2.4. BIOLOGICAL SURVEY DATA

2.4.1. Data Processing

Biological data was amalgamated from 16 data sets from surveys in the Torres Strait study area conducted for a variety of purposes ranging from stock assessment of tropical rock lobster and Bech-de-Mer to environmental seabed mapping surveys. The surveys were carried out between 1987 and 2002. The information has been extracted from diver or video transects: divers recorded habitat components & coverage and video was coded into habitat facies in 'real-time' — both can be regarded as "Rapid Assessment" techniques. The video data included the frequency of major habitat facies such as: substratum type (mud, silt, sand, rubble, rock); epifaunal gardens (combinations of composition & density); and algae, seagrass, and bivalve shell beds. The diver data included similar estimates of percent cover of these (Pitcher *et al*., 1992) and some additional components in more detail (some to genus or species). Altogether, the datasets contained 527 individual variables collected at over 1500 sites in the area of interest. Theses were amalgamated and combined to produce 26 standard variables that were considered as suitable for describing broad-scale environmental characteristics of the study area.

Table 2.4-1. Coverage of information common across the widest spatial coverage of Project and Cruises. (A) Widest coverage/simplest data, (B) slightly narrower coverage/more information (see section 2.1.1 for definitions of these fields). Scale: S = Small, M = Medium and L = Large.

(A) Dataset	Scale	Sub_Code	Bio_Code	Seagrass	Sub-Comp	Algae
Warraber	S	✓	✓			
KulasiWest	M	✓	✓	PA		PA
Lobster'89	L	✓	✓	PA		PA
PNG Lob'98	M	✓	✓	PA		PA
Lobster'02	L	✓	✓	PA		PA
IPC'96	L	✓	✓	PA		PA
Chevron'97	L	✓	✓	PA		PA

(B) Dataset	Scale	Sub_Code	Bio_Code	Seagrass	Sub-Comp	Algae
Warraber	S	✓	✓		✓ %	
KulasiWest	M	✓	✓	Tot%		
Lobster'89	L	✓	✓	Tot%	✓ %	Tot%
PNG Lob'98	M	✓	✓	Tot%	✓ %	Tot%
Lobster'02	L	✓	✓	Tot%	✓ %	Tot%
IPC'96	L	✓	✓		✓ %	Tot%
Chevron'97	L	✓	✓		✓ %	Tot%

The final dataset had a wide coverage for more general habitat descriptions such as coded substrate (1452/1501) and epibenthos abundance (1336/1501) categories (Table 2.4-1A). A somewhat more detailed substratum description that included the relative cover of several substrate categories were possible at 968/1501 sites, and algal & seagrass percent cover was possible at 490/1501 sites (Table 2.4-1B). The resulting biological dataset had a broad spatial coverage, north-south and east-west (Figure 2.4-1) which provided a wide match with the range of variability of physical covariates (section 2.3.2) to examine bio-physical relationships (section 2.7).

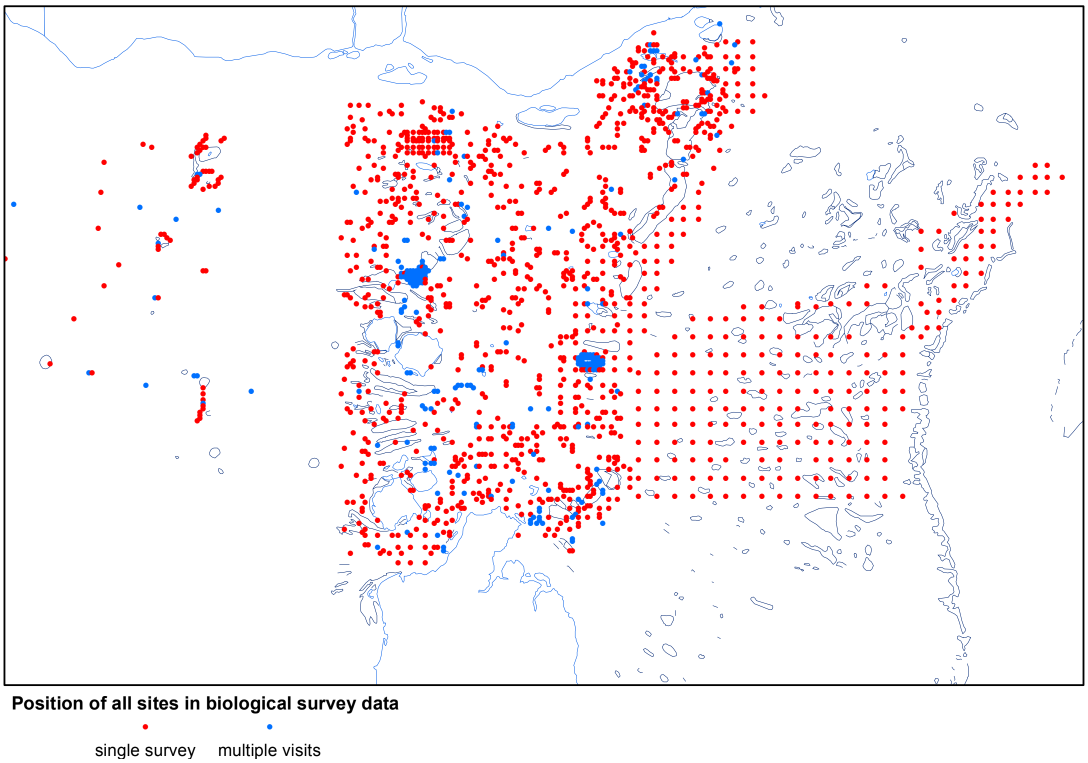


Figure 2.4-1 Map of the location of all seabed biological survey sites included in the data collated for this project (sources, see section 2.1.1).

2.4.2. Maps of Biological Survey Data

2.4.2.1. Substratum Type

The substratum type observed during dive or towed-video surveys was recorded in a number of forms: SUB_CODE (see section 2.1.1 for definition) was the most basic form and was common to all included surveys (Figure 2.4-2). During most towed-video surveys, the video was observed in real time and occurrence of mud, sand, rubble, rocks, reef, along with other seabed habitats was entered directly into a computer ("Tappity" data) an later summarized to percent occurrence on transects (Figure 2.4-4)and, with intermediate spatial coverage, estimated substratum proportions of mud, sand, rubble, consolidated and hard (Figure 2.4-4). These patterns correspond well with the sediment attribute data, with the same processes affecting the observed distributions (see section 2.3.2.2).

2.4.2.2. Epibenthos Habitat

The epibenthos habitat type observed during dive or towed-video surveys was also recorded in a number of forms: BIO_CODE (see section 2.1.1 for definition) was the most basic form and was common to all included surveys comprising a rough estimate of the abundance of epibenthos on a transect (Figure 2.4-2); "Tappity" data was entered in real time during most towed-video surveys and included sparse, medium and dense epibenthos gardens, along with other seabed habitats that were summarized to percent occurrence on transects (Figure 2.4-4); and one survey estimated the percent cover of epibenthos and coral (Figure 2.4-3). The patterns of epibenthos distribution correspond well with that of harder substratums (2.4.2.1) and with high seabed current stress (2.3.2.3).

2.4.2.3. Algae and Seagrass

The algae and seagrass observed during dive or towed-video surveys were recorded as presence/ absence most broadly (Figure 2.4-5) and, in somewhat fewer surveys, as estimated percent cover (Figure 2.4-6). Algae were almost ubiquitous across Torres Strait and harder, higher current areas tended to have higher cover of algae (cf. 2.4.2.1 and 2.3.2.3). Of course there were many species of algae and different species occurred in different physical environments. Seagrass was more restricted to the north-west half of Torres Strait and there were few sites that had very high cover of seagrass.

2.4.2.4. Seabed fishes

During the mid 1980's a series of cruises was undertaken in the central eastern Torres Strait as part of a study of the effects of trawling in the region. The purpose of these cruises was to characterise the fish bycatch of the Torres Strait prawn trawl fishery and to compare fish assemblages in trawled and adjacent untrawled areas. Specifically, seven cruises were conducted between February 1985 and September 1986. The vessel towed 4 × 5 fathom prawn trawls for approximately 30 minutes at each site. The fish species in the catch were identified, weighed and measured. Catch data from each trawl was standardised as kg/ha and for this report, mean catch rates of each species at each site were calculated across all cruises. The location of the sampling sites was largely restricted to the area of the trawl fishery and a more limited area in the west of Warrior Reef closure (Figure 2.4-7).

In general, the sites inside the area closed to trawling, to the west of the Warrior Reef complex had more fish species (mean = 21.4 ±*sd* 10.7 species; range: 13 – 29.5) than those open to trawling (mean = 14.7 ±*sd* 7.9 species; range: 10.4 – 28.7) (p<0.001) (Figure 2.4-8). However, catch rates were similar inside the closure (mean = 7.1 ±*sd* 2.2 kg/ha; range:4.2 – 11.3 kg/ha) compared with outside the closure (mean = 6.3 ±*sd* 4.8 kg/ha; range: 2.4 – 32.4 kg/ha) (p<0.84), although one site within the trawl ground had a very high catch rate of fish (Figure 2.4-9).

The fish assemblages of this part of Torres Strait were characterised by clustering the sites and then these clusters were examined in relation to the physical characteristics using Linear Discriminant Function analysis (see Section 2.7.3).

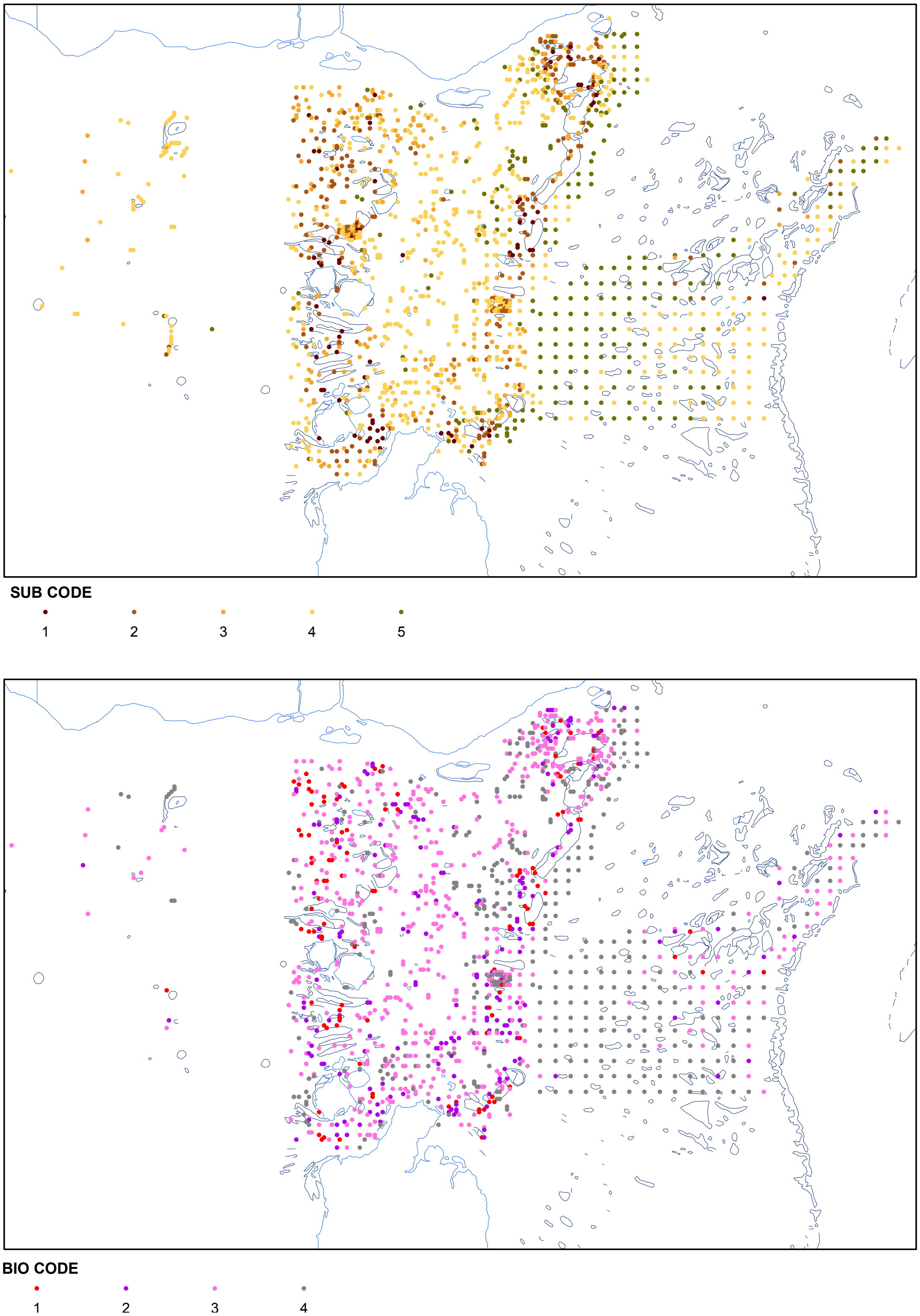
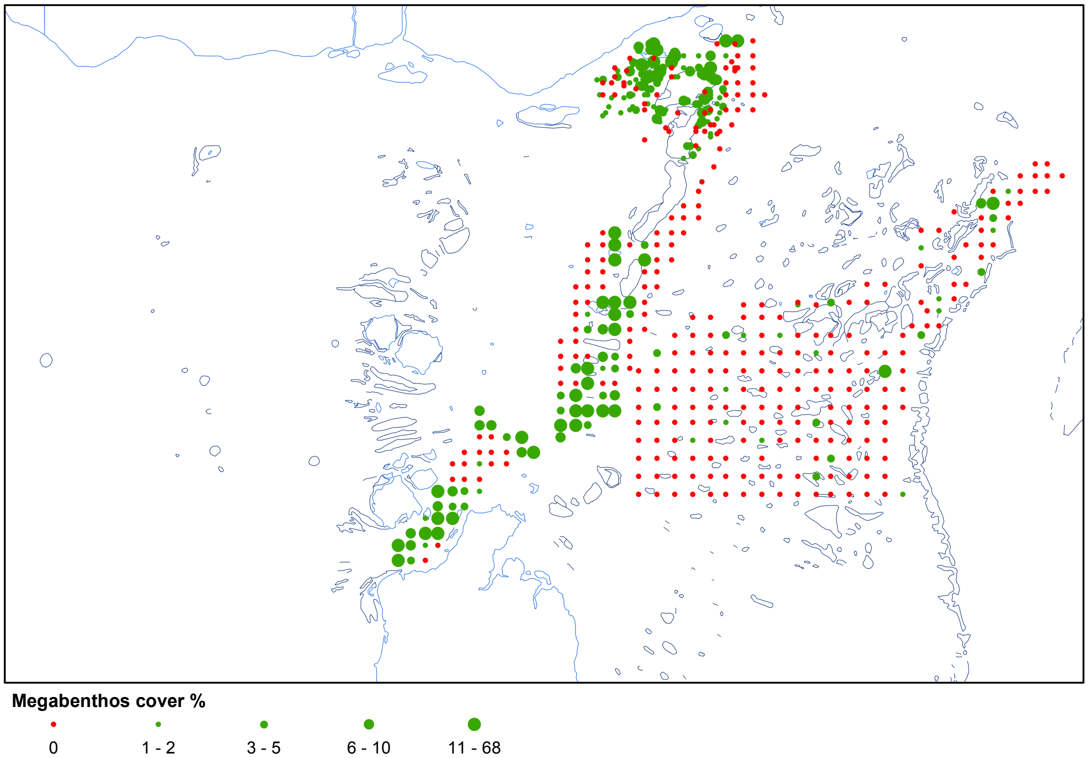


Figure 2.4-2 Maps of substratum type and epibenthic gardens from biological survey sites: Substrate classification: 1= >50% hard substrate, 2= 10-50% hard substrate, 3= rubble (<10% hard & >15% rubble), 4= sand (muddy sand & sand), 5= mud (mud & sandy mud); Epibenthos classification: 1= dense fauna, 2= sparse fauna, 3= very sparse fauna, 4= no fauna (sources, see section 2.1.1).



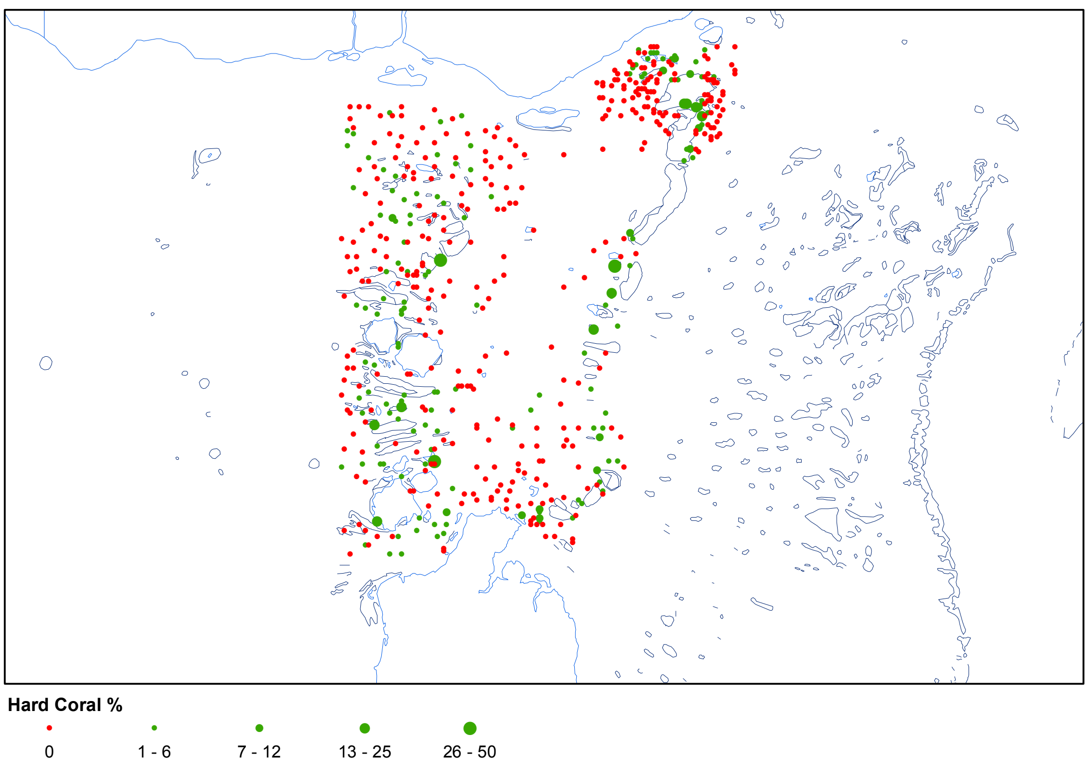
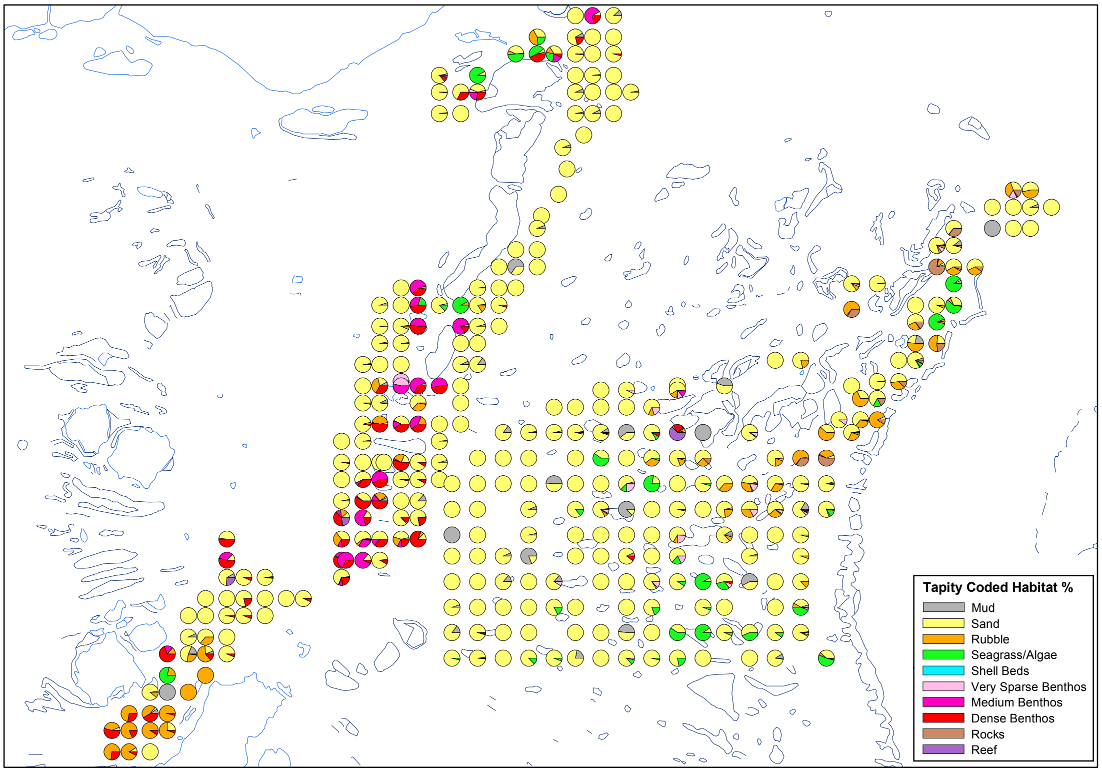


Figure 2.4-3 Maps of percent cover of epibenthos gardens and hard coral gardens of 500 m transects, from biological survey sites (sources, see section 2.1.1).



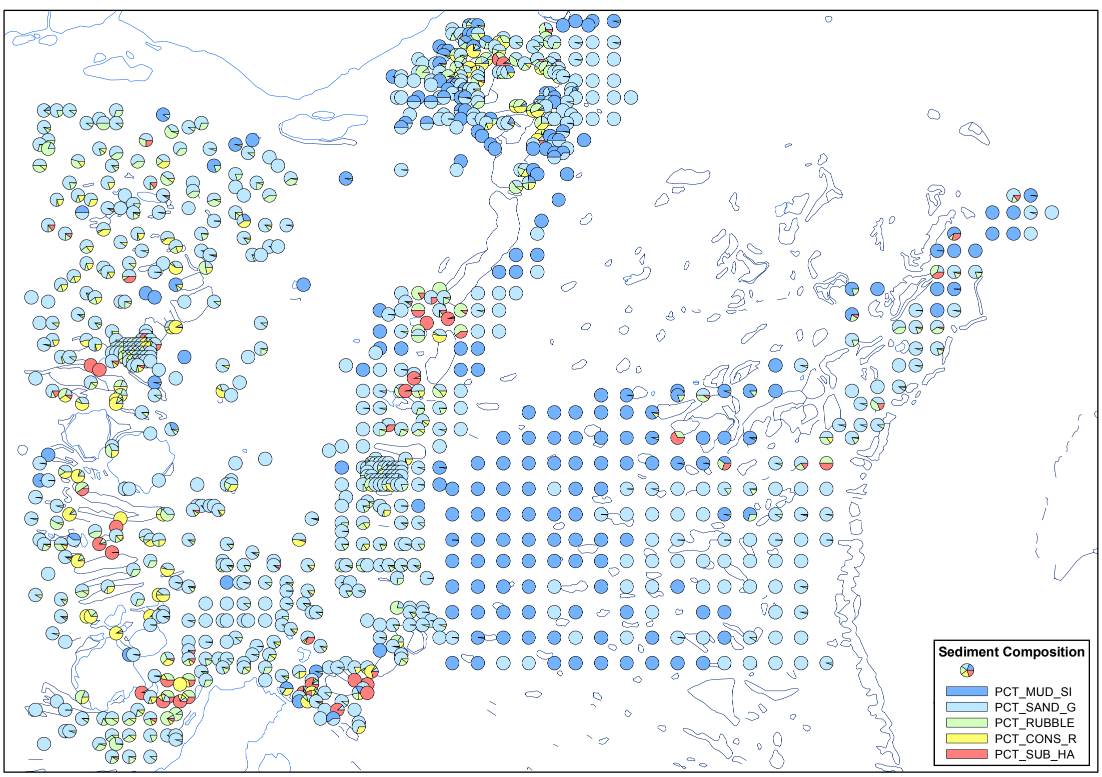


Figure 2.4-4 Maps of summary underway 'tappity' coding of seabed habitats and observed substratum composition, as a percentage of 500 m transects, from biological survey sites (sources, see section 2.1.1).

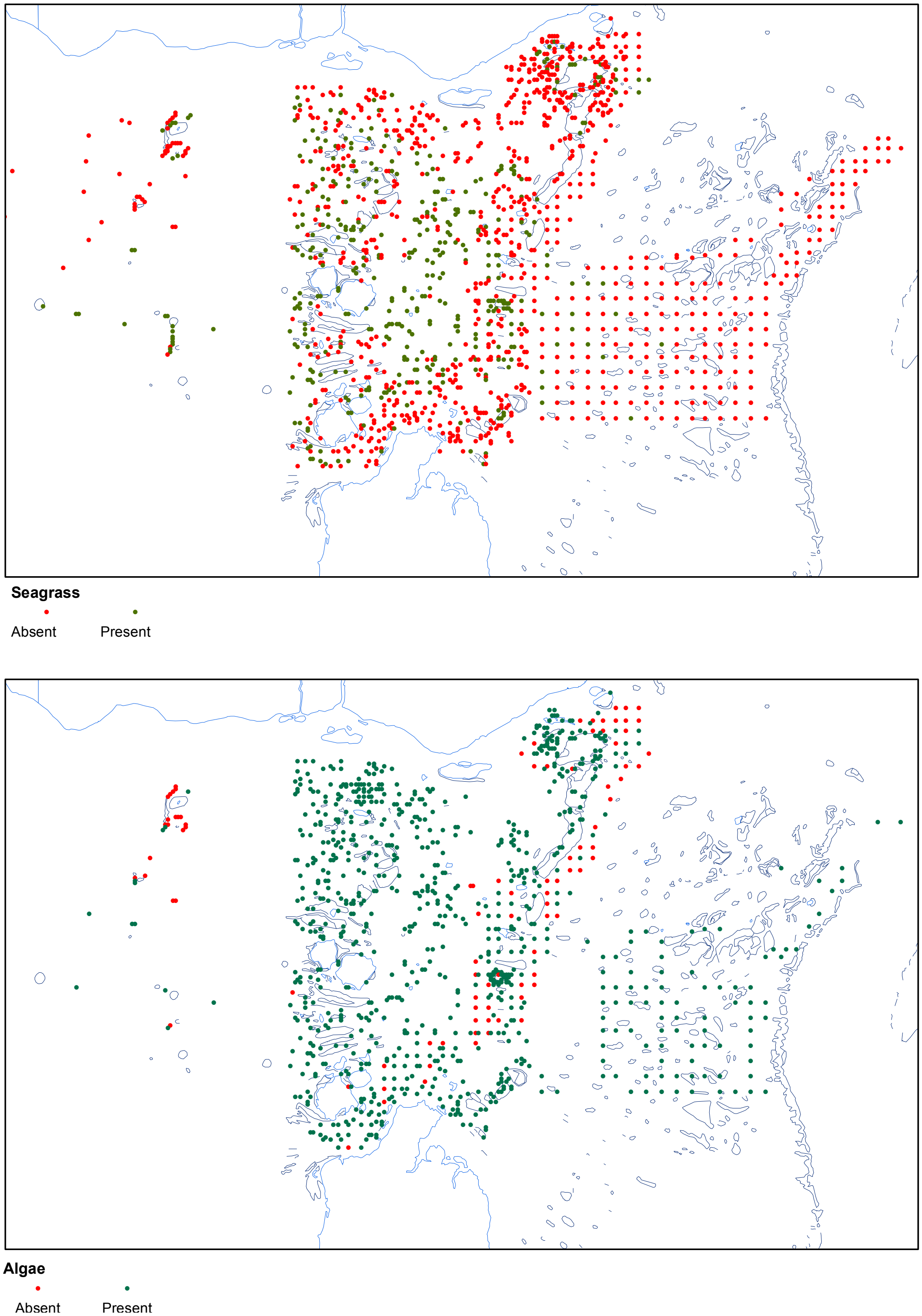
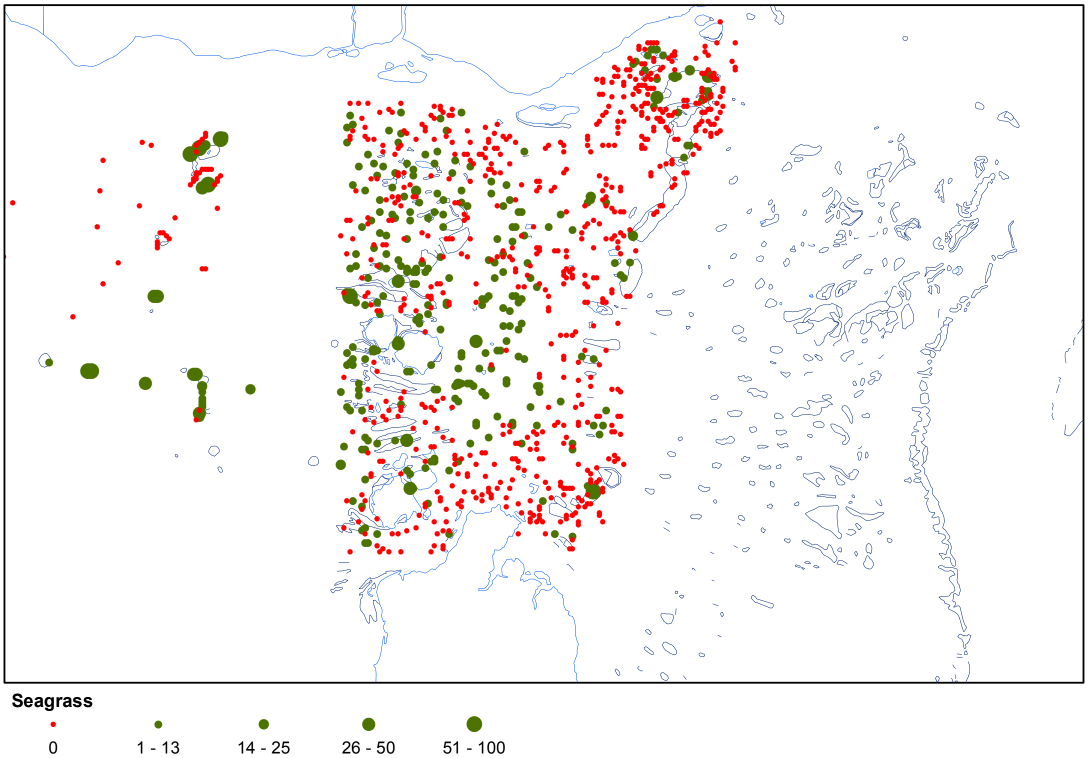


Figure 2.4-5 Maps of seagrass presence/absence and algal presence/absence from biological survey sites (sources, see section 2.1.1).



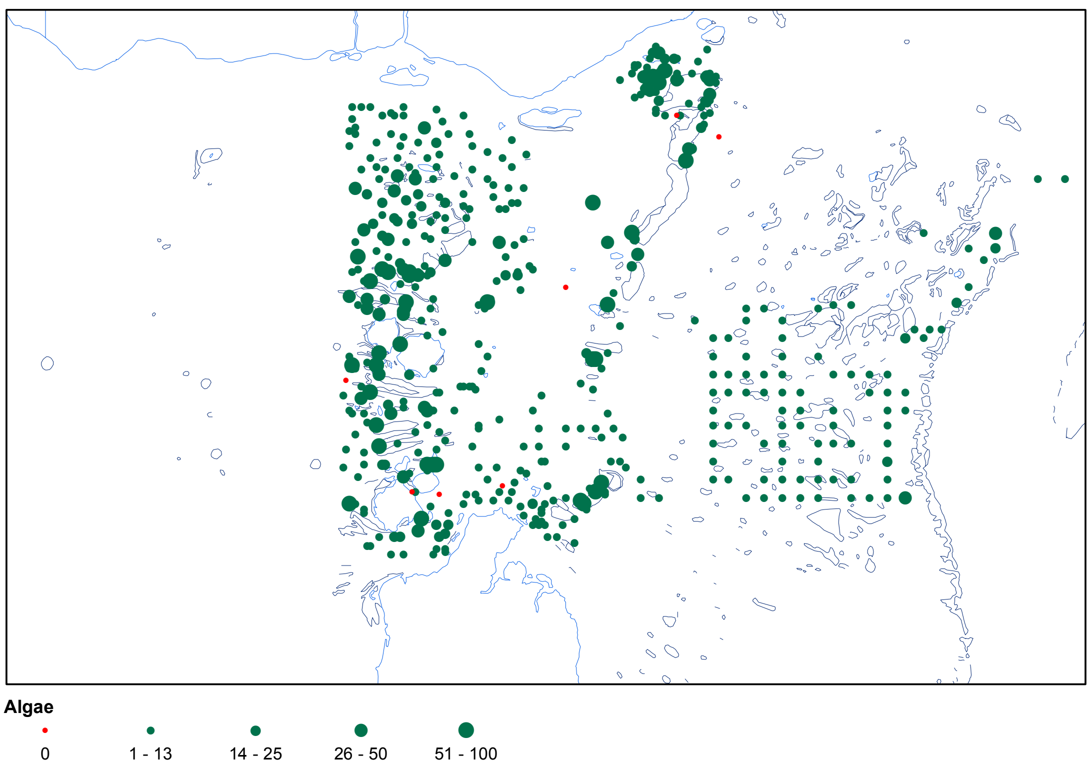


Figure 2.4-6 Maps of seagrass percent cover and algal percent cover of 500 m transects, from biological survey sites (sources, see section 2.1.1).

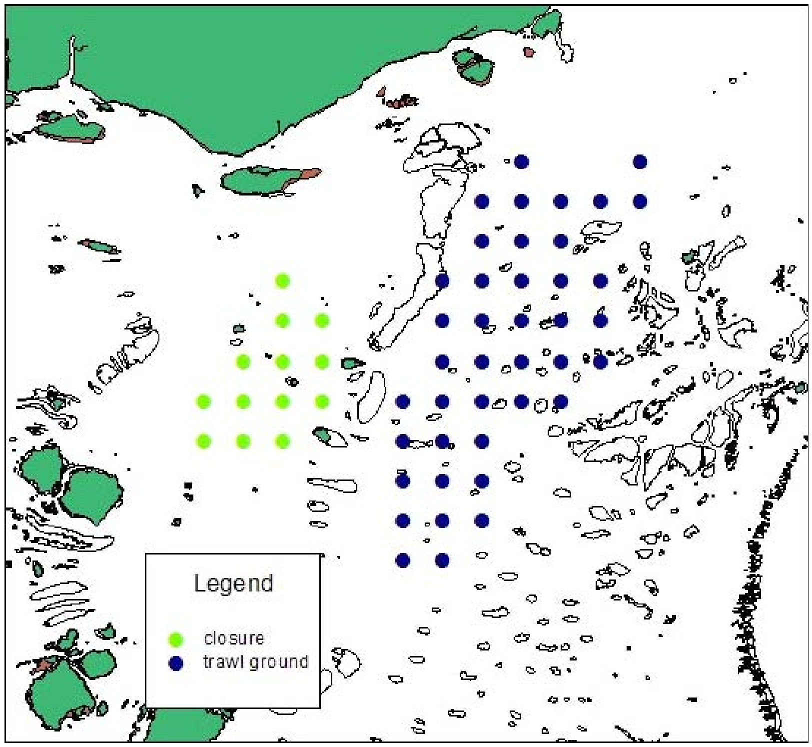


Figure 2.4-7 Map of central Torres Strait showing the sites sampled by the *FV Jacqueline-D* during the "Effects-of-Trawling" surveys of 1985-6 (sources, see section 2.1.1).

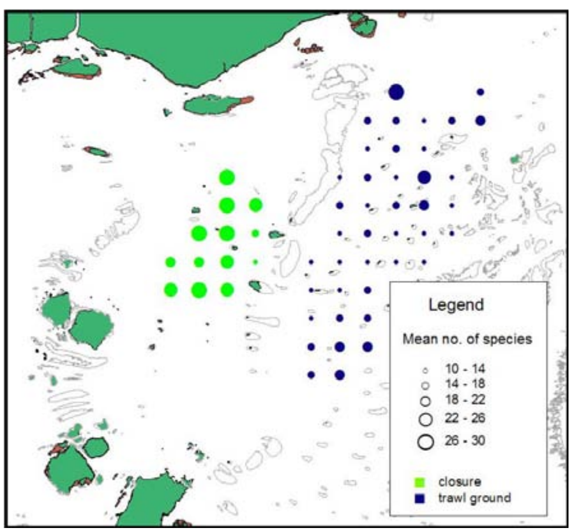


Figure 2.4-8 Map of central Torres Strait showing the mean number of fish species sampled by the *FV Jacqueline-D* during the "Effects-of-Trawling" surveys of 1985-6 (sources, see section 2.1.1).

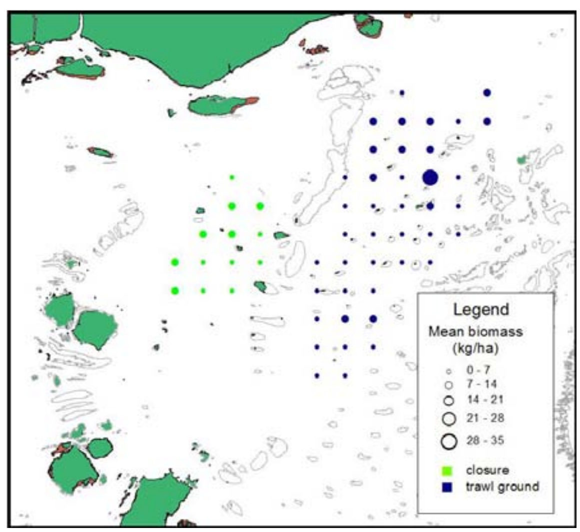


Figure 2.4-9 Map of central Torres Strait showing the mean catch rate (kg/ha) of fish bycatch sampled by the *FV Jacqueline-D* during the "Effects-of-Trawling" surveys of 1985-6 (sources, see section 2.1.1).

2.5. ACOUSTIC SURVEY DATA

Underwater acoustic sound pulses emitted from an echosounder, then reflected by the seafloor and collected by the seabed classification instrument RoxAnn provided acoustic information about the seabed in Torres Strait. This section outlines the acoustic information available to the project, and the analyses used to produce a biological and substrate stratification of the Torres Strait based on this acoustic information.

2.5.1. Data Processing

Seabed acoustic roughness and hardness data was collected continuously along with depth information (approximately twice a second), on seventeen cruises which either conducted research in Torres Strait or transited through the area collecting data. Over the 10 cruises a total of 1633625 individual RoxAnn samples were recorded. The earliest cruise was conducted on the Southern Surveyor in 1993 (ss0193), and the most recent being the 2002 annual Torres Strait rock lobster abundance survey. The vessel track for RoxAnn acoustic sampling is shown in Figure 2.5-1 below.

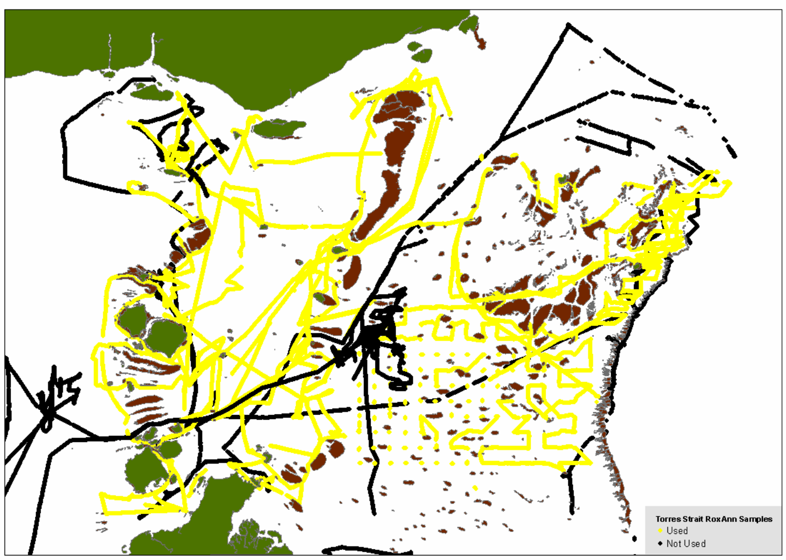


Figure 2.5-1 The vessel track for RoxAnn acoustic sampling in Torres Strait (samples used shown in yellow).

Given that there was no dedicated acoustics expertise assigned to the collection of the data during most of these cruises, and the acoustic sampling was carried out as an adjunct to existing research programs, data quality was an issue on some of these cruises. With the stratification analysis of this data being conducted post-hoc, careful data quality checking was required at a course "cruise" level to investigate systematically biased data (e.g. due to changing echosounder or RoxAnn settings over a cruise and between cruises). Initial "cruise" level data quality checks precluded some sections of acoustic data completely from further analysis. In fact six of the ten cruises (in their entirety) were not used in further analysis, as the data collected was of such poor quality, or the cruise provided little useful data. However the remaining four cruises with usable data accounted for 87% of the total available RoxAnn data and had a broad geographic coverage (see figure above, highlighting usable RoxAnn data in yellow, with unusable data in black). Therefore the “cruise level” data quality filtering of acoustic information did not significantly compromise subsequent analyses and the stratification objectives.

Even with the broad geographic coverage of usable acoustic data in Torres Strait, it should be noted that there are some information gaps. For example, there are gaps in the geographic coverage of RoxAnn data in western and north-eastern Torres Strait. These areas should be targeted for any future acoustic sampling.

The total number of RoxAnn samples collected for each of the four cruises is shown in Table 2.5-1 below. Three cruises; the 1996 Pipeline, 1997 Chevron, and 2002 Beche de Mer cruises account for the vast majority (and similar proportions) of the total available data with only a minor contribution from the earlier 1993 Rock Lobster cruise.

Table 2.5-1 Total RoxAnn samples collected.

Cruise	RoxAnn Samples
1993 – Torres Strait Rock Lobster	53381
1996 – Torres Strait Pipeline	362858
1997 – Chevron	486623
2002 – Torres Strait Beche de Mer	512211

Two cruises; the 1993 Rock Lobster, and 2002 Beche de Mer cruises covered the central area of Torres Strait from Thursday Island in the south-west to the Warrior Reefs in the north-east. The 1997 Chevron cruise covered central-eastern Torres Strait, again around the Warrior Reefs, Sassie Island and south-west to Thursday Island. While there was some data available for eastern Torres Strait from the 2002 Beche de Mer cruise (especially in the north-east); the 1996 Pipeline cruise accounted for the vast majority of data, covering the 10000 km^2 area from Kagar Reef in the south-west to Don Cay in the north-east. The geographic coverage for each of the four cruises is shown in Figure 2.5-2 below.

Careful filtering and interpretation (Pitcher et al., 1999; Kloser et al., 2001) was required to remove erroneous data (e.g. due to the effect of poor weather influencing the acoustic information, and depth outliers due to noise spikes in readings) within each cruise prior to further analysis; and take into account possible influencing relationships within the data if possible (e.g. bias effects of depth and vessel speed).

The RoxAnn instrument provided relative information on acoustic roughness and hardness. These measures (a unitless numerical index between 0 – 2048) are not absolute or easily calibratable. Also, the RoxAnn readings may be influenced by a number of variables, most importantly the echo sounder system used, including transceiver settings and transducer type. Though, once calibrated over known seabed type and with appropriate “ground truthing”, these indices can provide continuous classification of basic physical seabed types for a cruise. As different echosounders and echosounder settings were used in the collection of RoxAnn data, and there was no way to calibrate the system between cruises, the data from each cruise was treated separately in subsequent analyses. On completion of these separate analyses the acoustic information was stratified into the same biological and substrate classes.

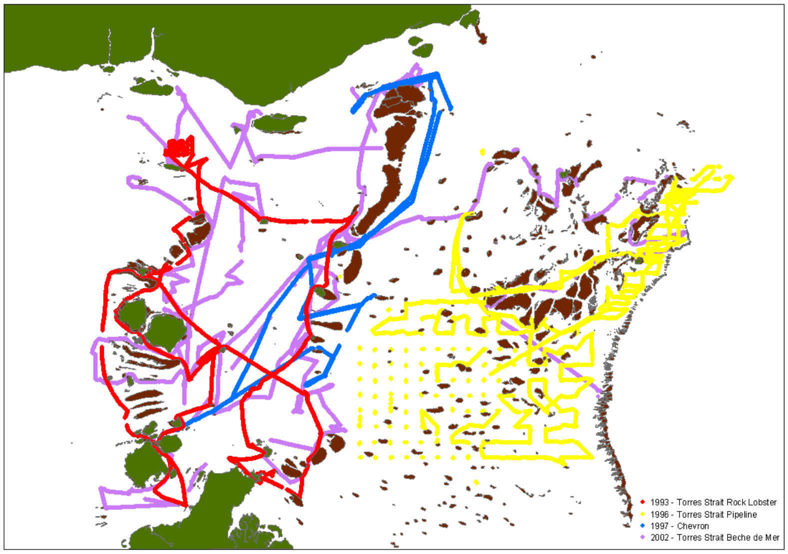


Figure 2.5-2 The geographic coverage for each of the four RoxAnn acoustic sampling cruises.

Ground truth information was available, at the scale of 0.001 (or approximately 110 m^2) decimal degrees, on the biological nature of the seabed, as well as the type of seabed substrate encountered. The biological ground truth information was categorized, on the density of biological organisms, into four classes of seabed; no benthos, very sparse benthos, sparse benthos, and dense benthos. While, the substrate ground truth information was categorized into five classes of seabed; mud (or silt), sand, rubble, sparse rock, and dense rock.

The RoxAnn acoustic depth, roughness and hardness information was aggregated on a similar grid scale (used throughout this project) of 0.01 decimal degrees (or approximately 1.1 km^2). The acoustic information within each 0.01 decimal degree grid cell was aggregated and five basic statistics generated. Statistics of minimum, maximum, median, mean and standard deviation were produced for each acoustic measure (depth, roughness, and hardness), for each of the four cruises.

A diagrammatic summary of the data used to predict seabed strata is shown in Figure 2.5-3 below.

Maps of the raw acoustic index statistics available at the 0.01 degree grid scale are shown in Figure 2.5-4 below; mean depth is rendered in blue (min = 3 m, maximum = 150 m), mean acoustic roughness is rendered in red (minimum = 83, maximum = 2379), and mean acoustic hardness is rendered in green (minimum = 57, maximum = 1412). Though the information is presented here on the same scale it is important to note that the roughness and hardness statistics for the four different cruises are not directly comparable.

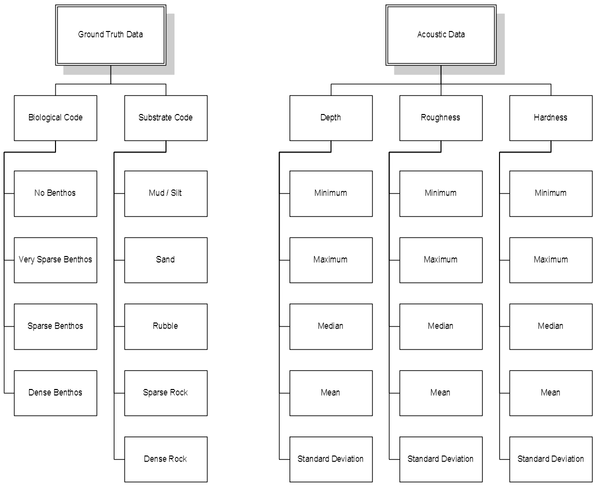


Figure 2.5-3: Schematic of Ground Truth classes and RoxAnn Acoustic predictor variables available (at the 0.01 degree grid scale), and used in the stratification analysis for each cruise.

2.5.2. Classification of Acoustics Data

To predict seabed stratum type from the grided acoustic information, the acoustic data were first merged by location with the benthic biological code and substrate code ground truth data (see data schematic Figure 2.5-3 above). This matched up areas where both acoustic and ground truth data were available. For each of the four cruises, the acoustic grided depth, roughness and hardness and roughness statistics were then classified, based on the benthic biological code and, in a separate analysis, substrate code ground truth data, using a linear discriminant function analysis that maximally separated the biological and substrate types based on the acoustic statistics (i.e. a total of eight discriminant function analyses were conducted). The linear discriminant function analysis, which was based on the four point biological and five point substrate classification systems, indicated that the acoustic data were able to discriminate satisfactorily among the seabed types.

The performance of the linear discriminant analysis was assessed and compared between cruises using the cross-validated error rate. The cross-validated error rate was calculated using the leave-one-out method of evaluation and gives the overall posterior probability of strata membership. The error rate performance is shown in the Table 2.5-2 below for each cruise (and graphically in Figure 2.5-5 below). Overall it was noted that the error rates were consistent across the cruises analysed; ranging from 13.6% to 33.9% for biological strata and from 21.1% to 41.6% for substrate strata.

Table 2.5-2 Cross-validated classification error rates by cruise (Note - expressed as % incorrectly classified).

Cruise (Individual DF Analysis)	Overall Error Rate	
	Biological	Substrate
1993 - Torres Strait Rock Lobster	35.3%	41.6%
1996 - Torres Strait Pipeline	13.6%	35.3%
1997 - Chevron	28.5%	21.1%
2002 - Torres Strait Beche de Mer	33.9%	40.4%

Mean Depth

Mean Roughness

Mean Hardness

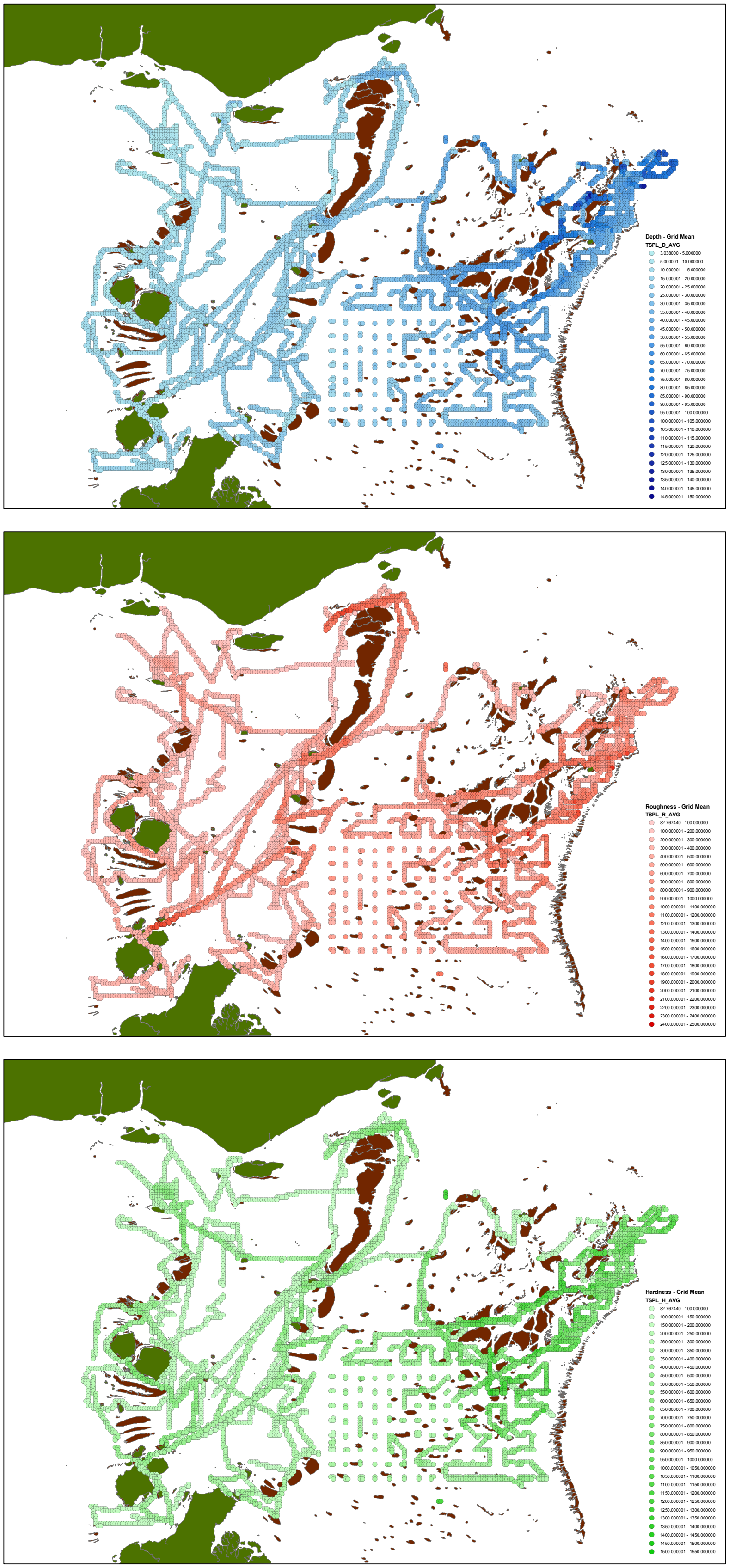


Figure 2.5-4 Maps of acoustic index statistics available at the 0.01 degree grid scale.

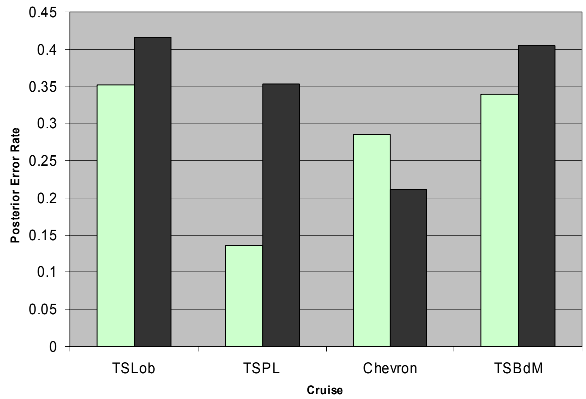


Figure 2.5-5 Discriminant function classification error rate performance for Sub_Code (■) and Biocode (□), by cruise.

An indication of which classes were the best described by the acoustic information, is given in Figure 2.5-6 below, comparing the mean classification performance across the cruises for each seabed strata. Overall, biological type strata were better described by the acoustics than substrate type strata were. This result was due to the Torres Strait Pipeline cruise biological classification performing significantly better than any of the other seven discriminant function analyses. Of the biological strata, the very sparse benthos and dense benthos classes were poorly described by the discriminant function; while sparse benthos was the best described with an error rate of 14.0%. Of the substrate strata, the sand and rubble substrate classes had poor error rates compared with the sparse rock class which was the best described class with an error rate of 11.1%.

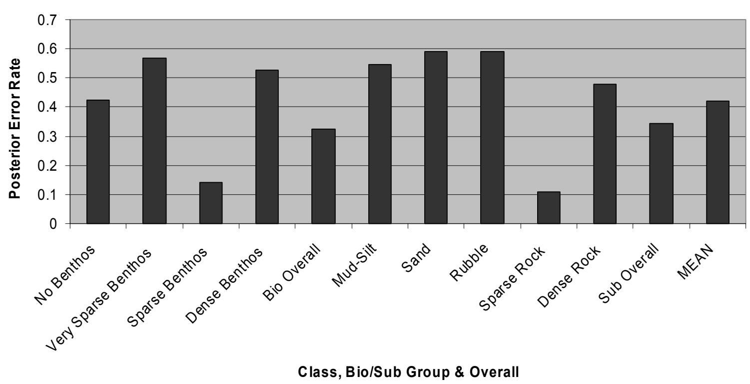


Figure 2.5-6 Mean discriminant function error rates by classification stratum class.

Table 2.5-3 and Table 2.5-4 (below) show the average classification characteristics for the linear discriminant functions across the four cruises. The tables show how the data was assigned from one classification strata into another. The values in the table are expressed as a percentage of the total data assigned to each particular stratum. The proportions of correctly classified data are shown on the diagonal of the table – where one class of data is correctly classified into that same class by the discriminant function. These classification tables highlight the performance of individual classes in the discriminant function analysis. It shows the classes where incorrectly classified data is assigned to, and particularly, where data is not assigned to adjoining strata, but, to classes that are dissimilar.

Erroneous classification results from discriminant function analysis may take either of two forms: The data may be wrongly assigned from the correct class into another incorrect class, this is knows as “miss-classification”; or the data may assigned from an incorrect class into the correct class, this is known as “false-classification”. The proportions of data which have been miss-classified are shown as the first (%) value in the table and may be read across the rows of the table (totalling 100%); while the proportions of data which have been falsely-classified are shown as the second (%) value in the table and may be read down the columns of the table (totalling 100%).

Table 2.5-3 Detailed biological classification table. Shows average proportion of data classified from one strata into another across all cruises, and highlights whether data was miss-classified (reading % totals across table rows) or falsely classified (reading % totals down table columns).

Miss Classification		**To**			
	False Classification	**No Benthos**	**V Sparse Benthos**	**Sparse Benthos**	**Dense Benthos**
From	**No Benthos**	50.3% 79.4%	0.0% 0.0%	47.9% 6.2%	0.0% 0.0%
	V Sparse Benthos	1.3% 3.9%	35.8% 66.5%	51.3% 11.9%	11.6% 5.6%
	Sparse Benthos	1.1% 10.6%	4.6% 28.9%	82.9% 65.3%	11.5% 18.9%
	Dense Benthos	0.9% 6.0%	1.1% 4.6%	30.8% 16.6%	67.3% 75.5%

Table 2.5-4 Detailed substrate classification table. Shows average proportion of data classified from one strata into another across all cruises, and highlights whether data was miss-classified (reading % totals across table rows) or falsely classified (reading % totals down table columns).

Miss Classification		**To**				
	False Classification	**Mud / Silt**	**Sand**	**Rubble**	**Sparse Rock**	**Dense Rock**
From	**Mud / Silt**	40.4% 76.1%	16.0% 9.4%	11.4% 8.3%	32.2% 3.7%	0.0% 0.0%
	Sand	1.5% 6.5%	45.9% 61.5%	5.5% 9.1%	47.0% 12.4%	0.0% 0.0%
	Rubble	1.0% 4.3%	0.0% 0.0%	34.7% 55.8%	61.8% 16.0%	2.5% 1.8%
	Sparse Rock	1.0% 13.1%	6.1% 25.7%	4.4% 22.8%	71.9% 59.7%	16.7% 37.4%
	Dense Rock	0.0% 0.0%	2.1% 3.3%	2.1% 4.1%	25.8% 8.3%	70.1% 60.8%

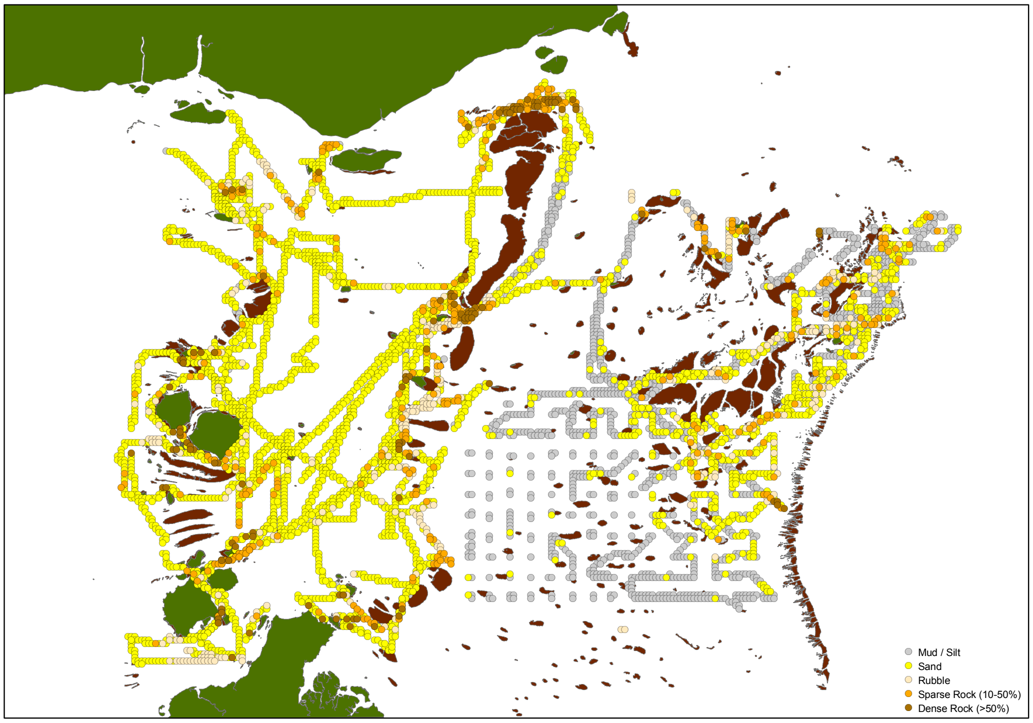
Classification results from the biological stratification of Torres Strait based on acoustic information (see Table 2.5-3 above) show that in terms of miss-classification performance the “No Benthos” and “Very Sparse Benthos” classes are poorly described, with 49.7% of the data from the “No Benthos” class being miss-classified as “Sparse Benthos”, and more “Very Sparse Benthos” being miss-classified as “Sparse Benthos” (51.3%) than “Very Sparse Benthos”. The “Sparse Benthos” and “Dense Benthos” classes have acceptable miss-classification performance with the majority of data being classified correctly (82.9% and 67.3% respectively) and a large proportion of miss-classified data falling into adjacent strata. When looking at false-classification performance all of the biological classes performed well with correct classification rates at 65% or above, with other data being falsely classified into adjacent strata. The classification table also highlights that the “Sparse Benthos” class is by far the most likely class resulting from any type of incorrect classification. The classification table for substrate stratification (see Table 2.5-4 above) shows that both of the “Rock” classes (Sparse and Dense) are well described in terms of miss-classification error with less than 30% of data being miss-classified. The other three substrate classes do not have good miss-classification error rates with less than 50% of data correctly classified. Of these three classes most of the data was miss-classified as “Sparse Rock”; with 32.2% of “Mud/Silt”, 47.0% of “Sand”, and 61.8% of Rubble miss-classified into the “Sparse Rock” strata. As was the case with the biological classification, the substrate classes are better performed in terms of false-classification, with all classes having more than 55% of their data correctly assigned.

For each cruise the acoustic depth, roughness and hardness statistics were classified using the discriminant function to assign seabed biological and seabed substrate stratum types to the remaining unclassified. This procedure classified the seabed in areas where only acoustic data existed, producing a map of seabed strata for the full length of each cruise track. With the predicted seabed strata as the basis, the analyses conducted for each of the four cruises were fused to produce a single seabed type stratification for all Torres Strait cruise tracks based on acoustic information.

Maps of the resulting biological and substrate stratification of the Torres Strait based on acoustic information are shown in Figure 2.5-7 below. The substrate stratification map highlights large areas of sand in central western Torres Strait with the presence of rock and rubble area in the channels between reefs (such as to the north and south of Warrior Reef). There is also an extensive area of south eastern Torres Strait which is predominated by mud and silt substrate graduating to sand, rubble and rock substrates in the north east (also through many inter-reefal channels). The biological stratification map outlines large areas of very sparse benthos in central western Torres Strait with increasing density in channels between the Warrior Reef complex. There is a large section of south eastern Torres Strait which has been classified as having no benthos, which corresponds to the large mud strata shown in the substrate map.

The acoustic stratification is consistent with the results of analyses based on other physical covariates. This confirms that acoustic information is a useful surrogate for seabed biological and substrate types and should be a data collection objective for any forthcoming fieldwork.

Substratum



Epi-benthos

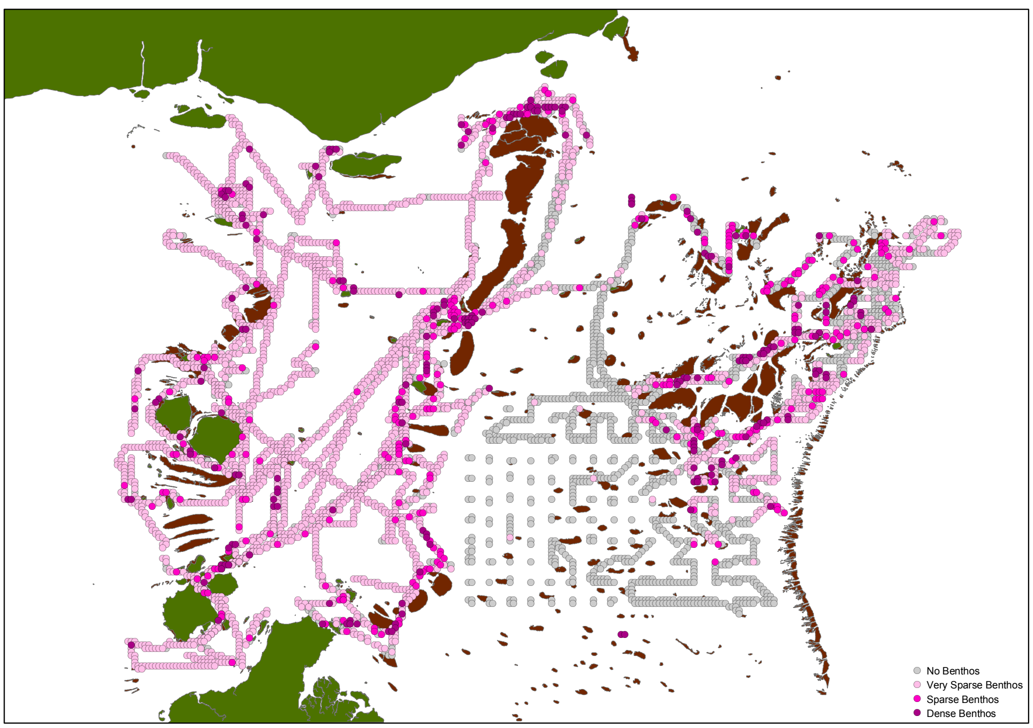


Figure 2.5-7 Maps of classified seabed substrate and biological strata for Torres Strait based on acoustic information.