

Improved knowledge of Torres Strait seabed biota and reef habitats

Marine and Atmospheric Research

M.D.E. Haywood, M. Browne, T. D. Skewes, W.A. Rochester, I.M. McLeod,
C.R. Pitcher, D.M. Dennis, J.R. Dunn, S.J. Cheers, T.J. Wassenberg
Marine and Tropical Sciences Research Facility
Project 1.3.1



Australian Government

Department of the Environment,
Water, Heritage and the Arts

© James Cook University and The University of Queensland.

National Library of Australia Cataloguing-in-Publication entry:

Author: Haywood, M.D.E (Michael Donald Edmund) 1958-

Title: Improved knowledge of Torres Strait seabed biota and reef habitats / authors, M.D.E. Haywood [et al.].

Edition: 1st ed.

Publisher: Cleveland, Qld.: CSIRO Marine and Atmospheric Research, 2007.

ISBN: 9781921424014 (pbk.)
9781921424038 (CD-ROM)
9781921424021 (pdf)

Notes: Bibliography.

Subjects: Marine ecology--Torres Strait.
Marine biodiversity--Torres Strait.
Reefs--Torres Strait--Environmental aspects.
Torres Strait--Environmental aspects.

Other Authors: CSIRO. Marine and Atmospheric Research.

Dewey Number: 577.7476

This report should be cited as:

Haywood, M. D. E., Browne, M., Skewes, T., Rochester, W., McLeod, I., Pitcher, R., Dennis, D., Dunn, J., Cheers, S., Wassenberg, T. (2007) *Improved knowledge of Torres Strait seabed biota and reef habitats*. Report to the Marine and Tropical Sciences Research Facility. Reef and Rainforest Research Centre Limited, Cairns (115 pp.).

Made available for download by the Reef and Rainforest Research Centre on behalf of the Australian Government's Marine and Tropical Sciences Research Facility.

The Marine and Tropical Sciences Research Facility (MTSRF) is part of the Australian Government's Commonwealth Environment Research Facilities programme. The MTSRF is represented in North Queensland by the Reef and Rainforest Research Centre Limited (RRRC). The aim of the MTSRF is to ensure the health of North Queensland's public environmental assets – particularly the Great Barrier Reef and its catchments, tropical rainforests including the Wet Tropics World Heritage Area, and the Torres Strait – through the generation and transfer of world class research and knowledge sharing.

This publication is copyright. The Copyright Act 1968 permits fair dealing for study, research, information or educational purposes subject to inclusion of a sufficient acknowledgement of the source.

The views and opinions expressed in this publication are those of the authors and do not necessarily reflect those of the Australian Government or the Minister for the Environment, Water, Heritage and The Arts.

While reasonable effort has been made to ensure that the contents of this publication are factually correct, the Commonwealth does not accept responsibility for the accuracy or completeness of the contents, and shall not be liable for any loss or damage that may be occasioned directly or indirectly through the use of, or reliance on, the contents of this publication.

This report is available for download from the
Reef and Rainforest Research Centre Limited website:
<http://www.rrrc.org.au/publications/>



February 2007

Contents

1.	SUMMARY	1
2.	BACKGROUND	3
3.	INTER-REEFAL SEABED MAPPING	4
3.1.	Introduction	4
3.2.	Methodology.....	5
3.2.1.	<i>Inter-reefal seabed mapping</i>	<i>5</i>
	<i>Sampling design</i>	<i>5</i>
3.2.2.	<i>Field sampling</i>	<i>5</i>
	<i>Vessel description</i>	<i>5</i>
	<i>Towed video camera</i>	<i>6</i>
	<i>Tasks, events and navigation data logging</i>	<i>7</i>
3.2.3.	<i>Video data processing</i>	<i>9</i>
3.2.4.	<i>Data analysis.....</i>	<i>11</i>
	<i>Physical co-variate data</i>	<i>11</i>
3.2.5.	<i>Accounting for aggregation / dispersion.....</i>	<i>12</i>
3.2.6.	<i>Random forests</i>	<i>13</i>
3.2.7.	<i>Transformation in non-parametric estimation.....</i>	<i>14</i>
3.3.	Results and Discussion	16
3.3.1.	<i>Estimation of spatial distribution of counts</i>	<i>16</i>
3.3.2.	<i>Estimation of spatial distribution of cover types</i>	<i>26</i>
3.3.3.	<i>Comparison of video biota with that from sled and trawl</i>	<i>30</i>
4.	CORAL REEF CHARACTERISATION	38
4.1.	Introduction	38
4.2.	Methodology.....	39
4.2.1.	<i>Basemap</i>	<i>39</i>
4.2.2.	<i>Reef datasets</i>	<i>40</i>
4.2.3.	<i>Survey methods.....</i>	<i>41</i>
4.2.4.	<i>Sample design.....</i>	<i>41</i>
4.2.5.	<i>Field sampling</i>	<i>41</i>
4.2.6.	<i>Cluster analysis</i>	<i>44</i>
	<i>Analysis datasets.....</i>	<i>44</i>
	<i>Species and habitat clustering.....</i>	<i>44</i>
	<i>Habitat-forming species and environmental factors</i>	<i>45</i>
	<i>Fishes and environmental factors</i>	<i>47</i>
4.2.7.	<i>Gap analysis.....</i>	<i>47</i>
4.3.	Results	48
4.3.1.	<i>Physical covariate datasets</i>	<i>48</i>
4.3.2.	<i>Maps of survey data</i>	<i>56</i>
	<i>Substrate maps</i>	<i>56</i>
	<i>Coral growth forms</i>	<i>64</i>
	<i>Reef fish distribution</i>	<i>73</i>
	<i>Holothurian distribution.....</i>	<i>76</i>
	<i>Seagrass distribution</i>	<i>78</i>
4.3.3.	<i>Clustering the maximum number of samples (a).....</i>	<i>79</i>

4.3.4.	<i>Clustering by the maximum number of species (b)</i>	81
4.3.5.	<i>Clustering by habitat-forming species (c)</i>	87
4.3.6.	<i>Clustering by habitat attributes (d)</i>	89
4.3.7.	<i>Association between clusters from different analyses</i>	90
4.3.8.	<i>Habitat-forming species and fetch</i>	91
4.3.9.	<i>Habitat-forming species and environmental factors</i>	92
4.3.10.	<i>Fishes and environmental factors</i>	94
4.3.11.	<i>Gap analysis</i>	96
4.4.	Discussion	97
4.4.1.	<i>Clustering</i>	97
4.4.2.	<i>Threats</i>	99
4.4.3.	<i>Gap analysis</i>	99
4.4.4.	<i>Identification of areas for future work</i>	100
5.	REFERENCES	101

List of Figures

Figure 1. The 19.6 m James Cook University research vessel RV James Kirby.....	6
Figure 2. The Drop-Cam system recovered after completion of a 500 m video transect and the surface real-time monitoring, control and data acquisition system.	7
Figure 3. Operator analysing seabed video.	9
Figure 4. Screen dump of the front end of the custom Delphis software module used to view and analyse the towed underwater video.	10
Figure 5. Estimated spatial distribution of Anthozoa <i>Turbinaria</i> sp.	17
Figure 6. Estimated spatial distribution of Actinopterygii 'bonyfish' sp.	18
Figure 7. Partial effect curves for Anthozoa <i>Turbinaria</i> sp. See Table 3 for the definitions of the covariate abbreviations.	19
Figure 8. Partial effect curves for Actinopterygii 'bonyfish' sp. See Table 3 for the definitions of the covariate abbreviations.	20
Figure 9. Estimated spatial distribution (count) of Holothuroidea <i>Holothuria</i> sp.	22
Figure 10. Estimated spatial distribution (count) of Anthozoa <i>Junceella fragilis</i>	23
Figure 11. Estimated spatial distribution (count) of Anthozoa <i>Solenocaulon</i> sp. (bare).	24
Figure 12. Estimated spatial distribution (count) of Anthozoa branching sp.	25
Figure 13. Estimated spatial distribution (percent cover) of bioturbation sediment.	26
Figure 14. Estimated spatial distribution (percent cover) of "sponge (globular)".	27
Figure 15. Estimated spatial distribution (percent cover) of "seagrass strap-like".	28
Figure 16. Estimated spatial distribution (percent cover) of "sponge digitate".	29
Figure 17. Torres Strait: Map showing sites sampled with various gears during the CRC Torres Strait project "Mapping and physical characterisation of key biotic and physical attributes of the Torres Strait Ecosystem" (Pitcher et al. 2007).	30
Figure 18. Predicted distributions of the anthozoan, <i>Sphenopus marsupalis</i> generated from (a) catches obtained using sled and trawl (Pitcher et al. 2007); and (b) towed underwater video count data.	36
Figure 19. Predicted distributions of the sponge, <i>Ianthella flabiformis</i> generated from (a) catches obtained using sled and trawl (Pitcher et al. 2007) and (b) towed underwater video count data.	37
Figure 20. Map of the Torres Strait showing reefs digitised from AUSLIG 1:100, 000 and 1:250, 000 topographic map series with Landsat Thematic Mapper imagery.	39
Figure 21. Map of Torres Strait showing reef survey sites by cruise.	42
Figure 22. Schematic diagram of the database structure for the biological survey data.	43
Figure 23. Map showing the positions of reef survey sites.	44
Figure 24. Depth (m) collated from digital soundings, satellite imagery and LADS data.	48
Figure 25. Modelled seabed current stress.	49
Figure 26. Weighted average surface nitrogen, oxygen, phosphate, silicate, temperature and salinity from the CARS dataset.	50
Figure 27. Mean annual chlorophyll a concentrations derived from SeaWiFS ocean colour.	52
Figure 28. Standard deviation of the annual Chlorophyll a concentrations derived from SeaWiFS ocean colour.	52
Figure 29. Mean annual K490 diffuse attenuation coefficient at wavelength 490nm derived from SeaWiFS ocean colour.	53
Figure 30. Standard deviation of the mean annual K490 diffuse attenuation coefficient at wavelength 490nm derived from SeaWiFS ocean colour.	53
Figure 31. Modelled mean wave height generated by the Wave Model (WAM) data supplied by the Bureau of Meteorology.	54
Figure 32. Modelled mean wave period generated by the Wave Model (WAM) data supplied by the Bureau of Meteorology.	55

Figure 33. Modelled mean wave direction generated by the Wave Model (WAM) data supplied by the Bureau of Meteorology.	55
Figure 34. Mean percentage cover of soft sediments (mud and sand) on Torres Strait Reef tops and edges.	56
Figure 35. Mean percentage cover of rubble on Torres Strait Reef tops and edges.	57
Figure 36. Mean percentage cover of boulders on Torres Strait Reef tops and edges.	58
Figure 37. Mean percentage cover of consolidated rubble on Torres Strait Reef tops and edges.	59
Figure 38. Mean percentage cover of pavement on Torres Strait Reef tops and edges.	60
Figure 39. Mean percentage cover of live coral on Torres Strait Reef tops and edges.	61
Figure 40. Mean percentage cover of dead standing hard coral on Torres Strait Reef tops and edges.	62
Figure 41. Percent cover of branching corals on Torres Strait reef tops and edges.	64
Figure 42. Percent cover of digitate corals on Torres Strait reef tops and edges.	65
Figure 43. Percent cover of encrusting corals on Torres Strait reef tops and edges.	66
Figure 44. Percent cover of foliose corals on Torres Strait reef tops and edges.	67
Figure 45. Percent cover of fungiid corals on Torres Strait reef tops and edges.	68
Figure 46. Percent cover of massive corals on Torres Strait reef tops and edges.	69
Figure 47. Percent cover of sub-massive corals on Torres Strait reef tops and edges.	70
Figure 48. Percent cover of tabulate corals on Torres Strait reef tops and edges.	71
Figure 49. Percent cover of columnar corals on Torres Strait reef edges (none were recorded on the reef tops).	72
Figure 50. Percent cover of corymbose corals on Torres Strait reef edges (none were recorded on the reef tops).	72
Figure 51. Species richness of reef fish on all 231 reef edge sites surveyed during 1995-96.	73
Figure 52. Plot of the tree model of reef species richness in terms of the physical covariates. For a full list of the physical covariates included in the analysis see Section 4.3.1.	74
Figure 53. Map showing the distribution and abundance of the 11 fish families surveyed on Torres Strait reefs during 1995-96. The pie charts show the mean numbers of fish per ha of each family recorded on the reefs.	75
Figure 54. Distribution and abundance of genera of holothurians throughout the Torres Strait.	77
Figure 55. Distribution and abundance of seagrasses throughout Torres Strait. Seagrass cover is expressed as log (percentage cover).	78
Figure 56. Sample clusters from clustering the maximum number of samples.	79
Figure 57. Sample clusters from clustering by the maximum number of species.	81
Figure 58. Sample clusters from clustering by habitat-forming species.	87
Figure 59. Sample clusters from clustering by habitat attributes.	89
Figure 60. Wind roses for Horn Island showing windspeed and direction for 15:00 h for (a) January and (b) July. Copyright © Bureau of Meteorology, Commonwealth of Australia 2007.	91
Figure 61. Average (a) south-east and (b) north-west fetch for sample clusters from clustering by habitat-forming species. The error bars indicate standard errors. Means and standard errors were calculated with log-transformed data.	92
Figure 62. CCA plot for habitat-forming species and all candidate environment variables except latitude, longitude, temperature seasonality and silicate. The grey symbols indicate sample WA scores. The coloured symbols indicate species scores (red, corals; green, algae; blue, seagrasses). The arrows indicate biplot scores, which indicate the correlations between the environment variables and the sample LC scores.	93
Figure 63. Sample (a) LC and (b) WA scores of the first axis of a CCA of habitat-forming species and all candidate environment variables except latitude, longitude, temperature seasonality and silicate. The sample LC scores are the linear combinations of the environment variables. The WA scores are the weighted averages of the species scores. The correspondence between map (b) and map (a) indicates how well variation in species composition along the axis (b) represents variation in the environment variables along the axis (a).	94

Figure 64. CCA plot for fishes and all candidate environment variables except latitude, longitude, temperature seasonality and silicate. The grey symbols indicate sample WA scores (the weighted averages of the species scores). The coloured symbols indicate species scores (coloured by family). The arrows indicate biplot scores, which indicate the correlations between the environment variables and the LC sample scores (the linear combinations of the environment variables) for each axis.	95
Figure 65. Sample (a) LC and (b) WA scores of the first axis of a CCA of fishes and all candidate environment variables except latitude, longitude, temperature seasonality and silicate. The sample LC scores are the linear combinations of the environment variables. The WA scores are the weighted averages of the species scores. The correspondence between map (b) and map (a) indicates how well the species composition data (b) can represent the variation in the environmental data (a).	95
Figure 66. Map of the Torres Strait showing cluster membership of the reefs. K-means clustering was performed on the first 4 dimensions of an nMDS based on a selection of biologically important physical covariates.	96

List of Tables

Table 1. The number of sites sampled with each sampling device during the CRC Torres Strait project “Mapping and physical characterisation of key biotic and physical attributes of the Torres Strait Ecosystem”.	4
Table 2. Substratum and Biological habitat types and animal events types entered in real time to annotated the video transect.	8
Table 3. Description of physical co-variables used in modelling the “species” distributions of video-identified biota.	11
Table 4. Numbers of individuals in each broad taxonomic category sampled throughout Torres Strait with the trawl or video in the (Pitcher et al. 2007) study compared to those observed on the towed video.	31
Table 5. Comparison of all the OTUs identified in sufficient numbers to enable distribution modelling from identified from the video with their possible matches identified from the epibenthic sled and trawl from the (Pitcher et al. 2007).	32
Table 6. CMAR surveys on the reefs of Torres Strait since 1995.	40
Table 7. Environment variables available and selected for canonical correspondance analysis. Details of the data sources are described above.	46
Table 8. Species of holothurians identified throughout the Torres Strait during diver surveys conducted from 1995 to 2005.	76
Table 9. Association of species groups with sample clusters from clustering the maximum number of samples. For each cluster and species group the table shows the frequency of the species group and the number of samples in which the species group was observed as a percentage of the number of samples from all clusters in which the species group was observed. Zeros in the table represent small non-zero values rounded to zero. True zeros are represented by blanks.	80
Table 10. Association of species groups with sample clusters from clustering by the maximum number of species. For each cluster and species group the table shows the frequency of the species group and the number of samples in which the species group was observed as a percentage of the number of samples from all clusters in which the species group was observed. Zeros in the table represent small non-zero values rounded to zero. True zeros are represented by blanks.	82
Table 11. Association of species groups with sample clusters from clustering by habitat-forming species. For each cluster and species group the table shows (a) the frequency of the species group, (b) the average percentage cover of the species group, (c) the number of samples in which the species group was observed as a percentage of the number of samples from all clusters in which the species group was observed, and (d) the total cover of the species group as a percentage of the total cover of the species group in samples all clusters (calculated as if	

sample areas were all equal). Zeros in the table represent small non-zero values rounded to zero. True zeros are represented by blanks.....	88
Table 12. Association of habitat elements with sample clusters from clustering by habitat attributes. For each cluster and habitat element the table shows (a) the frequency of the habitat element, (b) the average percentage cover of the habitat element, (c) the number of samples in which the habitat element was observed as a percentage of the number of samples from all clusters in which the habitat element was observed, and (d) the total cover of the habitat element as a percentage of the total cover of the habitat element in samples all clusters (calculated as if sample areas were all equal). Zeros in the table represent small non-zero values rounded to zero. True zeros are represented by blanks.....	90
Table 13. Associations among clusters from different analyses. The table shows the percentage of samples in each cluster from analyses (a) (maximum number of samples) and (d) (habitat attributes) that were in each cluster from analysis (b) (maximum number of species).....	90
Table 14. Area and number of sites surveyed (all surveys combined) within each of the clusters produced by a K-means clustering on the first 4 dimensions of an nMDS based on a selection of biologically important physical covariates. RRI = Reef Resource Inventory.....	97

Acronyms Used In This Report

AFMA	Australian Fisheries Management Authority
CART	Classification and Regression Trees
CCA	Canonical Correlation Analysis
CDF	Cumulative Distribution Function
CMAR	Division of Marine and Atmospheric Research (CSIRO)
CRC	Cooperative Research Centre
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CTD	Conductivity Temperature Depth
ESE	East-South-East
ETS	Eastern Torres Strait
GBR	Great Barrier Reef
GPS	Geographic Positioning System
INB	In-bag
nMDS	non-metric Multi Dimensional Scaling
OOB	Out-of-bag
OTUs	Operational Taxonomic Units
PAM	Partitioning Around Medoids
PAR	Photosynthetically Active Radiation
PNG	Papua New Guinea
RRI	Reef Resource Inventory
SD	Standard Deviation
TS	Torres Strait
TSPZ	Torres Strait Protected Zone
VMS	Vessel Monitoring System
WNW	West-north-west

Acknowledgements

Federic Saint-Cast generously provided wave height, period and direction data generated from his model. Thanks to the RAN for providing the LADS bathymetry data and to Geoscience Australia for the sediment data. Many thanks to Lea Crosswell for designing the front cover.

1. Summary

The CRC Torres Strait project “Mapping and physical characterisation of key biotic and physical attributes of the Torres Strait Ecosystem” (Pitcher et al. 2007) mapped inter-reefal habitats and biota throughout the Torres Strait. Data was collected using trawls, benthic sleds and a towed underwater video camera. Biota recovered from the sled and trawl catches were combined, modeled and used to provide comprehensive distribution and abundance information on the benthic species and assemblages found at the sites sampled with these devices. Video data was sub-sampled and analysed primarily for habitat characterisation rather than the identification, measurement and enumeration of biota. One of the aims of this project was to reanalyze the video data, without subsampling and use the resulting data to map the distribution of the biota that could be identified from the images.

Most taxa were not able to be identified as comprehensively from the video as those collected with the extractive devices. The catches from the epibenthic sled and trawl were merged to give a total of diverse dataset comprised of 2,372 Operational Taxonomic Units (OTUs) consisting of sponges, molluscs, crustaceans, coral, bryozoans, echinoderms, fishes and algae (Pitcher et al. 2007). Of these 258 OTUs were present in sufficient samples to enable modeling of their spatial distribution and abundance. In the present study, our analysis of the towed video collected during that field study we were able to distinguish only 62 distinct OTUs for the count analysis and 30 OTUs for the percent cover analysis that were observed often enough to enable distribution modeling. The video is more efficient at capturing information on some of the larger, sessile benthic taxa such as corals, gorgonians and sponges. Smaller, mobile and cryptic biota such as crustaceans, bivalves and gastropods are not well documented by the video.

The video has limitations when used under poor visibility (as is common in Torres Strait) and in recording mobile, cryptic or small biota. Nevertheless, in areas that are too rough to sample with sled or trawl or for some taxa that are not well sampled by sled or trawl (e.g. the gorgonian whips *Junceella fragilis* and *J. juncea*) the video complements data collected with the other devices. Ideally all three devices should be used together in order to gain the most comprehensive picture of seabed biodiversity.

Torres Strait is a unique environment being a shallow, submerged land bridge between PNG and the northern tip of Cape York. With the Coral Sea to the east and the Arafura Sea to the west, strong physical drivers operate throughout the region and are probably responsible for shaping the character of the >1,200 coral reef communities of the region.

During the past twelve years CMAR has conducted a number of surveys on the reefs of Torres Strait. The primary objective of these surveys was generally to facilitate the stock assessment of various fishery target species such as beche-de-mer or finfish, however ancillary habitat data was also collected albeit at a coarser level of taxonomic resolution. The primary objective of this part of the study was to characterise the reefs of Torres Strait in terms of the biological and habitat data collected during these surveys and to relate the nature of the reefs to the available physical data.

It was necessary to conduct a range of analyses on subsets of the complete dataset because many combinations of species-site were absent due to the variable nature of the objectives of the original surveys. Four cluster analyses were performed: (a) clustering of samples by species with a dataset that included as many samples as possible, (b) clustering of samples from reef edges by as many species as possible, (c) clustering of samples from reef edges by habitat-forming species, and (d) clustering of samples by habitat attributes (substrate and habitat-forming species). In general most clusters tended to occur either in a particular area,

e.g. the Ribbon Reefs, occur in a general area, e.g. east or west, although some clusters tended to be widespread across the region.

One theme common to virtually all of the biological datasets collated for this study is that the principle axis of contrast runs from east-southeast to west-northwest across the Torres Strait. Percentage cover of all the growth forms of live hard coral was higher in the east, as were reef fish and beche-de-mer diversity and abundance. Seagrasses on the other hand were more abundant in the west.

Information collated for this study included comprehensive datasets on reef fish, beche-de-mer and seagrass; most of which was identified to species level. Information on corals and algae was based on visual estimates of percentage cover morphotypes. Little data on the other reef biota was collected. Given that the coral reefs of the TS are of immense cultural, economic and aesthetic value and that they may be threatened by climate change and Crown-of-thorns starfish, it is imperative that we have comprehensive information on the status and natural variability of reef populations. We recommend that a comprehensive survey be conducted throughout the reefs of the area, in a similar manner to that conducted for the inter-reefal areas by the CRC Torres Strait (Pitcher *et al.* 2007), covering a wide range of taxonomic groups. From this initial survey, a subset of representative reefs could be chosen for future monitoring of key groups of organisms to assist in managing this valuable resource.

2. Background

The Torres Strait is an area of shallow continental shelf (maximum depth 20-60 m) covering ~60,000 km² and is bounded by the Papua New Guinea and Australian mainlands to the north and south respectively. Bathymetry throughout the area is complex, with approximately 800 islands and numerous reefs, shoals and areas of submarine sand dunes.

Research on benthic habitats in the region has been fairly limited although a recently completed study characterised the inter-reefal benthic habitats of Torres Strait added greatly to our knowledge. However, it is important to understand the limitations of our existing knowledge and to this end this project aims to perform a desktop gap analysis of the information collected to date and, where possible, analyse existing data to fill gaps.

Similarly, although there are a diverse range of coral reef communities in the region, they have not been well studied. Over the past twelve years CMAR has undertaken numerous diver surveys on various reefs throughout Torres Strait, generally with a focus on the stock assessment of important fishery species such as sea cucumbers or reef fish. Ancillary habitat data was also collected during these surveys. We have collated the reef associated data in an attempt to classify and characterise the reefs of the region in terms of some of the conspicuous biota and their physical environment.

This report is divided into two sections. The first deals with an analysis of benthic inter-reefal video data and describes how it has assisted in information gap-filling, while in the second we present a collation of Torres Strait reef habitat data, a characterisation of the reefs and examine the biophysical relationships between reef type and a range of environmental parameters.

3. Inter-reefal Seabed Mapping

3.1. Introduction

Between 2003 and 2006 the CRC Torres Strait project “Mapping and physical characterisation of key biotic and physical attributes of the Torres Strait Ecosystem” (Pitcher *et al.* 2007) mapped inter-reefal habitats and biota throughout the Torres Strait providing managers with baseline information to assist in ensuring the environmental sustainability of fisheries, regional marine planning and ecosystem understanding for multiple use management. This project collected comprehensive information on seabed biodiversity from 216 inter-reefal sites during two one-month long voyages. Data was collected using trawls, benthic sleds and a towed underwater video camera. Biota recovered from the sled and trawl catches were used to provide comprehensive distribution and abundance information on the benthic species assemblages found at the sites sampled with these devices. Video data was sub-sampled and analysed primarily for habitat characterisation rather than the identification, measurement and enumeration of biota.

Of the 216 sites visited, only 108 were sampled with all three devices (Table 1). Some sites were not sampled because of gear breakdown, but many (17) were only sampled with the video as the seabed was too rough to trawl or sled.

Consequently, two possible reasons for gaps in information on the distribution and abundance of seabed biota surveyed during the previous project are that there may have been taxa that were not sampled by the sled or trawl, but were recorded on video; secondly, the 17 sites that were too rough to sample with either sled or trawl may have biota that were not recorded at sites that were sampled using the extractive devices.

One of the objectives of the present project was to attempt to fill some of the information gaps by reanalysing the video data and to, as far as possible, identify all biota and to create maps of their distribution and abundance by modelling the relationship of various physical predictors to each taxa’s sampled distribution. However, because of the inherent limitations of video (e.g. poor visibility, movement of the target animals and inability to distinguish fine details necessary for accurate identification etc.) the taxonomic resolution obtained from the video data will not be as comprehensive as that from devices such as the trawl and sled identified.

Table 1. The number of sites sampled with each sampling device during the CRC Torres Strait project “Mapping and physical characterisation of key biotic and physical attributes of the Torres Strait Ecosystem”.

Sampling device	No. of Sites
Sled only	2
Trawl only	32
Video only	17
Sled and Trawl	9
Sled and Video	48
Video and Trawl	0
Trawl, Sled and Video	108

3.2. Methodology

3.2.1. Inter-reefal seabed mapping

Sampling design

The rationale behind the sampling design was to collate all available biological, physical and bottom water data for the Torres Strait; determine any bio-physical relationships; identify the environmental variables important in explaining the distribution and abundance of benthos and then to stratify the TSPZ seabed based on these variables, weighted by their importance. The sample design (Pitcher et al. 2004) collated 17 datasets of physical and biological data for Torres Strait. For the physical environment these data included: bathymetry, sediment grain-size and composition, water attributes, water chemistry and ocean colour. The biological data included: basic seabed habitats, seagrass and algae cover and some trawl samples. After checking quality and redundancy among sources, 32 physical variables were identified and mapped as potentially useful for modelling and stratification. A 0.01 degree resolution (~1.1 km) grid was established for analyses and sampling design, and the physical variables were re-sampled to this grid and mapped (interpolated where required), to provide a consistent set of full-coverage covariates at ~45,000 grid cells for the project.

Biological data was sourced from multiple legacy projects, each with different objectives, and was reconciled to useable common-denominator formats. The broad-scale physical factors important in structuring patterns in the biological data were identified. Seabed current stress was the most important variable, and others included: chlorophyll, turbidity, oxygen, salinity, nutrients, sediment grain size, and depth. These bio-physical relationships were used to predict and map the available biological data to the whole Torres Strait region, with an estimate of the uncertainty. The Torres Strait region was characterised by weighting each physical covariate by its biological importance, then grouping the 0.01° grid cells into strata that had similar physical attributes. The stratification was mapped and represented an interim surrogate characterisation of Torres Strait. Sites for sampling by this Mapping Task were selected from the bio-physical strata to provide representative coverage of the Torres Strait environment (see Pitcher et al. 2004 for further details).

While the original sampling design provided for 440 benthic sampling sites, of which 324 were identified for sampling by research trawl, the available funds for fieldwork would permit up to half those numbers of sites to be sampled distributed over about two-thirds of Torres Strait, given ideal seagoing conditions in terms of weather, navigation and breakdowns.

3.2.2. Field sampling

Vessel description

The video survey was conducted from the RV James Kirby; a 19.6 m long, steel-hulled vessel owned by the James Cook University in Townsville. The design is based on a commercial prawn trawler, has a small laboratory and a large rear deck area clear of obstructions for deploying numerous sampling gears and instruments, with a series of winches, through the stern A-frame. The vessel has a variable pitch three-blade propeller giving excellent low-speed control for towed video work. It generally has a crew of two, Master and Mate and can accommodate up to six scientific staff.



Figure 1. The 19.6 m James Cook University research vessel RV James Kirby.

The voyage was conducted from 24 March to 20 April 2005, of which three days each way were needed transiting between TS and Townsville. In addition, several days were required for fitting and trialing of gear (video and cameras, hydraulic winch and epibenthic sled) on the vessel in Townsville prior to departure for Torres Strait. One hundred and seventy-three sites were videoed and all videoing was done during daylight.

Towed video camera

An underwater video camera system (Drop-Cam,) was towed just above the seabed at each site wherever possible, for a distance of ~500 m, to characterise habitats and visible biota. The Drop-Cam system consisted of cameras, frame, fibre-optic towing cable, cable winch, hydraulic crane, CTD instrument, Control and data logging computers, video recorders and display monitors.

Video and still cameras and a CTD instrument were mounted within a galvanised steel frame. The video cameras were twin 3-chip Panasonic E300 digital video cameras fitted with 2.8 mm Fujinon lens. The video field of view was illuminated with 2 x 500 and 2 x 250 W lights. A Canon 20D 8.2 mega pixel digital still camera fitted with a 4 GB memory card, Canon EF 17-40 mm auto focus lens and two Speedlite 550EX strobes recorded still photographs of the seabed every 5 s during the camera transects. Twin lasers, spaced 28.5 cm apart, were fitted in front of the field of view of the cameras. When visible on the seabed these enabled measurements to be made of the benthos. All cameras, lights and strobes were housed in custom built housings rated to 3000 m.

The CTD was a Seabird SBE 19plus Seacat Profiler fitted with sensors for conductivity, temperature, pressure, oxygen, chlorophyll, turbidity and PAR. Data from the sensors was recorded at 0.25 s intervals and logged onto a computer database on board. Data from a

PAR sensor fitted to the vessel was also recorded and logged to enable comparison between surface and underwater light levels.

Also mounted in the frame were two pressure housings; one for the power supply and the other for a computing system for data and video collection. Apart from the digital still images, all data and video were sent to the vessel, in real time during the transect, from the Drop-Cam via a fibre link; the digital still images were downloaded at the end of each transect.

Data and video were converted to fibre optic media through a Focal 903 multiplexer and sent to the demultiplexer surface unit in the vessel via a single optical fiber. At the demultiplexer the signals were separated into data and video. Control of the Drop-Cam system (cameras, lights, lasers, CTD) and all processing and logging of data was done using in-house custom software on Pentium PCs, with the video recorded onto Panasonic DVC-Pro tapes.

The general procedure upon arrival at each site was as follows. The video camera was deployed and lowered to within approximately 0.5 m of the seabed. The CTD was set to switch on upon contact with seawater and recorded data throughout the deployment, transect and recovery. The vessel was then driven at approximately 1.5 knots, towing the camera frame for a distance of 500 m. Position and distance towed was recorded by differential GPS every 0.1 s. Video was displayed in real time on a monitor in the vessel laboratory, enabling scientific staff to raise and lower the camera, with remote joystick control of the winch, in order to maintain altitude above the seabed during the transect (Figure 2).



Figure 2. The Drop-Cam system recovered after completion of a 500 m video transect and the surface real-time monitoring, control and data acquisition system.

Tasks, events and navigation data logging

Custom software modules acquired navigation, video and Drop-Cam data on several computers. Independent modules acquired and logged date/time and position data from differential GPS, depth data from the vessel echosounder and heading data from the vessel gyrocompass on a navigation computer and shared these data onto a local area network. A custom map module also ran on the navigation computer to plot site waypoint and vessel position and constantly checked current position against the designed site waypoint list and also shared these data onto the network to ensure all modules indexed logged data against

the correct site number and date/time. A second computer ran modules for Drop-Cam system control and data (depth, pitch, roll, altitude) acquisition/logging, CTD data acquisition/logging and surface PAR data acquisition/logging. A third computer ran modules for control and data acquisition/logging of the DVC-Pro VTRs. A fourth computer ran the control and logging module for recording information on each camera deployment and recovery (Tasks module) and real-time summary of the video data (Events) logging module. The Tasks module was manually operated, with touch screens for recording start and end operations for the Drop-Cam transects. Each task record included site number, date/time, position, and depth data shared onto the network by other modules, as well as task type. The Tasks module also measured tow lengths, and the touch of the Drop-Cam start task switched on logging for all other modules.

The Events module was also manually operated – a scientific staff member used a modified keyboard to enter a real-time summary of the video imagery: seabed substratum type, biological habitats and conspicuous individual animals. A lookup table transcribed keyboard scan codes into the seabed types listed in Table 2, which were logged and prefixed with date/time, position, and depth data for each seabed event recorded. Note that the position recorded was that of the GPS antenna on the vessel, whilst the feature being recorded could be 20-25 m behind that because the camera was being towed from the stern.

Table 2. Substratum and Biological habitat types and animal events types entered in real time to annotated the video transect.

Substratum	Biological habitat	Animal events
Soft Mud	Bioturbated	Anemone
Silt (Sandy-Mud)	Bivalve Shell Beds	Ascidian
Sand	Alcyonarians:	Bryozoan
Coarse Sand	Sparse	Commercial Fish
Sand Waves / Dunes	Medium	Crab
Rubble (5-50 mm)	Dense	Crinoid
Stones (50-250 mm)	Whip Garden:	Gastropod
Rocks (> 250 mm)	Sparse	Holothurian
Bedrock / Reef	Medium	Hydroid
	Dense	Non Commercial Fish
	Gorgonian Garden:	Sea Pen
	Sparse	Solitary Coral
	Medium	Starfish
	Dense	Urchin
	Sponge Garden:	
	Sparse	
	Medium	
	Dense	
	Hard Coral Garden:	
	Sparse	
	Medium	
	Dense	
	Live Reef Corals	
	Flora:	
	Seagrass	
	Algae:	
	Caulerpa	
	Halimeda	
	No BioHabitat	

3.2.3. Video data processing

After the voyage the video was transferred from the DVC-Pro digital tapes via firewire to removable computer hard drives for storage and archiving. Each site's camera tow was saved as a separate .avi video file, a process which was automated using batch files scripted from VTR time-code data logged in the field. The site identity of each video clip was cross-checked during the video capture process, using the audio data stream from the vessel GPS, which had been recorded onto the audio tracks of the video tape during each camera tow. A decoder box and software translated the audio signal from each video recording into GPS data and displayed the date-time and positional information and calculated the corresponding site number for cross reference against the data logged in the field.

The objective of the analysis of the towed video was to describe sessile biota from the video recordings, and visually estimating their approximate size. A pair of red lasers was fitted to the camera frame 20 cm apart and their reflection on the seabed or epibenthos provided a reference to assist in size estimation. The operators who performed the video analysis underwent training before the process began and continued to consult together regularly throughout to maintain consistency. Procedural manuals were developed to assist with consistent decision making and identification.



Figure 3. Operator analysing seabed video.

Video files were copied from the removable hard drives onto work folders on the video analysis workstation computers (Figure 3), from where copies were deleted once analysed. A custom Delphi software application was used to view the videos and enter the data directly into ORACLE.

A total of 177 video transects each of between ten and twelve minutes duration were analysed for the presence of mobile or sessile animals. Sessile animals such as sponges, corals and hydroids were identified as far as possible and counted. For other biota that were difficult to count e.g. algae, seagrass areas of bioturbation etc. the operators made estimates of their percent cover across the whole transect. Any new objects were identified and an image was captured and saved to the database. A total of 19,914 records were made.

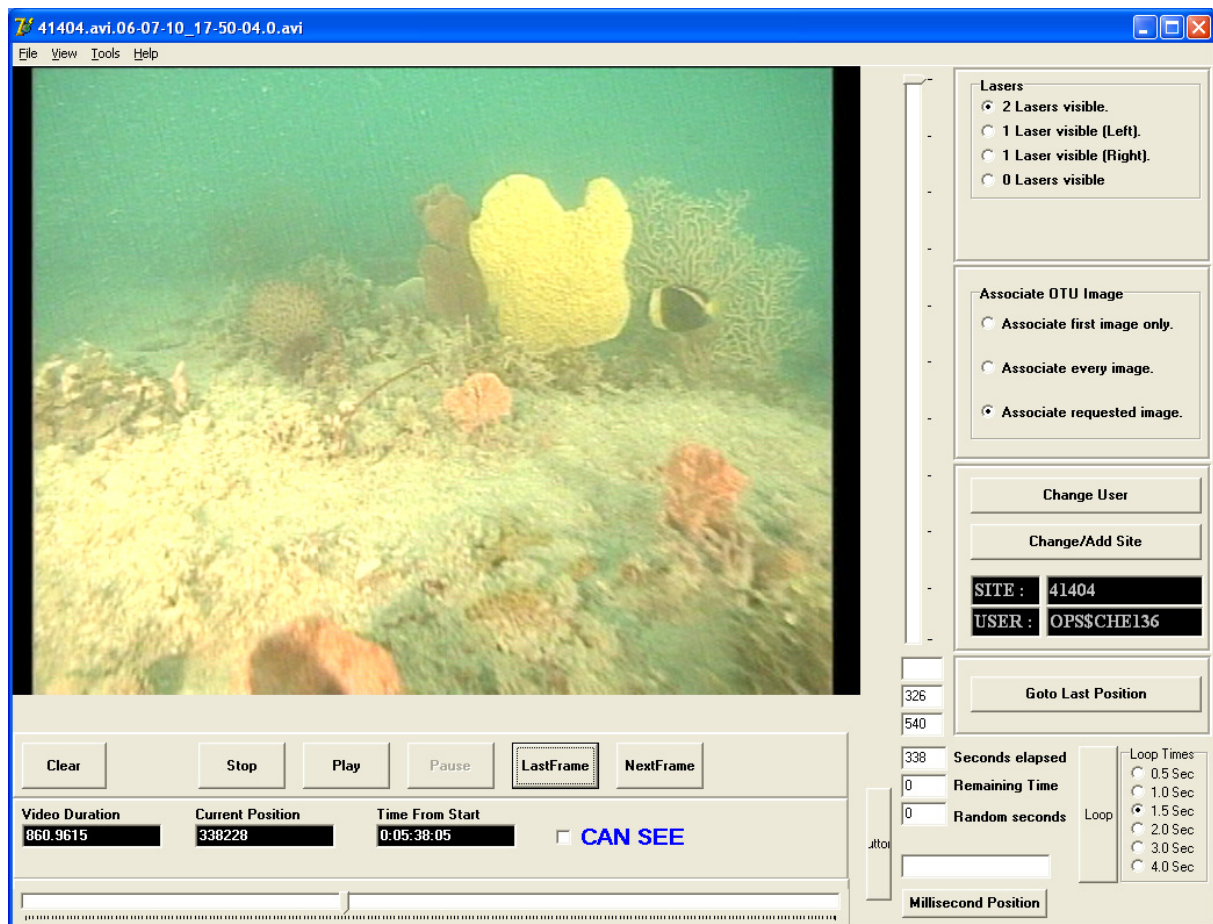


Figure 4. Screen dump of the front end of the custom Delphis software module used to view and analyse the towed underwater video.

Because of varying levels of clarity or turbidity and the altitude of the camera above the seabed, it was not always possible to identify the objects on the seabed; these biota were assigned to an unidentified category. Also the majority of sponges and ascidians from video are impossible to identify beyond the phylum level and so these were categorized based on gross morphology, e.g. barrel, club, branching etc.

3.2.4. Data analysis

Physical co-variate data

Physical co-variate data were collated from various internal and external sources and have been detailed. The following co-variables used in relation to the video data are given in Table 3.

Table 3. Description of physical co-variables used in modelling the “species” distributions of video-identified biota.

Co-variate name	Co-variate description
DEM4_BATHY	Bathymetry
DEM4_ASPECT	Aspect
DEM4_SLOPE	Slope
M_BSTRESS	Benthic stress
GA_CRBNT	% Carbonate
GA_GRAVEL	% Gravel
GA_SAND	% Sand
GA_MUD	% Mud
CARS_NO3_AV	Annual mean nitrate (μM) levels at the seabed
CARS_NO3_SD	Standard deviation of the annual mean nitrate levels at the seabed
CARS_O2_AV	Annual mean oxygen (mL/L) levels at the seabed
CARS_O2_SD	Standard deviation of the annual mean oxygen levels at the seabed
CARS_PO4_AV	Annual mean phosphate (μM) levels at the seabed
CARS_PO4_SD	Standard deviation of the annual mean phosphate levels at the seabed
CARS_SI_AV	Annual mean silicate (μM) levels at the seabed
CARS_SI_SD	Standard deviation of the annual mean silicate levels at the seabed
CARS_S_AV	Annual mean salinity (PSU) at the seabed
CARS_S_SD	Standard deviation of the annual mean salinity at the seabed
CARS_T_AV	Annual mean temperature ($^{\circ}\text{C}$) at the seabed
CARS_T_SD	Standard deviation of the annual mean temperature at the seabed
SW_CHLA_YAV	SEAWIFS Annual mean Chlorophyll a (mg/m^3) levels at the sea surface
SW_CHLA_YSD	SEAWIFS Standard deviation of the annual mean Chlorophyll a levels at the sea surface
SW_K490_YAV	SEAWIFS Annual mean K490 diffuse attenuation co-efficient at the sea surface
SW_K490_YSD	SEAWIFS Annual standard deviation of the K490 diffuse attenuation co-efficient at the sea surface
SW_K_B_IRR	Relative benthic irradiation derived from SEAWIFS
TRWL_EFF_I	Index of trawling effort produced from fishery logbook and VMS data

The video data were used to generate “species” distributions for each category identified. Two forms of data derived from the video tows were used: benthic *coverage* data (expressed as percent cover per site) and *count* data (expressed as the total number of individuals observed per site). The methods used to generate and collate these data sets are discussed in Section 3.2.3. The basic approach to distribution estimation involved modelling the

relationship of various physical predictors to each taxa's sampled distribution. These models are then used to predict on a continuous grid covering the whole of the Torres Strait. In order to generate reliable models, sufficient samples must exist in the sampled data set. Therefore, we did not attempt to model taxa that were observed at less than 10 distinct sites. After excluding these rare taxa, 30 coverage and 61 count taxa were considered for modelling.

Species distribution modelling is a developing and active field of research, and there are many alternatives with regards to methodology. Our strategy, for both coverage and count data sets, was to apply a non-parametric estimation method 'random forests' (discussed below) with empirical estimation of an optimal transform coefficient for the estimate. While we would not attempt to argue that this strategy is absolutely superior to alternative strategies (such as generalised linear modelling), it does have properties that make it appropriate in this context:

1. It can handle a large number of predictor variables and interactions effectively without overfitting;
2. It is less susceptible to model variability (such as that introduced by stepwise variable selection procedures in linear modelling);
3. It extrapolates conservatively (via the property of local averaging in predictor space) when making estimates at new sites; and
4. It includes a robust measure of generalisation performance (through 'out-of-bag estimates') enabling objective evaluation of model efficacy.

The follow technical sections detail and justify the use of random forests with empirical estimation of the prediction scale for species distribution estimation. We outline the well-known issue of aggregation in ecological data, and examine how this issue relates to non-parametric estimation, random forests in particular. Our technique for addressing aggregation in non-parametric ecological estimation is introduced and justified.

3.2.5. Accounting for aggregation / dispersion

In ecology it is well known that the observations of individuals of the same in space or time are not independent events due to tendency of biota to aggregate. By definition, the 'aggregation effect' is separate from variability due to physical, temporal, or other parameters, and therefore is manifested during modeling as inhomogenous error variance. Given realistic sampling effort in marine ecology, after the variability arising from covariates has been taken into account, this is generally manifested in the form of (relatively) few data points with disproportionately large residuals.

In a parametric context, when these residuals exceed those predicted by a model (e.g. the Poisson), this is termed over-dispersion. The degree of aggregation must be taken into account in order for valid model optimization and interpretation. The modern paradigm in generalised linear modeling is to adopt an appropriate model for the error distribution (Venables and Ripley 1994), a sophisticated and flexible approach often applied in ecology (e.g. Kendal 2002). However, when the over-dispersion is extreme and/or unknown, this can present challenges. For example, since the negative binomial distribution for count data often incorporates an estimated dispersion parameter, estimation of the deviance depends on accurate specification of the dispersion. This can present difficulties, as the dispersion parameter depends on estimation of the residuals, which may depend on model selection criteria that in turn depend on appropriate specification of the distribution. Utilising the 'quasi-likelihood' approach can lead to difficulties in model selection, since model comparison criteria generally require explicit definition of the likelihood. Effective specification of the error distribution can also be cumbersome when the empirical distribution is not well described by any of the standard models in the generalised linear family. For example, although the

presence of many zeros in ecological data does not necessarily imply zero-inflation in count data (Warton 2005), abundances quantified on a continuous scale often demand the use of specialized or composite error distributions. Practical implementation, in our experience, can sometimes lead to sub-optimal results.

Finally, there is increasing interest in a variety of pattern-recognition or machine-learning (i.e. non-parametric) based methods in empirical ecological modeling (Thuiller et al. 2003; Phillips et al. 2004; Segurado and Araujo 2004; Pearson et al. 2007). There are many aspects of estimation in the ecological context driving this trend. For example, researchers may be dealing with a large number of potential predictors (possibly exceeding the number of biota presences) and require flexible methods for searching this space without over-fitting. The structural effects from the covariates may be non-linear or characterised by complex and largely poorly understood interactions. Also investigators may wish to leverage information from the distribution of other species to improve the estimation of rare species, demanding a multivariate perspective (Elith et al. 2006; Ferrier and Guisan 2006; Ferrier et al. 2007). In this non-parametric context, explicit modeling of the stochastic components does not usually take place. When, for instance, the variance of the estimand has some power relationship to the mean, it is desirable to ensure, at the very least, that this property does not interfere with efficient estimation of the systematic effects.

3.2.6. Random forests

First proposed by (Breiman 2001), random forests are a powerful non-parametric estimation method. Random forests may be described as bagging classification and regression trees (CART¹). That is, many bootstrap samples are selected from the data and a binary decision tree is fit to each. When splits are made at each node in each CART, a subset of predictors are chosen at random for evaluation. In regression, the predictor and the split point is chosen so as to minimise a sum absolute or square residual criterion.

Trees constituting the random forest are not pruned using cross-validation, but are instead fully grown (i.e. over-fit). These measures act to increase the random variability of the constituent trees within the forest so as to reduce correlation in the errors of the ensemble of estimators. Given

- a) sufficient independence between estimators,
- b) some ability of the CARTs to approximate the underlying function, and
- c) a reasonably large number of trees,

then the final estimate produced by averaging estimates over the ensemble can be expected to have good bias / variance properties (Boinee, 2005). As an estimator, they are based on sound bootstrapping principles (Davison and Hinkley, 2003), and a growing body of empirical results suggests random forests are a very effective applied estimation tool. By averaging estimates produced by a large set of CART, they avoid the instability, model uncertainty, and much of the bias associated with any single CART. A significant attraction of random forests is the manner in which they satisfy the competing goals of high model flexibility and strong generalization ability.

Random forests are essentially an instance of familiar aggregating methods (Breiman 1996; Schapire and Singer 1999) applied to a family of weak but flexible learners (in the form of

¹ CART is a standard tool in statistical pattern recognition and will not be described in detail here. Readers may refer to a text in pattern recognition or machine learning, e.g. Duda RO, Hart PE, Stork DG (2001) Pattern Classification. Wiley, New York for an introduction.

overfit and randomized CART). Much of the power of random forests is derived from the bootstrap aggregation procedure, in which random samples (of the same size as the original dataset) are drawn with replacement. Due to the replacement property, a bootstrap sample will incorporate ~ two-thirds ('in-bag' (INB)) unique samples from the original data. Thus, ~ one-third of the CART will not have 'seen' any given datum. Utilising only that subset of CART in calculating the average random forest estimate yields an 'out-of-bag' (OOB) estimate, which is similar to having an independent 'validation set' for un-inflated estimates of model performance.

Although the precise mapping implemented by the random forest is difficult to specify, a number of diagnostics are available. For example, *variable importance scores* can be calculated by randomly permuting each predictor, and calculating the average decrease in the proportion of explained variance. *Partial effects* of each predictor can be calculated by integrating the estimates with respect to each predictor individually.

Despite the attention paid to random forests in neuroscience (Lehmann et al. 2007), sensory analysis (Granitto et al. 2007), and remote sensing (Pal 2005; Gislason et al. 2006), relatively little attention has been paid to random forests in ecography, species distribution and biodiversity estimation. For instance, random forests were not included in a recent comprehensive comparative study (Elith et al. 2006) and review of species modeling approaches and issues (Araujo and Guisan 2006), where GLM, GAM, neural networks, CART and other more specialized methods were considered. However, interest in random forests in ecology has recently been increasing. Random forests have been considered by (Prasad et al. 2006) and were applied by (Garzon et al. 2006; Garzon et al. 2007) for estimating species distribution and habitat suitability, as well as (Schwartz et al. 2006) for predicting future species distributions.

As mentioned above, the CART constituting the random forest estimator are based on optimising sums of squares or sums of absolute deviations. From standard regression theory (Box and Cox 1964) it is known that departures from normality decrease the efficiency of such estimators. In realistic ecological surveys, outliers often exist, and the variation of the response may be related to the mean estimates. Non-constant variation gives greater weight to data with higher variation, and therefore, as with other forms of regression, it is often desirable to transform the response variable when applying CARTs (De'Ath 2002).

3.2.7. Transformation in non-parametric estimation

One way to improve the efficiency of sum of squares – type regressors is to consider monotone transformations of the response. A power transformation of the estimand is one approach to better satisfy the assumption of normal and constant variance, to 'tame' large outliers, or more generally, to find the optimal scale in which the fixed effects influence the response (Rubin 1984). A given power transformation and its associated model may be compared with other possible transformations by some measure of model efficacy (e.g. heteroscedacity of model residuals). We propose that the efficiency of modern non-parametric regression estimators can be improved by implementing this well-established technique. In particular, we expect that random forest estimation of ecological distributions can be more effective using empirical transformation of the biological response.

In classical statistics, the issue of an inhomogenous error distribution has been approached by transformation (Box and Cox 1964) or weighting of the response (Han 1987). Given an arbitrary response vector $\mathbf{y} = \mathbf{y}_u = (y_1, y_2, \dots, y_n) \geq 0$, the power transform may be defined as:

$$y_u^{(\phi)} = \begin{cases} \frac{y_u^\phi - 1}{\phi} & \phi \neq 0 \\ \log y_u & \phi = 0 \end{cases}.$$

We are concerned with ecological count data, for which outliers and over-dispersion are the primary issues, and which by definition does not contain records in the region $(0,1)$. In this context, it is sufficient to simply consider the transformations

$$y_u^{(\phi)} = y_u^\phi \quad \phi \in (0,1]$$

It was also shown that given a linear model of the transformed response

$$y_u^{(\phi)} = \mathbf{x}_u \beta + e_u$$

where the response is a linear combination of the covariates $\mathbf{x}_u$, and e_u is independent and drawn from a known distribution, then ϕ may be optimised simultaneously with the parameters β so as to maximise the likelihood in relation to the untransformed observations (Box and Cox 1964).

This method of optimising ϕ cannot be applied to random forests for a number of reasons, first among them being the fact that a likelihood criterion does not exist. We can imagine a much more general model

$$y_u^{(\phi)} = g(\mathbf{x}_u) + e_u$$

where $g(\mathbf{x}_u)$ is a random forest estimate (or more generally, the output of some arbitrary non-parametric non-linear estimator), and where the error term has an unknown distribution. In place of likelihood, we consider applying Kendall's rank-correlation to the response and the model estimates

$$\tau(y_u^{(\phi)}, g(\mathbf{x}_u)) = \frac{4k}{n(n-1)} - 1,$$

where n is the number of observations, and k is the sum, over all observations, of observations ranked after the given observation, by both rankings (Kendall 1938). Considering the situation where $g(\cdot)$ is linear with respect to finite parameters, (Han 1987) shows that as the true e_u departs from normality, maximizing Kendall's τ criterion leads to an estimator with greater efficiency than the corresponding likelihood criterion. Further it is shown that ϕ uniquely maximises τ in large samples. We do not have corresponding theory showing that these results hold when $g(\cdot)$ is a non-linear stochastic estimator. However, given that $g(\cdot)$ is optimised so as to minimise the residual (square) deviations, we have a strong expectation that similar properties should apply. The approach of maximizing Kendall's τ criterion for a random forest estimator will be tested empirically in this study. Although we are lacking methods to estimate ϕ and $g(\cdot)$ simultaneously, the speed of

random forest estimation, combined with modern computing resources, entails that iterative estimation of ϕ and $g(\cdot)$ is relatively undemanding.

3.3. Results and Discussion

3.3.1. Estimation of spatial distribution of counts

As described above, the random forest estimator was applied with empirical transformation of the response to both the benthic coverage and video count data. Ten values of ϕ were used, equally spaced between 0.05 and 1, leading to a set of model estimates for each species. The optimum ϕ^* is estimated via a cubic polynomial (also described in (Browne et al. 2007)) and the final model estimates generated. This section discusses in some detail the transform-estimated random forest models for the two example taxonomic categories the hard coral *Turbinara* (a genus containing 13 species which have been recorded in Torres Strait) and the broad group of fish Actinopterygii (a Class containing several hundred species which have been identified in the Torres Strait). These will be discussed as typical examples of relatively efficacious models with different levels of ϕ^* . They are discussed here to provide some insight into how the empirical distributions are accounted for by the analysis method.

Actinopterygii 'bony fish' sp. is an example where the optimal power transformation ϕ^* is relatively larger, whilst Anthozoa *Turbinara* sp. represents a case where the empirical optimal transform was determined to be close to 0. Figure 5 and Figure 6 show the estimated distribution of *Turbinara* sp. and 'bony fish', respectively, with estimates coloured using a rainbow colour-map. The colour scale is scaled relative to each species so that in each case, red corresponds to the maximum estimated value, and blue corresponds to an estimate of zero. This scaling is necessary to provide visual contrast so as to provide clear maps for all taxonomic types. However, it is very important to remember (a) that this is a relative colouring scale and (b) that the validating of inferring real differences between "high" (red) and "low" (blue) estimates is entirely dependent on the out-of-bag rank-order correlation statistic (provided at the bottom of each map) and the range of the estimates (provided by the colour legend). The correlation statistic provides an important measure of prediction uncertainty, and interpretation of estimated distributions as absolute indicators of abundance should only be done with a clear understanding of this issue.

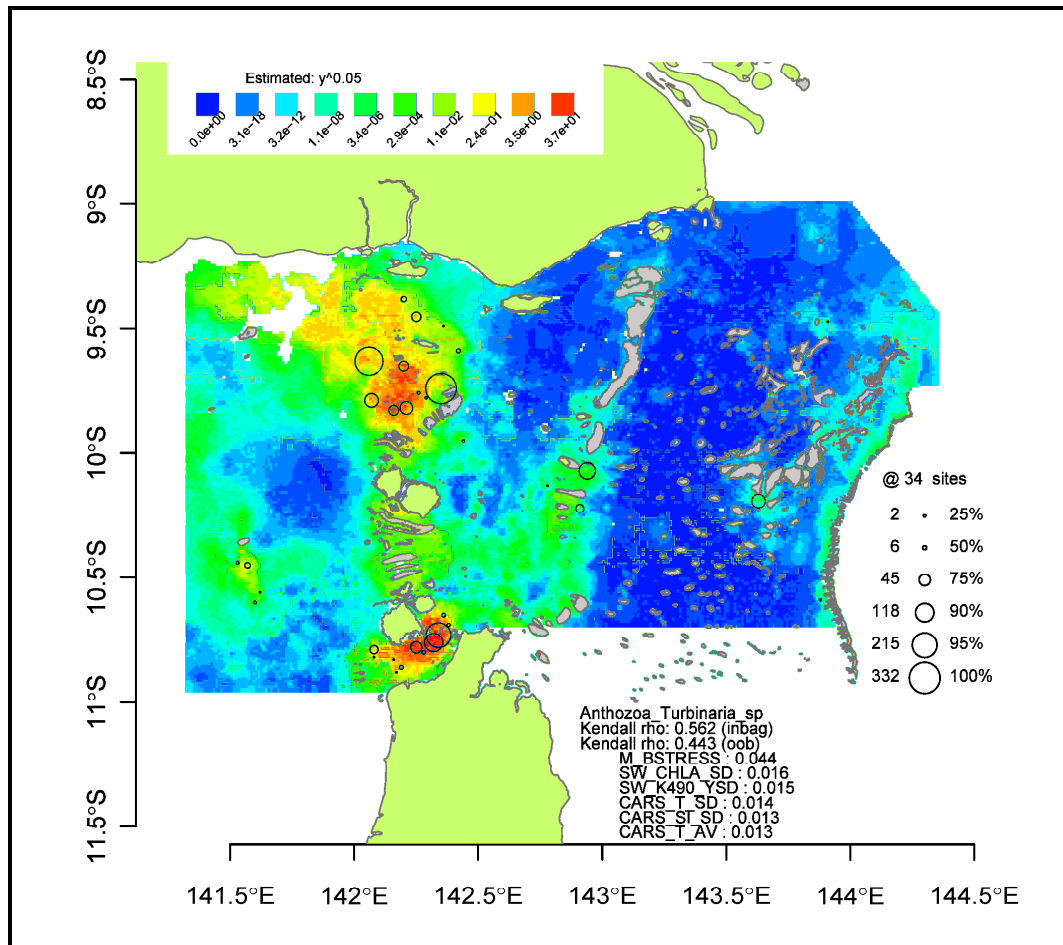


Figure 5. Estimated spatial distribution of *Anthozoa Turbinaria* sp.

The legend at the top of each figure shows the colours corresponding to estimates transformed back to the original scale of the data. The plotted open black circles indicate the spatial locations where that taxonomic category was actually observed on the video, with circle area proportional to the number of individuals observed. The legend on the right-hand side indicates the total number of sites at which the taxon was present, and provides a key to counts at various percentiles from the empirical distribution. The percentiles denoted on the right correspond to the truncated (presence-only) data: i.e. fifty percent corresponds to the median non-zero count. The text at bottom summarizes the τ criterion obtained INB and OOB. The variable importance measure for the top six predictors (ordered) is also displayed.

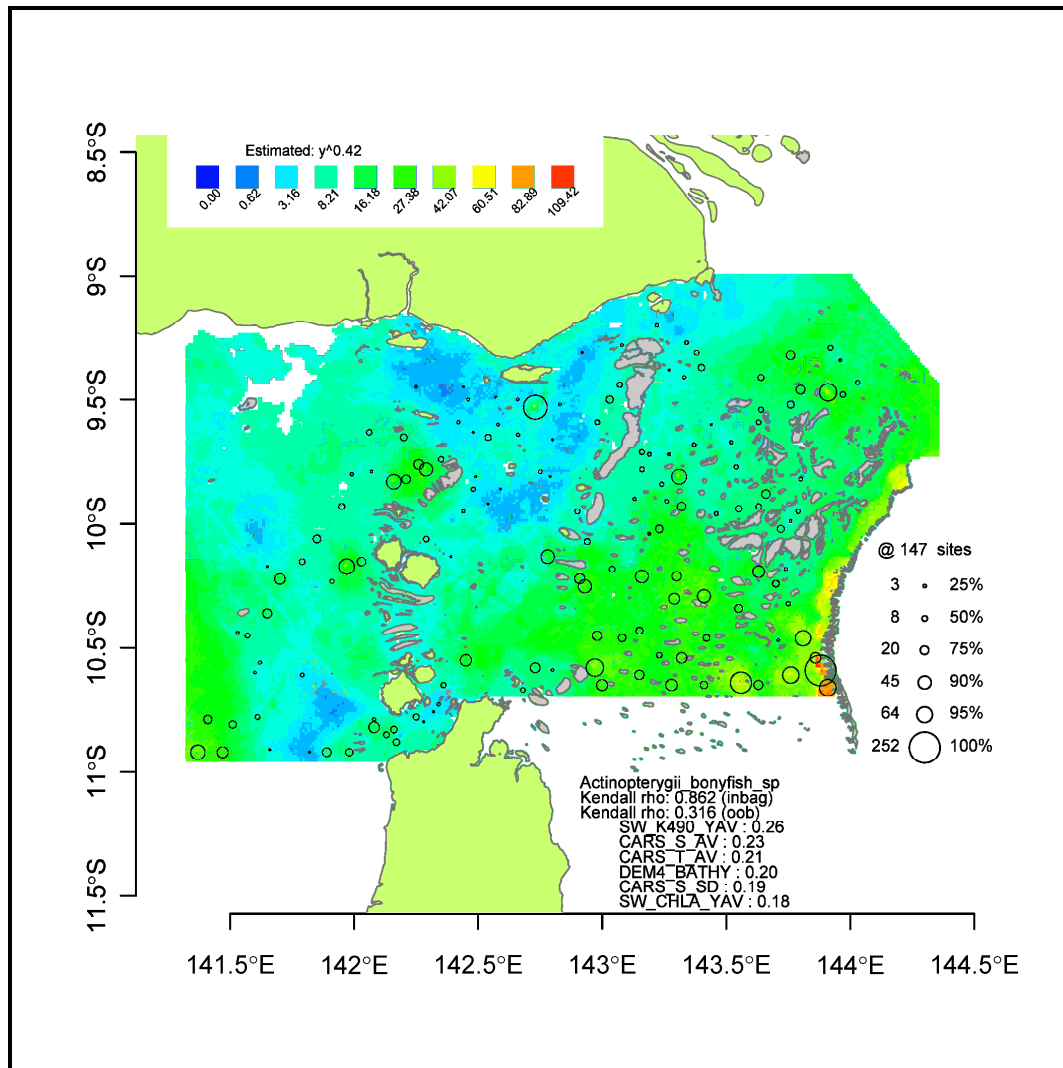


Figure 6. Estimated spatial distribution of Actinopterygii 'bonyfish' sp.

The partial dependence of the estimates on individual predictor variables is of significant interest to ecologists wishing to understand the environmental and habitat preferences of particular species. Figure 7 and Figure 8 plot the partial effects, for the twelve most important predictors, for *Turbinara* sp. and 'bony fish', respectively. Partial plots are ordered in importance, starting from the bottom left and running left-right. The x-dimension measures the predictor score and the y-dimension measures the partial effect on the response. The interpretation of partial effects with respect to predictors should take into account the distribution of data across the observed range of each predictor. Therefore, the partial effect curve is plotted according to percentiles with respect to each predictor. The nine circles on each plot correspond to 10 percentile increments from 10 to 90, and therefore can be used to gauge the quantity of data that segments of the partial effect correspond to.

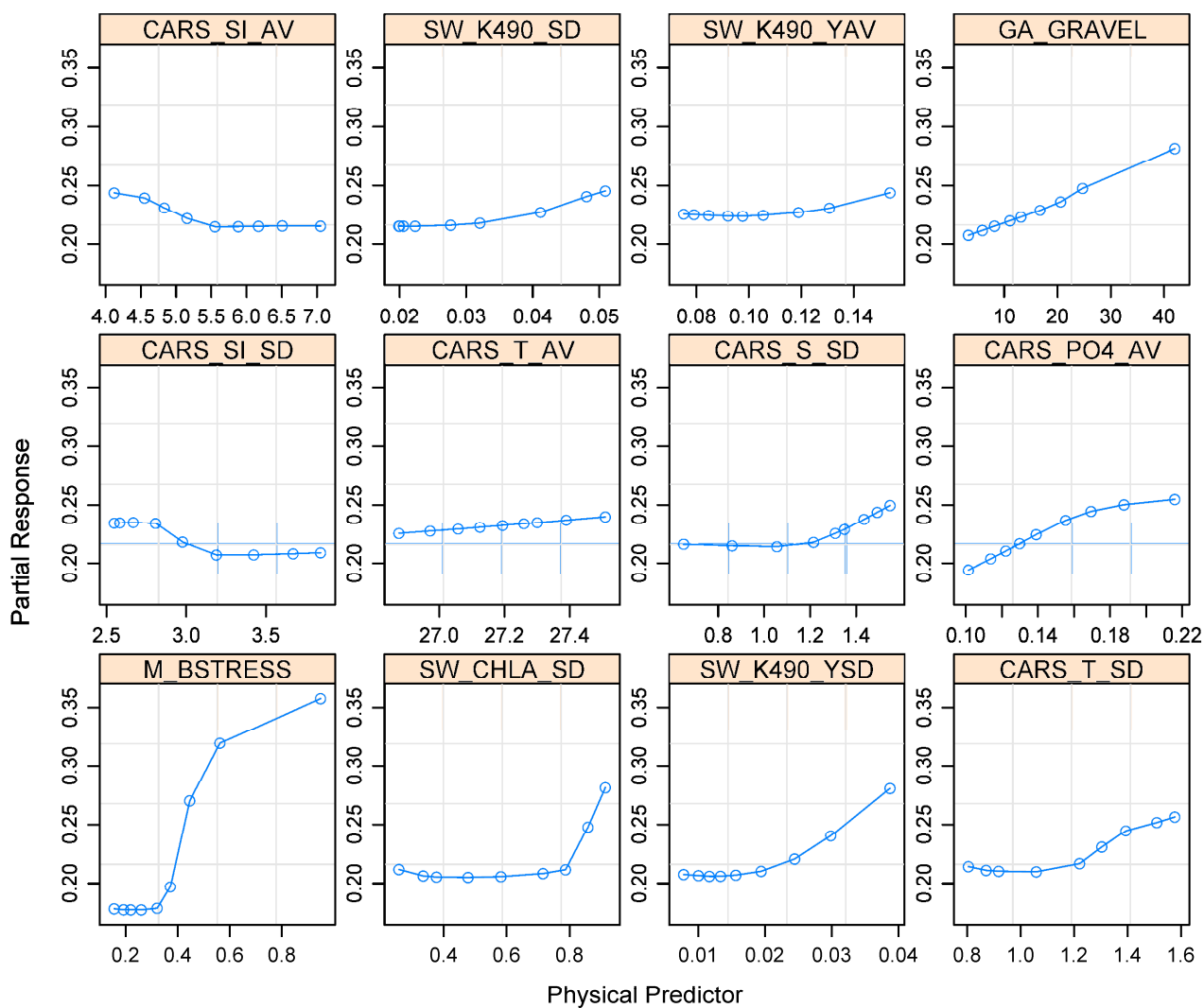


Figure 7. Partial effect curves for *Anthozoa Turbinaria* sp. See **Table 3** for the definitions of the covariate abbreviations.

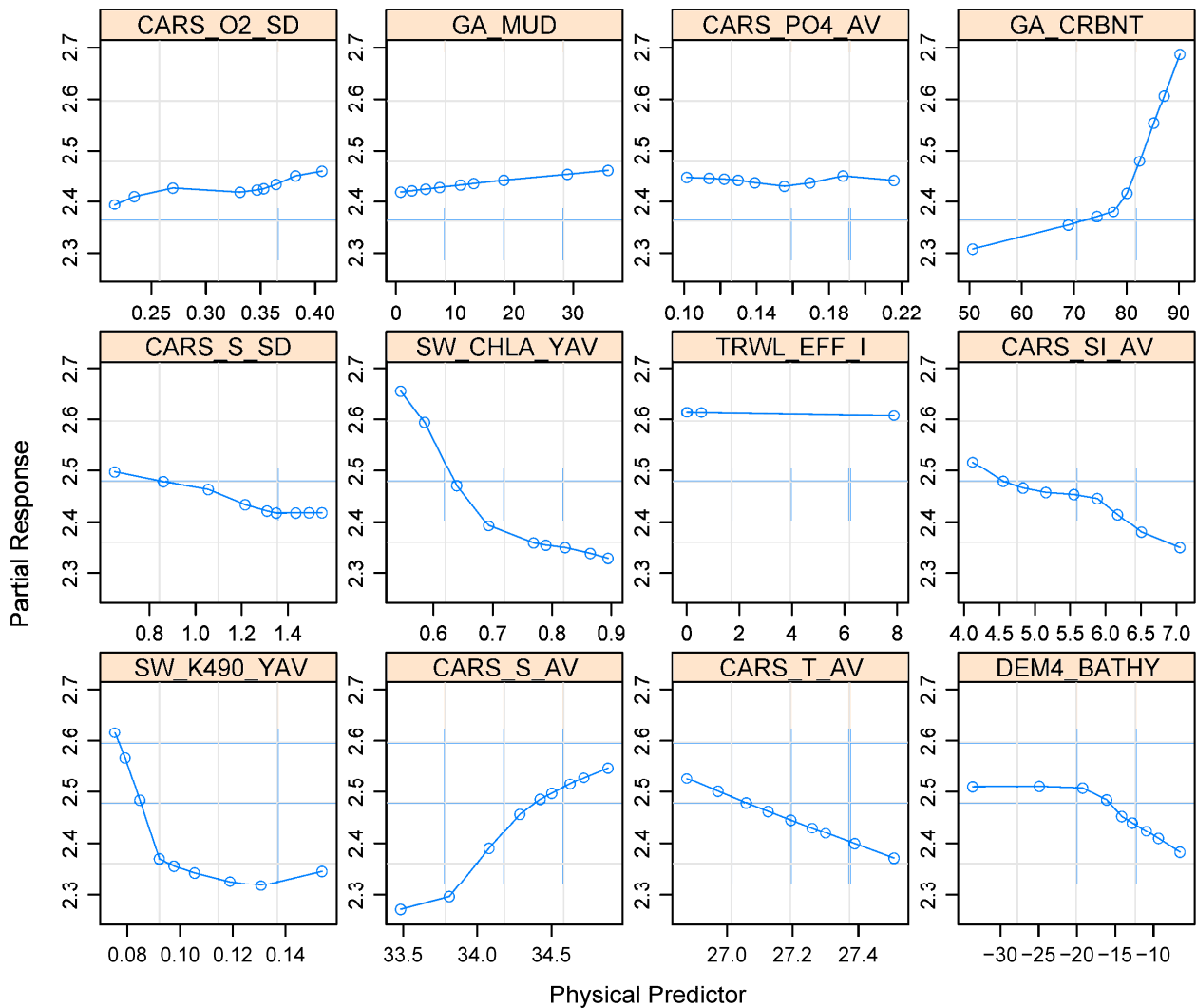


Figure 8. Partial effect curves for Actinopterygii ‘bonyfish’ sp. See **Table 3** for the definitions of the covariate abbreviations.

Anthozoa *Turbinara* sp. is a category of hard coral identified to Genus level. The apparently high specificity of habitat preference for *Turbinara* in Figure 5 may be explained as being partly due to the fact that it is a sessile taxon that attaches to hard bottom types, and partly due to the relatively higher level of taxonomic resolution compared to the broad category Actinopterygii. These properties would be expected to contribute to fewer presences (relative to a mobile, non-specific taxonomic category), and a greater degree of aggregation in the count data. Aggregation in the biological response arises directly from spatial autocorrelation of benthic characteristics e.g. patches of hard substrate suitable for the settlement of coral will tend to occur together on the seabed. We might expect the physical covariates to be relatively good predictors of whether local habitat is potentially suitable for *Turbinara*. However, there are issues of scale in estimating quantity of *Turbinara* from relatively coarsely resolved covariates (0.1 degree), considering the natural scales of variation in sedimentary conditions. That is, whether a particular video run happened to occur on a highly suitable or marginally suitable benthic substrate was to a large extent unpredictable. This would lead us to expect to do less well in modeling variation in counts. A further numerical reason for expected poorer performance on estimating variation in quantity is simply that relatively fewer presences provides less support for generating a model that is able to generalise well (due to higher model variability).

These characteristics of *Turbinara* are consistent with the functioning of the random forest estimator with an empirically estimated transformation. Compared to other species models, the random forest estimator produced a localised distribution. Although a relatively small proportion of the Torres Strait seabed appeared suitable for *Turbinara*, given suitability, there was a good chance of observing relatively large quantities of the coral (Figure 7). This is consistent with the cumulative distribution function (CDF) for this species², with relatively few presences, but a relatively large preponderance of larger counts when there was presence. Benthic stress exceeding 0.5 (well known to lead to rocky benthic conditions) was found to strongly correspond with this Genus's response (Figure 25), and was estimated to be the best predictor by a large margin (see variable importance scores in Figure 5). The optimal power transform for *Turbinara* was the minimum considered ($\phi^* = .05$), resulting in a very strong reduction in the relative variance of the quantity component relative to the presence / absence component of the count data. This tells us that, relative to the presence / absence of *Turbinara*, little or none of the variability in quantities was predictable. By deduction, the relatively high overall criterion obtained (by marine ecology standards) of $\tau^{\text{out}} = .44$ indicates that the presence of *Turbinara* was relatively well explained by the covariates.

Actinopterygii 'bony fish' sp. is a broad category of fish identified to Class level. The non-specific taxonomic resolution, combined with the mobility characteristic of this taxa, would lead us to expect a less aggregated distribution. This is confirmed in Figure 6. The empirically estimated power coefficient in this case was 0.42, which corresponds to retaining a significant proportion of the variation in counts, as opposed to presence / absence. Given the diffuse distribution of 'bony fish', variation in predictors would be more likely to lead to gradual changes in abundance rather than, say, the sharply delineated distribution of *Turbinara* sp. Thus, in both cases of 'bony fish' and *Turbinara* sp., the optimal empirical power appears to be related to the sampled distribution of the taxa.

Figure 9, Figure 10 and Figure 11 show some further examples of taxa distribution estimates for; Holothuroidea *Holothuria* sp, Anthozoa *Junceella fragilis*, and Anthozoa *Solenocaulon* sp. (bare), respectively. The Holothurian taxa is an example of a relatively rarely occurring taxa with a dispersed distribution. The model is quite poor, with an out-of-bag rank correlation of 0.2. The uncertainty is reflected in the small scale of the estimates (as recorded in the colour legend). Due to the intrinsic stochasticity involved in sampling marine biota, and the uncertainty introduced by issues of scale, and adequacy of covariates, this level of model adequacy is not uncommon. We note that another reason for weaker models is simply the lack of information available for building models of rare taxa. We recommend that estimates of species distribution generated by models with lower criterion scores are treated with caution, and interpreted primarily as plausible hypotheses of distribution, rather than firm conclusions. Whilst the model criterion of Anthozoa *Junceella fragilis* is slightly higher than Holothuroidea *Holothuria* sp., it also falls into this category of model. The distribution Anthozoa *Solenocaulon* sp (bare), which has a visually distinct distribution; occurring almost exclusively in the eastern TS, a relatively strong out-of-bag criterion, can be interpreted with more confidence. These species all had estimated power coefficients of 0.05, which suggests that the effect of physical covariates on the presence or absence of these taxa could be estimated, but that variability in quantity could not. Anthozoa branching sp (Figure 12) provides a counter example, where although the overall model adequacy was relatively weak, the estimated power coefficient was relatively high (0.41). In this case, it appears that variability in quantity was to some extent estimatable from the covariates.

² The 'circle legend' on the right of Figure 5 provides a good indication of the CDF for Anthozoa *Turbinara* sp.

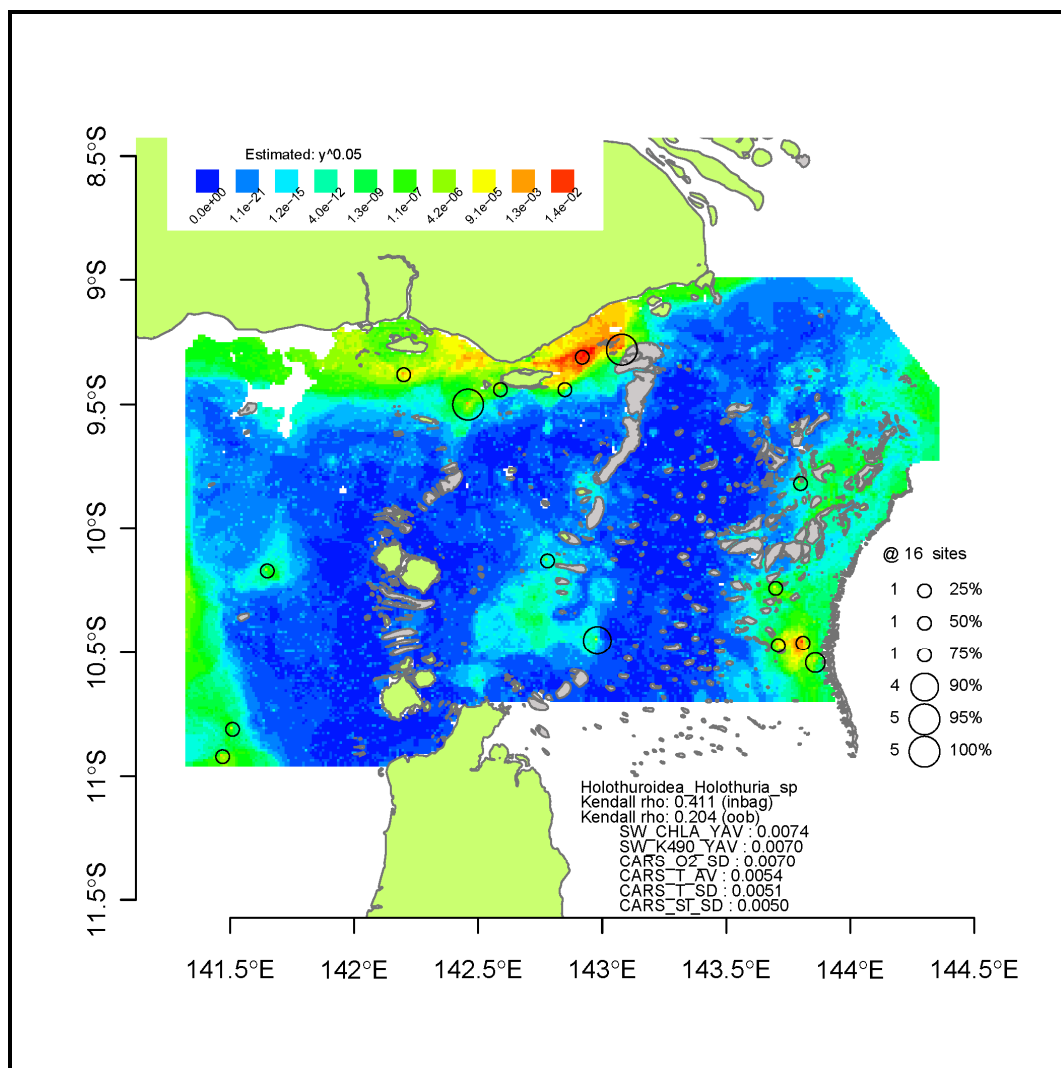


Figure 9. Estimated spatial distribution (count) of *Holothuroidea Holothuria* sp.

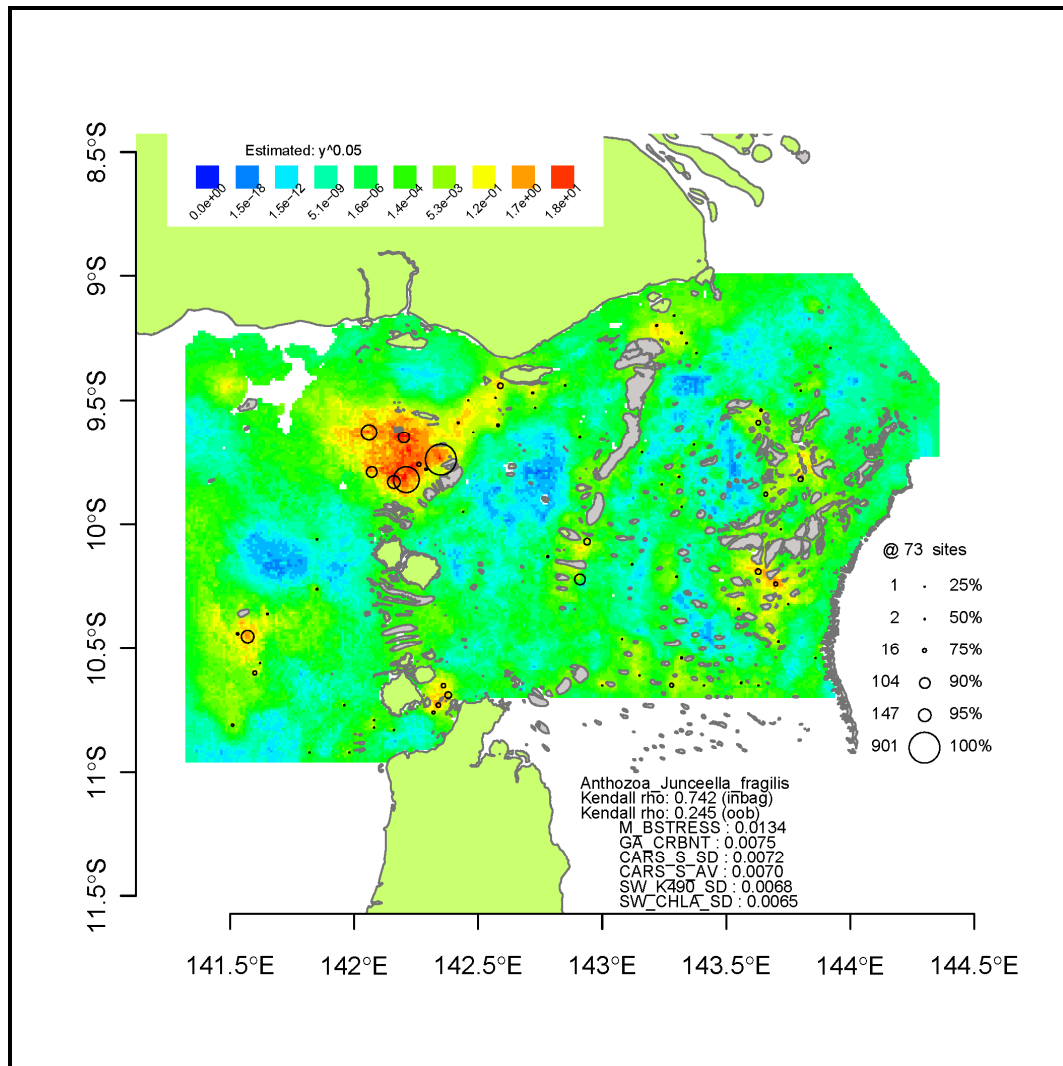


Figure 10. Estimated spatial distribution (count) of *Anthozoa Junceella fragilis*.

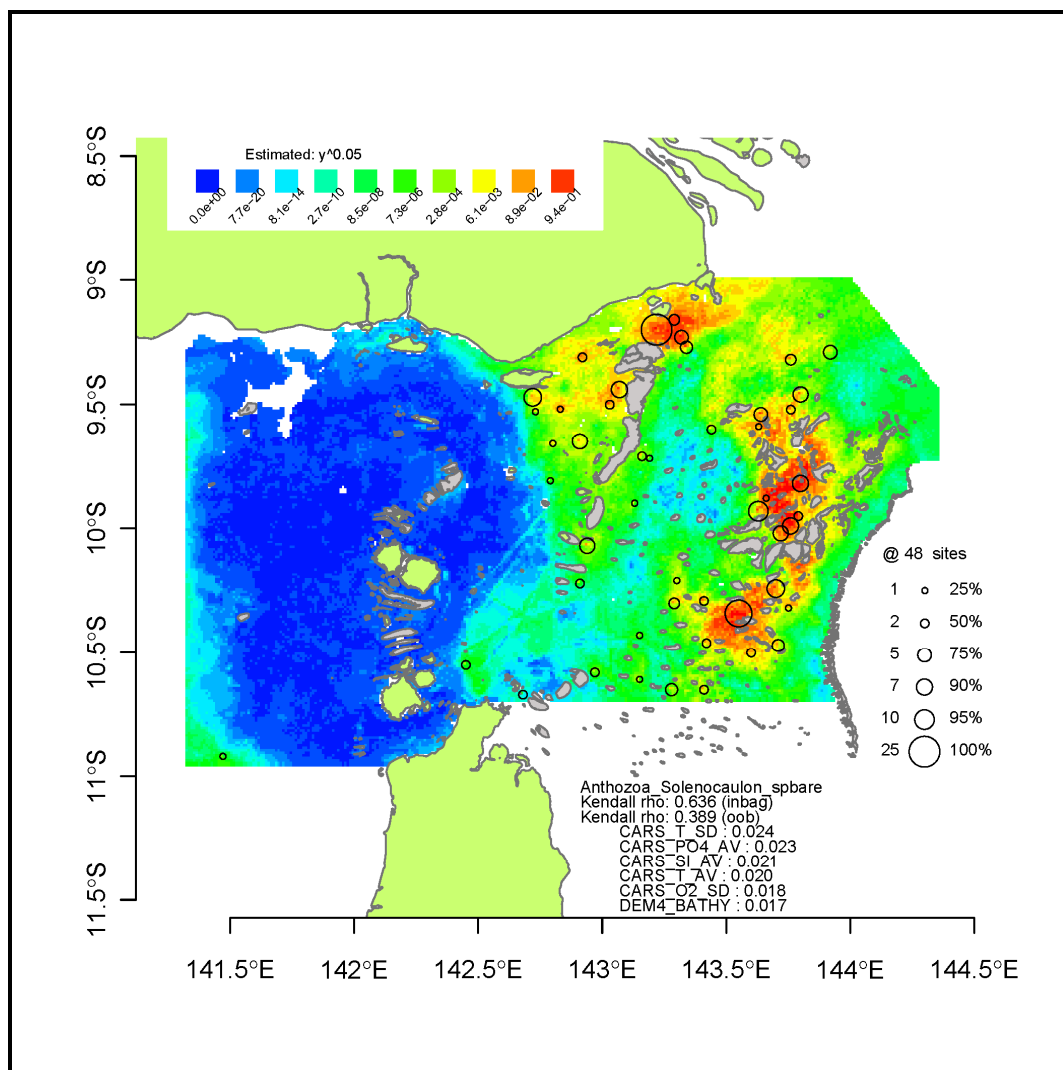


Figure 11. Estimated spatial distribution (count) of Anthozoa *Solenocaulon* sp. (bare).

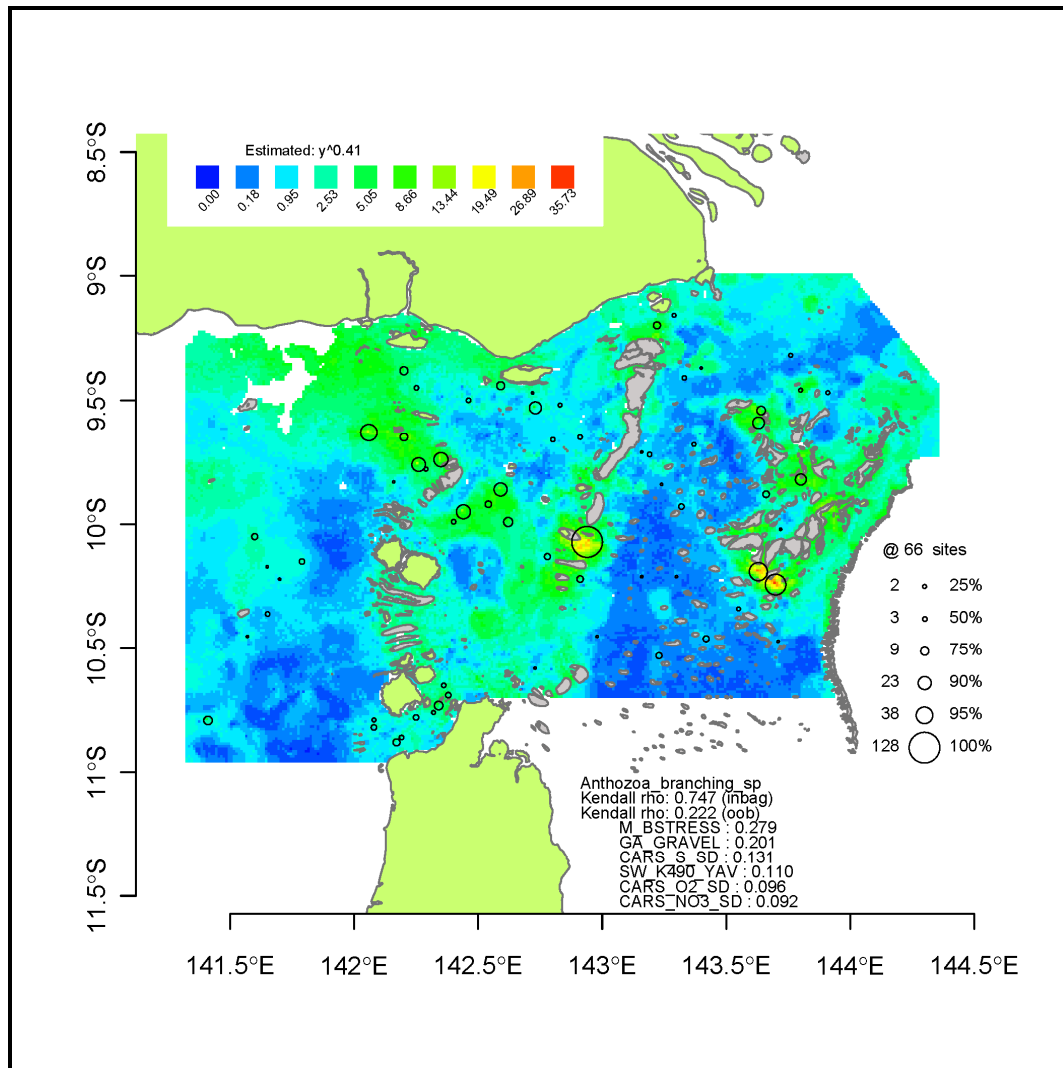


Figure 12. Estimated spatial distribution (count) of Anthozoa branching sp.

3.3.2. Estimation of spatial distribution of cover types

As stated earlier, the methods used for estimation of cover types were identical to those used for estimation of the taxa counts. Therefore, the points made in the sections above regarding model evaluation and interpretation applies to cover types as well. From the point of view of estimation, the main difference is that the quantity estimated in the former is a count quantity, and the quantity estimated in this case is a proportion (expressed in percentages). Although the maximum quantities in this case are obviously restricted to 100(%), the empirical distributions were practically similar to count data, with a preponderance of zeros, small percent cover observations, and a few larger coverages $\ll 100$. Thus, despite the fact that the theoretical distributions are clearly different, from the perspective of non-parametric estimation were able to treat them similarly. We will restrict the discussion to a consideration of estimated distributions for representative benthic cover types.

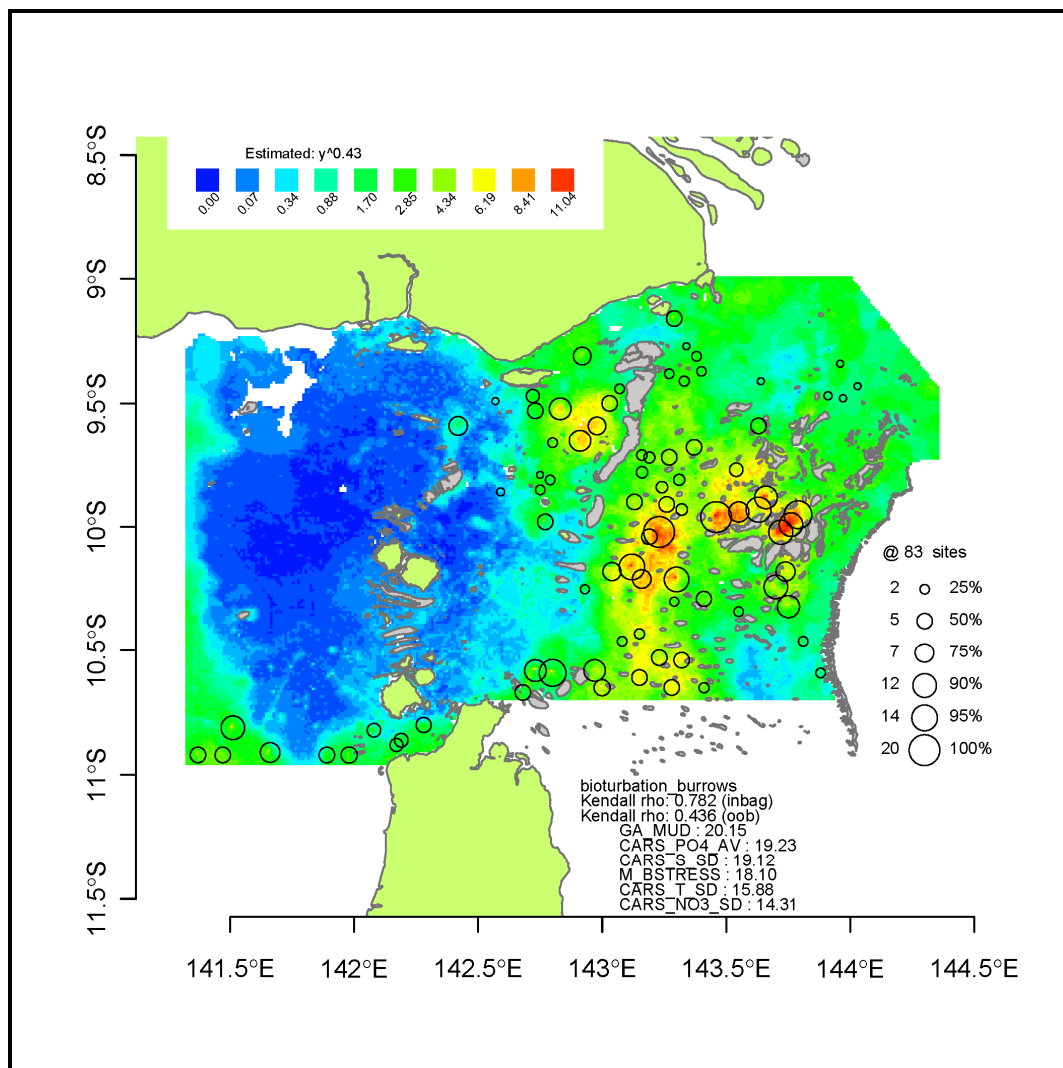


Figure 13. Estimated spatial distribution (percent cover) of bioturbation sediment.

Figure 13 shows a very common benthic cover type, “bioturbated sediment” with a distinct spatial distribution that is well explained by the physical predictors. The estimated power coefficient of 0.43 suggests that the model was able to predict not only whether any proportion of a site would be characterised by burrows, but also the degree of cover

observed. Some of the efficacy of this model is certainly due to the availability of relatively more observations (83) for model optimisation.

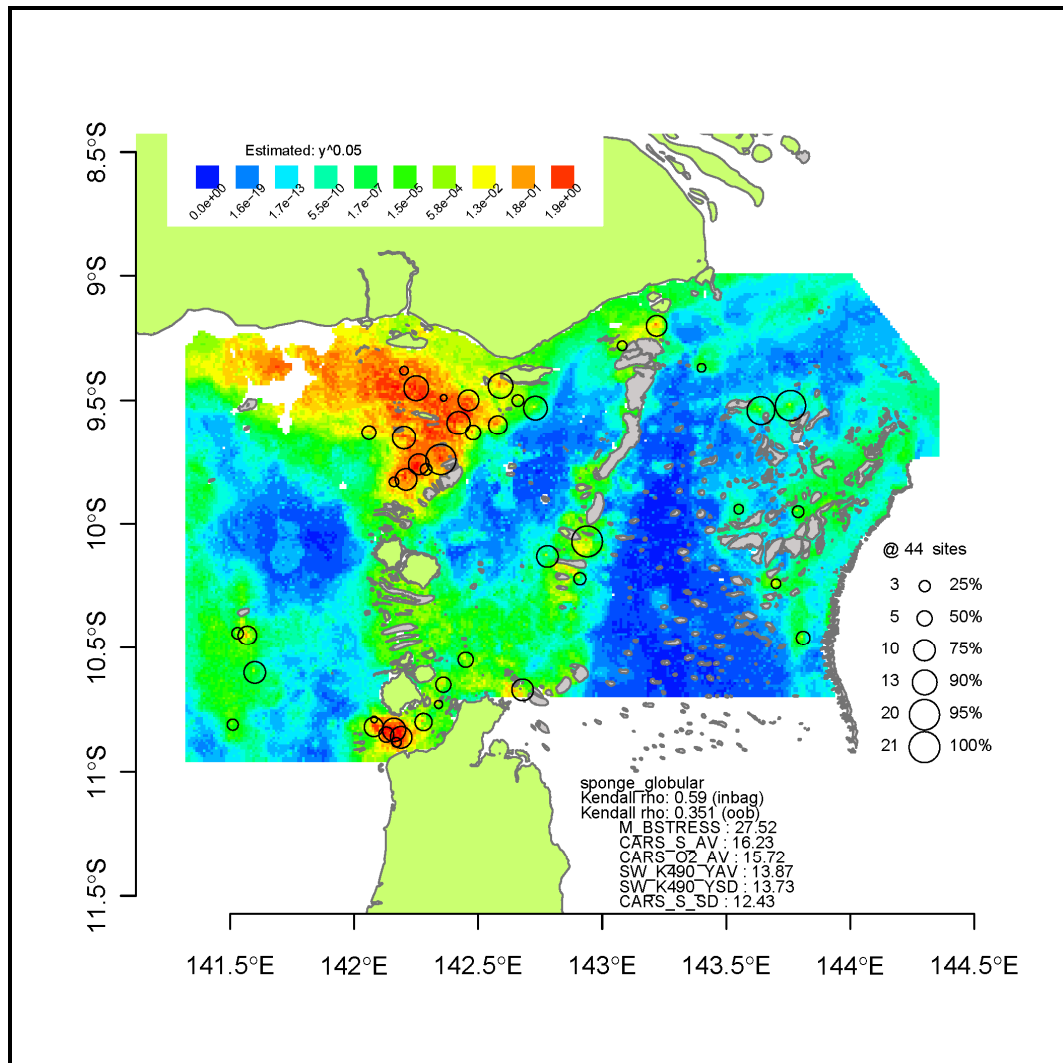


Figure 14. Estimated spatial distribution (percent cover) of “sponge (globular)”.

The “sponge (globular)” cover (Figure 14), another relatively common type (44 site observations), was predicted quite well from the covariates (criterion of 0.35). However in this case, it was the presence / absence of the cover type at a site that was predictable, rather the proportion of cover observed. Seagrasses with a “strap-like” form (Figure 15) were an example of a relatively rarer category with cover proportion nevertheless well estimated from the covariates. The distribution of this category of seagrasses apparently has a well localised distribution, both spatially, and in terms of habitat preference. By way of contrast, we include the map of the rare cover type “sponge digitate” (Figure 16) which was characterised by a poor model fit. We reiterate our cautions for interpreting distribution maps for models with lower validation criteria.

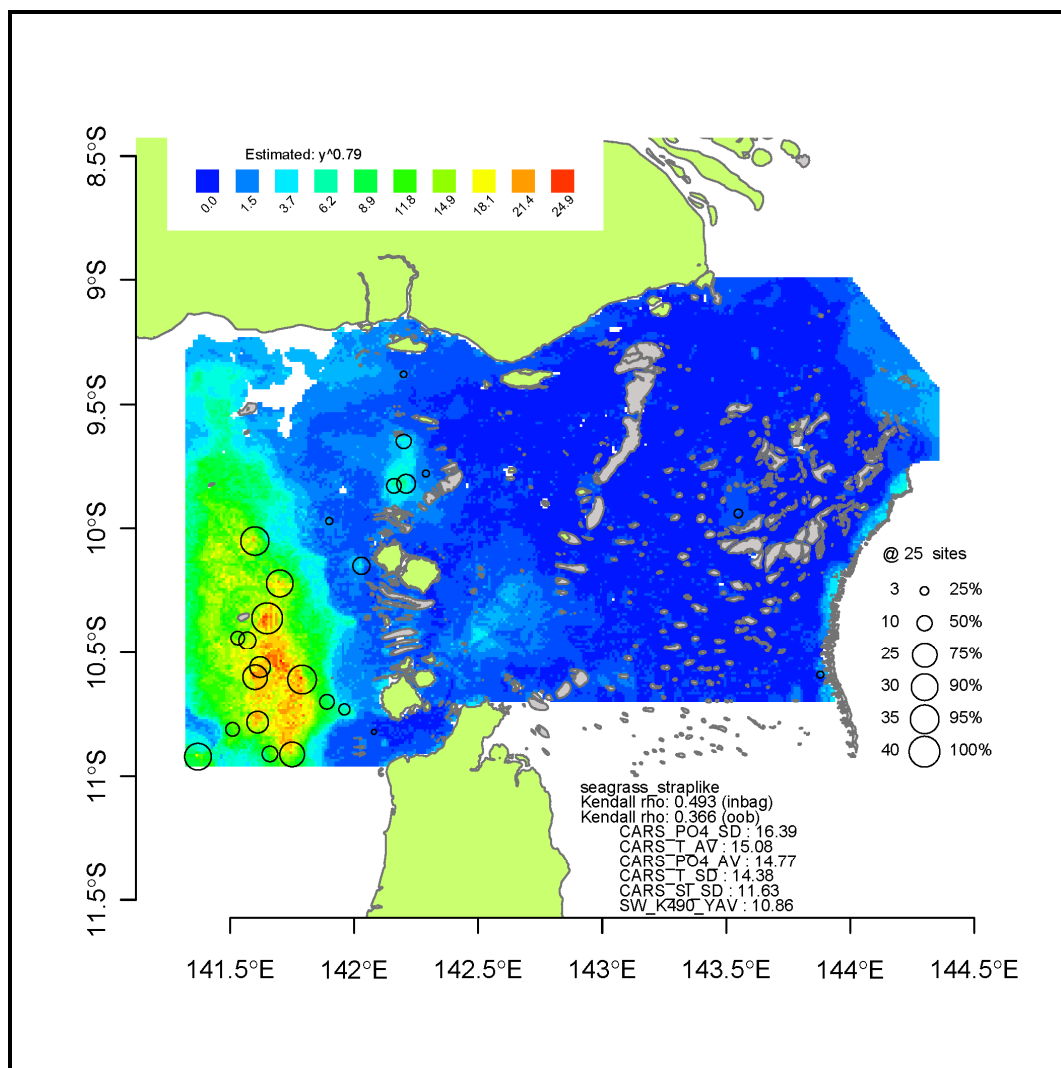


Figure 15. Estimated spatial distribution (percent cover) of “seagrass strap-like”.

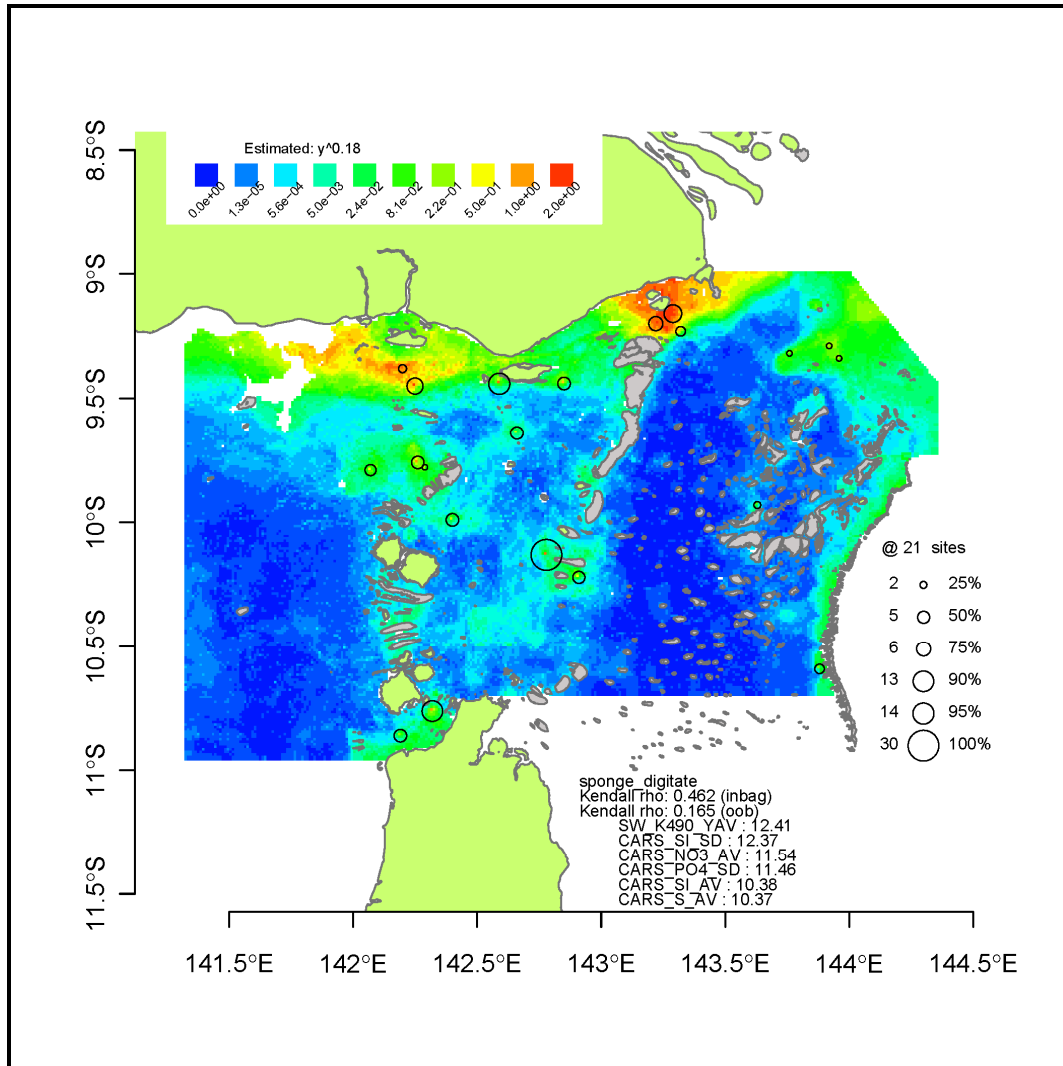


Figure 16. Estimated spatial distribution (percent cover) of “sponge digitate”.

3.3.3. Comparison of video biota with that from sled and trawl

Half of the 216 sites visited were sampled with all three devices (Figure 17). Some sites were not sampled because of gear breakdown, but many (17) were not able to be sampled with the trawl or sled because of the presence reef or rocks which may have fouled the sampling gear. Most of the sites that were sampled by video only were to the west of the Orman Reefs while the remainder were spread out across most of the study area: to the southwest of Badu, in the Endeavour Strait, around Darnley Island and to the south of the Hibernia Passage.

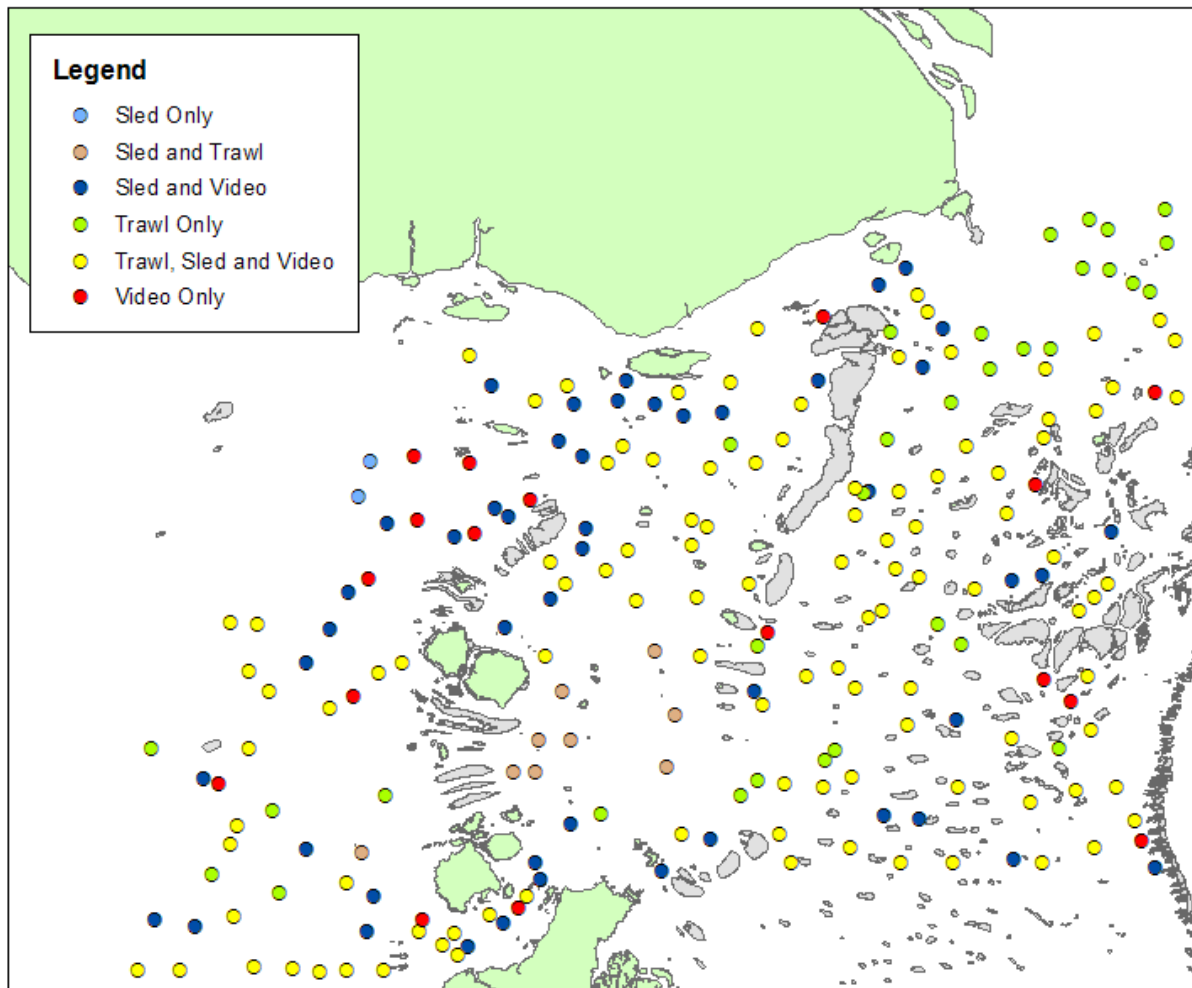


Figure 17. Torres Strait: Map showing sites sampled with various gears during the CRC Torres Strait project “Mapping and physical characterisation of key biotic and physical attributes of the Torres Strait Ecosystem” (Pitcher *et al.* 2007).

During the (Pitcher *et al.* 2007) study the catches from the epibenthic sled and trawl were merged to give a total of diverse dataset comprised of 2,372 Operational Taxonomic Units (OTUs) consisting of sponges, molluscs, crustaceans, coral, bryozoans, echinoderms, fishes and algae. Of these 258 OTUs were present in sufficient samples to enable modeling of their spatial distribution and abundance. In our analysis of the towed video collected during that field study we were able to distinguish only 62 distinct OTUs for the count analysis and 30 OTUs for the percent cover analysis that were observed often enough to enable distribution modeling. In addition to this, the biota collected from the sled and trawl were generally able

to be identified to a much higher degree than those examined on the video (mean percentage of videoed taxa identified to genus per site = 28.4%; mean percentage of sled/trawl-caught taxa identified to genus per site = 80.7%). This makes it difficult to conduct a formal comparison of the information gained by the towed video with that from the sled and trawl. However, in a tentative attempt to map the taxa identified from the video against those present in the sled/trawl catches, two key points become apparent. Firstly, the video appears to be more efficient at capturing information on some of the larger, sessile benthic taxa like corals and gorgonians (Anthozoa) and sponges (Demospongia; Table 4). Smaller, mobile and cryptic biota such as crustaceans, bivalves and gastropods are not well documented using the video compared to the extractive devices.

Table 4. Numbers of individuals in each broad taxonomic category sampled throughout Torres Strait with the trawl or video in the (Pitcher et al. 2007) study compared to those observed on the towed video.

Class	Sled/trawl	Video
Actinopterygii	5893	4084
Annelida	61	0
Anthozoa	1179	19533
Ascidacea	416	799
Asteroidea	845	158
Bivalvia	1027	8
Bryopsidophyceae	537	0
Bryozoa	809	139
Calcarea	45	0
Cephalopoda	457	8
Chlorophyceae	20	0
Chlorophyta	90	0
Chondrichthyes	75	9
Cladophorophyceae	5	0
Crinoidea	269	9453
Crustacea	3417	48
Ctenophora	0	2
Cubozoa	1	0
Demospongiae	3837	7262
Echinoidea	667	454
Florideophyceae	106	0
Gastropoda	870	24
Holothuroidea	380	76
Hydrozoa	411	1378
Inarticulata	0	1
Mammalia	0	4
Mollusca	10	0
Nemertea	72	0
Ophiuroidea	530	32
Phaeophyceae	250	0

Class	Sled/trawl	Video
Phaeophyta	41	0
Plantae	16	0
Polychaeta	64	75
Reptilia	3	0
Rhodophyceae	54	0
Rhodophyta	15	0
Scyphozoa	9	0
Sipunculida	1	0
Thaliacea	0	3
Ulvophyceae	3	0

Secondly, the many broad categories (e.g. *Actinopterygii_bonyfish_sp*) used in the video analysis may have many possible equivalents in the sled/trawl dataset (Table 5). In some cases biota observed on the video were able to be identified to species level and therefore there was a one-to-one match between the OTUs from the different devices (e.g. *lanthella flabelliformis*, *Sphenopus marsupalis* and *Halophila ovalis*), but this was not commonly the case. In rare cases, such as the gorgonian whips (*Junceela fragilis* and *J. juncea*) biota were able to be identified to species from video imagery, but were never caught in the sled or trawl. Presumably animals such as these having a flexible stem and a firm attachment to the seabed are not sampled effectively with devices like the epibenthic sled or trawl.

Table 5. Comparison of all the OTUs identified in sufficient numbers to enable distribution modelling from identified from the video with their possible matches identified from the epibenthic sled and trawl from the (Pitcher et al. 2007).

Video analysis type	Video OTU	Possible equivalent sled/trawl caught OTUs
Count	Actinopterygii_bonyfish_sp	80 species of fish were identified in sufficient numbers to model
	Actinopterygii_burrowingfish_burrowingfish	
	Actinopterygii_callionymid_sp	<i>Repomucenus belcheri</i>
		<i>Repomucenus sublaevis</i> (Fem)
		<i>Calliurichthys grossi</i>
		<i>Dactylopus dactylopus</i>
	Actinopterygii_solesflounders_sp	<i>Grammatobothus polyophthalmus</i>
		<i>Cynoglossus maculipinnis</i>
		<i>Pseudorhombus elevatus</i>
		<i>Engyprosopon grandisquama</i>
		<i>Pseudorhombus spinosus</i>
		<i>Pseudorhombus diplospilus</i>
		<i>Zebrias craticula</i>
		<i>Cynoglossus</i> sp 1 punctate
		<i>Pseudorhombus argus</i>
	Anthozoa_alcyoniid_sp	<i>Alertigorgia orientalis</i>
		<i>Carijoa</i> sp1
		<i>Dichotella</i> sp1

Video analysis type	Video OTU	Possible equivalent sled/trawl caught OTUs
		<i>Dendronephthya</i> spp
		<i>Subergorgia suberosa</i>
	Anthozoa_branching_sp	<i>Subergorgia suberosa</i>
	Anthozoa_bushy_sp	<i>Dendronephthya</i> spp
	Anthozoa_Carijoasp_sp	
	Anthozoa_Cirripathes_sp	
	Anthozoa_Ctenocella_sp	
	Anthozoa_Dichotella_sp	
	Anthozoa_gorgoniid_sp	<i>Subergorgia suberosa</i>
	Anthozoa_Junceella_fragilis	
	Anthozoa_Junceella_junceae	
	Anthozoa_Nephtheid_sp	<i>Dendronephthya</i> spp
	Anthozoa_Porites_sp	
	Anthozoa_Pteroides_sp	
	Anthozoa_Sarcophyton_sp	
	Anthozoa_scleractinian_encrusting	
	Anthozoa_scleractinian_massive	
	Anthozoa_seafan_sp	<i>Subergorgia suberosa</i>
		<i>Alertigorgia orientalis</i>
		<i>Carijoa</i> sp1
		<i>Dichotella</i> sp1
	Anthozoa_Sinularia_sp	
	Anthozoa_softcoral_bushy	
	Anthozoa_Solenocaulon_spbare	
	Anthozoa_Solenocaulon_spencrusting	
	Anthozoa_Sphenopus_marsupialis	<i>Sphenopus marsupalis</i>
	Anthozoa_Subergorgia_sp	<i>Subergorgia suberosa</i>
	Anthozoa_Turbinaria_sp	<i>Turbinaria</i> spp/ colonies massive columnar laminar
	Anthozoa_Turbinaria_sp_map	
	Anthozoa_Turbinaria_sp_pe	
	Anthozoa_Virgularia_sp	
	Ascidacea_ascidian_encrusting	
	Ascidacea_ascidian_pink	
	Ascidacea_ascidian_solitary	
	Ascidacea_Hypodistoma_deerratum	
	Asteroidea_Iconaster_longimanus	<i>Iconaster longimanus</i>
	Asteroidea_seastar_sp	<i>Astropecten zebra</i>
		<i>Luidia maculata</i>
		<i>Metrodora subulata</i>
		<i>Asterinidae</i> sp2
		<i>Stellaster equestris</i> cf

Video analysis type	Video OTU	Possible equivalent sled/trawl caught OTUs
		<i>Pentacaster gracilis</i>
		Goniasteridae sp5
		<i>Goniodiscaster rugosus</i> cf
		<i>Anthenea</i> sp1
		<i>Euretaster insignis</i>
	Crinoidea_crinoid_sp	unidentified crinoids
	Demospongiae_Cymbastela_coralliophila	
	Demospongiae_lanthella_basta	
	Demospongiae_lanthella_flabelliformis	<i>lanthella flabelliformis</i>
	Demospongiae_lanthella_sp	<i>lanthella quadrangulata</i>
		<i>lanthella flabelliformis</i>
	Demospongiae_Ircinia_sp	<i>Ircinia</i> 1255
	Demospongiae_Oceanapia_sp	
	Demospongiae_Pseudoceratina_sp	
	Echinoidea_urchin_sp	<i>Prionocidaris bispinosa</i>
		<i>Peronella lesueuri</i>
		<i>Peronella orbicularis</i> cf
		<i>Brissopsis luzonica</i>
		<i>Lovenia elongata</i>
		<i>Breynia australasiae</i>
		<i>Salmacis belli</i>
		<i>Temnotrema bothryoides</i>
		<i>Nudechinus</i> spp
		<i>Temnopleurus alexandri</i>
	Holothuroidea_Holothuria_sp	<i>Holothuria ocellata</i>
	Hydrozoa_hydroid_feathery	
	Hydrozoa_hydroid_linearclump	
	Malacostraca_crab_sp	
	sponge_sponge_barrel	
	sponge_sponge_branching	
	sponge_sponge_bushy	
	sponge_sponge_club	
	sponge_sponge_digitate	
	sponge_sponge_encrusting	
	sponge_sponge_erect	
	sponge_sponge_fan	
	sponge_sponge_foliose	
	sponge_sponge_globular	
	sponge_sponge_lobate	
	sponge_sponge_vase	
	sponge_sponge_wrinklyfan	
Percent cover	alcyonarian_other_bushy	<i>Dendronephthya</i> spp

Video analysis type	Video OTU	Possible equivalent sled/trawl caught OTUs
	alcyonarian_scleraxonian_branching	<i>Subergorgia suberosa</i>
		<i>Alertigorgia orientalis</i>
		<i>Carijoa</i> sp1
		<i>Dichotella</i> sp1
	alcyonarian_scleraxonian_Solenocaulonencrusted	
	alcyonarian_scleraxonian_whip	
	algae_branching	
	algae_bushy	
	algae_Caulerpa	<i>Caulerpa</i> sp1
		<i>Caulerpa racemosa</i>
		<i>Caulerpa taxifolia</i>
	algae_filamentous	
	algae_Halimeda	<i>Halimeda discoidea</i>
		<i>Halimeda opuntia</i>
		<i>Halimeda cylindracea</i>
	algae_Sargassum	<i>Sargassum flavicans</i>
		<i>Sargassum decurrens</i>
		<i>Sargassum</i> sp12
		<i>Sargassum spinifex</i>
	bioturbation_burrows	
	bioturbation_mounds	
	bioturbation_pits	
	flora_erect	<i>Caulerpa</i> sp1
		<i>Caulerpa racemosa</i>
		<i>Caulerpa taxifolia</i>
		<i>Halimeda discoidea</i>
		<i>Halimeda opuntia</i>
		<i>Halimeda cylindracea</i>
		<i>Sargassum flavicans</i>
		<i>Sargassum decurrens</i>
		<i>Sargassum</i> sp12
		<i>Sargassum spinifex</i>
	flora_prostrate	
	hydroid_feathery	
	seagrass_compound	<i>Halophila spinulosa</i>
	seagrass_ovoid	<i>Halophila ovalis</i>
	seagrass_straplike	<i>Cymodocea serrulata</i>
	sponge_branching	
	sponge_digitate	
	sponge_encrusting	
	sponge_fan	
	sponge_globular	
	UFO_branching	
	UFO_bushy	

Video analysis type	Video OTU	Possible equivalent sled/trawl caught OTUs
	UFO_digitate	
	UFO_encrusting	
	UFO_globular	<i>Oceanapia</i>

For the few taxa that were identified to species in both the video and the catches from the sled and trawl it is interesting to compare the predicted distributions obtained. It must be kept in mind that the data used from the sled and trawls were biomass compared to the video data which was expressed as counts of individuals or individual colonies.

Sphenopus marsupalis

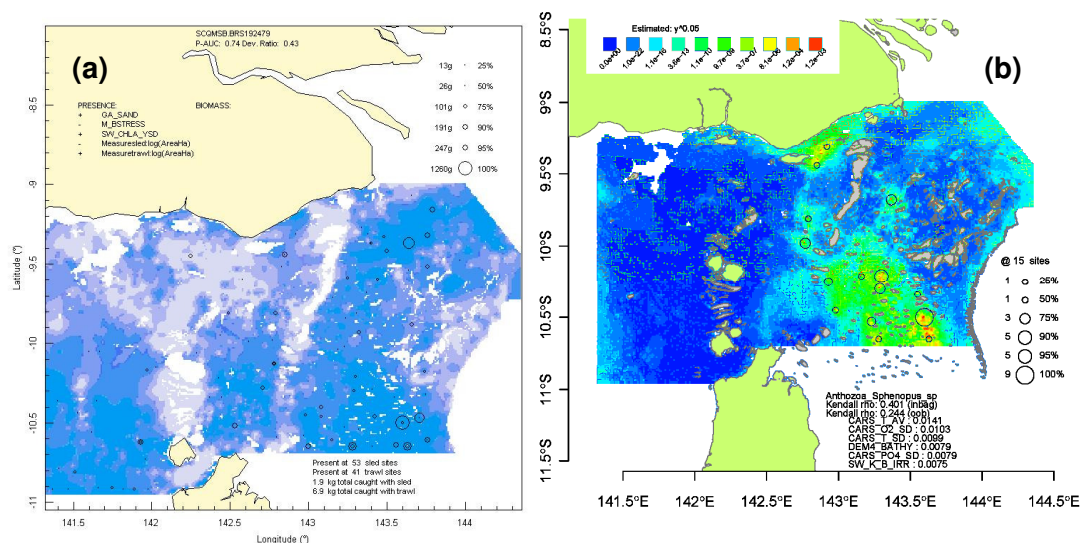


Figure 18. Predicted distributions of the anthozoan, *Sphenopus marsupalis* generated from (a) catches obtained using sled and trawl (Pitcher et al. 2007); and (b) towed underwater video count data.

In the case of the anthozoan, *Sphenopus marsupalis* catches from the sled and trawl indicated a fairly widespread distribution although dominating in the east. The distribution generated from the video data confirmed the higher numbers in the east, but they were not observed at all in the western Torres Strait (Figure 18). *S. marsupalis* are quite small (generally < 5 cm across) and are often coloured similar to the surrounding sediment. Given that visibility in the western Torres Strait is often poor it is likely that they could not be seen on many of the video transects although they were obviously present since they were caught in the trawl and or sled.

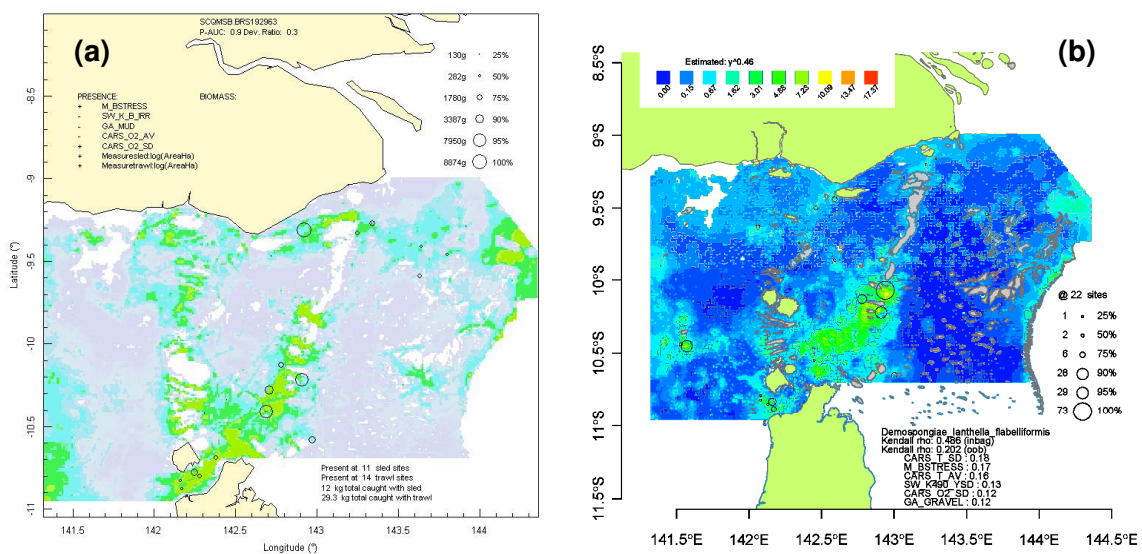
lanthella flabiformis

Figure 19. Predicted distributions of the sponge, *lanthella flabiformis* generated from (a) catches obtained using sled and trawl (Pitcher et al. 2007) and (b) towed underwater video count data.

In contrast, the sponge *lanthella flabiformis* grows to be up to 60-70 cm across and is quite conspicuous underwater as it grows erect from the seabed. The predicted and observed distributions generated from both methods seem to be in reasonable agreement (Figure 19) with the majority occurring around the reefs to the south of the Warrior Reef complex and the Endeavour Passage.

Biodiversity data collected using the video obviously has some limitations in the sense that many taxa are not able to be identified as comprehensively as those collected with extractive devices. Also it has limitations when observing under poor visibility and in recording mobile, cryptic or small biota. Never-the-less, in areas that are too rough to sample with sled or trawl or for some taxa that are not well sampled by sled or trawl (e.g. the gorgonian whips *Junceella fragilis* and *J. juncea*) the video complements data collected with the other devices. Ideally all three devices should be used together in order to gain the most comprehensive picture of seabed biodiversity.

4. Coral Reef Characterisation

4.1. Introduction

The eastern border of Torres Strait is marked by the edge of the continental shelf where the ribbon reefs of the Great Barrier Reef extend along a north-south orientation. The central region is characterised by a broad ridge of reefs and islands, extending north from Cape York Peninsula. To the west there are fewer islands and reefs; much of the seabed in the northwest is uncharted and characterised by shallow shoals and strong tidal currents which have created sand waves. The waters throughout the strait are generally shallow (< 20 m) although the bathymetry is complex with numerous reefs and shoals and some deeper (up to 60 m) channels caused by tidal scouring (Harris 2001; Harris *et al.* 2005). There are many reefs throughout the Torres Strait (1,295 individual reefs and about 283 reef complexes) ranging in size from 0.2 ha to 16,500 ha. These reefs cover a total area of 2,734 km² and constitute a vital part of the region's biodiversity and support a number of important local fisheries, including sea cucumber, trochus, finfish and tropical rock lobster. The largest reef in Torres Strait is the Warrior reef complex. Located in the central-north part of the strait close to Papua New Guinea, it is approximately 70 km long and averages 7 km wide. Almost all of the 275 islands within the Strait have developed reef flats surrounding them.

During the last glacial maximum (20,000 years ago) Torres Strait was above sea level (Chappell 1983) and remained so until about 7000 years ago (Hopley 1983). Surface and subsurface geologic investigations and radioactive carbon dating suggest that reef formation began at 6 m depth at this time (Woodroffe *et al.* 2000). The corals grew until reaching the present sea level about 5,000 years ago and then extended laterally (in many cases aligned east-west) in response to the dominant current direction (Jones 1995). Many present day platform reefs in Torres Strait grew on the remnants of these Pleistocene coral reefs (Heap *et al.* 2004). In contrast, most of the fringing reefs around the granitic islands have formed over a granite substrate e.g. the reefs around Hammond and Yam Islands (Woodroffe *et al.* 2000). In other areas where there are high levels of mud, the fringing reefs seem to have formed as a thin veneer overlaying a mud wedge (Woodroffe *et al.* 2000).

Currents are one of the dominant forces responsible for determining the nature of the reefs of Torres Strait. At its narrowest point, the strait is only 150 km wide and the tidal regimes on either side of the strait (Gulf of Papua/Coral Sea and the Gulf of Carpentaria/Arafura Sea) are out of phase with each other (Amin 1978). This results in a pressure gradient across the strait (Bode and Mason 1994) with tides of up to 3 m and current speeds of up to 5 knt in the central and western parts of the strait where velocities are accelerated between the islands and reefs. Wind induced currents (particularly during the south east trade season) combine with the tidally-induced currents to produce turbid water in the central Torres Strait (Harris and Baker 1991).

Sedimentation is also a factor determining the formation and continued survival of coral reefs. The Fly River in Papua New Guinea drains to the northeast of Torres Strait, exporting an estimated 10 million tonnes of suspended sediment per annum (Harris *et al.* 1993). While most of this water disperses to the east, away from Torres Strait (Robertson *et al.* 1993; Hemer *et al.* 2004), occasional intrusions of turbid Fly River water have been observed extending into the northern parts of Torres Strait (Ayukai and Wolanski 1997). Sea level differences between the east and western sides of Torres Strait generate occasional westward currents and may be responsible for this westerly movement, especially when the south-easterly trade winds subside (Ayukai and Wolanski 1997). Strong tides and storm events are responsible for significant increases in the suspended sediment load in the Strait and the high levels may persist for up to three weeks or more (Margvelashvili and Saint-Cast 2007).

It is important to study the coral reefs of Torres Strait because of their unique nature, high biodiversity and commercial values for tourism and indigenous fisheries. Coral reefs in Torres Strait may be threatened by a range of anthropogenic influences, including: overfishing, increased sediment input from logging and mining operations in Papua New Guinea and climate change. In this study we plan to characterise and classify the reefs of Torres Strait and to relate the nature of the reefs to their surrounding physical environment.

4.2. Methodology

4.2.1. Basemap

A basic requirement for this project was a comprehensive map of reef outlines for the region. This was done by augmenting an existing map of the reefs and islands of Torres Strait digitized from AUSLIG 1:100,000 and 1:250,000 topographic map series (Long et al. 1995) with reef outlines from classified Landsat Thematic Mapper imagery. The following images were used: Scene path 99, Row 67 acquired 10:00 01/08/1993, Scene path 98, Row 66 acquired 10:00 31/10/1998, Scene path 98, Row 67 acquired 10:00 11/11/1992 and Scene path 98, Row 68 acquired 10:00 27/04/1990.

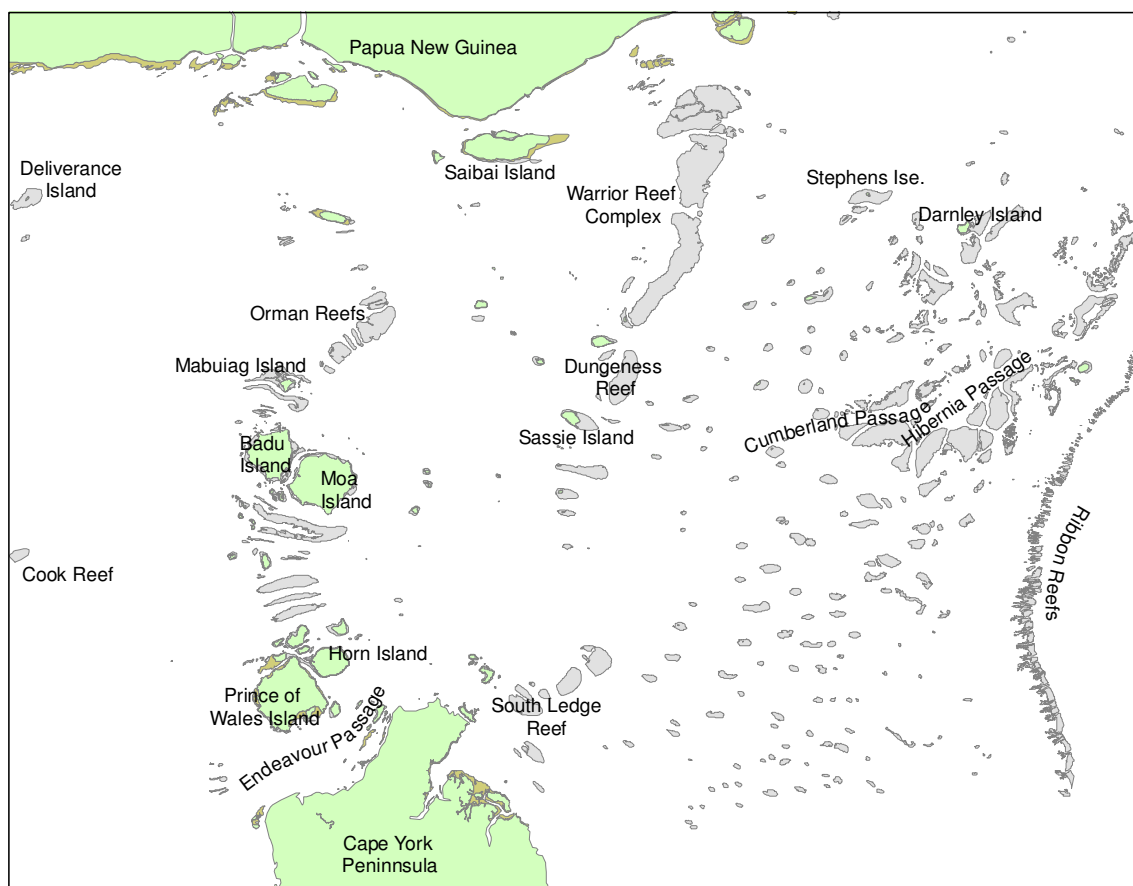


Figure 20. Map of the Torres Strait showing reefs digitised from AUSLIG 1:100, 000 and 1:250, 000 topographic map series with Landsat Thematic Mapper imagery.

4.2.2. Reef datasets

CMAR has conducted a number of surveys on the reefs of Torres Strait over the past 12 years (Table 6). These surveys have been conducted for a variety of reasons including sea cucumber population estimation and reef resource inventories. However, in addition to counts of the particular target species being made during each survey, visual estimates of percent cover of various substrate and macrobenthic components have also been recorded.

Table 6. CMAR surveys on the reefs of Torres Strait since 1995.

Survey name	Date	Coverage	Sites	Survey components
Reef Resource Inventory (RRI) over 4 surveys (HE01, HE02, JK01, JK02) (Long et al. 1997)	1995/96	Full coverage of Torres Strait reefs, including PNG	1510 sites: • 1200 reef top; and • 310 reef edge.	• Reef macroinvertebrate density (~30 species) • Visual estimates of % cover of substrate components (6 categories) • Visual estimates of % cover of macrobenthic components (~20 species, ~20 genus or higher categories) • Finfish UVC (100+ species)
East Torres Strait lobster breeding ground survey (Pitcher et al. 1997)	1996	Darnley, Murray and outer Barrier Reef	139 sites	• Reef top and reef edge (to 15 m) • Reef macroinvertebrate density (~40 species) • Visual estimates of % cover of substrate components (6 categories) • Visual estimates of % cover of macrobenthic components (~20 species, ~20 genus or higher categories)
2002 East Torres Strait sea cucumber survey (Skewes et al. 2003)	2002	East of Warrior Reef to the outer Barrier Reef	424 sites	• Reef top and reef edge (to 20 m) • Reef macroinvertebrate density (~40 species) • Visual estimates of % cover of substrate components (10 categories) • Visual estimates of % cover of macrobenthic components (~20 species, ~20 genus or higher categories)
2005 East Torres Strait sea cucumber survey (Skewes et al. 2006)	2005	East of Warrior Reef to the outer Barrier Reef	123 sites (subset of above)	As above
Warrior Reef monitoring surveys (Skewes et al. 2006)	1995/96, 1998, 2000, 2002, 2004	Warrior Reef reeftop	49 sites (additional sites in some years)	• Reef macroinvertebrate density (~40 species) • Visual estimates of % cover of substrate components (10 categories) • Visual estimates of % cover of macrobenthic components (~20 species, ~20 genus or higher categories)

4.2.3. Survey methods

The methods used to collect the reef data during the surveys varied between surveys to some extent; however the general approach was as follows. The reef map used for the sample design was as explained above. It was generally stratified into zones based on available catch information, likely abundance of species of interest (in most cases, holothurians) and physiographic characteristics of the fishery habitats. These zones formed the basis for the sample design and stratified analysis. Each zone was further divided into at least two habitat strata; the reef edge and the reef top. During later surveys, a third strata was identified, the reef top buffer stratum, being a 200 m wide buffer around the inside of the reef margin. The reef top buffer was identified as an area likely to contain higher quantities of targeted species of sea cucumber.

4.2.4. Sample design

The sampling density was varied in each stratum due to predicted abundance and variance estimates, sampling logistics and time constraints. The density of sampling on the reef top stratum was generally about 1 site per 2 km², and the reef top buffer stratum was about 1 site every 1 km² (Figure 21). With an equivalent effort spent on the reef edge, this meant a sampling density of about one site every 4 km around the reef edge (Figure 21). This resulted in sampling effort, in terms of the number of sites sampled, being roughly divided 60:40 between reef top and reef edge habitats. This allocation of sampling effort was based on the experiences of previous reef resource surveys and has been shown to be the best compromise for efficient sampling of reef resources, while allowing for habitat mapping of reef top areas.

Survey sites were assigned by dividing each stratum on the reef top into 1 km² or 2 km² grids (depending on the strata type), and the reef edge into sections 4 km long. Sample sites were then located within the grids/sections at random. For the reef top, this meant that each sample site was randomly allocated within each grid. For the reef edge, the sample sites were selected at random from points spaced 150 m along the edge within each of the sections.

Generally, the sampling effort was randomly allocated to each habitat stratum within each zone in a representative way, except in cases where the sample areas were too spatially dispersed, such as the Great North East Channel and south eastern Torres Strait. In these areas, we allocated samples randomly within an area that could be sampled by the two teams of divers within a day. This means that this area in particular is sampled in a clumped fashion. While this limits some of the conclusions drawn from the data in these areas, we feel that the results are probably representative of the zone in which the sampling occurred. Generally, where areas were sampled at a different density, they were treated separately in the analysis, nullifying the potential bias that this could cause.

4.2.5. Field sampling

Field work was undertaken by two teams of divers operating from dinghies and survey sites were located using a portable GPS. On the reef top, a diver swam along a 40 m transect and recorded resource and habitat information 2 m either side of the transect line. Holothurians and other benthic fauna of commercial or ecological interest were counted, and where possible, returned to the dinghy and measured (weight and length where possible). At each site, the substrate was described in terms of the percentage of sand, rubble, consolidated rubble, pavement and live coral. The growth forms and dominant taxa of the live coral component and the percentage cover of all other conspicuous biota such as seagrass and algae were also recorded.

On the reef edge, at each site two divers swam adjacent transects perpendicular to the reef edge from the reef crest to a depth of 20 m or a distance of 100 m, whichever came first. One diver counted species of interest and recorded habitat information using the same protocol as the reef top, and the second diver counted species of interest and collected them for measuring back in the dinghy.

During the fish surveys a SCUBA diver swam approximately 100 m perpendicular to the reef crest at about mid depth (usually between 3 to 7 m) counting and identifying fish 5 m on either side of the transect. Counts were made of all a suite of 197 species which had been surveyed by AIMS during regular surveys on the northern GBR for over 15 years (Williams 1982b).

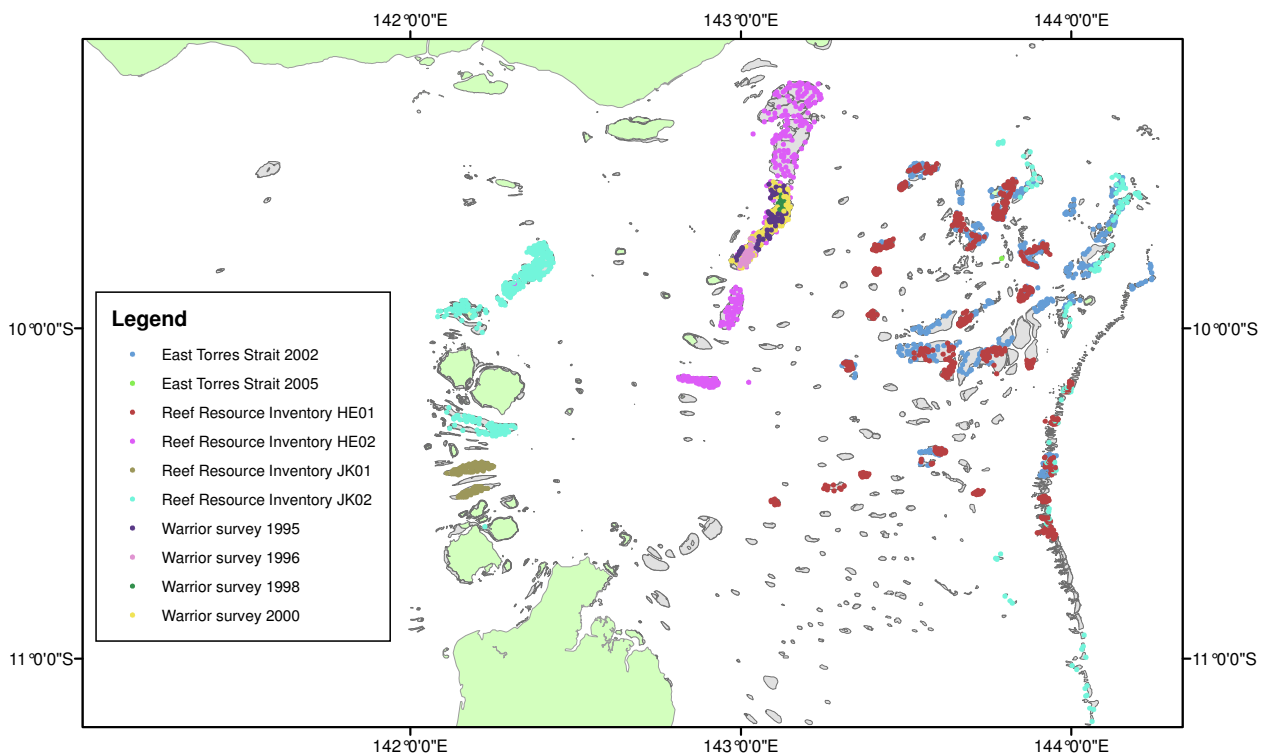


Figure 21. Map of Torres Strait showing reef survey sites by cruise.

The data from the various surveys were resolved to a common denominator format and collated into a single Oracle database (Figure 22).

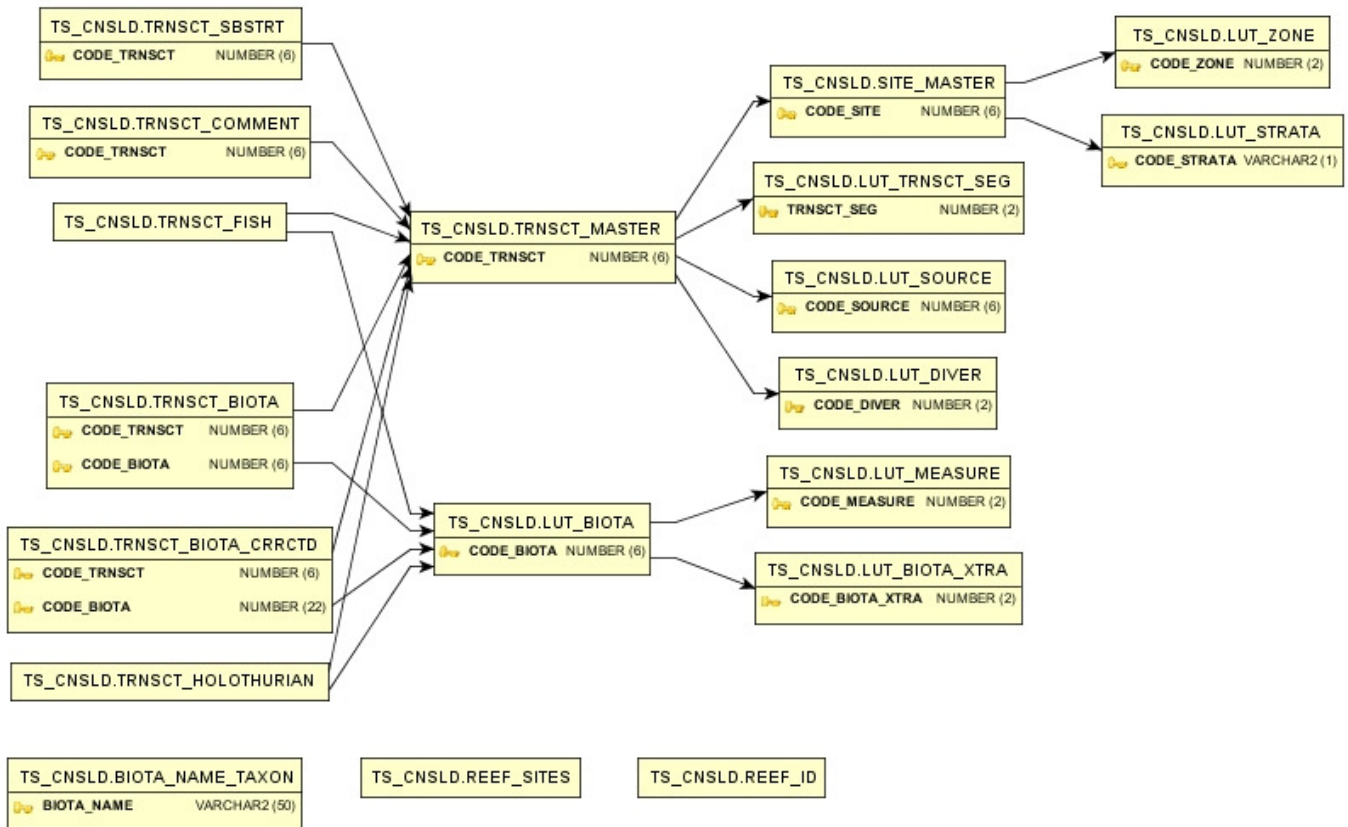


Figure 22. Schematic diagram of the database structure for the biological survey data.

The database included information on: site latitude and longitude; date and time surveyed; depth; visibility; transect length and width; percentage cover of dominant substrate types – soft substrate, hard substrate, gravel, rubble, consolidated rubble, boulders or platform; percentage cover of algae, seagrass, total live, dead standing and soft corals; percentage cover of the different live coral morphotypes (branching, sub-massive, digitate, massive, foliose, columnar, corymbose, tabulate, encrusting and fungids); counts of reef macrofauna such as clams, lobsters, holothurians etc.; for the surveys where underwater visual censusing for fish were undertaken, fish were identified to species (only species of a fixed suite of 197 species that had been monitored on the GBR for by AIMS for more than 15 years (Williams 1982a; Williams 1982b), counted and total length was estimated to the nearest 5 cm.

When combined, these data provided information from 2,280 sites on 385 reefs. Most reefs in the central Torres Strait and many of those in the northeast have been surveyed, but there are gaps in the southeast and many of the Ribbon Reefs on the outer GBR have not been surveyed (Figure 23). The majority of sites (1,946) have only been surveyed once; 334 sites have been surveyed more than once.

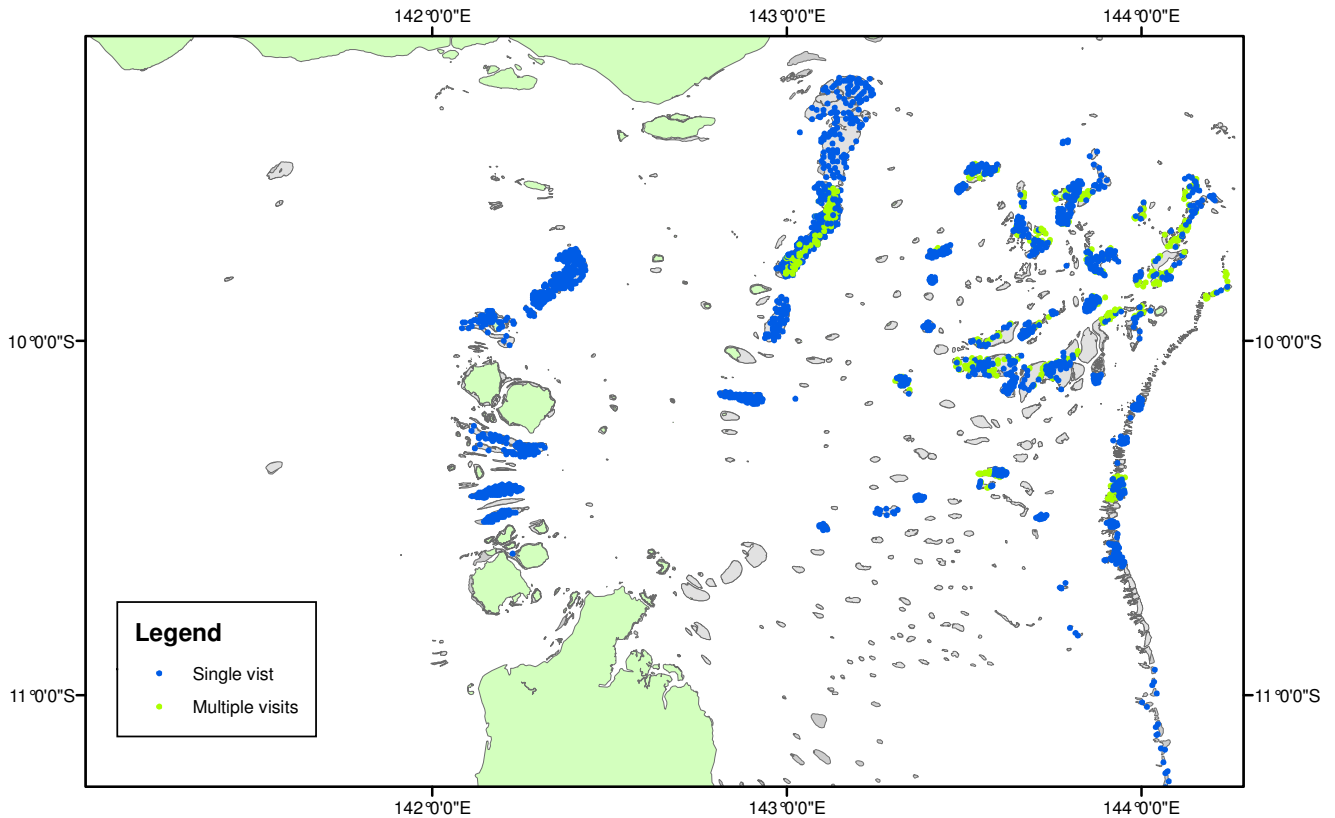


Figure 23. Map showing the positions of reef survey sites.

4.2.6. Cluster analysis

Analysis datasets

The collated reef dataset contained data for 2,551 samples from 2,196 sites on 192 reefs. Observed organisms were identified as species groups that corresponded to taxonomic groups (mostly species) or, less commonly, as morphotypes. The 335 species groups were in the broader groups: fish (196), ascidians (1), rock lobsters (4), corals (23), hydroids (1), sea cucumbers (25), starfishes (6), urchins (1), crinoids (1), clams (7), pearl oysters (2), trochus (2), sponges (7), algae (49) and seagrasses (10).

Because the reef data were collected for different purposes on different voyages, the species recorded varied among the collated datasets. To create complete sample by species matrices for multivariate analysis, subsets of samples and/or species were selected as appropriate for each analysis.

Species and habitat clustering

Four clustering analyses were performed: (a) clustering of samples by species with a dataset that included as many samples as possible, (b) clustering of samples from reef edges by as many species as possible, (c) clustering of samples from reef edges by habitat-forming species, and (d) clustering of samples by habitat attributes (substrate and habitat-forming species).

Analysis (a) aimed to cluster as many samples as possible. When selecting data to create a complete sample by species matrix, emphasis was therefore placed on retaining samples

rather than species. The 2,220 retained samples were from 1,887 sites on 177 reefs. All 10 datasets were included. The 37 retained species codes were for sea cucumbers (12), algae (10), seagrasses (9), *Panulirus* (rock lobsters), Fungiidae (mushroom corals), *Acanthaster planci* (crown of thorns starfish), Tridacnidae (clams), *Pinctada* (pearl oysters) and *Trochus niloticus* (trochus). The analysis was performed on presence/absence data.

Analysis (b) clustered samples with as many taxa as possible. Emphasis was therefore placed on retaining species rather than samples. The 229 retained samples were from 229 sites on 57 reefs. Only reef edges from the RRI datasets were included. The 224 retained species codes were for fish (185), sea cucumbers (8), coral morphotypes (8), algae (8), seagrasses (9), *Panulirus* (rock lobsters), Fungiidae (mushroom corals), *Acanthaster planci* (crown of thorns starfish), Tridacnidae (clams), *Pinctada* (pearl oysters) and *Trochus*. The analysis was performed on presence/absence data.

Analysis (c) clustered samples by habitat-forming species. The 560 samples were from 522 sites on 137 reefs. Only reef edge samples from the ETS and RRI datasets were included. The 28 species codes were for coral morphotypes (7), Fungiidae, algae (11) and seagrasses (9). The analysis was performed on quantitative data (percent cover).

Analysis (d) clustered samples by substrate and a subset of habitat-forming species cover. The 2,546 samples were from 2,193 sites on 191 reefs. All 10 datasets were included. The 10 percentage cover variables were for boulders, rubble, soft sediment, consolidated rubble, hard substrate, pavement, live hard coral, dead standing hard coral, algae and seagrass. The analysis was performed on quantitative data (percent cover).

Samples were clustered with the partitioning around medoids (PAM) method (Kaufman and Rousseeuw 1990), which is an unsupervised, non-hierarchical clustering method similar to the k-means method. The analysis was performed with the PAM package of the R computer program (Team 2007). The distance metric was the Jaccard index for presence/absence data and the Ružička index (the quantitative equivalent of the Jaccard index) for quantitative data. With the PAM method the number of clusters must be chosen in advance. Silhouette plots were therefore used to identify the best number near the readily interpretable number of five clusters. PAM clusters have no intrinsic nesting or ordering. For display purposes the clusters were sorted in the tree order of a hierarchical clustering of the cluster medoids.

Habitat-forming species and environmental factors

Relationships between habitat-forming species composition and environmental factors were examined with canonical correspondance analysis (CCA). The sample by species matrix was the same as that used for clustering (excluding nine samples for which environmental data were missing). The environment variables were a subset of the available variables described above (Table 7). Variables were excluded from the analysis if they were highly correlated with other variables or showed little variation within Torres Strait. Skewed variables were transformed with log or arcsine transformations. Standardisation was not necessary because CCA is not affected by linear transformations of the environment variables.

CCA was performed with the vegan package of the R computer program (R Development Team 2007). The first CCA was an inclusive search for environmental gradients related to species composition and included all candidate environment variables. Latitude and longitude were initially included in case they were surrogates for environment factors we had not considered, but on the condition that they would be preferentially dropped if shown to be reasonably redundant. Additional analyses were then performed to examine relationships between species composition and specific environment variables of interest.

CCA produces species scores and two types of sample scores: LC scores, which are linear combinations of the environment variables, and WA scores, which are weighted averages of the species scores (the weights for a sample are the abundances of the species in the sample). The key difference is that LC scores are calculated only from the environment variables, whereas the WA scores are calculated only from the species composition data. CCA maximises the correlation between the WA and LC scores. Correspondance between WA scores and LC scores indicates how well variation in species composition along the axis represents variation in the environment variables along the axis.

Table 7. Environment variables available and selected for canonical correspondence analysis. Details of the data sources are described above.

Variable	Description	Units	Source	Used
latitude	Latitude			✓
longitude	Longitude			✓
boulders	Boulders cover	%	Survey	✓
rubble	Rubble cover	%	Survey	✓
soft_sed	Soft sediment cover	%	Survey	✓
cons_rub	Consolidated rubble cover	%	Survey	✓
hard_sub	Hard substrate cover	%	Survey	✓
pavement	Pavement cover	%	Survey	✓
fetch_se	South-east fetch	m	Calculated	✓
fetch_nw	North-west fetch	m	Calculated	✓
max_depth	Maximum depth	m	Calculated	✓
current_stress	Current stress	N/m ²	Bode and Mason	✓
nitrogen	Surface nitrogen	µM	CARS	
oxygen	Surface oxygen	mL/L	CARS	
phosphate	Surface phosphate	µM	CARS	
silicate	Surface silicate	µM	CARS	✓
salinity	Surface salinity	PSU	CARS	✓
temp	Surface temperature	°C	CARS	✓
si_sd	Silicate seasonality	µM	CARS	
s_sd	Salinity seasonality	PSU	CARS	
t_sd	Temperature seasonality	°C	CARS	✓
chl_a_avg	Chlorophyll a	mg/m ³	SeaWIFS	
chl_a_sd	Chlorophyll a seasonality	mg/m ³	SeaWIFS	
k490_avg	Turbidity	/m	SeaWIFS	✓
k490_sd	Turbidity seasonality	/m	SeaWIFS	
wave_height	Wave height	m	WAM	✓
wave_dir_se	Wave direction south-easterly component	–	WAM	✓
wave_period	Wave period	s	WAM	✓

Fishes and environmental factors

Relationships between fish species composition and environmental factors were examined with CCA using the same methods and environmental variables as those used for habitat-forming species. The sample by species matrix was the same as that used for cluster analysis (b), except that only fish species were included. The main aim of the fish analysis was to indicate whether the environment variables correlated with habitat-forming species composition are also likely to be correlated with the species composition of other taxonomic groups.

4.2.7. Gap analysis

Our approach to conducting a gap analysis of our knowledge of the reefs of Torres Strait was threefold. The first was simply the spatial extent of areas that had not been surveyed at all; secondly we considered the rather limited range of data collected during the surveys; and, thirdly we attempted to document gaps identified through the biophysical analyses. The assumption made in this later analysis was that nature and biodiversity of the reefs was shaped primarily by environmental covariates and that by sampling the full range of covariate space we could be sure that we had comprehensively sampled the full range of coral reefs biotypes.

Previous analyses had identified the covariates that explained the highest degree of variation in the biological patterns among reefs (annual mean K490, annual mean surface salinity, annual surface salinity standard deviation, annual mean surface temperature, annual surface temperature standard deviation, wave period, wave height, wave direction and epibenthic current stress) and only these were included in the gap analysis. In order to reduce the number of variables in the analysis to make the computations more manageable and to ensure the clustering dataset was orthogonal, we performed a non-metric Multi Dimensional Scaling (nMDS) of the covariates prior to clustering.

A point coverage of a regular grid of 0.01° spacing was created in the GIS and intersected with the various covariate grids so that each point had a set of attributes corresponding to the underlying values of the covariates. All points overlying reef were then subsetting and a Euclidean distance matrix was generated with all covariate values standardized to have a mean of 0 and standard deviation of 1. The distance matrix was then used to generate a nMDS and the first 4 dimensions were clustered using K-means clustering. Eight clusters were chosen because examination of silhouette plots indicated this to be a reasonable number and it was readily interpretable.

The clusters were then loaded back into the GIS as points and a 0.01° degree grid was created using the cluster id as the grid value. The reef sites were then overlaid on the grid to determine the number of sites sampled within each cluster throughout the TS.

4.3. Results

4.3.1. Physical covariate datasets

The physical datasets used in this study were collated from a variety of internal and external sources. Many were used in the sample design for the Torres Strait mapping and characterisation of key biota project (Pitcher *et al.* 2004; Pitcher *et al.* 2007). Several of the internally held datasets were updated with recently acquired data and all were then mapped onto a 0.01° (approximately 1.11 km) grid and the survey sites were overlaid in order to append the various attributes.

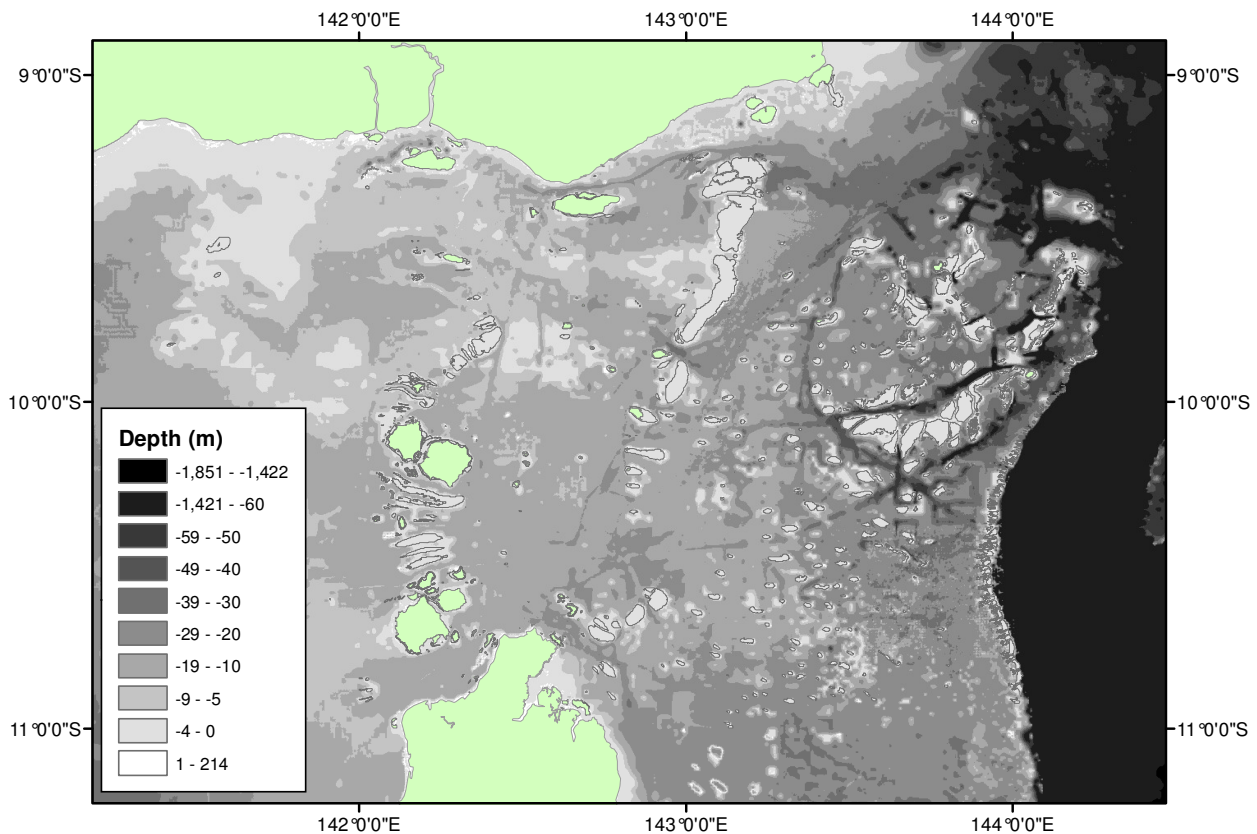


Figure 24. Depth (m) collated from digital soundings, satellite imagery and LADS data.

Bathymetry: This dataset includes data from CMAR (digital soundings and imagery) and GA sources and for this project was updated with new Laser Airborne Depth Sounder (LADS) data provided by the Australian Hydrographic Service (Figure 24). Although this data is at a very high spatial resolution, it was only available for a limited area in the southeastern Strait, along the eastern face of the Warrior Reef complex and to the west of Prince of Wales Island.

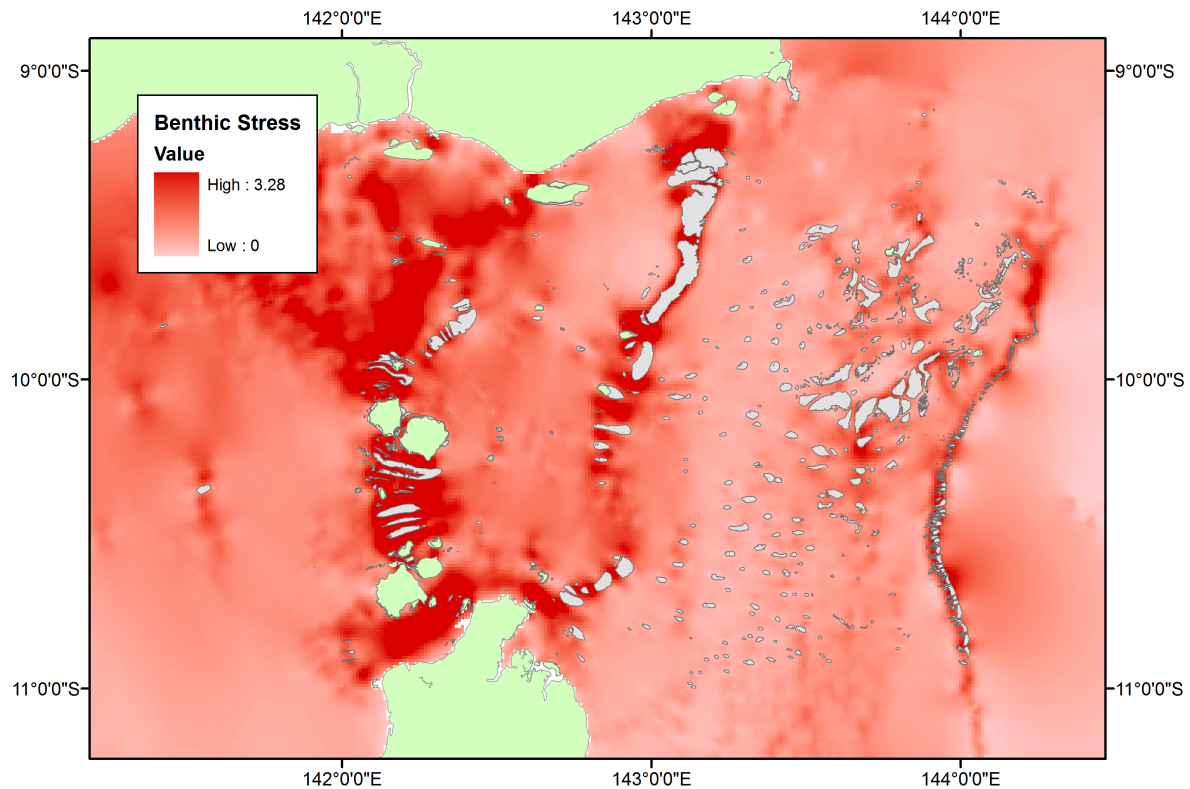


Figure 25. Modelled seabed current stress.

Seabed current stress: This dataset was provided by Bode and Mason (JCU / CRC Reef). The data is Root Mean Square (RMS) stress (Pascals (N.m^{-2})) output from a circulation model run over period of approximately 6 months. The modelled coverage is for the entire Torres Strait region, at 1 minute of arc resolution (~ 1.8 km), but is dependent on bathymetry data (which is incomplete) and other model assumptions.

Seabed current stress is defined as the frictional force imposed on water near the seabed. This force is important in redistributing sediments and may thereby influence the distribution of benthic biota. Currents have also been implicated in influencing the shape of reefs forming in the Torres Strait (Jones 1995). Generally the highest areas of current stress occur where the water is forced between narrow passes between islands or reefs (Figure 25).

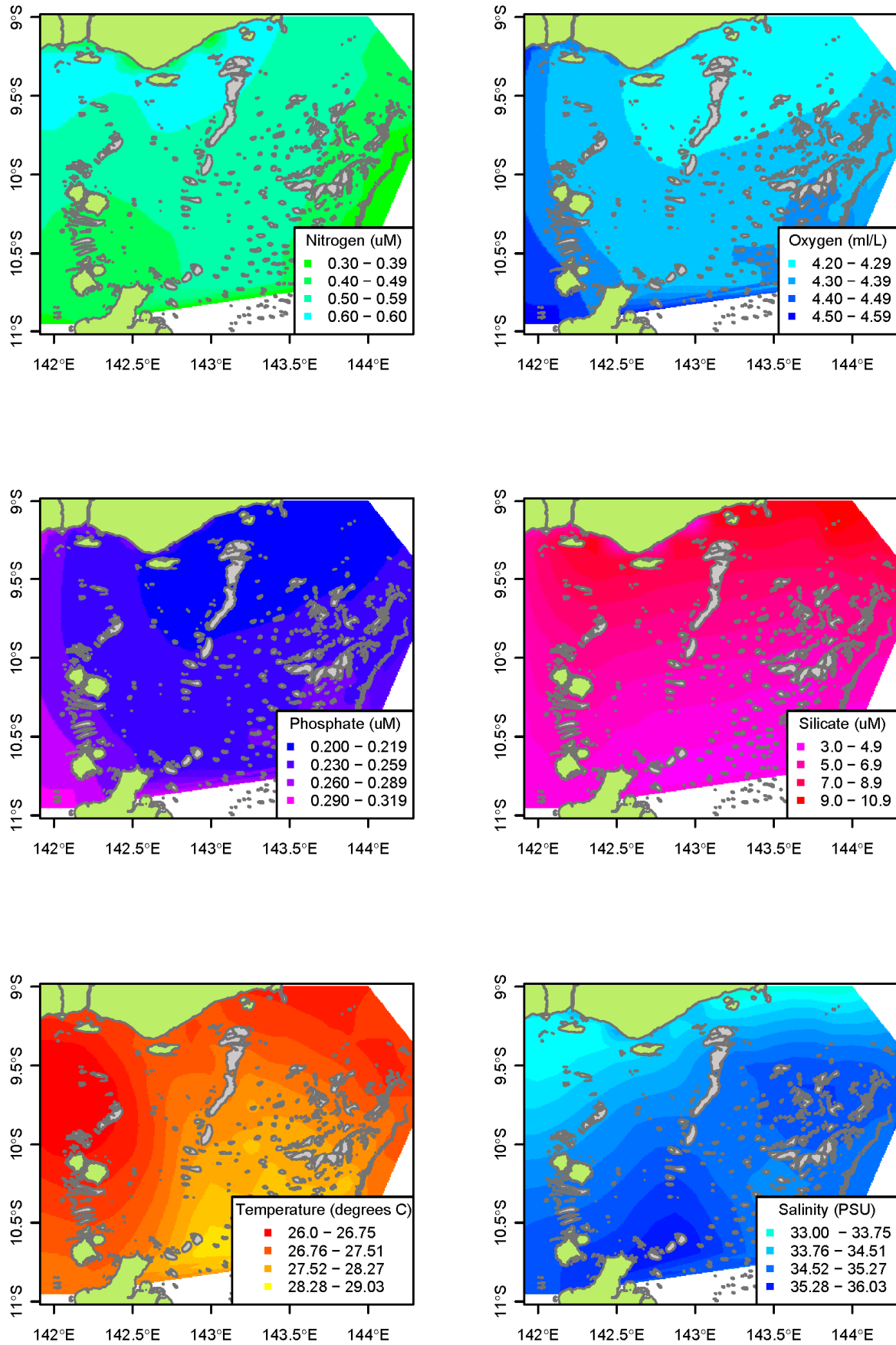


Figure 26. Weighted average surface nitrogen, oxygen, phosphate, silicate, temperature and salinity from the CARS dataset.

CSIRO Atlas of Regional Seas (CARS; (Ridgway and Dunn 2002)): This dataset is an Australia wide database of water-column physical and chemical attributes maintained by CMAR. The dataset included temporal series at fixed stations in Torres Strait and additional data collected during the CRC Torres Strait mapping and Characterisation (Pitcher et al. 2007) project has been collated to provide broader spatial coverage. The data have been modelled and interpolated at 0.01 degree resolution to provide a full coverage of the study area in 4 dimensions (spatially, through the water column and time; Figure 26). For the purposes of this project, this has been summarised by calculating the annual mean and standard deviation for each parameter estimated at the surface of the water column. The mean values for each parameter are shown in Figure 26; maps of the standard deviation are not shown. Because the standard deviation measure is over the modeled values for a full year, it may be considered as an index of seasonality for each parameter. Water properties evaluated included:

- Temperature: degrees C
- Salinity: psu
- Oxygen: ml.l^{-1}
- Silicate: μM
- Phosphate: μM
- Nitrate: μM

SeaWiFS Ocean Colour: This dataset includes estimates of mean and standard deviations for chlorophyll a concentration and turbidity, processed by CMAR based on SeaWiFS satellite data and calibration and validation algorithms. Relative benthic irradiance has been calculated, based on K490, latitude, and depth.

- Chlorophyll-a (mg.m^{-3}) concentration, mean and standard deviations
- K490 diffuse attenuation coefficient at wavelength 490nm, mean and standard deviations.

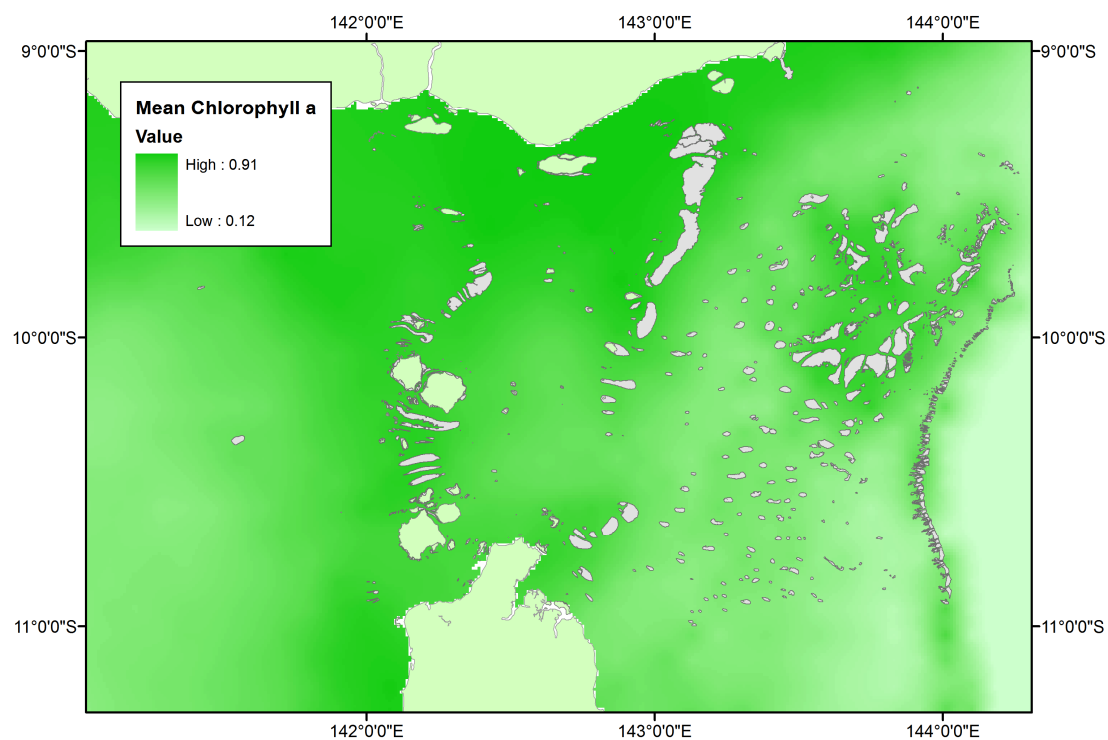


Figure 27. Mean annual chlorophyll a concentrations derived from SeaWiFS ocean colour.

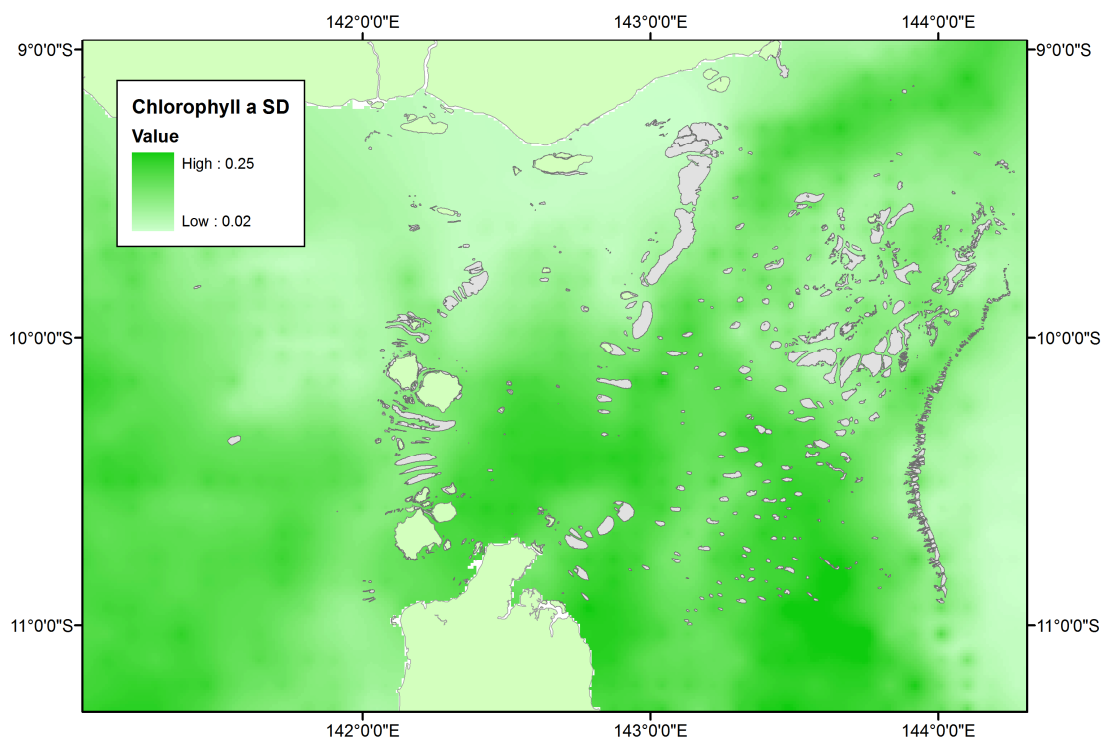


Figure 28. Standard deviation of the annual Chlorophyll a concentrations derived from SeaWiFS ocean colour.

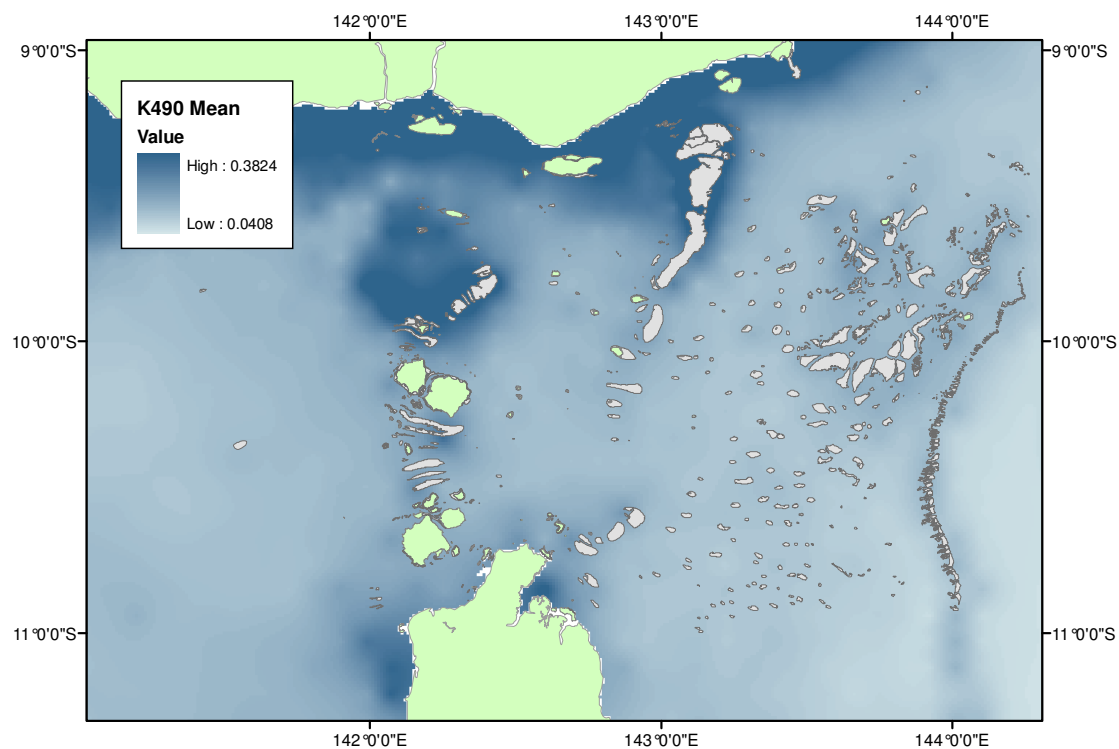


Figure 29. Mean annual K490 diffuse attenuation coefficient at wavelength 490nm derived from SeaWIFS ocean colour.

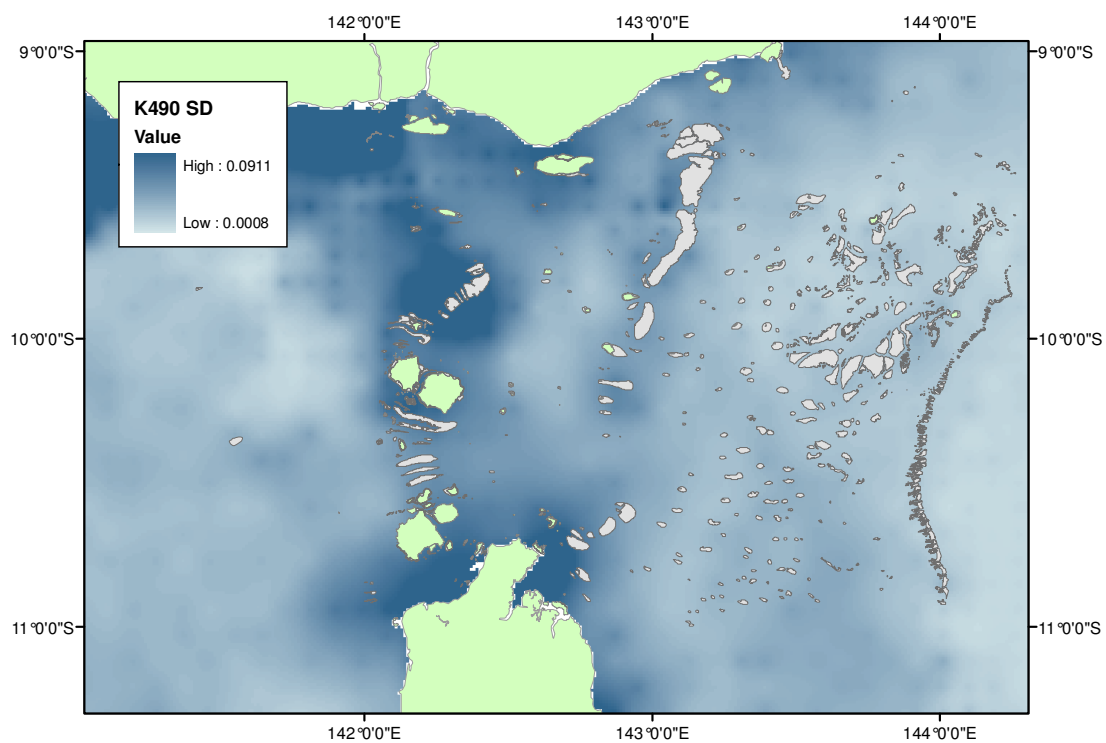


Figure 30. Standard deviation of the mean annual K490 diffuse attenuation coefficient at wavelength 490nm derived from SeaWIFS ocean colour.

Fetch: Fetch was estimated for each reef site to provide an index of exposure. Winds in Torres Strait are predominately from the southeast during the trade season (May to October) and from the northwest during the monsoon season (December to March; (Walsh and Nitttrouer 2003)). We calculated a southeasterly and northwesterly fetch for each reef site by estimating the distance to the nearest reef, island or mainland coastline both 135° (southeast) and 315° (northwest) from the site. This was done in an automated way in the GIS, but the values were manually checked and adjusted if necessary. For instance in some cases a particular site may lie in the lee of an extensive array of fragmented reefs which essentially provide shelter from the prevailing winds and seas. However, a line at 135° from the site may by chance pass through gaps between the reefs suggesting that the site has a large fetch. In cases such as these the fetch values were adjusted manually to reflect more accurately the level of exposure at the site.

Wave height, period and direction: The Torres Strait shelf is protected from the oceanic swells generated in the Coral Sea by the Ribbon Reefs and so despite an extended trade wind season, wave heights in the Strait are generally less than 3.5 m during the trade season and less than 1.5 m in the monsoon season (McMillan 1982). Wave height, period and direction were modelled using the Wave Model (WAM) data supplied by the Bureau of Meteorology and Geoscience Australia at a resolution of 0.1 degrees. Annual means for each wave field were calculated and mapped to each site.

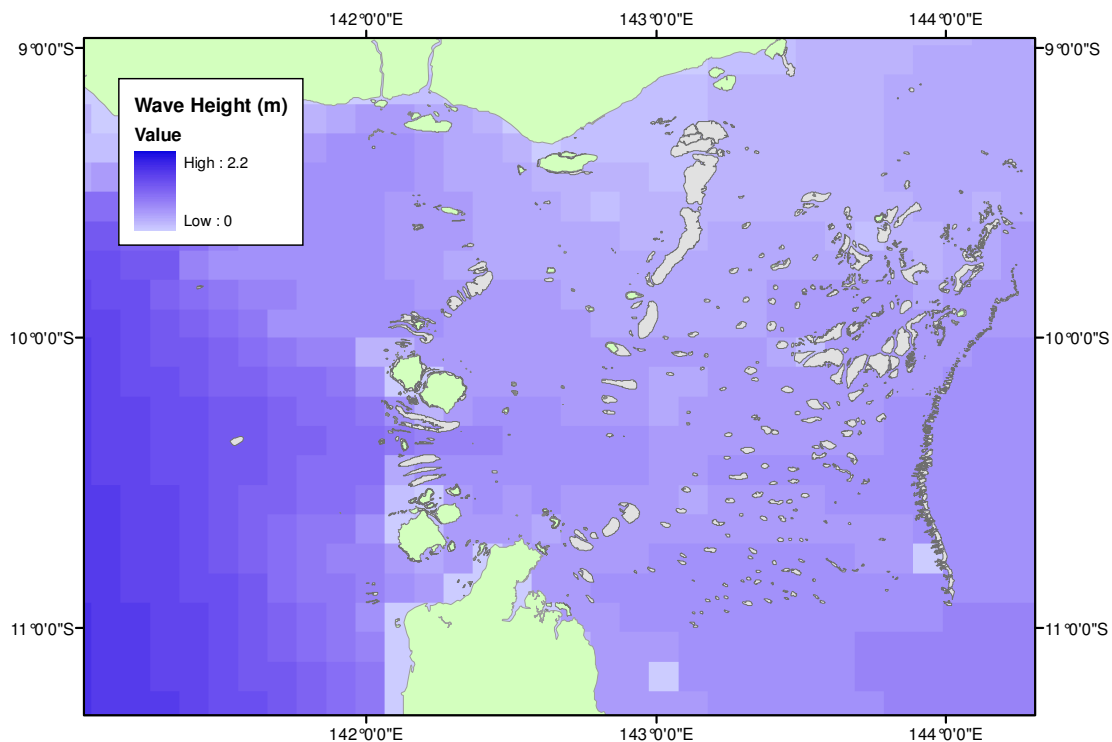


Figure 31. Modelled mean wave height generated by the Wave Model (WAM) data supplied by the Bureau of Meteorology.

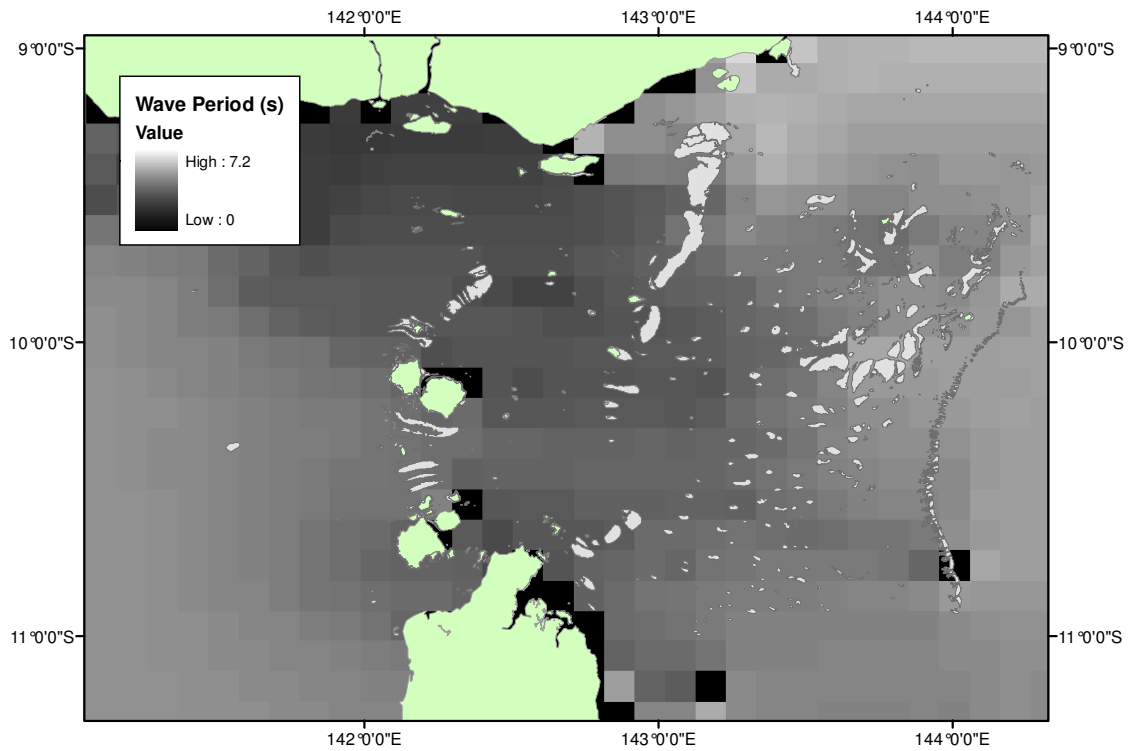


Figure 32. Modelled mean wave period generated by the Wave Model (WAM) data supplied by the Bureau of Meteorology.

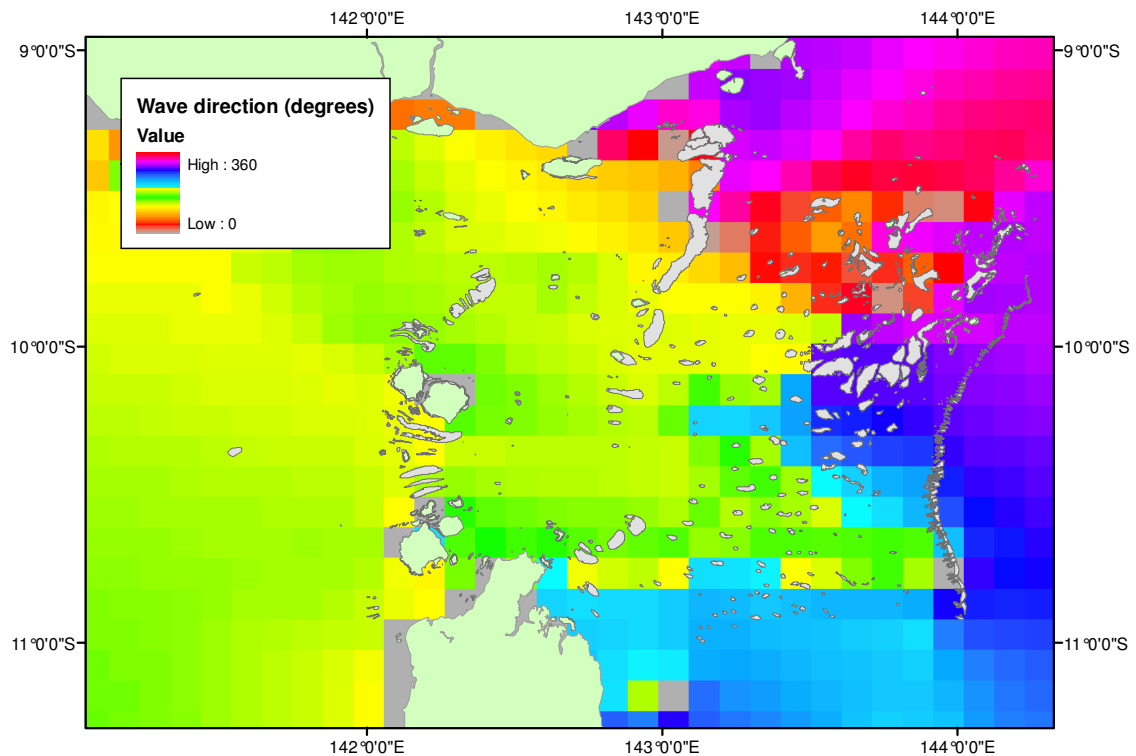


Figure 33. Modelled mean wave direction generated by the Wave Model (WAM) data supplied by the Bureau of Meteorology.

4.3.2. Maps of survey data

In order to summarise the data collected for each reef the following maps represent the mean value of each variable for each reef surveyed. The substrate and biotic characteristics of reef edges compared to reef tops are likely to be quite disparate and so separate maps are shown for reef edge and top characteristics.

Substrate maps

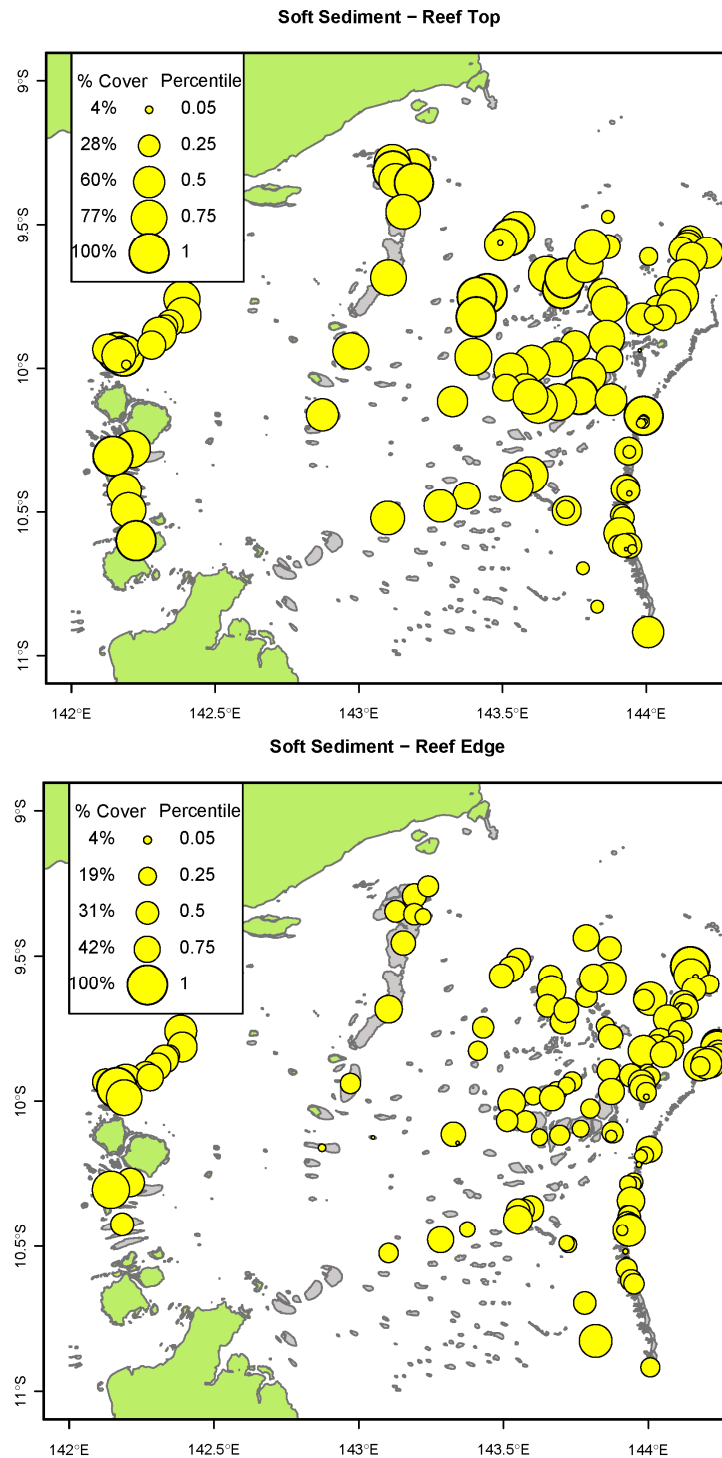


Figure 34. Mean percentage cover of soft sediments (mud and sand) on Torres Strait Reef tops and edges.

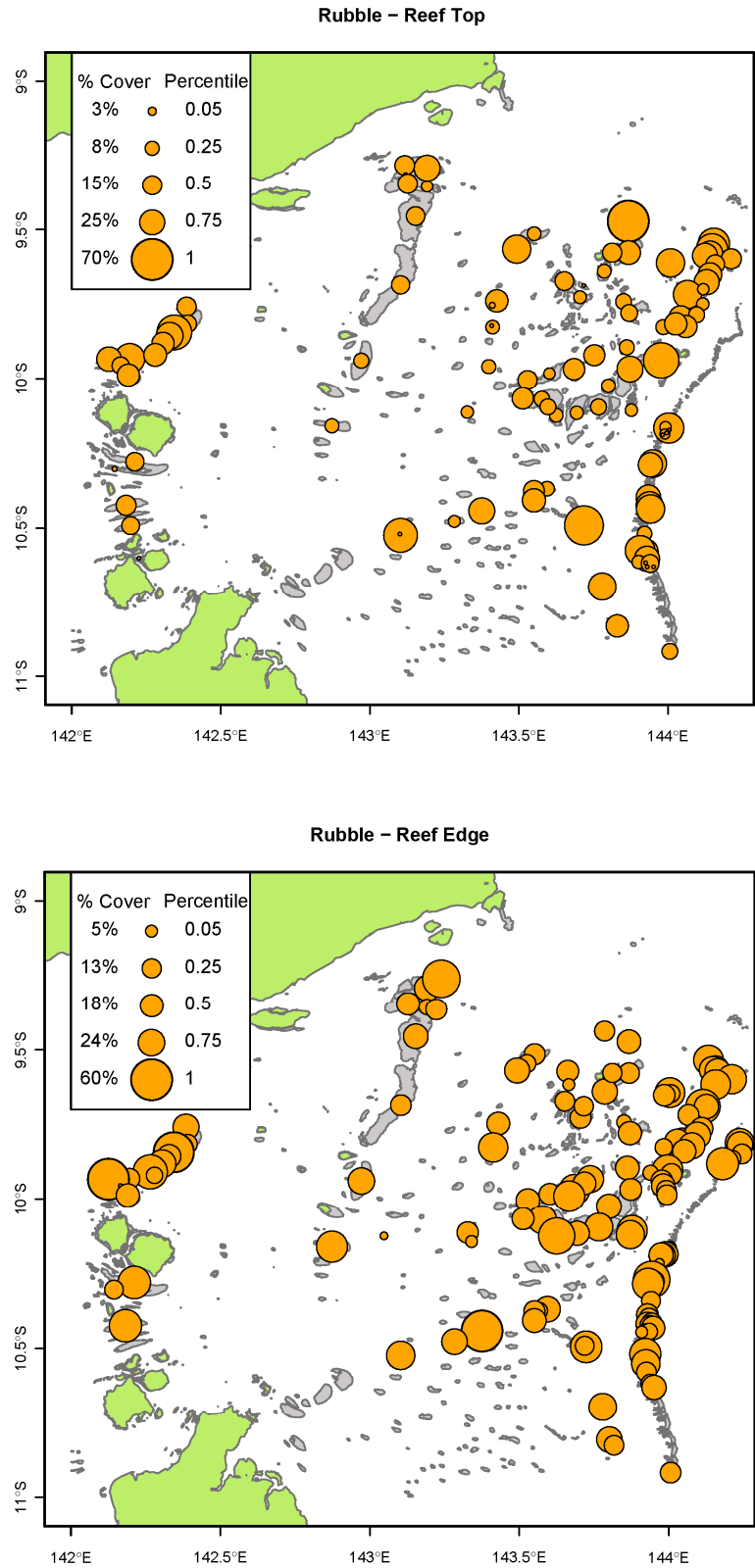


Figure 35. Mean percentage cover of rubble on Torres Strait Reef tops and edges.

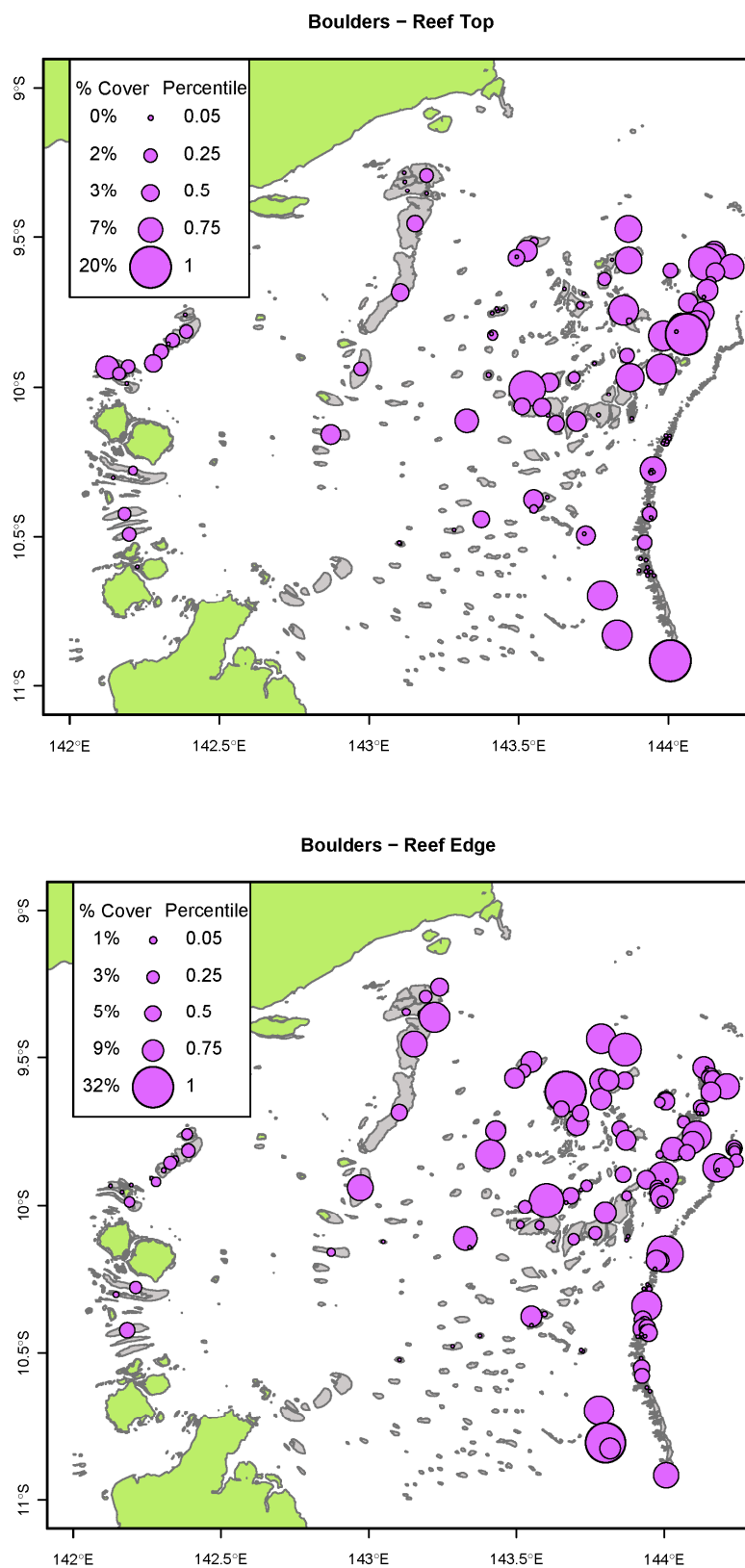


Figure 36. Mean percentage cover of boulders on Torres Strait Reef tops and edges.

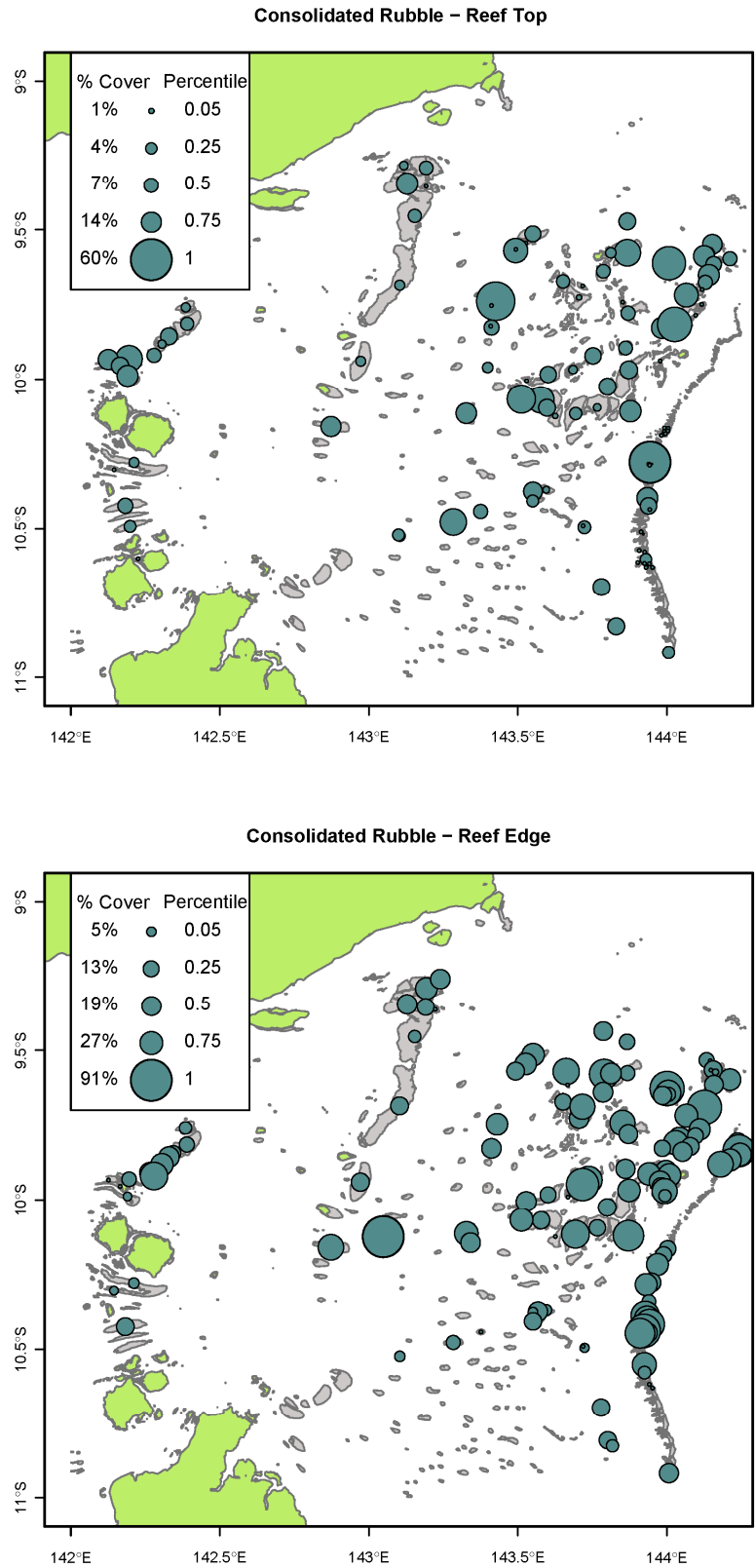


Figure 37. Mean percentage cover of consolidated rubble on Torres Strait Reef tops and edges.

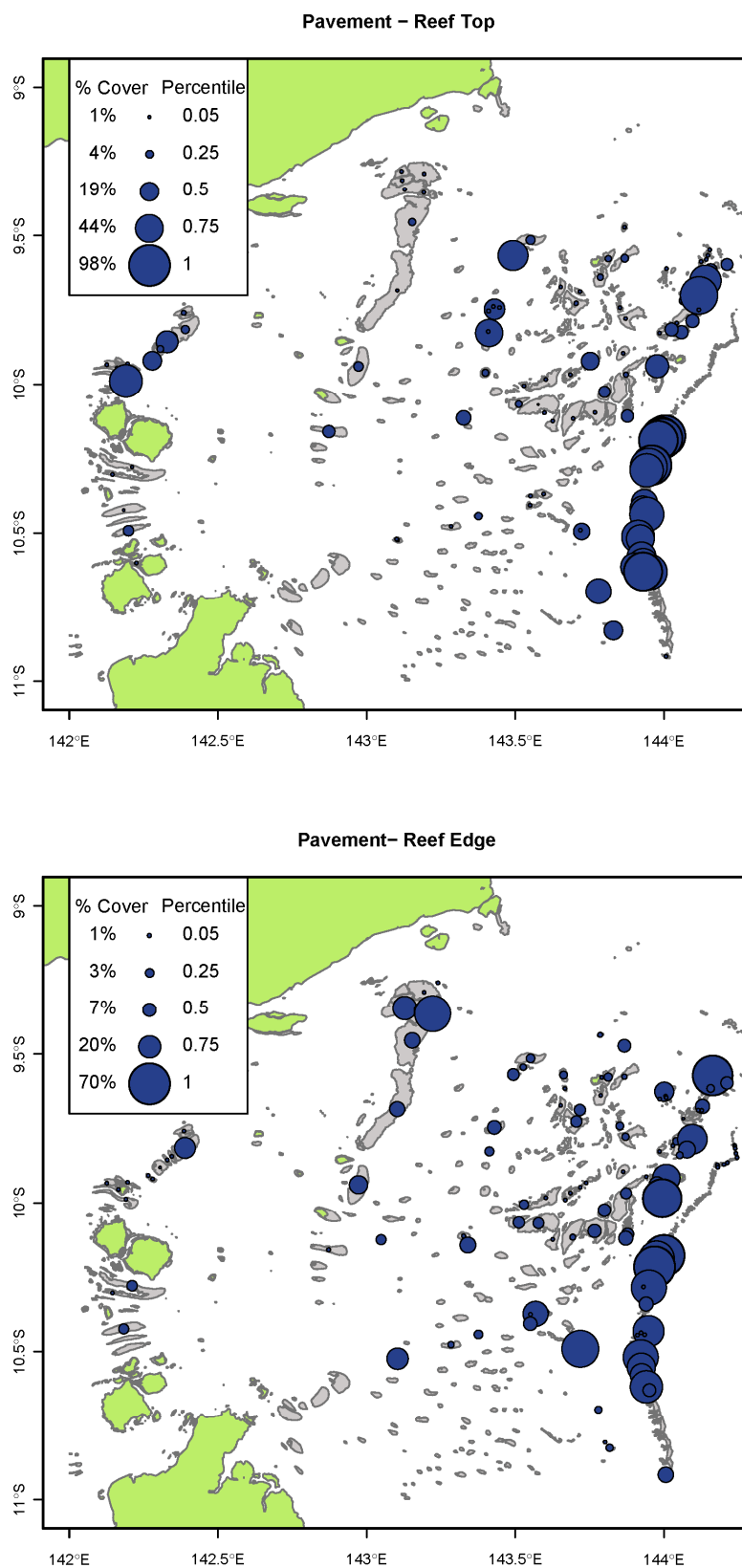


Figure 38. Mean percentage cover of pavement on Torres Strait Reef tops and edges.

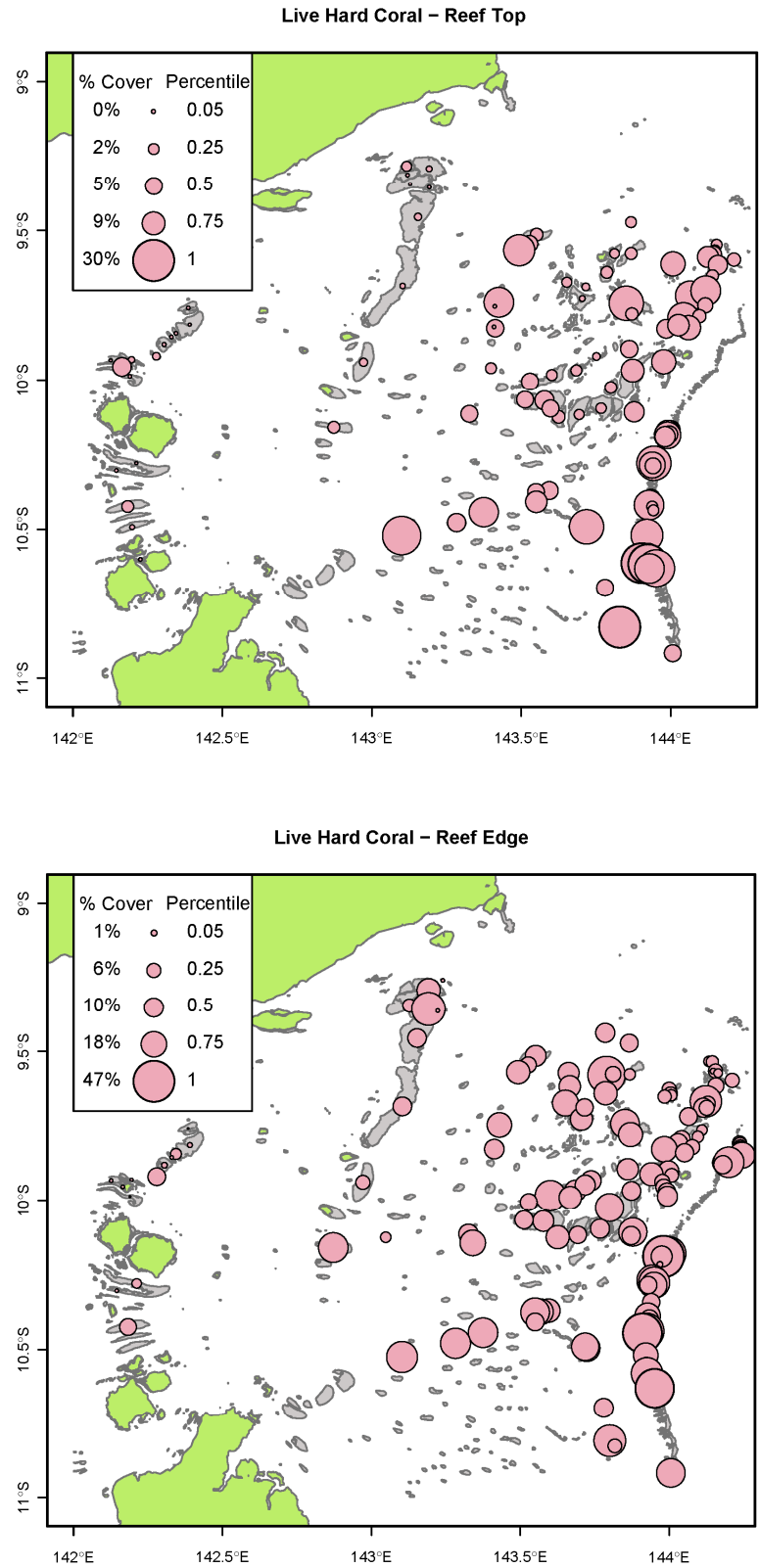


Figure 39. Mean percentage cover of live coral on Torres Strait Reef tops and edges.

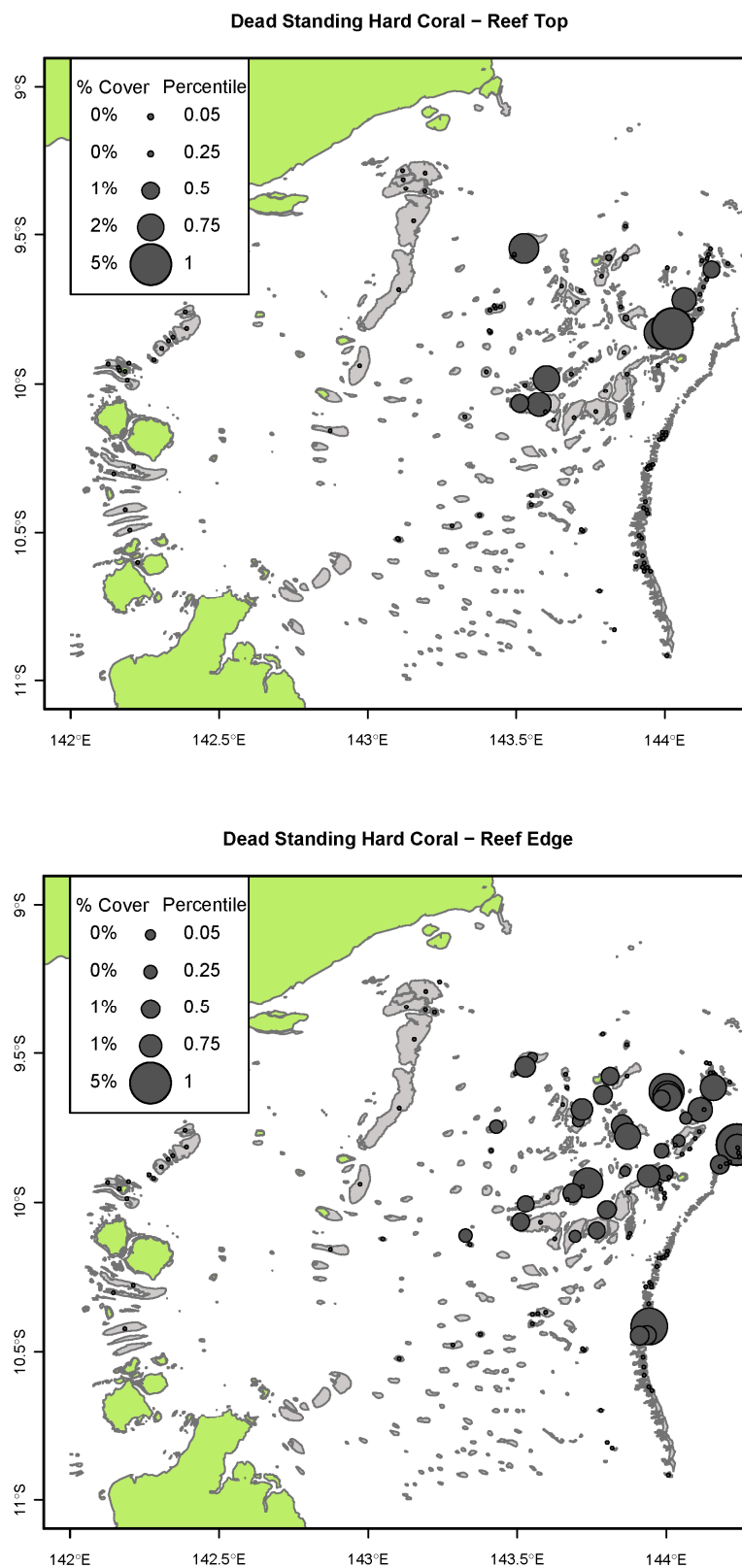


Figure 40. Mean percentage cover of dead standing hard coral on Torres Strait Reef tops and edges.

Soft sediments are more prevalent on the reef tops than edges and although they are common throughout the region, tend to be lowest on the eastern ribbon reefs (Figure 34). Rubble also appears to be fairly ubiquitous throughout the reefs surveyed and is more abundant on the reef edges, particularly on reefs around the Cumberland Passage in the eastern Strait (Figure 35). Boulders are also more abundant on the eastern reefs and more commonly found on edge habitats compared to the reef tops (Figure 36). Consolidated rubble is relatively uncommon on the reef tops, but found throughout on the reef edges (Figure 37). Pavement is generally characteristic of exposed reef crests and so it is not surprising to note that it is abundant and common on the top and shallow edges of the Ribbon reefs in the eastern Torres Strait (Figure 38). It is also found on the edges of some of the northern Warrior Reefs and the reefs around Mabuiag Island and the Orman Reefs. Mean live coral cover is generally below 30% on the reef tops and is highest on the eastern Ribbon Reefs and very low on the western reefs (Figure 39). Live coral cover is generally greater on the reef edges throughout, although very low on the westernmost reefs. Dead standing hard coral was only recorded in the eastern Torres Strait and even there is relatively uncommon with a mean cover $\leq 5\%$ (Figure 40).

Coral growth forms

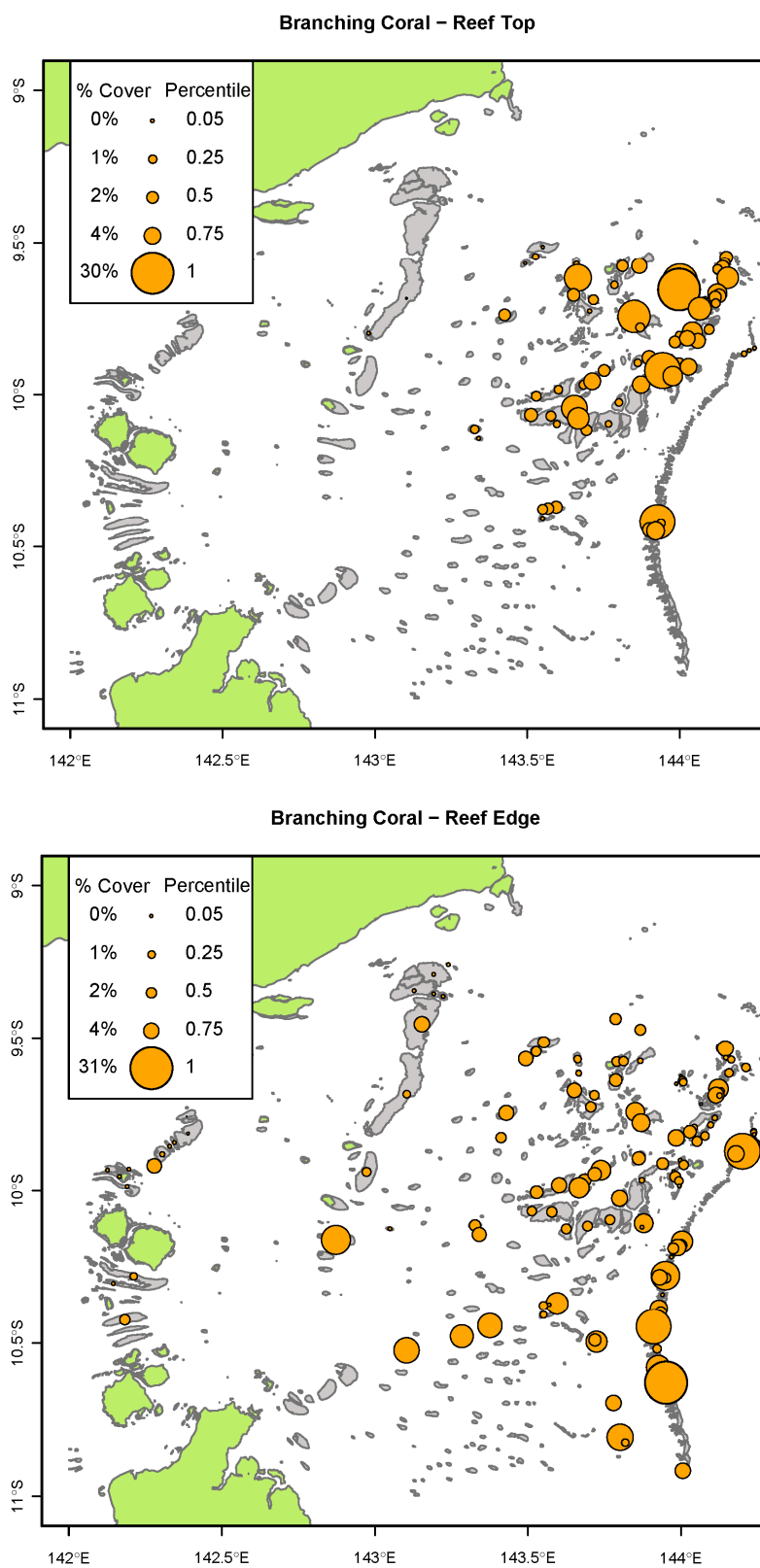


Figure 41. Percent cover of branching corals on Torres Strait reef tops and edges.

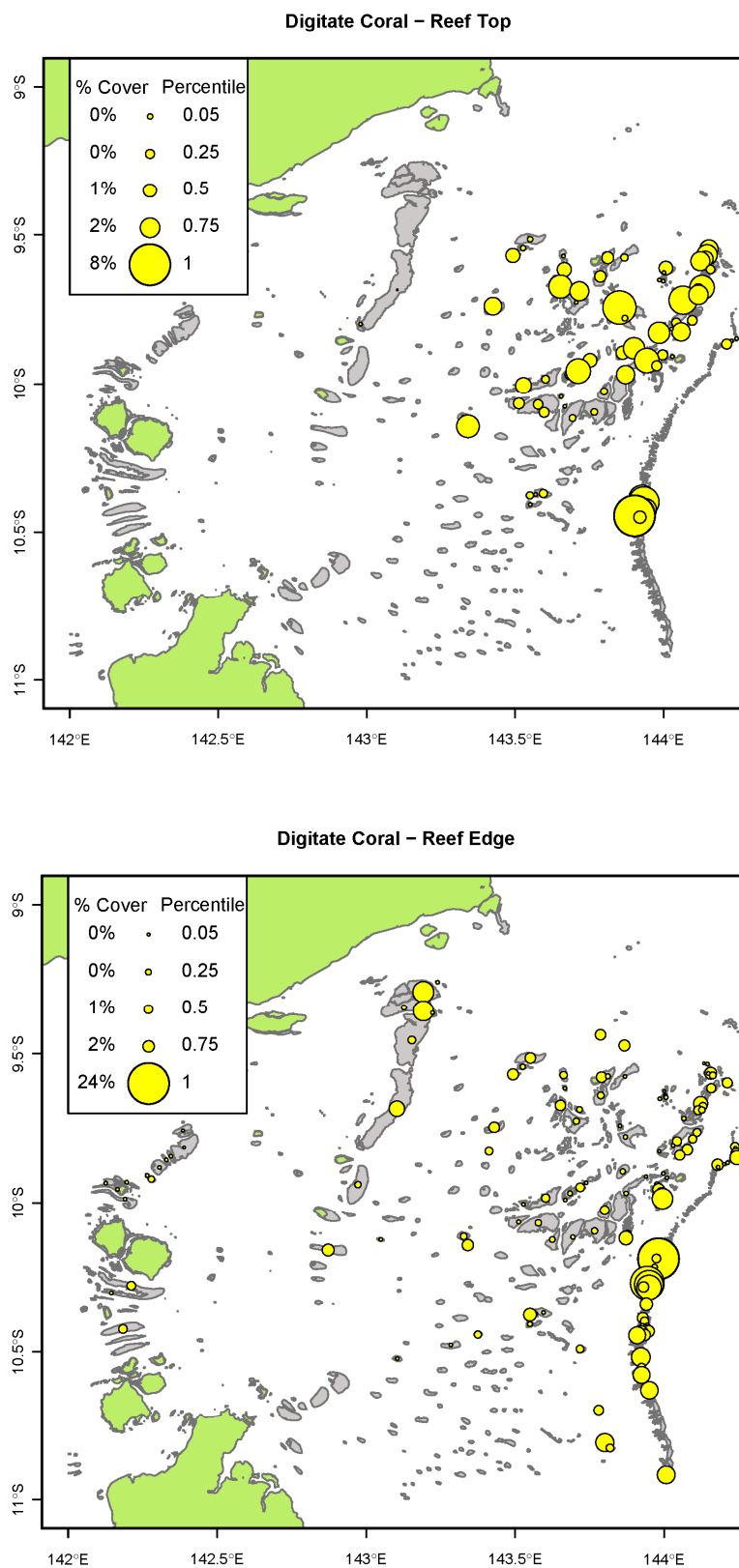


Figure 42. Percent cover of digitate corals on Torres Strait reef tops and edges.

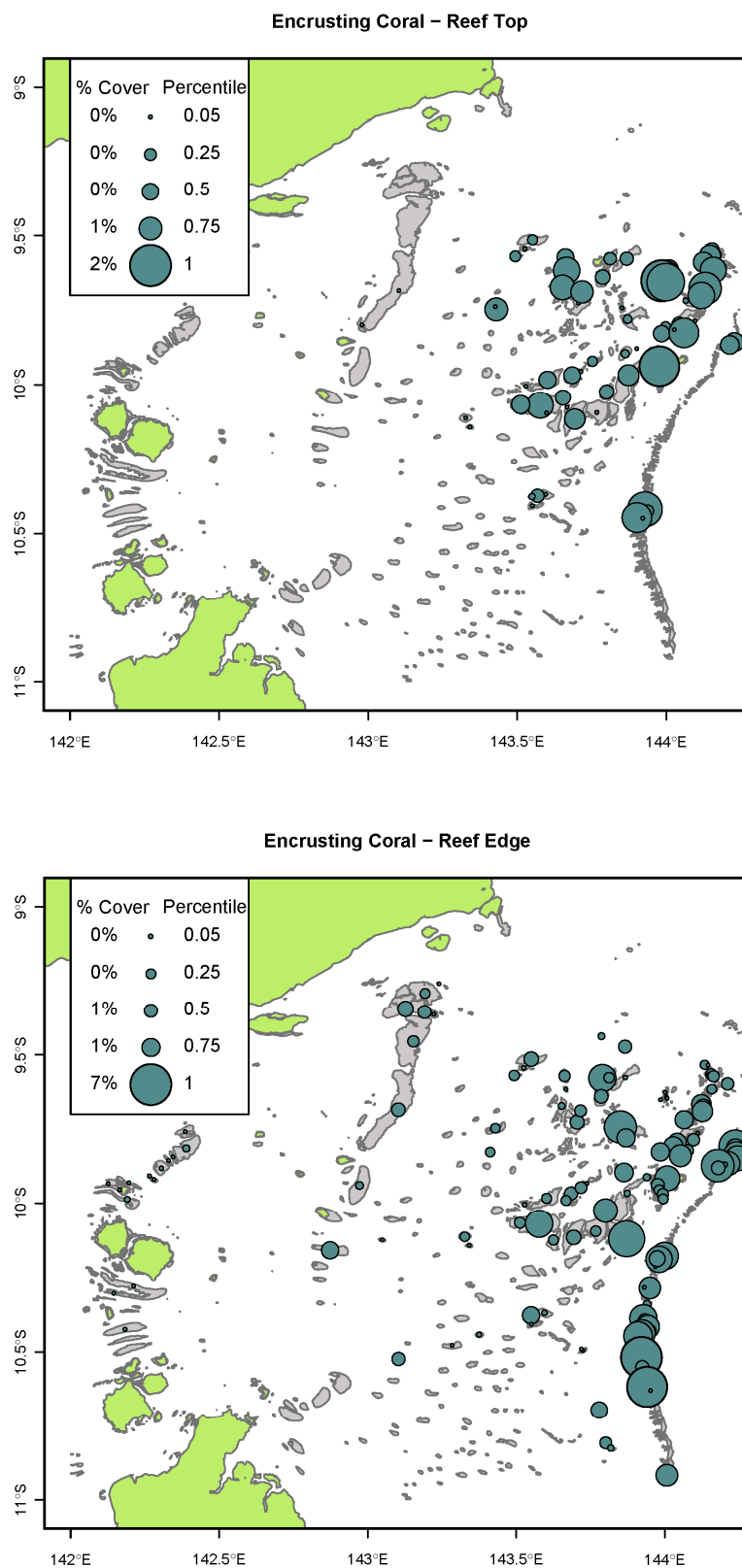


Figure 43. Percent cover of encrusting corals on Torres Strait reef tops and edges.

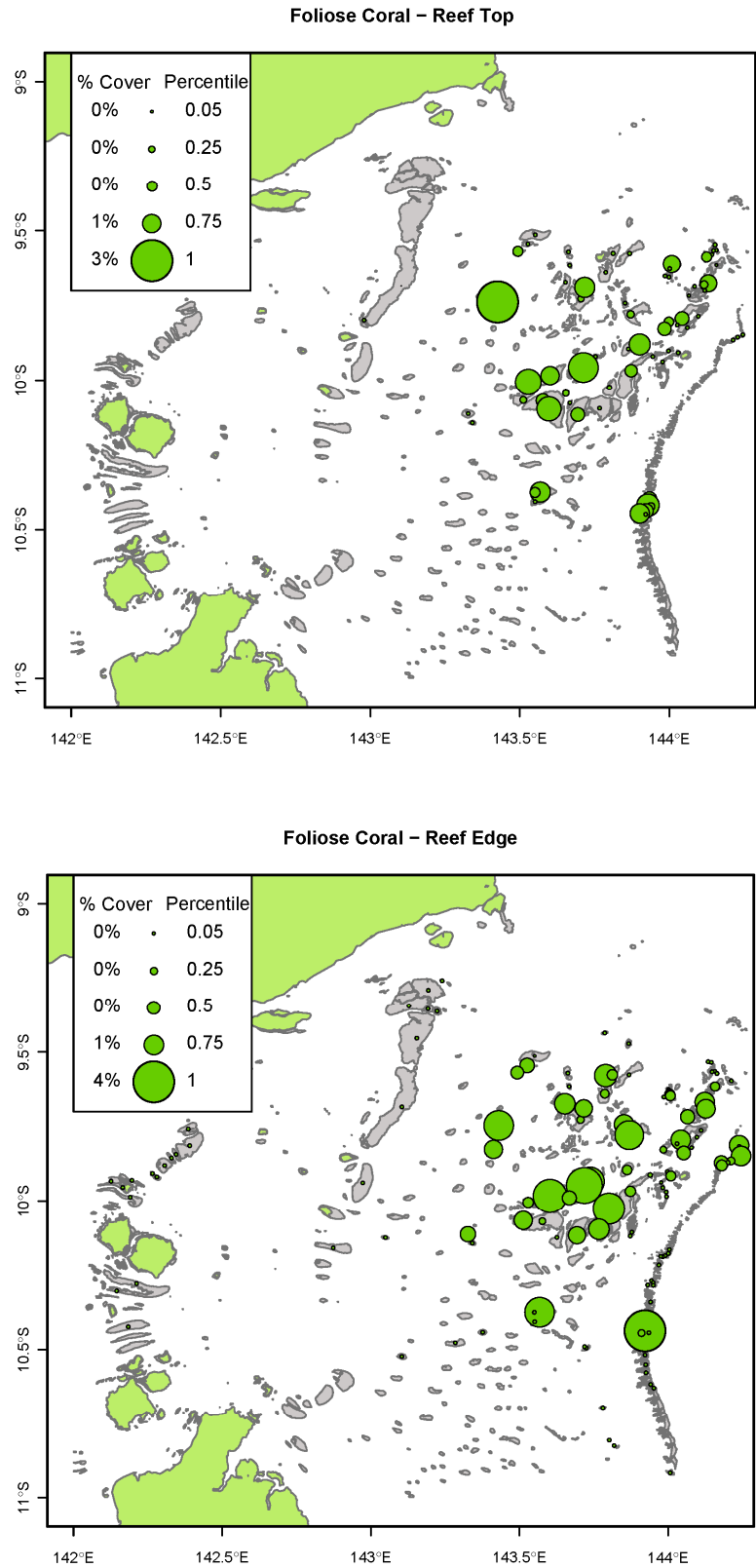


Figure 44. Percent cover of foliose corals on Torres Strait reef tops and edges.

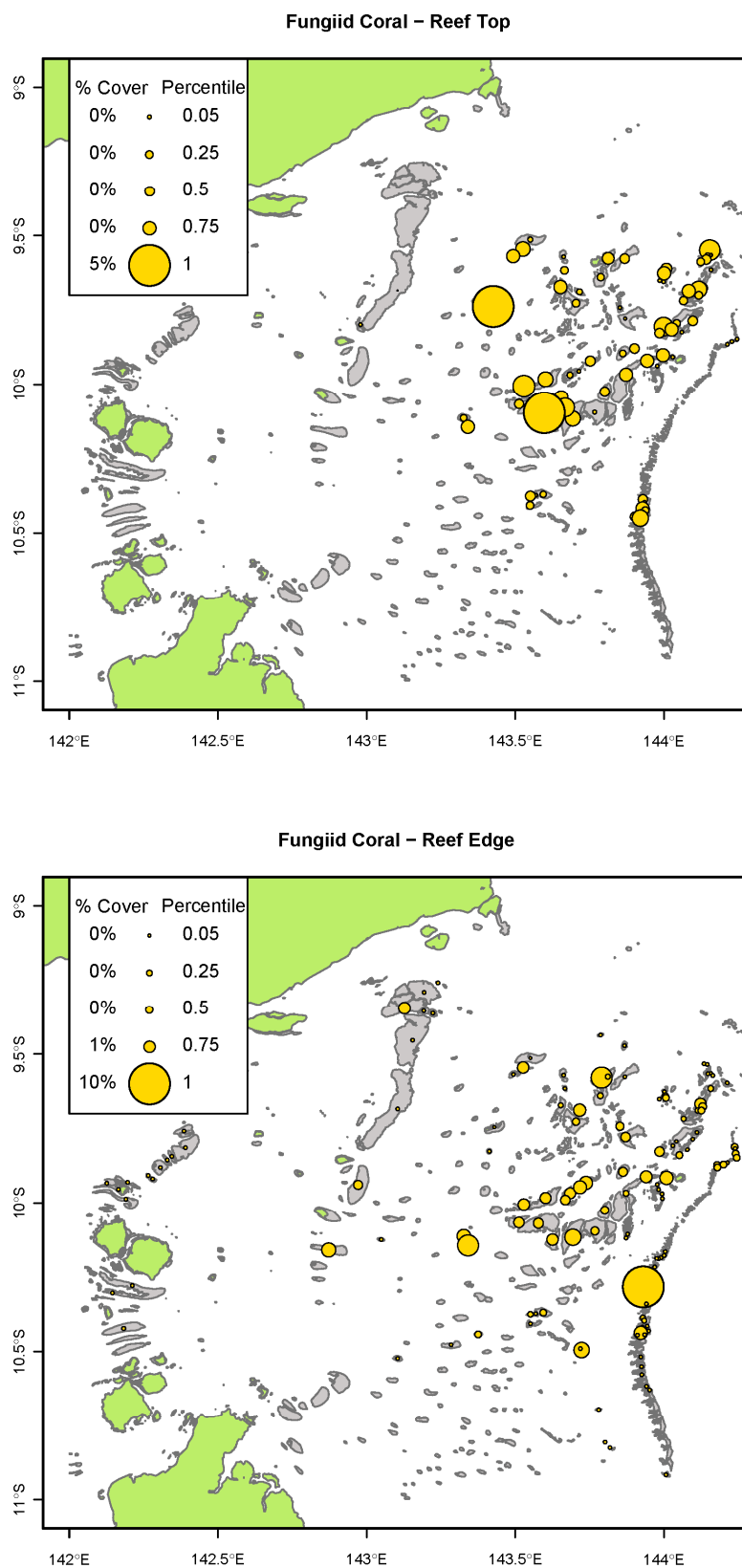


Figure 45. Percent cover of fungiid corals on Torres Strait reef tops and edges.

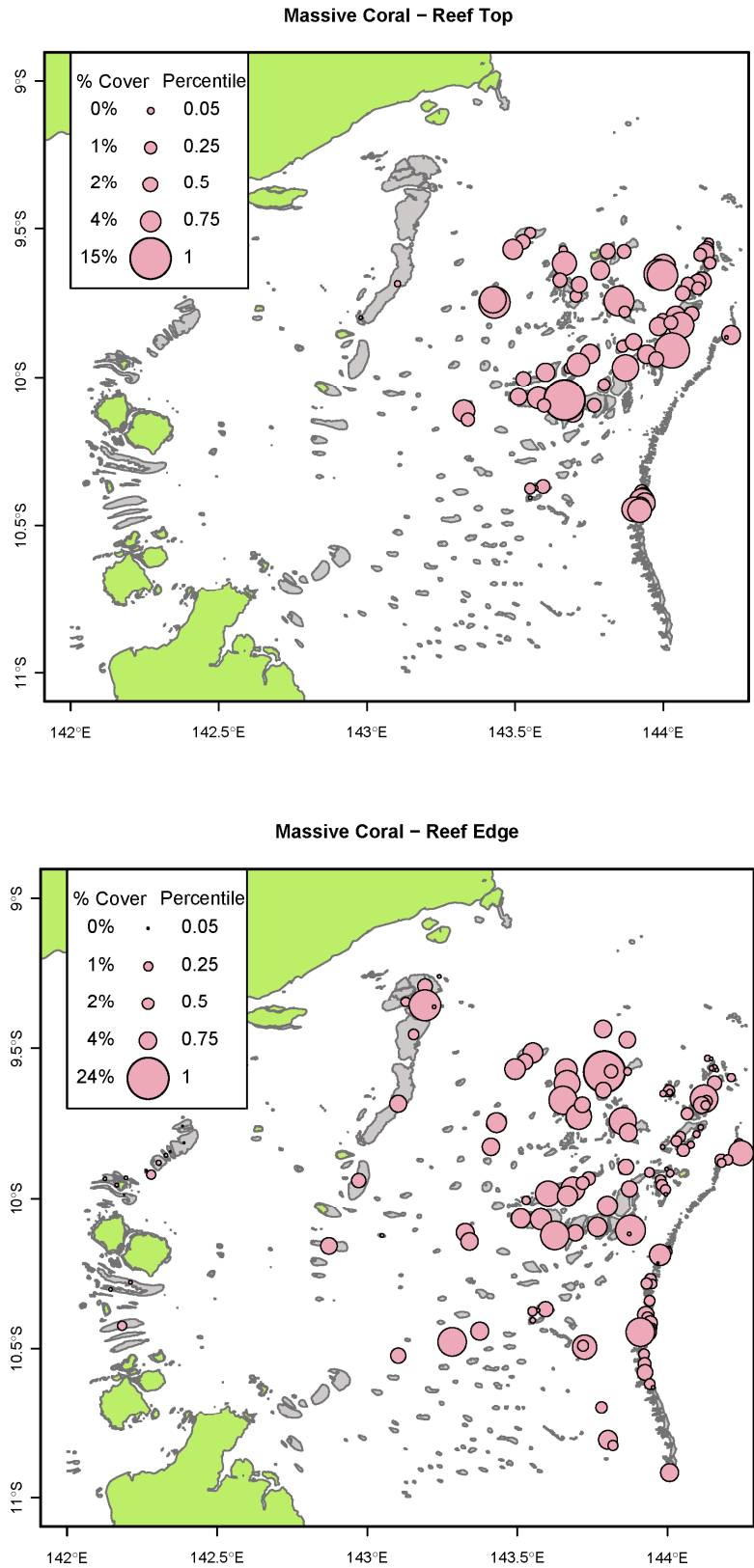


Figure 46. Percent cover of massive corals on Torres Strait reef tops and edges.

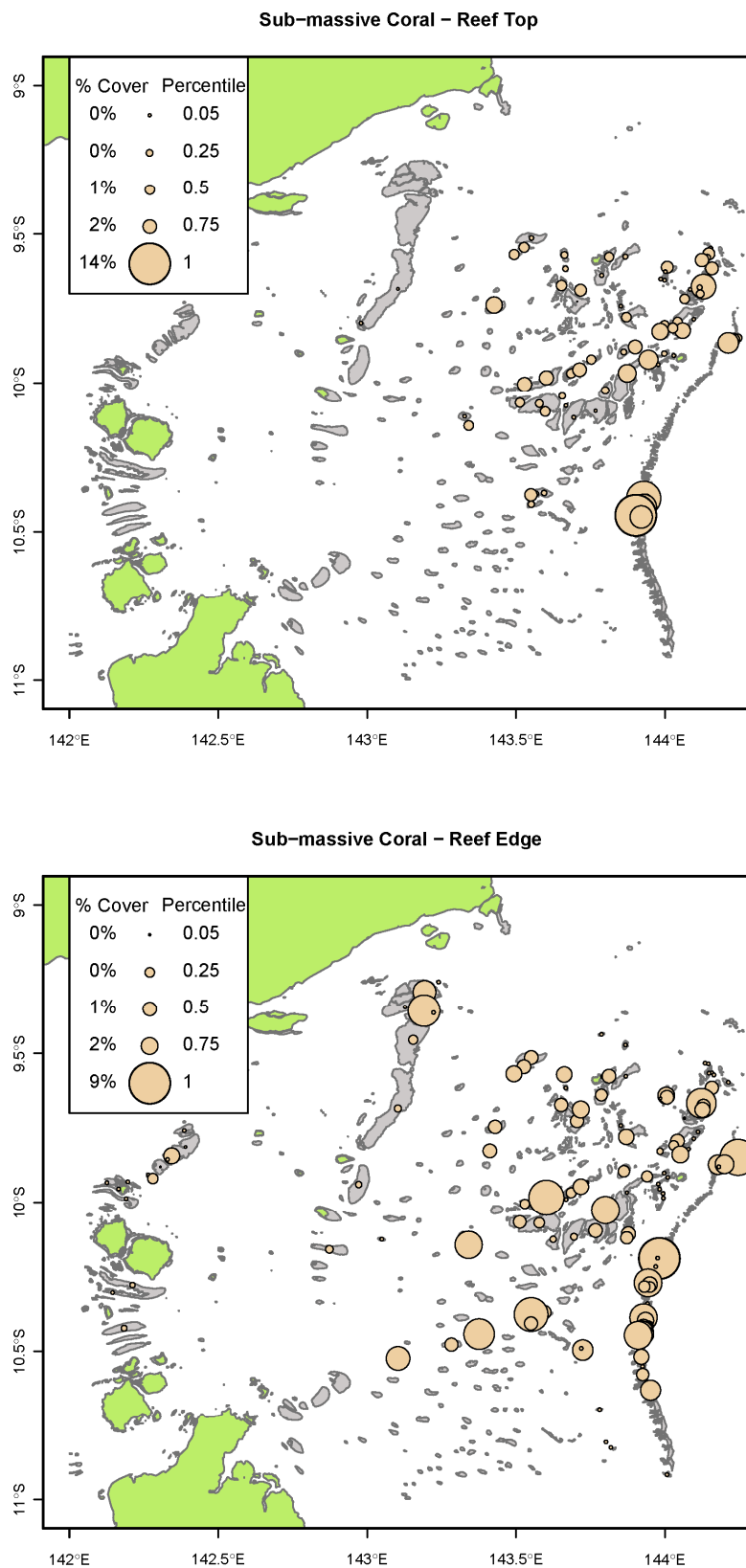


Figure 47. Percent cover of sub-massive corals on Torres Strait reef tops and edges.

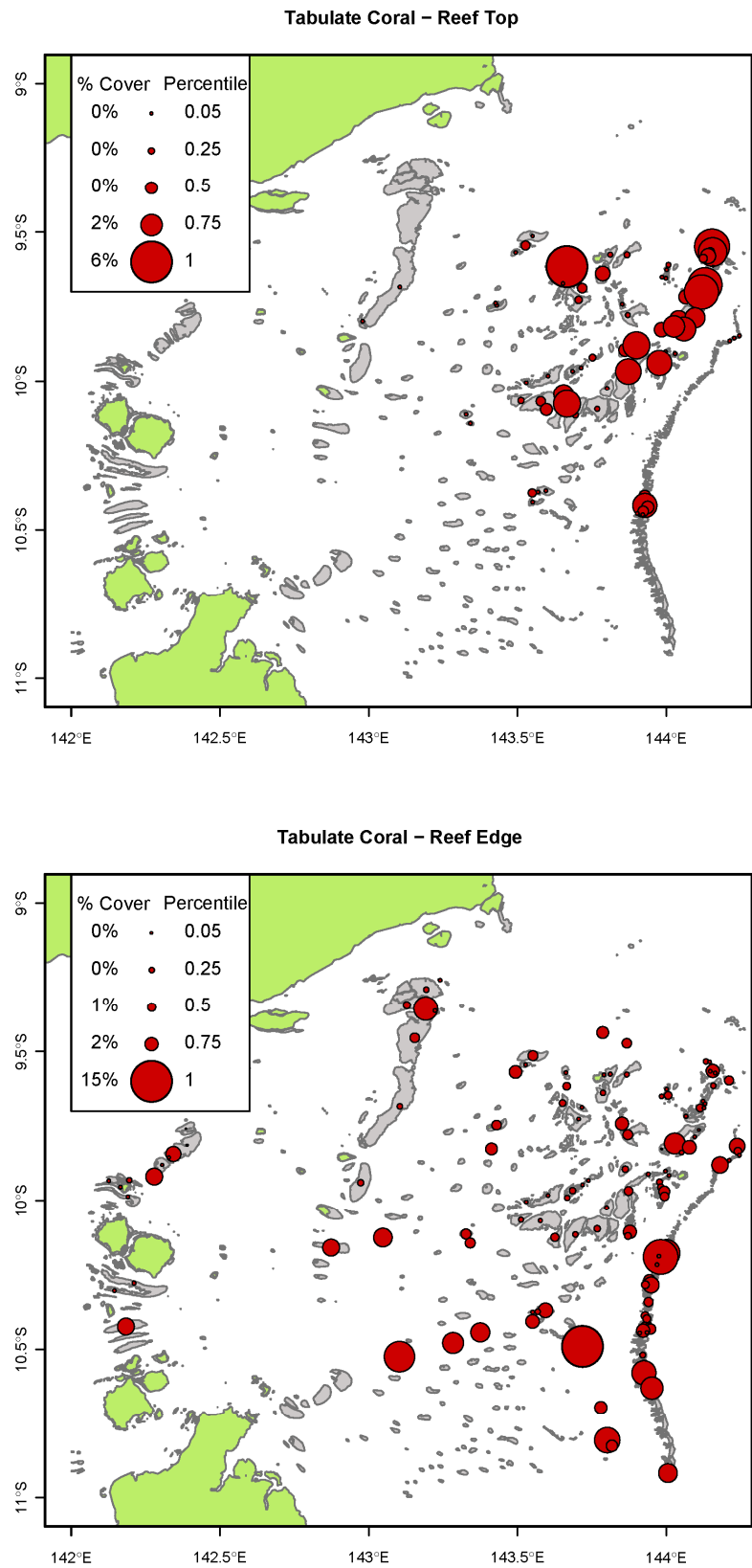


Figure 48. Percent cover of tabulate corals on Torres Strait reef tops and edges.

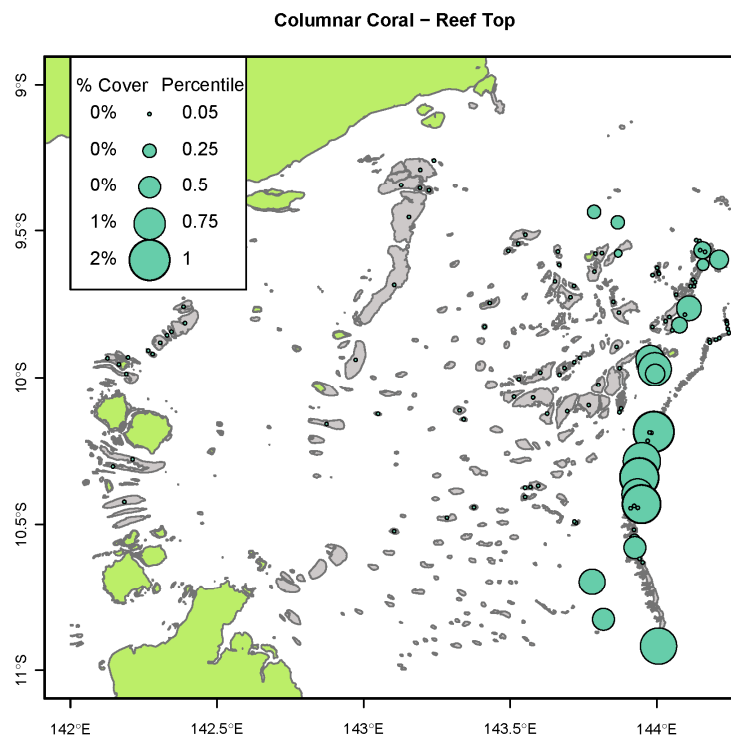


Figure 49. Percent cover of columnar corals on Torres Strait reef edges (none were recorded on the reef tops).

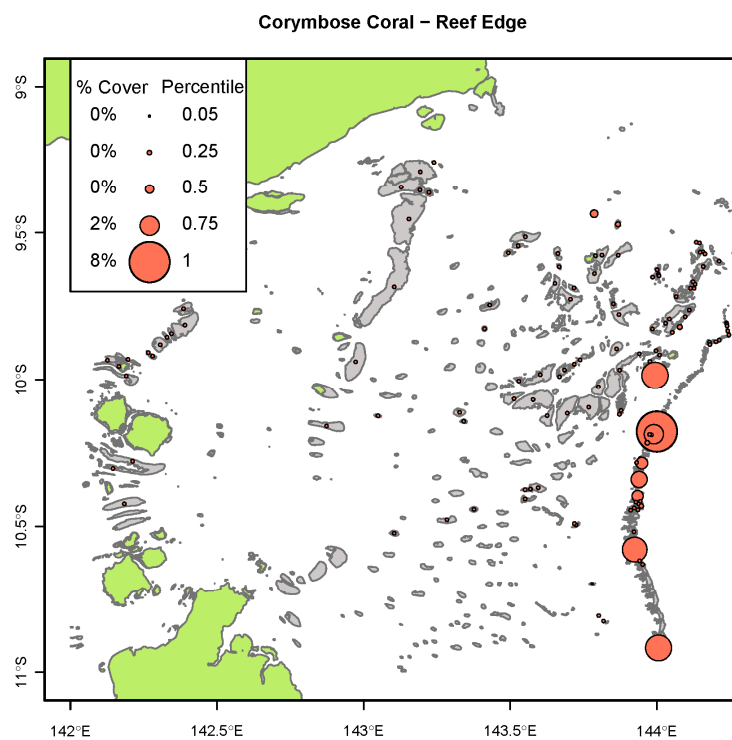


Figure 50. Percent cover of corymbose corals on Torres Strait reef edges (none were recorded on the reef tops).

Branching corals were found in reasonably high proportions (up to ~ 30% total cover of the seabed) on both the reef tops and edges, primarily in the eastern parts of Torres Strait (Figure 41). Digitate corals were generally more abundant on the reef tops on the reefs around the Cumberland and Hibernia Passages and more common on the edges on the ribbon reefs (Figure 42). Encrusting and foliose corals did not form a high proportion of the total live coral cover, but were more common in the east. Similarly, fungiid corals were also uncommon and only featured on one or two reefs (Figure 43 and Figure 44). Massive corals covered up to 24% of the seabed on the reef edges and were one of the more dominant growth forms found on the Warrior Reefs (Figure 46). Massive corals were less common on the ribbon reefs compared to the reefs around the Cumberland and Hibernia Passages. Sub-massive corals were also found on the Warrior Reefs and were most common on the ribbon reef tops, but were more widespread on the reef edges (Figure 47). Tabulate corals were generally more abundant on the reef tops on the reefs around the Cumberland and Hibernia Passages, but uncommon elsewhere on the tops (Figure 48). They featured on reef edges on the Warrior, Orman and the Ribbon Reefs and also on the reefs in the south eastern parts of the Strait. Columnar and corymbose corals were not recorded on reef tops and were primarily found in small amounts on the edges of the ribbon reefs (Figure 49 and Figure 50).

Reef fish distribution

A total of 231 reef edge sites were surveyed on 57 different reefs extending from the Ribbon reefs in the east to the reefs north of Prince of Wales Island in the west (Figure 51, Figure 53).

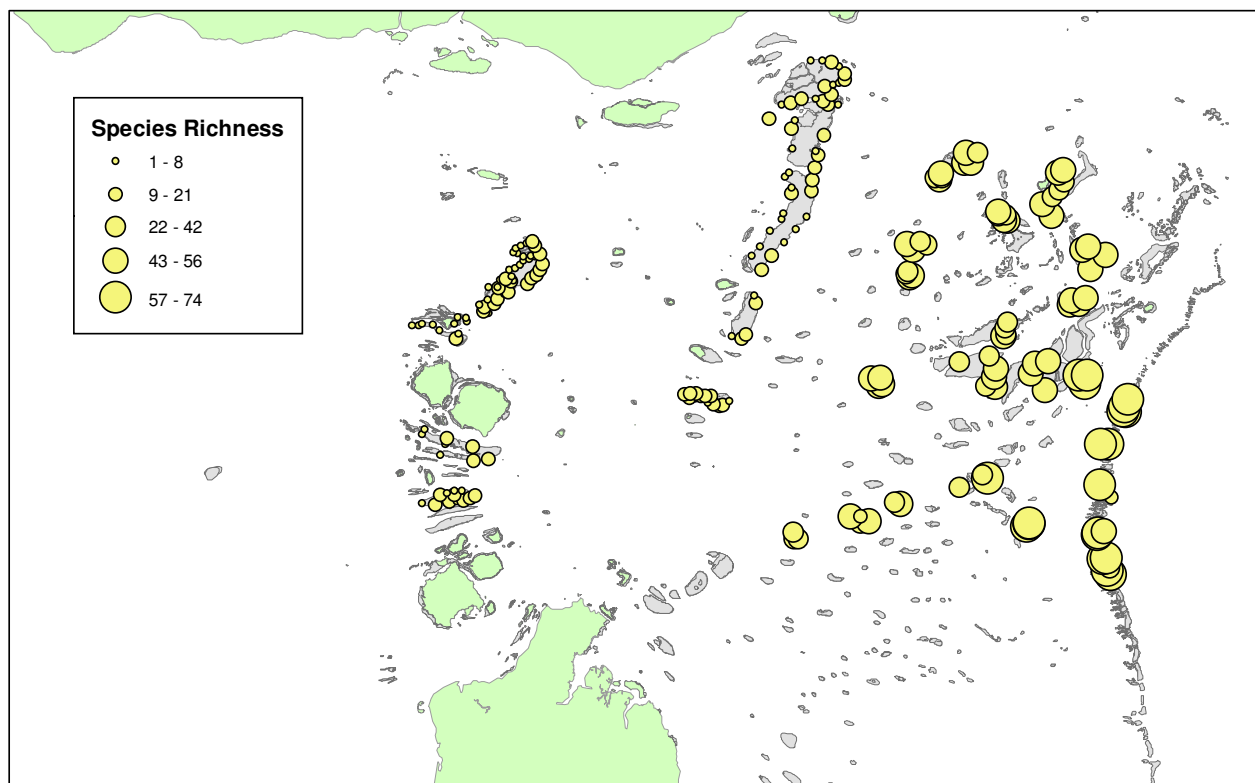


Figure 51. Species richness of reef fish on all 231 reef edge sites surveyed during 1995-96.

Species richness on the eastern reefs was significantly greater than that on the reefs of the central and western Strait (Figure 51). A tree model was fitted to the species richness data using all the available physical covariates (Section 4.3.1). The standard deviation (SD) of surface salinity (standard deviation over the full historical dataset) was the most important covariate in explaining the variation of reef fish species richness across the Torres Strait (Figure 52).

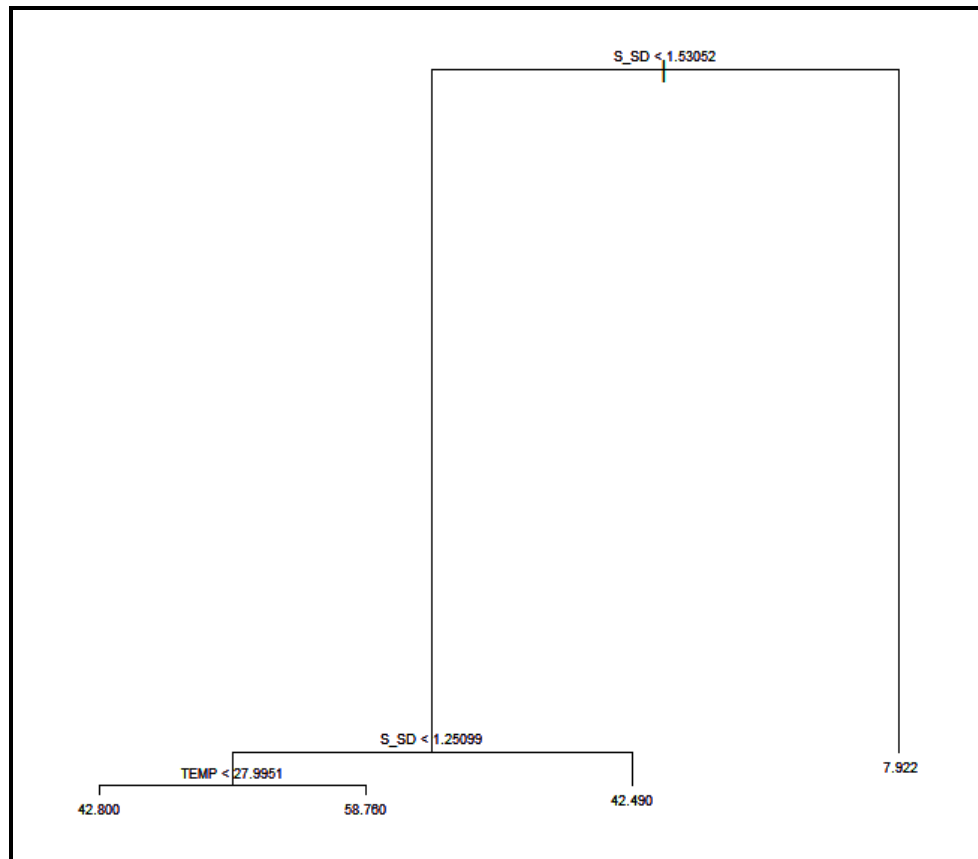


Figure 52. Plot of the tree model of reef species richness in terms of the physical covariates. For a full list of the physical covariates included in the analysis see Section 4.3.1.

The initial split was between sites having a relatively high salinity SD (≥ 1.53052) and those with a lesser salinity SD (Figure 52). The sites with a high SD had low mean species richness (7.922) and none of the other physical covariates explained a significant amount of the variation in the species richness among these sites. Of the sites where the salinity SD was < 1.53052 there was a further split at a threshold of a salinity SD of 1.25099. The group of sites having the lowest salinity SD were then split at a temperature threshold of 27.9951°C . The length of the stems on the later two splits indicate that the contribution of these explanatory variables was much less than that of the first split at a salinity SD of 1.53052. Note that all site groups at the first split of salinity SD < 1.53052 had relatively high mean species richness estimates of > 42.4 .

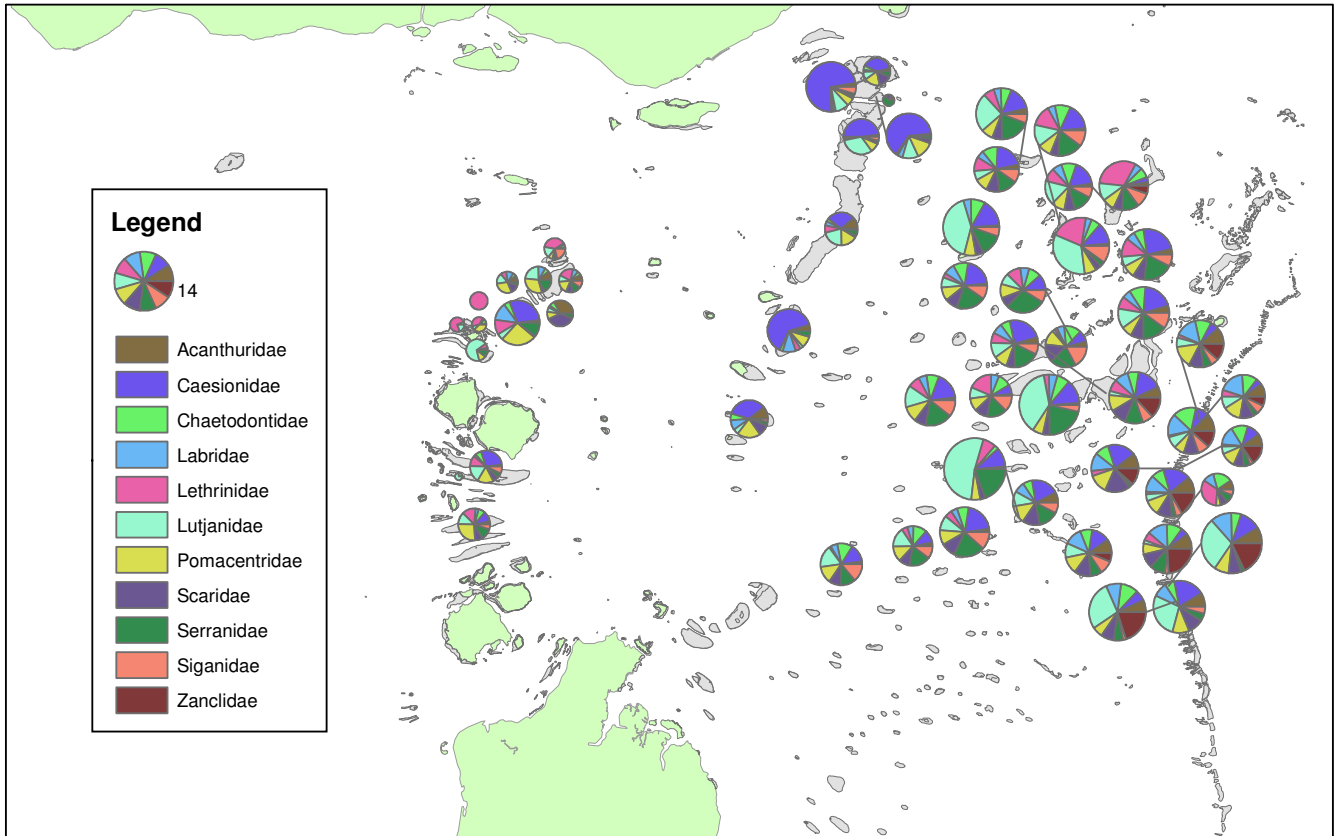


Figure 53. Map showing the distribution and abundance of the 11 fish families surveyed on Torres Strait reefs during 1995-96. The pie charts show the mean numbers of fish per ha of each family recorded on the reefs.

Overall fish abundance was greater on the eastern reefs and there were clear differences in the dominant families between east and west (Figure 53). Lutjanidae and Zanclidae were dominant on the outer Ribbon reefs, while Lutjanidae and Serranidae were common on the mid –shelf reefs, particularly in the south-east. In the north-east, Lutjanidae, Caesionidae and Lethrinidae were the dominant families. In the central Torres Strait, on the Warrior Reef complex, Caesionidae were generally dominant across all reefs. The western reefs did not appear to show any consistent pattern with variable contributions from a variety of families.

Holothurian distribution

Over all surveys a total of 2196 sites have been surveyed for holothurians as part of AFMA (Australian Fisheries Management Authority) funded stock assessment surveys. Twenty-five different taxa were identified, 23 of these to species (Table 8).

Table 8. Species of holothurians identified throughout the Torres Strait during diver surveys conducted from 1995 to 2005.

Taxon	Common name	Commercial value
<i>Actinopyga echinites</i>	Deep-water redfish	Medium
<i>Actinopyga lecanora</i>	Stonefish	Low
<i>Actinopyga mauritiana</i>	Surf redfish	Low to medium
<i>Actinopyga miliaris</i>	Blackfish	Medium
<i>Bohadschia argus</i>	Tigerfish	Low
<i>Bohadschia graeffei</i>	Flowerfish	Low
<i>Bohadschia marmorata</i>	Brown sandfish	Low
<i>Bohadschia similus</i>	Chalkfish	Low
<i>Holothuria atra</i>	Lolly fish	Low
<i>Holothuria coluber</i>	Snakefish	Low
<i>Holothuria edulis</i>	Pinkfish	Low
<i>Holothuria fuscogilva</i>	White teatfish	High
<i>Holothuria fuscopunctata</i>	Elephant's trunk fish	Low
<i>Holothuria hilla</i>		Non-commercial
<i>Holothuria leucospilota</i>		Non-commercial
<i>Holothuria nobilis</i>	Black teatfish	High
<i>Holothuria scabra</i>	Sandfish	High
<i>Stichopus chloronotus</i>	Greenfish	Low
<i>Stichopus horrens</i>	Peanutfish	Medium
<i>Stichopus variegatus</i>	Curryfish	Low to medium
<i>Stichopus vastus</i>	Curryfish	Medium
<i>Synaptidae</i>		Non-commercial
<i>Thelenota ananus</i>	Prickly redfish	High
<i>Thelenota anax</i>	Amberfish	Low

Holothurians were generally only found in the eastern and central parts of the Torres Strait with the highest densities on reefs around the Hibernia Passage (Figure 54). Holothurians have been grouped to genus level in this figure for clarity.

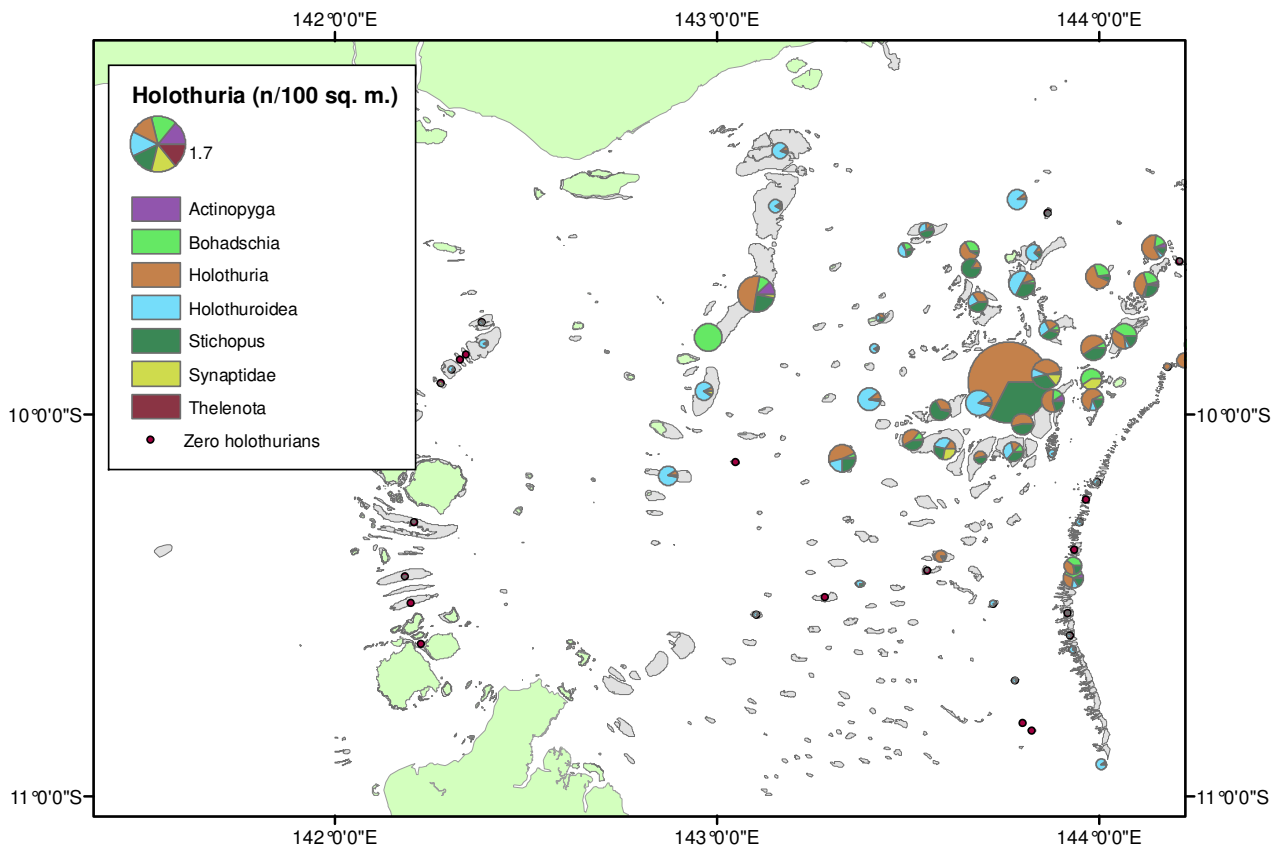


Figure 54. Distribution and abundance of genera of holothurians throughout the Torres Strait.

Seagrass distribution

Seagrasses were identified to species and percentage cover of each transect was estimated by divers. Seagrasses were only found in the western and central parts of the Strait on the Warrior Reef complex. In general, wherever seagrass was found, *Thalassia hemprichii* tended to be the most prevalent species. *Thalassodendron ciliatum* was dominant on the southern Warrior Reefs whereas seagrass communities tended to be more diverse on the northern Warriors with *Thalassia hemprichii* being the dominant species. *Cymodocea rotundata* and *T. hemprichii* dominated on the reefs to the north of Prince of Wales Island while the communities were more diverse to the north around the Orman Reefs.

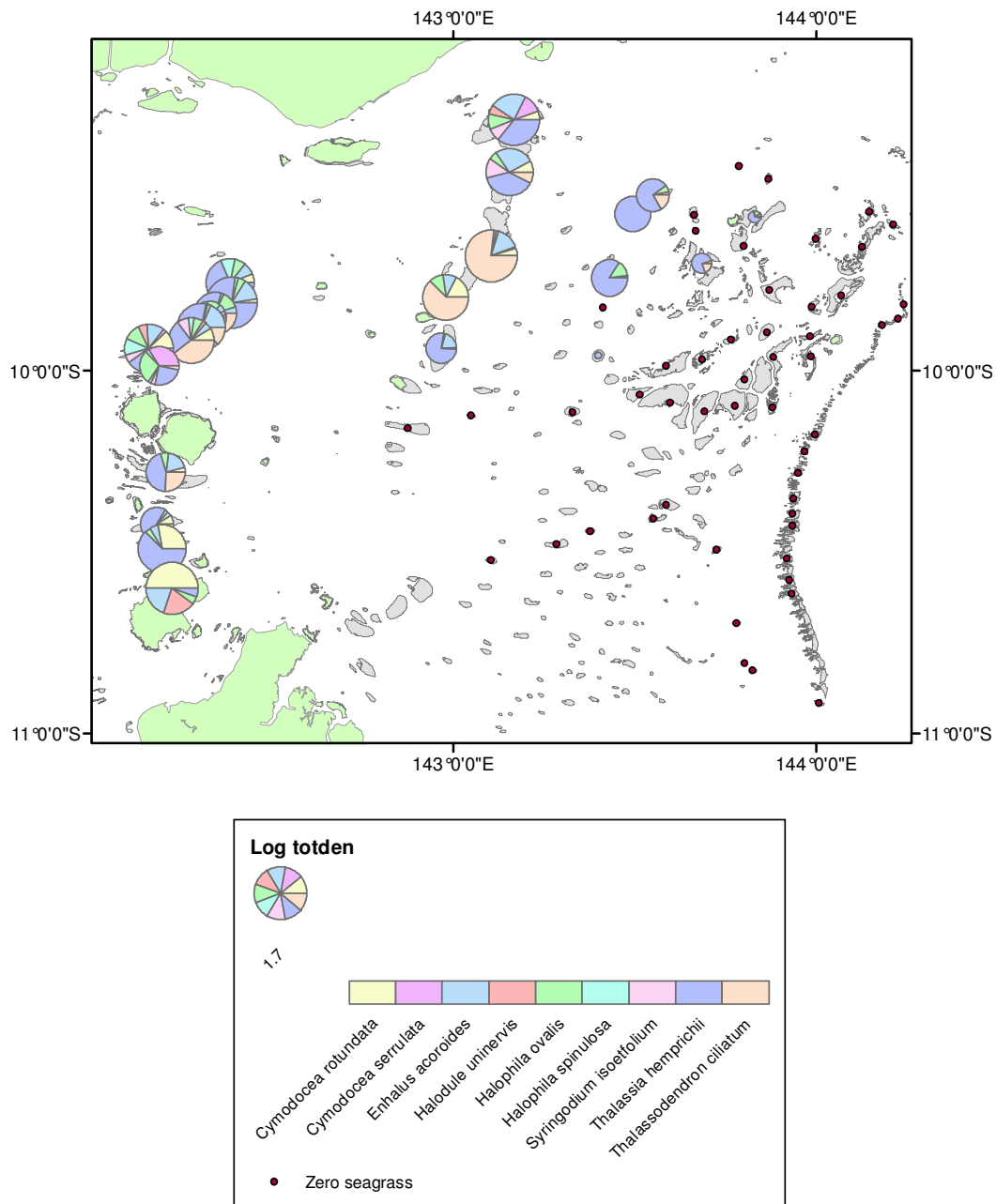


Figure 55. Distribution and abundance of seagrasses throughout Torres Strait. Seagrass cover is expressed as log (percentage cover).

4.3.3. Clustering the maximum number of samples (a)

When the maximum number of samples was clustered by species, the samples were sorted into six clusters with different, though overlapping, geographic distributions (Figure 56). The clusters varied in the taxonomic groups of species associated with them (Table 9). Clusters 1-3 were concentrated in the east. Cluster 1 was characterised by a high occurrence of clams, fungiids and sea cucumbers (particularly *Bohadschia argus* and *Holothuria edulis*); Cluster 2 featured algae and sea cucumbers; and Cluster 3 featured *Halimeda* algae only. Cluster 4 was widespread throughout the central and western Torres Strait, and featured algae (especially *Sargassum* spp.) and seagrass (especially *Thalassia hemprichii*). Cluster 5 was mostly confined to western Torres Strait and Warrior Reef, and featured well mixed communities of seagrasses and algae. Cluster 6 was concentrated on Warrior Reef, and featured a different suite of algal communities and a high occurrence of seagrass (*Enhalus acoroides*). Among the less frequently sampled species groups, crown of thorns starfish (*Acanthaster planci*) and pearl oysters (*Pinctada* spp.) were concentrated in the east in Cluster 1.

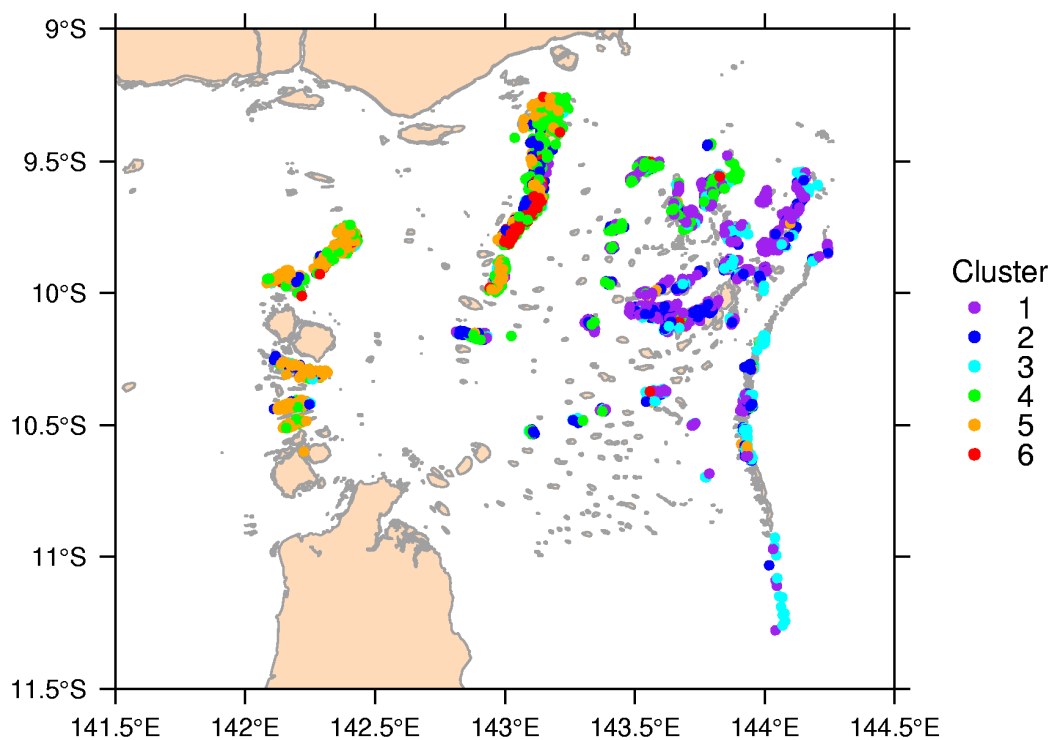


Figure 56. Sample clusters from clustering the maximum number of samples.

Table 9. Association of species groups with sample clusters from clustering the maximum number of samples. For each cluster and species group the table shows the frequency of the species group and the number of samples in which the species group was observed as a percentage of the number of samples from all clusters in which the species group was observed. Zeros in the table represent small non-zero values rounded to zero. True zeros are represented by blanks.

Species group	Class	Frequency						Percent. species group samples					
		1	2	3	4	5	6	1	2	3	4	5	6
All samples		100	100	100	100	100	100	19	15	12	20	20	15
<i>Panulirus</i>	Malacostraca	3	2	2	2	0		38	21	15	24	3	
Fungiidae	Anthozoa	57	5	2	2	0	0	88	7	2	3	1	0
<i>Acanthaster planci</i>	Asteroidea	5	1			0		85	12			4	
Synaptidae	Holothuroidea	1	1		2	2	1	18	11		29	32	11
<i>Actinopyga echinites</i>	Holothuroidea	2	1	1	2	2	3	25	10	5	20	18	22
<i>Actinopyga miliaris</i>	Holothuroidea	3	2	1	1	1	3	27	18	7	11	14	23
<i>Bohadschia argus</i>	Holothuroidea	19	5	2	1	2	1	70	13	4	4	6	2
<i>Bohadschia marmorata</i>	Holothuroidea	0			1	0	2	13			27	13	47
<i>Holothuria atra</i>	Holothuroidea	23	22	13	20	10	12	26	19	9	23	12	11
<i>Holothuria coluber</i>	Holothuroidea	6	2	2	4	3	5	31	8	5	21	14	22
<i>Holothuria edulis</i>	Holothuroidea	38	2	2	1	1	0	88	4	3	2	2	1
<i>Holothuria leucospilota</i>	Holothuroidea	2	12	3	7	3	2	8	36	9	29	11	8
<i>Holothuria nobilis</i>	Holothuroidea	6	4	2	1	0		53	27	10	6	4	
<i>Stichopus chloronotus</i>	Holothuroidea	12	14	6	8	5	1	29	27	10	20	14	1
<i>Stichopus variegatus</i>	Holothuroidea	32	4	2	4	3	4	70	6	2	9	7	6
Tridacnidae	Bivalvia	70	9	7	7	2	1	76	7	5	8	3	1
<i>Pinctada</i>	Bivalvia	7	1					94	6				
<i>Trochus</i>	Gastropoda	3	2	1	1		0	50	25	8	12		4
<i>Caulerpa</i>	Chlorophyceae	9	81		45	67	12	4	32		24	35	5
<i>Halimeda</i>	Chlorophyceae	39	64	100	19	65	19	15	20	24	8	27	6
<i>Udotea</i>	Chlorophyceae	2	2	0	1	16	2	7	5	1	6	74	6
<i>Boodlea</i>	Chlorophyceae	0	1	0	3	9	19	1	3	1	12	33	50
<i>Dictyota</i>	Phaeophyceae	3	1	2	3	10	4	13	4	6	13	49	14
<i>Padina</i>	Phaeophyceae	3	5	3	52	16	29	3	3	2	53	17	22
<i>Sargassum</i>	Phaeophyceae	6	9	9	82	27	27	4	4	4	56	19	14
<i>Turbinaria ornata</i>	Phaeophyceae	4	6	5	20	5	8	9	11	7	46	12	14
<i>Laurencia</i>	Rhodophyceae	3	5	3	6	55	68	2	3	1	5	46	42
<i>Gracilaria</i>	Rhodophyceae	3	5	8	5	51	13	3	5	6	7	66	12
<i>Enhalus acoroides</i>	Liliopsida		4	1	9	53	82		3	0	7	42	48
<i>Halophila ovalis</i>	Liliopsida	7	2	3	15	51	5	8	2	3	19	64	5
<i>Halophila spinulosa</i>	Liliopsida	1	1	0	3	6		8	6	2	27	56	
<i>Thalassia hemprichii</i>	Liliopsida	0	5	3	49	73	2	0	3	1	37	56	1
<i>Cymodocea rotundata</i>	Liliopsida		2		3	16	18		5		9	48	38
<i>Cymodocea serrulata</i>	Liliopsida		0		2	6	2		2		24	61	13
<i>Halodule uninervis</i>	Liliopsida	0	1		2	5	2	2	9		20	51	18
<i>Syringodium isoetifolium</i>	Liliopsida		1		2	9	1		4		17	76	4
<i>Thalassodendron ciliatum</i>	Liliopsida	1	9	3	12	20	93	1	6	2	11	18	63

4.3.4. Clustering by the maximum number of species (b)

When samples were clustered by the maximum number of species, the samples were sorted into five clusters with distinct geographical distributions (Figure 57). Cluster 1 was concentrated between Warrior Reef and the Ribbon Reefs, Cluster 2 was concentrated on the Ribbon Reefs, Clusters 3 and 4 were western, and Cluster 5 was concentrated around Warrior Reef. Fishes were most frequent in Clusters 1 and 2, and were infrequent in Clusters 3 and 5 (Table 10). Cluster 1 featured *Chelmon marginalis* (marginated coralfish), *Pomacentrus taeniometopon* (brackish damsel), *Chrysiptera rollandi* (bluehead demoiselle), *Scarus flavipectoralis* (yellowfin parrotfish), *Scarus rivulatus* (surf parrotfish) and *Plectropomus leopardus* (common coral trout). Cluster 2 featured *Halichoeres hortulanus* (checkerboard wrasse), *Acanthurus nigrofusus* (surgeonfish), *Zanclus cornutus* (moorish idol), *Gomphosus varius* (birdnose wrasse) and *Naso tuberosus* (squarenose unicornfish). Cluster 3 featured *Lethrinus harak* (thumbprint emperor). Cluster 4 featured *Pomacentrus pavo* (peacock damsel), *Chelmon rostratus* (beaked coralfish), *Lutjanus carponotatus* (stripey snapper), *Abudefduf bengalensis* (Bengal sergeant) and *Plectropomus maculatus* (barcheek coral trout). Cluster 5 featured *Acanthurus blochii* (dark surgeonfish), *Scarus ghobban* (blue-barred parrotfish) and *Abudefduf bengalensis* (Bengal sergeant). Corals were concentrated in Cluster 1 and infrequent in Cluster 3. Clams, pearl oysters and crown of thorns starfish were concentrated in Cluster 1. Sea cucumbers were most frequent in Clusters 1 and 4, and were absent from Cluster 5. Algae were concentrated in Clusters 3-5, although *Halimeda* was most frequent in Cluster 2. Seagrasses were concentrated and frequent in Cluster 3, frequent in Cluster 4 and rare in the remainder.

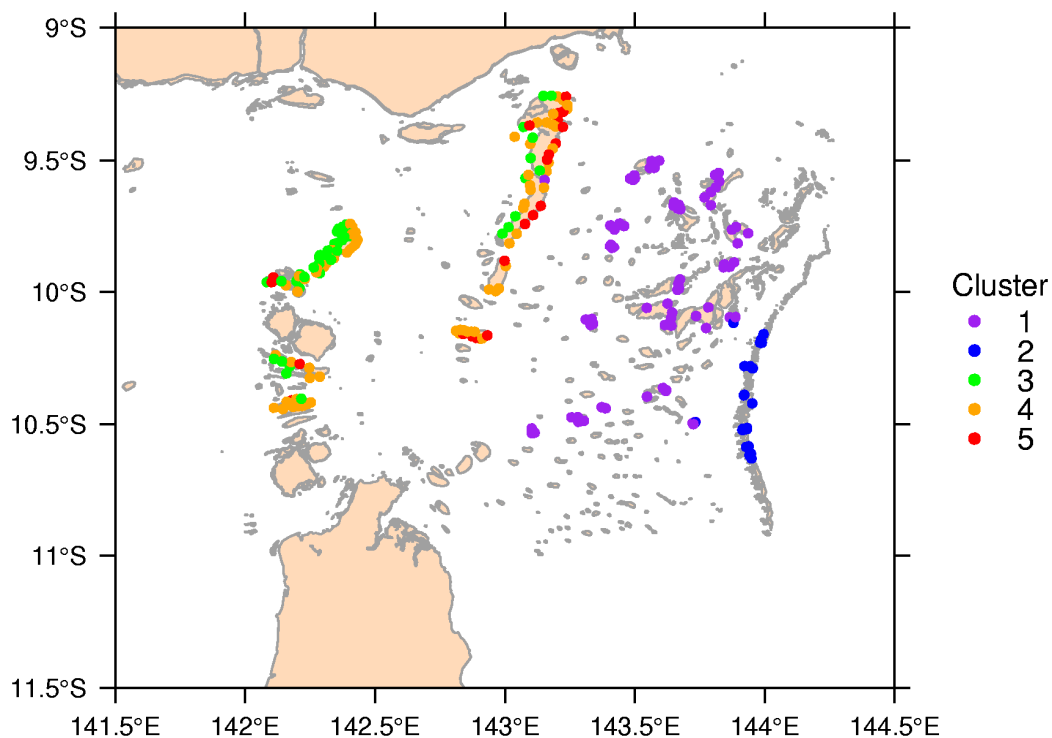


Figure 57. Sample clusters from clustering by the maximum number of species.

Table 10. Association of species groups with sample clusters from clustering by the maximum number of species. For each cluster and species group the table shows the frequency of the species group and the number of samples in which the species group was observed as a percentage of the number of samples from all clusters in which the species group was observed. Zeros in the table represent small non-zero values rounded to zero. True zeros are represented by blanks.

Species group	Class	Family	Frequency					Percent. sp. group samples				
			1	2	3	4	5	1	2	3	4	5
All samples			100	100	100	100	100	31	9	19	31	9
<i>Panulirus</i>	Malacostraca	Palinuridae	3		5	10	10	15		15	54	15
<i>Acanthurus blochii</i>	Actinopterygii	Acanthuridae	93	52	9	71	67	46	7	3	35	10
<i>Acanthurus dussumieri</i>	Actinopterygii	Acanthuridae	6					100				
<i>Acanthurus lineatus</i>	Actinopterygii	Acanthuridae	4	29		1		30	60		10	
<i>Acanthurus mata</i>	Actinopterygii	Acanthuridae	1	5				50	50			
<i>Acanthurus nigricans</i>	Actinopterygii	Acanthuridae		5					100			
<i>Acanthurus nigricauda</i>	Actinopterygii	Acanthuridae	32	67		1		61	37		3	
<i>Acanthurus nigrofuscus</i>	Actinopterygii	Acanthuridae	4	81				15	85			
<i>Acanthurus olivaceus</i>	Actinopterygii	Acanthuridae	3	48				17	83			
<i>Acanthurus pyroferus</i>	Actinopterygii	Acanthuridae		38					100			
<i>Acanthurus thompsoni</i>	Actinopterygii	Acanthuridae		5					100			
<i>Acanthurus triostegus</i>	Actinopterygii	Acanthuridae		10					100			
<i>Acanthurus xanthopterus</i>	Actinopterygii	Acanthuridae	8	19				60	40			
<i>Ctenochaetus binotatus</i>	Actinopterygii	Acanthuridae	21	57				56	44			
<i>Ctenochaetus striatus</i>	Actinopterygii	Acanthuridae	22	76				50	50			
<i>Naso annulatus</i>	Actinopterygii	Acanthuridae	11	24				62	38			
<i>Naso brevirostris</i>	Actinopterygii	Acanthuridae	6	48				29	71			
<i>Naso lituratus</i>	Actinopterygii	Acanthuridae	3	48				17	83			
<i>Naso tuberosus</i>	Actinopterygii	Acanthuridae		62					100			
<i>Naso unicornis</i>	Actinopterygii	Acanthuridae	40	38				78	22			
<i>Naso vlamingii</i>	Actinopterygii	Acanthuridae	6	10				67	33			
<i>Paracanthurus hepatus</i>	Actinopterygii	Acanthuridae		29					100			
<i>Zebrasoma scopas</i>	Actinopterygii	Acanthuridae	15	76				41	59			
<i>Zebrasoma veliferum</i>	Actinopterygii	Acanthuridae	25	43				67	33			
<i>Siganus argenteus</i>	Actinopterygii	Siganidae	15	33				61	39			
<i>Siganus corallinus</i>	Actinopterygii	Siganidae	38	33	2			77	20	3		
<i>Siganus doliatus</i>	Actinopterygii	Siganidae	83			18		82			18	
<i>Siganus javus</i>	Actinopterygii	Siganidae	4			1		75			25	
<i>Siganus lineatus</i>	Actinopterygii	Siganidae	54		2	11		81		2	17	
<i>Siganus puellus</i>	Actinopterygii	Siganidae	53	14		6		84	7		9	
<i>Siganus punctatus</i>	Actinopterygii	Siganidae	50	43		1		78	20		2	
<i>Siganus spinus</i>	Actinopterygii	Siganidae		5					100			
<i>Siganus vulpinus</i>	Actinopterygii	Siganidae	8	29				50	50			
<i>Zanclus cornutus</i>	Actinopterygii	Zanclidae	4	81				15	85			
<i>Cheilinus undulatus</i>	Actinopterygii	Labridae	14	67				42	58			
<i>Choerodon fasciatus</i>	Actinopterygii	Labridae	54	19		6		83	9		9	
<i>Coris aygula</i>	Actinopterygii	Labridae		43		3			82		18	
<i>Coris gaimard</i>	Actinopterygii	Labridae	3	62				13	87			
<i>Epibulus insidiator</i>	Actinopterygii	Labridae	72	62				80	20			
<i>Gomphosus varius</i>	Actinopterygii	Labridae	4	81				15	85			
<i>Halichoeres hortulanus</i>	Actinopterygii	Labridae	6	95				17	83			
<i>Halichoeres trimaculatus</i>	Actinopterygii	Labridae	1	29				14	86			
<i>Hemigymnus fasciatus</i>	Actinopterygii	Labridae	21	43		1		60	36		4	

Species group	Class	Family	Frequency					Percent. sp. group samples				
			1	2	3	4	5	1	2	3	4	5
<i>Hemigymnus melapterus</i>	Actinopterygii	Labridae	96	67		50	10	57	12		30	2
<i>Thalassoma hardwicke</i>	Actinopterygii	Labridae	11	76				33	67			
<i>Thalassoma janseni</i>	Actinopterygii	Labridae	1	57				8	92			
<i>Thalassoma lunare</i>	Actinopterygii	Labridae	82	62	5	32		61	13	2	24	
<i>Thalassoma lutescens</i>	Actinopterygii	Labridae		29					100			
<i>Abudefduf bengalensis</i>	Actinopterygii	Pomacentridae	40		14	64	43	32		7	51	10
<i>Abudefduf sexfasciatus</i>	Actinopterygii	Pomacentridae	11	24				62	38			
<i>Abudefduf vaigiensis</i>	Actinopterygii	Pomacentridae	4	10				60	40			
<i>Abudefduf whitleyi</i>	Actinopterygii	Pomacentridae	12			4		75			25	
<i>Acanthochromis polyacanthus</i>	Actinopterygii	Pomacentridae	88	81	5	57	5	51	14	2	33	1
<i>Amblyglyphidodon curacao</i>	Actinopterygii	Pomacentridae	93	33		4		87	9		4	
<i>Amblyglyphidodon leucogaster</i>	Actinopterygii	Pomacentridae	61	62				77	23			
<i>Chromis acares</i>	Actinopterygii	Pomacentridae		24					100			
<i>Chromis amboinensis</i>	Actinopterygii	Pomacentridae	4	14				50	50			
<i>Chromis atripectoralis</i>	Actinopterygii	Pomacentridae	43	67				69	31			
<i>Chromis atripes</i>	Actinopterygii	Pomacentridae		29					100			
<i>Chromis chrysurus</i>	Actinopterygii	Pomacentridae		10					100			
<i>Chromis iomelas</i>	Actinopterygii	Pomacentridae		5					100			
<i>Chromis lepidolepis</i>	Actinopterygii	Pomacentridae	1	62		1		7	87		7	
<i>Chromis margaritifer</i>	Actinopterygii	Pomacentridae		52		4			79		21	
<i>Chromis nitida</i>	Actinopterygii	Pomacentridae	18					100				
<i>Chromis retrofasciata</i>	Actinopterygii	Pomacentridae	3	14				40	60			
<i>Chromis ternatensis</i>	Actinopterygii	Pomacentridae	8	62				32	68			
<i>Chromis vanderbilii</i>	Actinopterygii	Pomacentridae		14					100			
<i>Chromis viridis</i>	Actinopterygii	Pomacentridae	4	10		1		50	33		17	
<i>Chromis weberi</i>	Actinopterygii	Pomacentridae	6	43				31	69			
<i>Chromis xanthura</i>	Actinopterygii	Pomacentridae	1	14	2			20	60	20		
<i>Chrysiptera cyanea</i>	Actinopterygii	Pomacentridae		10	2	15	5		13	7	73	7
<i>Chrysiptera flavipinnis</i>	Actinopterygii	Pomacentridae	32	19				85	15			
<i>Chrysiptera rex</i>	Actinopterygii	Pomacentridae		10		1			67		33	
<i>Chrysiptera rollandi</i>	Actinopterygii	Pomacentridae	89	19		1		93	6		1	
<i>Chrysiptera talboti</i>	Actinopterygii	Pomacentridae	32	52		1		66	31		3	
<i>Dascyllus aruanus</i>	Actinopterygii	Pomacentridae	11	43				47	53			
<i>Dascyllus reticulatus</i>	Actinopterygii	Pomacentridae	24	71				53	47			
<i>Dascyllus trimaculatus</i>	Actinopterygii	Pomacentridae	6	43				31	69			
<i>Neopomacentrus azysron</i>	Actinopterygii	Pomacentridae	76	10	2	7		87	3	2	8	
<i>Neopomacentrus bankieri</i>	Actinopterygii	Pomacentridae	71		2	29	10	68		1	28	3
<i>Neopomacentrus cyanomos</i>	Actinopterygii	Pomacentridae	39					100				
<i>Plectroglyphidodon dickii</i>	Actinopterygii	Pomacentridae		29		1			86		14	
<i>Plectroglyphidodon johnstonianus</i>	Actinopterygii	Pomacentridae		29					100			
<i>Plectroglyphidodon lacrymatus</i>	Actinopterygii	Pomacentridae	7	81				23	77			
<i>Pomacentrus amboinensis</i>	Actinopterygii	Pomacentridae	92	67	14	12	14	67	14	6	9	3
<i>Pomacentrus bankanensis</i>	Actinopterygii	Pomacentridae	1	71	7	10		4	58	12	27	

Species group	Class	Family	Frequency					Percent. sp. group samples				
			1	2	3	4	5	1	2	3	4	5
<i>Pomacentrus brachialis</i>	Actinopterygii	Pomacentridae	88	67				82	18			
<i>Pomacentrus chrysurus</i>	Actinopterygii	Pomacentridae	4		7	11		21		21	57	
<i>Pomacentrus coelestis</i>	Actinopterygii	Pomacentridae	56	52		3	5	74	20		4	2
<i>Pomacentrus grammorhynchus</i>	Actinopterygii	Pomacentridae	29	14				88	12			
<i>Pomacentrus lepidogenys</i>	Actinopterygii	Pomacentridae	36	67		1		63	34		2	
<i>Pomacentrus moluccensis</i>	Actinopterygii	Pomacentridae	99	57		71	5	53	9		38	1
<i>Pomacentrus nagasakiensis</i>	Actinopterygii	Pomacentridae	79		5	8		88		3	9	
<i>Pomacentrus pavo</i>	Actinopterygii	Pomacentridae	3	10	12	44	5	5	5	12	76	2
<i>Pomacentrus philippinus</i>	Actinopterygii	Pomacentridae	1	14				25	75			
<i>Pomacentrus reidi</i>	Actinopterygii	Pomacentridae	12	14				75	25			
<i>Pomacentrus taeniometopon</i>	Actinopterygii	Pomacentridae	89	10				97	3			
<i>Pomacentrus vaiuli</i>	Actinopterygii	Pomacentridae		14	9	10			21	29	50	
<i>Pomacentrus wardi</i>	Actinopterygii	Pomacentridae	38		7	39	10	45		5	47	3
<i>Stegastes apicalis</i>	Actinopterygii	Pomacentridae	24			3	5	85			10	5
<i>Stegastes fasciolatus</i>	Actinopterygii	Pomacentridae		29					100			
<i>Stegastes nigricans</i>	Actinopterygii	Pomacentridae	3	43		3		15	69		15	
<i>Cetoscarus bicolor</i>	Actinopterygii	Scaridae	38	52				71	29			
<i>Chlorurus bleekeri</i>	Actinopterygii	Scaridae	40	10		3		88	6		6	
<i>Chlorurus japanensis</i>	Actinopterygii	Scaridae	3	5				67	33			
<i>Chlorurus microrhinos</i>	Actinopterygii	Scaridae	72	57				81	19			
<i>Chlorurus sordidus</i>	Actinopterygii	Scaridae	53	86				68	32			
<i>Hipposcarus longiceps</i>	Actinopterygii	Scaridae	15	52				50	50			
<i>Scarus altipinnis</i>	Actinopterygii	Scaridae	24	62				57	43			
<i>Scarus chameleon</i>	Actinopterygii	Scaridae	11	62				38	62			
<i>Scarus dimidiatus</i>	Actinopterygii	Scaridae	8	19				60	40			
<i>Scarus flavipectoralis</i>	Actinopterygii	Scaridae	88	14		1		94	4		1	
<i>Scarus forsteni</i>	Actinopterygii	Scaridae		19		1			80		20	
<i>Scarus frenatus</i>	Actinopterygii	Scaridae	11	81	2	7	5	25	53	3	16	3
<i>Scarus ghobban</i>	Actinopterygii	Scaridae	90	14	14	25	48	64	3	6	18	10
<i>Scarus globiceps</i>	Actinopterygii	Scaridae	15	43		11	5	38	31		28	3
<i>Scarus niger</i>	Actinopterygii	Scaridae	69	71				77	23			
<i>Scarus oviceps</i>	Actinopterygii	Scaridae	6	52				27	73			
<i>Scarus psittacus</i>	Actinopterygii	Scaridae	14	67	2	28	19	20	29	2	41	8
<i>Scarus rivulatus</i>	Actinopterygii	Scaridae	82	5				98	2			
<i>Scarus rubroviolaceus</i>	Actinopterygii	Scaridae		33		1	5		78		11	11
<i>Scarus schlegeli</i>	Actinopterygii	Scaridae	22	48		1		59	37		4	
<i>Scarus spinus</i>	Actinopterygii	Scaridae	3	57				14	86			
<i>Caesio caerulea</i>	Actinopterygii	Caesionidae	40	24				85	15			
<i>Caesio cuning</i>	Actinopterygii	Caesionidae	82	5		22		78	1		21	
<i>Caesio teres</i>	Actinopterygii	Caesionidae		14		21	10		15		75	10
<i>Pterocaesio marri</i>	Actinopterygii	Caesionidae	79	62		3		79	18		3	
<i>Pterocaesio tile</i>	Actinopterygii	Caesionidae		29					100			
<i>Pterocaesio trilineata</i>	Actinopterygii	Caesionidae		10					100			
<i>Chaetodon aureofasciatus</i>	Actinopterygii	Chaetodontidae	97	10		44	14	65	2		30	3
<i>Chaetodon auriga</i>	Actinopterygii	Chaetodontidae	31	62		6		56	33		10	
<i>Chaetodon baronessa</i>	Actinopterygii	Chaetodontidae	40	52		1		71	27		2	
<i>Chaetodon bennetti</i>	Actinopterygii	Chaetodontidae	3					100				

Species group	Class	Family	Frequency					Percent. sp. group samples				
			1	2	3	4	5	1	2	3	4	5
<i>Chaetodon citrinellus</i>	Actinopterygii	Chaetodontidae		52					100			
<i>Chaetodon ephippium</i>	Actinopterygii	Chaetodontidae	4	38				27	73			
<i>Chaetodon kleinii</i>	Actinopterygii	Chaetodontidae	3	62				13	87			
<i>Chaetodon lineolatus</i>	Actinopterygii	Chaetodontidae	21	19		1		75	20		5	
<i>Chaetodon lunula</i>	Actinopterygii	Chaetodontidae	1	5				50	50			
<i>Chaetodon melannotus</i>	Actinopterygii	Chaetodontidae	39	48		1	5	70	25		2	2
<i>Chaetodon ornatissimus</i>	Actinopterygii	Chaetodontidae		14					100			
<i>Chaetodon pelewensis</i>	Actinopterygii	Chaetodontidae		52					100			
<i>Chaetodon plebeius</i>	Actinopterygii	Chaetodontidae	62	95		25	5	54	24		21	1
<i>Chaetodon rafflesii</i>	Actinopterygii	Chaetodontidae		10					100			
<i>Chaetodon rainfordi</i>	Actinopterygii	Chaetodontidae	28			3		91			9	
<i>Chaetodon speculum</i>	Actinopterygii	Chaetodontidae	10	19				64	36			
<i>Chaetodon trifascialis</i>	Actinopterygii	Chaetodontidae	4	71				17	83			
<i>Chaetodon trifasciatus</i>	Actinopterygii	Chaetodontidae	42	76		11		56	30		15	
<i>Chaetodon ulietensis</i>	Actinopterygii	Chaetodontidae	14	29				62	38			
<i>Chaetodon unimaculatus</i>	Actinopterygii	Chaetodontidae	1	14		3		17	50		33	
<i>Chaetodon vagabundus</i>	Actinopterygii	Chaetodontidae	39	52				72	28			
<i>Chelmon marginalis</i>	Actinopterygii	Chaetodontidae	96	10				97	3			
<i>Chelmon muelleri</i>	Actinopterygii	Chaetodontidae			7					100		
<i>Chelmon rostratus</i>	Actinopterygii	Chaetodontidae	6		47	61	29	5		27	59	8
<i>Coradion chrysozonus</i>	Actinopterygii	Chaetodontidae	39					100				
<i>Forcipiger flavissimus</i>	Actinopterygii	Chaetodontidae		33					100			
<i>Forcipiger longirostris</i>	Actinopterygii	Chaetodontidae		5					100			
<i>Parachaetodon ocellatus</i>	Actinopterygii	Chaetodontidae	10		7	3		58		25	17	
<i>Lethrinus atkinsoni</i>	Actinopterygii	Lethrinidae	39	48		1		72	26		3	
<i>Lethrinus erythropterus</i>	Actinopterygii	Lethrinidae	53	10				95	5			
<i>Lethrinus harak</i>	Actinopterygii	Lethrinidae	1		42	14	5	3		60	33	3
<i>Lethrinus laticaudis</i>	Actinopterygii	Lethrinidae	40	5		17		69	2		29	
<i>Lethrinus lentjan</i>	Actinopterygii	Lethrinidae	8		5	1		67		22	11	
<i>Lethrinus obsoletus</i>	Actinopterygii	Lethrinidae	1	5	2	3		20	20	20	40	
<i>Lethrinus olivaceus</i>	Actinopterygii	Lethrinidae	7	10		1		62	25		12	
<i>Lethrinus ornatus</i>	Actinopterygii	Lethrinidae	1					100				
<i>Lethrinus rubrioperculatus</i>	Actinopterygii	Lethrinidae		24					100			
<i>Lethrinus xanthochilus</i>	Actinopterygii	Lethrinidae		19					100			
<i>Lutjanus biguttatus</i>	Actinopterygii	Lutjanidae		5					100			
<i>Lutjanus bohar</i>	Actinopterygii	Lutjanidae	12	81				35	65			
<i>Lutjanus carponotatus</i>	Actinopterygii	Lutjanidae	96	5	44	78	24	46	1	13	37	3
<i>Lutjanus fulviflamma</i>	Actinopterygii	Lutjanidae	24	24	2	10	5	55	16	3	23	3
<i>Lutjanus fulvus</i>	Actinopterygii	Lutjanidae		5		1			50		50	
<i>Lutjanus gibbus</i>	Actinopterygii	Lutjanidae	8	24				55	45			
<i>Lutjanus kasmira</i>	Actinopterygii	Lutjanidae		5					100			
<i>Lutjanus monostigma</i>	Actinopterygii	Lutjanidae	1					100				
<i>Lutjanus quinquelineatus</i>	Actinopterygii	Lutjanidae	7	14				62	38			
<i>Lutjanus rivulatus</i>	Actinopterygii	Lutjanidae		14					100			
<i>Lutjanus sebae</i>	Actinopterygii	Lutjanidae	17			3		86			14	
<i>Lutjanus vitta</i>	Actinopterygii	Lutjanidae	6				5	80				20
<i>Symphorus nematophorus</i>	Actinopterygii	Lutjanidae	21			1		94			6	
<i>Chromileptes altivelis</i>	Actinopterygii	Serranidae	31	10		3		85	8		8	

Species group	Class	Family	Frequency					Percent. sp. group samples				
			1	2	3	4	5	1	2	3	4	5
<i>Plectropomus areolatus</i>	Actinopterygii	Serranidae	17					100				
<i>Plectropomus laevis</i>	Actinopterygii	Serranidae	8	43				40	60			
<i>Plectropomus leopardus</i>	Actinopterygii	Serranidae	90	29		1		90	8		1	
<i>Plectropomus maculatus</i>	Actinopterygii	Serranidae	54		14	65	24	40		6	48	5
<i>Variola louti</i>	Actinopterygii	Serranidae		52					100			
Fungiidae	Anthozoa	Fungiidae	12	10		6		60	13		27	
coral branching	Anthozoa		96	67	2	47	24	56	11	1	28	4
coral digitate	Anthozoa		53	76	7	49	43	38	16	3	35	9
coral encrusting	Anthozoa		26	43	2	28	24	35	17	2	37	9
coral foliose	Anthozoa		21					100				
coral heliopora	Anthozoa		1	5				50	50			
coral massive	Anthozoa		92	43	9	54	52	51	7	3	30	9
coral submassive	Anthozoa		67	76		29	14	55	18		24	3
coral tabulate	Anthozoa		53	67		47	48	40	15		35	10
<i>Acanthaster planci</i>	Asteroidea	Acanthasteridae	7	5	2			71	14	14		
Synaptidae	Holothuroidea	Synaptidae			2	3				33	67	
<i>Actinopyga miliaris</i>	Holothuroidea	Holothuriidae	3			3		50			50	
<i>Bohadschia argus</i>	Holothuroidea	Holothuriidae	7					100				
<i>Holothuria atra</i>	Holothuroidea	Holothuriidae	1		2			50		50		
<i>Holothuria coluber</i>	Holothuroidea	Holothuriidae				1					100	
<i>Holothuria edulis</i>	Holothuroidea	Holothuriidae	10			1		88			12	
<i>Stichopus chloronotus</i>	Holothuroidea	Holothuriidae	3	10				50	50			
<i>Stichopus variegatus</i>	Holothuroidea	Holothuriidae	7			3		71			29	
Tridacnidae	Bivalvia	Tridacnidae	12	10		11	10	43	10		38	10
<i>Pinctada</i>	Bivalvia	Pteriidae	4					100				
<i>Trochus</i>	Gastropoda	Trochidae	17				10	86				14
<i>Caulerpa</i>	Chlorophyceae	Caulerpaceae	3	19	70	39	19	3	6	44	41	6
<i>Halimeda</i>	Chlorophyceae	Udoteaceae	25	76	26	29	33	25	22	15	29	10
<i>Padina</i>	Phaeophyceae	Dictyotaceae	4		30	39	76	5		22	47	27
<i>Sargassum</i>	Phaeophyceae	Sargassaceae	1		49	43	62	2		32	47	20
<i>Turbinaria ornata</i>	Phaeophyceae	Sargassaceae	12	5	7	22	71	20	2	7	36	34
<i>Laurencia</i>	Rhodophyceae	Rhodomelaceae			23	14	10			45	45	9
<i>Amphiroa</i>	Rhodophyceae	Corallinaceae				3					100	
<i>Gracilaria</i>	Rhodophyceae	Gracilariaceae				1					100	
<i>Enhalus acoroides</i>	Liliopsida	Hydrocharitaceae			26	6	5			69	25	6
<i>Halophila ovalis</i>	Liliopsida	Hydrocharitaceae	3		60	10		6		74	20	
<i>Halophila spinulosa</i>	Liliopsida	Hydrocharitaceae			47	3				91	9	
<i>Thalassia hemprichii</i>	Liliopsida	Hydrocharitaceae			65	17				70	30	
<i>Cymodocea rotundata</i>	Liliopsida	Cymodoceaceae			16					100		
<i>Cymodocea serrulata</i>	Liliopsida	Cymodoceaceae			21	4				75	25	
<i>Halodule uninervis</i>	Liliopsida	Cymodoceaceae			5					100		
<i>Syringodium isoetifolium</i>	Liliopsida	Cymodoceaceae			19					100		
<i>Thalassodendron ciliatum</i>	Liliopsida	Cymodoceaceae			12	10	5			38	54	8

4.3.5. Clustering by habitat-forming species (c)

When samples were clustered by habitat-forming species they were sorted into five clusters that were geographically interspersed except for Cluster 1, which was concentrated in the west, and Cluster 2, which was concentrated to a lesser degree in the Ribbon Reefs and far east (Figure 58). Cluster 1 featured a high cover of seagrasses and algae (excluding *Halimeda*), and a relatively low cover of corals (Table 11). Cluster 2 featured a high cover of *Halimeda*. Cluster 3 featured a high cover corals. Cluster 4 shared a similar species mix, but with generally lower cover. Cluster 5 featured low covers of all species.

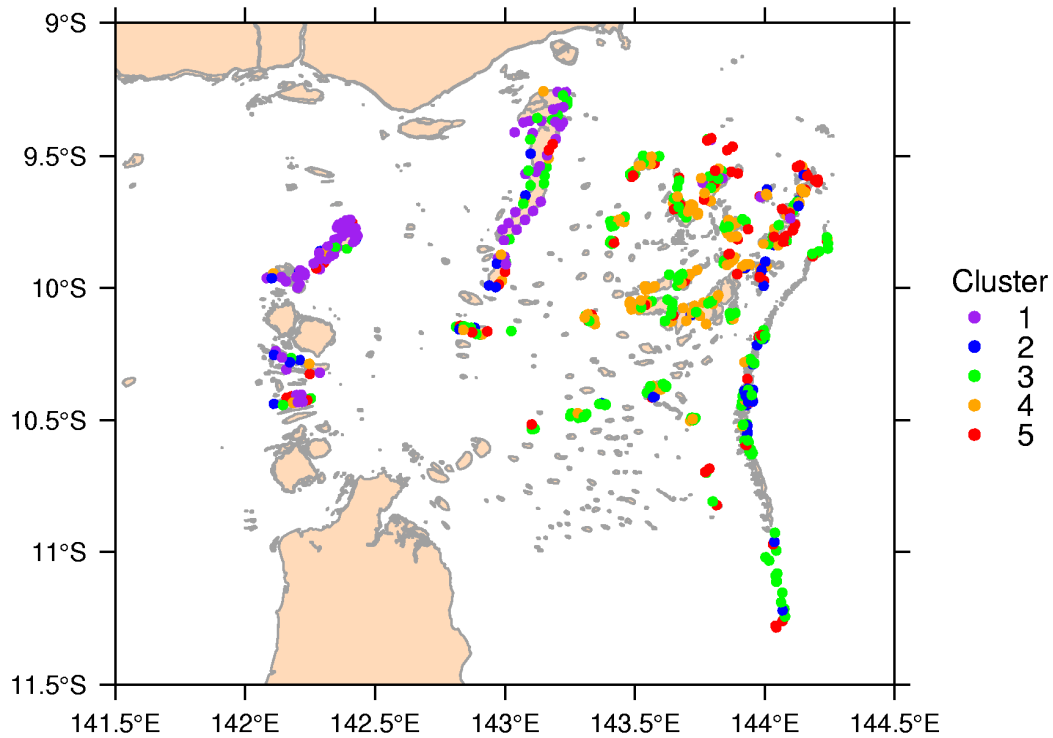


Figure 58. Sample clusters from clustering by habitat-forming species.

Table 11. Association of species groups with sample clusters from clustering by habitat-forming species. For each cluster and species group the table shows (a) the frequency of the species group, (b) the average percentage cover of the species group, (c) the number of samples in which the species group was observed as a percentage of the number of samples from all clusters in which the species group was observed, and (d) the total cover of the species group as a percentage of the total cover of the species group in samples all clusters (calculated as if sample areas were all equal). Zeros in the table represent small non-zero values rounded to zero. True zeros are represented by blanks.

Species group	Class	Frequency					Average cover					Pcnt sp. grp samples					Pcnt sp. grp cover				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
All samples		100	100	100	100	100						16	10	31	27	17					
Fungiidae	Anthozoa	1	5	20	66	6	0	0	0	1	0	1	2	24	69	4	0	4	35	58	3
live coral branching	Anthozoa	13	50	85	92	85	0	1	7	2	1	3	7	36	34	20	1	3	70	19	8
live coral digitate	Anthozoa	15	70	62	54	72	0	1	4	1	1	4	13	35	26	22	1	6	74	9	10
live coral encrusting	Anthozoa	5	57	58	47	55	0	1	2	0	0	2	12	39	28	20	1	12	67	12	9
live coral foliose	Anthozoa		11	17	54	8		0	0	0	0		5	24	66	6		5	49	45	1
live coral massive	Anthozoa	22	70	84	88	89	0	1	7	3	2	5	9	35	32	20	1	3	65	22	8
live coral submassive	Anthozoa	2	30	54	76	15	0	0	2	1	0	1	7	39	48	6	0	3	63	32	2
live coral tabulate	Anthozoa	11	38	64	28	65	0	0	3	0	1	4	9	45	18	25	1	3	81	5	11
<i>Caulerpa</i>	Chlorophyceae	54	46	13	17	10	4	2	0	0	0	36	20	18	19	7	69	22	5	3	1
<i>Chlorodesmis</i>	Chlorophyceae		14	5	5	2		0	0	0	0		32	32	28	8		48	40	9	2
<i>Halimeda</i>	Chlorophyceae	22	100	35	54	27	0	8	1	1	0	8	23	25	33	10	2	65	17	12	4
<i>Udotea</i>	Chlorophyceae	1		2	3		0		0	0		11		33	56		5		49	46	
<i>Dictyota</i>	Phaeophyceae	1	4	1	4	2	0	0	0	0	0	8	15	15	46	15	12	17	13	25	32
<i>Padina</i>	Phaeophyceae	46	14	7	10	5	2	0	0	0	0	50	10	15	19	6	71	7	6	5	11
<i>Sargassum</i>	Phaeophyceae	64	11	8	7	8	13	0	0	0	0	60	6	15	11	8	92	1	3	2	1
<i>Turbinaria ornata</i>	Phaeophyceae	20	12	8	9	12	2	1	0	0	0	27	11	23	21	18	42	11	20	16	11
<i>Laurencia</i>	Rhodophyceae	20	5	1	3	1	2	0	0	0	0	61	11	7	18	4	88	7	2	1	2
<i>Amphiroa</i>	Rhodophyceae	1	5	2	1	1	0	0	0	0	0	9	27	36	18	9	21	36	23	9	11
<i>Gracilaria</i>	Rhodophyceae		2		1			0		0			50		50			76		24	
<i>Enhalus acoroides</i>	Liliopsida	15	5	1			0	0	0			76	18	6			74	26	0		
<i>Halophila ovalis</i>	Liliopsida	49	7	5	14	5	3	0	0	0	0	52	5	11	26	6	91	3	2	3	0
<i>Halophila spinulosa</i>	Liliopsida	30	2	1	2	1	2	0	0	0	0	81	3	3	9	3	98	2	0	0	0
<i>Thalassia hemprichii</i>	Liliopsida	41	7	1			4	0	0			88	10	2			95	5	1		
<i>Cymodocea rotundata</i>	Liliopsida	8					2					100					100				
<i>Cymodocea serrulata</i>	Liliopsida	13	2				1	0				92	8				98	2			
<i>Halodule uninervis</i>	Liliopsida	5					1					100					100				
<i>Syringodium isoetifolium</i>	Liliopsida	9					1					100					100				
<i>Thalassodendron ciliatum</i>	Liliopsida	9	4		2	2	1	0		0	0	53	13		20	13	66	25		3	5

4.3.6. Clustering by habitat attributes (d)

When samples were clustered by habitat attributes they were sorted into five clusters that were interspersed but had some geographical segregation (Figure 59). Clusters 1 and 2 were concentrated in the east. Cluster 3 was concentrated around Warrior Reef. Clusters 4 and 5 were widespread. Cluster 1 featured pavement and algae (Table 12). Cluster 2 featured a mixture of soft sediment, boulders and rubble and the highest average cover of live coral. Cluster 3 featured soft sediment and seagrass. Cluster 4 featured rubble, soft sediment and algae. Cluster 5 featured soft sediment with little cover of the other habitat elements.

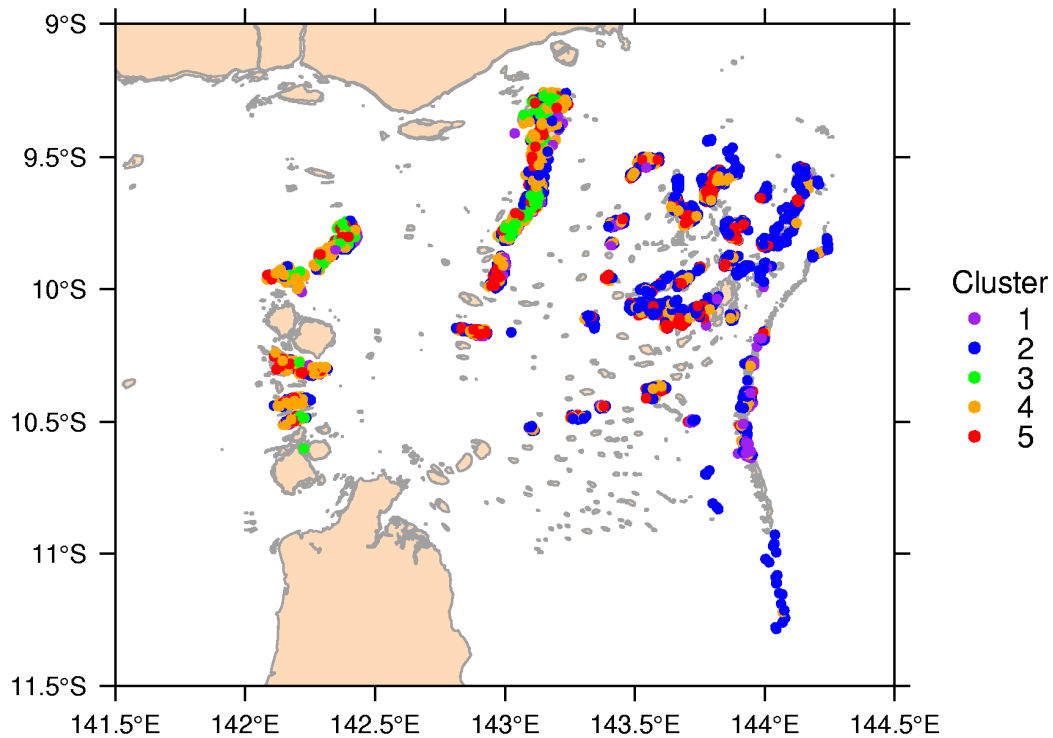


Figure 59. Sample clusters from clustering by habitat attributes.

Table 12. Association of habitat elements with sample clusters from clustering by habitat attributes. For each cluster and habitat element the table shows (a) the frequency of the habitat element, (b) the average percentage cover of the habitat element, (c) the number of samples in which the habitat element was observed as a percentage of the number of samples from all clusters in which the habitat element was observed, and (d) the total cover of the habitat element as a percentage of the total cover of the habitat element in samples all clusters (calculated as if sample areas were all equal). Zeros in the table represent small non-zero values rounded to zero. True zeros are represented by blanks.

Habitat element	Frequency					Average cover					Pcnt element samples					Pcnt element cover				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
All samples	100	100	100	100	100						8	33	14	25	20					
boulders	11	60	17	20	11	1	8	1	1	0	3	65	8	17	8	2	80	4	11	3
rubble	39	89	72	89	45	6	20	9	25	2	4	39	14	30	12	3	44	8	42	3
soft sediment	60	94	99	98	100	11	29	87	63	95	5	32	15	26	22	1	17	21	28	34
consolidated rubble	19	74	22	26	10	3	17	3	7	1	4	65	8	18	5	3	69	6	22	1
hard substrate	5	28	4	6	5	2	6	0	1	0	3	72	5	12	8	6	81	1	10	3
pavement	99	24	4	8	5	69	3	0	1	0	40	40	3	11	5	77	16	0	5	1
live coral	81	95	7	37	26	9	13	0	1	1	12	58	2	18	10	13	80	0	5	3
dead standing coral	2	9		0	0	0	0		0	0	4	91		2	2	4	93		2	1
seagrass	11	12	100	63	38	1	1	66	8	3	2	9	33	38	18	0	3	76	17	4
algae	85	80	95	97	82	31	9	19	38	7	8	30	15	28	19	13	15	14	50	8

4.3.7. Association between clusters from different analyses

When clustering analyses are performed independently on different sets of samples and/or species as above, the samples will be sorted into a different set of clusters in each analysis unless species distributions are highly correlated with one another. Clusters from the four analyses presented here were often associated with one another, but were not equal. For example, clusters from analyses (a) (maximum number of samples, but a reduced set of species) and (d) (habitat attributes) were often associated with those from analysis (b) (maximum number of species). The majority of samples in a given cluster from analyses (a) and (d) were always in 1–2 clusters from analysis (b) (Table 13). The associations were often logical. For example, Cluster 1 in analysis (a) and Cluster 1 in analysis (b) featured clams, fungiids and sea cucumbers; and Clusters 5 and 6 in analysis (a) and Cluster 3 in analysis (b) featured seagrasses and algae. Similarly, Cluster 2 in analysis (d) and Cluster 1 in analysis (b) featured live coral; and Cluster 3 in analysis (d) and Cluster 3 in analysis (b) featured seagrasses.

Table 13. Associations among clusters from different analyses. The table shows the percentage of samples in each cluster from analyses (a) (maximum number of samples) and (d) (habitat attributes) that were in each cluster from analysis (b) (maximum number of species).

Max. spp. clusters	Max. samples clusters						Habitat attribute clusters				
	1	2	3	4	5	6	1	2	3	4	5
1	81	35	44	1	5	0	5	50	0	7	20
2	4	12	35	0	0	0	35	11	0	0	0
3	0	12	0	25	84	67	0	1	92	46	50
4	15	32	12	52	11	17	35	34	8	33	10
5	0	8	9	21	0	17	25	5	0	14	20

4.3.8. Habitat-forming species and fetch

Throughout the Strait, waves are generally wind-driven rather than generated by oceanic swells. This is because of the shallow bathymetry of the region and the fact that the eastern edge of the Strait is protected from Coral Sea swell by the Ribbon reefs. The winds in Torres Strait are predominantly from the southeast during the winter trade season and from the northwest during the summer monsoon (Figure 60). The trade winds are generally stronger (mean July wind speed at 15:00 h for Horn Island = 30.1 kmh^{-1} ; Bureau of Meteorology 2007) than those during the monsoon (mean January wind speed at 15:00 h for Horn Island = 16.8 kmh^{-1} Bureau of Meteorology 2007).

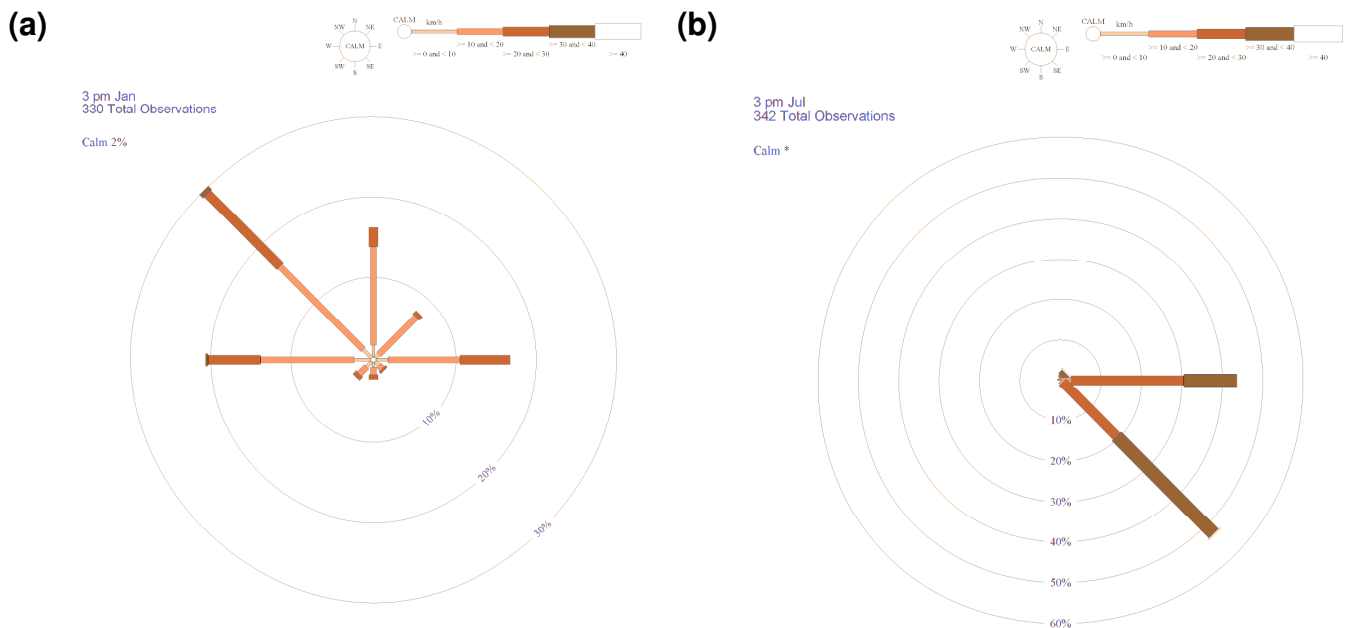


Figure 60. Wind roses for Horn Island showing windspeed and direction for 15:00 h for (a) January and (b) July. Copyright © Bureau of Meteorology, Commonwealth of Australia 2007.

For each of the reef sample sites in our database, we estimated fetch from the southeast and northwest to act as surrogates for exposure in winter and summer respectively. If the abundance of habitat-forming species such as corals, algae and seagrasses are affected by exposure, then sample clusters based on the cover of these species might be associated with high or low values of these indices. Southeast fetch varied significantly among the sample clusters from analysis (c) (habitat-forming species – reef edge sites only) (ANOVA, $F = 6.84$, $p < 0.001$, $n = 560$), but northwest fetch did not ($F = 1.32$, $p = 0.26$, $n = 560$). Clusters 2 and 5 were associated with the highest values of southeast fetch (Figure 61). These clusters featured the lowest covers of habitat-forming species other than *Halimeda* in Cluster 2.

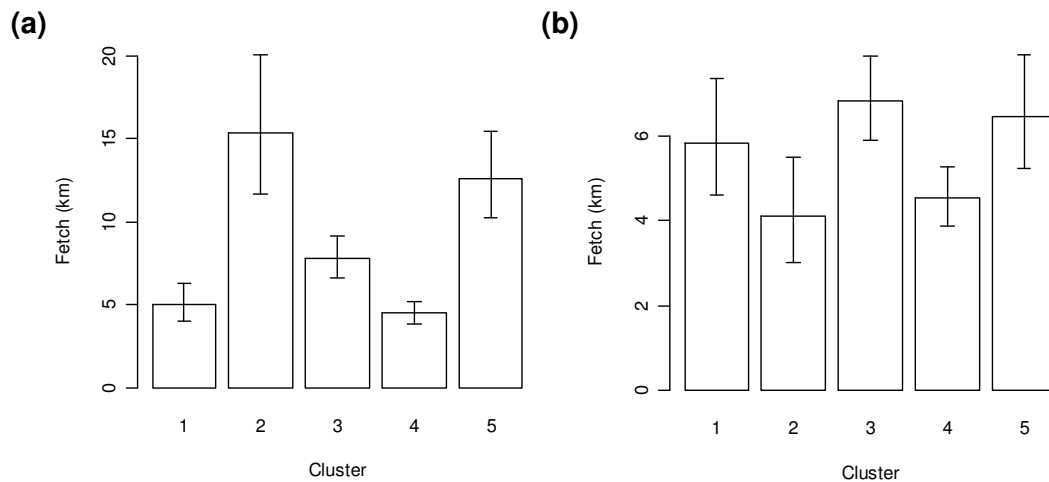


Figure 61. Average (a) south-east and (b) north-west fetch for sample clusters from clustering by habitat-forming species. The error bars indicate standard errors. Means and standard errors were calculated with log-transformed data.

4.3.9. Habitat-forming species and environmental factors

In a canonical correlation analysis (CCA) including all candidate environment variables the inertia of the first three constrained axes accounted for 38%, 21% and 11% (total 70%) of the total inertia of all constrained axes (1.87). Note that these percentages are for comparing among constrained (environmental) axes; they do not indicate the percentage of the total variation in species composition. This analysis is referred to below as the 'full model'.

In the full model, latitude, longitude, silicate and temperature seasonality were relatively redundant and could be dropped with very little effect on sample or species scores. The constrained inertia was then 1.68 (90% of that of the full model). In this model the environment variables most correlated with the first axis were turbidity and soft sediment (positive); and temperature, salinity and wave period (negative) (Figure 62). Those most correlated with the second axis were soft sediment and wave height (positive); and pavement and consolidated rubble (negative). The first axis separated seagrasses (positive) from corals (negative). The second axis tended to separate seagrasses (positive) from algae (negative) (Figure 62). Maps of the LC and WA scores indicated that geographic variation of the composition of habitat-forming species along the first axis effectively represented the general geographic variation of the environment variables along the axis (Figure 63).

The primary relationships revealed by CCAs on specific subsets of the environment variables in the full model were consistent with those revealed by the full model. In a CCA including just latitude and longitude the constrained inertia was 0.71 (38% of that of the full model). The first axis represented a line running from ESE to WNW and separated corals (ESE) from seagrasses (WNW).

In a CCA including just substrate variables the constrained inertia was 0.82 (44% of that of the full model). The first axis represented a gradient from soft sediment to the hard substrates. It separated seagrasses (soft sediment) from corals (hard substrates).

In a CCA including just south-east and north-west fetch the constrained inertia was a low 0.14 (8% of that of the full model). The first axis represented a gradient from north-west fetch to south-east fetch. It separated the seagrasses *Halodule uninervis* and *Thalassodendron ciliatum* (south-east fetch) from the other seagrasses.

In a CCA including only turbidity the constrained inertia was 0.55 (30% of that of the full model). It separated seagrasses (high turbidity, generally soft substrate) from corals (low turbidity, hard substrate).

In a CCA including only the hydrology variables (temperature, salinity and silicate) the constrained inertia was 0.87 (47% of that of the full model). The first axis separated seagrasses (high silicate) from corals (high temperature and salinity).

In a CCA including only hydrodynamics variables (current stress, wave height, wave period, wave direction) the constrained inertia was 0.75 (40% of that of the full model). The first axis separated coral (long wave period, south-easterly waves, low current stress) from seagrasses (short wave period, north-westerly waves, high current stress). The second axis was correlated with wave height.

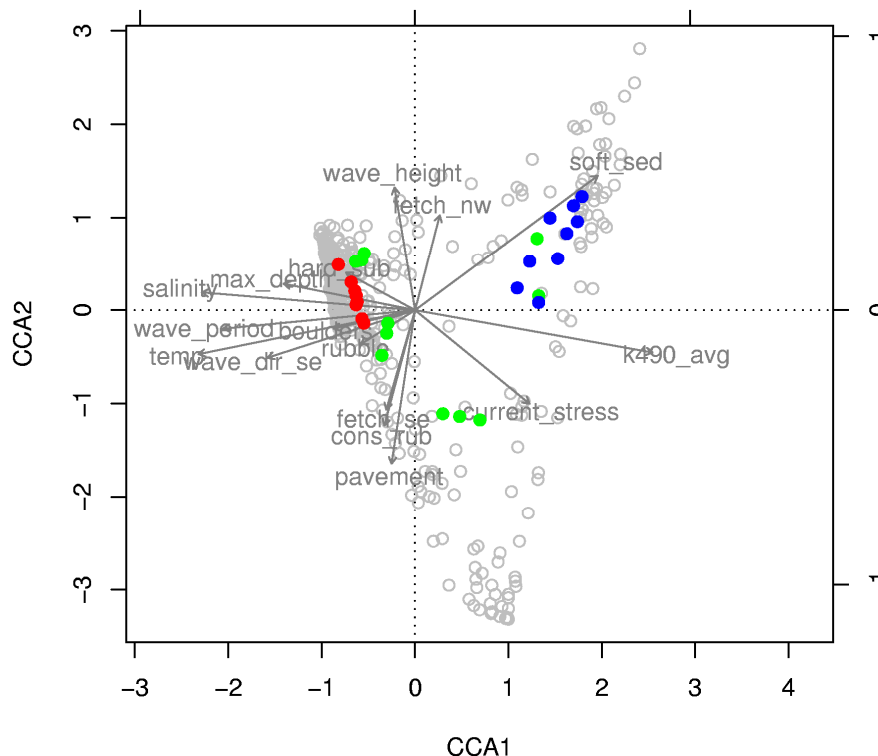


Figure 62. CCA plot for habitat-forming species and all candidate environment variables except latitude, longitude, temperature seasonality and silicate. The grey symbols indicate sample WA scores. The coloured symbols indicate species scores (red, corals; green, algae; blue, seagrasses). The arrows indicate biplot scores, which indicate the correlations between the environment variables and the sample LC scores.

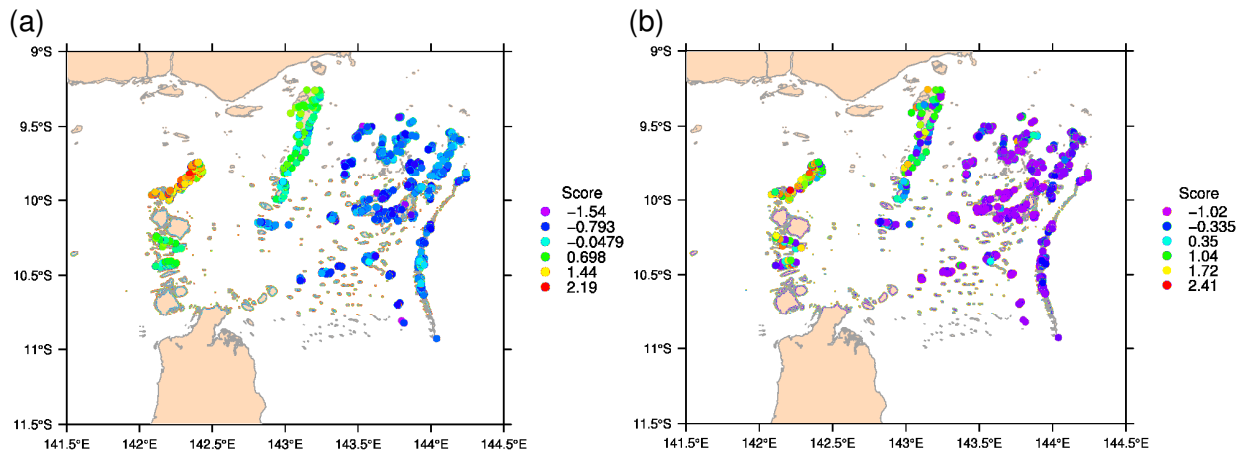


Figure 63. Sample (a) LC and (b) WA scores of the first axis of a CCA of habitat-forming species and all candidate environment variables except latitude, longitude, temperature seasonality and silicate. The sample LC scores are the linear combinations of the environment variables. The WA scores are the weighted averages of the species scores. The correspondence between map (b) and map (a) indicates how well variation in species composition along the axis (b) represents variation in the environment variables along the axis (a).

4.3.10. Fishes and environmental factors

In a CCA of fishes that included all candidate environment variables except latitude, longitude, temperature seasonality and silicate (see above), the inertia of the first three axes accounted for 36%, 26% and 9% (total 71%) of the total inertia of all constrained axes. The variables most correlated with the first axis were turbidity (positive); and wave period and wave direction (negative) (Figure 64). The environmental gradient was correlated with latitude and longitude: sample LC scores increased from the south-east to the north-west (Figure 65). The environment variables most correlated with the second axis were current stress (positive); and salinity (negative). The most obvious trend in species composition was a general decrease in fish frequency from the negative (south-east) to the positive (north-west) end of the axis (Figure 64). There was no general segregation of fishes along the axis at the family level, but there was some variation. For example, acanthurid and labrid species tended to be towards the negative (south-east) end, whereas Lethrinidae species tended to be at the centre or positive (north-west) end (Figure 64).

The CCAs for fish and habitat-forming species shared the first-axis gradient from warm, clear water and long wave period to cool, turbid water, short wave period and soft sediment (although the soft sediment relationship was weaker for fish). The major differences were correlation with salinity for habitat-forming species but not fish, and correlation with pavement for fish but not habitat-forming species. The second axes of the two analyses were less alike, although they both shared a gradient in wave height.

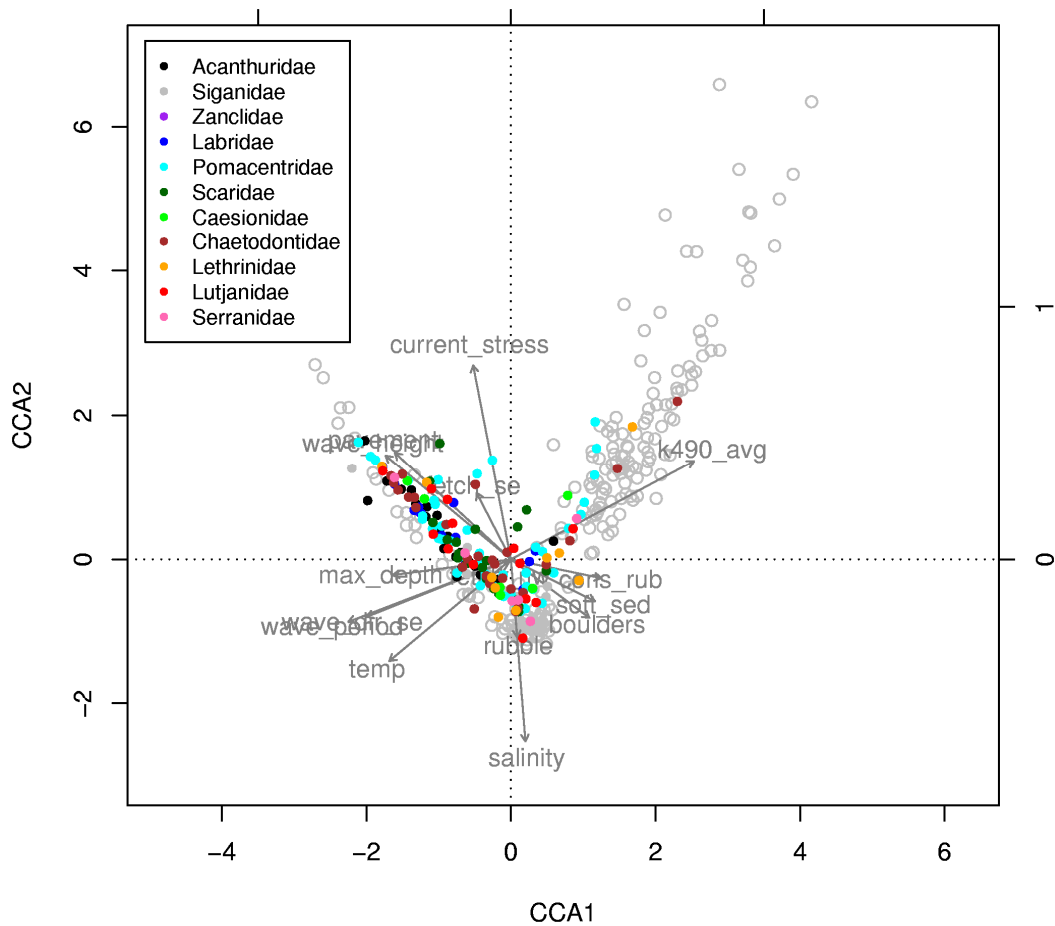


Figure 64. CCA plot for fishes and all candidate environment variables except latitude, longitude, temperature seasonality and silicate. The grey symbols indicate sample WA scores (the weighted averages of the species scores). The coloured symbols indicate species scores (coloured by family). The arrows indicate biplot scores, which indicate the correlations between the environment variables and the LC sample scores (the linear combinations of the environment variables) for each axis.

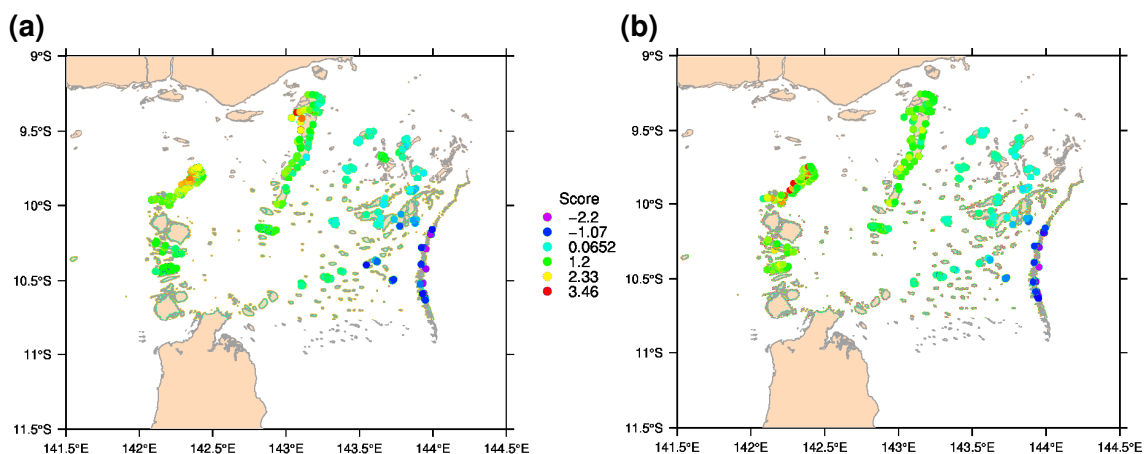


Figure 65. Sample (a) LC and (b) WA scores of the first axis of a CCA of fishes and all candidate environment variables except latitude, longitude, temperature seasonality and silicate. The sample LC scores are the linear combinations of the environment variables. The WA scores are the weighted averages of the species scores. The correspondence between map (b) and map (a) indicates how well the species composition data (b) can represent the variation in the environmental data (a).

4.3.11. Gap analysis

From the cluster analysis of the biologically important physical covariates, Cluster 1 was concentrated primarily around the Orman Reefs in the west; Cluster 2 was a very small cluster almost exclusively on Cook Reef in the southwest TS Cluster 3 was centered on the northern part of the Warrior reef complex; Cluster 4 occurred on the southern Warrior Reefs, Dungeness Reef and the elongated reefs in between Moa and Hammond islands. Cluster 5 occurred in the central section of the Warrior Reef complex; Cluster 6 was rather surprising in that its members occurred in the western TS on fringing reefs around Moa Island, but also on some of the southern Ribbon Reefs; Cluster 7 was the largest cluster and was the dominant cluster in the eastern TS occurring on most of the Ribbon Reefs, but also on most of the reefs just to the west of the Ribbons; Cluster 8 dominated the reefs in the central zone of TS extending from the north in PNG waters to the eastern tip of Cape York in the south.

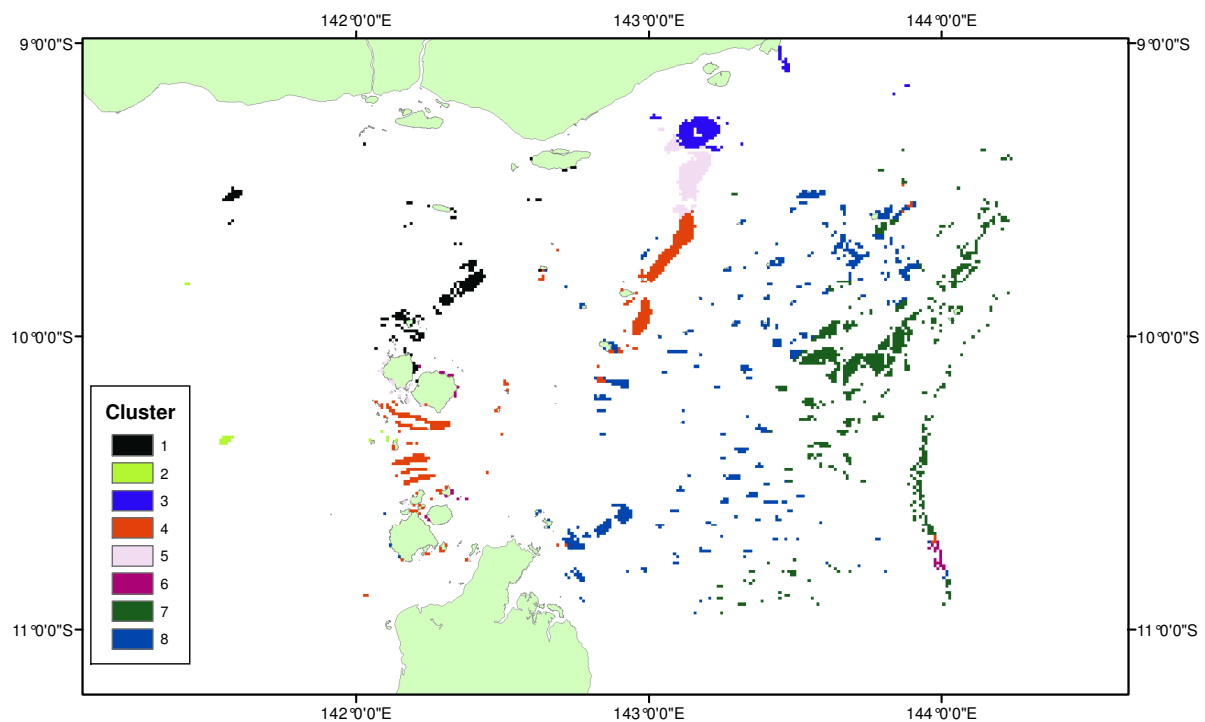


Figure 66. Map of the Torres Strait showing cluster membership of the reefs. K-means clustering was performed on the first 4 dimensions of an nMDS based on a selection of biologically important physical covariates.

Two of the small clusters (Clusters 2 [some of the fringing reef around Moa Island and a small section of the southern part of the Ribbon Reefs] and 6 [Cook Reef]) have not been surveyed at all, whereas the two largest clusters (Clusters 7 and 8) have a large number of surveyed sites within them, whereas the remaining clusters are somewhat intermediate in terms of sampling intensity (Table 14).

Table 14. Area and number of sites surveyed (all surveys combined) within each of the clusters produced by a K-means clustering on the first 4 dimensions of an nMDS based on a selection of biologically important physical covariates. RRI = Reef Resource Inventory.

Cluster id	Area (km ²)	Number of sites (all surveys)	Number of sites (RRI surveys only)
1	188	178	178
2	18	0	0
3	147	56	56
4	376	451	238
5	184	82	61
6	35	0	0
7	729	767	322
8	569	536	302

The surveys that were the most comprehensive in terms of taxa identified were the Reef Resource Inventory (RRI) surveys conducted in 1994/95 (Table 6). When we examine the distribution of sites surveyed in the RRI surveys in relation to the clusters, the pattern is similar, except that the total number of sites surveyed in Clusters 4, 7 and 8 was approximately half that surveyed when all surveys were considered (Table 14).

4.4. Discussion

4.4.1. Clustering

Coral reef sites were clustered by four sets of species and/or habitat attributes represented as presence/absence or percentage cover. Analysis (a) aimed for geographic coverage and clustered as many samples as possible by the restricted set of species that were recorded in all datasets. Analysis (b) aimed for the broadest taxonomic coverage and clustered a restricted set of samples by the large number of species that was available for these samples. Analysis (c) clustered reef edge samples by the cover of habitat-forming species, providing clusters and a dataset that could be used to test for relationships between the cover of the species and environmental variables. Analysis (d) clustered samples by substrate and a subset of habitat-forming species.

The resulting clusters varied among being concentrated in a particular area (e.g. Warrior Reef or the Ribbon Reefs), occurring in a general area (e.g. east or west) and being widely dispersed. However, a number of general patterns emerged. Sites associated with corals, *Halimeda*, trochus, pearl oysters, crown of thorns starfish, sea cucumbers and most fish species were most frequent in the east, whereas sites associated with seagrasses were most frequent in the west. In the east the Ribbon Reefs featured a composition of fish species different from that of the remainder of the east and had the highest frequency and cover of *Halimeda*. In the west, the Warrior Reef seagrass assemblage differed from that of the remainder of the west by a high frequency of *Thalassodendron ciliatum* and *Enhalus acoroides*. Interspersed throughout Torres Strait were areas where habitat-forming species comprised mainly algae or where the cover of all habitat-forming species was low.

The cover of habitat-forming species (corals, seagrasses and algae) was shown to vary significantly with the south-east fetch associated with each site on the reef edge. Reef edges in Torres Strait are exposed to waves and the associated turbulence induced by south-

easterly trade winds in the winter and north-westerly monsoon winds in the summer, however, north-west fetch did not appear to influence the cover of habitat-forming species. This may be because on average, the trade winds are more consistent and intense. Reef geomorphology on Cocos Island, Torres Strait and Lady Elliot Island reefs (Kench and Brander 2006) and habitat variability and coral morphology and cover on Hawaiian reefs (Dollar 1982; Gibbs *et al.* 2006) have all been shown to be moulded by wave exposure. In some cases, wave exposure can improve the degree of coral cover. In Guam, coral recruitment is limited by substrate availability and under calm conditions cyanobacteria cover the substrate, preventing the settlement of coral larvae. Exposure to monsoonal winds removes the cyanobacteria which is replaced by crustose coralline algae making the substrate more suitable for coral settlement (Becerro *et al.* 2006). However, this does not appear to be the case on Torres Strait reefs, the clusters with the highest exposure to the trade winds featured the lowest covers of habitat-forming species (seagrasses, algae, corals).

Variation in the species composition of habitat-forming species was also related to variation in the physical environment. There was a clear relationship with latitude and longitude: a gradient from corals to seagrasses ran from ESE to WNW. However, we do not need to rely on latitude and longitude, which are surrogate environment variables, to represent this pattern because it can be represented by a combination of true environment variables such as temperature, salinity, turbidity and soft sediment.

The overall trend was for seagrasses to be associated with soft sediment and high turbidity, and for corals to be associated with harder substrates and clear, warm, saline water. Algal species tended to occur at various positions along this gradient. Secondary trends included an association between species relatively tolerant to disturbance (*Halodule uninervis* and *Thalassodendron ciliatum*) and south-east fetch, which is a surrogate for high energy environments and the associated disturbance.

There is at least some generality in the environmental factors that are correlated with species composition. Fishes and habitat-forming species shared a species composition gradient from ESE to WNW and a correlation between species composition and a gradient from warm, clear water and long wave period to cool, turbid water, short wave period and soft sediment. Fish and beche-de-mer species richness followed a similar pattern being greater in the east than the west. Seagrasses showed a similar trend, although in the opposite direction with species richness being greater in the west.

These patterns observed on coral reefs in Torres Strait are similar to those documented for fish and corals in the Great Barrier Reef (GBR). In the central GBR there is a strong cross-shelf gradient running from west to east. Here, coral reefs are found in environments ranging from turbid, productive inshore waters having relatively low wave energy to outer reefs on the edge on the continental shelf where the Coral Sea waters are clear, low in nutrients and have relatively high wave energy (Done 1982). This environmental gradient is reflected by the coral communities across the shelf with significant differences occurring between inshore, mid-shelf and outer-slope coral communities (Done 1982). Although corals were only classified as morphotypes in the present study and not identified to species as they were in the GBR study (Done 1982), it is interesting to note that the trend is similar.

Significant changes in reef fish communities have also been found across the mid-shelf of the GBR (Williams 1982b). There were differences between reefs within each zone (inshore, mid-shelf and outer-slope), but these were not as great as those across the shelf. In a related study, species richness was highest on the mid-shelf reefs, intermediate on the outer slope and lowest on the inshore reefs (Williams and Hatcher 1983). This contrasts with the present study where species richness was highest on the outer slope (Ribbon) reefs, lowest on the western reefs and intermediate on the mid-shelf reefs. It is difficult to give reasons for this discrepancy, although the low species richness found on the western reefs appeared to be

related to the degree of seasonal variation in salinity. A previous study found that fish communities in Torres Strait were influenced by the distance from sources of terrestrial runoff; the most influential of these was the Mai Kussa River in Papua New Guinea (Milton and Long 1997). This freshwater input from the north may distort the general pattern described for the fish communities of the GBR.

4.4.2. Threats

The coral reefs in Torres Strait are vulnerable to the effects of raised sea level temperatures and other climate significant changes related to climate. Higher temperatures have caused widespread coral bleaching to the southern reefs of the Great Barrier Reef in 1998, 2002 and 2004 (Wilkinson 2004), and localised bleaching of corals has been observed in Torres Strait during lobster surveys in the mid 1990s (CSIRO, unpublished data). Under current scenarios (Hoegh-Guldberg 2006) coral bleaching appears to be a significant threat to the reefs of Torres Strait. The downstream ecological consequences of coral bleaching are not known, however, it would likely reduce the diversity and abundance of many species groups including other demersal invertebrate communities, benthic, demersal and pelagic fish. Under these conditions, the reef is likely to be dominated by macroalgae, echinoderms and selected herbivorous fish.

Crown of thorns occur in Torres Strait reef in low densities (1.41 Ha^{-1} in 2002 and 2.03 Ha^{-1} in 2005). There has been some localised reef degradation attributed to crown of thorns outbreaks in the past (Skewes et al. 2004), however it has not been observed in a widespread fashion, even during times of outbreaks on the GBR.

Artisinal and subsistence fishing for several species takes place on the shallow reefs, primarily for finfish, sea cucumber and trochus. These fisheries can cause localised and depletions and at least two species of sea-cucumber, *Holothuria scabra* (sandfish) and *H. whitmaei* (black teatfish) are considered as overfished and are closed to fishing (Skewes et al. 2004). Several traditional food species are also fished on the shallow reefs, however, this is largely considered to be at sustainable levels (Harris et al. 1994).

Torres Strait contains several major (Prince of Wales Channel) and minor (Great North East Channel) shipping lanes, and it contains half of the 'marine environment high risk areas' (MEHRAs) for oil spills in Queensland waters (Rasheed and Thomas 2005). Between 1985 and 2000, there were 40 major shipping incidents in the Great Barrier Reef and Torres Strait (Anon. 2003) consisting of 26 groundings and 14 collisions. Shipping through Torres Strait is expected to increase in the future, and although new procedures and legislation is reducing risk, there is still a reasonable probability of shipping accidents within Torres Strait. Groundings represent the main source of risk to the Great Barrier Reef and Torres Strait, with the potential for ecological damage due to pollution from oil or chemical spills or anti-fouling paints, as well as physical damage to habitats.

4.4.3. Gap analysis

In terms of overall spatial coverage a high proportion of the reefs of TS have been surveyed by CMAR over the past 12 y. There are some obvious gaps: the reefs to the west of the main TS island chain, the mid-shelf reefs of the southeast, the large reefs to the east of the tip of Cape York and there are gaps in the coverage of the Ribbon Reefs on the outer GBR. By examining the reefs in terms of combinations of physical covariates that had been identified in shaping the biological nature of the reefs we identified that most classes of reefs had been sampled except for two minor categories (Clusters 2 and 6). In terms of area these reefs only represent 2.4% of the total area of coral reefs in the TS.

Although the spatial coverage of sampling has been quite extensive in TS, it must be remembered that the range of data collected on these surveys has been quite limited. Most surveys were focused on sampling to assist in stock assessment of particular fishery species or habitat related to fishery species (seagrass) and so while precise counts and species level identifications were performed for the target species, ancillary data collected on other invertebrates and algae for instance were either in terms of broad taxonomic groups (or morphotypes for corals and sponges). So while the spatial extent of sampling has been quite broad, there are large gaps in our knowledge of the biodiversity of the reef biota of the TS. It is also worth noting that the most comprehensive survey (RRI) was only conducted once and that was 12 years ago.

The CRC TS mapping project identified 2,372 OTUs in the inter-reefal areas throughout TS and it is highly likely that the coral reefs of the region support a biota equally as rich. From a biological point of view, TS is interesting because it appears to be a biogeographic boundary; sea level fluctuations during the Pleistocene would have resulted in the Strait (along with the Malacca and Sunda Straits) being emergent for significant periods, acting as a barrier to gene flow between the Pacific and Indian Oceans (Chenoweth *et al.* 1998). This resulted in the formation of different species on either side of the barrier; when sea levels subsequently rose and the barrier was flooded, some incipient species merged, but others remained as distinct species e.g. the coral reef fish *Chaetodon lineolatus* and *C. oxycephalus* (Randall 1999). Given that the coral reefs of the Torres Strait have been so little studied, hold a wealth of biodiversity and are potentially at risk due to the threats posed by climate change, it would be timely to increase our knowledge of the TS reef biota.

4.4.4. Identification of areas for future work

The coral reefs of the TS are of immense cultural, economic and aesthetic value as they support a range of artisanal and commercial fishing industries and a developing tourism industry. Information on the status and natural variability of reef populations is vital in the management of these resources. AFMA regularly funds stock assessment surveys for reef associated target fishery species such as bech-de-mer and tropical rock lobster, but little is known of the status of other reef organisms. The corals of Torres Strait have only been identified to morphotypes in this study and there appears to have been little in the way of a systematic description of the coral communities of the region. Climate change and Crown-of-thorns starfish (CoT) represent potential threats to the coral reefs of TS and while the Australian Institute of Marine Science have conducted long term monitoring for CoTs on the Great Barrier Reef since the mid 1980's, their monitoring sites only extend to Cape Grenville, well south of the Torres Strait. CMAR surveys have identified some local reef degradation resulting from CoTs, but the full extent is unknown.

We recommend that a comprehensive survey be conducted throughout the reefs of the area, in a similar manner to that conducted for the inter-reefal areas by the CRC Torres Strait (Pitcher *et al.* 2007), covering a wide range of taxonomic groups. From this initial survey, a subset of representative reefs could be chosen for future monitoring of key groups of organisms.

5. References

- Amin M (1978) A statistical analysis of storm surges in the Torres Strait. *Australian Journal of Marine and Freshwater Research* 29: 479-496
- Anon. (2003) Great Barrier Reef & Torres Strait shipping impact study.
- Araujo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography* 33: 1677-1688
- Ayukai T, Wolanski E (1997) Importance of Biologically Mediated Removal of Fine Sediments from the Fly River Plume, Papua New Guinea. *Estuarine, Coastal and Shelf Science* 44: 629-639
- Becerro MA, Bonito V, Paul VJ (2006) Effects of monsoon-driven wave action on coral reefs of Guam and implications for coral recruitment. *Coral Reefs* 25: 193 - 199
- Bode L, Mason LB (1994) Numerical modelling of tidal currents in Torres Strait and the Gulf of Papua, Report to the Victorian Institute of Marine Science, 65 pp.
- Box GEP, Cox DR (1964) An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)* 26: 211-252
- Breiman L (1996) Bagging predictors. *Machine Learning* 24: 123-140
- Breiman L (2001) Random forests. *Machine Learning* 45: 5-32
- Browne M, Haywood M, Ellis N, Venables RPB (2007) Improving non-parametric species distribution models with a power transformation. *Ecography*: submitted
- Chappell J (1983) A revised sea-level record for the last 300, 000 years from Papua New Guinea. *Search* 14: 99-101
- Chenoweth SF, Hughes JM, Keenan CP, Lavery S (1998) When oceans meet: a teleost shows secondary intergradation at an Indian-Pacific interface. *Proceedings: Biological Sciences* 265: 415-420
- De'Ath G (2002) Multivariate regression trees: a new technique for modeling species-environment relationships. *Ecology* 83: 1105-1117
- Dollar SJ (1982) Wave stress and coral community structure in Hawaii. *Coral Reefs* 1: 71-81
- Done TJ (1982) Patterns in the distribution of coral communities across the central Great Barrier Reef. *Coral Reefs* 1: 95-107
- Duda RO, Hart PE, Stork DG (2001) *Pattern Classification*. Wiley, New York
- Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129-151
- Ferrier S, Guisan A (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology* 43: 393-404

- Ferrier S, Manion G, Elith J, Richardson K (2007) Using generalised dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* 13: 252-264
- Garzon MB, Blazek R, Neteler M, de Dios RS, Ollero HS, Furlanello C (2006) Predicting habitat suitability with machine learning models: The potential area of *Pinus sylvestris* L. in the Iberian Peninsula. *Ecological Modelling* 197: 383-393
- Garzon MB, de Dios RS, Ollero HS (2007) Predictive modelling of tree species distributions on the Iberian Peninsula during the Last Glacial Maximum and Mid-Holocene. *Ecography* 30: 120-134
- Gibbs A, Cochran S, Storlazzi C, Grossman E, Field M (2006) Coral Habitat Variability in the Kona Coast National Parks EOS, Transactions, American Geophysical Union. American Geophysical Union, 2000 Florida Ave., N.W. Washington DC 20009 USA, [URL:http://www.agu.org/pubs/agu_joureos.html]
- Gislason PO, Benediktsson JA, Sveinsson JR (2006) Random Forests for land cover classification. *Pattern Recognition Letters* 27: 294-300
- Granitto PM, Gasperi F, Biasioli F, Trainotti E, Furlanello C (2007) Modern data mining tools in descriptive sensory analysis: A case study with a Random forest approach. *Food Quality and Preference* 18: 681-689
- Han AK (1987) A Non-parametric analysis of transformations. *Journal of Econometrics* 35: 191-209
- Harris AN, Dews G, Poiner IR, Kerr J (1994) The traditional and island based catch of the Torres Strait Protected Zone. Final report to the Scientific Advisory Committee of the Torres Strait Protected Zone
- Harris PT (2001) Environmental management of Torres Strait: a marine geologist's perspective. In: Gostin VA (ed) *Gondwana to Greenhouse: Environmental Geoscience - An Australian perspective*. Geological Society of Australia Special Publication, Adelaide, pp 317-328
- Harris PT, Baker EK (1991) The nature of sediments forming the Torres Strait turbidity maximum. *Australian Journal of Earth Sciences* 38: 65-78
- Harris PT, Baker EK, Cole AR, Short SA (1993) A preliminary study of sedimentation in the tidally dominated Fly River delta, Gulf of Papua. *Continental Shelf Research* 13: 441-472
- Harris PT, Heap A, Passlow V, Hughes M, Daniell J, Hemer M, Anderson O (2005) Tidally incised valleys on tropical carbonate shelves: An example from the northern Great Barrier Reef, Australia. *Marine Geology* 220: 181-204
- Heap A, Daniell J, Mazen D, Harris PT, Sbaffi L, Fellows M, Passlow V (2004) Geomorphology and sedimentology of the Northern Marine Planning Area of Australia. Review and synthesis of relevant literature in support of Regional Marine Planning. Geoscience Australia, geoscience Australia Record 2004/11, Canberra
- Hemer MA, Harris PT, Coleman R, Hunter J (2004) Sediment mobility due to currents and waves in the Torres Strait-Gulf of Papua region. *Continental Shelf Research* 24: 2291-2316

- Hoegh-Guldberg O (2006) Impacts of climate change on coral reefs. In: Office CMaARRttAG (ed) Impacts of Climate Change on Australian Marine Life. Department of the Environment and Heritage
- Hopley D (1983) Evidence of 15, 000 years of sea level change in tropical Queensland. In: Hopley D (ed) Australian sea levels in the last 15,000 years: a review. department of Geography, James Cook University, pp 93-104
- Jones MR (1995) The Torres Reefs, North Queensland, Australia -- strong tidal flows a modern control on their growth. Coral reefs. Heidelberg 14: 63-69
- Kaufman L, Rousseeuw PJ (1990) Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, New York
- Kench PS, Brander RW (2006) Wave Processes on Coral Reef Flats: Implications for Reef Geomorphology Using Australian Case Studies. Journal of Coastal Research 22: 209-223
- Kendal WS (2002) Spatial aggregation of the Colorado potato beetle described by an exponential dispersion model. Ecological Modelling 151: 261-269
- Kendall M (1938) A New Measure of Rank Correlation. Biometrika 30: 81-89
- Lehmann C, Koenig T, Jelic V, Prichep L, John RE, Wahlund LO, Dodge Y, Dierks T (2007) Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). Journal of Neuroscience Methods 161: 342-350
- Long BG, Skewes T, Taranto TJ, Jacobs D, Dennis D (1997) Torres Strait Reef Resource Inventory and Reef Habitat Mapping. CSIRO Marine Research, MR-GIS 97/6, Cleveland, Australia
- Long BG, Skewes TD, Poiner IR (1995) Torres Strait marine geographic information system. Recent Advances in Marine Science and Technology '94'. (Eds O. Bellwood, JH Choat and N. Saxena.) 1994: 231-239
- Margvelashvili N, Saint-Cast F (2007) Modelling of fine sediment transport in Torres Strait. Cooperative Research Centre for Torres Strait. Torres Strait Research Program
- McMillan J (1982) A Global Atlas of GEOS-3 Significant Wave height Data and Comparison of the Data with National Buoy Data(156882)
- Milton DA, Long BG (1997) Influence of coastal processes on large scale patterns in reef fish communities of Torres Strait, Australia. CSIRO Marine Research Cleveland, Q. 4163, Australia
- Pal M (2005) Random forest classifier for remote sensing classification. International Journal of Remote Sensing 26: 217-222
- Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography 34: 102-117
- Phillips S, Dud M, Robert K, Schapire E (2004) A maximum entropy approach to species distribution modeling Proceedings of the Twenty-First International Conference on Machine Learning

Pitcher CR, Condie S, Ellis N, McLeod I, Haywood M, Gordon SR, Skewes TD, Dunn J, Dennis D, Cotterell E, Austin M, Venables W, Taranto T (2004) Torres Strait seabed and water-column data collation, biophysical modeling and characterisation. CSIRO Marine Research. pp. 117, Final Report to the National Oceans Office

Pitcher CR, Haywood M, Hooper J, Coles R, Bartlett C, Browne M, Cannard T, Carini G, Carter A, Cheers S, Chetwynd D, Colefax A, Cook S, Davie P, Ellis N, Fellegara I, Forcey K, Furey M, Gledhill D, Hendriks P, Jacobsen I, Johnson J, Jones M, Last P, Marks S, McLeod I, Sheils J, Sheppard J, Smith G, Strickland C, Van der Geest C, Venables W, Wassenberg T, Yearsley G (2007) Mapping and Characterisation of Key Biotic & Physical Attributes of the Torres Strait Ecosystem. CSIRO/QM/QDPI CRC Torres Strait Task Final Report.

Pitcher R, Bishop M, Dennis D, Skewes T (1997) The Torres Strait lobster fishery. Stock assessment report, Torres Strait Fisheries Assessment Group. Australian Fisheries Management Authority, Canberra, Australia

Prasad AM, Iverson LR, Liaw A (2006) Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems* 9: 181-199

R Development Team (2007) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Randall JE (1999) Zoogeography of coral reef fishes of the Indo-Pacific region. Proceedings of the 5th Indo-Pacific Fish conference, Noumea, New Caledonia, pp 23-26

Rasheed MA, Thomas R (2005) Critical marine habitats adjacent to Torres Strait shipping lanes and ports: Interim Atlas - Prince of Wales shipping channel and Port of Thursday Island - 2004. DPI&F, Northern Fisheries Centre, Cairns

Ridgway KR, Dunn JR (2002) Ocean interpolation by four-dimensional weighted least squares - application to the waters around Australia. *Journal of Atmospheric and Oceanic Technology* 19: 1357-1375

Robertson AI, Daniel PA, Dixon P, Alongi DM (1993) Pelagic biological processes along a salinity gradient in the Fly delta and adjacent river plume (Papua New Guinea). *Continental Shelf Research* 13: 205-224

Rubin DB (1984) The Analysis of Transformed Data: Comment. *Journal of the American Statistical Association* 79: 309-312

Schapire RE, Singer Y (1999) Improved boosting algorithms using confidence-rated predictions. *Machine Learning* 37: 297-336

Schwartz MW, Iverson LR, Prasad AM, Matthews SN, O'Connor RJ (2006) Predicting extinctions as a result of climate change. *Ecology* 87: 1611-1615

Segurado P, Araujo MB (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography* 31: 1555-1568

Skewes T, Dennis D, Toutsoukos A, Haywood M, Wassenberg T, Austin M (2003) Stock survey and sustainable harvest strategies for Torres Strait beche-de-mer. AFMA Project Number R01/1343. CSIRO Marine Research Final Report.

Skewes T, Taylor S, Dennis D, Haywood M, Donovan A (2006) Sustainability Assessment of the Torres Strait Sea Cucumber Fishery, CRC-TS Project Task Number: T1.4. CSIRO Marine and Atmospheric Research Final Report. , pp 50

- Skewes TD, Dennis D, Koutsoukos AD, Haywood MDE, Wassenberg T, Austin M (2004) Stock survey and sustainable harvest strategies for Torres Strait beche-de-mer. Australian Fisheries Management Authority. CSIRO Marine Research, 1876996617 Subject Fishery resources Research Torres Strait. Fisheries Research Torres Strait. Marine resources Torres Strait., Cleveland, Qld.
- Thuiller W, Araujo MB, Lavorel S (2003) Generalised models vs. classification tree analysis: Predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science* 14: 669-680
- Venables WN, Ripley BD (1994) *Modern Applied Statistics with S-Plus*. Springer-Verlag, New York
- Walsh JP, Nittrouer CA (2003) Contrasting styles of off-shelf sediment accumulation in New Guinea. *Marine Geology* 196: 105-125
- Warton DI (2005) Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. *Environmetrics* 16: 275-289
- Wilkinson C (2004) *Status of coral reefs of the world: 2004*. Australian Institute of Marine Science, Townsville, Queensland, Australia.
- Williams D, McB. (1982a) Patterns in the distribution of fish communities across the central Great Barrier Reef. *Coral reefs*. Heidelberg 1: 35-43
- Williams DM (1982b) Patterns in the distribution of fish communities across the central Great Barrier Reef. *Coral Reefs* 1: 35-43
- Williams DM, Hatcher AI (1983) Structure of fish communities on outer slopes of inshore, mid-shelf and outer shelf reefs of the Great Barrier Reef. *Marine Ecology Progress Series* 10: 239-250
- Woodroffe CD, Kennedy DM, Hopley D, Rasmussen CE, Smithers SG (2000) Holocene reef growth in Torres Strait. *Marine Geology* 170: 331-346