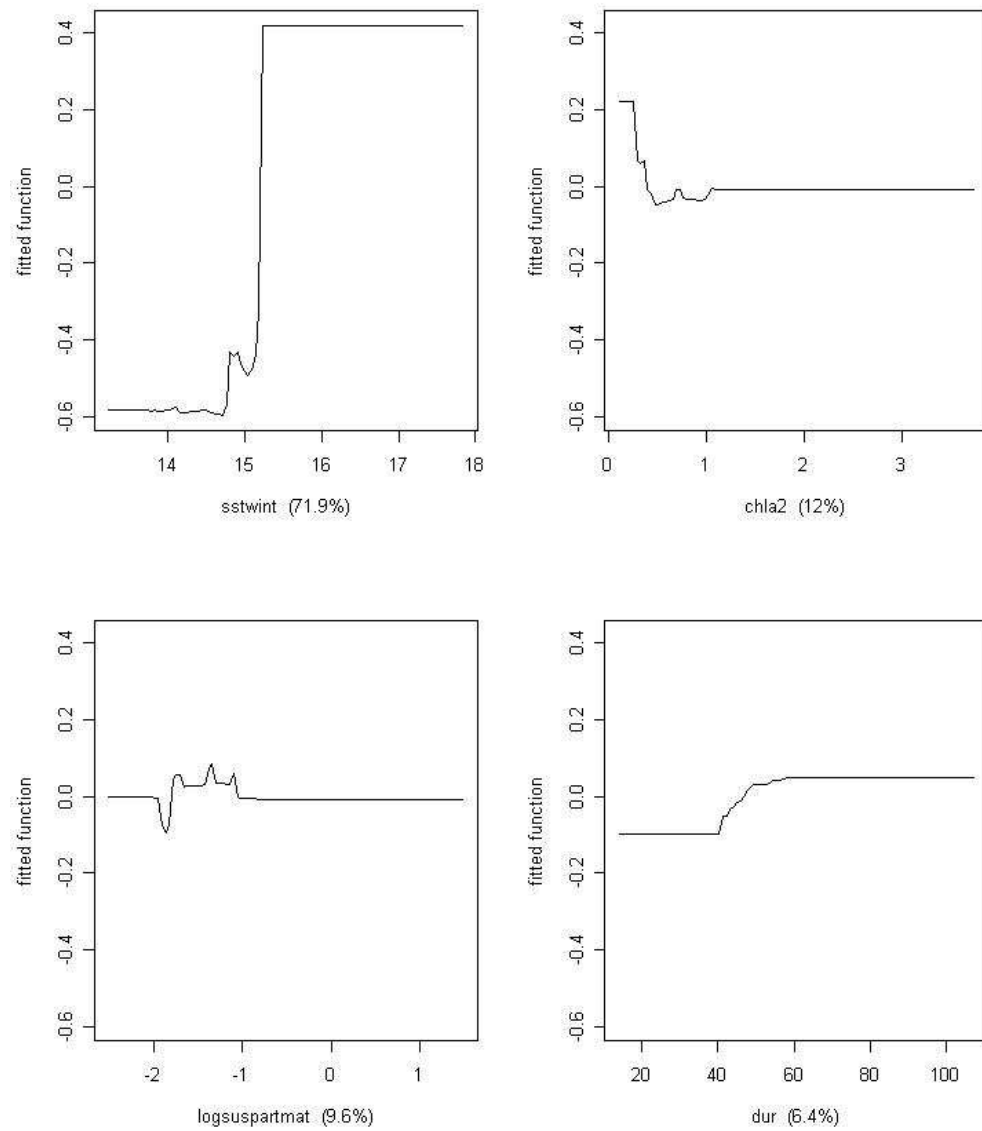
Notolabrus cinctus : Girdled wrasse



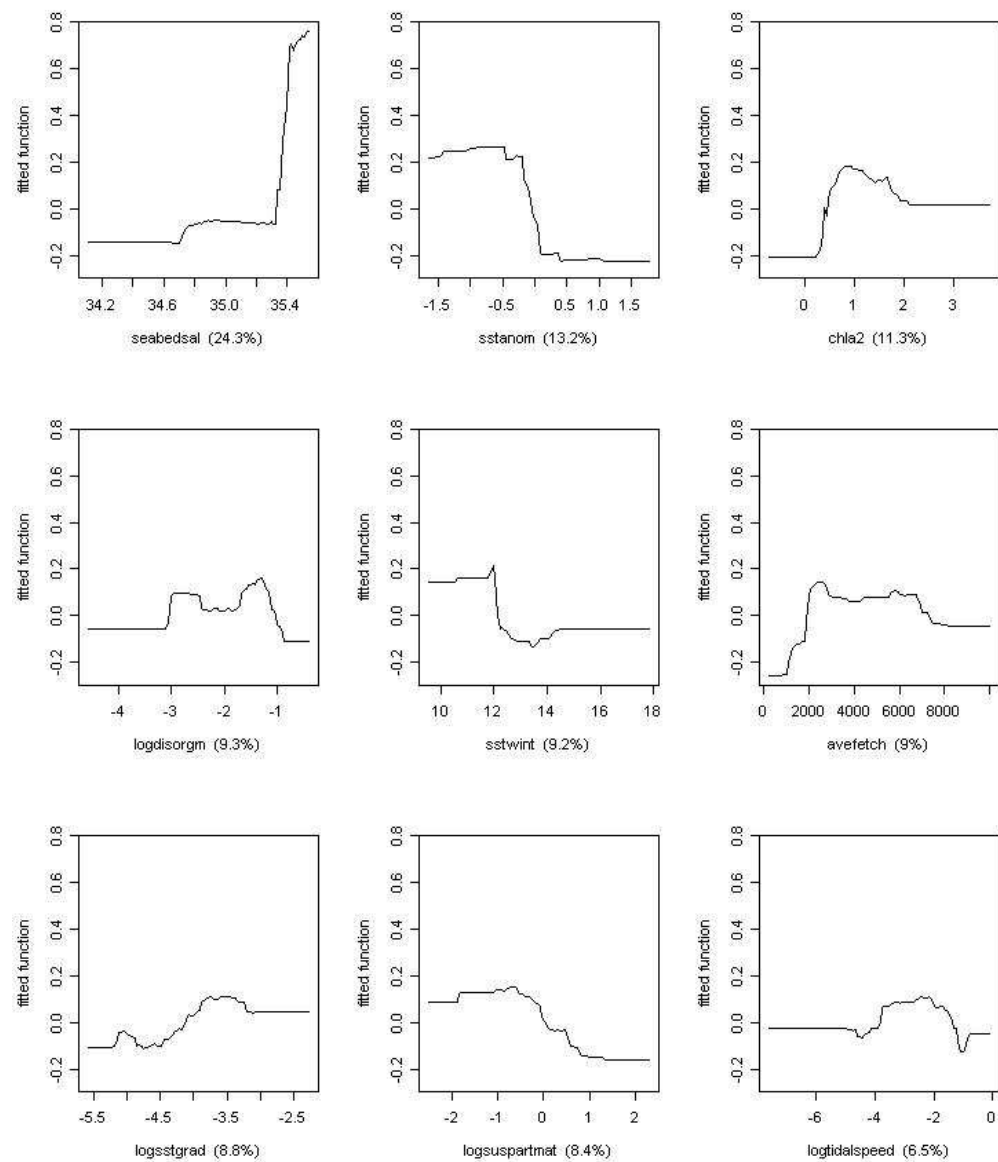
Notolabrus fucicola : Banded wrasse



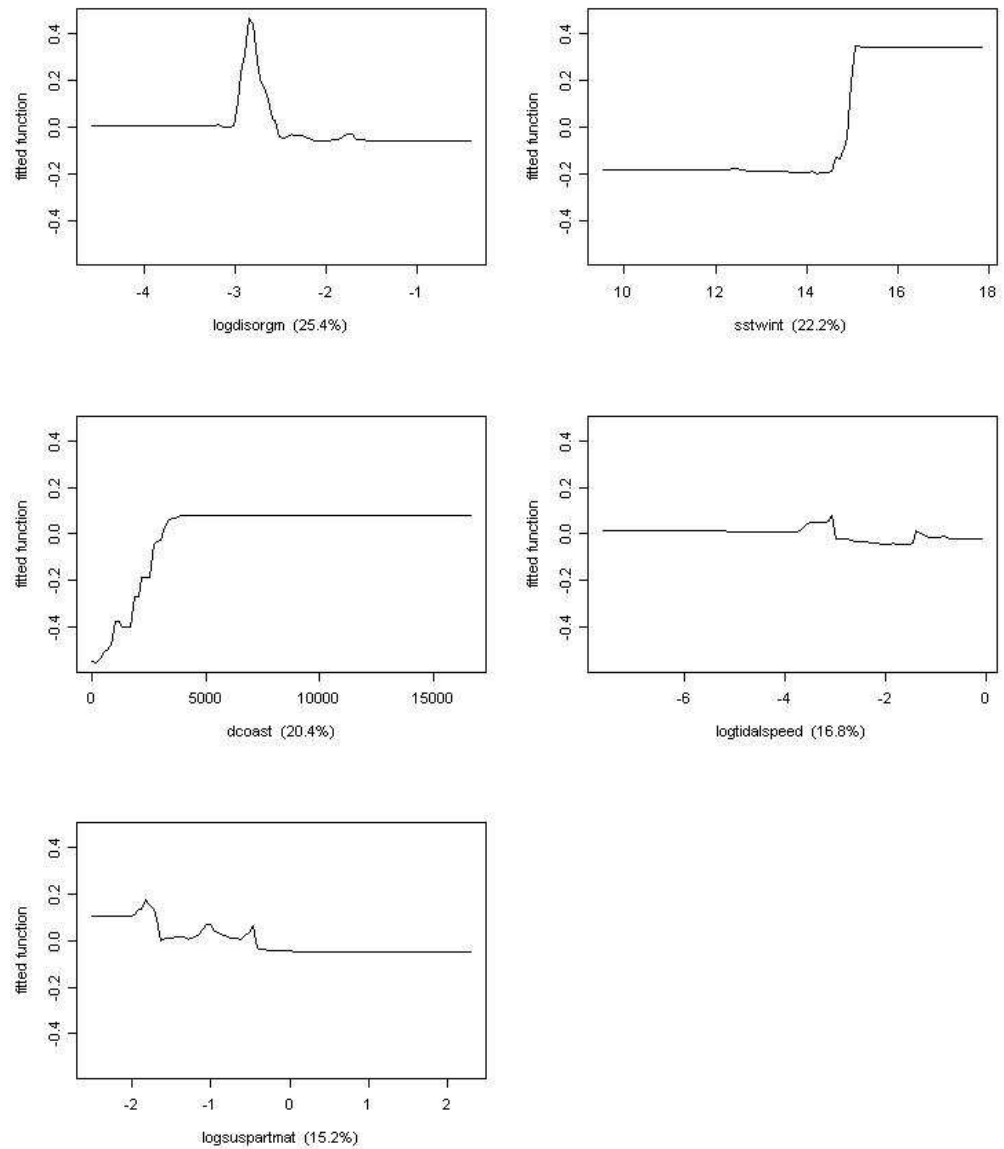
Notolabrus inscriptus : Green wrasse



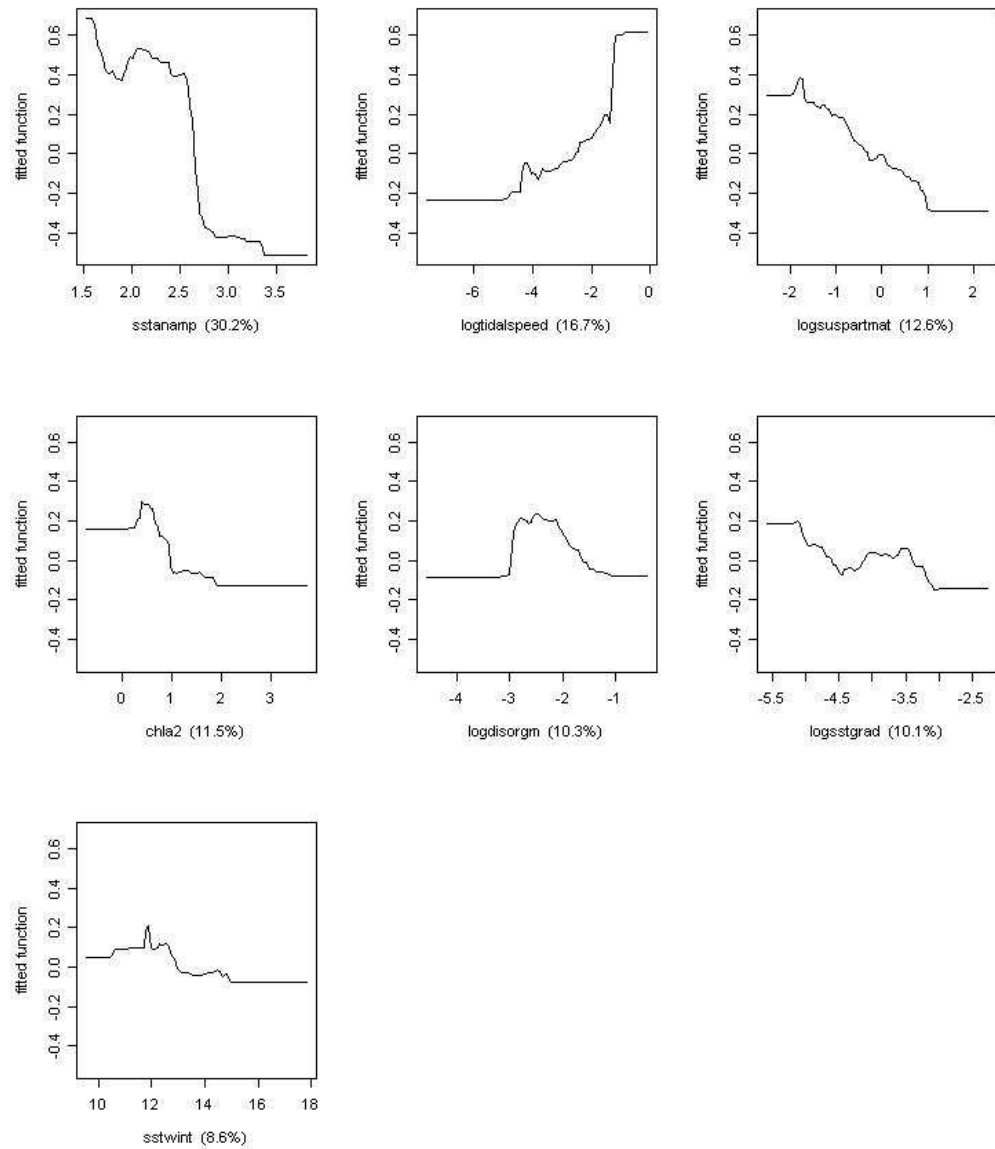
Notoclinops segmentatus : Blue-eyed triplefin



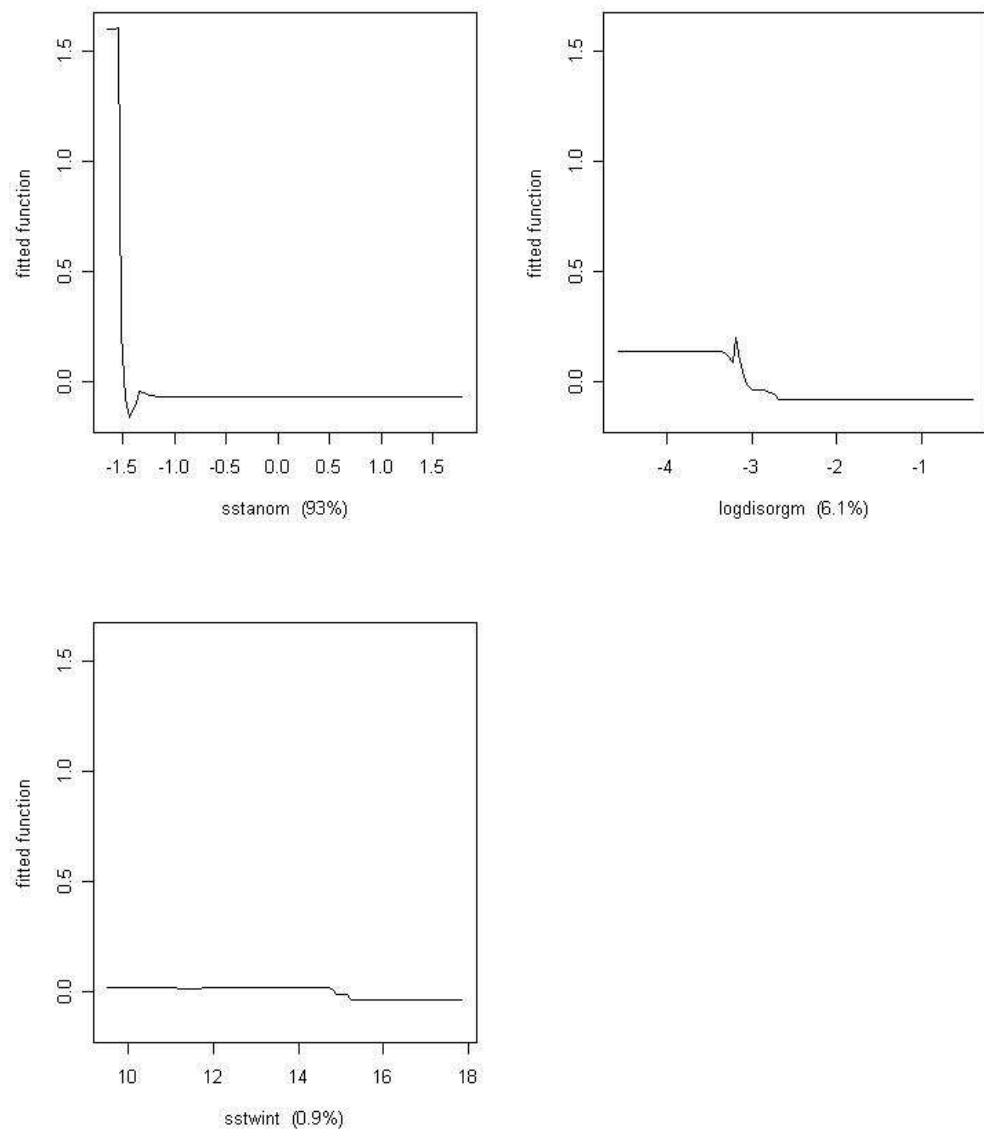
Notoclinops yaldwyni : Yaldwyn's triplefin



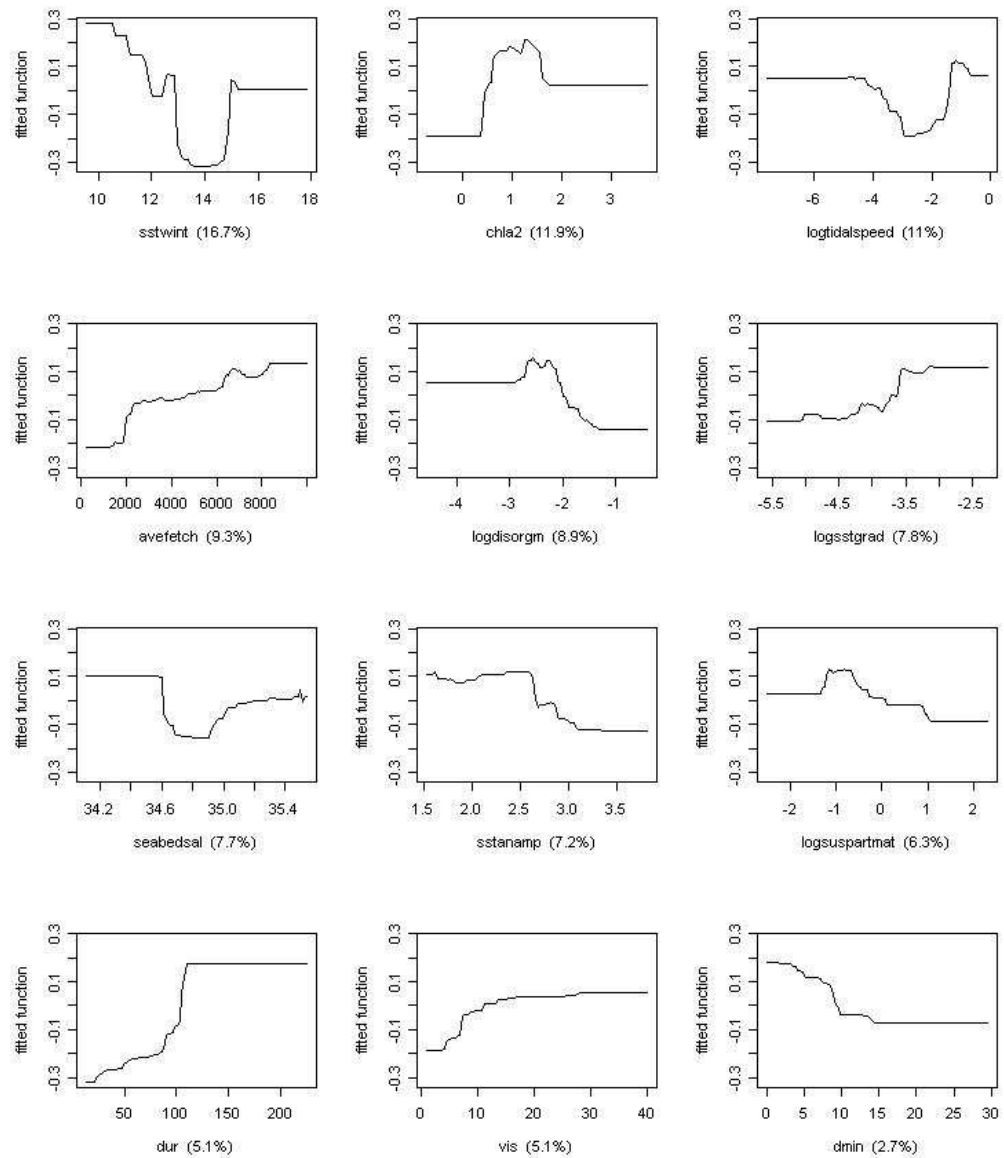
Obliquichthys maryannae : Oblique-swimming triplefin



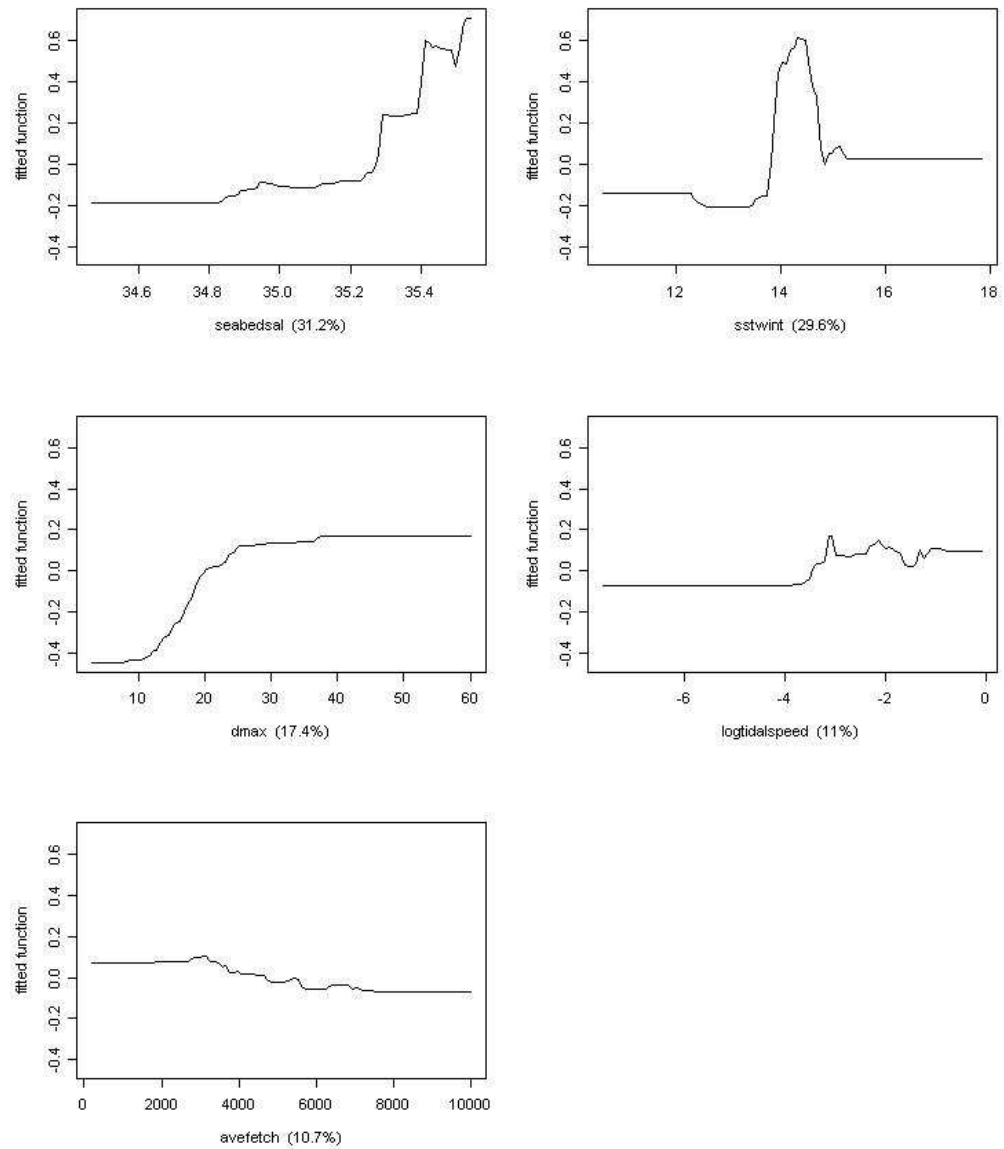
Odax cyanoallix : Blue-finned butterflyfish



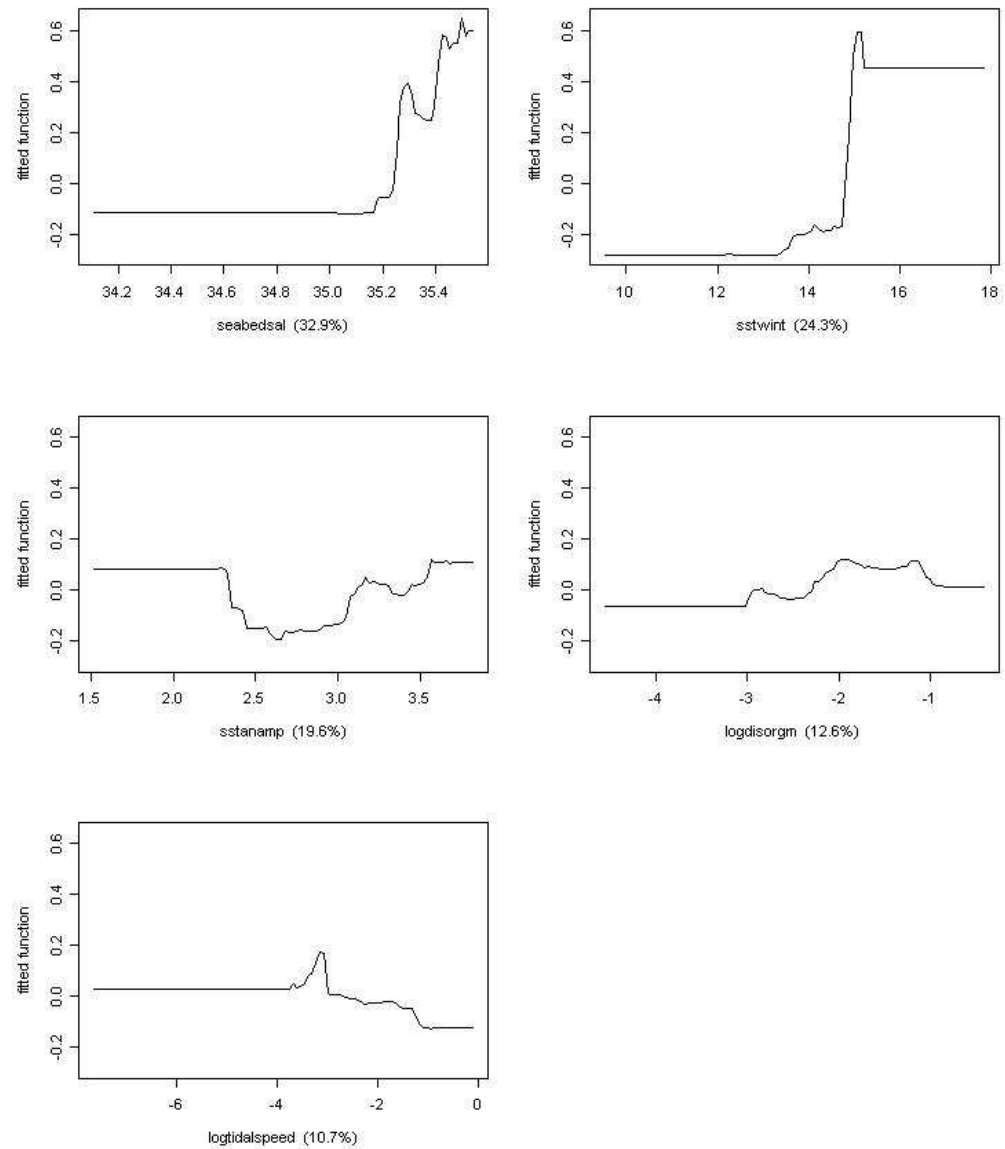
Odax pullus : Butterfish



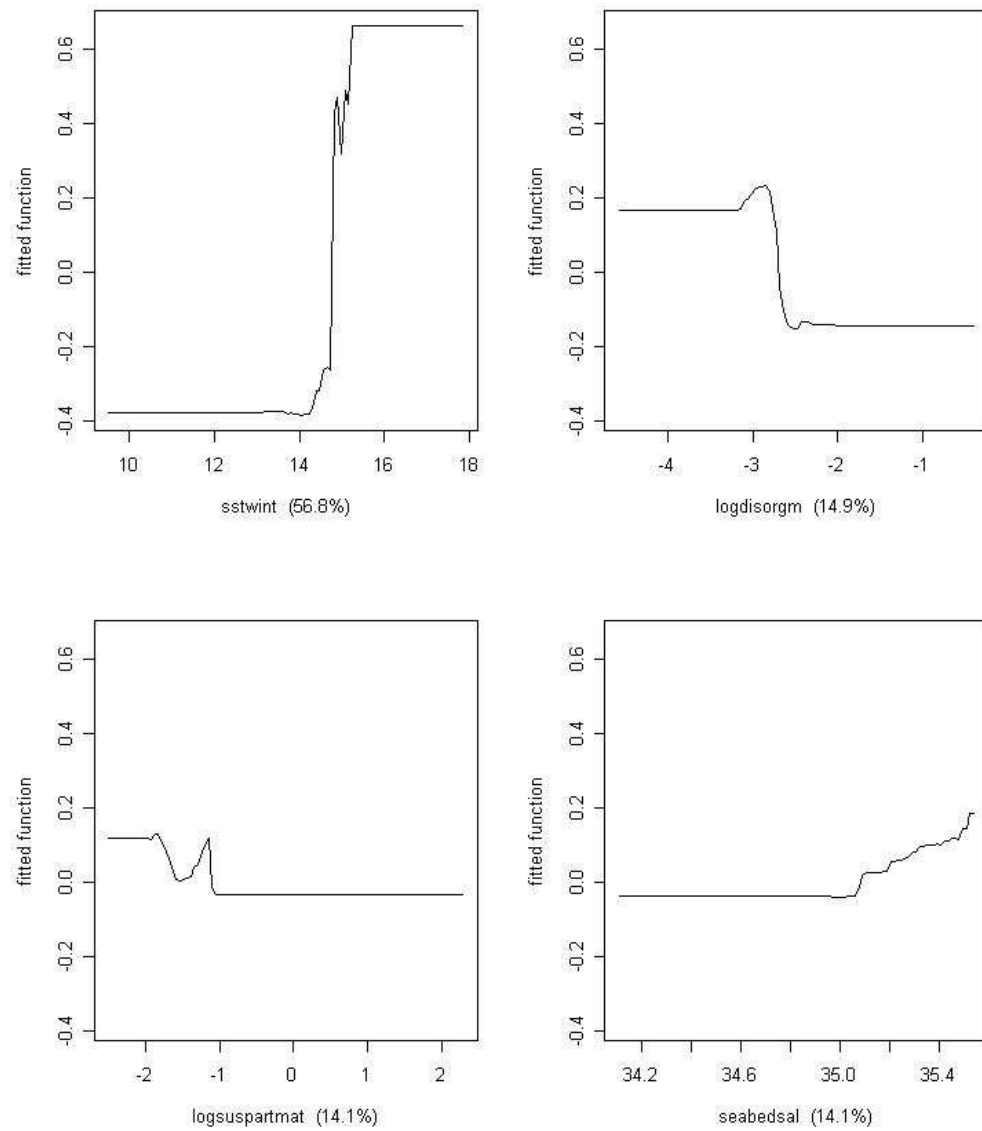
Optivus elongatus : Slender roughy



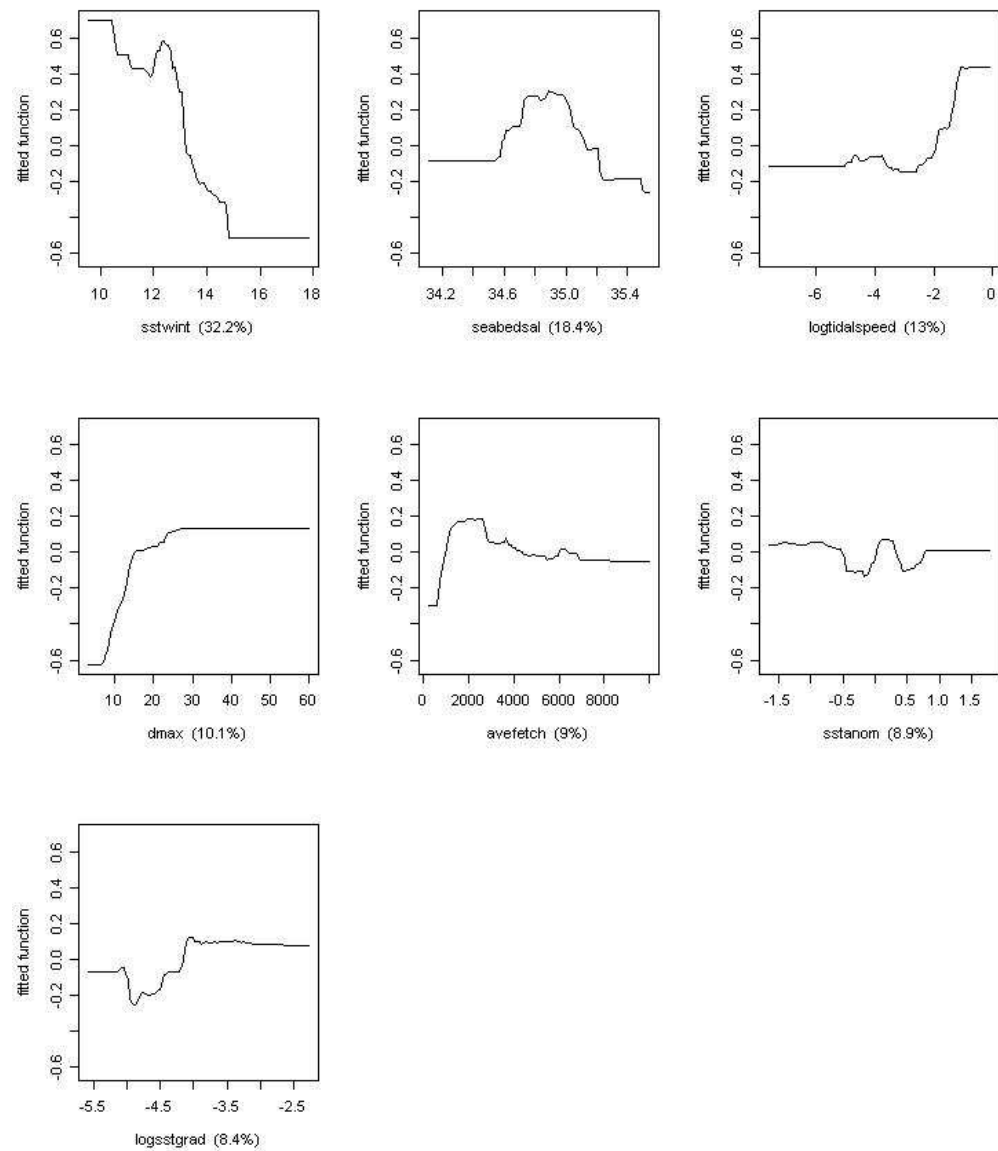
Pagrus auratus : Snapper



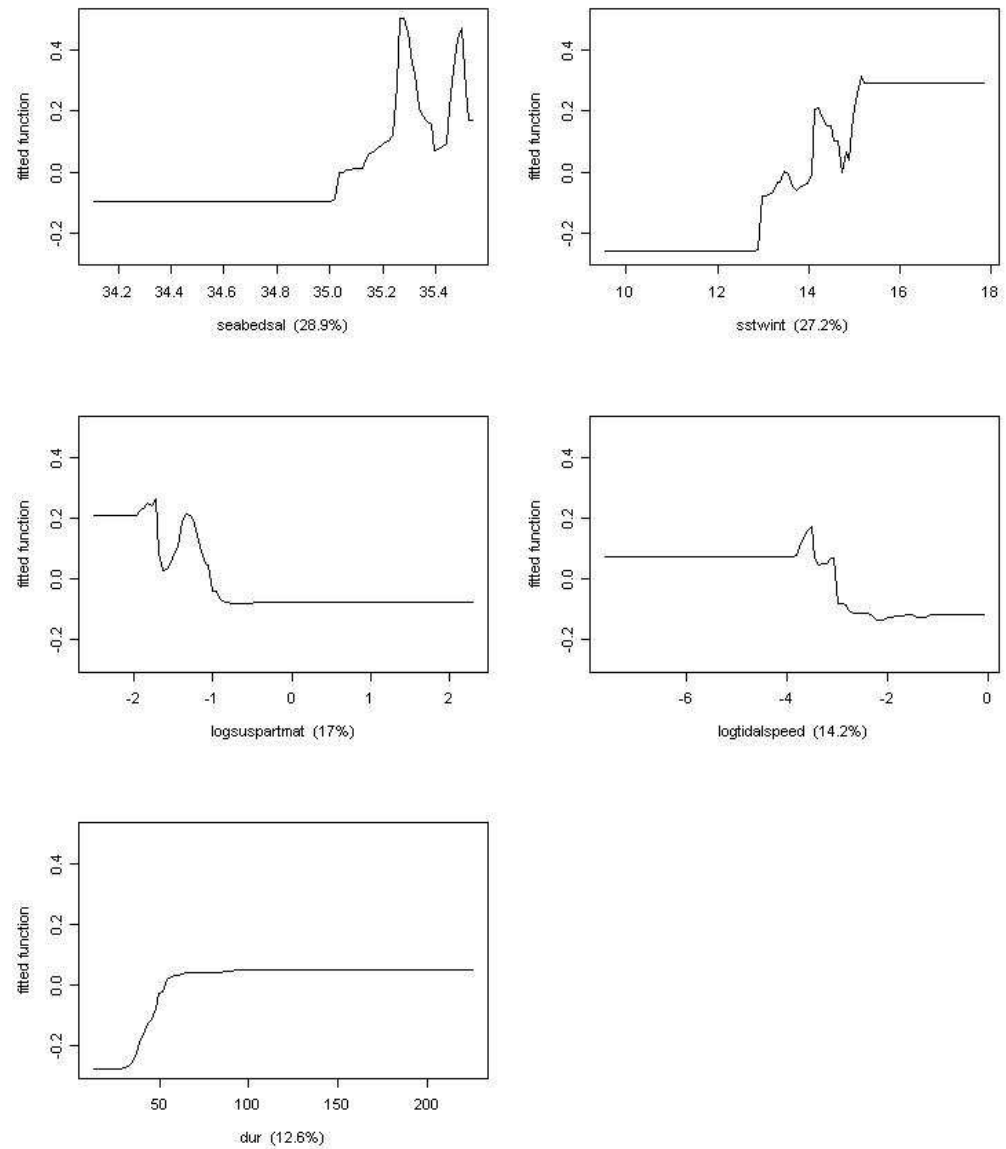
Parma alboscaphularis : Black angelfish



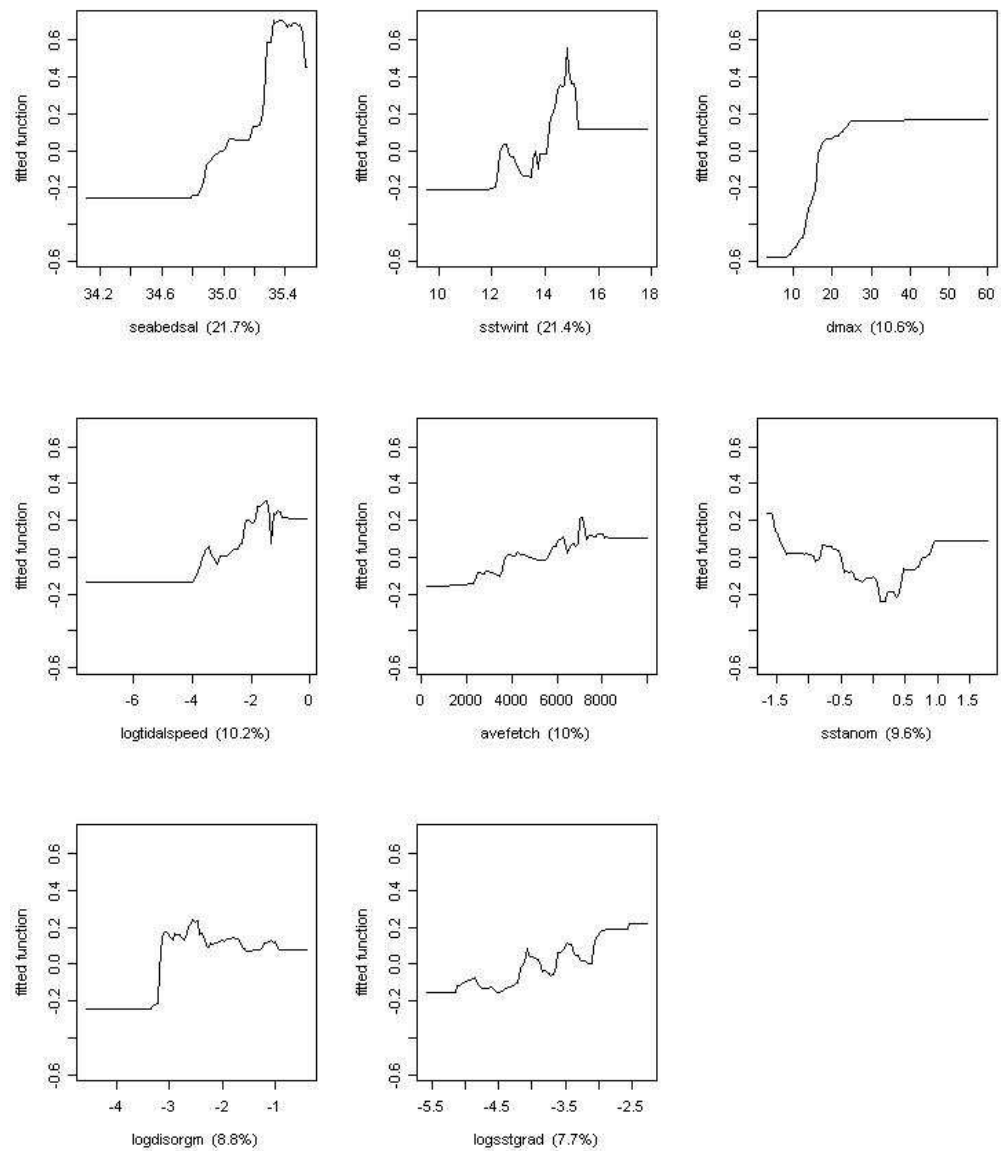
Parapercis colias : Blue cod



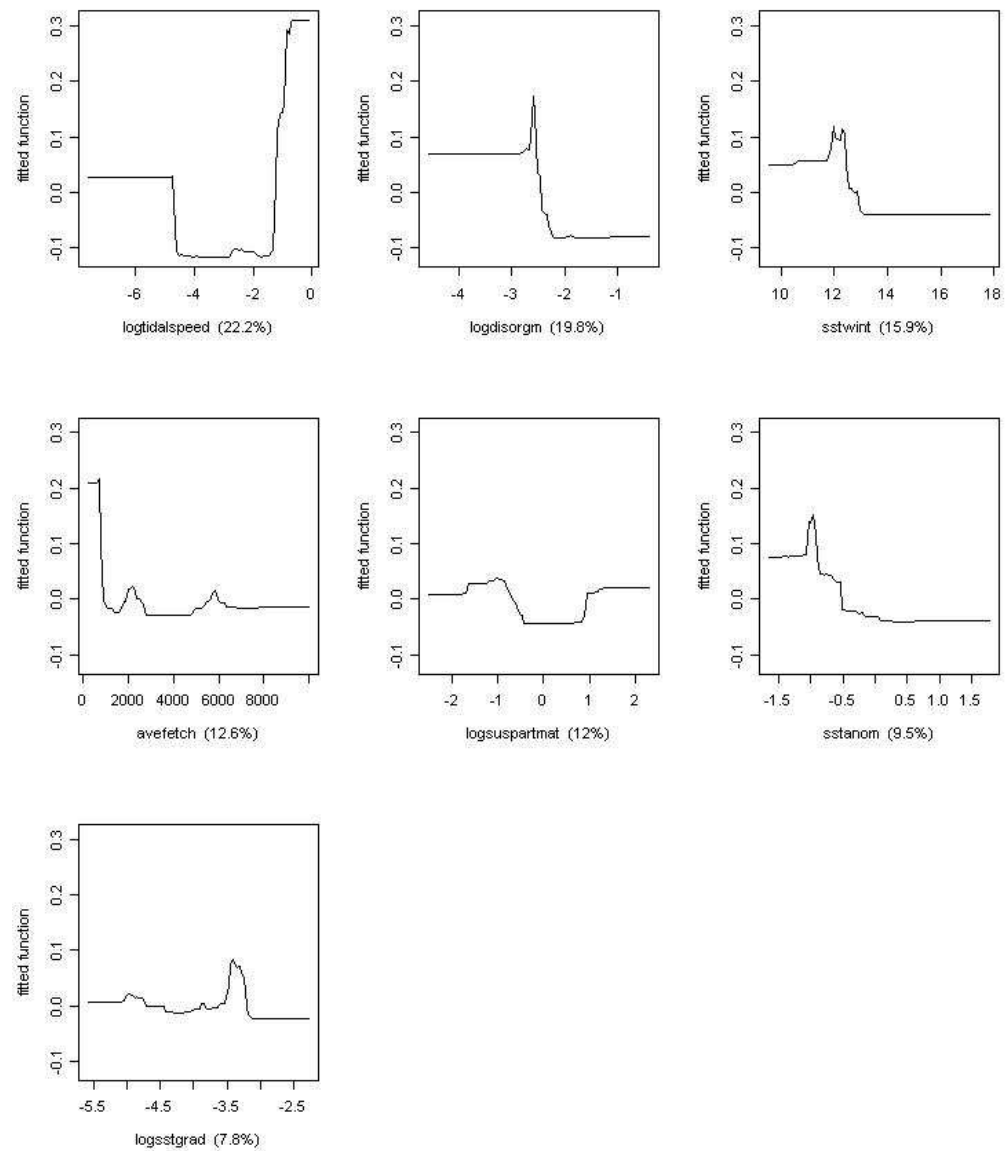
Parablennius laticlavius : Crested blenny



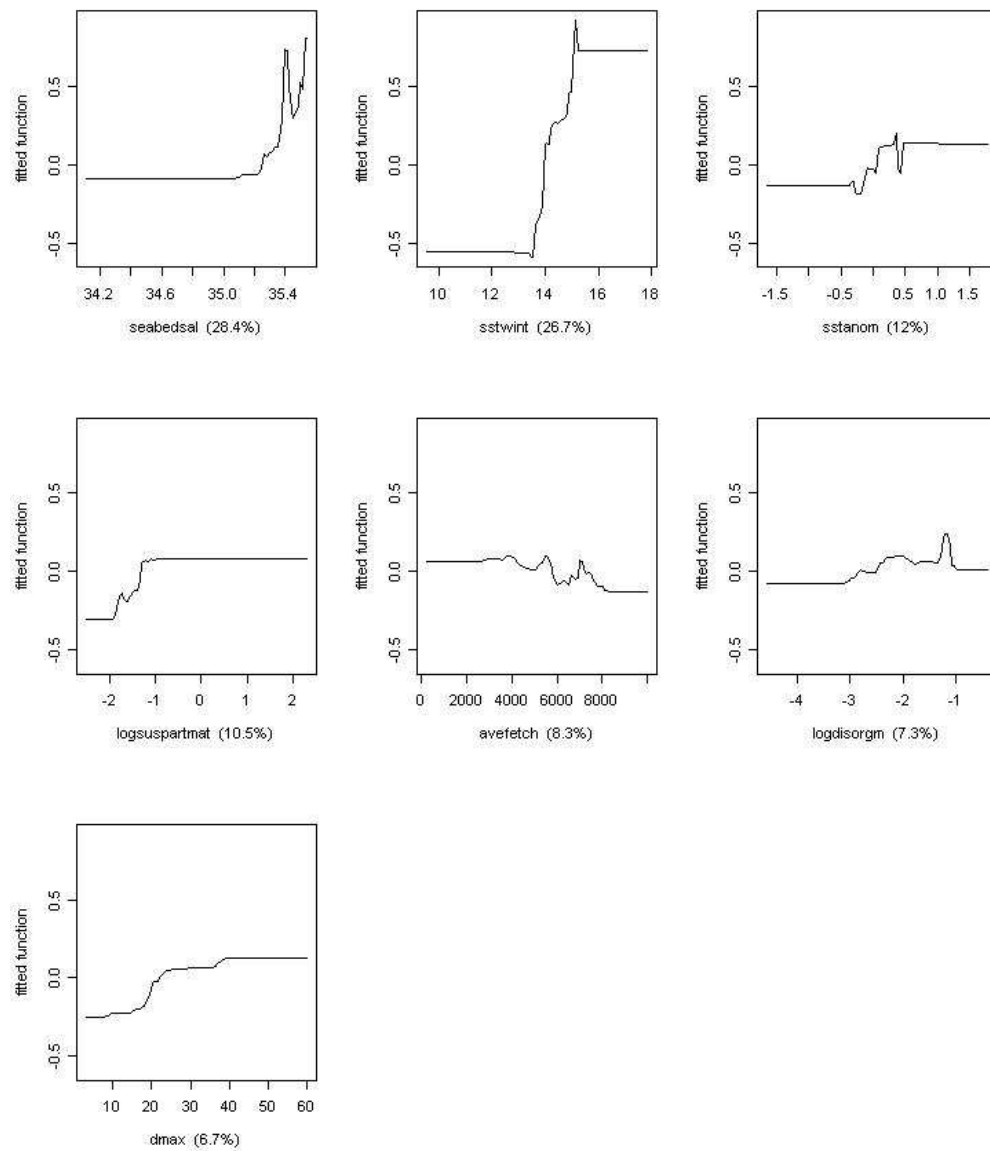
Parika scaber : Leatherjacket



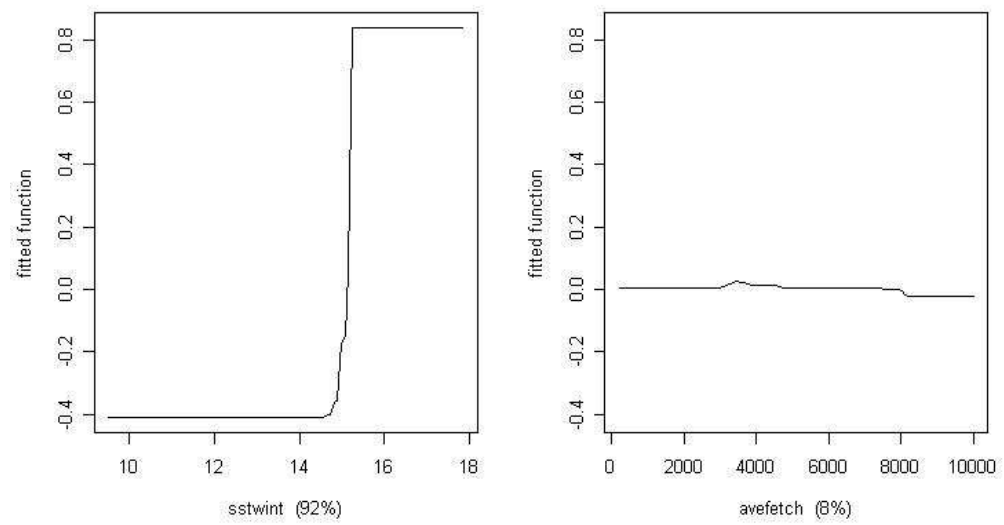
Paratrachichthys trailli : Common roughy



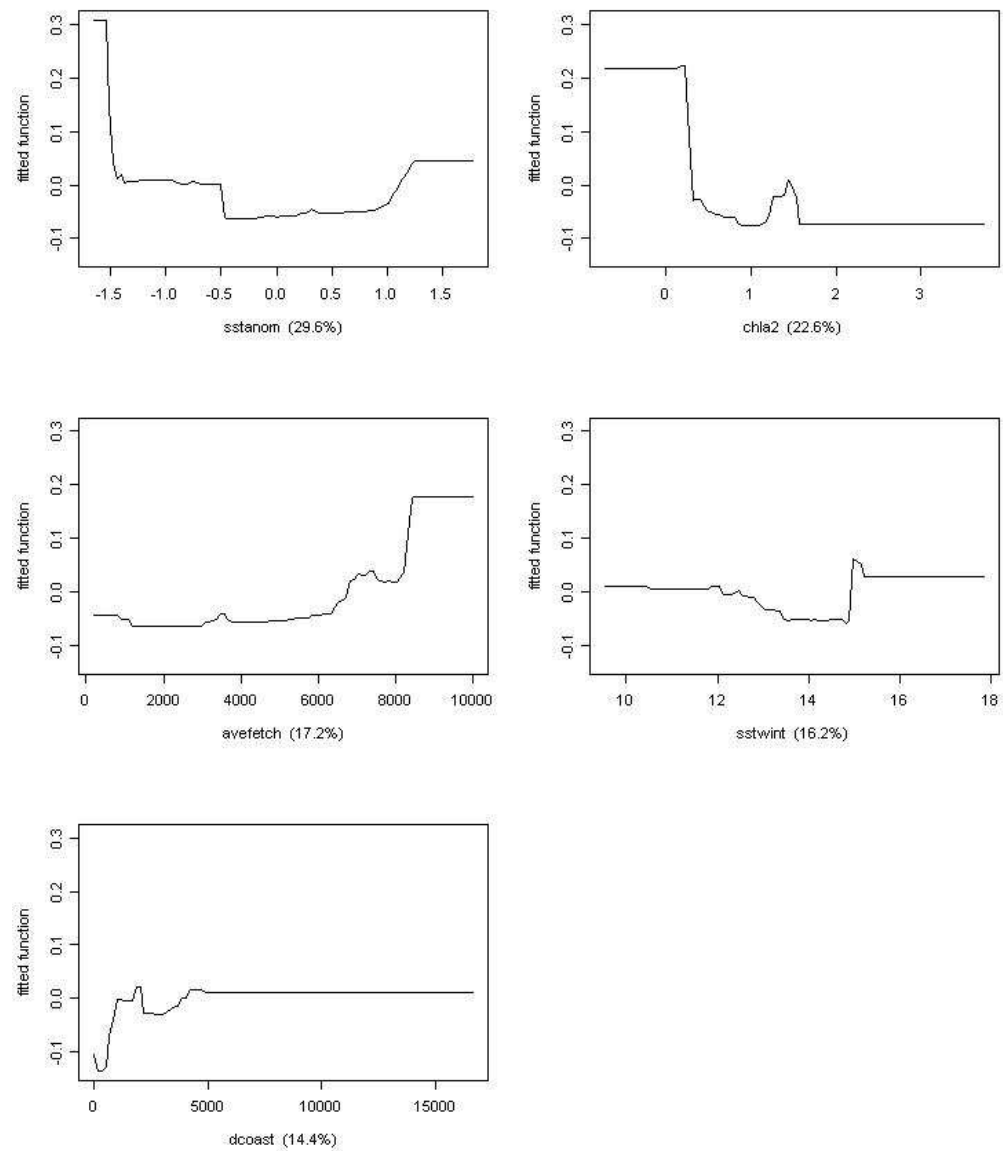
Pempheris adspersa : Bigeye



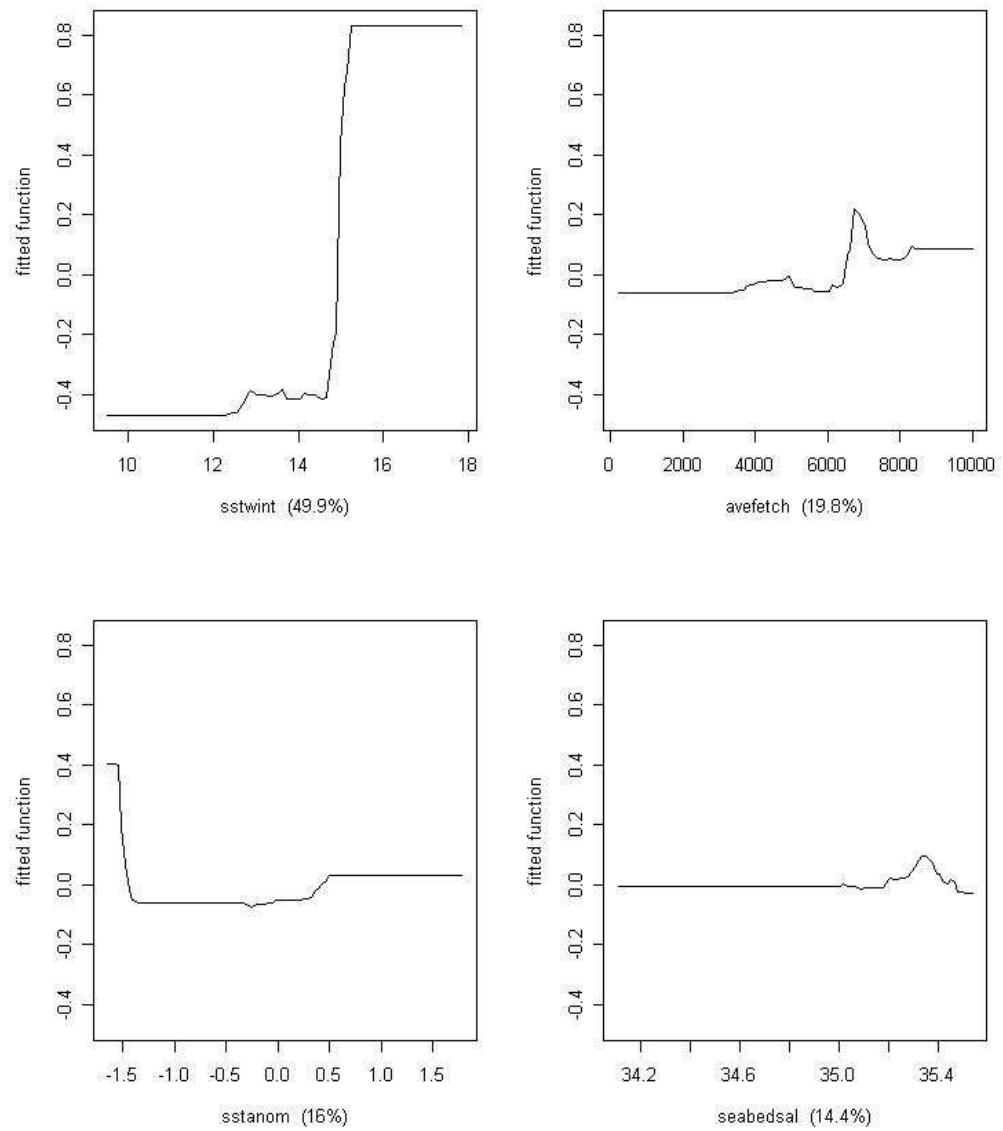
Plagiotremus tapeinosoma : Mimic blenny



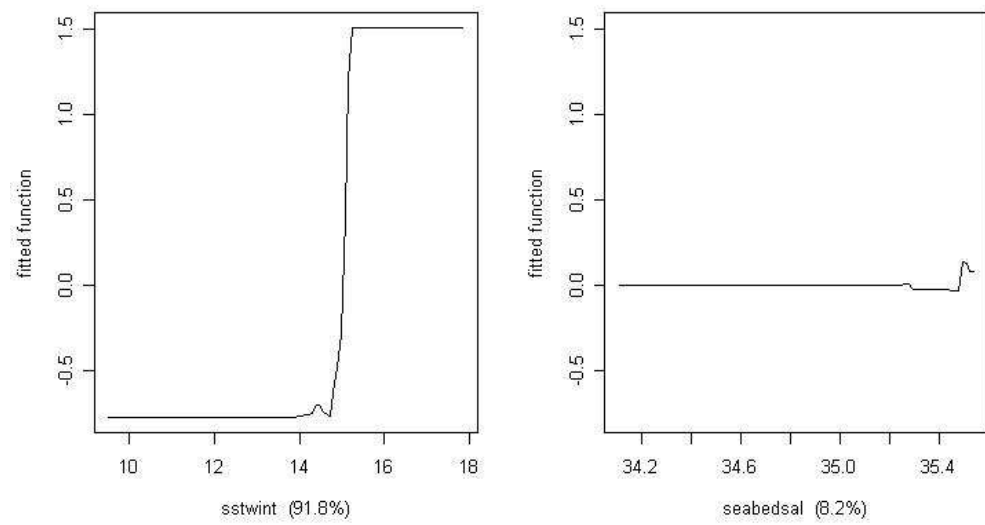
Pseudophycis barbata : Southern bastard cod



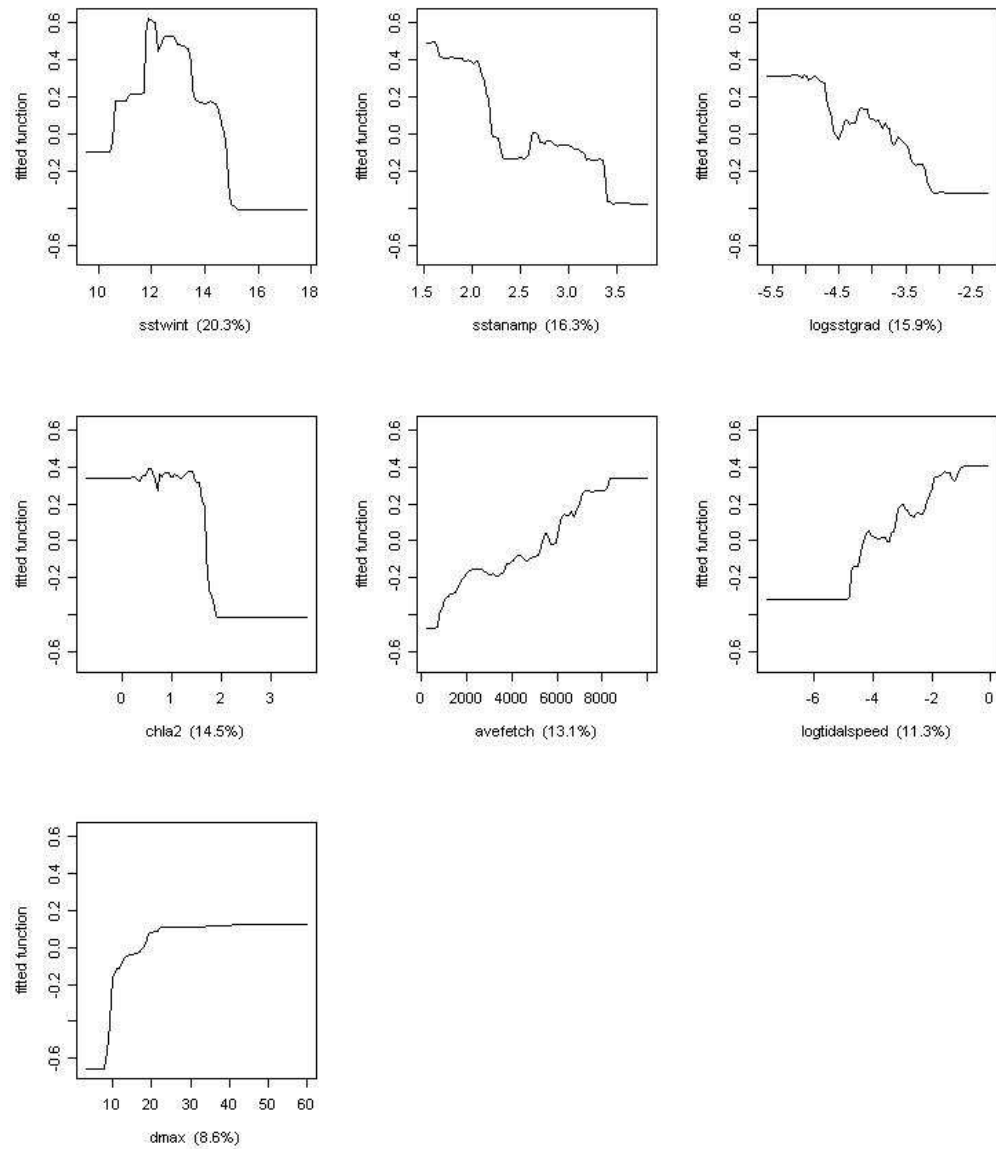
Pseudocaranx dentex : Trevally



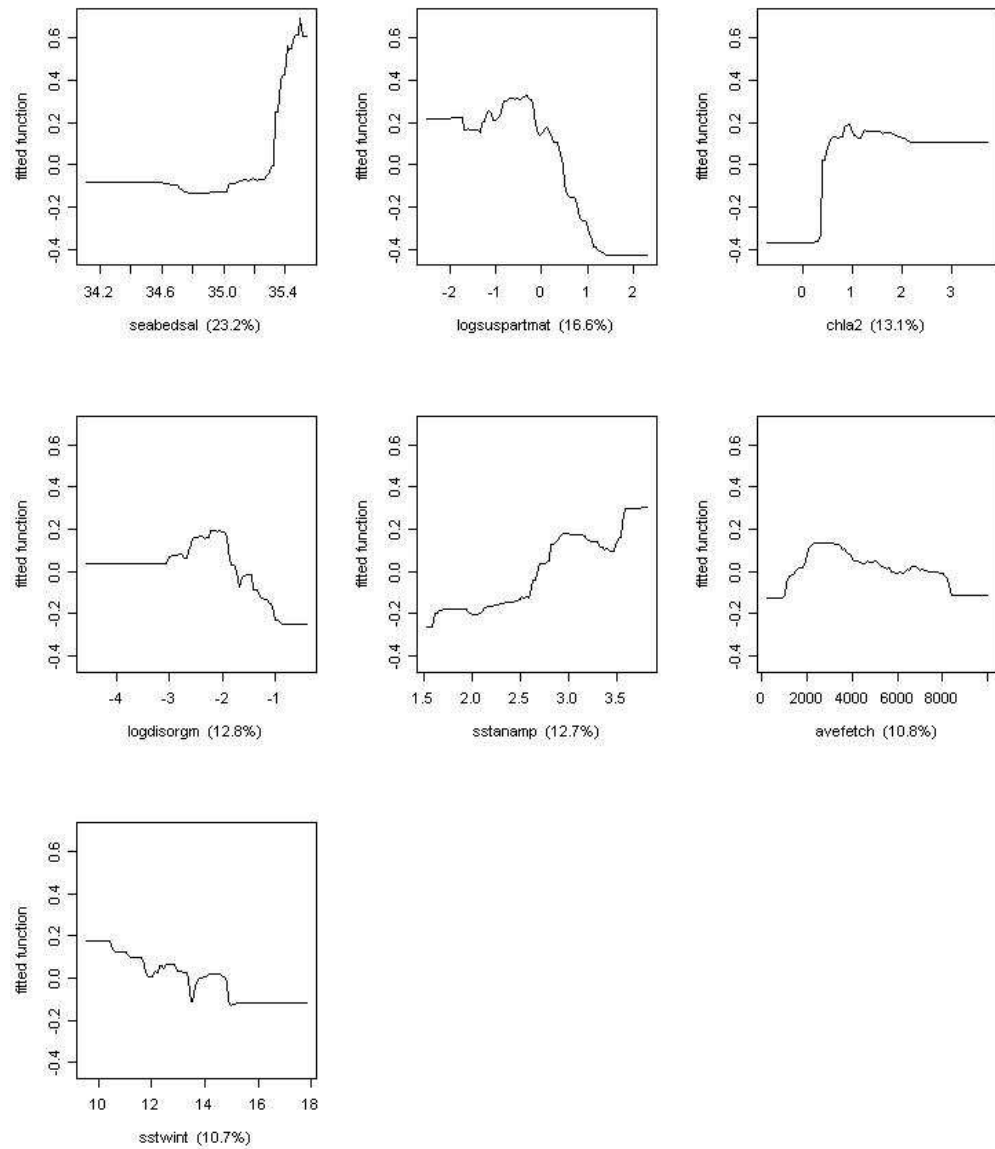
Pseudolabrus luculentus : Orange wrasse



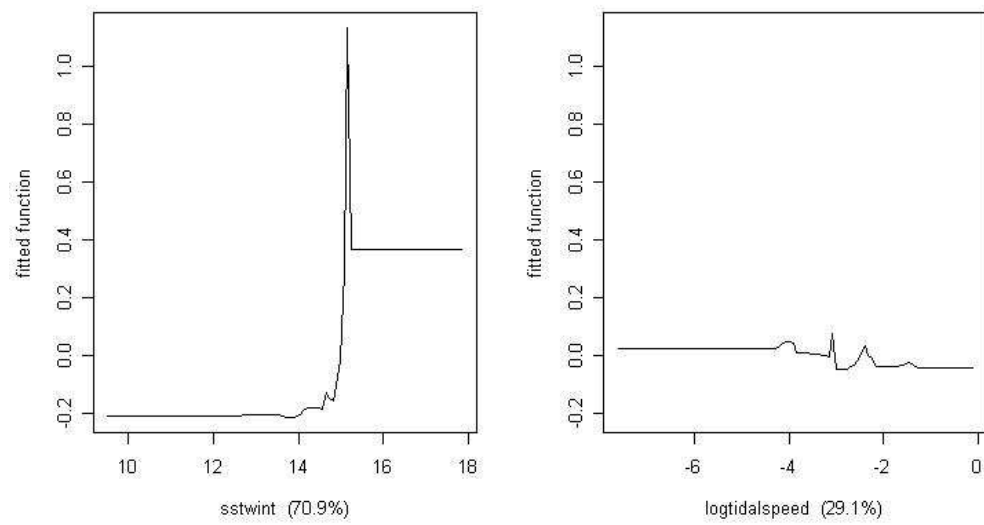
Pseudolabrus miles : Scarlet wrasse



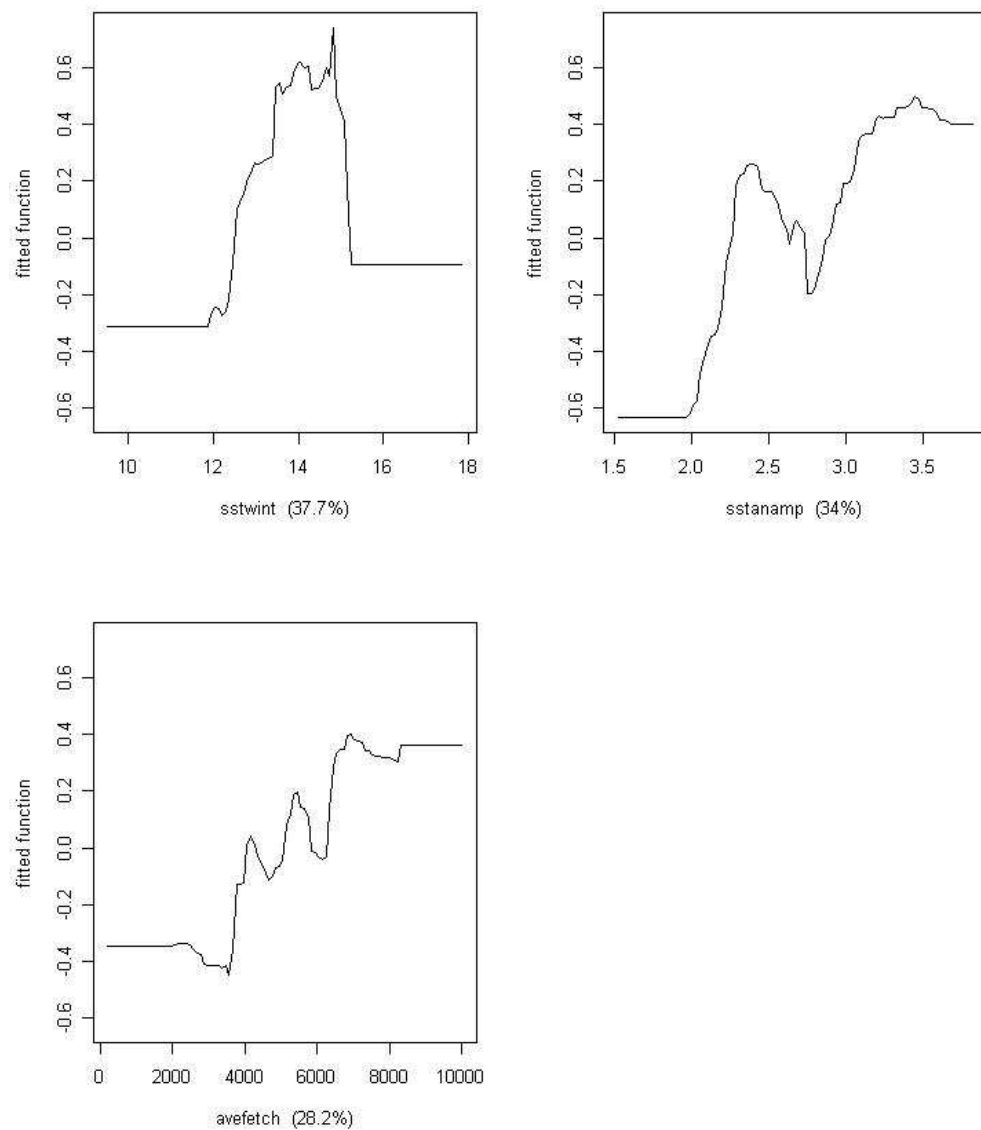
Ruanoho whero : Spectacled triplefin



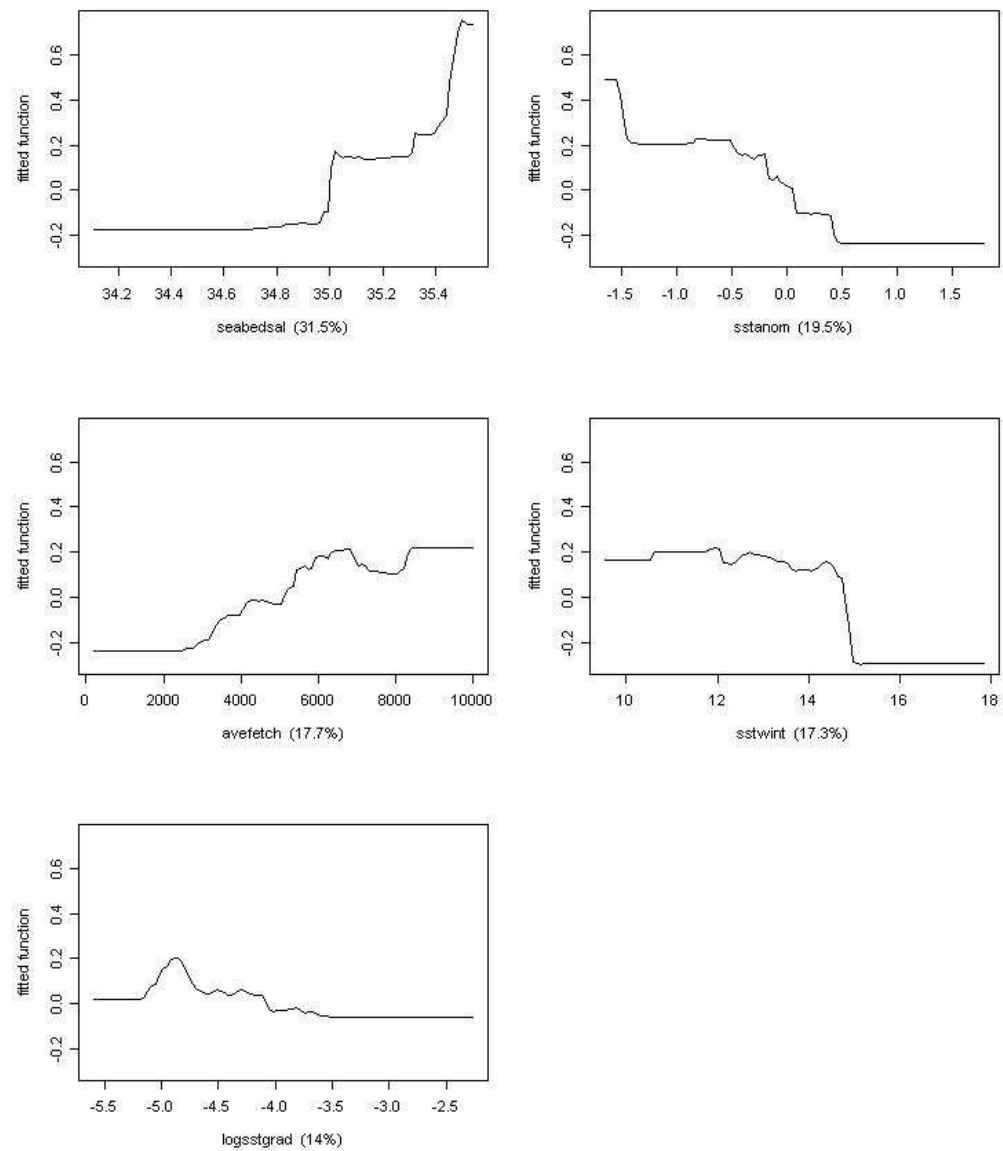
Scorpaena cardinalis : Northern scorpionfish



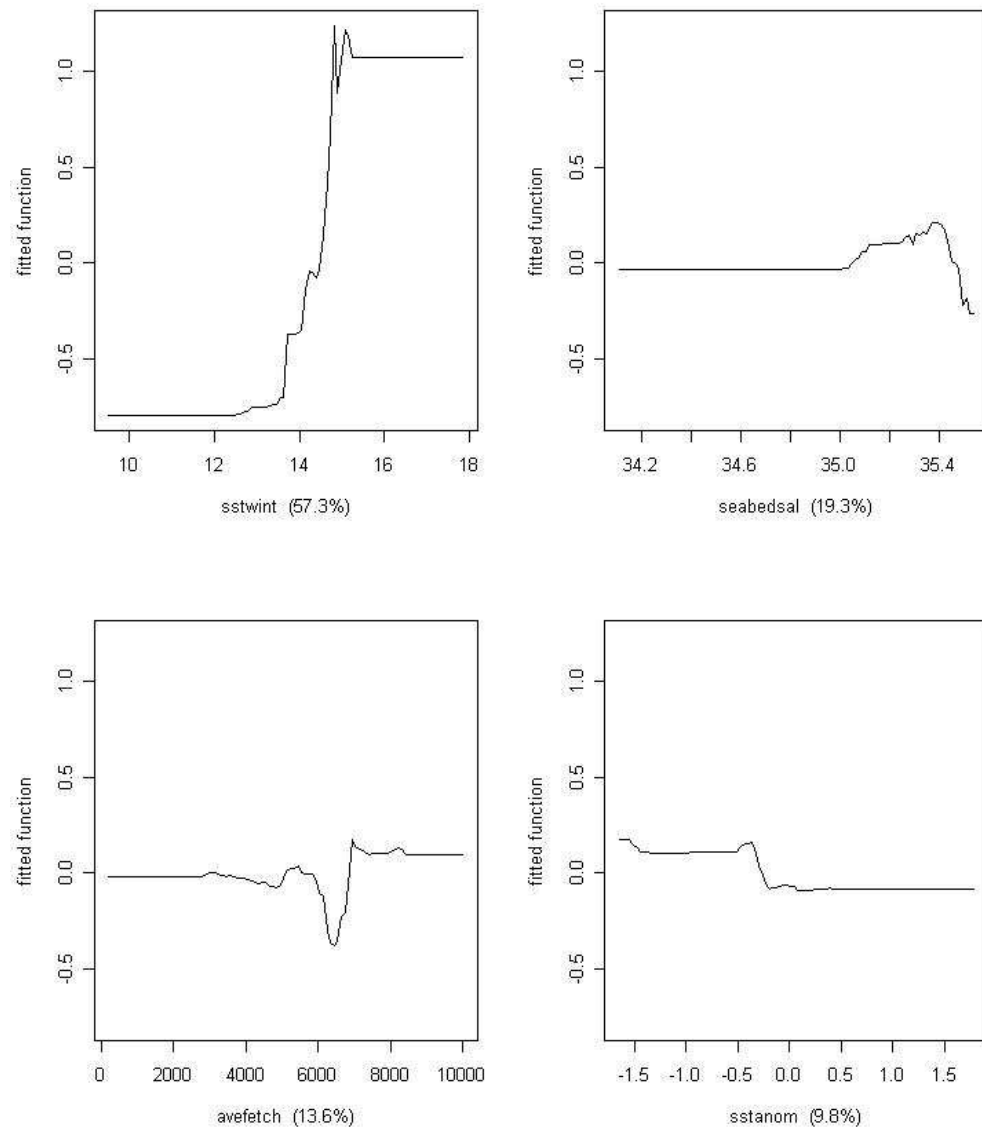
Scorpius lineolatus : Sweep



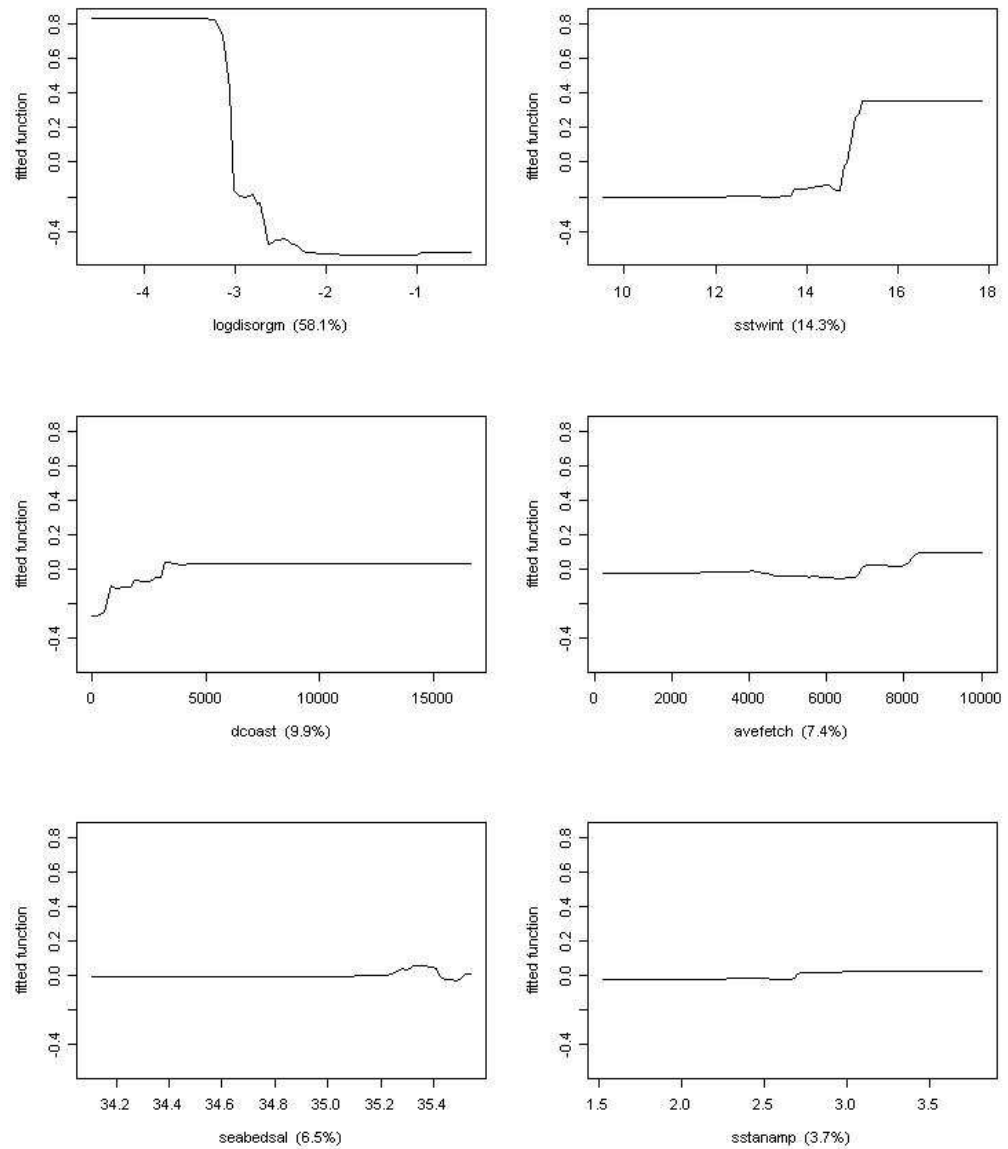
Scorpaena papillosus : Dwarf scorpionfish



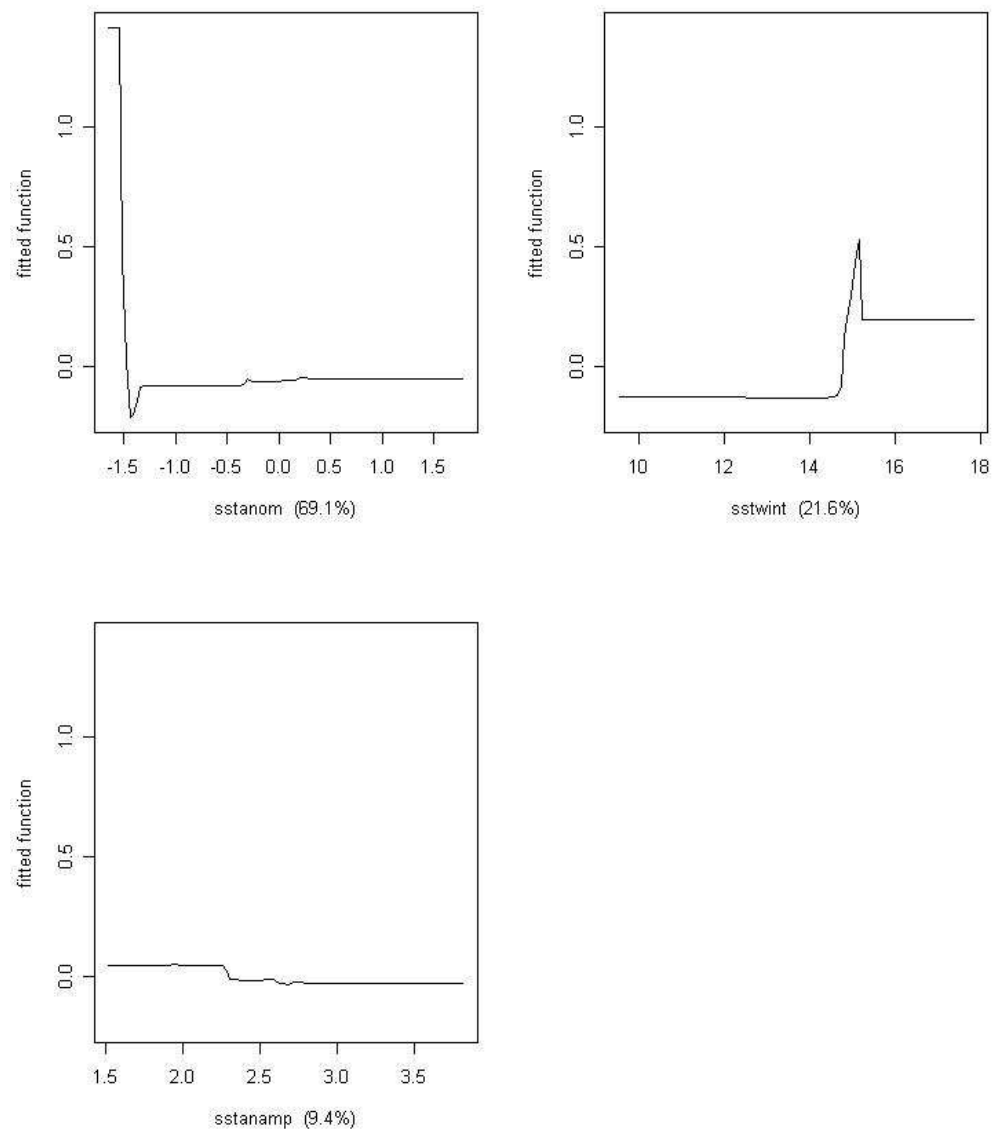
Scorpius violaceus : Blue maomao



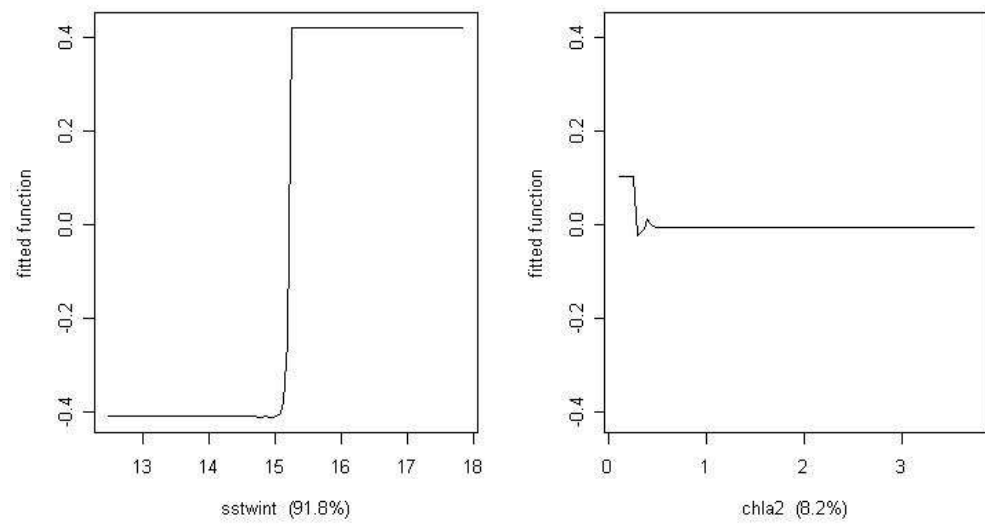
Seriola lalandi : Kingfish



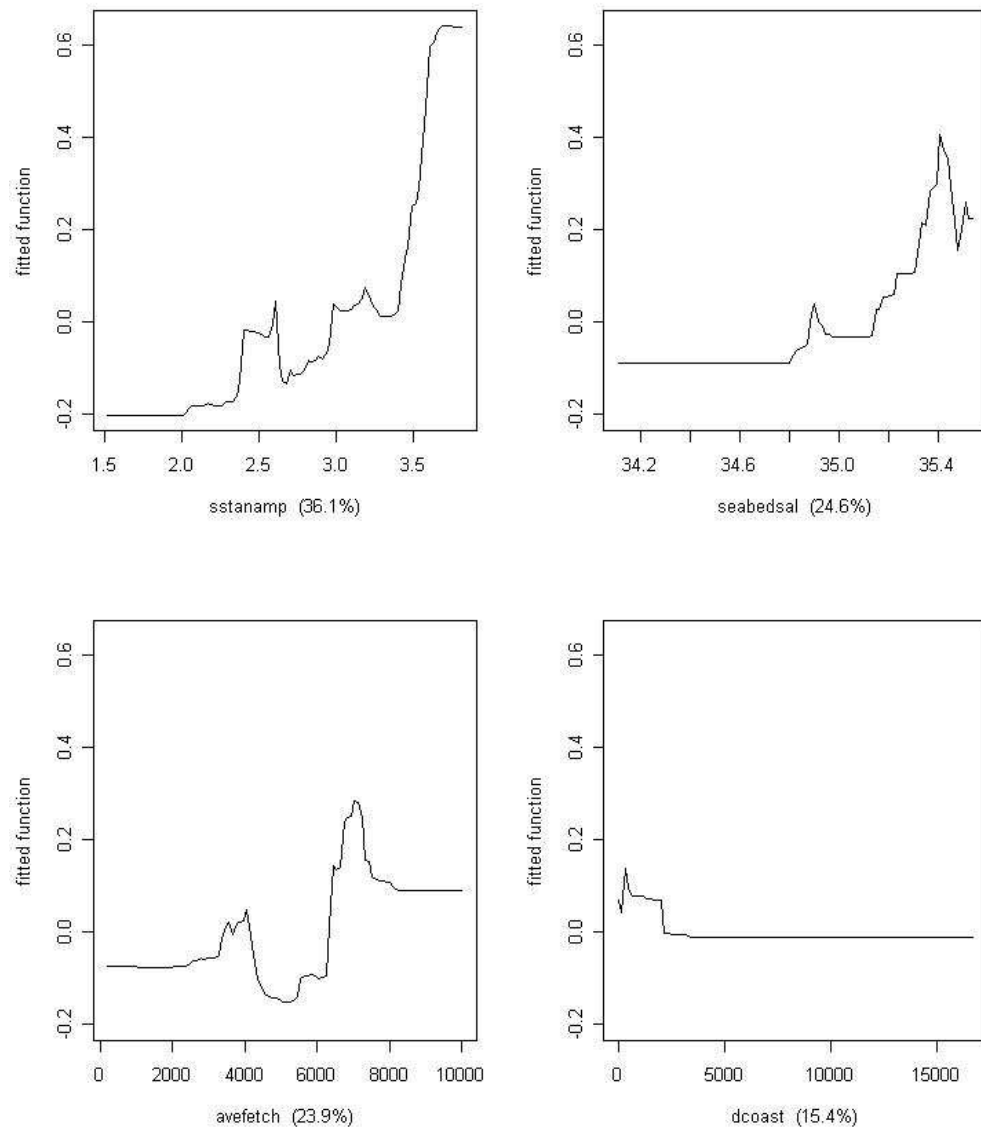
Suezichthys aylingi : Crimson cleanerfish



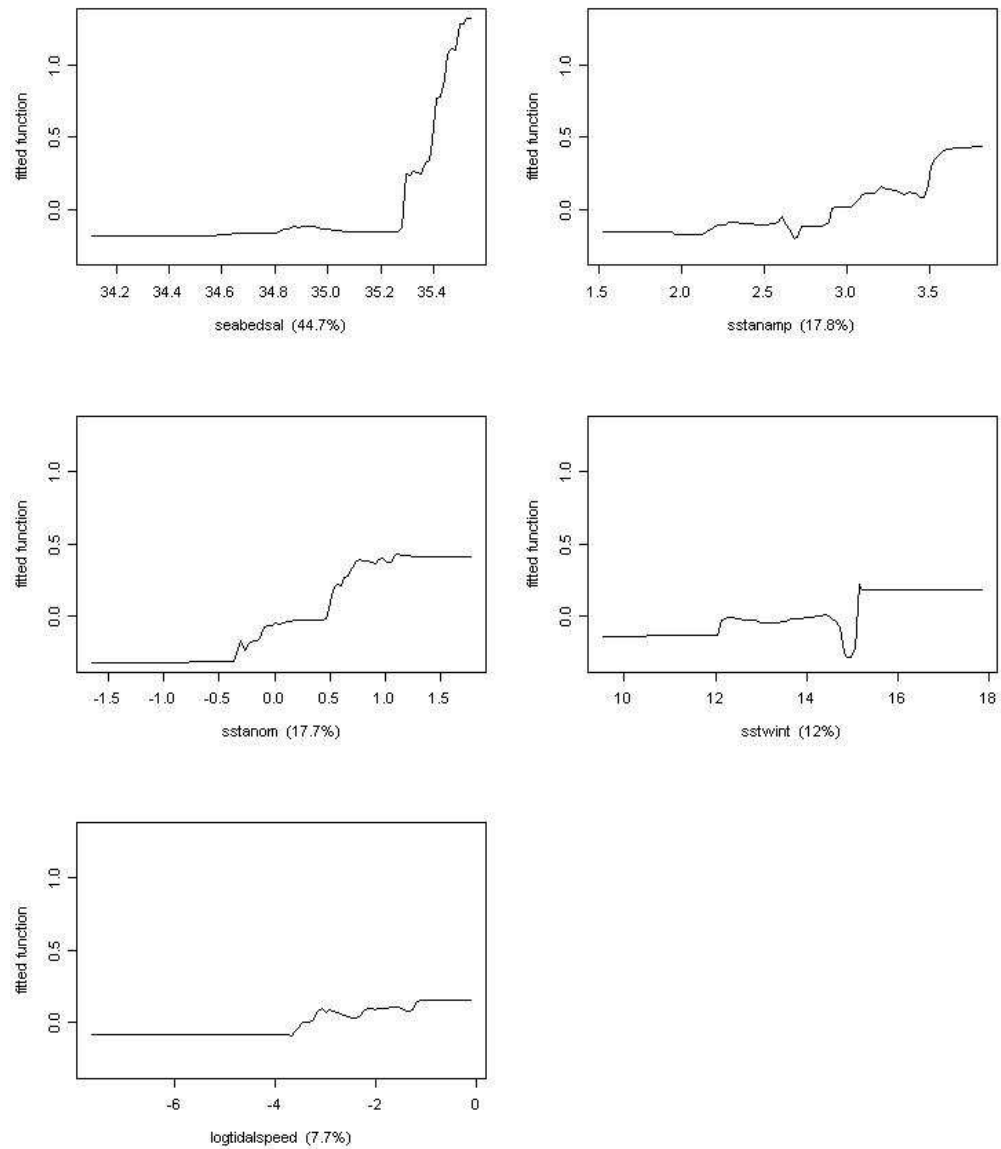
Trachypoma macracanthus : Toadstool grouper



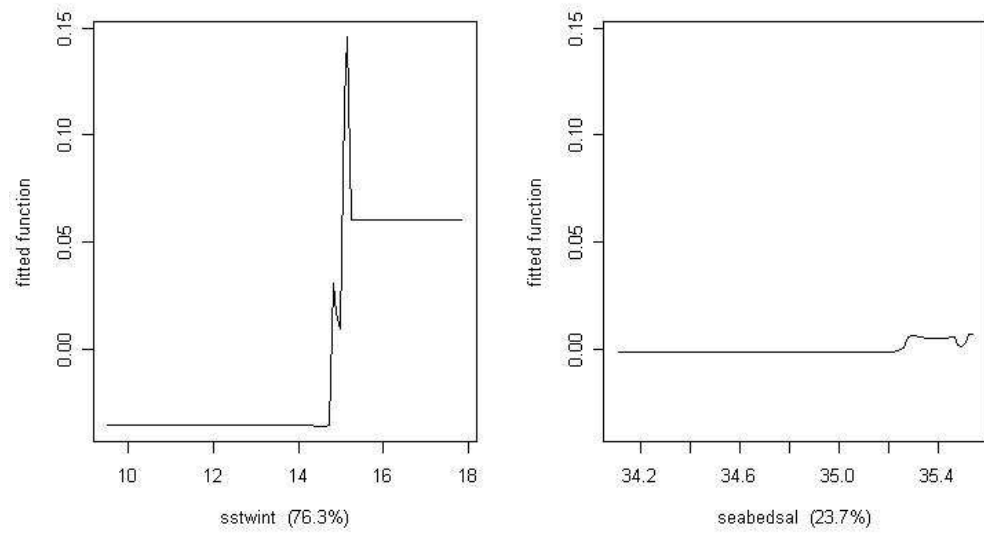
Trachurus novaezelandiae : Jack mackerel



Upeneichthys lineatus : Goatfish



Zanclistius elevatus : Long-finned boarfish



Zeus faber : John dory

