

**School of Applied Science**

**An investigation of the relationship between seabed type and benthic and  
benthic-pelagic biota using acoustic techniques**

**Paulus Justiananda Wisatadjaja Siwabessy**

**This thesis is presented as part of the requirements for  
the award of the Degree of Doctor of Philosophy  
of the Curtin University of Technology**

**August 2001**

## Declaration

I hereby declare that this submission is my own work, and that to the best of my knowledge contains no material previously published or written by another person, or material which to a substantial extent has not been accepted for the award of any other degree of a university or other institute of higher learning, except where due acknowledgment is made in the text.

---

Paulus Justiananda W. Siwabessy

## **Abstract**

A growing recognition of the need for effective marine environmental management as a result of the increasing exploitation of marine biological resources has highlighted the need for high speed ecological seabed mapping. The practice of resource mapping making extensive use of satellite remote sensing and airborne platforms is well established for terrestrial management. Marine biological resource mapping however is not readily available except in part from that derived for surface waters from satellite based ocean colour mapping. Perhaps the most fundamental reason is that of sampling difficulty, which involves broad areas of seabed coverage, irregularities of seabed surface and depth. Conventional grab sampling techniques are widely accepted as a standard seabed mapping methodology that has been in use long before the advent of acoustic techniques and continue to be employed. However, they are both slow and labour intensive, factors which severely limit the spatial coverage available from practical grab sampling programs. While acoustic techniques have been used for some time in pelagic biomass assessment, only recently have acoustic techniques been applied to marine biological resource mapping of benthic communities. Two commercial bottom classifiers available in the market that use normal incidence echosounders are the RoxAnn and QTC View systems. Users and practitioners should be cautious however when using black box implementations of the two commercial systems without a proper quality control over raw acoustic data since some researchers in their studies have indicated problems with these two bottom classifiers such as, among others, a depth dependence.

In this thesis, an alternative approach was adopted to the use of echosounder returns for bottom classification. The approach used in this study is similar to that used in the commercial RoxAnn system. In grouping bottom types however, multivariate analysis (Principal Component Analysis and Cluster Analysis) was adopted instead of the allocation system normally used in the RoxAnn system, called RoxAnn squares. In addition, the adopted approach allowed for quality control over acoustic data before further analysis was undertaken. As a working hypothesis, it was assumed that on average  $\frac{\partial E1}{\partial R_0} = 0$  and  $\frac{\partial E2}{\partial R_0} = 0$  where E1 and E2 are the roughness and hardness indices, respectively, and  $R_0$  is the depth. For roughness index (E1), this was achieved by introducing a constant angular integration interval to the tail of the first bottom returns whereas for hardness index (E2), this was achieved by introducing a constant depth integration interval. Since three different frequencies, i.e. 12, 38 and 120 kHz, were operated, Principal Component Analysis was used here to reduce the dimensionality of roughness and hardness indices, formed from the three operated frequencies, separately. The *k*-means technique was applied to the first principal component of roughness index and the first principal component of hardness index to produce separable seabed types. This produced four separable seabed types, namely soft-smooth, soft-rough, hard-smooth and hard-rough seabeds. Principal Component Analysis was also used to reduce the dimensionality of the area backscattering coefficient  $s_A$ , a relative measure of biomass of benthic mobile biota.

The bottom classification results reported here appear to be robust in that, where independent ground truthing was available, acoustic classification was generally congruent with ground truth results. When investigating the relationship between derived bottom type and acoustically assessed total biomass of benthic

mobile biota, no trend linking the two parameters, however, appears. Nevertheless, using the hierarchical agglomerative technique applied to a set of variables containing the average first principal component of the area backscattering coefficient  $s_A$ , the average first principal component of roughness and hardness indices, the centroids of the first principal component of roughness and hardness indices associated with the four seabed types and species composition of fish group of the common species in trawl stations available, two main groups of quasi acoustic population are observed in the North West Shelf (NWS) study area and three groups are observed in the South East Fisheries (SEF) study area. The two main groups of quasi acoustic population in the NWS study area and the three main groups of quasi acoustic population in the SEF study area are associated with the derived seabed types and fish groups of the common species.

# Acknowledgments

There are many people whom I would like to acknowledge for their assistance throughout this project. Firstly, I would like to thank my principal supervisor Prof. John Penrose for his help and guidance during this project since early 1997, and also for his assistance with this thesis. I am greatly indebted to him.

I would like also to thank my associate supervisors Dr. David Fox of CSIRO Mathematical and Information Sciences, Perth for his invaluable assistance with geostatistics and spatial modelling, and Mr Rudy Kloser of CSIRO Marine Research, Hobart for his help, input and discussion during the course of the project. His expertise in fisheries acoustics was very valuable. Thanks are also extended to;

My associate supervisor Mr. Alec Duncan for his help with digital signal processing and my associate supervisor Mr. Rob McCauley for his assistance with biological and ecological aspects of this project.

Dr. Keith Sainsbury of CSIRO Marine Research, Hobart for making the acoustic and biological data from two surveys ss895 and ss797 in the North West Shelf region of Western Australia available for me to use in this project, for allowing me to participate in the ss797 fishery survey and for providing financial support during my last visit to CSIRO Marine Laboratories in Hobart.

Mr. Wade Whitelaw and Mr. Clive Stanley of CSIRO Marine Research, Hobart for their assistance in arranging for me to participate in the ss797 fishery survey and for providing the biological data needed.

Dr. Pavel Sakov and Mr. Tim Ryan of CSIRO Marine Research, Hobart for their assistance with the operational of the ECHO software and in particular with the algorithm to derive seabed roughness and hardness indices.

CSIRO Marine Research for providing technical assistance and facilities during the many visits to the Marine Laboratories in Hobart.

Prof. Brian O'Connor, the head of the School of Applied Science, for making available to me the excellent school facilities.

All members of the School and the Centre for Marine Science and Technology for making my stay such a pleasant one.

AusAID for making available financial support to conduct my PhD program at Curtin University of Technology, Perth, and AusAID Liaison Officer Mrs. Deb Pyatt and former AusAID Liaison Officer Mrs. Liz Campbell for their assistance since 1997.

Finally, I would like to thank my family (Lenny and Justin Siwabessy), my parents (Paul and Deetje Siwabessy) and friends for their support, encouragement and understanding during the course of the project.

This thesis is dedicated to Lenny, Justin, and my parents,

Paul and Deeje



# Table of contents

	Page
Declaration .....	i
Abstract .....	ii
Acknowledgments.....	v
Table of contents.....	viii
List of figures.....	xi
List of tables .....	xvi
Glossary of symbols.....	xviii
Chapter 1 Introduction .....	1
1.1. Background and significance.....	3
1.2. Research objectives.....	7
1.3. Thesis structure .....	9
1.4. Commercial bottom classifiers, RoxAnn and QTC View .....	10
1.4.1. Methods of operation .....	11
1.4.2. Problem identification.....	12
1.5. Classification methodology .....	16
1.6. Spatial methodology .....	18
Chapter 2 Physical concepts of multiple bottom reflection .....	22
2.1. Reflection of acoustic wave from the bottom surface .....	28
2.2. Acoustic scattering and reflection at rough surfaces .....	31
2.3. Backscatter of the first and second acoustic bottom returns.....	35
2.3.1. First acoustic bottom returns.....	38
2.3.2. Second acoustic bottom returns .....	40
2.4. Selection of integration interval.....	42
2.4.1. Constant angular integration interval for the tail of the first bottom returns.....	42

	2.4.2. Constant depth integration interval for the entire second bottom returns .....	56
Chapter 3	Research data collection and analysis.....	58
	3.1. Acoustic data.....	58
	3.1.1. Acoustic calibration .....	60
	3.1.2. Acoustic data quality control .....	62
	3.1.3. Acoustic data analysis.....	66
	3.1.3.1. Seabed parameters.....	68
	3.1.3.2. Relative measure of biomass ( $s_A$ ).....	69
	3.2. Net data .....	71
	3.2.1. Catch collection.....	71
	3.2.2. Net data analysis .....	71
	3.3. Underwater photographs.....	75
Chapter 4	Comparison between acoustic and catch estimates .....	78
	4.1. Introduction.....	78
	4.2. Species composition .....	80
	4.3. Acoustic estimates .....	82
	4.4. Acoustically equivalent net estimates.....	88
	4.5. Discussion.....	92
	4.6. Conclusions.....	99
Chapter 5	Results of bottom classifications for North West Shelf and South East Fisheries regions .....	102
	5.1. Characteristics of study areas .....	102
	5.2. First principal component of roughness and hardness parameters .....	110
	5.3. Training sets.....	120
	5.4. Along-track seabed classification .....	142
	5.5. Ground truth versus acoustically derived seabed type.....	156
	5.6. Kriged spatial modelling.....	166
Chapter 6	Relationship between common species and acoustically derived seabed type.....	188

6.1. Introduction.....	188
6.2. PCA of the area backscattering coefficient $s_A$ .....	189
6.3. The "quasi acoustic population" .....	193
6.3.1. The first principal component of roughness and hardness indices versus the first principal component of area backscattering coefficient .....	199
6.3.2. Fish group versus seabed type .....	206
Chapter 7 Conclusions .....	216
Appendix A Principal Component Analysis and Cluster Analysis.....	224
A.1. Principal Component Analysis (PCA).....	224
A.2. Cluster Analysis (CA).....	228
A.2.1. Measure of similarity .....	228
A.2.2. Hierarchical agglomerative methods .....	230
A.2.3. Iterative relocation ( <i>k</i> -means) technique.....	232
Appendix B Variogram and Kriging techniques.....	236
Appendix C Listing of source code of the employed MATLAB functions.....	240
Appendix D Shell and gawk scripts .....	243
Appendix E List of publications .....	251
References .....	252

## List of figures

	Page
2.1. Plot of the dependence of coherence parameter on the Rayleigh parameter for various rms heights of the rough surface.....	23
2.2. Dependence of the backscattering strength upon the incident angle from an abyssal plain bottom at various frequencies.....	25
2.3. Dependence of the backscattering strength upon the incident angle from a very rough bottom at various frequencies and from various bottom types.....	27
2.4. Geometry of the reflection at an interface between two media. ....	29
2.5. Representation of the region of integration and the associated vectors. ....	33
2.6. Geometry of the monostatic scattering arrangement. ....	36
2.7. Geometry of acoustic bottom returns.....	39
2.8. The average backscattering strength versus depth at a nominal bottom depth of 28m. ....	46
2.9. The average backscattering strength versus depth at a nominal bottom depth of 50m. ....	47
2.10. The average backscattering strength versus depth at a nominal bottom depth of 74m. ....	48
2.11. The average backscattering strength versus depth at a nominal bottom depth of 98m. ....	49
2.12. Scatterplot of roughness index against depth for three operated frequencies. Start angle of $8^\circ$ plus 3 m offset corresponds to the oblique depth at which backscattering strength is 35 dB below the maximum peak. ....	51
2.13. Scatterplot of roughness index against depth for three operated frequencies. Start angle of $14^\circ$ plus 3 m offset corresponds to the oblique depth at which backscattering strength is 45 dB below the maximum peak. ....	52
2.14. Scatterplot of roughness index against depth for 12 kHz data set. ....	54
2.15. Plot of oblique depth at which backscattering strength is 45 dB below the maximum peak of the first bottom echo for the start angle versus bottom depth and the horizontal distance off the axis resulting from the start angle versus bottom depth.....	55

3.1.	Flow diagram of data collection and processing system from various samplers used in this study.....	59
3.2.	Representative example of echogram for 120 kHz data set.....	63
3.3.	Example of corrupted echogram and procedure to correct spiky pings due to a time jitter.....	64
3.4.	A snapshot of the MS Access relationship between data tables within the database to extract the GPS data and import to the acoustic indices integrated data to fill the missing Longitude and Latitude data.....	67
4.1.	Schematic explanation of nominal position of trawl and positions of trawl shifted backward and forward around the nominal one.....	83
4.2.	Comparisons between the acoustic $s_A$ estimates at the locations provided from the vessel's logbook (on x-axis) and those from several lagged locations of trawls (on y-axis).....	84
4.3.	Scatterplot of acoustic $s_A$ estimates (abscissa) and net $s_A$ estimates (ordinate) between 0.5 to 2.5 m above the seabed in 71 trawl stations.....	87
4.4.	Beamwidth effect on a highly patchy distribution of fish.....	93
4.5.	Scatterplot of log transformed acoustic $s_A$ estimates and log transformed net $s_A$ estimates of single dominant species.....	95
4.6.	Scatterplot of acoustic $s_A$ estimates with a correction for the dead zone (abscissa) and net $s_A$ estimates (ordinate) 71 trawl stations.....	98
4.7.	Scatterplot of the squared difference between both $s_A$ estimates prior to and after the introduction of the ADZ compensation.....	98
4.8.	Scatterplot of log transformed acoustic $s_A$ estimates and log transformed net $s_A$ estimates of species with length frequency measurements available.....	100
5.1a.	Map of the Australian continent with the two continental shelves, north-west (NWS) and south-east (SEF), and the bathymetry around the continent from 50 to 500 m (light to dark blue).....	103
5.1b.	Map of the north-west continental shelf of Australia together with the bathymetry around the continental shelf from 50 to 500 m (light to dark blue).....	104
5.1c.	Map of the south-east continental shelf of Australia together with the bathymetry around the continental shelf from 50 to 500 m (light to dark blue).....	105
5.2a.	Seabed relief of the north west shelf continental shelf of Western Australia, known as the North West Shelf (NWS) region, from the coastline extending up to 180-m depth contour.....	107

5.2b.	Surface sediments of the south east continental shelf of Australia, known as the South East Fisheries (SEF) region. ....	111
5.3.	Scatterplot of acoustical seabed indices for the NWS study area. ....	112
5.4.	Scatterplot of acoustical seabed indices for the SEF study area. ....	113
5.5.	Scatterplot of RoxAnn roughness and hardness indices with depth. ....	115
5.6.	Scree plot of roughness parameter and hardness parameter for NWS study area. ....	117
5.7.	Scree plot of roughness parameter and hardness parameter for SEF study area. ....	117
5.8.	Representative examples of seabed images taken by a 35 mm Photosea 1000 camera system. ....	121
5.9.	Map of the study areas, the complete track, the useful track where E1 and E2 indices are available, the coastline and the bathymetry. ....	122
5.10.	Representative examples of echograms of derived seabed class that corresponds to Figure 5.8. ....	123
5.11.	Dendrogram showing four derived seabed classes from photographic records. ....	131
5.12.	Scatterplot of hardness index versus roughness index of the training sets comprising 4 different classes of the seabed for the SEF study area. ....	135
5.13.	Scatterplot of hardness index versus roughness index of the training sets comprising 4 different classes of the seabed for the NWS study area. ....	136
5.14.	Scatterplot of first principal component of roughness index versus first principal component of hardness index of the refined seabed classes. ....	141
5.15.	Scatterplot of first principal component of E1 index (PC1_E1) versus first principal component of E2 index (PC1_E2) of the entire data together with the four acoustically derived seabed classes. ....	144
5.16a.	Map of acoustically derived seabed types along the track, bathymetry, coastline and benthic habitat types for the NWS study area. ....	147
5.16b.	Map of acoustically derived seabed types along the track, bathymetry and coastline for the SEF study area. ....	148
5.17.	Seabed images representing five benthic habitat types in the NWS study area proposed by Sainsbury <i>et al.</i> (in prep). ....	157
5.18.	Histograms of acoustically derived bottom types and benthic habitat types for the NWS study area. ....	158
5.19.	Histograms of acoustically derived bottom types for the SEF study area in reference sites of known seabed types. ....	164

5.20.	Map of all the straight transects extracted from the useful track used in the autocorrelation analysis and the interpolation into unsampled areas by Geostatistics. ....	167
5.21.	Representative example of series of roughness index in one of the straight transects.....	169
5.22.	Representative example of series of hardness index of the similar straight transect in Figure 5.21.....	170
5.23.	The scree plot produced by applying the principal component analysis into a set of E1 and E2 given in Figures 5.21 and 5.22.....	171
5.24.	Series of the first principal component of roughness index and hardness index that correspond to Figures 5.21 to 5.23.....	173
5.25.	Plot of autocorrelation function of the first principal component of roughness index of the series given in Figure 5.24(a) and hardness index of the series shown in Figure 5.24(b).....	174
5.26.	Variogram of the first principal component of roughness and hardness indices for the NWS study area and the SEF study area.....	175
5.27.	Map of region of Kriging within the autocorrelation characteristic length of the straight lines shown in Figure 5.20.....	177
5.28a.	Contour plot of bottom roughness structure and bottom hardness structure using Kriging technique.....	178
5.28b.	Contour plot of standard errors for (i) bottom roughness and (ii) bottom hardness associated with Figure 5.28a for the NWS region by using the Kriging technique.....	179
5.28c.	Plot of Krige estimated index versus actual index for (i) bottom roughness and (ii) bottom hardness for the NWS region using the Kriging technique.....	180
5.29a.	Contour plot of bottom roughness structure and bottom hardness structure using Kriging technique.....	181
5.29b.	Contour plot of standard errors for (i) bottom roughness and (ii) bottom hardness associated with Figure 5.29a for the SEF region by using the Kriging technique.....	182
5.29c.	Plot of Krige estimated index versus actual index for (a) bottom roughness and (b) bottom hardness for the SEF region using the Kriging technique.....	183
5.30.	Contour plot of seabed type derived from Kriged PC1_E1 and Kriged PC1_E2 indices given in Figures 5.28 and 5.29 for (a) the NWS study area (b) the SEF study area.....	186
6.1.	Scree plot of area backscattering coefficient, $s_A$ , formed from the three operated frequencies.....	190

6.2.	Series of acoustical $s_A$ estimates at three operated frequencies showing the beamwidth effect on a highly patchy distribution of fish for the two observed patterns in the component loadings. ....	192
6.3.	Series of acoustical $s_A$ estimates at three operated frequencies for those first principal components having the highest variance. ....	194
6.4a.	Pictorial plot of four derived groupings of fish species given as pie charts overlaid into four seabed types given as a Cartesian plot of PC1_E2 versus PC1_E1 for the NWS study area. ....	200
6.4b.	Pictorial plot of four derived groupings of fish species given as pie charts overlaid into four seabed types given as a Cartesian plot of PC1_E2 versus PC1_E1 for the SEF study area. ....	201
6.5a.	Scatterplot of the first principal component of roughness index and hardness index with the first principal component of area backscattering coefficient averaged within the trawl transect ( $\approx 30$ minutes). ....	203
6.5b.	Scatterplot of the first principal component of roughness index and hardness index with the first principal component of area backscattering coefficient averaged within a 15-minute interval from all the acoustic records. ....	204
6.5c.	Scatterplot of the first principal component of roughness index and hardness index with the first principal component of area backscattering coefficient averaged within a 5-minute interval from all the acoustic records. ....	205
6.6.	Dendrogram of the quasi acoustic population for the NWS and the SEF study areas. ....	207
6.7a.	Map of four main groups of quasi acoustic population and fish group composition in % by weight (pie charts) overlaid to the acoustically derived seabed type of the NWS study area. ....	209
6.7b.	Map of four main groups of quasi acoustic population and fish group composition in % by weight (pie charts) overlaid to the acoustically derived seabed type of the SEF study area. ....	213



## List of tables

		Page
2.1.	Bottom acoustic absorption coefficient for various bottom types .....	44
3.1.	Calibration settings for the SIMRAD EK 500 echosounder.....	61
3.2.	List and description of variables scored from seabed images.....	77
4.1.	A summary of catches in 72 trawl stations. ....	81
4.2.	The correlation coefficients between the acoustic $s_A$ estimates in the nominal position and those in the 6 different lagged positions.....	86
4.3.	A summary of the statistical test assessing the difference of the correlation coefficient between the acoustic $s_A$ estimates in the nominal position and those in the 6 different lagged positions.....	86
4.4.	Proportion of species with and without length-frequency measurements, the number of species with length-frequency measurements, the number of total species and the number of different types of swimbladder present at 72 trawls. ....	90
4.5.	$s_A$ estimate, mean length, bladder type, proportion and weight of selected species with and without length-frequency measurements available. ....	91
5.1.	Eigenvalues and total variances of principal components for roughness and hardness indices at NWS and SEF study areas. ....	119
5.2.	Component loadings of roughness and hardness parameters for NWS and SEF study areas. ....	119
5.3.	A set of variables used in the hierarchical agglomerative clustering technique to derive distinguishable seabed classes from photographic records. ....	129
5.4.	Description of overall seabed types in trawl stations available shown as open circle in Figure 5.9.....	130
5.5.	Centroids of four seabed types.....	134
5.6.	Euclidean distances between four seabed classes estimated for single frequency and multi-frequency. ....	138
5.7.	Percentage of the assigned class derived by using the $k$ -means technique applied to the training sets of the SEF and NWS study areas, separately, using the centroids given in Table 5.5. ....	139
5.8.	Centroids after all data points have been assigned to the four seabed classes for the SEF study area and the NWS study area separately.....	145

5.9.	Euclidean distances between four seabed classes estimated from the centroids after all data points have been assigned to the four seabed classes for the SEF study area and the NWS study area.....	145
5.10.	The total area, the area of acoustic coverage and the area of photographic coverage in m <sup>2</sup> within four derived areas (i), (ii), (iii) and (iv) shown in Figures 5.1 and 5.16a. ....	150
5.11.	Confusion matrix between acoustically derived seabed types and benthic habitat types for the NWS study area.....	161
5.12.	Confusion matrix between acoustically derived seabed types and seabed types of reference sites for the SEF study area. ....	161
6.1.	Parameters adopted to form quasi acoustic population in the NWS study area. ....	196
6.2.	Parameters adopted to form quasi acoustic population in the SEF study area. ....	197

# Glossary of symbols

## Roman letter symbols.

$a, b$	regression parameters of the length-weight relationship of fish.
$A$	insonified area.
$A_t$	towing area.
<b>B</b>	between-cluster covariance matrix used in Cluster Analysis.
$c$	speed of sound.
$c_\tau$	offset derived from simple regression of the bottom depth and the oblique depth.
<b>C</b>	sample covariance used in Principal Component Analysis.
$d_{\theta a}$	oblique depth at start angle of the integration interval for the tail sector of the first acoustic bottom return.
$d_{\theta b}$	oblique depth at stop angle of the integration interval for the tail sector of the first acoustic bottom return.
$d(P,Q)$	distance between points P and Q used in Cluster Analysis.
$d(x,y)$	Euclidean distance used in Cluster Analysis.
$d_a$	start depth of the integration interval for the entire second acoustic bottom return.
$d_b$	end depth of the integration interval for the entire second acoustic bottom return.
$d_f$	end depth for integration interval that corresponds to $d_{\theta b}$ and $d_b$ .
$d_s$	start depth for integration interval that corresponds to $d_{\theta a}$ and $d_a$ .
$\delta d$	acoustic sampling interval.
<b>D</b>	matrix of similarity or distance measures used in Cluster Analysis.
$\mathbf{e}_j^T$	transposed eigenvectors.
$\mathbf{e}_j$	eigenvectors used in Principal Component Analysis.

<b>E</b>	matrix of series of eigenvectors used in Principal Component Analysis.
E1	roughness index.
E2	hardness index.
<i>G</i>	Green's function.
$G_o, G(\theta)$	transducer gain.
<i>h</i>	lag distance used in variogram and Kriging techniques.
$H_{dz}$	effective lost height due to acoustic dead zone.
<i>j</i>	$\sqrt{-1}$ .
<i>k</i>	acoustic wave number.
$\mathbf{I}_j$	eigenvalues used in Principal Component Analysis.
<i>L</i>	fish length.
<b>L</b>	diagonal matrix of series of eigenvalues $\mathbf{I}_j$ used in Principal Component Analysis.
<i>m</i>	slope derived from simple regression of the bottom depth and the oblique depth.
$m_a(\theta)$	scattering coefficient of the sea surface.
$m_s$	acoustic scattering coefficient.
$n_l$	number of integration layers.
$n_p$	number of pings.
$n_s$	number of $s_A$ values.
$n_{sl}$	number of catch.
<i>N</i>	number of data.
$N(h)$	number of any possible pairs of data points used in variogram and Kriging techniques.
<i>p</i>	acoustic pressure.
$p_c$	received acoustic pressure due to coherent returns.
$p_{ic}$	received acoustic pressure due to incoherent returns.
<i>P</i>	power.
PC1_E1	first principal component of roughness index.
PC1_E2	first principal component of hardness index.

PC1_SA	first principal component of area backscattering coefficient $s_A$ .
PC <sub><i>j</i></sub>	$j^{\text{th}}$ principal component.
<b>R</b>	correlation matrix used in Principal Component Analysis.
$r_o$	reference range (= 1 m).
$R_o$	bottom depth.
$R, r$	range.
$s_A$	area backscattering coefficient.
$\bar{s}_A$	average area backscattering coefficient.
$s_{A(\text{catch})}$	net-derived area backscattering coefficient.
$s_V$	volume backscattering coefficient.
$S_V$	volume reverberation signal ( $10 \log_{10}(s_V)$ ) in dB.
$S_n$	two-sided power spectra density (PSD).
$T$	correlation length of the surface roughness.
$TS$	target strength.
$u$	magnitude of the particle velocity.
<b>u</b>	vectors of particle velocities.
$U$	inbound wave.
$\mathbf{v}_j$	column vector of constants used in Principal Component Analysis.
$\mathbf{v}_j^T$	transposed column vector of constants $\mathbf{v}_j$ .
<b>V</b>	matrix of eigenvectors used in Principal Component Analysis.
$w$	weights.
$W$	fish weight.
$W_{\text{inf}}$	asymptotic weight of fish.
<b>W</b>	within-cluster covariance matrix used in Cluster Analysis.
<b>W<sup>k</sup></b>	matrix of weights.
$x$	series of data, i.e. series of the first principal component of E1 (PC1_E1) or E2 (PC1_E2).
$\bar{x}$	arithmetic mean of series of data $x$ .
$x'$	mean-removed series of data $x$ .

$x''$	mean-removed, detrended series of data $x$ .
$x_n'''$	bandpassed, mean-removed, detrended series of data $x$ with zero padded.
$X_n$	spectrum in positive frequency range.
$X_n^*$	complex conjugate of $X_n$ .
<b>X</b>	matrix data of the original data set, i.e. matrix data formed separately by E1, E2 or $s_A$ values from the three operated frequencies, used in Principal Component Analysis.
$Y$	regionalised variable used only in variogram and Kriging techniques.
$Y^k(x_o)$	estimated Krige value at unsampled point $x_o$ .
$Z$	acoustic impedance.

### **Greek letter symbols.**

$\alpha$	seawater absorption coefficient.
$\Delta$	sampling time unit.
$\Gamma(h)$	semi-variogram used in variogram and Kriging techniques.
$\varphi$	angle of the incident acoustic wave towards the vertical plane.
$\varphi_n$	autocorrelation (inverse Cosine transform of the PSD).
$\lambda$	acoustic wavelength.
$\theta_o$	half beamwidth.
$\theta_a$	start angle of the integration interval for the tail sector of the first acoustic bottom return.
$\theta_b$	stop angle of the integration interval for the tail sector of the first acoustic bottom return.
$\rho$	density of the media.
$\sigma$	rms deviation of the surface irregularities.
$\sigma_k$	Kriging variance.
$\sigma_{sl}$	backscattering cross section.
$\tau$	pulse length.

$\psi$	equivalent beamwidth.
$\mathfrak{R}$	acoustic pressure reflection coefficient.
$\langle \mathfrak{R} \rangle_G$	theoretical Gaussian surface value.
$\mathfrak{R}_c$	coherent reflection coefficient.
$\mathfrak{R}_{ic}$	incoherent reflection coefficient.
$\mathfrak{S}$	scattering function.

# Chapter 1

## Introduction

The practice of resource mapping making extensive use of satellite remote sensing and airborne platforms is well established for terrestrial management. Marine biological resource mapping however is not equivalently available except in part from that derived for surface waters from satellite based ocean colour mapping. Perhaps the most fundamental reason is the sampling difficulty, which involves broad areas of seabed coverage, irregularities of seabed surface and depth. Except in very shallow waters the absorption of electromagnetic waves by seawater significantly limits the applicability of space and airborne sensors.

Underwater acoustics has received a special place in fisheries research. It has shown its power in estimates of the abundance of fish stocks, pelagic in particular. Yet, underwater acoustic technology needs further development before it will be fully utilised in fisheries research. The development, though, is underway, for instance a cooperative research program in fisheries by means of acoustics by EC countries, with the intention that standardised techniques and devices will be established, applicable to their whole waters. Active underwater acoustic devices provide return information not only from fish but from the seabed as well. However, relating acoustic data to yield relationships between fish communities and epibenthos has received less attention (Foote, 1996). Only a few papers on this have been published, notably Scott (1982), Burns *et al.*, (1989), Walsh (1992), Magorrian *et al.*, (1995), Greenstreet *et*



*al.*, (1997), Collins and McConnaughey (1998) and Bax *et al.*, (1999). CSIRO Marine Research is very interested in developing a way of relating fish community and epibenthos from acoustic returns. Advantages of using acoustic techniques over conventional trawl methods are fourfold. Firstly, a surface vessel especially during net deployment may influence biota by virtue of its radiated noise. Secondly, an echosounder is probably unlikely to influence biota because its bandwidth is outside the sensitivity ranges of most biota. Thirdly, a bottom trawl will alter benthic structures and, to some extent or other, must be selective. Lastly, acoustic techniques offer a high speed of survey in comparison to conventional net-based methods.

Fishers have traditionally used acoustic bottom echoes from nominally normal incidence echosounders for seabed characterisation. Only recently has attention been given to the use of quantified acoustic bottom echoes from echosounders for seabed classification in marine ecological applications. Two commercial bottom classifiers available in the market that use normal incidence echosounders are the RoxAnn and QTC View systems. Both systems use shape and energy features contained in range corrected acoustic bottom signals. Orłowski (1984) and Chivers *et al.* (1990) have used the energy features contained in the first and second acoustic bottom echoes as seabed descriptors, and Heald and Pace (1996) provide a theoretical background of relationships between energy features of the two echoes and seabed parameters. Lurton and Pouliquen (1992) and Collins *et al.* (1996) on the other hand use only a detailed analysis of the first acoustic bottom echoes. Only recently have studies on marine biological resource mapping of benthic communities used these acoustic techniques. Examples include Magorrian *et al.* (1995), Greenstreet *et al.* (1997), Kaiser *et al.* (1998), Sorensen *et al.* (1998) using the commercial RoxAnn system, Prager *et al.* (1995) using the commercial QTC-view system, and Bax *et al.* (1999),

Siwabessy *et al.* (1999, 2000) and Kloser *et al.* (2001) using the RoxAnn-like technique of the energy features of the first and second acoustic bottom returns.

### **1.1. Background and significance**

In 1978, a year prior to the declaration of the 200 nmi Australian fishing zone, CSIRO Marine Research began studying the fisheries of the North West Shelf (NWS) of Western Australia. Since 1986, there have been eight management programs in the NWS. Each program was a continuation of earlier management programs and aimed to measure changes in benthic community and abundance in the intervening years. Similarly, CSIRO Marine Research has established an active management program in the South East Fisheries (SEF) area off the Victoria and New South Wales coasts. Apart from biological samples taken in both areas of investigation, acoustic measurements have also been conducted. A scientific SIMRAD EK 500 echosounder has been used to collect the acoustic data in both areas.

Previous results from the NWS indicated that the multispecies fish community in the area of investigation was habitat dependent and this might be possibly applicable to other tropical regions of northern Australia (Sainsbury *et al.*, 1997). In addition, Sainsbury *et al.*, (1997) indicated that the historical changes in species composition in the NWS partly resulted from trawl-induced modification of the epibenthic habitat.

In many studies of multispecies fish communities in relation to their habitat, the habitat has often been defined exclusively in terms of either physical oceanography parameters such as temperature, salinity, depth and mixed layer depth or chemical oceanography parameters such as some specific chemical contents. In addition, these studies mainly covered pelagic fish rather than demersal fish and other

components of benthic system just above the seabed. Through such studies, the relationship between the community and the so-defined habitat has become better understood. For instance, Blaber *et al.*, (1994) indicated a relationship between fish community and seawater depth. The scope of habitat in fact is not limited only to physical and chemical factors but also covers biological factors such as the nature of the biotic community that occupies seabed surfaces. Biological habitat factors in relation to demersal fish communities appear to have received less attention than physical and chemical factors. However, Sullivan *et al.*, (1991) and Sainsbury (1991) have pointed out the dependence of some fish on epibenthic organisms. The probability of occurrence of *Lethrinus* and *Lutjanus* on one hand was significantly higher in areas with large (>25 cm) epibenthic organisms than in areas with no large epibenthos and on the other hand that of *Nemipterus* and *Saurida* showed otherwise (Sainsbury, 1991).

The NWS is one of the northern Australian tropical regions which support a diverse Indo-West Pacific fish fauna (Sainsbury *et al.*, 1997). Some of the high-valued fish from the genera *Lethrinus* and *Lutjanus* in particular have made the NWS their home. The area has been fished mostly by foreign distant-water fleets, particularly prior to the declaration of the 200 nmi Australian fishing zone. Between 1959 and 1963, the Japanese introduced stern trawlers targeting the large stocks of the commercially high-valued fish of the genus *Lethrinus* and this marked the first major commercial fishing taking place on the NWS (Sainsbury *et al.*, 1993). Taiwanese trawlers were the second foreign distant-water fleets coming to and fishing in the area. Taiwanese pair trawlers began fishing in the area in 1972 (Sainsbury *et al.*, 1993). After the declaration of the 200 nmi Australian fishing zone, the Taiwanese continued

fishing in the area under a licence agreement with the Australian Government until 1989 (Sainsbury *et al.*, 1993).

Although a small domestic trawl operation had also become established by the time the Taiwanese withdrew fishing from the area, changes in biotic community in the area had taken place (Sainsbury *et al.*, 1993; Sainsbury *et al.*, 1997). The composition of the fish catch changed from having been dominated by the genera *Lethrinus* and *Lutjanus* to being dominated by the less-valued fish of the genera *Nemipterus* and *Saurida* (Sainsbury *et al.*, 1993; Sainsbury *et al.*, 1997). In line with the replacement of the dominant fish catch, the next change observed was the decline of the quantity of the epibenthic organisms caught (Sainsbury *et al.*, 1993).

The SEF is the Australia's most important fishery for domestic scalefish markets (Tilzey, 1994). The area has been fished since the early 1900s. The SEF trawl sector stretches from Sydney southwest around Tasmania to Kangaroo Island in South Australia (Caton *et al.*, 1997). Trawlers originally fished continental shelf species in depths of less than 200 m and targeted mainly *Neoplatycephalus richardsoni*. After 1930, the subsequent decline of *Neoplatycephalus richardsoni* led to increasing capture of other continental shelf species. From 1915 to 1950, the SEF was dominated by steam trawlers. Danish seiners then became dominant in the SEF from the early 1950s to the early 1970s. After the early 1970s, the SEF has been dominated by modern trawlers. Due to the virtual absence of management restrictions, the NSW-based Danish seine fleet expanded into Victoria in the 1950s and into deeper waters in the 1970s and eventually spread as far as Tasmania. By the early 1980s, the number of the NSW-based fleet had increased to double that of 1970. Despite various management interventions aimed at limiting fleet expansion, fishing

effort continues to increase. The quota system was first introduced in 1988 in the SEF for eastern gemfish. In 1990, the quota system was also introduced for orange roughy. In 1992, it expanded into other 15 major commercial species or species groups in the SEF. These quota species account for more than 80% of the recorded landed catch weight.

In their study to identify and map different habitat types and their relationship to invertebrate and fish communities in one mesohabitat (an area measured in km<sup>2</sup>) within megahabitat (an area measured in 100s or greater of km<sup>2</sup>) in the northeastern remote zone of the SEF, Bax *et al.* (1999) concluded that the discrimination of habitats from visual inspection of echograms and limited verification with physical sampling could provide sufficient information for spatial management. From visual inspection of echograms, they divided this mesohabitat into three distinct macrohabitats namely soft, rough and hard habitats. They also post processed the stored acoustic data to produce two descriptors of the RoxAnn system, i.e. E1 and E2 which were also adopted in this study and are fully defined in sections 1.4, 2.3.1 and 3.1.3.1, and equations 2.6 and 3.3. The subjective RoxAnn square was adopted to divide the scatterplot of E1 and E2 into different habitat types. They found a one to one correspondence between habitats determined by these two techniques. In addition, they found that ten species associated most with the dissimilarity between macrohabitats were *Mustelus antarcticus*, *Latridopsis forsteri*, *Neoplatycephalus richardsoni*, *Scomber australasicus*, *Caesioptera lepidoptera*, *Pristiophorus cirratus*, *Squalus megalops*, *Pseudophycis bachus*, *Trachurus declivis* and *Pseudocaranx dentex*. Most of these discriminating species were common to three macrohabitats but present at markedly different abundances. Two species (*Neoplatycephalus richardsoni* and *Pseudophycis bachus*) occurred only in the soft habitat and two other

species (*Caesioptera lepidoptera* and *Latridopsis forsteri*) associated with rough habitat were scarce on the soft habitat (Bax *et al.*, 1999).

## **1.2. Research objectives**

The main objective of this study was to establish the relationship, if any, between seabed types and associated demersal fish groups in areas of the North West Shelf (NWS) and the Southeast Fishery (SEF) regions. This calls for an optimal post-classification technique that gives robust seabed classification results. A combination of RoxAnn-like techniques and multivariate statistical tools were used for this study. Training sets were established from ground truthing and were tested against the classification technique adopted. Results from these training sets were then applied to the remaining data. This study also called for a biomass assessment of mobile benthos and thus, a comparison between acoustically derived biomass and that derived from nets. The importance of this was to decide whether the acoustically derived biomass or the net-derived biomass was suitable to be included when investigating the relationship between acoustically derived seabed types and demersal fish groups above them.

The results of this study are considered to be useful for monitoring fisheries management and spatial management of marine fisheries. Acoustic data processing techniques adopted in this study provide an alternative to the available commercial bottom classifiers. Like the commercial bottom classifiers, the techniques presented here are applicable for all normal incident echosounders as long as they are able to provide first and second bottom echoes that can be stored in digital formats. The techniques presented here provide post processing analysis that enables users/practitioners to carry out data quality control before conducting further analysis.

Although the technique used in this study for seabed classification is based on the existing RoxAnn-like system, this current study has contributed in enhancing the technique to involve of several acoustic frequencies using multivariate analysis, in particular Principal Component Analysis (PCA). In addition, this study has contributed in developing a classification procedure using Cluster Analysis, in particular the iterative relocation (*k*-means) technique. This study has also contributed in producing contour maps of derived seabed types in the NWS and the SEF regions. Lastly, this current study has contributed in seeking links between benthic habitats and derived seabed types for the areas covered experimentally.

The two study areas discussed in this work are situated off the north-west coast of Western Australia (NWS study area) between 114.75° to 119°E and 18.5° to 21°S, and off the south-east of mainland Australia (SEF study area) between 36° and 39°S. The shelf in the NWS region stretches from 10 m depth to 200 m depth, and extends to 170-200 m depth in the SEF study area. The NWS study area is subdivided into three management zones, Barrow Island, Legendre and Hedland zones, based on the history of fishing activities. Based on the current zonation by the Australian Fisheries Authority, the SEF study area includes parts of the north and east management zones. Data used in this study come from 2 different surveys in the NWS region between August and September 1995 (ss895) and 1997 (ss797) and from a single survey in the SEF study area between November and December 1996 (ss696). Acoustic data were collected in both studies by the FRV *Southern Surveyor* using a calibrated SIMRAD EK500 scientific echosounder with hull-mounted transducers of three different frequencies, 12 (single beam), 38 (split beam) and 120 kHz (split beam).

### 1.3. Thesis structure

This thesis is divided into seven chapters. In the first chapter, background and significance, and research objectives are presented. Commercial bottom classifiers and some issues concerning them are also discussed in this chapter. The first chapter concludes with an introduction of the classification and spatial methodology (which are reviewed in Appendices A and B) adopted in this study.

Chapter 2 reviews the physical principles of first and second acoustic bottom returns. Interpretation of the first and second acoustic bottom returns in relation to seabed parameters is discussed. It also introduces the technique used to select the angular integration interval for the useful sector of the tail of the first bottom echo and the constant depth integration interval for the complete second bottom echo.

Chapter 3 describes research data collection and analysis. It includes acoustic data, net data and seabed photographs. This chapter also describes acoustic calibration and acoustic data quality control used in the present work.

In Chapter 4, a comparison between acoustically derived density estimates and that derived from nets is presented. Previous studies on this topic are also mentioned. The objective of this chapter is to explore the possibility of the inclusion of net derived  $s_A$  estimates to form the so-called "quasi acoustic population" in order to establish the relationship between seabed types and fish groups.

Chapter 5 presents the results of seabed classification in the NWS and SEF study areas. A comparison between acoustically derived seabed types and that derived from photographs and that based on reference sites is discussed.



Chapter 6 explores the relationship between the first principal component of roughness and hardness indices and the first principal component of the relative measure of biomass from acoustic data within trawl transects and from all the acoustic records along the track. Chapter 6 concludes with a discussion of the relationship between acoustically derived seabed types, total near-bottom mobile biomass and groups of common species of demersal fish.

The thesis concludes with Chapter 7 that presents final remarks.

#### **1.4. Commercial bottom classifiers, RoxAnn and QTC View**

The use of acoustic techniques for seabed classifications has become a widely accepted tool, especially on continental shelves. Perhaps the simplest implementation is to attach a commercial bottom classifier to the transducer of the existing echosounder. This allows for automatic data collection along the vessel's track. Relying solely on commercial bottom classifiers and letting them work as a black box for seabed classification may lead to misclassification. This is because the acoustical parameters derived by the systems are empirical, and quantitative physical theory that explains directly the relationship between inferred acoustical parameters and bottom interaction is lacking (Hamilton *et al.*, 1999). Two commercial bottom classifiers available in the market are the RoxAnn and QTC-View systems. The former is designed and manufactured by Marine Micro Systems Ltd. Of Aberdeen, Scotland, UK, and the latter by Quester Tangent Corporation, Sidney, Canada. In this section, these two commercial bottom classifiers are described. The discussion includes principles of operation and problem identification. Particular attention is given to the RoxAnn classifier as essentially a RoxAnn(-like) technique is adopted in this study.

#### 1.4.1. Methods of operation

Due to wavefront curvature, acoustic waves excited by a transducer experience three different situations on the seabed. Progressively, in time sequence, they will, on a flat, horizontal seabed, ensound a circle with annuli of increasing radii and reducing grazing angle on the seabed as time increases. The returned acoustic envelope at the transducer comprises two components; the total specular reflection and backscatter returns from particular annuli. A peak in the first part of the returned acoustic envelope is due to coherent components from specular reflection and a decaying tail is due principally to incoherent contributions from successive annuli. The shape of the returned acoustic envelope is a function of acoustic roughness; the rougher the seabed, the longer the tail will be. The shape of the returned acoustic envelope is also a function of acoustic hardness and depends upon frequency, pulse length, beamwidth and the characteristic acoustic impedance of the seabed. The term hardness is widely used in the literature and may be viewed as a descriptor of the contrast in acoustic impedance offered by the water-seabed interface.

The RoxAnn and QTC-View systems employ different principles. The RoxAnn system uses a multiple echo approach whereas the QTC-View uses an echo shape approach on a single echo. In the RoxAnn system, the first and second acoustic bottom returns are of interest. The first acoustic bottom return is a direct reflection from the seabed whereas the second acoustic bottom return has reflected twice at the seabed and once at the sea surface. Two parameters used in the RoxAnn system are E1 and E2. The E1 parameter is a measure of the energy in the tail of the first acoustic bottom return and the E2 parameter is a measure of the total energy of the complete second acoustic bottom return (Burns *et al.*, 1989). The E1 factor is derived by integrating only the tail sector of the first echo and provides an index of seabed

roughness whereas the E2 factor is obtained by integrating the complete second echo and gives an index of seabed hardness (Burns *et al.*, 1989). A scatterplot of E1 versus E2 is then produced. Seabed classifications are conducted by dividing the scatterplot into a set of the so-called RoxAnn squares; each of them is held to represent a particular seabed type or substrate.

Unlike the RoxAnn system, the QTC-View system examines the shape characteristic of the first bottom echo only although later versions of QTC-View plan to use both shape and energy parameters. The QTC-View system extracts 166 features of the returned envelope of the first bottom echo. The QTC-View parameters known as Q-values (Q1, Q2 and Q3) are chosen automatically by QTC software using principal component analysis to reduce the dimensionality of the extracted features (166 features plus combinations of them) into *low*-dimension features, Q-values (Prager *et al.*, 1995; Anon, 1995; Collins and McConnaughey, 1998). These Q-values are believed to represent most of the variation in the data. Q1, Q2 and Q3 are then plotted in 3-dimensional space for visual inspection of clustering. In supervised classification mode, class assignments are based on the multivariate distances between echoes to be classified and clusters that represent the acoustic properties of the seabed chosen by the operator. In unsupervised mode, the QTC-View software automatically provides classifications. In addition, the software provides a confidence estimate of each ping in choice of class.

#### 1.4.2. Problem identification

Despite the claim by the manufacturer that the RoxAnn system is not dependent on vessel speed, Hamilton *et al.* (1999) in their study to compare the performance of the RoxAnn and QTC-View found that E2 was inversely related to

vessel speed. They also found that E1 sometimes experienced change in synchronisation with E2. Similarly, Schlagentweit (1993) observed a consistent seabed classification by the RoxAnn system only at constant speed. He suggested that this might be related to changes in aeration and engine noise. Hamilton *et al.* (1999) also found that the occasional engine noise as the vessel was held on station or manoeuvred gave unexpected result i.e. the RoxAnn system did not function well when the vessel was essentially stationary where it was expected to be more reliable. In contrast, Hamilton *et al.* (1999) observed that the QTC data were not obviously dependent on vessel speed. They found that classification was consistent and did not change regardless of speed or even when the vessel was stationary or manoeuvred.

Kloser *et al.* (2001) found a depth bias in their RoxAnn data that could not be explained by differences in bottom type as determined from sediment and photographic samples. They found that both E1 and E2 parameters increased with increasing depth before clipping respectively at 130 and 70 m. When the depth trend prior to data clipping was removed from the E1 and E2 results, the resulting data compared favourably with the data derived from a CSIRO developed processing algorithm. Similarly, Voulgaris and Collins (1990) experienced a depth dependence in their RoxAnn data within a depth interval between 2.7 and 31 m. Hamilton *et al.* (1999), however, did not observe depth dependence in their RoxAnn and QTC data. Nonetheless, they have warned that RoxAnn and QTC data might vary with depth and water column properties because water column absorption and scattering are not allowed for by either system. In addition, they noticed that the RoxAnn system when sampling seabed with great slopes or depth changes could have problems detecting the second echo and part of it could be included in the first echo.

In their program to classify seabed types in the Australian southeast fishery by using normal incident acoustic techniques, Kloser *et al.* (2001) observed the dependency of seabed classifications on acoustic frequency. For the same seabed feature, different roughness indices were observed for two different frequencies they used. This highlights a potential misclassification by using only one frequency.

Hamilton *et al.* (1999) and Kloser *et al.* (2001) noticed a bias due to slope or a sudden rise or drop of the seabed in their RoxAnn data. High slopes or sudden rises or drops of the seabed normally produce long tails in the first bottom echo which thus provide large apparent acoustic roughness index estimates. For a sudden rise or drop of the seabed, this bias can be easily noticed in the echograms. Similarly, this bias can be picked up easily in the echograms if the vessel steams normal to the high slopes. If on the other hand the vessel is transecting parallel to the slope, this bias can only be interpreted once seabed types were plotted on the correspondence bathymetric map. Hamilton *et al.* (1999) observed a similar problem in their QTC data. This bias however can be used as a unique indication to identify such seabed types or areas (Greenstreet *et al.*, 1997; Hamilton *et al.*, 1999; Kloser *et al.*, 2001). In addition, Kloser *et al.* (2001) found that a narrow beamwidth was more sensitive to the slopes than a wider one.

The RoxAnn manufacturer recommends the use of RoxAnn squares, introduced by Burns *et al.* (1989), to assess seabed classifications and encourages users to adopt it. A number of problems in using the RoxAnn squares, however, have been noted by Voulgaris and Collins (1990), Greenstreet *et al.* (1997) and Hamilton *et al.* (1999). Greenstreet *et al.* (1997) stated that the rectangular nature of the RoxAnn squares was unlikely to be the best representation of each seabed or substrate type.

They observed inconsistency in the allocation system of the RoxAnn squares in boundaries between different seabed types. They then proposed the use of image processing methods as employed in the *IDRISI* software package to implement seabed classification. Voulgaris and Collins (1990), Greenstreet *et al.* (1997) and Hamilton *et al.* (1999) also found that E1 and E2 parameters were not independent but linearly related. Hamilton *et al.* (1999) found that their E1 and E2 parameters were linearly related such that data form an elongated roughly elliptical envelope inclined to the E1 and E2 axes. Using principal component analysis, Voulgaris and Collins (1990) found high correlation levels for their E1 and E2 data. Since E1 and E2 are not orthogonal in RoxAnn space, Hamilton *et al.* (1999) argued that the RoxAnn squares cut across the data trend. Their finding was confirmed by their QTC-to-RoxAnn converted data. They proposed the concept of classification polygons to replace the simplistic notion of RoxAnn squares. They used parallelograms with two sides parallel to the data trend with an argument that classes on the same trend line may have similar general bottom properties. Because no allowance for absorption and scattering is accounted for, the trend is an artefact and is not numerically meaningful (Hamilton *et al.*, 1999).

As mentioned earlier, the QTC-View system extracts 166 features from the received envelope of the first bottom echo. Principal Component Analysis is applied to the *high*-dimensional matrix formed by the extracted features to produce three parameters, presumably coming from the first three principal components having most of the variation of the original features. The physical and mathematical meaning of the extracted features as well as the three reduced features (Q-values) though are unknown to the user. Principal Component Analysis is applicable if variables (features) in a data matrix are intercorrelated. If all the variables are independent or in

other words orthogonal from one to another, all or most of the principal components are required to account for most of the variation in the original data. A very fundamental problem in the QTC-View system is that information as to whether or not intercorrelation exists in the extracted features is unknown. Furthermore, how well the three reduced features (Q-values) account for the variation of the data is not revealed to the user.

### **1.5. Classification methodology**

The approach used in this study is similar to that used in the commercial RoxAnn system. Like the RoxAnn system, the technique used in this study produced the so-called E1 and E2 parameters. E1, representing seabed roughness, was derived by integrating the tail sector of the first bottom echo and E2, describing seabed hardness, was obtained by integrating the complete second bottom echo. Since three frequencies were used, there were 3 sets of E1 and 3 sets of E2, each set corresponds to each frequency. In grouping bottom types however, multivariate analysis (Principal Component Analysis and Cluster Analysis) was adopted instead of the allocation system normally used in the RoxAnn system, called RoxAnn squares. The main reason for not using the RoxAnn squares has been discussed in the previous section.

The main objectives in using Principal Component Analysis (PCA) in this study were twofold. Firstly, it was used to reduce the dimensionality of the features sets but still retain a classification ability similar to that of the original parameter set. Secondly, it was used to transform separately a set of E1 parameters and a set of E2 parameters such that the transformed components (principal components) were

orthogonal such that any trend that might exist towards one particular frequency or more is being removed or at least minimised.

Clustering or grouping is an exploratory procedure that searches for natural structure within a data set. This procedure involves sorting the data cases or objects into groups or clusters so that objects in the same cluster are more like one another than they are like objects in other clusters. Sorting occurs on the basis of similarities calculated from the data; no assumptions about the structure of the data are made. Clustering is useful for developing classification schemes and suggesting or supporting hypotheses about the structure of the data.

Two techniques adopted here were the hierarchical agglomerative and iterative relocation (*k*-means) techniques. The former was adopted to establish the relationship between acoustically derived seabed types and fish groups using the Ward's minimum variance linkage method and squared Euclidean distances. The latter was on the other hand used for seabed classification.

Appendix A reviews Principal Component Analysis and Cluster Analysis, in particular hierarchical agglomerative and iterative relocation (*k*-means) techniques, as used in this work. Useful textbook treatments of aspects of this section are to be found in Chapter 8 of Johnson and Wichern (1998), Chapter 9 of Wickens (1995), section 6.3.1 of Bailey and Gatrell (1995), Chapter 12 of Rencher (1995), Chapter 6 of Harris (1975) and Chapter 6 of Timm (1975) for principal component analysis and Chapter 12 of Johnson and Wichern (1998), Chapters 2 to 4 of Gordon (1981), Diday and Simon as Chapter 3 of Fu (1976), Chapters 2 to 3 of Everitt (1974) and Chapters 5 to 7 of Anderberg (1973) for cluster analysis.



A summary of techniques adopted here for classification purposes is as follows. PCA was used to reduce the dimensionality of roughness and hardness indices prior to the introduction of the  $k$ -means technique used for seabed classification in the training set and in the entire data (section 5.2). PCA was also applied to the  $s_A$  estimates to reduce the dimensionality of data coming from the three operated frequencies prior to the introduction of the hierarchical agglomerative clustering technique to form the quasi acoustic population (section 5.7). In addition, PCA was applied to standardised roughness and hardness indices, and  $s_A$  estimates to avoid the scale-dependent problem. The  $k$ -means technique was used for seabed classification in the training sets and the entire data (sections 5.3 and 5.4). The  $k$ -means technique was applied to the first principal components of roughness and hardness indices. For the training sets, the feedback system (see section A.2.3) was used to minimise the effect of initial centroids on the classification results. For the remaining data however, the  $k$ -means technique adopted the centroids produced from the training sets and data points were assigned to the seabed class corresponding to the adopted centroids without iterative process. The hierarchical agglomerative clustering technique was adopted here to establish the relationship between acoustically derived seabed types and fish groups (section 5.7). The similarity measure and the linkage method adopted here in the hierarchical agglomerative clustering technique were the square Euclidean distance and Ward's minimum variance linkage method.

## **1.6. Spatial methodology**

Attention has also been given to the extent to which seabed roughness (E1) and hardness (E2) derived along the vessel track can be extended into unsampled areas. To explore the possibility of extending the vessel track into unsampled areas,

spatial autocorrelation, variogram and Kriging techniques were used in this study (section 5.6) and are reviewed in Appendix B.

Spatial autocorrelation was used for two purposes. Firstly, it was used in a descriptive manner to describe the spatial dependency of the data (direction and strength of autocorrelation). Secondly, it was used to infer a spatial correlation model which underpins the spatial interpolation technique of Kriging.

To achieve the objectives and to easily monitor any directional information that may exist, the vessel track in both areas was separately divided into straight transects as many as possible in any possible directions. Roughness and hardness indices of the three operated frequencies ( $E1_{12\text{kHz}}$ ,  $E2_{12\text{kHz}}$ ,  $E1_{38\text{kHz}}$ ,  $E2_{38\text{kHz}}$ ,  $E1_{120\text{kHz}}$ ,  $E2_{120\text{kHz}}$ ) at these straight transects were extracted. Since the series of roughness and hardness indices at each straight transect was not equally spaced, the roughness and hardness indices at the three operated frequencies were resampled separately using the linear interpolation method prior to the introduction of PCA. PCA was afterwards applied to sets of interpolated E1 and E2 indices separately to produce the first principal component of roughness and hardness indices ( $PC1_{E1}$  and  $PC1_{E2}$ ) at each straight transect. An autocorrelation function (ACF) was then applied to them separately.

The algorithm adopted to estimate the spatial ACF of the series of  $PC1_{E1}$  and  $PC1_{E2}$  is as follows:

1. Suppose that the series of the bottom roughness or the bottom hardness index is  $x_n$  and the number of data points is  $N$ .

2. Calculate the mean value of the data and subtract the mean value from the data,

i.e. remove any DC offset from the data:  $\bar{x} = \frac{1}{N} \sum_{n=0}^{N-1} x_n$  and  $x'_n = x_n - \bar{x}$ .

3. Estimate the linear trend (slope) and detrend it from the mean-removed data  $x'_n$ , and assign the mean-removed, detrended data as  $x''_n$ .

4. Bandpass the mean-removed, detrended data  $x''_n$  to smooth and remove the high-order trend from the data.

5. Add zeros; as many as the number of data  $N$  at the end of the data. Call this modified series  $x'''_n$ . The number of data of the modified data is therefore  $2N$ . To avoid confusion, let  $2N = N'$  as the new number of data.

6. Compute the DFT of  $N'$  point series using the FFT algorithm (Brook and Wynne, 1988; Siwabessy, 1995):

$$X_n = \frac{1}{N'} \sum_{k=0}^{N'-1} x'''_k e^{-j2\pi kn / N'} \quad (1.1)$$

7. Retain only the first  $N'/2$  coefficients of the spectrum in the positive frequency range of  $X_n$  and compute the PSD at each frequency component. The two-sided PSD is given by  $S_n = N'\Delta X_n X_n^*$  (Brook and Wynne, 1988), where  $N'\Delta = T$  and  $X_n^*$  is the conjugate of  $X_n$ . However, since  $X_n$  has only been computed for the positive half of the frequency range, the PSD becomes the single sided spectral density given by (Brook and Wynne, 1988)

$$S_n = 2N'\Delta X_n X_n^* \quad (1.2)$$

8. Compute the inverse Cosine transform of the PSD,  $S_n$ , to produce the ACF

(Brook and Wynne, 1988):

$$\varphi_n = \sum_{k=0}^{N'/2-1} S_k \cos\left(\frac{2\pi nk}{N'/2}\right) \quad (1.3)$$

The autocorrelation characteristic lengths of PC1\_E1 and PC1\_E2 were estimated by taking a distance where the first sign transition occurred. Using the above algorithm, simple MATLAB codes were written and are listed in Appendix C.

The acoustic samples show two primary characteristics, which are typical of the regionalised variable: (1) an apparently stochastic process, which gives random variability in space and (2) a similar probability density function (PDF) at all places, which allows mapping. Many techniques in classical statistics used to deal with these two major characteristics attempt to remove the characteristic of the spatial distribution (the second characteristic). In contrast however, the Kriging technique has at least one advantage over these techniques; it explicitly involves the spatial autocorrelation between samples in the analysis. Like other methods on the other hand, the Kriging technique assumes stationarity of the distribution.

## Chapter 2

### Physical concepts of multiple bottom reflection

The sea surface and seabed are highly complex boundaries as they are usually not smooth for underwater acoustic wavelengths of interest. In addition, the seabed is essentially not a homogeneous medium. Scattered waves arise when a sound wave travelling in the water encounters a rough sea surface or seabed. These scattered waves consist of coherent and incoherent components. The latter is commonly referred to as reverberation. The coherent component represents a wave propagating in the direction of the specular reflection. The ratio between the amplitude of the coherent reflection and that of the incident wave describes the coherent reflection coefficient. For a perfectly flat, fully reflecting surface, the coherent reflection coefficient for plane waves has a magnitude of unity. The coherent reflection coefficient is however less than unity if the acoustic wave is incident on a rough surface even if the surface material provides for complete reflectivity. The coherent reflection coefficient decreases as the scale of the surface irregularity relative to the acoustic wavelength increases. The incoherent component on the other hand corresponds to the individual components scattered from irregular surfaces. A simple parameterisation of surface roughness and its effect on the coherence parameter is provided by the Rayleigh parameter, which describes roughness in terms only of surface displacements perpendicular to an average surface plane. Figure 2.1 shows the dependence of the coherent reflection  $\mathfrak{R}$  on the Rayleigh parameter  $(k\sigma \cos\phi)^2$ ,

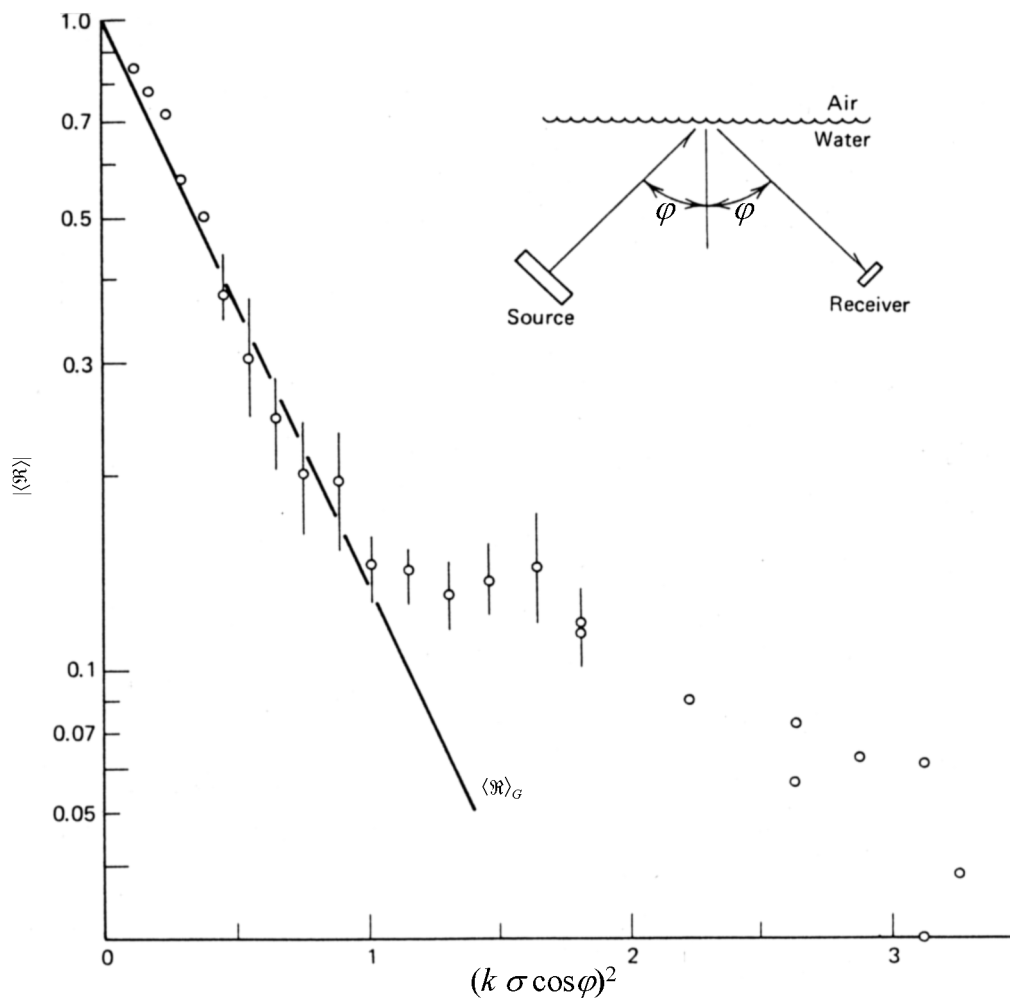


Figure 2.1. Plot of the dependence of coherence parameter on the Rayleigh parameter for various rms heights of the rough surface. After Clay and Medwin (1977).

where  $k$  is the acoustic wave number,  $\sigma$  is the root mean square (rms) height of the surface irregularities,  $\varphi$  is the angle of the incident acoustic wave towards the vertical plane and  $\langle \mathfrak{R} \rangle_G$  is the theoretical Gaussian surface value. Figure 2.1 suggests that the coherent reflection parameter  $\mathfrak{R}$  approaches zero for a large value of the Rayleigh parameter. Suppose, for illustration, that the rms height,  $\sigma$ , is 10 cm (0.1 m) and the grazing angle,  $\varphi = 0$ , i.e. for incidence perpendicular to the horizontal plane. For a frequency of 12 kHz (the longest wavelength operated in the work described here), having an acoustic wave number of 50.27, the Rayleigh parameter becomes  $(50.27 \times 0.1)^2 = 25.27$ . This suggests that the incoherent component will be very significant for the work reported here. Figure 2.1 shows that the coherent reflection parameter  $\mathfrak{R}$  agrees well with the theoretical Gaussian surface value  $\langle \mathfrak{R} \rangle_G$  only for a Rayleigh parameter less than 1. The simple model of surface roughness is not sufficient for the conditions involved in the present work.

While the sea surface, essentially a perfectly reflecting surface if near surface bubbles are ignored, only scatters the incident acoustic wave, the seabed not only scatters the incident acoustic wave but absorbs it as well. A part of the energy of the incident acoustic wave penetrates below the seabed and some of this component contributes to the return signal at the transducer. At low frequencies, the acoustic bottom reflection is determined primarily by the parameters of the bottom sediment. On the other hand, bottom relief rather than bottom sediment plays an important role in acoustic bottom reflection/scattering at frequencies above a few kilohertz. The signal reflected specularly from a hard, rough seabed such as a very rough rocky seabed might be less than that from a smooth, soft seabed such as muddy sediment. However, the reverse is expected if both bottom surfaces are smooth. Figure 2.2

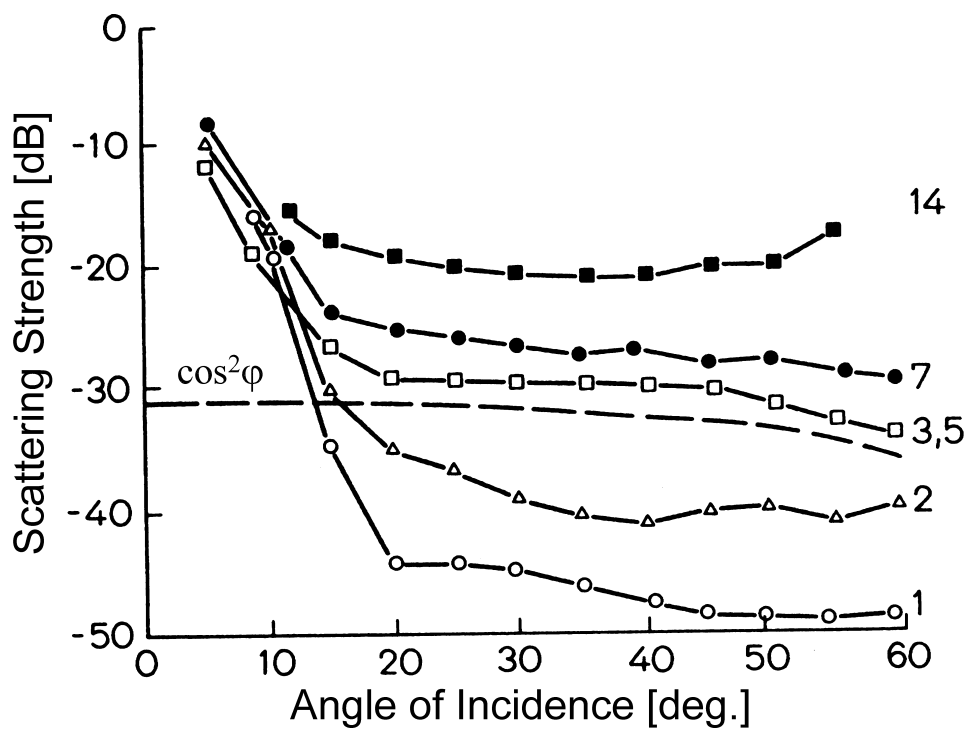
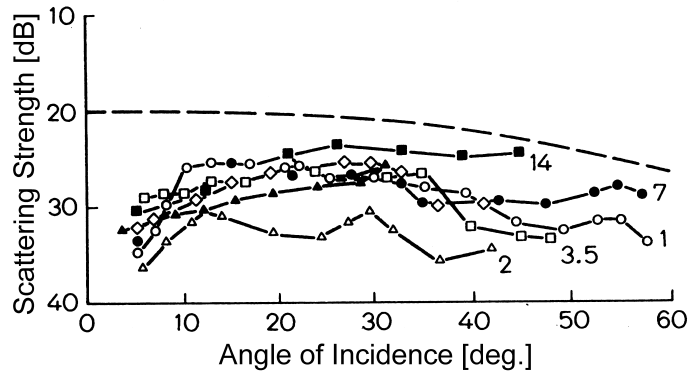


Figure 2.2. Dependence of the backscattering strength upon the incident angle from an abyssal plain bottom at various frequencies: 1, 2, 3.5, 7, 14 kHz. A dashed line ( $\cos^2 \varphi$ ) is the Lambert's law representation. After Brekhovskikh and Lysanov (1982).

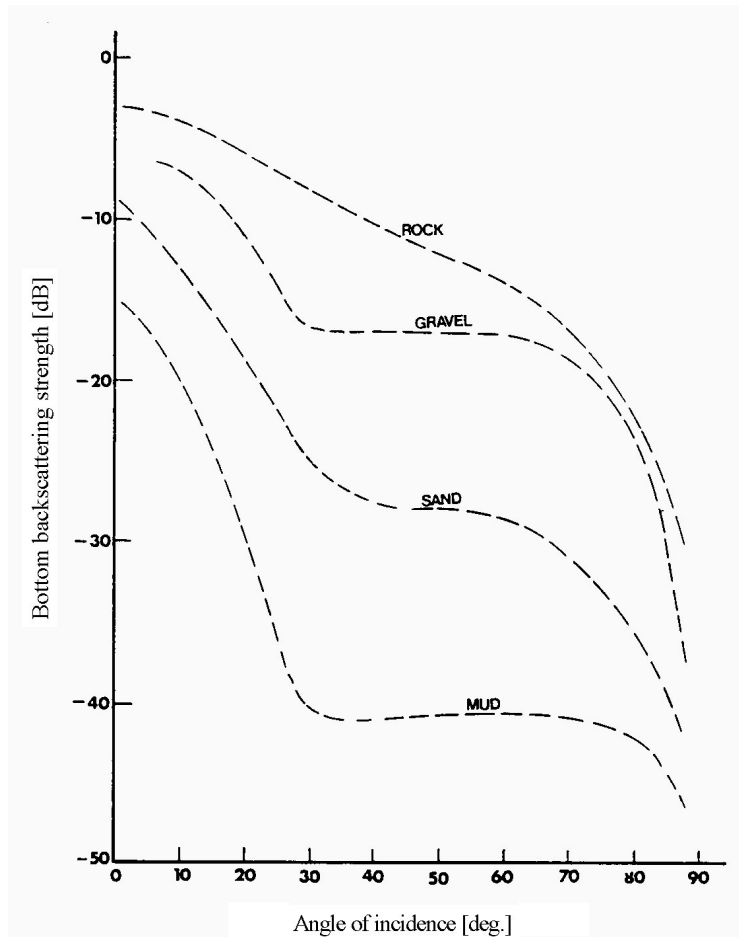


shows the dependence of the backscattering strength (logarithm of the backscattering coefficient  $m_s$ , a non-dimensional quantity, which is equal to the ratio of the acoustic power scattered by the unit surface per unit solid angle to the incident wave intensity) upon the incident angle at various frequencies from an abyssal plain bottom. Figure 2.2 also shows the form of an approximation commonly used to describe rough surface scattering i.e. Lambert's law given as a dashed line ( $\cos^2\phi$ ) in Figure 2.2. Figure 2.3a on the other hand reveals that the angular dependence on the backscattering strength from a very rough bottom is weak for all incident angles. As a frequency increases, a linear decrease of the scattering strength in Figure 2.2, in particular at angles of incidence less than  $20^\circ$ , fades away and almost disappears at 14 kHz. Figure 2.3a, on the other hand, does not possess the characteristic just discussed for Figure 2.2. Using data taken from Urick (1954), McKinney and Anderson (1964), and Wong and Chesterman (1968), Chivers *et al.* (1990) reproduced the plot of bottom backscattering strength at various angles of incidence for four different marine sediment types namely rock, gravel, sand and mud (Figure 2.3b). An approximate 10 dB difference between the four marine sediment types is observed at an angle of incidence of  $30^\circ$ . This difference is partly due to the different characteristic acoustic impedances of each sediment type.

The literature on acoustic scattering from rough surfaces is very substantial. In this chapter, some key features of such scattering analysis are presented to inform discussion of the data processing approach taken in the present work. The discussion begins with the fundamentals of acoustic plane wave reflection at an interface. It also describes the notion of the energy of the acoustic bottom returns for the evaluation of seabed types. It then concludes with a discussion of the adopted approach for the evaluation of seabed types. Useful treatments of aspects of this chapter are to be



(a)



(b)

Figure 2.3. Dependence of the backscattering strength upon the incident angle from (a) a very rough bottom at various frequencies: 1, 2, 3.5, 7, 14 kHz. (Brekhovskikh and Lysanov, 1982) and (b) various bottom types (Chivers, *et al.*, 1990).

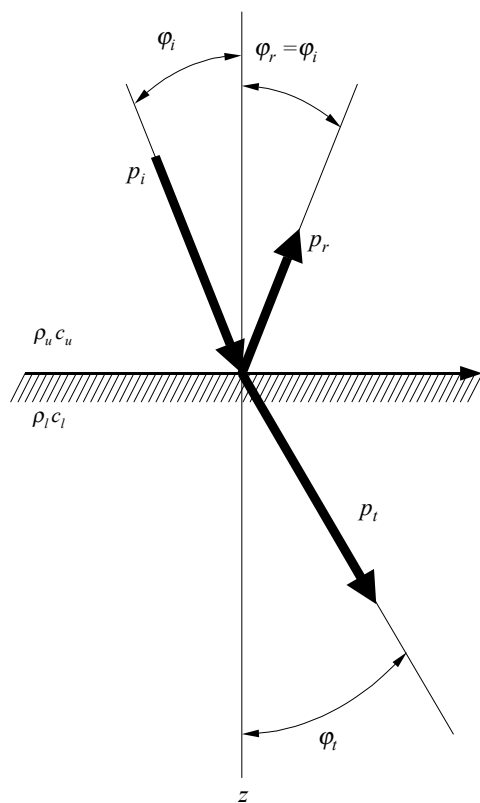
found in Tolstoy and Clay (1966), Tucker and Gazey (1966), Clay and Medwin (1977), Brekhovskikh and Lysanov (1982), Orłowski (1984), Heald and Pace (1996), and Medwin and Clay (1998).

## 2.1. Reflection of acoustic wave from the bottom surface

A first approximation to the interpretation of reflection/scattering from real water-seabed interfaces may be made by neglecting the effect of volume scattering within the seabed, and by assuming that plane waves are involved. Figure 2.4 shows a ray representation of such a process, showing incident, reflected and transmitted rays. Figure 2.4 assumes that the interface between the two media is horizontal. Acoustic waves incident on the seawater-seabed boundary involve reflection and scattering at the boundary and transmission in the second medium. This process is determined primarily by the acoustic impedance ( $Z = \rho c$ ) mismatch between the two involved media. At the boundary, it is assumed that there cannot be an excess of pressure on one side or the other, i.e.  $p_i + p_r = p_t$ , and the two media must maintain contact at the boundary, i.e.  $u_i \cos \varphi_i + u_r \cos \varphi_r = u_t \cos \varphi_t$  where  $u$  is the magnitude of the particle velocity and related to the acoustic pressure by  $p = \rho c u$ . In the simplest case of plane, normal incidence waves, the acoustic pressure reflection coefficient  $\mathfrak{R}$  is defined as

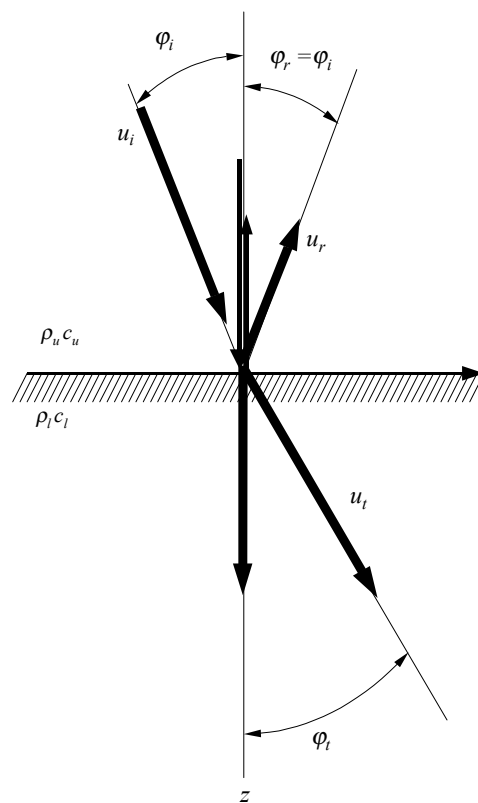
$$\mathfrak{R} = \frac{p_i}{p_r} = \frac{Z_l - Z_u}{Z_l + Z_u} = \frac{\rho_l c_l - \rho_u c_u}{\rho_l c_l + \rho_u c_u} \quad (2.1)$$

where  $p_i$  and  $p_r$  are respectively the incident and reflected wave pressures,  $Z$  is the acoustic impedance,  $\rho$  is the density of the media,  $c$  is the sound speed, and  $u$  and  $l$  denote upper and lower media respectively. Although this formulation is based on



(a) First boundary condition.

$$p_i + p_r = p_t$$



(b) Second boundary condition.

$$u_i \cos \phi_i + u_r \cos \phi_i = u_t \cos \phi_t$$

Figure 2.4. Geometry of the reflection at an interface between two media.  $\rho_u$  is the density of the upper medium,  $\rho_l$  is the density of the lower medium,  $c_u$  is the sound speed in the upper medium and  $c_l$  is the sound speed in the lower medium.  $\phi_i$  is the incident angle,  $\phi_r$ , which is equal to  $\phi_i$ , is the reflected angle and  $\phi_t$  is the refracted angle. The angles  $\phi_i$  and  $\phi_t$  are related by Snell's law.  $\mathbf{u}_i$ ,  $\mathbf{u}_r$  and  $\mathbf{u}_t$  are vectors of the incident, reflected and transmitted particle velocities, respectively. After Clay and Medwin (1977).

and valid for the liquid-gas interface, it is still applicable at normal incidence to the liquid-solid boundary and is the first, simplest approximation for the seawater-seabed interface (at normal incidence). At normal incidence, the shear (transverse) waves, which are absent in the liquid and gaseous media, are not excited. At angles other than normal incidence, both longitudinal and shear waves are excited in the solid medium and travel with different velocities. Consequently, the total (input) impedance of the solid comprises the longitudinal and the shear solid impedances, i.e.  $Z_{in} \equiv Z_l \cos^2 2\varphi_s + Z_s \sin^2 2\varphi_s$  where  $Z_{in}$  is the input impedance of the solid,  $Z_l$  and  $Z_s$  are the longitudinal and the shear impedances, respectively, and  $\varphi_s$  is the transmitted angle of the shear wave. At normal incidence,  $\varphi_i = \varphi_s = 0$  and the input impedance is simply the longitudinal impedance,  $Z_{in} = Z_l$ .

The extension of this analysis to real water-seabed systems and to realistic beam geometries involves a number of refinements. Kloser *et al.*, (2001) have listed a number of factors causing the reflected bottom signals to be different from the incident acoustic pulses. These are;

- (1) Acoustic impedance mismatch of the seawater-seabed interface leading to surface scattering of the main pulse.
- (2) Acoustic parameters of the instrument.
- (3) Acoustic signal penetration into the seabed leading to volume scattering of the main pulse.
- (4) Directional reflections at the seawater-seabed interface because of seabed roughness.

- (5) Time delay of oblique returns because of spherical spreading with changing depth.
- (6) Scattering response from the sea surface, subsurface bubbles and vessel hull for the second acoustic bottom return.
- (7) Seabed slopes.
- (8) Seawater acoustic absorption.
- (9) Acoustic noise.

Neglecting acoustic absorption, the following relationship holds for normal incident echosounders (Brekhovskikh and Lysanov, 1982; Orłowski, 1984)

$$4\pi \int_0^{\theta_0} m_s(\theta, \omega) \sin 2\theta d\theta \leq \frac{\langle p^2 \rangle}{p_0^2} \leq \mathfrak{R}^2 \quad (2.2)$$

where  $m_s(\theta, \omega)$  is the acoustic scattering coefficient,  $\omega$  is the angular frequency,  $\theta$  is the incident angle of the acoustic wave on the bottom,  $\theta_0$  is the half beamwidth,  $\langle p^2 \rangle$  is the average square of the received pressure,  $p_0$  is the received acoustic pressure of an ideal reflecting surface and  $\mathfrak{R}$  is the acoustic pressure reflection coefficient of a smooth boundary.

## 2.2. Acoustic scattering and reflection at rough surfaces

For acoustic frequencies of interest in the present work, seabed surfaces are in general rough. The definition of the term roughness though is somewhat arbitrary and ambiguous. The roughness of a scattering surface is not solely an inherent property of its own but is dependent also upon the viewing condition. A scattering surface will appear rougher as an insonifying wavelength decreases. For instance, a particular

surface that may well be smooth to ultrasound, may be rough to visible light. It is therefore critical to consider the scale of the roughness with reference to the insonifying wavelength of the incident radiation in considering how the incident energy will be scattered.  $k\sigma$  (as in Figure 2.1), where  $k$  is the acoustic wave number and  $\sigma$  is the rms deviation of the surface irregularities, is a common expression to scale the surface roughness where scale lengths in the plane of the surface are not considered. For  $k\sigma \ll 1$ , the normal incident backscattered return is coherent and its amplitude is determined directly by the reflection coefficient. For  $k\sigma \gg 1$ , the magnitude of the coherent returns is much reduced. In addition, the distribution of pressure amplitudes arising from an ensemble of return signals from rough interfaces varies from Gaussian ( $k\sigma \ll 1$ ) to Rayleigh ( $k\sigma \gg 1$ ) and to higher order distributions, for the case where the insonified area is translated across a rough interface.

Many formulations have been developed to describe scatter from rough surfaces. They are in general approximations and many only produce results close to experimental data in limited ranges of application. A useful starting point is provided by the Helmholtz-Kirchhoff formulation. Derivation of the Helmholtz integral equation starts first with separating variables of the wave equation into time and space dependent components. The time independent wave equation resulting from this process is known as the Helmholtz equation. This is a simplification of the more general Helmholtz integral equation which itself is a mathematical derivation of the Huygens' principle. The notion of the Helmholtz-Kirchhoff formulation is depicted in Figure 2.5. Here, a source  $Q$  insonifies a surface and each incremental area of the surface becomes a new source of a Huygens' wavelet. This leads the wavelets to spread spherically as far as a certain point at a particular time. The acoustic pressure

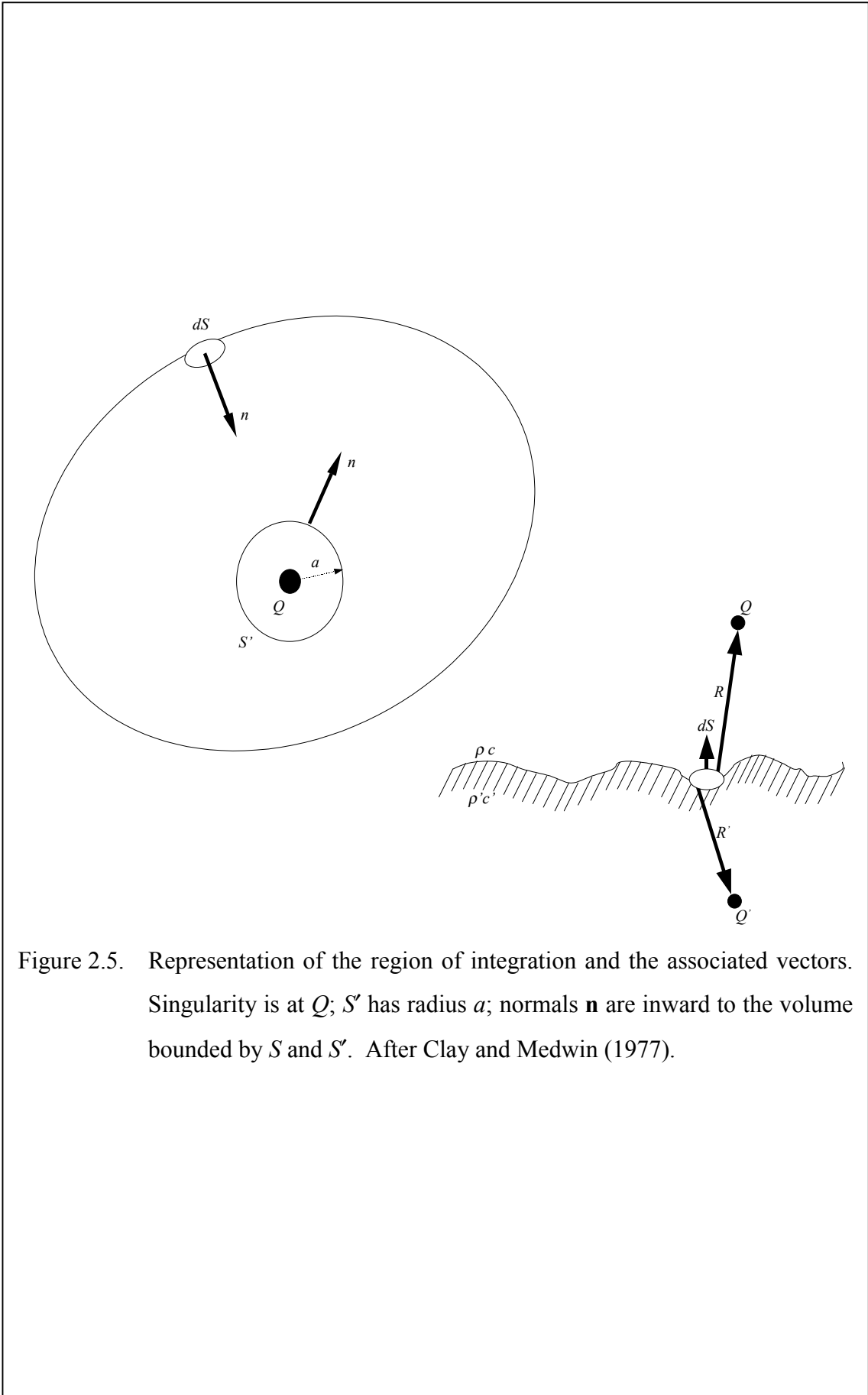


Figure 2.5. Representation of the region of integration and the associated vectors. Singularity is at  $Q$ ;  $S'$  has radius  $a$ ; normals  $\mathbf{n}$  are inward to the volume bounded by  $S$  and  $S'$ . After Clay and Medwin (1977).



at that point is then an integral over all the wavelets. The integral solution of the Helmholtz equation involves the Green's function that essentially uses Gauss' theorem. This solution is called the Helmholtz integral equation and is given by

$$U(Q) = \frac{1}{4\pi} \int_s \nabla(UG) \cdot d\underline{S} \quad (2.3)$$

where  $G$  is the Green's function and is the outbound spherical wave given and  $U$  is the inbound wave evaluated on the surface and  $U(Q)$  is the scattered field only at  $Q$ . The net field at  $Q$  would be  $U_{\text{net}}(Q) = U_{\text{inc}}(Q) + U(Q)$ . Two approximations are needed to simplify the Helmholtz integral equation. The first approximation is the Kirchhoff approximation and assumes that the pressure reflection coefficient  $\mathfrak{R}$  derived for an infinite plane boundary is applicable and useable at every point of a rough surface. This approximation, that  $\mathfrak{R}$  does not vary as a result of local, detected variations in incident angle arising from the varying inclinations of the facets of the rough surface, is also essentially adopted in the signal processing described later in this document. Thus the descriptors "hardness" and "roughness" used below do not acknowledge interaction between these parameters. Replacing the product  $\nabla U \cdot d\underline{S}$  to the normal derivative of  $U$  at  $dS$ ,  $-\frac{\partial U}{\partial n} dS$ , and substituting  $-\mathfrak{R} \frac{\partial U_s}{\partial n}$  into  $\frac{\partial U}{\partial n}$ , the resulting Helmholtz-Kirchhoff integral equation thus becomes

$$U(Q) = \frac{\mathfrak{R}}{4\pi} \int_s \frac{\partial}{\partial n} \left( U_s \frac{e^{-ikR}}{R} \right) dS \quad (2.4)$$

where  $U_s$  is the incident wave field,  $e^{-ikR}/R$  is the point source of the Green's function ( $G$  in equation 2.3) for the Huygens' wavelets,  $R$  is the distance between the incremental area of the surface  $dS$  and the point  $Q$  and  $\partial/\partial n(\ )$  is the derivative along

the normal to the surface evaluated at  $dS$ . It is important to note that equation (2.4) assumes that the source and receiver are in the same medium. Equation (2.4) shows that the field  $Q$  at the surface is, as expected, proportional to the source pressure.

The scattering function is commonly used to express the results of scattering experiments. The geometry of monostatic scattering is shown in Figure 2.6. In the spirit of the notion of equation (2.4) and assuming that the received acoustic pressure  $p$  is proportional to the incident acoustic pressure, the mean square of the received acoustic pressure for the monostatic arrangement (as in Figure 2.6), i.e. the source and receiver are coincident ( $R_{receiver}=R_{source}=R$ ), may be written as

$$\langle p^2 \rangle = \left( R_0^2 \int_{T_1} p_0^2 dt \right) \cdot \frac{A\mathfrak{S}}{R^4} \quad (2.5)$$

where  $p_0$  is the average acoustic pressure at  $R_0$ ,  $R$  is the average range to the source/receiver,  $\mathfrak{S}$  is the scattering function and depends on a combination of properties of the seabed and the acoustic system, and  $T_1$  is the duration of all arrivals from insonified area  $A$ . In the present work,  $\varphi \approx 0$ , depending on vessel attitude in roll and pitch, and local seabed slope over the insonified area.

### 2.3. Backscatter of the first and second acoustic bottom returns

In general, signals of the acoustic bottom return comprise coherent and incoherent components (Brekhovskikh and Lysanov, 1982; Pace and Ceen, 1982; Orłowski, 1984; Medwin and Clay, 1998) and the average of the returned acoustic pressure values squared may be written as

$$\langle p^2 \rangle = p_c^2 + \sum_s p_{ic(s)}^2 \quad (2.6)$$

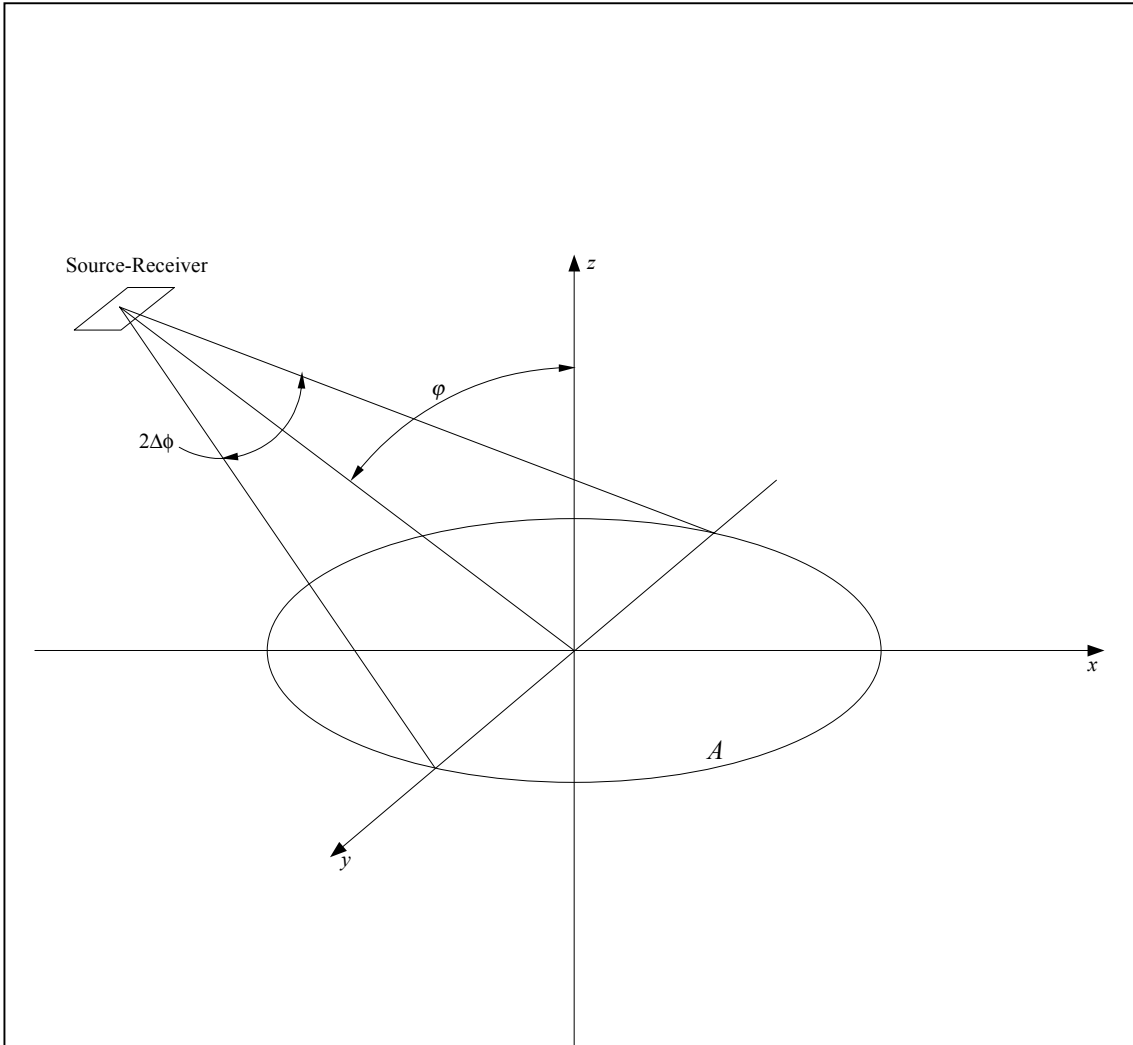


Figure 2.6. Geometry of the monostatic scattering arrangement.  $A$  is the insonified area;  $R$  is the distance to source and receiver  $\Delta\phi$  is the  $e^{-1}$  half beam widths of the directional source. In the present work,  $\phi \approx 0$ , depending on vessel attitude in roll and pitch, and local seabed slope over the insonified area. After Clay and Medwin (1977).

where  $p_c$  is the received acoustic pressure due to the coherent returns, i.e. the reflected components and  $p_{ic}$  is the received acoustic pressure due to incoherent returns, i.e. the "scattered" component of the return signals. The coherent component is proportional to the square of the averaged complex pressure  $\langle p \rangle^2$  over a number of realisations of the surface scattering area and the incoherent component is proportional to the difference between the average of the squared pressure values  $\langle p^2 \rangle$  and the averaged complex pressure  $\langle p \rangle^2$ . A tail present in the received signals significantly longer than the transmitted signal may be attributed to the incoherent component (Pace and Ceen, 1982). Applying the concept described in equation (2.2) into equation (2.6), Orłowski (1984) defines the total acoustic pressure reflection coefficient as

$$\langle \mathfrak{R}^2 \rangle = \mathfrak{R}_c^2 + 4\pi \int_0^{\theta_0} m_s(\theta) \sin \theta d\theta \quad (2.7)$$

$$\langle \mathfrak{R}^2 \rangle = \mathfrak{R}_c^2 + \mathfrak{R}_{ic}^2 \quad (2.8)$$

where  $\mathfrak{R}_c$  is the coherent reflection coefficient and  $\mathfrak{R}_{ic}$  is the incoherent reflection coefficient. For an ideally, perfectly flat, smooth surface, all the energy transmitted normally to the seabed would return to the transducer and energy at other angles would be reflected away. This means that the second term of equation (2.8) becomes zero and the coherent reflection (the first term) depends on the seabed acoustic absorption.

Two assumptions commonly applied to the range dependence of the intensity of the acoustic backscatter from a rough surface are that it varies as  $(R_0+R_1)^2$  or  $(R_0R_1)^2$ , where  $R_0$  and  $R_1$  are the distances of the source and receiver from the

scattering surface, respectively. When the source and receiver are coincident, the round trip loss in decibel form becomes  $20\log 2R_0$  for the former and  $40\log R_0$  for the latter. The total backscattered intensity becomes predominantly coherent when the rms surface deviations are small in comparison with the insonifying wavelength ( $\sigma \leq 0.02\lambda$ ), where  $\sigma$  and  $\lambda$  are the rms surface deviation and the wavelength, and the range dependence will approach  $20\log 2R_0$  (Pace and Al-Hamdani, 1985). When the incoherent component becomes dominant ( $\sigma \geq 0.25\lambda$ ), the scattered signal intensity reduces as the square of the distance from the rough surface and the range dependence would require  $40\log R_0$  spreading factor (Pace and Al-Hamdani, 1985). For the echo integration that was adopted in this study, the time-varied-gain (TVG) specified by the manufacturer for volume scattering compensation is  $20\log R_0 + 2\alpha R_0$  (Anon, 1993). This applies to incoherent scattering from a volume. On the other hand, this does not apply to all parts of acoustic reflected pulse from the seabed. Since the round-trip spreading loss, and the time-varied-gain (TVG) used to offset depend in detail on the scattering regime, a measure of uncertainty is necessarily associated with the influence of range on the values of recorded field data.

### 2.3.1. First acoustic bottom returns

Focusing on the second term of equations (2.7) and (2.8), Heald and Pace (1996) endeavored to relate energy features from the first acoustic bottom returns and roughness parameters. Figure 2.7 shows the geometry of the first and second backscatter return from the seabed. For the incremental area  $dA_1$  in the farfield of the scattering patch, the first backscatter return becomes incoherent (Pace and Al-Hamdani, 1985). Total backscatter return is subject to the sum of all backscatter

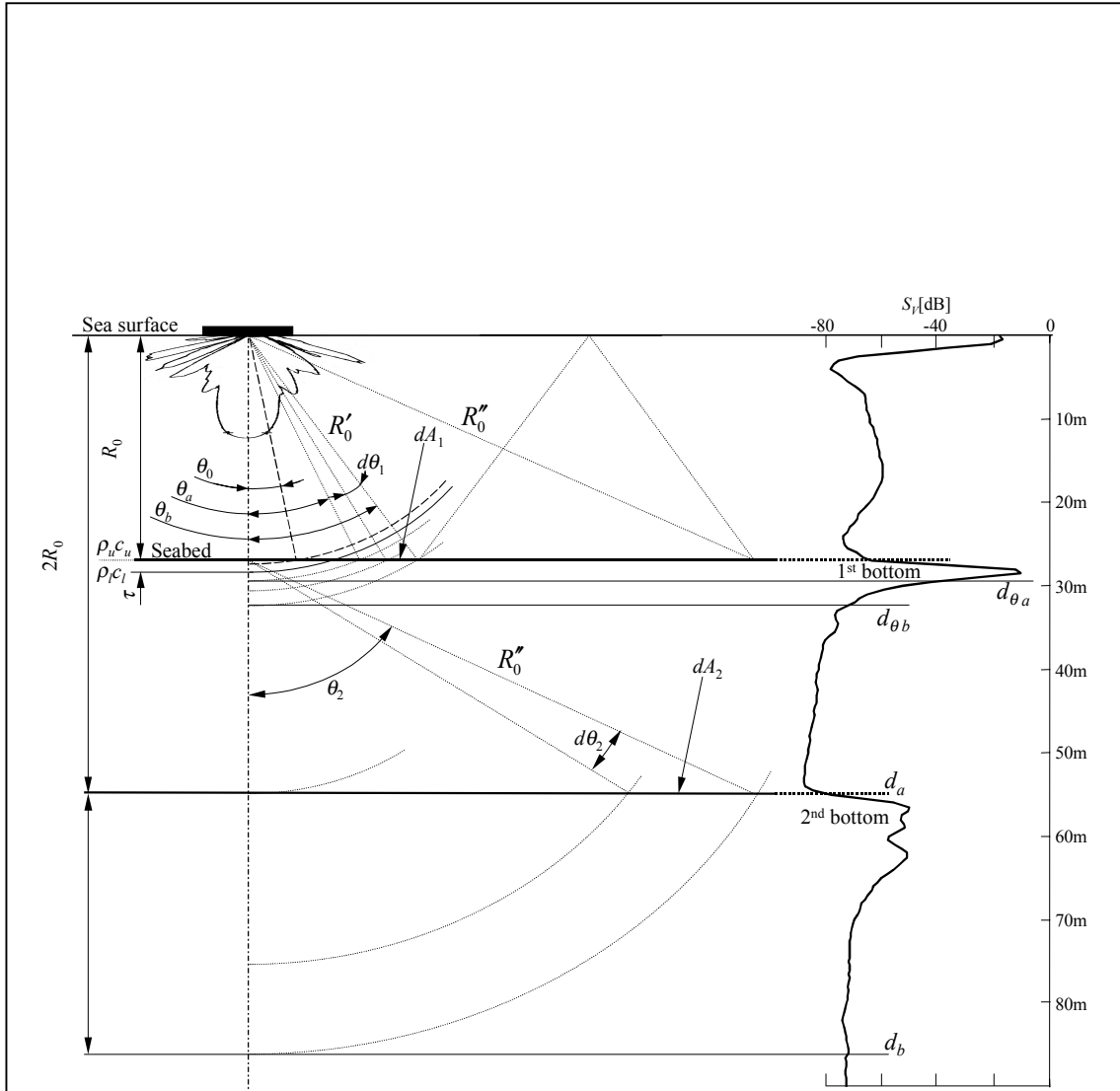


Figure 2.7. Geometry of acoustic bottom returns. Note here that  $\varphi$ , as defined in Figure 2.6, is equal to 0.  $\theta_i$ , ranging here from  $\theta_0 \rightarrow \theta_2$ , represents incident angles for off-axis beam components;  $d_{\theta_a}$  and  $d_{\theta_b}$  are the oblique depths corresponding to  $\theta_a$  and  $\theta_b$ , respectively,  $d_a$  and  $d_b$  are the constant depth interval for the integration of the second acoustic bottom return.

$$dA_1 = 2\pi R_0'^2 \tan \theta_1 d\theta_1. \quad R_0' = R_0 \sqrt{1 + \tan^2 \theta_1}.$$

$$dA_2 = 2\pi R_0''^2 \tan \theta_2 d\theta_2. \quad R_0'' = R_0 \sqrt{1 + \tan^2 \theta_2}. \quad \theta_2 = \tan^{-1}(3 \tan \theta_1).$$

return from all areas. Following Heald and Pace (1996), the received acoustic pressure may be expressed as

$$\begin{aligned}
 p_{bs1}^2 &= p_0^2 \int_{\theta_a}^{\theta_b} \frac{m_s(\theta_1) G^2(\theta_1)}{(R'_0)^4} dA_1 \\
 p_{bs1}^2 &= 2\pi p_0^2 \int_{\theta_a}^{\theta_b} \frac{m_s(\theta_1) G^2(\theta_1) \sin\theta_1 \cos\theta_1}{R_0^2} d\theta_1
 \end{aligned} \tag{2.9}$$

where  $p_0$  is the source pressure at a distance of 1 m from the source  $dA_1 = 2\pi R_0'^2 \tan\theta_1 d\theta_1$ ,  $R'_0 = R_0 \sqrt{1 + \tan^2 \theta_1}$ ,  $G(\theta_1)$  is the transducer gain and  $m_s(\theta_1)$  is the acoustic scattering coefficient;  $m_s(\theta_1) \propto \mathfrak{R}^2$  and  $m_s(\theta_1) \propto (\sigma/T)^2$  where  $\sigma$  is the rms height of the surface roughness and  $T$  is the correlation length of the surface roughness. Heald and Pace (1996) further suggest that the integration limit of the intensity envelope of the first backscatter return from the seabed is in the region where the insonified area is an annulus when  $ct/2 > c\tau/2$ , i.e.  $\sqrt{c(t-\tau)/R_0} \leq \theta_1 \leq \sqrt{ct/R_0}$ . In practice, however, this integration limit needs to take into account the pulse shape and the Q of electronics and transducer.

### 2.3.2. Second acoustic bottom returns

Orlowski (1984) used a monostatic geometry for treatments of the second backscatter return from the seabed whereas Heald and Pace (1996) used an on axis bistatic geometry. In the present work, monostatic geometry was used throughout. The geometry of the second backscatter return from the seabed is shown in Figure 2.7. Figure 2.7 shows that the second acoustic bottom arises from those rays that have been specularly reflected twice at the seabed and obliquely back reflected once from the sea surface. The received acoustic pressure from the second backscatter return

from the seabed must include coherent and incoherent components, i.e. the integration limit includes the complete returned envelope, and may be expressed as

$$\begin{aligned}
 p_{bs2}^2 &= p_0^2 \int_{\theta_a}^{\theta_b} \frac{m_s^2(\theta_1) m_a(\theta_1) G^2(\theta_1)}{(R_0'')^4} dA_2 \\
 p_{bs2}^2 &= 18\pi p_0^2 \int_{\theta_a}^{\theta_b} \frac{m_s^2(\theta_1) m_a(\theta_1) G^2(\theta_1) \frac{\tan \theta_1}{\cos^2 \theta_1}}{R_0^2 (1 + 9 \tan^2 \theta_1)^2} d\theta_1
 \end{aligned} \tag{2.10}$$

where  $m_a(\theta_1)$  is the sea surface scattering coefficient,  $dA_2 = 2\pi R_0''^2 \tan \theta_2 d\theta_2$ ,  $R_0'' = R_0 \sqrt{(1 + \tan^2 \theta_2)}$  and  $\theta_2 = \tan^{-1}(3 \tan \theta_1)$ . For the second backscatter return from the seabed, the complete returned envelope is required and is obtained when  $ct/2 \leq c\tau/2$  and  $ct/2 > c\tau/2$ , where  $t=0$  corresponds to its onset at the receiver (Heald and Pace, 1996), i.e.  $0 \leq \theta_2 \leq 2\sqrt{ct/R_0}$ . The magnitude of the double reflection signal appearing in the second bottom backscatter depends on the square of the acoustic scattering coefficient  $m_s(\theta_1)$  of the seabed and on the magnitude of the acoustic pressure returning to the transducer. The introduction of a scattering coefficient of the sea surface  $m_a(\theta_1)$ , which is absent in the received acoustic pressure of the first bottom backscatter accounts for the rough surface scattering from the air-water interface. Like the normal reflection from the seabed for the first bottom return, the specular reflection at the seabed for the second return may give a direct measure of the characteristic acoustic impedance of the seabed relative to the surrounding seawater. In this following discussion, the term ‘‘hardness’’ of the seabed is used as a surrogate for the seabed acoustic impedance.



## 2.4. Selection of integration interval

Acoustic signals returning to the transducer after being reflected once from the seabed (first bottom return) and twice from the seabed (second bottom return) are simply a time series. The amplitude of these signals at a particular instance arises from a number of mechanisms such as bottom reverberation and oblique surface scattering. It represents the signal interference corresponding to these mechanisms after convolution with the acoustic pulse transmitted in the appropriate direction.

The angular characteristics of the transducer beam play a primary role in determining the area of the seabed from which signals are received. This area does not have a well-defined boundary. On the one hand, the tendency to define the beamwidth of a transducer by an arbitrary fixed figure, such as the -3dB or -6dB power points, is a common practice. On the other hand, the seabed area that contributes to received signals at the transducer arises in fact from the range of angles from which the acoustic signals coming out from the transducer are above the signal-to-noise ratio (SNR) of the system. This then relies upon the excitation level and the beamwidth of the transducer, and the characteristics of the seabed as has been pointed out by Kloser *et al.*, (2001).

### 2.4.1. Constant angular integration interval for the tail of the first bottom returns

As mentioned earlier, the tail present in the received signals may be attributed to the incoherent component. In the discussion on the incoherent component earlier, the oblique back reflection is assumed to be the sole contributor to the return signal. In fact, not only does the oblique back reflection contribute to the incoherent component in the first bottom echo, but so does the competing mechanism of sub-

bottom reverberation (Clay, 1966; Chivers, *et al.*, 1990). The angular beam pattern and the natural characteristic of the seabed, including for instance bottom acoustic absorption (see Table 3.1), determine the relative magnitude of the oblique backscattering strength and the sub-bottom reverberation in the first bottom echo. It then becomes necessary to isolate as much as possible the incoherent component due only to the oblique back reflection as much as possible in order to obtain a measure of roughness. This might be achieved by using a transducer with an adequate beamwidth to allow for minimising the potential contribution of the sub-bottom reverberation (Chivers, *et al.*, 1990). The time, after the start of the first normal reflection, by which the sub-bottom reverberation has decayed to a negligible level becomes important.

The constant angular integration interval for the first bottom echo consists of two different incident angles off the beam axis (see equation (2.9) and Figure 2.7). The first one ( $\theta_a$  in Figure 2.7) is the start off-axis angle at which the sub-bottom reverberation is estimated to have decayed to a negligible level and the second angle ( $\theta_b$  in Figure 2.7) to ensure that the proportion of the tail sector being integrated is similar regardless of depth. The assumption used for the selection of the first and second angles is that the roughness index is depth-independent. Suitable choice of  $\theta_a$  has the effect of removing the initial normal back reflection which is often 10 or 20 dB above the level of the “tail” signal. The removal also permits the available dynamic range of the system to be more effectively used for ground discrimination (Chivers, *et al.*, 1990). In this current study, three different frequencies, 12, 38 and 120 kHz, were operated. The nominal beamwidths were  $16^\circ/17.5^\circ$ ,  $7.1^\circ$  and  $11.2^\circ$  for

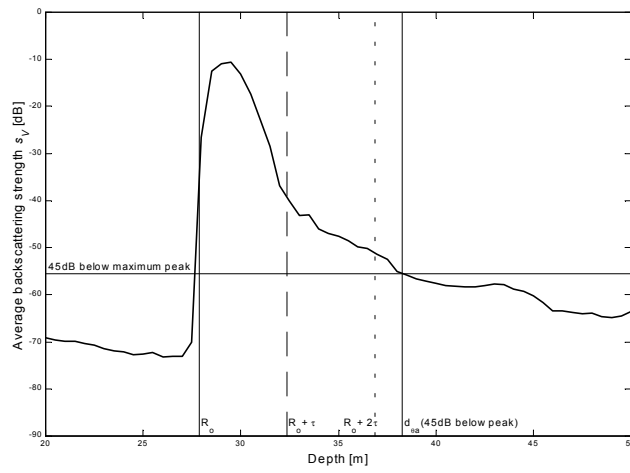
Table 2.1. Bottom acoustic absorption coefficient for various bottom types.  $v_p$  = compressional velocity in  $\text{m s}^{-1}$  and  $\alpha_p$  = compressional bottom acoustic absorption coefficient.

Bottom type	Frequency [kHz]	$v_p$ [ $\text{m s}^{-1}$ ]	$\alpha_p$ [ $\text{dB m}^{-1}$ ]	Reference
Clayey Silt	14	1457	1	Hamilton <i>et al.</i> (1970)
	25	1457	4	Hamilton (1972)
	100	1464	18	Hamilton (1972)
Clay	12	1500	1.6	Jensen <i>et al.</i> (1994)
	38		5.1	Jensen <i>et al.</i> (1994)
	120		16	Jensen <i>et al.</i> (1994)
Silt	12	N/A	6.7	Jackson and Briggs (1992)
	12	1575	7.6	Jensen <i>et al.</i> (1994)
	25	1572	8	Hamilton (1972)
	38	1575	24.1	Jensen <i>et al.</i> (1994)
	100	1537	60.9	Hamilton (1972)
Sand	120	1575	76.2	Jensen <i>et al.</i> (1994)
	12	N/A	4.4	Jackson and Briggs (1992)
	12	1650	5.8	Jensen <i>et al.</i> (1994)
	14	N/A	5.2	Jackson and Briggs (1992)
		1650	6.8	Jensen <i>et al.</i> (1994)
		1798	7	Hamilton <i>et al.</i> (1970)
	25	1645	9.4	Hamilton (1972)
	38	1650	18.4	Jensen <i>et al.</i> (1994)
	100	1630	47.3	Hamilton (1972)
	120	1650	58.2	Jensen <i>et al.</i> (1994)
Gravel	12	1800	4	Jensen <i>et al.</i> (1994)
	38		12.7	Jensen <i>et al.</i> (1994)
	120		40	Jensen <i>et al.</i> (1994)
Moraine	12	1950	2.5	Jensen <i>et al.</i> (1994)
	38		7.8	Jensen <i>et al.</i> (1994)
	120		24.6	Jensen <i>et al.</i> (1994)
Chalk	12	2400	1	Jensen <i>et al.</i> (1994)
	38		3.2	Jensen <i>et al.</i> (1994)
	120		10	Jensen <i>et al.</i> (1994)
Limestone	12	3000	0.4	Jensen <i>et al.</i> (1994)
	38		1.3	Jensen <i>et al.</i> (1994)
	120		4	Jensen <i>et al.</i> (1994)
Basalt	12	5250	0.2	Jensen <i>et al.</i> (1994)
	38		0.7	Jensen <i>et al.</i> (1994)
	120		2.3	Jensen <i>et al.</i> (1994)
Nonbioturbated	3.5	1547	0.35	Richardson and Young (1980)
Bioturbated	3.5	1514	0.20	Richardson and Young (1980)

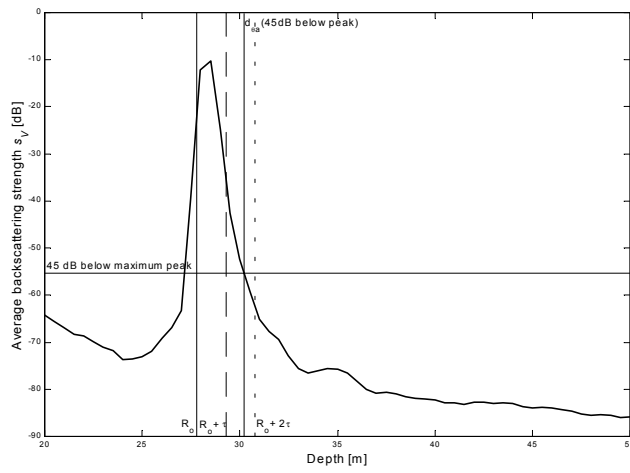
12, 38 and 120 kHz transducers. Figures 2.8 to 2.11 show the average backscattering coefficient  $s_V$  from the two adopted surveys (averaged over 1000 pulses) at four different depths for the three operated frequencies. Selection of the angular integration interval was based on the following assumption. It was assumed that a linear trend of E1 with depth was likely to be an artefact of beam geometry and choice of TVG. Thus, despite the fact that there is no law of nature which requires that  $\frac{\partial E1}{\partial R_0}$  must equal to 0, as a working hypothesis, it was assumed here that on average  $\frac{\partial E1}{\partial R_0} = 0$  where E1 is the roughness index and  $R_0$  is the depth. Lastly, the angular integration interval should include the similar tail sector regardless of the beamwidth of the operated transducer units. For the start angle of the integration interval  $\theta_a$  (equation (2.9) and Figure 2.7), it is assumed that the backscattering level after the peak of the first bottom echo at which the sub-bottom reverberation is deemed negligible is the same for the three operated frequencies and be selected arbitrarily such that the working hypothesis is met. Depths at which the sub-reverberation is negligible at the three frequencies are then plotted against bottom depths and a simple regression is then applied to the plot following this expression

$$d_{\theta_a} = mR_o + c_\tau \quad (2.11)$$

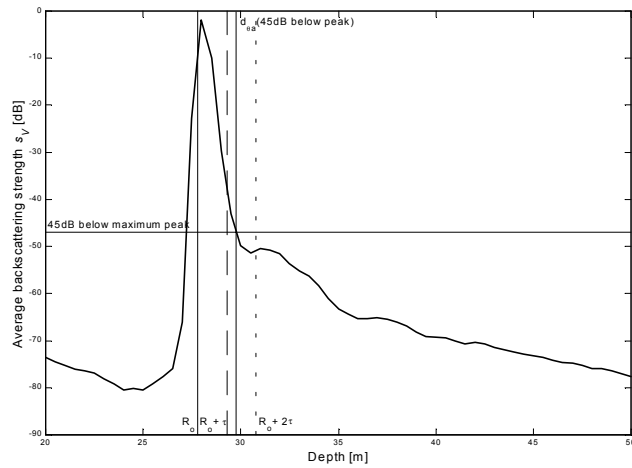
where  $m = \sec(\theta_a)$ ,  $R_o$  is the bottom depth and  $c_\tau$  is the offset. As a matter of convenience, the regression line is then forced to intercept at an offset  $c$  equal to the pulse length or a multiple of the pulse length and the new angle depicted in the slope  $m$  is reproduced accordingly.



(a) 12 kHz

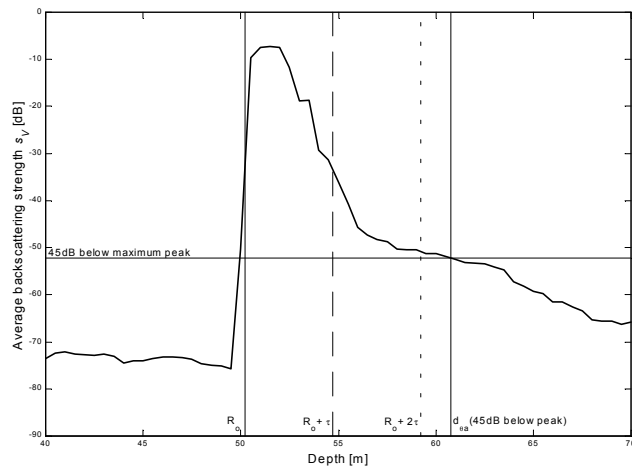


(b) 38 kHz

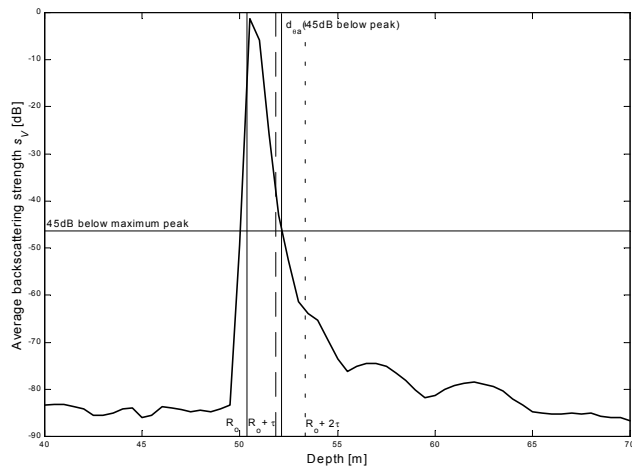


(c) 120 kHz

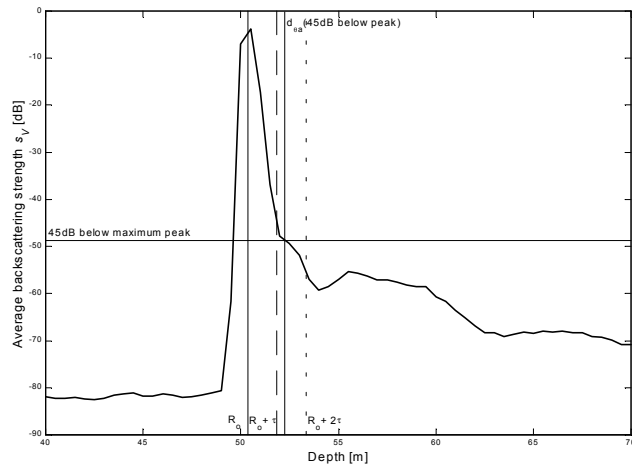
Figure 2.8. The average backscattering strength versus depth at a nominal bottom depth of 28m.  $R_o$ =bottom depth;  $d_{\theta_a}$ =oblique depth at start angle  $\theta_a$ ;  $\tau$ =pulse length (4.5 m for 12 kHz and 1.5 m for 38 and 120 kHz).



(a) 12 kHz

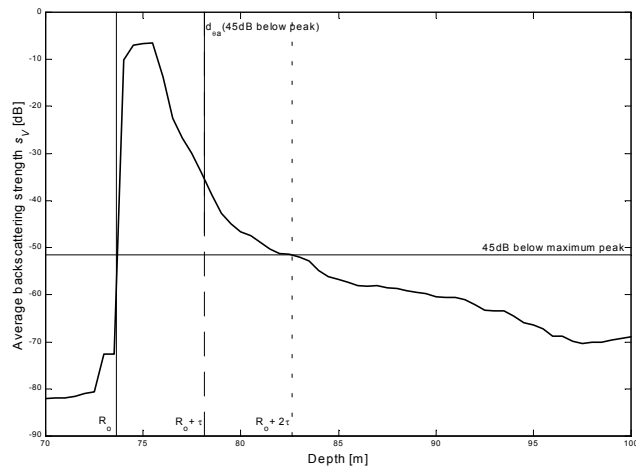


(b) 38 kHz

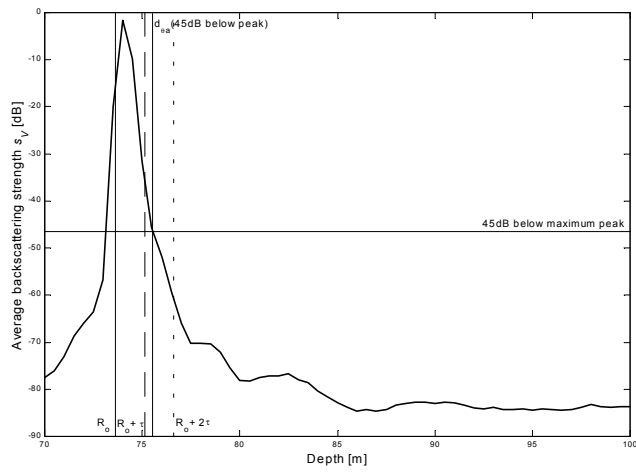


(c) 120 kHz

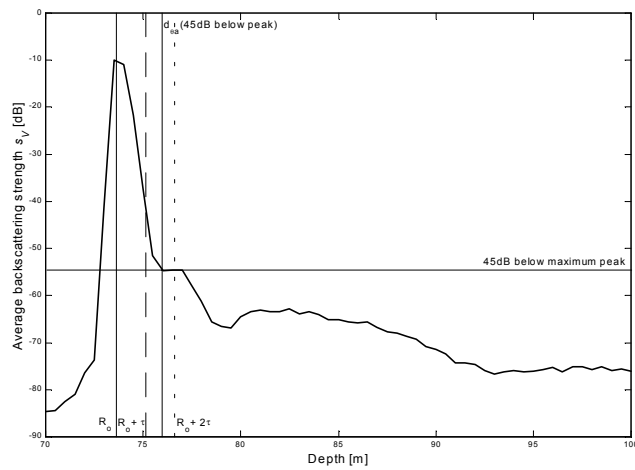
Figure 2.9. The average backscattering strength versus depth at a nominal bottom depth of 50m.  $R_o$ =bottom depth;  $d_{\theta_a}$ =oblique depth at start angle  $\theta_a$ ;  $\tau$ =pulse length (4.5 m for 12 kHz and 1.5 m for 38 and 120 kHz).



(a) 12 kHz

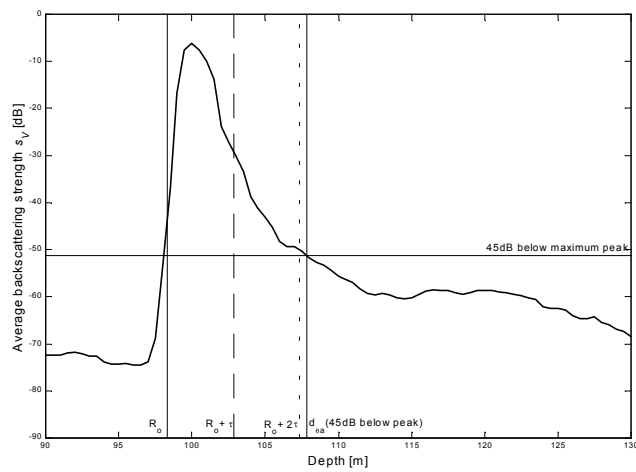


(b) 38 kHz

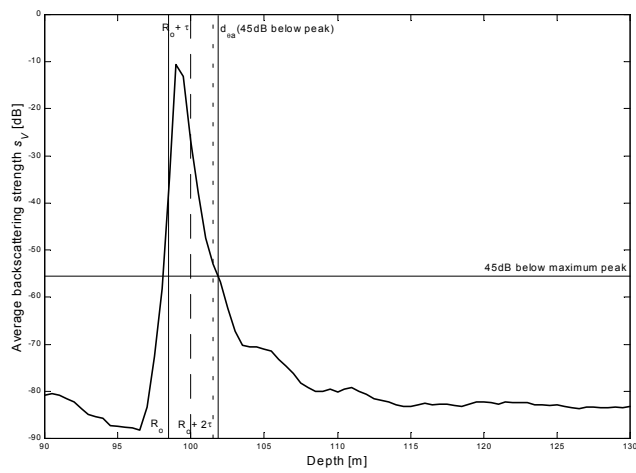


(c) 120 kHz

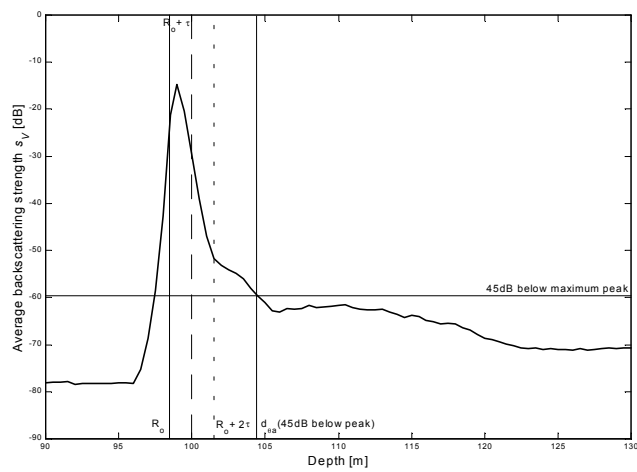
Figure 2.10. The average backscattering strength versus depth at a nominal bottom depth of 74m.  $R_o$ =bottom depth;  $d_{\theta_a}$ =oblique depth at start angle  $\theta_a$ ;  $\tau$ =pulse length (4.5 m for 12 kHz and 1.5 m for 38 and 120 kHz).



(a) 12 kHz



(b) 38 kHz



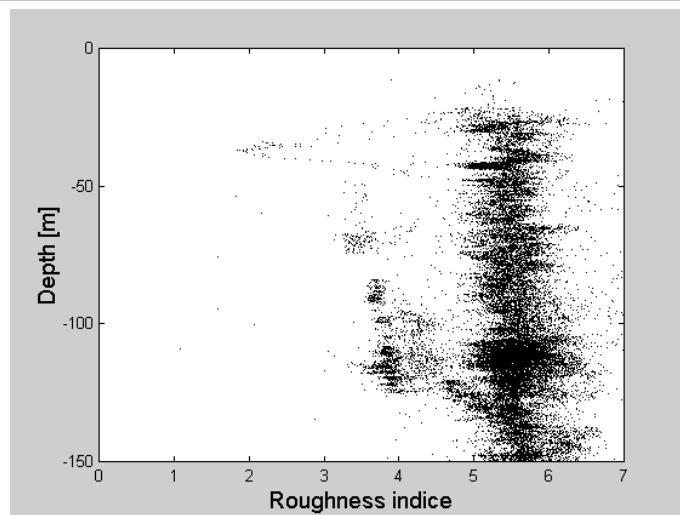
(c) 120 kHz

Figure 2.11. The average backscattering strength versus depth at a nominal bottom depth of 98m.  $R_o$ =bottom depth;  $d_{\theta_a}$ =oblique depth at start angle  $\theta_a$ ;  $\tau$ =pulse length (4.5 m for 12 kHz and 1.5 m for 38 and 120 kHz).

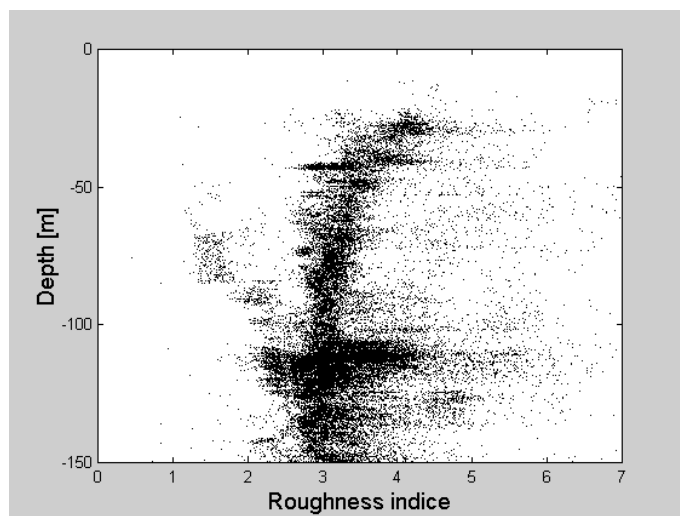


Since the second angle of the integration interval  $\theta_b$  (equation (2.9) and Figure 2.7) is selected to ensure that the tail sector being integrated is similar, selection of this angle is not as critical as selection of the start angle of the integration interval  $\theta_a$ . The second angle is selected such that the entire tail sector of the first acoustic bottom echo is included. To derive the second angle  $\theta_b$ , a similar procedure to that used to derive the first angle  $\theta_a$  was applied. Depths at which the average backscattering coefficient  $s_V$  starts to become constant or in other words the tail of the first acoustic bottom returns ends at the three frequencies are then plotted against corresponding bottom depths and a simple regression is applied following the expression given in equation 2.11. The regression line is forced to intercept at an offset  $c$  similar to that of the first angle.

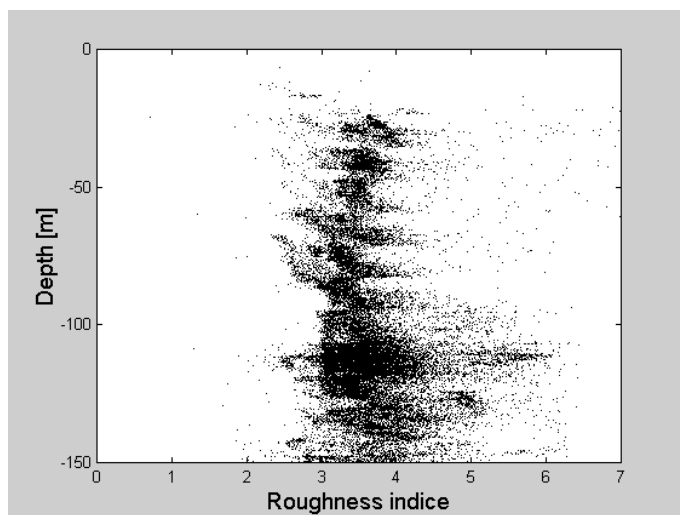
A number of backscattering levels at which the sub-bottom reverberation is assumed to be negligible within 35 dB and 45 dB below the maximum peak of the first bottom echo were evaluated. Representative examples of these plots for the NWS data are shown in Figure 2.12 for 35 dB below the maximum peak and Figure 2.13 for 45 dB below the peak for the three operated frequencies. Taking into consideration that the angular interval should include the similar tail sector regardless of the beamwidth of the operated transducer units, plots of tail integration with the start angles of the integration interval  $\theta_a$  commencing at the backscattering levels 35 dB and 45 dB below the maximum peak, respectively, were then made. It was found that start angle of the integration interval  $\theta_a$  commencing at the backscattering level 45 dB below the maximum peak allowed for almost similar tail sector being integrated regardless of the beamwidth and provided the integration results independent of depth in comparison to that at the backscattering level 35 dB below



(a) 12 kHz

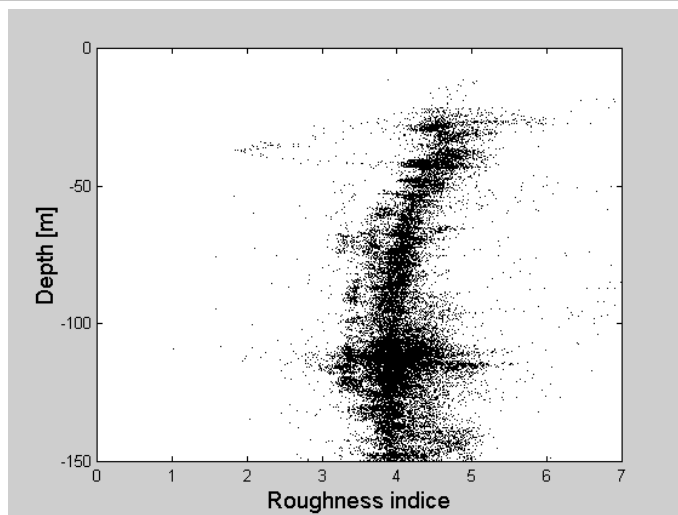


(b) 38 kHz

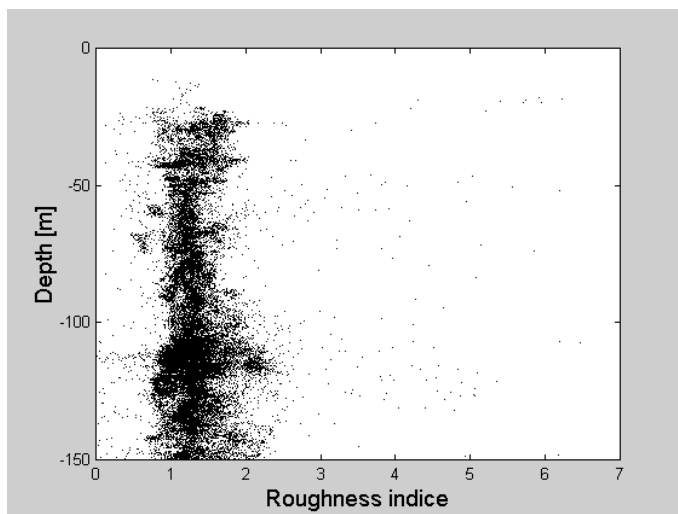


(c) 120 kHz

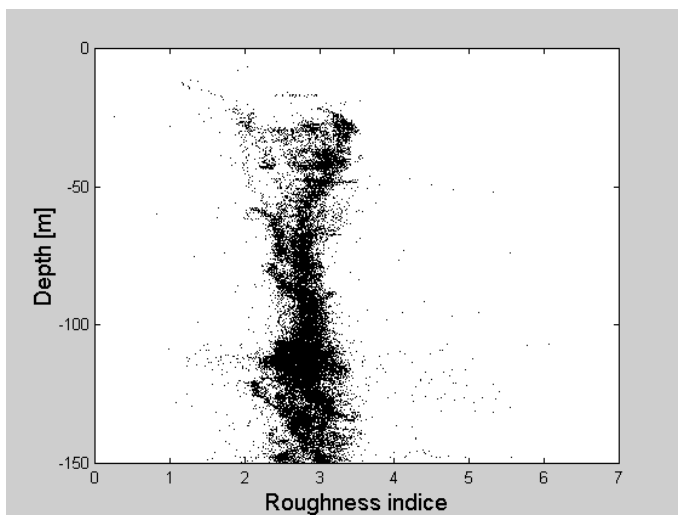
Figure 2.12. Scatterplot of roughness index against depth for three operated frequencies. Integration interval =  $8^\circ$  to  $40^\circ$  plus 3 m offset. Start angle of  $8^\circ$  plus 3 m offset corresponds to the oblique depth at which backscattering strength is 35 dB below the maximum peak.



(a) 12 kHz



(b) 38 kHz



(c) 120 kHz

Figure 2.13. Scatterplot of roughness index against depth for three operated frequencies. Integration interval =  $14^\circ$  to  $40^\circ$  plus 3 m offset. Start angle of  $14^\circ$  plus 3 m offset corresponds to the oblique depth at which backscattering strength is 45 dB below the maximum peak.

the maximum peak. The corresponding integration interval was between 14° and 40° with an offset of 3 m. The backscattering strength produced by using this angular integration interval is independent of depth for 38 and 120 kHz but for 12 kHz it is slightly dependent on depth less than 50 m. This is due to the fact that the 12 kHz transducer has a wider beamwidth and a longer pulse length than 38 and 120 kHz transducers. It was found that the appropriate angular integration interval for 12 kHz was between 0° and 35° and an offset of 9 m (Figure 2.14). The angular integration interval between 14° and 40° with an offset of 3 m for 38 and 120 kHz data (Figures 2.13(b) and(c)) and between 0° and 35° with an offset of 9 m for 12 kHz data (Figure 2.14) reflected the working hypothesis that on average  $\frac{\partial E1}{\partial R_0} = 0$ .

Figure 2.15(a) shows a scatterplot of the oblique depth at which backscattering strength is 45 dB below the maximum peak of the first bottom echo for the start angle versus the bottom depth along with regression lines. Figure 2.15(b) shows a scatterplot of the horizontal distance off the axis (at which the patch starts) resulting from the start angle versus the bottom depth. Since the start angle of the integration interval adopted here for 38 and 120 kHz, and for 12 kHz is different, the oblique depth associated with the corresponding start angle is also different as shown in Figure 2.15(a) as -- for 12 kHz and – – for 38 and 120 kHz, and so is the horizontal distance off the axis at which the patch starts as shown in Figure 2.15(b) as -- for 12 kHz and – – for 38 and 120 kHz. On average, the oblique depth at the start angle adopted for 38 and 120 kHz differs as far as 4.2 m from that adopted for 12 kHz (Figure 2.15(a)), and the start of patch derived from the start angle adopted for 38 and 120 kHz differs, on average, as far as 9 m from that adopted for 12 kHz (Figure 2.15(b)).

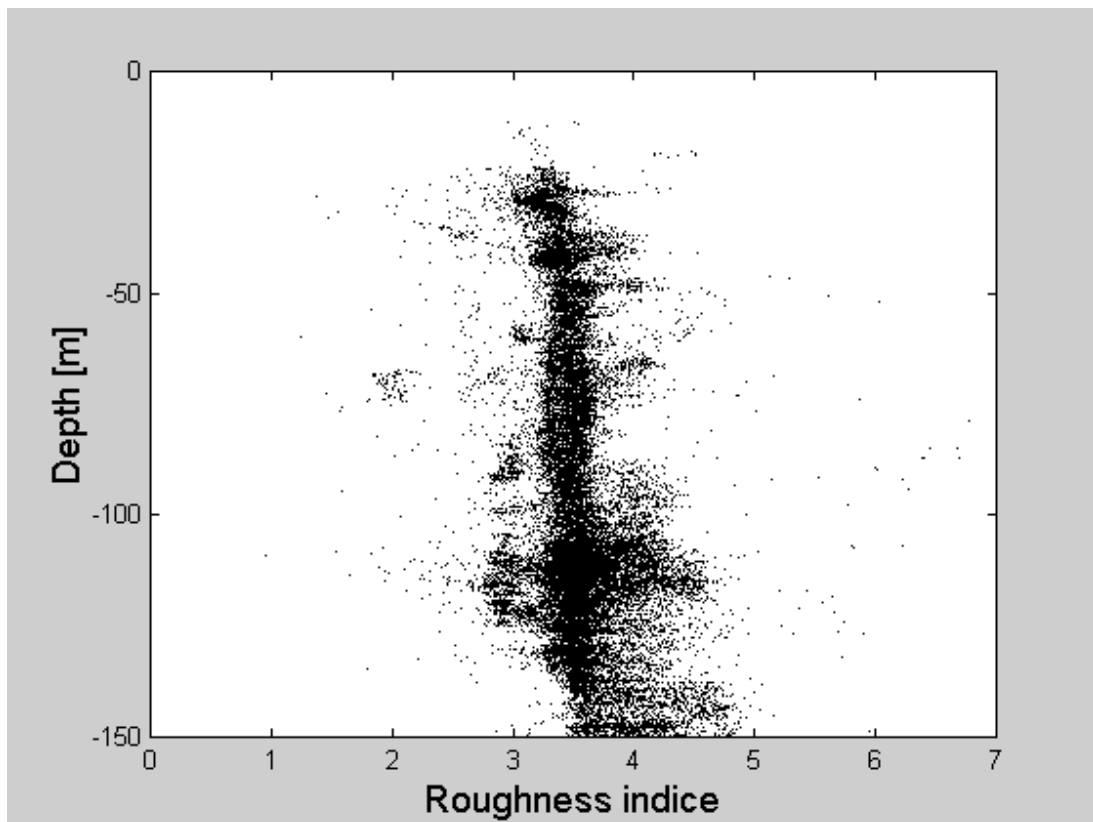
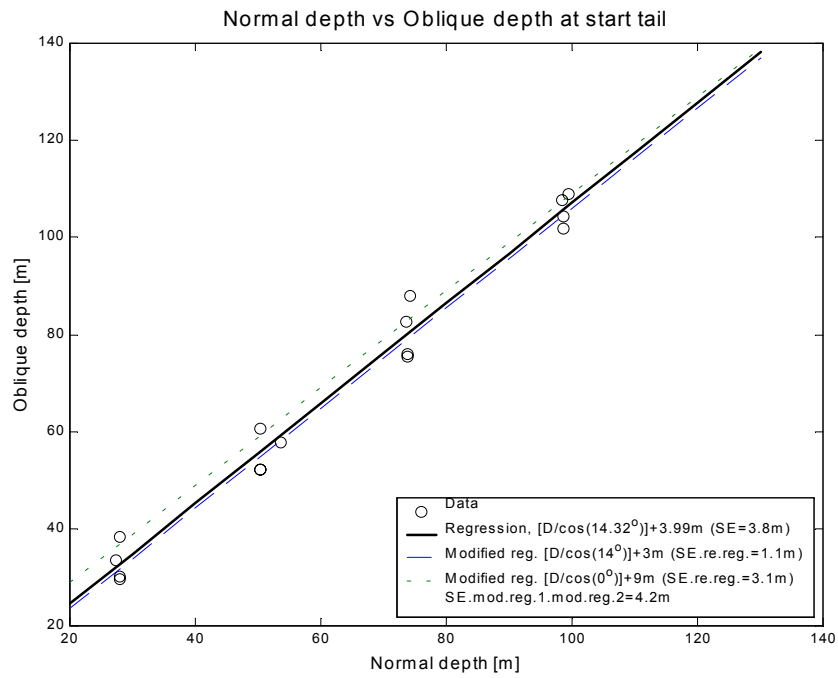
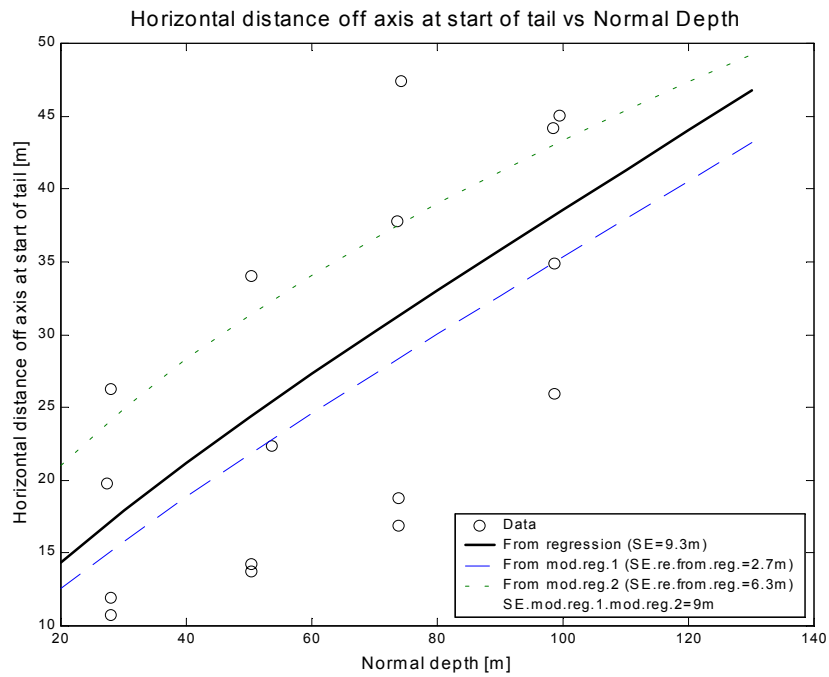


Figure 2.14. Scatterplot of roughness index against depth for 12 kHz data set.  
Integration interval =  $0^{\circ}$  to  $35^{\circ}$  plus 9 m offset.



(a)



(b)

Figure 2.15. Plot of oblique depth at which backscattering strength is 45 dB below the maximum peak of the first bottom echo for the start angle versus bottom depth (a) and the horizontal distance off the axis resulting from the start angle versus bottom depth (b).

#### 2.4.2. Constant depth integration interval for the entire second bottom returns

The coherent component becomes important when the second bottom echo is used as a measure of hardness. Unlike the tail sector of the first bottom echo, the initial sector of the second bottom echo due to the coherent component is merely dependent on depth. This is due to the fact that the coherent component in the initial sector of the bottom return is much bigger than the incoherent component in the tail component. The notion of adopting a constant angular integration interval in this case is to minimise the depth dependence of the results. However, since the initial sector of the second bottom echo is merely depth dependent, it is sufficient to adopt a constant depth integration interval. The key criterion is that the initial sector of the second bottom echo is selected such that the contribution of the coherent component is fully, or at least mostly, included. As a working hypothesis, again it was assumed here that on average  $\frac{\partial E2}{\partial R_0} = 0$  where E2 is the hardness index and  $R_0$  is the depth. The integration limit for the entire second bottom echo was defined as starting from twice the water depth ( $d_a$ ) and ending at twice the water depth plus 30 m ( $d_b$ ); see Figure 2.7 for the geometry.

An alternative to using the second bottom echo to determine the seabed hardness would be to use the leading edge of the first bottom echo. This would increase the number of data for further analysis where the second bottom echo is not available. Unless the entire hardness parameter set is obtained from the first bottom echo, mixing the two types of hardness parameters, one derived from the second bottom echo and the other derived from the first bottom echo, involves the use of estimates from signals of widely differing amplitude. Further analysis is required concerning how the hardness parameter derived from the second bottom echo and that

derived from the first bottom echo are related and this is beyond the scope of this study.



## Chapter 3

### Research data collection and analysis

Data sets used in this study are, acoustic data including returns from first and second bottom echoes and targets in the water column, biological data including catch composition and length-frequency measurement from net trawls, and seabed photographs. Data collection in the NWS and SEF study areas was conducted from the CSIRO's research vessel *RV Southern Surveyor*. A flow diagram of data collection and processing from the various samplers used in this study is shown in Figure 3.1. Acoustic data were collected using a SIMRAD EK 500 scientific echosounder operating three frequencies (12, 38 and 120 kHz). A Photosea 1000 underwater camera was used to collect seabed photographs. For biological samples, a McKenna demersal trawl was used. Details of these data sets and some analysis are described in this chapter.

#### 3.1. Acoustic data

A collection of acoustic returns were obtained from the *RV Southern Surveyor* using a SIMRAD EK 500 echosounder operating three different frequencies, 12, 38 and 120 kHz. The 12 kHz transducer was a single beam unit whereas the 38 and 120 kHz transducers were split beam transducers. The volume reverberation signal  $S_v$  in logarithmic form implemented in the SIMRAD EK 500 echosounder is as follows;

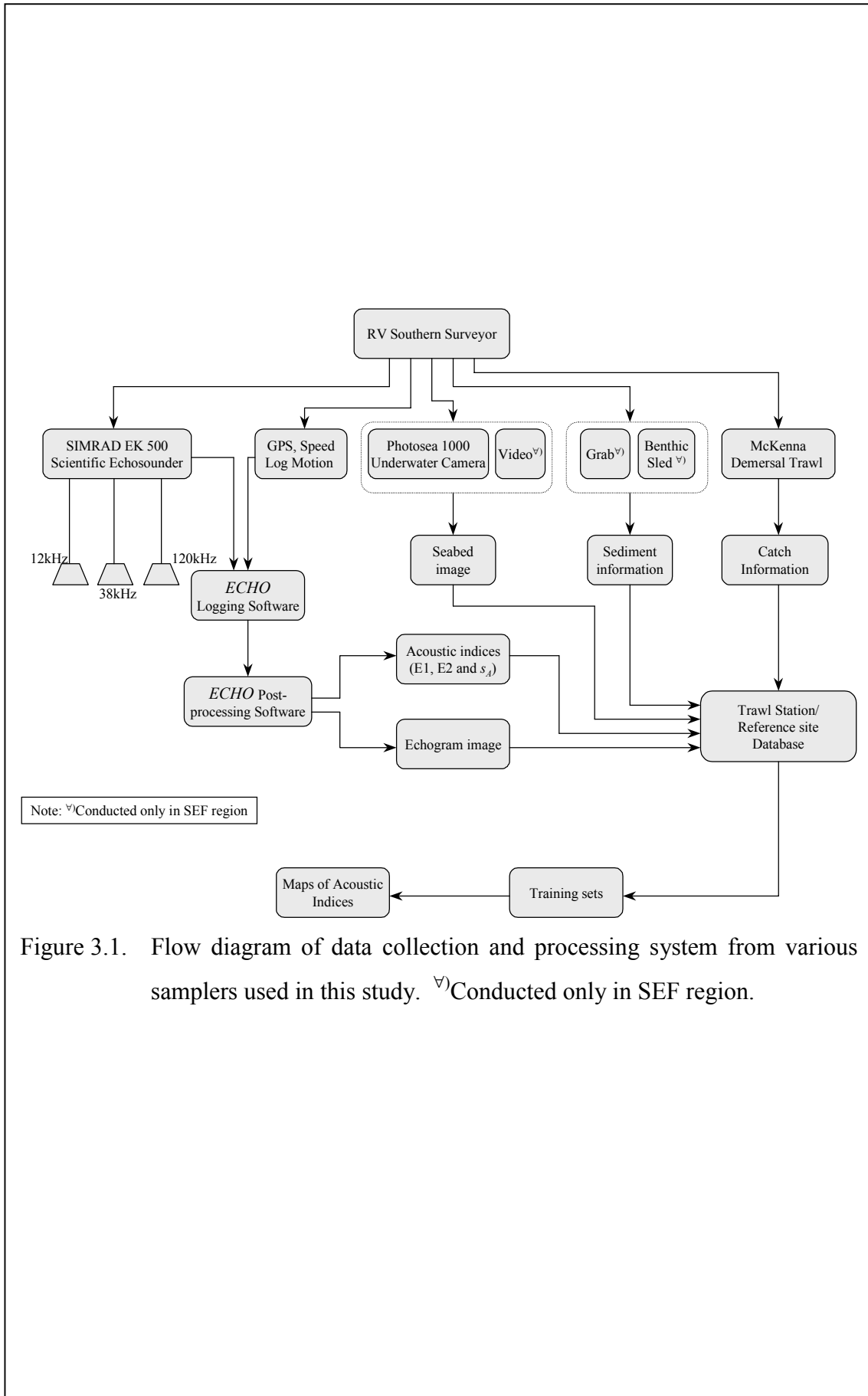


Figure 3.1. Flow diagram of data collection and processing system from various samplers used in this study. ∇)Conducted only in SEF region.

$$10\lg_{10}(s_v) = 10\lg_{10}(P_r) + 10\lg_{10}(r^2 10^{2\alpha r}) - 10\lg_{10}\left(\frac{P_t G_o^2 r_o^2 \lambda^2 c \tau \psi}{32\pi^2}\right) \quad (3.1)$$

where  $P_r$  is the received power,  $r$  is the range,  $\alpha$  is the seawater absorption coefficient,  $P_t$  is the transmitted power,  $G_o$  is the transducer gain,  $r_o$  is the reference range (= 1 m),  $\lambda$  is the acoustic wavelength,  $c$  is the sound speed in seawater,  $\tau$  is the pulse length and  $\psi$  is the equivalent beamwidth. Acoustic volume reverberation  $S_v$  data were continuously logged using *ECHO*, a software package developed by CSIRO Marine Research (Waring *et al.*, 1994; Kloser *et al.*, 1998).

### 3.1.1. Acoustic calibration

The acoustic system was routinely calibrated with a standard sphere, either a -33.6-dB, 60-mm copper sphere or a -42.35-dB 38.1-mm tungsten carbide sphere. Calibration procedures due to Foote (1982; 1983) were used as described in the operation manual of the Simrad EK500 echosounder (Anon., 1993). The standard sphere was placed on axis within the acoustic beam and the observed target strength ( $TS$ ) was compared to the target strength of the standard sphere used. The observed area backscattering coefficient ( $S_A$ ) was obtained by averaging the integration of the acoustic sphere returns over a short period of time and compared to the theoretical  $S_A$  derived from the expression given in the operation manual. Using expressions given in the operation manual, the corrected  $TS$  and  $S_v$  gains were calculated. Adjustment to these gains in the echosounder was then made as necessary. Table 3.1 provides the calibration results of the Simrad EK500 echosounder used in the survey (Kloser *et al.*, 2001). The 12 kHz transducer is the triangular one whereas others are the circular ones.

Table 3.1. Calibration settings for the SIMRAD EK 500 echosounder (After Kloser, 2001).

<b><i>Survey: ss696. Study Area: SEF region</i></b>	Frequency		
	12 kHz	38 kHz	120 kHz
Absorption (dB/km)	1	9	43
Pulse length (mS)	3	1	1
Bandwidth (kHz)	1.2	3.8	1.2
Calibration constant Svc	13.3	27.2	22.7
Beamwidth between -3 dB points (degrees)	16/17.5 <sup>¶)</sup>	7.1	11.2
Equivalent beam width (dB re 1 steradian)	-13	-20.7	-18.5
<b><i>Survey: ss696. Study Area: SEF region</i></b>			
<b><i>Survey: ss895. Study Area: NWS region</i></b>			
Absorption (dB/km)	1	7	42
Pulse length (mS)	3	1	0.3
Bandwidth (kHz)	1.2	3.8	12
Calibration constant Svc	13.4	27.3	22.8
Beamwidth between -3 dB points (degrees)	16/17.5 <sup>¶)</sup>	7.1	11.2
Equivalent beam width (dB re 1 steradian)	-13	-20.7	-18.5
<b><i>Survey: ss797. Study Area: NWS region</i></b>			
Absorption (dB/km)	1	7	42
Pulse length (mS)	1	1	1
Bandwidth (kHz)	1.2	3.8	1.2
Calibration constant Svc	13.2	27.1	22.6
Beamwidth between -3 dB points (degrees)	16/17.5 <sup>¶)</sup>	7.1	11.2
Equivalent beam width (dB re 1 steradian)	-13	-20.7	-18.5

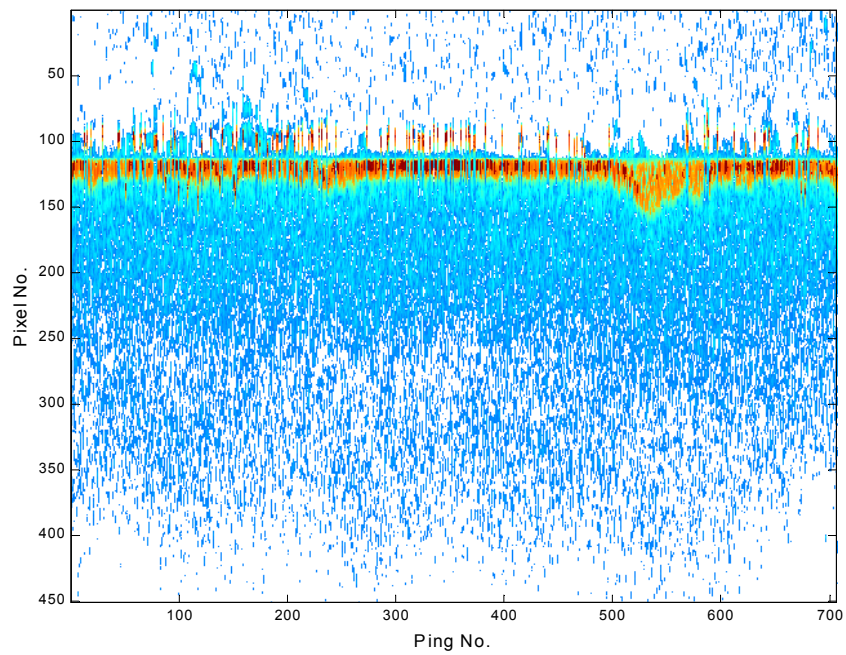
<sup>¶)</sup>A rectangular transducer.

### 3.1.2. Acoustic data quality control

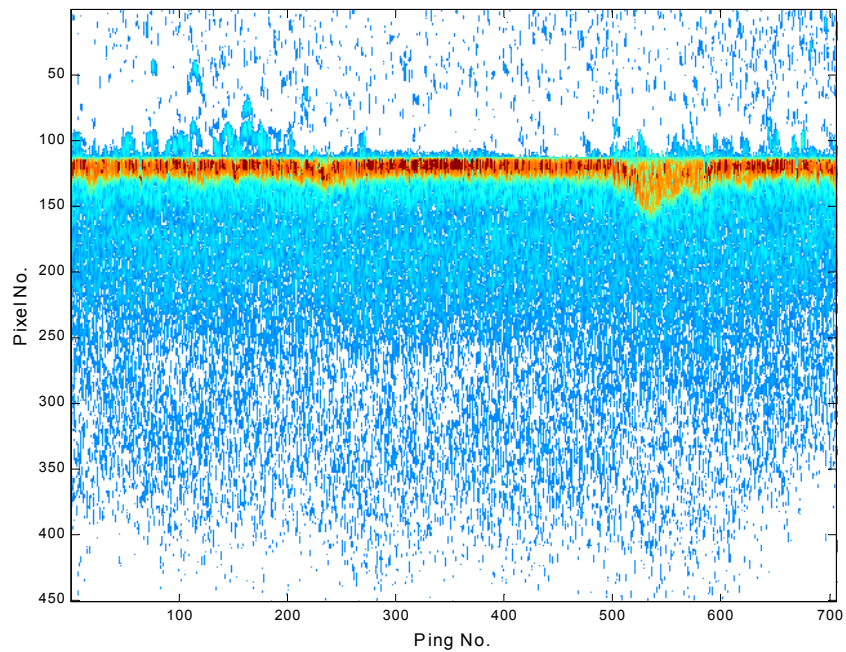
The quality control of the acoustic data was performed using *ECHO* software developed by the CSIRO Division of Marine Research on a Sun workstation (Waring *et al.*, 1994; Kloser *et al.*, 1998).

The software was used to edit and to analyse the acoustic data. This software enables the specification of background and spike noise thresholds, correction for calibration and absorption changes, removal of corrupted data and editing of bottom lines (Kloser *et al.*, 1996). Using *ECHO*, regions containing acoustic noise due to aeration and spike noise above the seabed due to a time jitter were excluded from the data set prepared for further analysis.

Substantial spike noise above the seabed due to a time jitter was found in the 120 kHz data set in the three cruises, two in the NWS region and one in the SEF region. A representative example of the spike noise is shown in Figure 3.2(a). Because of the semi-continuous occurrence of the spike noise, such records could not be simply marked bad and excluded from further analysis but instead were corrected. Simple MATLAB codes were created to explore the possibility of correcting this spike problem (Appendix C). It was found that spikes above the seabed were actually first bottom returns that had been shifted a few pixels up back in time, i.e. upwards on the echogram. It was also found that some first bottom returns had been shifted down as well. The procedure to correct the spike problem for 120 kHz bottom-locked data depicted in Figure 3.3 is as follows. First, a time interval within which the maximum bottom return,  $S_V$  [dB], occurs is defined. Once this time interval has been established, the next step is to find pings having maximum  $S_V$  values which fall outside this time interval. For fluency, these pings are called spiky pings. The third



(a)



(b)

Figure 3.2. Representative example of echogram for 120 kHz data set. (a) corrupted echogram due to spiky pings. (b) spiky-ping-corrected echogram after implementing the simple MATLAB codes given in Appendix C.

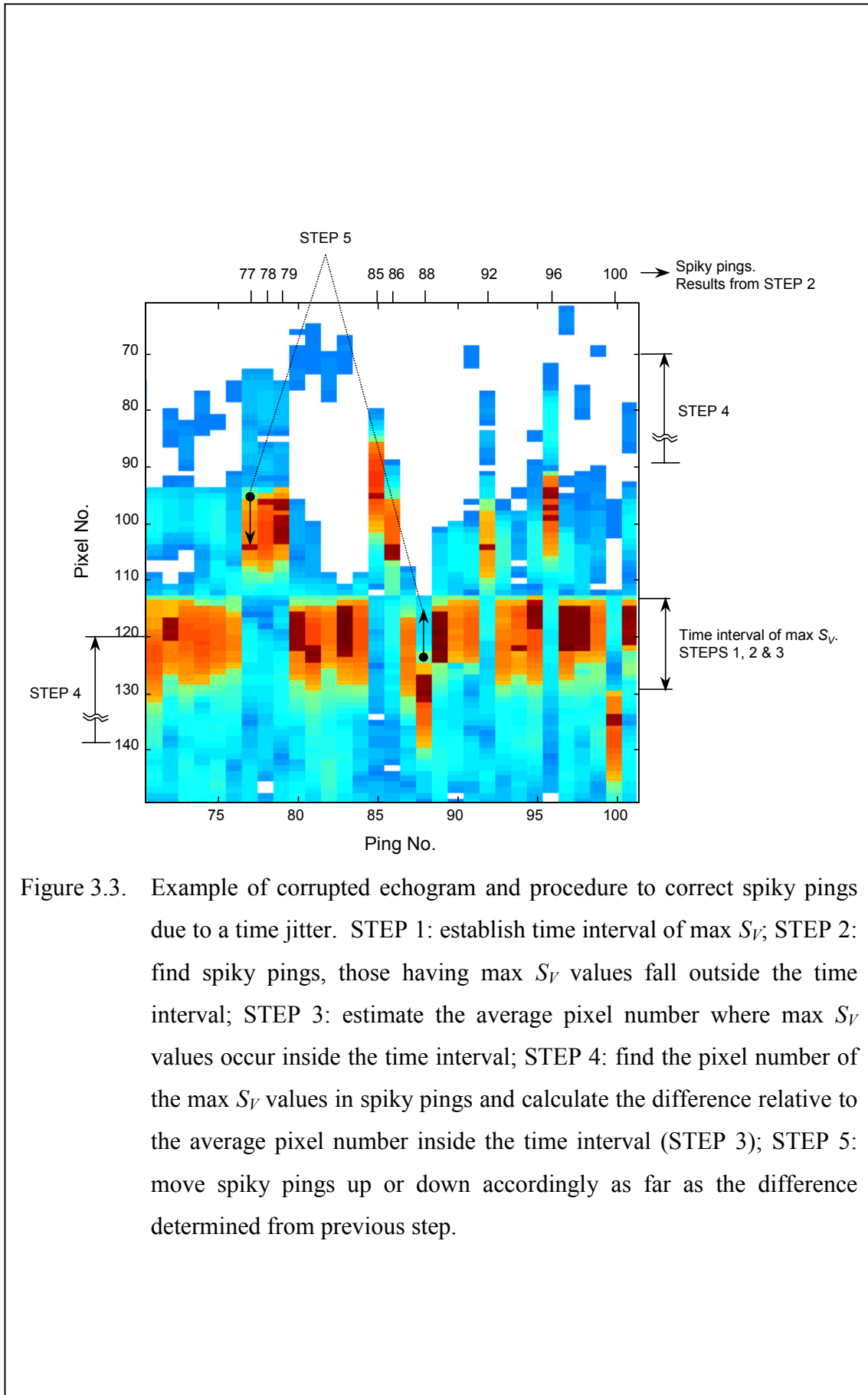


Figure 3.3. Example of corrupted echogram and procedure to correct spiky pings due to a time jitter. STEP 1: establish time interval of max  $S_V$ ; STEP 2: find spiky pings, those having max  $S_V$  values fall outside the time interval; STEP 3: estimate the average pixel number where max  $S_V$  values occur inside the time interval; STEP 4: find the pixel number of the max  $S_V$  values in spiky pings and calculate the difference relative to the average pixel number inside the time interval (STEP 3); STEP 5: move spiky pings up or down accordingly as far as the difference determined from previous step.

step is to find the average pixel number in which the maximum  $S_V$  values exist inside the time interval. Fourth, the MATLAB code searches in the spiky pings for the pixel number where the maximum  $S_V$  values exist and calculates the difference relative to the average pixel number of maximum  $S_V$  values inside the time interval. Finally, spiky pings are shifted up or down as far as the difference resulting from the previous step. Implementing this procedure to Figure 3.2(a) results in Figure 3.2(b). With a slight modification, this procedure was added to the *ECHO* software. While the simple MATLAB codes corrected the spiky pings by shifting them up and down, the algorithm in the *ECHO* software did not shift the spiky pings up and down but instead drew a new bottom line which moved with the time jitter.

Bottom editing was required for echoes from rough seabed area. The *ECHO* software allows redrawing a new bottom line at any corrupted bottom. For biomass assessment, the user raises the detected bottom depth to avoid integrating any unwanted seabed returns in the integration process. The intense seabed returns if included would greatly bias the calculation of the area backscattering coefficient arising from benthic biota. The estimated height of this zone can be calculated from the pulse length and the offset, the difference between the edited bottom signal depth and the true seabed depth. This represents the volume of water that is not directly sampled with acoustics but requires later corrections (Ona and Mitson, 1996; Kloser, 1996; Kloser *et al.*, 1996).

In addition to biomass assessment, the background noise and unwanted returns from plankton were also subtracted from the data set as outlined below in section 3.1.4.2. This noise removal however was not carried out within the *ECHO* software but rather was performed manually after the integration had been completed.



Although the SIMRAD EK 500 scientific echosounder is able to log GPS data simultaneously, GPS data are not available for the entire acoustic data from cruise ss895 and much of the area covered by the acoustic data from cruise ss797, both in the NWS region, due to technical reasons. Other instruments, however, recorded the underway GPS data with a time resolution of one minute and stored them in a digital format. Since, in this case, the time resolution and timestamp of the GPS records were different from those of the acoustic data, there was a need to first adjust the time resolution as well as timestamp of the acoustic data to match those of the GPS records. To match the time resolution and timestamp of the two data sets, the timestamp of the acoustic data was averaged to the nearest minute. Consequently, the acoustic data, which is the area backscattering coefficient  $s_A$  (see section 3.1.3), and other data such as depth and vessel log were averaged within one minute intervals. Due to the enormous amount of acoustic data, the averaging process was conducted by using a dedicated shell script within a Unix environment (Appendix D). Using the new timestamp of the acoustic data and the timestamp of the GPS records extracted from other instruments, the longitude and latitude from the GPS records were imported into the acoustic data using Microsoft Access (Figure 3.4).

### 3.1.3. Acoustic data analysis

Having finished quality control, three different parameters were produced by using the *ECHO* software. They are E1 and E2 parameters adopted from the RoxAnn system that represent seabed roughness and hardness, respectively, and the area backscattering coefficient,  $s_A$ , that represents a relative measure of biomass. In this section, derivations of E1, E2 and  $s_A$  parameters are described.

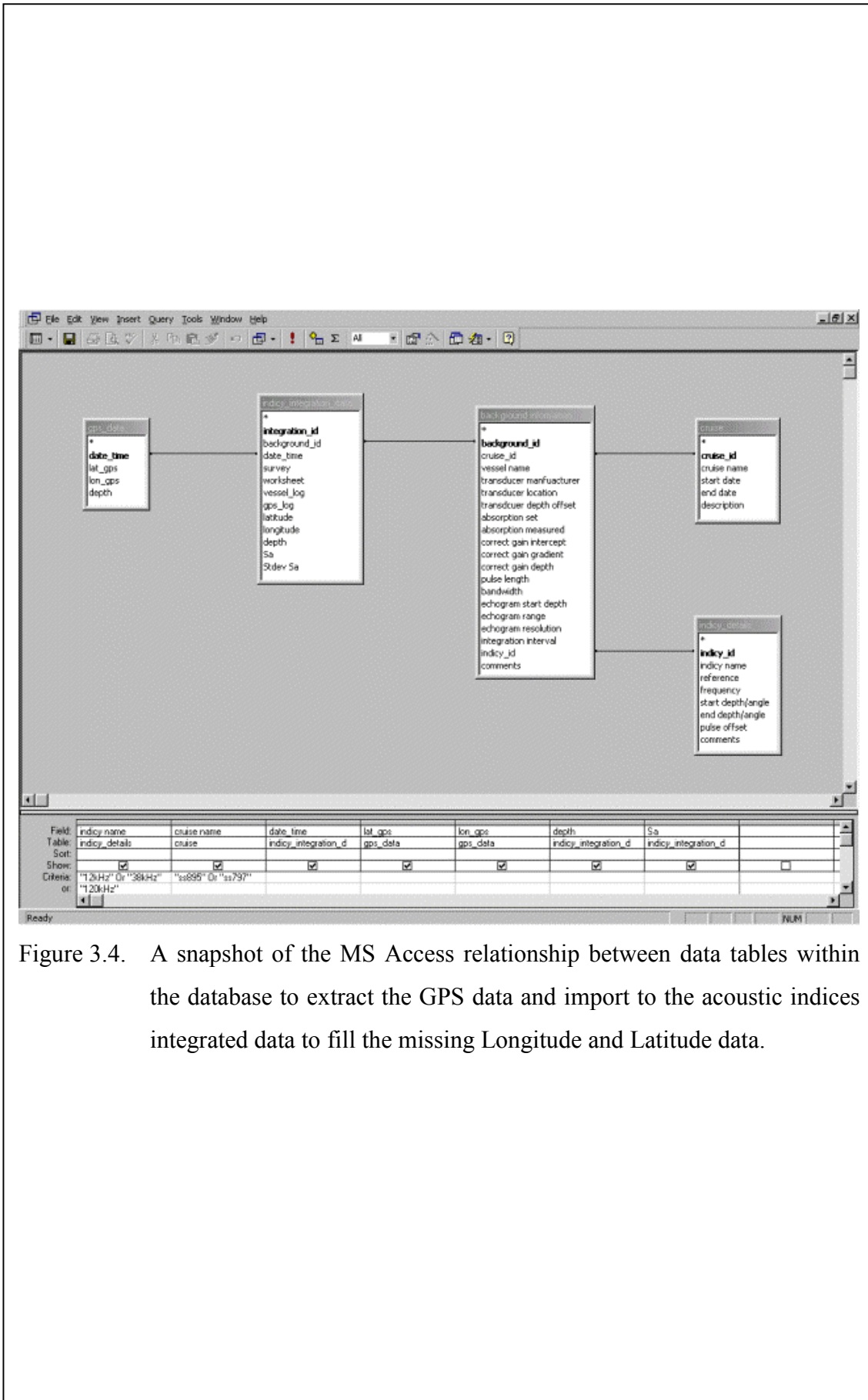


Figure 3.4. A snapshot of the MS Access relationship between data tables within the database to extract the GPS data and import to the acoustic indices integrated data to fill the missing Longitude and Latitude data.

### 3.1.3.1. Seabed parameters

Simple indices of seabed roughness (E1) and hardness (E2) were derived from the first and second acoustic bottom returns. The E1 parameters were derived by integrating the tail of the first acoustic bottom returns and the E2 parameters were derived by integrating the complete second bottom echoes (Orlowski 1984; Chivers *et al.*, 1990; Heald and Pace, 1996).

For analysis of acoustic bottom returns, the *ECHO* software provides several algorithms including a constant angular algorithm; see equation (2.6). This algorithm ensures that a constant angular sector of the incoherent field, irrespective of depth changes, is used for the integration of the first acoustic bottom backscatter. From section 2.4 in Chapter 2, the integration limit ( $\theta_a$  and  $\theta_b$  in equation (2.6) and Figure 2.7) after an offset of 9 m, was between  $0^\circ$  and  $35^\circ$  for 12 kHz data and between  $14^\circ$  and  $40^\circ$  after an offset of 3 m for 38 and 120 kHz data. Depths (after the bottom) corresponding to  $\theta_a$  and  $\theta_b$  varied with changing water depths and were estimated by

$$d_{\theta_i} = R_o / \cos \theta_i - R_o + c_\tau \quad (3.2)$$

where  $R_o$  is the bottom depth in meters and  $c_\tau$  is an offset in meters (9 m for 12 kHz data and 3 m for 38 and 120 kHz data). A constant depth algorithm was used for the integration of the complete envelope of the second acoustic bottom backscatter. The integration limit was defined as starting from twice the water depth ( $d_a$ ) and ending at twice the water depth plus 30 m ( $d_b$ ); see Figure 2.7. To reduce variability between pings in the backscatter returns and to standardise on a unit of length sampled, the integration was averaged over an along-track interval of 0.05 nmi. The integration of acoustic volume reverberation resulted in area backscatter coefficients that stem from

fisheries acoustics for biomass assessments and are adopted as a relative measure of acoustic energy for scattering from the seabed

$$\bar{s}_A = 1852^2 4\pi \frac{\sum_{p=1}^{n_p} \delta d \sum_{d=d_s}^{d_f} s_{v(dp)}}{n_p} \quad (3.3)$$

where  $s_v$  is the linear volume backscattering coefficient (equation 3.1)),  $\delta d$  is the acoustic sampling interval and  $n_p$  is the number of pings within 0.05-nmi along track interval. E1 and E2 parameters are obtained by taking the logarithm of the corresponding  $\bar{s}_A$ .  $d_s$  is the start depth for the integration that corresponds to  $\theta_a$  (equation 2.6 and Figure 2.7) by equation 3.2 for E1 and to  $d_a$  for E2, and  $d_f$  is the end depth for the integration that corresponds to  $\theta_b$  (equation 2.6 and Figure 2.7) by equation 3.2 for E1 and to  $d_b$  for E2.

### 3.1.3.2. Relative measure of biomass ( $s_A$ )

The acoustic returns,  $s_v$ , were integrated in 0.5 meter (m) depth layers between 0 and 5 m above the line defined as the seabed and averaged horizontally over 0.05 nautical mile (nmi) or 0.0926 km intervals. This gave the mean area backscattering coefficient,  $\bar{s}_{Aj}$ , of layer  $j$  over the horizontal interval. For layer  $j$ , the mean area backscattering coefficient was calculated by

$$\bar{s}_{Aj} = 4\pi 1000^2 \frac{\sum_{l=1}^{n_p} \left( 2 \sum_{k=1}^d s_{v(k,l)} \right)}{n_p} \quad [\text{m}^2 \text{km}^{-2}] \quad (3.4)$$

where  $s_v$  is the volume backscattering coefficient in linear scale,  $n_p$  is the number of pings and  $d$  is the number of volume backscattering coefficients within the layer  $j$ .

The removal of unwanted returns from plankton was carried out as follows. Minimum values of the mean area backscattering coefficient within each one-hour interval were first extracted. These minimum values were assumed to arise from plankton. The minimum value was then subtracted from all mean backscattering coefficients inside each one-hour interval. The one-hour interval was adopted to take into account the vertical migration of plankton in time but also to keep the distance between targets inside the block interval as reasonably close as possible.

Adjusting for the distance between the transducer and the trawl estimated from warp length data allowed a direct comparison of acoustic and net data, the area backscattering coefficient,  $\bar{s}_{Aij}$ , for trawl  $i$  and layer  $j$  was obtained by averaging the  $\bar{s}_{Aj}$  values within the duration of the trawl deployment:

$$s_{Aij} = \frac{1}{n_s} \sum_{n_s} \bar{s}_{Aj} \quad [\text{m}^2\text{km}^{-2}] \quad (3.5)$$

where  $n_s$  is the number of  $\bar{s}_{Aj}$  values available within the  $i^{\text{th}}$  trawl duration.

The area backscattering coefficient of the region swept by nets during the trawl duration resulted from summing up the area backscattering coefficients of all sequential layers falling in the swept area. The total height of these sequential layers in the vertical direction was therefore put identical to the headline height and was in general 2 m. The headline height was measured with Scanmar net sensors. The area backscattering coefficient for trawl  $i$  was then given by

$$s_{Ai} = \sum_{n_l} s_{Aij} \quad [\text{m}^2\text{km}^{-2}] \quad (3.6)$$

where  $n_l$  is the number of layers falling in the swept area.

### 3.2. Net data

The demersal trawl used in this survey had a mouth opening of around 40 m<sup>2</sup> with 20-m wingspread and 2-m headline height. Mesh sizes were 9 inch (in) at wings and belly, 6 in and then 4.5 in at the funnel leading to cod end, and 3.5 in at the cod end. Scanmar net sensors were operational for all of the trawls to measure the door spread, wingspread and headline height. These data were recorded on the vessel's logging system as well as the vessel's logbook.

#### 3.2.1. Catch collection

Immediately after each trawl, catches were identified on board and sorted to a species level. Only very few unidentified fish were retained for later identification. All sorted species were individually wet weighed. For the two surveys in the NWS region (ss895 and ss797), length-frequency data were collected for 14 pre-determined species of fish (*Saurida undosquamis*, *Saurida* sp. 2, *Epinephelus aerolatus*, *E. multinotatus*, *Lutjanus vittus*, *L. sebae*, *L. malabaricus*, *Nemipterus furcosus*, *N. celebicus*, *Diagramma labiosum*, *Lethrinus* sp., *L. nebulosus*, *L. genivittatus* and *Parupeneus heptacanthus*). The lengths of the fish were based on fork length. For the survey in the SEF study area, all species caught were length measured and counted.

#### 3.2.2. Net data analysis

In order to make the acoustic and net data comparable, a conversion is required either from the acoustically derived area backscattering coefficients into units to suit the net derived density estimates or from the density estimates into the area backscattering coefficients. The latter approach was used by Aglen (1996) and Siwabessy *et al.* (2000) and was adopted for this study. The procedure is as follows;

The density estimates were converted to theoretical  $s_A$  values using the backscattering cross section,  $\sigma_{sl}$ , for species  $s$  and length group  $l$ :

$$S_{A(catch)} = \sum n_{sl} \sigma_{sl} / A_t \quad [\text{m}^2 \text{km}^{-2}] \quad (3.7)$$

and

$$\sigma_{sl} = 4\pi 10^{TS_{sl}/10} \quad [\text{m}^2] \quad (3.8)$$

where  $n_{sl}$  and  $TS_{sl}$  are respectively the number of catch and the target strength for species  $s$  and length group  $l$ , and  $A_t$  is the towing area in  $\text{km}^2$ .

The target strength of fish, a parameter which indicates the magnitude of the fish echo, is a pivotal measure for it determines the accuracy of the conversion of the fish echo integration into fish density. The magnitude of the target strength of fish depends upon the acoustic impedance mismatch of the fish and on the physical area presented to the beam and the surrounding water; the greater the acoustic impedance mismatch, the stronger is the fish echo. The part of the fish body which has the greatest acoustic impedance mismatch with the water is the swimbladder for those fish possessing this organ, accounting for 90% or more of the energy of the acoustic fish return (MacLennan and Simmonds, 1992). The target strength however is highly variable due to changes in the orientation of the fish body, the presence and absence of the swimbladder, and the physiological state of the swimbladder. It is therefore appropriate to consider the target strength as a stochastic parameter whose mean value is derived from the average over a number of measurements.

The presence and type of swimbladder are species specific, hence net catches were categorised into 4 different groups namely physoclistous (closed bladder),

physostomous (open bladder), bladderless (no bladder) and "squid-like" species. As far as bladderless and squid-like species are concerned, a general equation of the target strength for each one of them, unlike that of physoclistous and physostomous species, is not available. Equations for target strength specific to some species of these sorts, though, are available in few publications such as for squid by Arnaya *et al.* (1989) and for mackerel by Edwards *et al.* (1984). The equations of the target strength of mackerel and squid were then used in this study to estimate the target strength of bladderless and squid-like species. The equations of the target strength used in this study for the four groups of species are then given by

$$\begin{aligned}
 TS &= 20\log_{10} L - 67.4 \ (\pm 2.3 \text{ dB}) \text{ for physoclist (Foote, 1987)} \\
 TS &= 20\log_{10} L - 71.9 \ (\pm 1 \text{ dB}) \text{ for physostome (Foote, 1987)} \\
 TS &= 20\log_{10} L - 84.9 \ (\pm 0.5 \text{ dB}) \text{ for bladderlessness (Edwards } et al., 1984) \\
 TS &= 20\log_{10} L - 75.4 \ (r = 0.91) \text{ for squid-like (Arnaya } et al., 1989)
 \end{aligned}
 \tag{3.9}$$

where  $L$  is the total length for the first three categories and is the mantle length for fourth in centimetres (cm).

The target strength adopted in this study was estimated from equation (3.9) rather than the relationship between acoustic data and net data since the *in-situ* measurement of the target strength was not conducted in the both surveys. In addition, it was hard to include all species from the tropical multi-species environment, as the current acoustic technology has not yet provided a positive identification of recovered organisms.

Target strength estimates for marine biota continue to be the subject of refinement (e.g. Pauly and Penrose (1998) concerning Antarctic krill and Kloser *et al.* (2000) concerning Black Oreo). It is clearly possible that the estimates shown in equation (3.9) may be refined in the future. The effect of a notional 5 dB increase in



$TS$ , applied to all the species listed above, would increase the  $S_A$  estimates of nets by factor of 3.

Crustacea, in particular sea bugs, were also caught on a very few occasions. Even though Penrose and Kaye (1979) in their review of target strengths of marine animals have shown that the empirical target strength formulation due to Love is applicable, in geometric region scattering, to some marine animals including crustacea, they were however excluded from the analysis, given the near seabed acoustic sampling problems, the low number of crustacea, and their low  $TS$ .

For species with length-frequency measurements unavailable from the survey, there was a need to predict the number of organisms in the catch and their mean length in a scientifically acceptable way. The procedure following Siwabessy *et al.* (2000) is as follows. Using the assumptions that the mean length and weight of an individual species were the same among a similar species within all trawl stations, the first step was to take the minimum weight of a similar species from all trawl stations and assume this minimum as a single target for that particular species. The next step was to check this minimum value to the asymptotic weight,  $W_{inf}$ , of this particular species. The asymptotic weight is the maximum weight possibly reached by one particular species for an infinite life expectancy. If the minimum weight was however greater than the asymptotic weight, the assumption, "a minimum weight for a single target", no longer held. A further assumption was then required. In this case, it was assumed that the minimum weight was accounted for more than 1 target. It could be 2, 3 or 4 targets depending on whether or not the weight of the individual target was below the asymptotic weight. Once this condition had been reached, the predicted

number of catch was given by the weight ratio. The mean length was then estimated by the weight-length relationship given by

$$W = aL^b \quad [\text{g}] \quad (3.10)$$

where  $a$  and  $b$  are parameters which are species dependent, and  $L$  is the length in cm. Parameters  $a$ ,  $b$  and  $W_{inf}$  were taken mainly from Froese and Pauly (1998) and from Benedito-Cecilio *et al.* (1997), Garcia-Arteaga *et al.* (1997), Kochzius (1997), Merella *et al.* (1997), and Ruiz-Ramirez *et al.* (1997).

Having finished predicting the number of catch and the mean length, a similar procedure to estimate the  $S_{A(catch)}$  values to that of the acoustic case was applied.

### 3.3. Underwater photographs

For ground truthing in the NWS study area, photographic surveys of the seabed were taken in trawl stations using a Photosea 1000 underwater camera. In the SEF study area, photographic surveys of the seabed and grab samples were taken only in reference sites of different seabed types.

The camera together with two strobes and a timer was attached to the mouth of the net. The elapsed time at each trawl station, from when the trawl reached the bottom until it left the bottom, was 30 minutes. Pictures were taken at 24-second intervals for the full duration of each trawl station, yielding around 75 pictures per 30-minute trawl or one picture per 47 metres. The pictures were processed on board to allow underway monitoring and adjustment of photographic parameters.

Only photographs from trawl stations in which E1 and E2 parameters were available were analysed. However, due to difficulties with camera malfunctions,

lighting, camera direction, gear entanglement and reflection of stirred up sediments, not all photographs in these trawl stations were interpretable. Several interpretable seabed images were selected from each of these trawl stations. Each selected seabed image represented similar physical attributes appearing within consecutive seabed images. Each interpretable seabed image that was selected was projected onto a screen and physical attributes, and types of epibenthic invertebrates (called benthos hereafter) were scored. Detailed descriptions of physical attributes and types of benthos are presented in Table 3.2. The score data (in percentages) were applied to the hierarchical agglomerative clustering technique (see section A.2.2). Representative seabed images are presented in section 5.3 of Chapter 5.

In the SEF region, the acoustic system intensively sampled the reference sites. The seabed and benthos types of these reference sites were based on the seabed images and grab samples. A range of reference sites was selected that provided distinct contrasts in seabed type and depth. These reference sites were located within mesohabitat areas known to have heterogeneous benthos (Kloser *et al.*, 2001; Bax and Williams, 2000).

Table 3.2. List and description of variables scored from seabed images.

Category	Variable	Description
Seabed morphology	Flat	Bottom is either flat or irregular due to sandwaves or bioturbation
	Ripple	All size of ripples (including bioturbation) typically 4 ripples up to less than 1 ripple per metre
	Ridge	Looks like a hill or a long tunnel in the image
	Rocky outcropping	Large individual rocks
Sediment type	Fine sand	
	Coarse sand	
	Mud	
Benthos type/shape	Lump	Lump shaped object; mostly sponge
	Cup	Cup shaped sponge
	Finger	Conical resembling finger; mostly sponge
	Gorgonian	Fan shaped, lacy, brittle looking organisms
	Grass	Grass looking organisms

## Chapter 4

### Comparison between acoustic and catch estimates

#### 4.1. Introduction

Acoustic surveys have now been widely accepted as a standard methodology for fisheries assessment. MacLennan and Simmonds (1992) provide a summary of the methods. Net surveys is a methodology which has been in use since long before the advent of acoustic techniques and continues to be employed.

The advent of acoustic techniques in fisheries assessment has led to comparisons with conventional, net techniques. In principle, comparisons between the two techniques can be simply derived from the catch data and the acoustic data along the towing track of a trawl. There are, however, few publications of such direct comparisons between the results of the two techniques. Examples include Aglen (1996) and Koslow *et al.* (1997) comparing the two techniques on fish, Everson and Bone (1986) and Everson (1987), both on krill.

The two techniques are different in nature. Nets are known to provide size selectivity arising from mesh dimension when used to sample marine organisms (ICNAF, 1963 (cited in MacLennan and Simmonds, 1992)). Such net results are more likely to provide underestimates of actual biomass. In addition, net avoidance by some marine organisms also gives rise to an underestimated result. Acoustic techniques may produce higher biomass estimates than net surveys. In a few cases on the other hand, acoustic techniques produce lower results than netting (Jin, 1990

(cited in Aglen, 1996); Sigurdsson, 1993 (cited in Aglen, 1996); Everson, 1987). An underestimated result from acoustics may arise from an overestimate of target strength to size relationship in the target population. Another source of underestimated results from acoustics is due to the specific behaviour of some fish. Cod, for instance, which are distributed in mid-water and might not be fully insonified, will dive just after the passage of a vessel and get caught in the trawl net towed on the seabed behind the vessel (Ona and Godo, 1990). Demersal nets in this case can produce higher biomass estimates than acoustics. Moreover, whereas net surveys discontinuously cover a relatively small volume of water, acoustic surveys on the other hand continuously sample a large volume of water along a vessel's track. Net techniques retain, however, a significant advantage over acoustic techniques in providing positive identification of recovered organisms.

In the case of demersal fish, there are zones in which returns of fish close to the bottom are indistinguishable from those of the bottom. This zone is called the Integrator Dead Zone, IDZ (Ona and Mitson, 1996) or Acoustic Shadow Zone (Kloser 1996, Kloser *et al.*, 1996). The inclusion of corrections for this zone in the echo integration results in higher values of the acoustic estimates. Ona and Mitson (1996) and Kloser (1996) provide a detailed discussion on this topic.

Although the echosounder operated hull-mounted transducers of three different frequencies, 12 kHz (single beam with 16/17.5° full angle of the triangular transducer), 38 kHz (split beam with 7.1° full angle) and 120 kHz (split beam with 11.2° full angle), in both study areas, this chapter will consider data from the 38 kHz hull mounted transducer collected in the NWS study area.

## 4.2. Species composition

Results from 72 trawl stations (see Table 4.1) with accompanying acoustic data are discussed. Where acoustic data were corrupted due to aeration effects, they were removed using the data editing function of *ECHO*. A summary of the 72 trawls is shown in Table 4.1.

The main trawl of interest was at trawl station 12 where about 80 kg of bonefish (*Albula neoguinaica*) were caught. This trawl was unique not only because bonefish made 51% by weight of the total catch but because they were found in an unexpected area. Normally, they are found in estuaries and mud-flats (Whitelaw, 1998). In trawl station 74, two different species from genus *Lutjanus* were caught totalling over 670 kg and accounting for 66% of the total catch. *Lethrinus* sp. were caught at trawl station 29 totalling over 630 kg and were 66% of the total catch.

Single species represented more than 50% of the total catch in 15 trawls namely 12, 15, 24, 28, 29, 31, 33, 37, 55, 64, 70, 74, 83, 112 and 113. Of these 15 trawls, 6 of them caught over 100 kg of a single dominant species. *Saurida undosquamis* was the single dominant species, accounting for more than 50% of the total catch, at trawl stations 28, 70, 83 and 112. Others such as *Nemipterus furcosus* and *Lethrinus* sp. were single dominant species respectively at trawl stations 15, 31 and 33 and trawl stations 29 and 64. Others such as *Albula neoguinaica*, *Saurida filamentosa*, *Leiognathus bindus*, *Pristotis jerdoni* and *Gnathanodon speciosus* were present as single dominant species only once at trawl stations 12, 24, 37, 55 and 113 respectively.

In trawl stations 42, 56, and 62, a single target accounted for more than 60% of the total catch. Trawl station 56 caught a single turtle of mass around 200 kg. This

Table 4.1. A summary of catches in 72 trawl stations (% by weight). <sup>1</sup>Length-frequency measurements available; <sup>2</sup>Over 100 kg; <sup>3</sup>Single target; \* Station excluded from analysis.

Trawl station	Catch [kg]	Dominant species	Proportion [%]	Mean Depth [m]	Trawl Station	Catch [kg]	Dominant species	Proportion [%]	Mean Depth [m]
8	69.82	n/a	n/a	159.9	70	85.18	<i>Saurida undosquamis</i> <sup>1</sup>	67	63.3
9	82.43	n/a	n/a	121	71	158.05	n/a	n/a	75.2
10	61.23	n/a	n/a	115	72	52.46	n/a	n/a	121.8
11	107.06	n/a	n/a	118.5	74	1022.53	<i>Lutjanus malabaricus</i> <sup>1,2</sup> + <i>Lutjanus quinquelineatus</i> <sup>2</sup>	66	120
12	154.5	<i>Albula neoguinaica</i>	51	78.5					
13	292.2	n/a	n/a	55.4	75	70.55	n/a	n/a	104.9
14	171.65	n/a	n/a	53.5	76	110.4	n/a	n/a	113.6
15	126.86	<i>Nemipterus furcosus</i> <sup>1,2</sup>	82	59.6	77	307.87	n/a	n/a	94.1
17	122.22	n/a	n/a	54.5	78	179.35	n/a	n/a	89.4
23	43.49	n/a	n/a	207.2	79	507.97	n/a	n/a	86
24	17.92	<i>Saurida filamentosa</i>	53	134	80	166.62	<i>Aetobatus narinari</i> <sup>3</sup>	51	72.9
26	121.97	n/a	n/a	72.2	81	177	n/a	n/a	68.5
27	225.49	n/a	n/a	70.7	82	168.46	n/a	n/a	65.1
28	38.56	<i>Saurida undosquamis</i> <sup>1</sup>	60	65.1	83	443.69	<i>Saurida undosquamis</i> <sup>1,2</sup>	59	62.8
29	962.48	<i>Lethrinus sp.</i> <sup>1,2</sup>	66	69.9	84	59.34	n/a	n/a	38.5
30	345.96	n/a	n/a	69.4	86	32.27	n/a	n/a	39.2
31	139.34	<i>Nemipterus furcosus</i> <sup>1</sup>	68	61.8	87	108	n/a	n/a	41.3
33	196.13	<i>Nemipterus furcosus</i> <sup>1,2</sup>	58	48.6	88	62.3	n/a	n/a	51.1
34	183.92	n/a	n/a	55.4	89	261.02	n/a	n/a	56.3
35	241.09	n/a	n/a	98.1	96	39.12	n/a	n/a	135.3
36	136.9	n/a	n/a	107.3	98	110.05	n/a	n/a	141.3
37	85.81	<i>Leiognathus bindus</i>	59	176.4	100	70.86	n/a	n/a	120.6
40	65.83	n/a	n/a	84.4	101	156.54	n/a	n/a	58.4
41	217.11	n/a	n/a	177.4	102	315.22	n/a	n/a	57.3
42	231.66	<i>Himantura fai</i> <sup>2,3</sup>	86	135.7	103	247.05	n/a	n/a	57.4
43	4.48	n/a	n/a	142.2	104	166.11	n/a	n/a	47.6
54	31.27	n/a	n/a	37.6	105	241.56	n/a	n/a	37.6
55	91.76	<i>Pristotis jerdoni</i>	52	31.3	106	146.26	n/a	n/a	31.9
56*	320.58	Turtle <sup>2,3</sup>	62	37.2	107	191.53	n/a	n/a	28.5
57	23.61	n/a	n/a	34.6	108	59.24	n/a	n/a	26.3
61	36.06	n/a	n/a	48	109	56.33	n/a	n/a	31.8
62	128.99	<i>Urogymnus asperrimus</i> <sup>2,3</sup>	78	47.6	110	336.3	n/a	n/a	39.7
64	210.87	<i>Lethrinus sp.</i> <sup>1,2</sup>	53	24.6	111	284.37	n/a	n/a	42.8
65	113.36	n/a	n/a	43	112	67.74	<i>Saurida undosquamis</i> <sup>1</sup>	51	68.2
66	253.63	n/a	n/a	43	113	243.8	<i>Gnathanodon speciosus</i> <sup>2</sup>	51	55.7
67	35.47	n/a	n/a	41.6	114	118.8	n/a	n/a	129.3
68	142.26	n/a	n/a	47.3					



trawl station was excluded when the two  $s_A$  estimates were compared one to another. Similarly, trawl stations 42 and 62 caught respectively a single *Himantura fai* of mass around 200 kg and a single *Urogymnus asperrimus* of about 100 kg. Unlike trawl station 42, these trawl stations were not excluded from the comparison. Lastly, a single *Aetobatus narinari* of mass 85 kg was caught in trawl station 80 and constituted 50% of the total catch mass.

### 4.3. Acoustic estimates

The net was dragged as close as possible to the seabed with the headline height varying from 1.5 to 2.5 m and the average close to 2 m. It was therefore assumed that the net swept a vertical range between 0.5 and 2.5 m above the seabed.

Despite the fact that times of trawl deployment and recovery, and warp lengths of all trawl stations were available from the survey, it was of interest to see whether or not there was any significant difference if these times were horizontally shifted backward (delay) and forward (ahead) around the nominal position (Figure 4.1). Consideration of this issue acknowledges the difficulty of assigning an exact value to the vessel-trawl separation distance. The nominal position was derived from the warp length and the water depth by a simple geometric calculation. 6 different lagged positions were compared to the nominal ones. Acoustic  $s_A$  estimates of trawl stations in the nominal position and those in lagged position are shown in Figure 4.2. It is evident that there is no difference of the acoustic  $s_A$  estimates between lagged positions in a range interval of 5 minutes ahead to 15 minutes delay, and the nominal ones. The Tukey-type test was however used to examine the difference between the correlation coefficients of the acoustic  $s_A$  estimates in different positions and those in the nominal ones  $r_{s_A(0)s_A(n)}$ , where n denotes the lagged position. The correlation

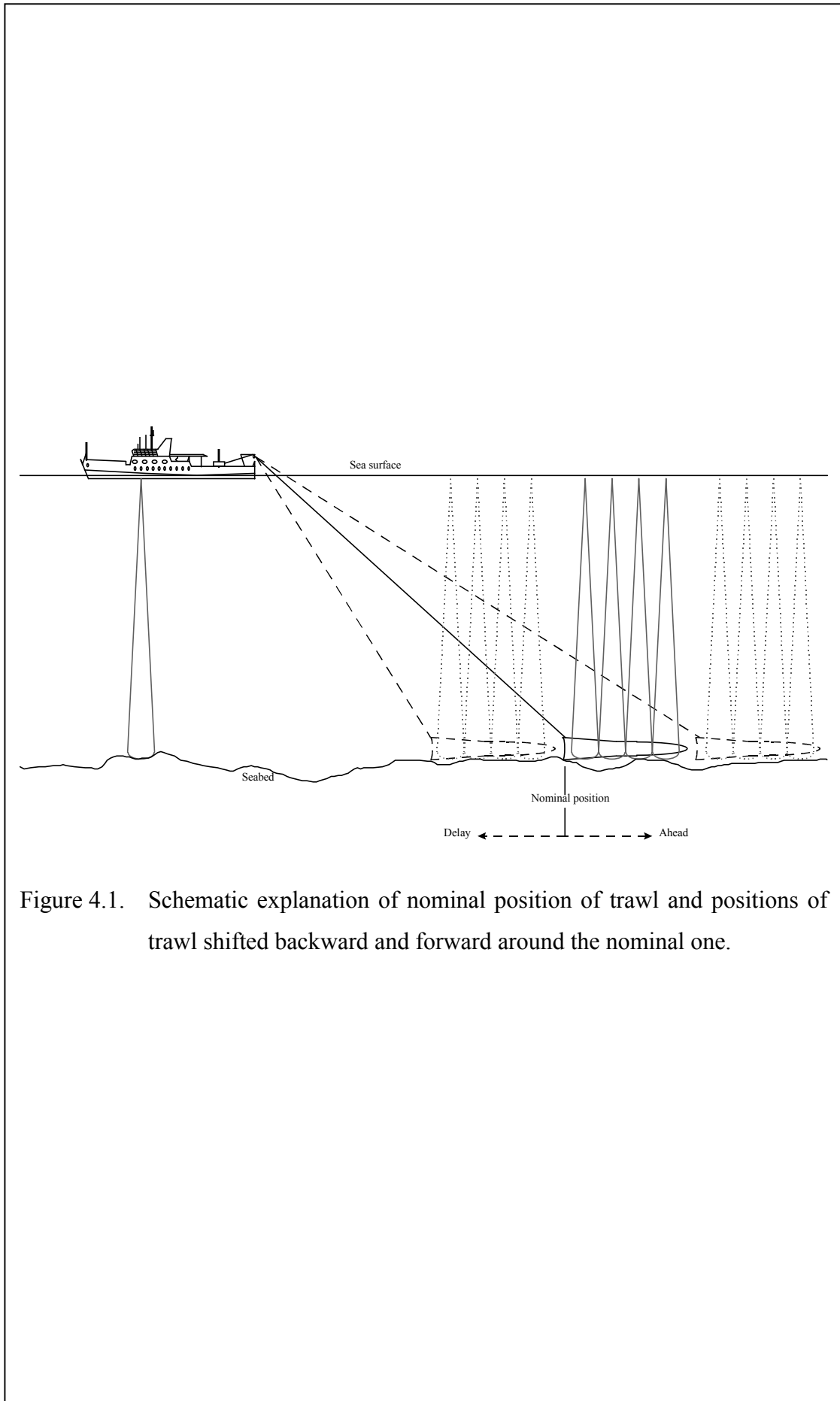


Figure 4.1. Schematic explanation of nominal position of trawl and positions of trawl shifted backward and forward around the nominal one.

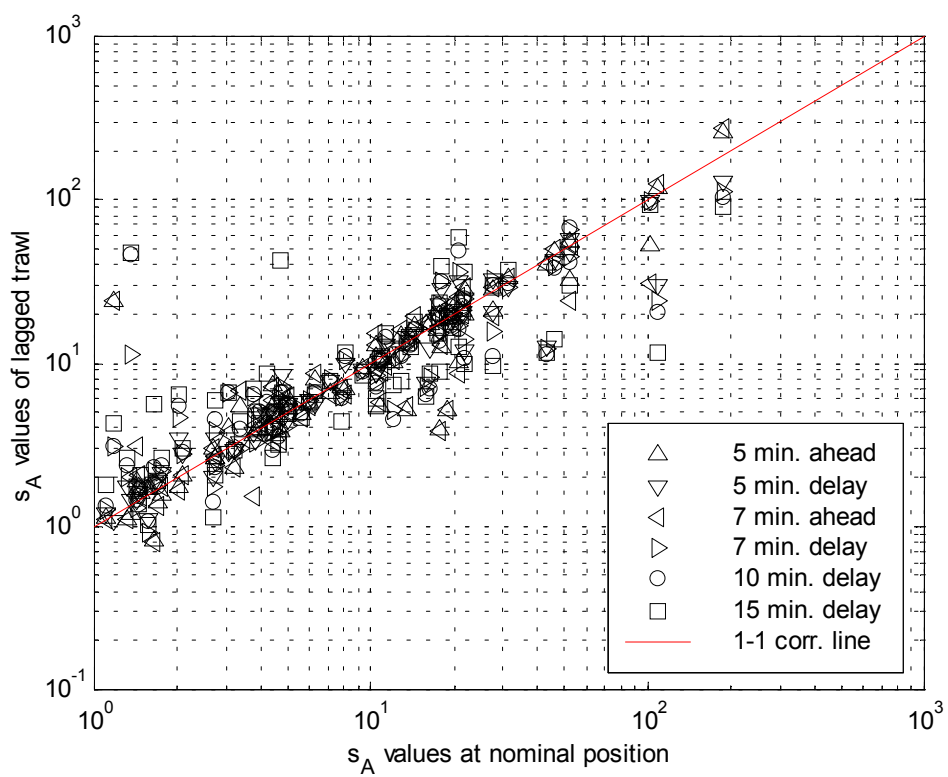


Figure 4.2. Comparisons between the acoustic  $s_A$  estimates at the locations provided from the vessel's logbook (on x-axis) and those from several lagged locations of trawls (on y-axis).

coefficients between the acoustic  $s_A$  estimates in the nominal position and those in the 6 different lagged positions are shown in Table 4.2. All the correlation coefficients given in Table 4.2 are significant at the 0.01 level ( $\alpha=0.01$ ). It is evident that the correlation coefficient between the acoustic  $s_A$  estimates in the nominal position and those in the lagged position  $r_{s_A(0)s_A(n)}$  decreases as the lagged position increases, so does the correlation coefficient of the acoustic  $s_A$  estimates between any pairs of the lagged positions. A summary of the statistical test to assess the difference of the correlation coefficients given in Table 4.2 is presented in Table 4.3. The overall conclusion drawn from Table 4.3 is that the difference of the correlation coefficients between the acoustic  $s_A$  estimates in the nominal position and those in the lagged positions up to 7 minutes ahead and 7 minutes delay is not significant at  $\alpha=0.01$ . This implies that the acoustic  $s_A$  estimates in the nominal position and those in the lagged positions up to 7 minutes ahead and 7 minutes delay are similar and those in lagged positions greater than 10 minutes delay are significantly different at  $\alpha=0.01$ .

The  $s_A$  estimates from acoustics and nets for all trawl stations available are shown in Figure 4.3. The variation between the two  $s_A$  estimates is up to one order of magnitude. It is evident that the nets provide on average slightly larger  $s_A$  estimates than the acoustics do. The slope  $b$  of the regression line (through the origin) fitted to data is  $0.905 \pm 0.076$  and is significant at  $\alpha=0.01$  ( $p \ll 0.01$ ). The coefficient of determination is significantly high ( $r^2 = 0.662$ ). Since the range of variation between the two  $s_A$  estimates was substantial, the two-tailed paired-sample  $t$ -test was used to assess the difference between means of the two estimates. The test suggested that the means of the two  $s_A$  estimates were not significantly different at  $\alpha=0.01$  ( $p=0.0194$ ). Sponges were often present in a large proportion and sometimes made up more than

Table 4.2. The correlation coefficients between the acoustic  $s_A$  estimates in the nominal position  $s_{A(0)}$  and those in the 6 different lagged positions  $s_{A(n)}$  ( $n$  denotes the lagged position) and between those within different lagged positions (significant at  $\alpha=0.01$ ).

Correlation coefficient, r						
	$s_{A(0)}$	$s_{A(5'ahead)}$	$s_{A(5'delay)}$	$s_{A(7'ahead)}$	$s_{A(7'delay)}$	$s_{A(10'delay)}$
$s_{A(5'ahead)}$	0.948					
$s_{A(5'delay)}$	0.919	0.815				
$s_{A(7'ahead)}$	0.920	0.996	0.771			
$s_{A(7'delay)}$	0.883	0.760	0.990	0.711		
$s_{A(10'delay)}$	0.819	0.691	0.941	0.642	0.971	
$s_{A(15'delay)}$	0.714	0.581	0.857	0.532	0.894	0.941

Table 4.3. A summary of the statistical test assessing the difference of the correlation coefficient between the acoustic  $s_A$  estimates in the nominal position  $s_{A(0)}$  and those in the 6 different lagged positions  $s_{A(n)}$  ( $r_{s_{A(0)}s_{A(n)}}$  where  $n$  denotes the lagged position) at  $\alpha=0.01$ .

	$r_{s_{A(0)}s_{A(5'ahead)}}$	$r_{s_{A(0)}s_{A(7'ahead)}}$	$r_{s_{A(0)}s_{A(5'delay)}}$	$r_{s_{A(0)}s_{A(7'delay)}}$	$r_{s_{A(0)}s_{A(10'delay)}}$
$r_{s_{A(0)}s_{A(7'ahead)}}$	$H_0 p > \alpha$ (p=0.1911)	$H_0 p > \alpha$ (p=0.1784)	$H_0 p > \alpha$ (p=0.0131)	$H_A p < \alpha$ (p=0.0001)	$H_A p < \alpha$ (p << 0.001)
$r_{s_{A(0)}s_{A(5'delay)}}$		$H_0 p > \alpha$ (p=0.9695)	$H_0 p > \alpha$ (p=0.2406)	$H_0 p > \alpha$ (p=0.0106)	$H_A p < \alpha$ (p << 0.001)
$r_{s_{A(0)}s_{A(7'delay)}}$			$H_0 p > \alpha$ (p=0.2562)	$H_0 p > \alpha$ (p=0.0118)	$H_A p < \alpha$ (p=0.0001)
$r_{s_{A(0)}s_{A(10'delay)}}$				$H_0 p > \alpha$ (p=0.1668)	$H_A p < \alpha$ (p=0.0037)
$r_{s_{A(0)}s_{A(15'delay)}}$					$H_0 p > \alpha$ (p=0.1289)

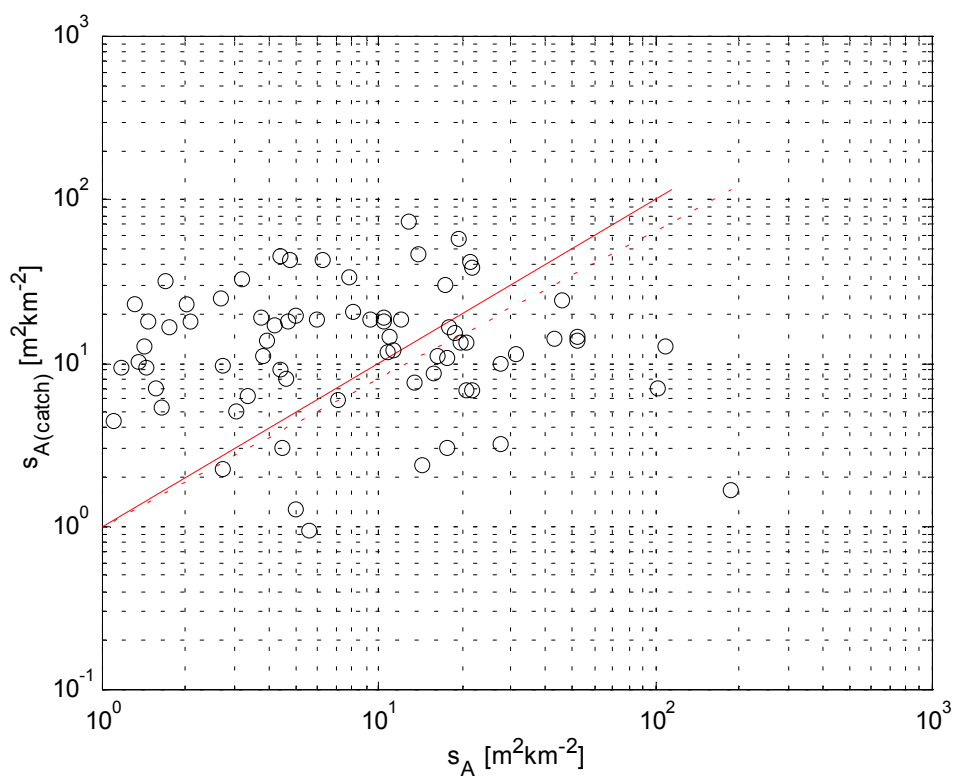


Figure 4.3. Scatterplot of acoustic  $s_A$  estimates (abscissa) and net  $s_A$  estimates (ordinate) between 0.5 to 2.5 m above the seabed in 71 trawl stations. — is the 1-1 correspondence line; - - is the regression line through the origin with the slope  $b$  of  $0.905 \pm 0.076$  ( $p \ll 0.01$ ) and coefficient of determination  $r^2=0.662$ .

50% of the catch. Attention has been given to the significance of the variation in  $s_A$  estimates due to the presence of sponges especially in a large proportion compared to other catches. A subsample comparison between acoustics and nets was prepared, where sponges accounted for 25% or less by weight of the catch. The result, however, gave no significant reduction in the range of results when both  $s_A$  estimates were compared one to another. The squared difference between both  $s_A$  estimates of the complete set as given in Figure 4.3 and that of the subsample one just mentioned above was prepared. There was a reduction in the mean of the squared difference after the subsample comparison. The  $t$ -test to examine the difference of variances of the two data sets revealed though that there was not a significant reduction in the variance at  $\alpha=0.01$  ( $p=0.551$ ) after the subsample comparison.

#### **4.4. Acoustically equivalent net estimates**

Species caught in each trawl stations were grouped according to the swimbladder type they belonged to. Almost all of the trawl station results comprised species of 3 different types of swimbladder namely physostomous, bladderless and squid-like types. Only in trawl station 5 were physoclistous species present. These species belonged to the family Clupeidae. Species of this group have lower target strength than physostomous species of the same length as shown in equation (3.9).

The proportions of species with and without length-frequency measurements available from the survey are shown in Table 4.4. In trawl stations 54 and 96, there were not any species with length-frequency measurements. In more than a half of the trawl stations, species with length-frequency measurements made less than 50% by weight of the total catch. Only 20 of the trawl stations were these species more than 50% of the total catch. Of 20, 5 trawl stations comprised a single majority (>75% by

weight of the total catch) of these species. The highest single majority of 85% occurred at trawl station 27 comprising 7 different species of this sort.

The total number of species caught and the number of species with length-frequency measurements in all trawl stations are also given in Table 4.4. Whereas trawl station 105 yielding 75 different species was the most diverse result, trawl stations 24 and 43 yielding only 13 different species provided the least diversity. Although trawl station 105 caught a total of 75 different species, it caught only 11 species of those where length-frequency measurements were conducted, whereas trawl station 89 having a total of 65 different species caught 13 species of this sort.

The estimated  $s_A$  values of selected species with and without length-frequency measurements, the mean length of species of the first sort, the estimated length of species of the second sort and the estimated number of species of the second sort are given in Table 4.5. Because the total number of species caught was more than 200, only a few selected species are noted in Table 4.5. It is shown in Table 4.5 that the estimated mean length of *Pentaprion longimanus*, *Leiognathus bindus*, *Dipterygonotus balteatus* and *Herklotsichthys koningsbergeri* is less than 11 cm. The predicted number of these species is respectively 2194, 5008, 20 and 1665. The predicted number of these species seems reasonable due to the fact that they are truly small in nature. The maximum length (in Total Length (TL)) of these species is 15, 14.8, 14 and 13 cm, respectively (Froese and Pauly, 1998). While the method used to predict the number and the mean length of species without length-frequency measurements seems workable for these species, it is not quite so for others such as *Nemipterus bathybius*, *Pterocaesio chrysozona* and *Sardinella gibbosa*. Although



Table 4.4. Proportion (% by weight) of species with and without length-frequency measurements, the number of species with length-frequency measurements, the number of total species and the number of different types of swimbladder present at 72 trawls. \*Station excluded from analysis.

Trawl station	Species without L-F available [%]	Species with L-F available [%]	Number of species with L-F available	Total Species	Number of different bladder type	Trawl Station	Species without L-F available [%]	Species with L-F available [%]	Number of species with L-F available	Total Species	Number of different bladder type
8	99.9	0.1	1	31	3	68	67.6	32.4	7	44	3
9	75.2	24.8	2	22	3	70	32.3	67.7	4	28	3
10	69.9	30.1	2	19	3	71	62.8	37.2	6	29	3
11	99.5	0.5	1	17	3	72	97.5	2.5	2	30	3
12	80.4	19.6	8	35	3	74	63.9	36.1	7	61	3
13	66.6	33.4	9	38	3	75	75	25	3	31	3
14	47.7	52.3	5	40	3	76	90.6	9.4	2	35	3
15	17.6	82.4	1	19	2	77	70.3	29.7	8	45	3
17	51.8	48.2	10	54	3	78	62.5	37.5	9	57	3
23	99.9	0.1	1	29	3	79	88.3	11.7	8	56	3
24	91.1	8.9	1	13	3	80	81.3	18.7	5	39	3
26	63	37	10	46	3	81	44.2	55.8	8	48	3
27	15	85	7	37	3	82	39.4	60.6	8	41	3
28	36.2	63.8	4	23	3	83	26.2	73.8	6	36	3
29	20.7	79.3	11	46	3	84	67.2	32.8	5	41	3
30	31.3	68.7	12	44	3	86	88.1	11.9	5	42	3
31	23	77	4	18	3	87	69.2	30.8	8	51	3
33	31	69	4	38	3	88	65.7	34.3	4	43	3
34	49.4	50.6	11	56	3	89	39.2	60.8	13	65	3
35	99.7	0.3	3	32	3	96	100	0	0	28	3
36	95.6	4.4	2	24	3	98	98.5	1.5	1	26	3
37	99.6	0.4	1	24	3	100	86.9	13.1	2	23	3
40	61.8	38.2	7	42	3	101	34.1	65.9	10	46	3
41	99.8	0.2	1	27	3	102	21.6	78.4	10	54	3
42	99.4	0.6	2	19	3	103	70.7	29.3	10	63	3
43	90.2	9.8	1	13	3	104	55.1	44.9	4	37	3
54	100	0	0	26	3	105	81.5	18.5	11	75	4
55	91	9	3	37	3	106	89.8	10.2	10	55	4
56*	88.6	11.4	6	45	3	107	93.4	6.6	6	51	4
57	75.1	24.9	4	27	3	108	88.6	11.4	5	41	4
61	67.5	32.5	4	27	3	109	79.3	20.7	6	49	3
62	95.1	4.9	3	38	3	110	80.6	19.4	9	73	4
64	26.3	73.7	9	48	3	111	47.1	52.9	8	47	3
65	25.3	74.7	8	32	3	112	47	53	4	24	3
66	59.1	40.9	12	57	3	113	86.3	13.7	7	42	3
67	78.2	21.8	5	26	3	114	68.7	31.3	2	25	3

Table 4.5.  $s_A$  estimate, mean length, bladder type, proportion (% by weight) and weight of selected species with and without length-frequency measurements available. Length is in TL (Total Length) for all fish and in ML (Mantle Length) for squid and cuttlefish.

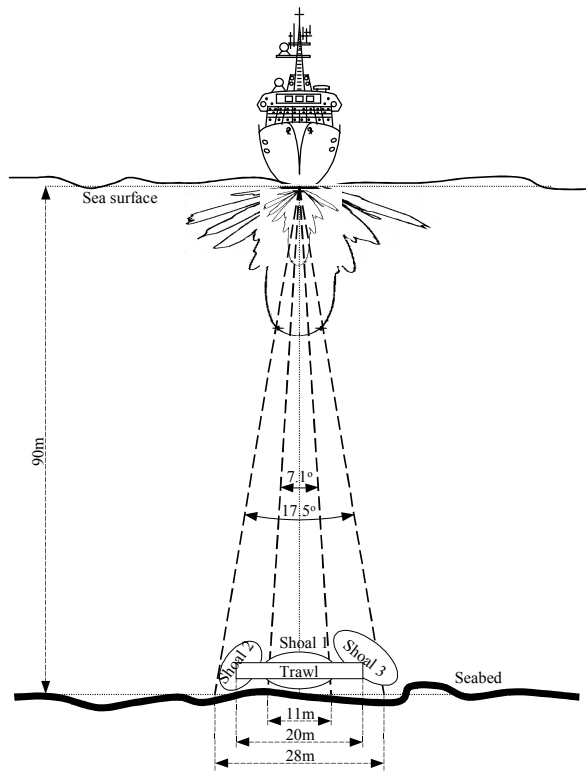
Trawl station	Species	Weight [kg]	Proportion [%]	Bladder Type	Number of Target	Mean Length [cm]	$s_A$ [ $m^2 km^{-2}$ ]
<u>Species without L-F measurements:</u>							
12	<i>Albula neoguinaica</i>	79.3	51.3	Physoclist	163	33.9	4.2220
33	Squid	19.2	9.8	Squid-like	954	9.9	0.3608
35	<i>Pentaprion longimanus</i>	44.5	18.5	Physoclist	2194	10.0	5.3974
	<i>Nemipterus bathybius</i>	48.3	20	Physoclist	4853	9.0	9.7521
37	<i>Leiognathus bindus</i>	50.4	58.7	Physoclist	5008	8.6	10.2076
41	<i>Ariomma indica</i>	2.7	1.2	Physoclist	56	15.3	0.2875
51	<i>Epinephelus rivulatus</i>	1.28	0.7	Physoclist	5	25.3	0.1187
66	<i>Abalistes stellaris</i>	34.6	13.6	Physoclist	335	14.7	2.3935
77	<i>Upeneus moluccensis</i>	106.5	34.6	Physoclist	2637	14.0	15.9497
78	Cuttlefish	15.2	8.5	Squid-like	756	6.7	0.1157
	<i>Hemigaleus microstoma</i>	0.72	0.4	Bladderlessness	1	43.6	0.0007
79	<i>Epinephelus maculatus</i>	3.6	0.7	Physoclist	26	19.9	0.2153
81	<i>Dipterygonotus balteatus</i>	0.2	0.1	Physoclist	20	10.5	0.0457
	<i>Pterocaesio chrysozona</i>	19.6	11.1	Physoclist	1960	10.8	4.8174
105	<i>Herklotsichthys koningsbergi</i>	16.5	6.8	Physostome	1665	10.1	1.5784
108	<i>Sardinella gibbosa</i>	0.67	1.1	Physostome	66	10.6	0.0669
110	<i>Rastrelliger kanagurta</i>	14.7	4.4	Bladderlessness	74	25.9	0.0196
<u>Species with L-F measurements:</u>							
12	<i>Epinephelus areolatus</i>	3.3	2.1	Physoclist	13	27.4	0.2230
13	<i>Epinephelus multinotatus</i>	4.5	1.5	Physoclist	11	66.0	1.3449
27	<i>Lutjanus vitta</i>	46.1	20.4	Physoclist	208	24.2	3.5029
29	<i>Lethrinus</i> sp.	639	66.4	Physoclist	310	28.3	5.2343
	<i>Lethrinus nebulosus</i>	36.3	3.8	Physoclist	30	43.4	1.1957
65	<i>Parupeneus heptacanthus</i>	19.5	17.2	Physoclist	102	23.4	1.9210
74	<i>Lutjanus malabaricus</i>	349	34.1	Physoclist	80	51.7	4.9046
77	<i>Nemipterus celebicus</i>	17.6	5.7	Physoclist	416	14.1	2.6246
83	<i>Saurida undosquamis</i>	261.2	58.9	Physoclist	271	37.2	8.6820
89	<i>Lutjanus sebae</i>	72.8	27.9	Physoclist	26	51.0	1.7275
	<i>Diagramma labiosum</i>	34.7	13.3	Physoclist	19	50.5	1.1888
100	<i>Saurida</i> sp.2	8	11.3	Physoclist	273	13.5	1.3292
102	<i>Nemipterus furcosus</i>	127.7	40.5	Physoclist	611	26.1	13.9266
111	<i>Lethrinus genivittatus</i>	84.2	29.6	Physoclist	528	17.8	4.2741

these species are not small in nature and are bigger than those mentioned previously, the mean length of these species predicted by this method is underestimated and the predicted number is overestimated. The maximum length (in TL) of these species is 30, 21 and 20 cm, respectively (Froese and Pauly, 1998).

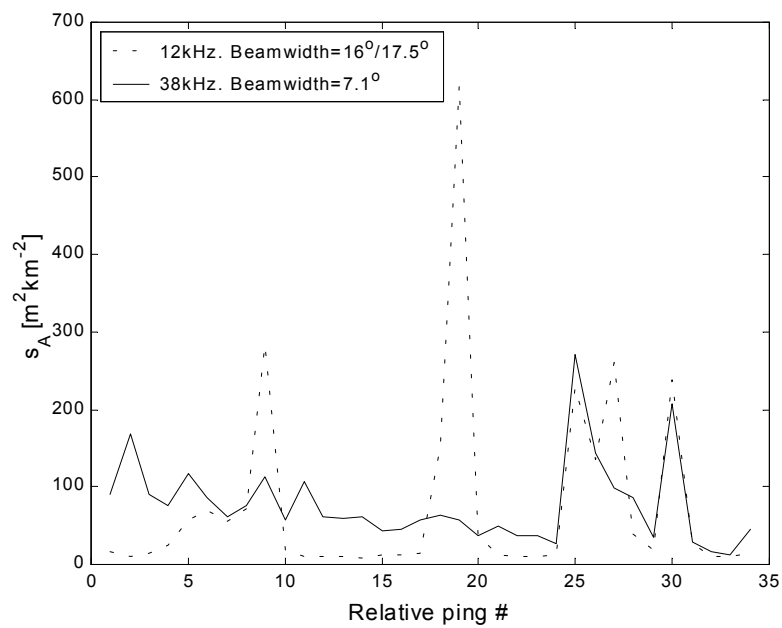
#### **4.5. Discussion**

The results indicate a substantial range of variation between acoustic estimates and net estimates. For one particular net estimate, a range of one order of magnitude is found in the corresponding acoustic estimate. The result from the sub-sampled data shows no significant reduction in the variation range of results when the two  $s_A$  estimates are compared one to another. The nets produce slightly larger  $s_A$  estimates than the acoustics do.

The substantial range of variation between the two  $s_A$  estimates may reflect the uncertainties in the parameters, notably the target strength used to determine the  $s_A$  estimates of the nets. Other possibilities influencing the differences between the two estimates could be due to the efficiency of the nets used in the survey and the influence of the detailed patchiness of the target population. Figure 4.4 shows for a depth of 90 m the cross-track acoustic coverage arising from the differing beamwidth of the transducer units and the width of the mouth opening of the trawl. If only shoal 1 shown in Figure 4.4(a) is present, the  $s_A$  estimates at 12 and 38 kHz, and from nets are comparable. In the presence of shoals 2 and 3 in Figure 4.4(a) however, the  $s_A$  estimates at 12 kHz are expected to be higher than those at 38 kHz (Figure 4.4(b)) and those from nets since the sampling scale of the 12 kHz transducer unit is higher than that of 38 kHz transducer unit and nets. Similarly, the  $s_A$  estimates from nets become higher than those at 38 kHz. Despite the fact that the difference between  $s_A$  estimates



(a) Schematic diagram of beamwidth effect



(b) Time series of  $s_A$  estimates for 12 and 38 kHz at station 79

Figure 4.4. Beamwidth effect on a highly patchy distribution of fish. (a) Schematic diagram of two different beamwidths ( $17.5^\circ$  for 12 kHz and  $7.1^\circ$  for 38 kHz) sensing a highly patchy distribution of fish shown as Shoals 1, 2 and 3. (b) Time series of  $s_A$  estimates for 12 kHz (---) and 38 kHz (—) at station 79.

from nets and acoustics is substantial, there is an overall linear trend. The relationship between the two  $s_A$  estimates for single dominant species is consistent only for *Nemipterus furcosus* (Figure 4.5). For *Saurida undosquamis* on the other hand, net  $s_A$  estimates decreases with an increase in acoustic  $s_A$  estimates. Elliot and Kloser (1994) demonstrated a consistently positive slope between acoustic and net estimates for sharks, morids and macrourids but a negative slope for eels and slickheads. Perhaps, this might point out different behaviour of different species.

Everson (1987) found the discrepancy between acoustic and net estimates of krill. The discrepancy was found to be twenty fold. Similar to the result of this study, his result revealed that acoustic techniques provided less biomass estimates than net techniques. An underestimate of acoustic estimates resulted from an overestimate of target strength used in his study to size relationships. He, however, encountered more net avoidance by day than at night similar to Everson and Bone (1986). In their study on demersal fish reaction to trawling noise, Ona and Godo (1990) suggested a pre-vessel avoidance by demersal fish during a trawl deployment was responsible for the lower acoustic estimates. The demersal fish escaped due to change of sound pressure when trawl deployment first occurred but then moved into the path of the trawl and got caught in the trawl later after the vessel had passed them.

In contrast, Pauly *et al.* (1997) in their study on acoustically and net derived krill biomass found that acoustic estimates were constantly larger by 3 orders of magnitude than net estimates. Sub-sampling approach applied in their study to trawls containing 90% krill by weight reduced the spread of variation by only one order of magnitude. Two orders of magnitude then remained in the spread of variation in their results. Similarly, Everson and Bone (1986) in their study on effectiveness of the

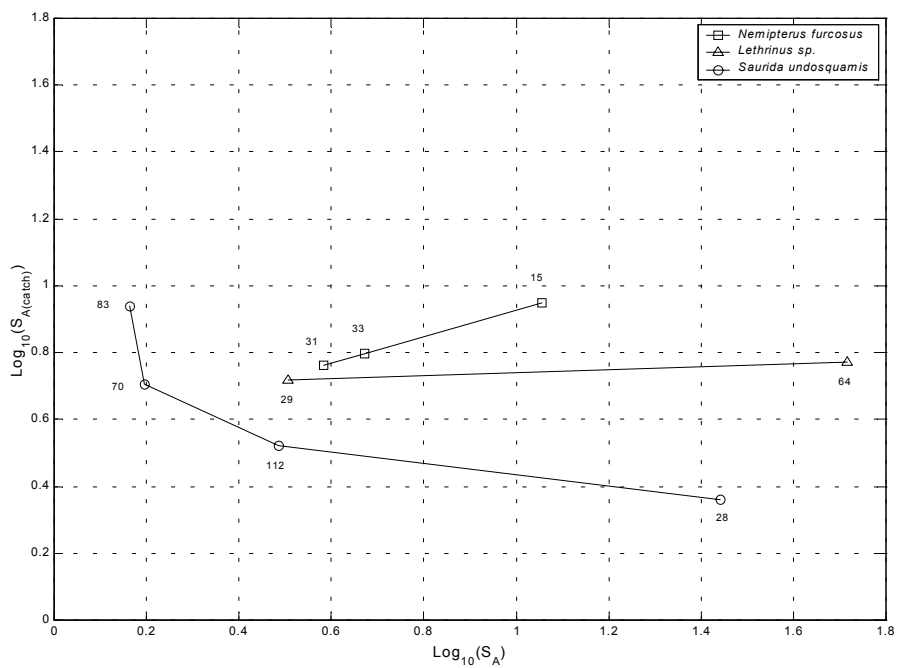


Figure 4.5. Scatterplot of log transformed acoustic  $s_A$  estimates and log transformed net  $s_A$  estimates of single dominant species. Lines are intended only to guide the eye. Numbers are trawl station numbers in Table 4.4.

rectangular midwater trawl for sampling krill assemblage found a discrepancy between acoustically and net derived biomass estimates by day but not at night. Results from their study for the daytime showed that nets caught fewer krill than they would have been predicted by the density estimates produced by the echosounder. Likewise, Koslow *et al* (1997) found that the acoustic estimate of pelagic fish biomass was higher than the net estimate by 7-fold. The discrepancy between the two estimates was due to net avoidance by larger, more active species, net extrusion by small species, and choice of inappropriate *TS*. However, the acoustic estimate was 3-fold less than the estimate derived from *in-situ TS* distributions rather than net data. The difference arose from higher representations of *TS* of bladderless fish and squid.

In this study however, slightly higher net estimates were expected due to the exclusion of the first 0.5 m depth layer off the bottom from the integration. The exclusion of this layer was made on purpose to avoid the inclusion of unwanted bottom returns in the integration process. Ona and Mitson (1996) developed an extrapolation approach to compensate this "lost zone" (known as dead zone). The principle of this approach is to estimate the effective lost height of the dead zone. The  $s_A$  lost due to the dead zone is then compensated by multiplying the effective lost height by the average  $s_A$  per metre of the immediate layer next to it. The effective lost height given by Aglen (1996) is

$$H_{dz} = \text{offset} + \frac{\tau}{4} + 0.004R_o \quad [\text{m}] \quad (4.1)$$

where offset is a bottom offset of 0.5 m,  $\tau$  is the pulse length of 1 ms and  $R_o$  is the water depth. Applying this approach to acoustic data in this study leads to an increase of acoustic  $s_A$  estimates by 56% on average (Figure 4.6). The similar statistical tests applied to the data set of Figure 4.3 were also introduced to the acoustic-dead-zone

(ADZ) corrected and net derived  $s_A$  estimates given in Figure 4.6. The slope  $b$  of the regression line (through the origin) fitted to data is  $0.905 \pm 0.076$  and is significant at  $\alpha=0.01$  ( $p \ll 0.01$ ). The significance test of the correlation coefficient suggested that the correlation coefficient was significantly different from zero at  $\alpha=0.01$  ( $p \ll 0.01$ ). The squared difference between both  $s_A$  estimates prior to the introduction of the ADZ compensation as given in Figure 4.3 and that after the introduction of the ADZ compensation as given in Figure 4.6 was prepared. The scatterplot of the squared difference between both  $s_A$  estimates prior to and after the introduction of the ADZ compensation into the acoustic  $s_A$  estimates is shown in Figure 4.7. It is evident that there is a strong linear relationship between the squared difference of both  $s_A$  estimates prior to and after the introduction of the ADZ compensation indicated by a high coefficient of determination ( $r^2=0.979$ ). This suggests that the variances of the two pair data sets are significantly the same. This is supported by the  $t$ -test showing that there is no significant difference in variance at  $\alpha=0.01$  ( $p=0.505$ ) after the introduction of the ADZ compensation.

Aglen (1996) on his study on the relationship between acoustic and swept-area estimates of fish showed a significant correlation between the two estimates. The agreement between the two estimates was established however because he tried to fit acoustic estimates of different height of the integrated layer above the seabed to net estimates. The correlation was found to best match net estimates at an integrated acoustic layer from 30 to 100 m above the seabed for haddock, saithe, cod and redfish and at the height of the trawl headline of 4 m for Norway pout. Unlike, Aglen (1996), Elliot and Kloser (1994) obtained a poor correlation between acoustic and net estimates of demersal fish. The poor correlation resulted partially from high



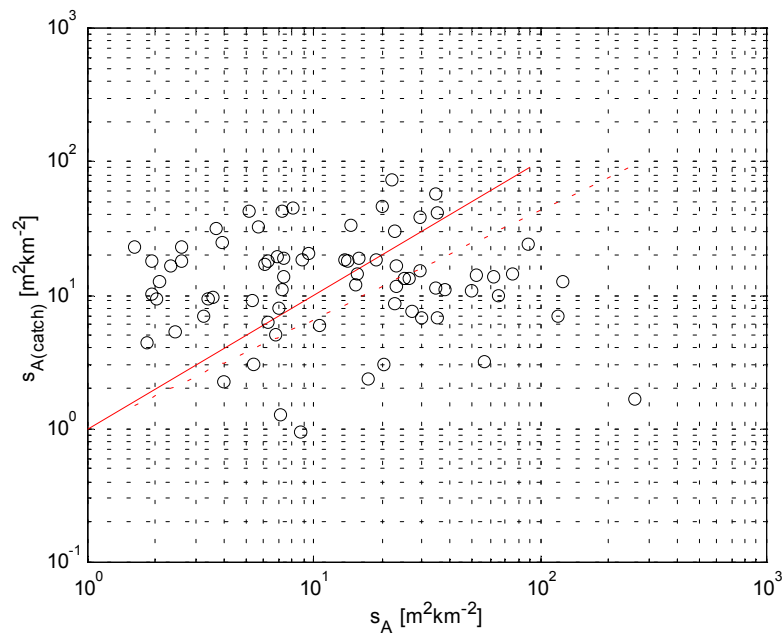


Figure 4.6. Scatterplot of acoustic  $s_A$  estimates with a correction for the dead zone (abscissa) and net  $s_A$  estimates (ordinate) 71 trawl stations. — is the 1-1 correspondence line; -- is the regression line through the origin with the slope of  $0.817 \pm 0.060$  ( $p \ll 0.01$ ) and coefficient of determination  $r^2 = 0.720$ .

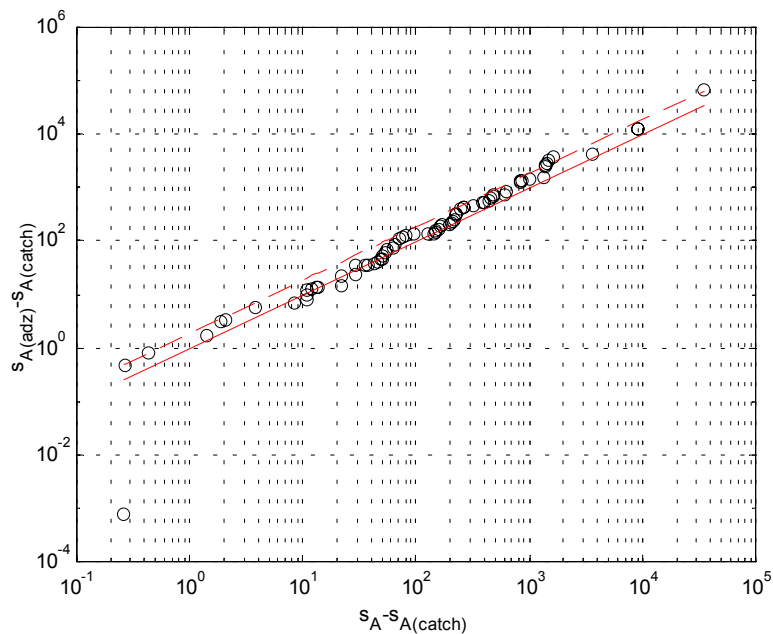


Figure 4.7. Scatterplot of the squared difference between both  $s_A$  estimates prior to and after the introduction of the ADZ compensation. — is the 1-1 correspondence line; -- is the regression line with the slope of  $1.858 \pm 0.032$  ( $p \ll 0.01$ ) and coefficient of determination  $r^2 = 0.979$ .

heterogeneity in the species composition and limitations on the acoustic system due mainly to the use of single beam system which led to a lack of any *TS* information.

The other possible reason of the discrepancy in the substantial range of variation between the two estimates in this study is the effectiveness of the method to predict the number and the mean length of species without length-frequency measurements, and hence the  $s_{A(\text{catch})}$  estimates. The method used here tends to underestimate the mean length of the fish. A slightly reduction of the variation between the two estimates is obtained when acoustic estimates are compared to the net estimates of species with length frequency measurements available (Figure 4.8). The variance though is not significantly different from that of the complete data set (Figure 4.3) at  $\alpha=0.01$  ( $p=0.139$ ) and that after the introduction of the ADZ compensation (Figure 4.6) at  $\alpha=0.01$  ( $p=0.254$ ).

#### **4.6. Conclusions**

One of the objectives of the comparison between the acoustically derived  $s_A$  estimates and the comparable estimates derived from demersal net was to explore the possibility of the inclusion of net derived  $s_A$  estimates when establishing the relationship between acoustically derived seabed types and fish groups. Despite the fact that there is an overall relationship between the two  $s_A$  estimates, the difference between them is substantial, which might be due to possibilities discussed above. This, therefore, leads to a conclusion that it is convenient to adopt only the acoustically derived  $s_A$  estimates rather than those from nets or both to form the so-called quasi acoustic population, which is discussed in section 5.7, in order to establish the relationship between seabed types and fish groups. In addition, it is

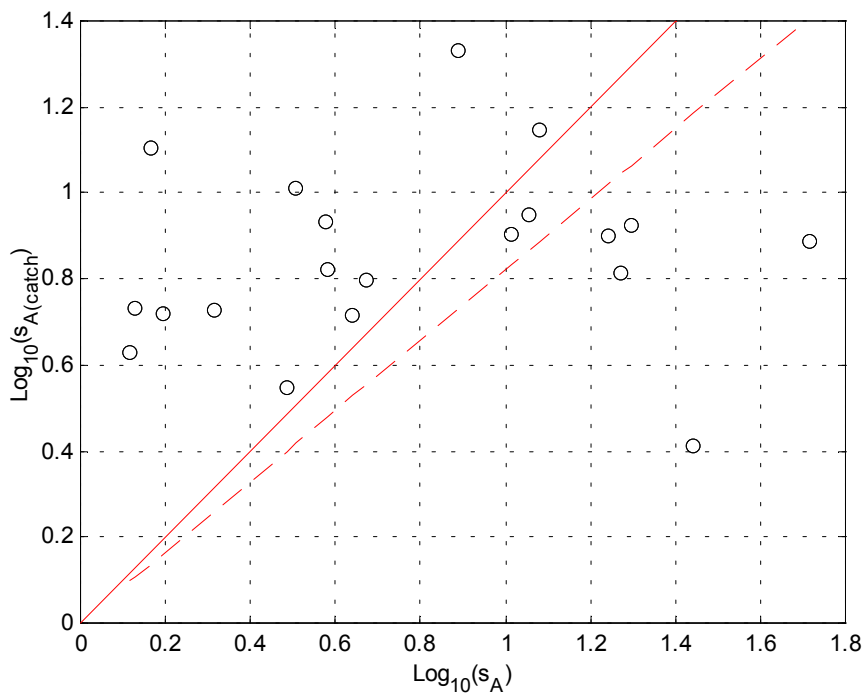


Figure 4.8. Scatterplot of log transformed acoustic  $s_A$  estimates and log transformed net  $s_A$  estimates of species with length frequency measurements available. — is the 1-1 correspondence line; -- is the regression line through the origin with the slope of  $0.822 \pm 0.119$  ( $p \ll 0.01$ ) and coefficient of determination  $r^2=0.701$ .

convenient to use the acoustically derived  $s_A$  estimates and acoustically derived roughness and hardness as they are synchronous.

Future work is therefore essential to investigate in detail whether or not the agreement between the acoustic estimates and net estimates can be improved. To achieve this, net data collections need to be revised. It is mandatory that the net data collection must include length-frequency measurements for all species or at least the most dominant species in order to minimise the uncertainties due to the number and size of species.

## Chapter 5

# Results of bottom classifications for North West Shelf and South East Fisheries regions

### 5.1. Characteristics of study areas

Information regarding characteristics of study areas provided in this section comes from previous studies conducted in the NWS and SEF regions. Figure 5.1a shows the location of the two regions within the Australian continent together with the bathymetry from 50 to 500 m around the continent. The former is located in the north-west continental shelf labelled as NWS in Figure 5.1a and the latter is situated in the south-east continental shelf given as SEF in Figure 5.1a. The bathymetry of the NWS and SEF regions in which the two study areas are located is shown in Figures 5.1b and 5.1c, respectively. The information on seabed types in particular comes from Jones (1973) for the NWS region, and Jones and Davies (1983) for the SEF region. The information of seabed features in both studies just mentioned was based on grab samples collected on a 10 nmi grid. Seabed types derived from this study in both regions are discussed in sections 5.3 to 5.6.

The NWS study area extends between longitudes 114° E and 123° E and between latitudes 14°S and 21°S to a depth of 200 metres. It varies in width from about 220 kilometres south of the Rowley Shoals to less than 15 kilometres off North West Cape in the south-west (Jones, 1973). The continental margin is bounded by the

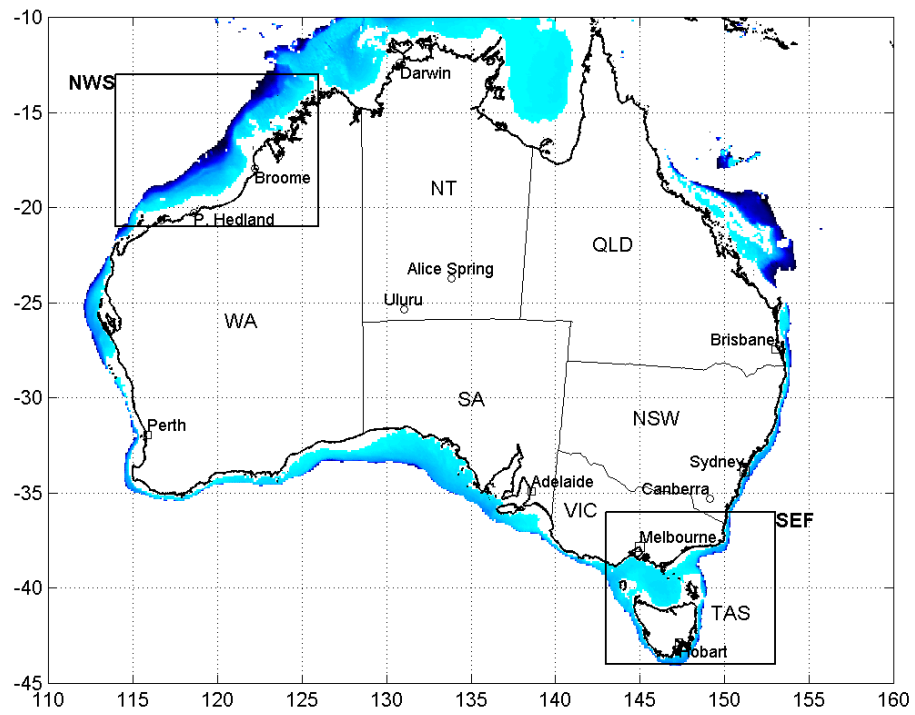


Figure 5.1a. Map of the Australian continent with the two continental shelves, north-west (NWS) and south-east (SEF), and the bathymetry around the continent from 50 to 500 m (light to dark blue).

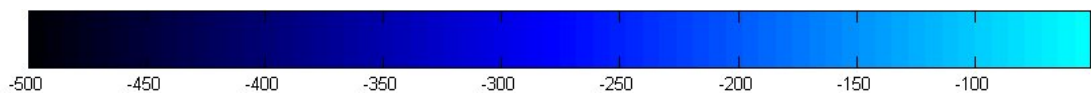
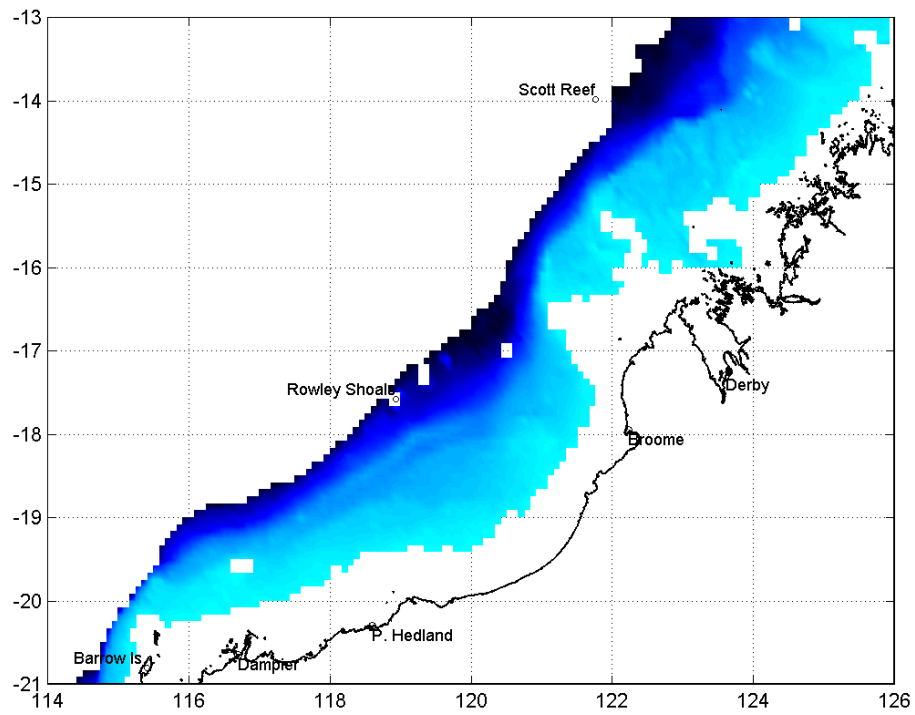


Figure 5.1b. Map of the north-west continental shelf of Australia together with the bathymetry around the continental shelf from 50 to 500 m (light to dark blue).

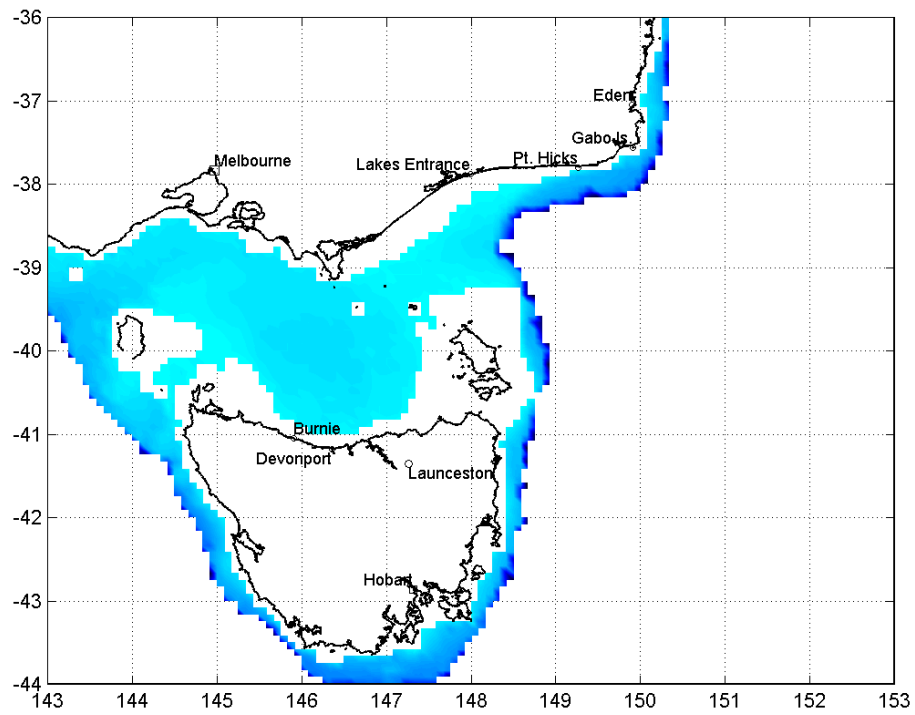


Figure 5.1c. Map of the south-east continental shelf of Australia together with the bathymetry around the continental shelf from 50 to 500 m (light to dark blue).



North Australian Basin and the floor of the basin that has an area of about 160 000 kilometre squares. It has been described as a true abyssal plain (Jones, 1973).

The area lies within the tropics (Wyrski, 1961) and has a monsoon climate, with a wet season during the north-west monsoon and a dry season during the south-east monsoon (Van Andel and Veevers, 1967). During the dry season, strong, gusty winds cause moderate and rough seas with the main swell coming from the south-east and during the wet season, the sea is usually smooth to moderate (Van Andel and Veevers, 1967).

The north-west continental shelf of Western Australia has a substrate consisting mostly of calcareous sands (McLoughlin and Young, 1985). Calcareous coral reef resides in restricted patches in depth less than 30 metres (Sainsbury, 1991). The continental shelf consists largely of an area of winnowing and sediment transport rather than deposition and over wide areas the flat surface is interrupted by minor irregularities that commonly consist of littoral features (Jones, 1973). The shelf is also characterised by large scale sand waves ranging from 4 to 10 metres high, small scale sand waves ranging from 1 to 2 metres high, and normal current ripples (Jones, 1973). Figure 5.2a taken from Jones (1973) shows minor relief of the continental shelf up to the 180-metre contour. Figure 5.2a shows that consistently irregular surfaces occur to the north in particular between 14° S and 16° S where the whole shelf seaward of the 40-m contour is rough and uneven. Although the irregular surface continues into the central part of the continental shelf southwest of latitude 16° S, smooth topography occurs dominantly near the outer edge of the continental shelf as well as inshore. Figure 5.2a was based on grab sample station spacing of 10 nmi (nautical miles) on a square grid. This station spacing, however, suffered

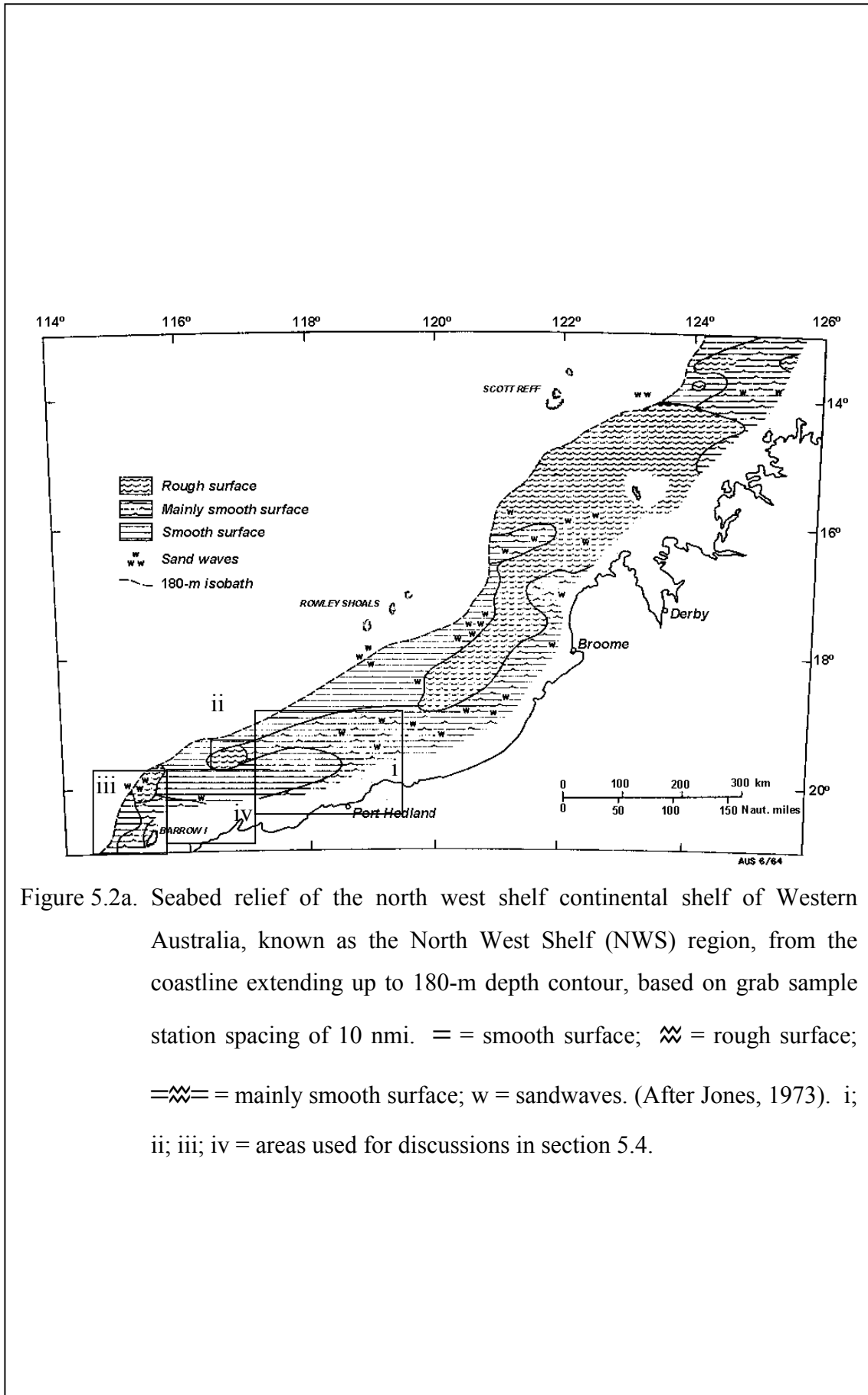


Figure 5.2a. Seabed relief of the north west shelf continental shelf of Western Australia, known as the North West Shelf (NWS) region, from the coastline extending up to 180-m depth contour, based on grab sample station spacing of 10 nmi. = = smooth surface; ≍ = rough surface; ≍≍ = mainly smooth surface; w = sandwaves. (After Jones, 1973). i; ii; iii; iv = areas used for discussions in section 5.4.

considerable distortion due to the largely unpredictable currents (Jones, 1973). In addition, position fixing was based mainly on celestial navigation (Jones, 1973), which necessarily depended on the sky visibility.

Oceanography is highly dynamic and complex. High current speeds and the vertical movement of isotherms through 50 metres are present in some periods and are attributed to strong semidiurnal tides and internal tide waves (Holloway, 1983, 1987). The El Nino Southern Oscillation plays an important role on the interannual variability of water temperature and sea level height (Pariwono *et al.*, 1986). Hydrological and productivity observations have shown some evidence of upwelling off the north-west Australian coast (Van Andel and Veevers, 1967; Holloway and Nye, 1985) that causes the water column to be well mixed between about May and October each year (Holloway and Nye, 1985). Although surface water circulation appears to favour upwelling when, in particular, south-east trades are active, these observations indicate that the upwelling present is minor (Jones, 1973). Nonetheless, it has contributed to a high biological productivity in the area (Tranter, 1962). This has made the north-west continental shelf one of the northern Australian tropical regions that support a diverse Indo-West Pacific fish fauna (Sainsbury *et al.*, 1997). This has also made the marine fish fauna of the north-west continental shelf probably richer in species than that of any other Australian region. For example, there exist over 600 species of only demersal fish (Sainsbury *et al.*, 1985). Coral reefs are well developed in a number of places on the north-west continental shelf, and patch and platform reefs are numerous on the inner part of the shelf (Bartlett, 1981). In the Dampier archipelago, a rich coral fauna of more than forty eight genera was found ranging from the most turbid inshore waters to relatively clear waters of the outer islands (Bartlett, 1981).

The SEF study area occupies an area between 36° and 39° S (Bax *et al.*, 1999). The continental shelf in this region extends to 170 to 200-m depth and varies in width from 25 kilometres in the north of the study area and 175 kilometres in the south. It lies within the temperate region and has become the Australia's most important fishery for domestic scalefish markets (Tilzey, 1994).

Riverine input is minimal. Upwelling was also observed in the south-east continental shelf (Edwards, 1990). The oceanography of the south-east continental shelf region is complex and causes intrusions of water from the continental slope onto the shelf during summer upwellings (Edwards, 1990; Bax *et al.*, 1999). The summer upwellings occur between mid January and the end of March because the wind blowing across the water produces surface currents, which move to the left of the wind direction in the Southern Hemisphere due to rotation of the earth (Edwards, 1990). This event however is intermittent and transitory, and occurs in the shallow water only (Edwards, 1990). Although one of the characteristics of this summer upwelling is a low level of nutrient enrichment, the effect on the biota at the lower end of the food chain is still reasonably strong because it occurs when both light and temperature are available for organisms to take advantage of the influx (Edwards, 1990).

Calcium carbonate is the dominant component of the surficial sediments that consist of recognisable skeletal debris mainly contributed by molluscs, bryozoans and foraminifera (Jones and Davies, 1983). Calcareous benthic foraminifera make up significant proportion of the shelf sediments in particular in eastern Bass Canyon where they commonly approach bryozoans and molluscs in abundance (Jones and Davies, 1983). The south-east continental shelf consists largely of an area of

winnowing and sediment transport rather than deposition (Jones and Davies, 1983). Coastal erosion and winnowing of the sea floor by currents and wave and swell action lift fine sediments into suspension in the near bottom water layers (Jones and Davies, 1983). Waves in the area average 1 to 3 metres high and penetrate to the seabed in 60-m depth (Jones and Davies, 1983).

Continental basement rock forms outcrops near the coast whereas coarse-grained sandstone dominated by quartz extends along the inner shelf and is consistent with sand bodies formed in a high energy coastal plain environment (Bax *et al.*, 1999). Fossiliferous limestones consisting largely of bivalve and bryozoan clasts are common on the outer shelf (Bax *et al.*, 1999). The shelf break is marked by structurally complex features formed by the necks of the Bass Canyon (Kloser *et al.*, 2001). The muddy sediments of central Bass Canyon occur in water depths ranging from 44 to 212 meters and are bounded by mainly terrigenous sand landwards and by mainly relict sand and gravel seawards where they are on the open shelf (Jones and Davies, 1983). Figure 5.2b taken from Jones and Davies (1983) shows the summary of sediment types in the south-east continental shelf. Figure 5.2b was based on samples (365 bottom grab samples) collected on a 10 nmi grid though stations in some areas of the east Tasmanian shelf were more closely spaced (Jones and Davies, 1983).

## **5.2. First principal component of roughness and hardness parameters**

Based on the constant angular integration interval and constant depth integration interval outlined in section 2.3 and acoustic data analysis for seabed parameters outlined in section 3.1.3.1, bottom roughness (E1) and bottom hardness (E2) indices for the three operated frequencies data were derived. Figures 5.3 and 5.4

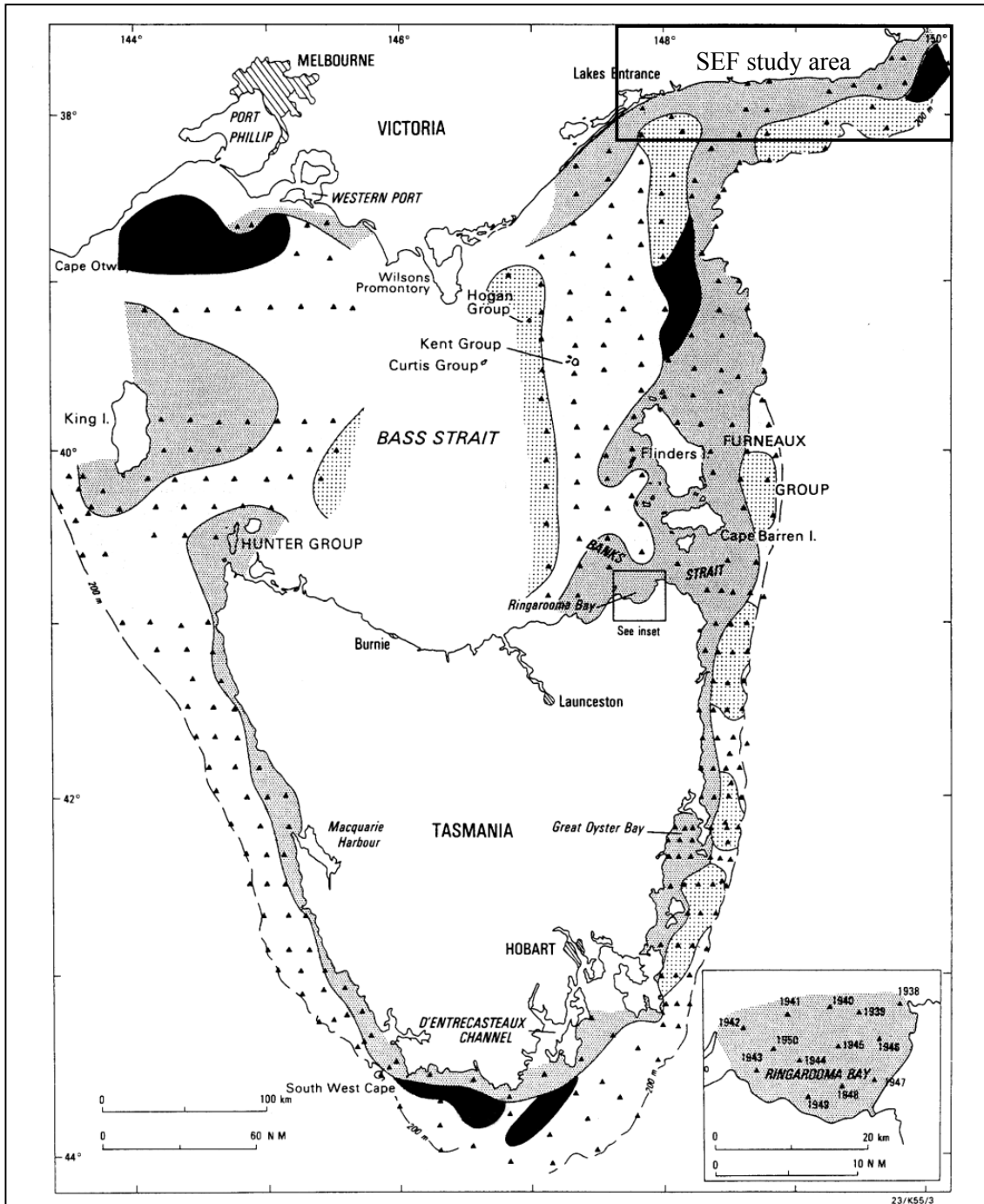






Figure 5.2b. Surface sediments of the south east continental shelf of Australia, known as the South East Fisheries (SEF) region, based on grab sample station spacing of 10 nmi.  = mainly quartz sands with variable amounts of modern shell debris;  = muddy quartzose and calcareous sediments;  = fine-grained shelly sands, slightly quartzose, with modern and relic components;  = bryozoan sands and gravels, mainly relict in origin. After Jones and Davies (1983).

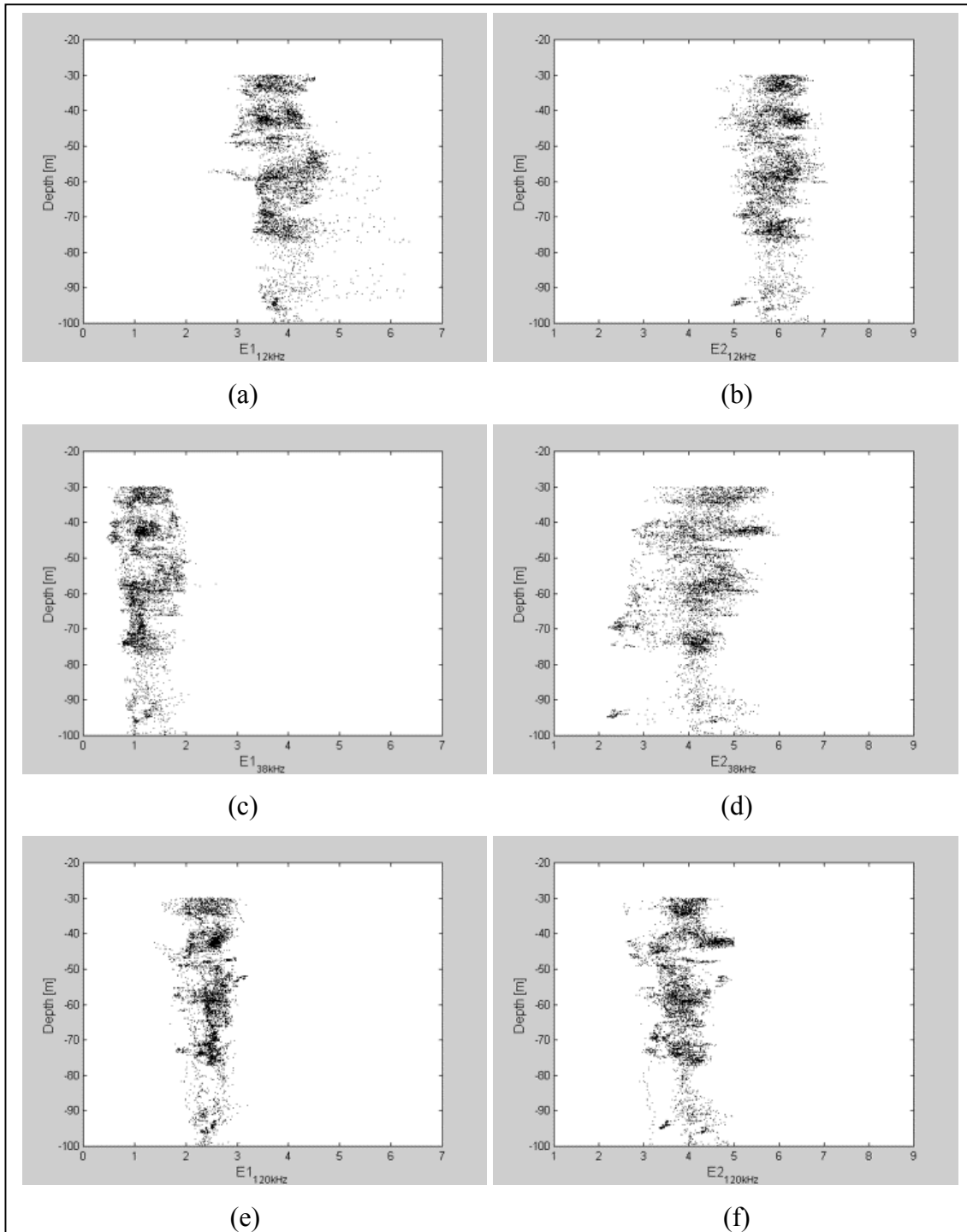


Figure 5.3. Scatterplot of acoustical seabed indices for the NWS study area. (a) roughness index vs depth at 12 kHz. (b) hardness index vs depth at 12 kHz. (c) roughness index vs depth at 38 kHz. (d) hardness index vs depth at 38 kHz. (e) roughness index vs depth at 120 kHz. (f) hardness index vs depth at 120 kHz.

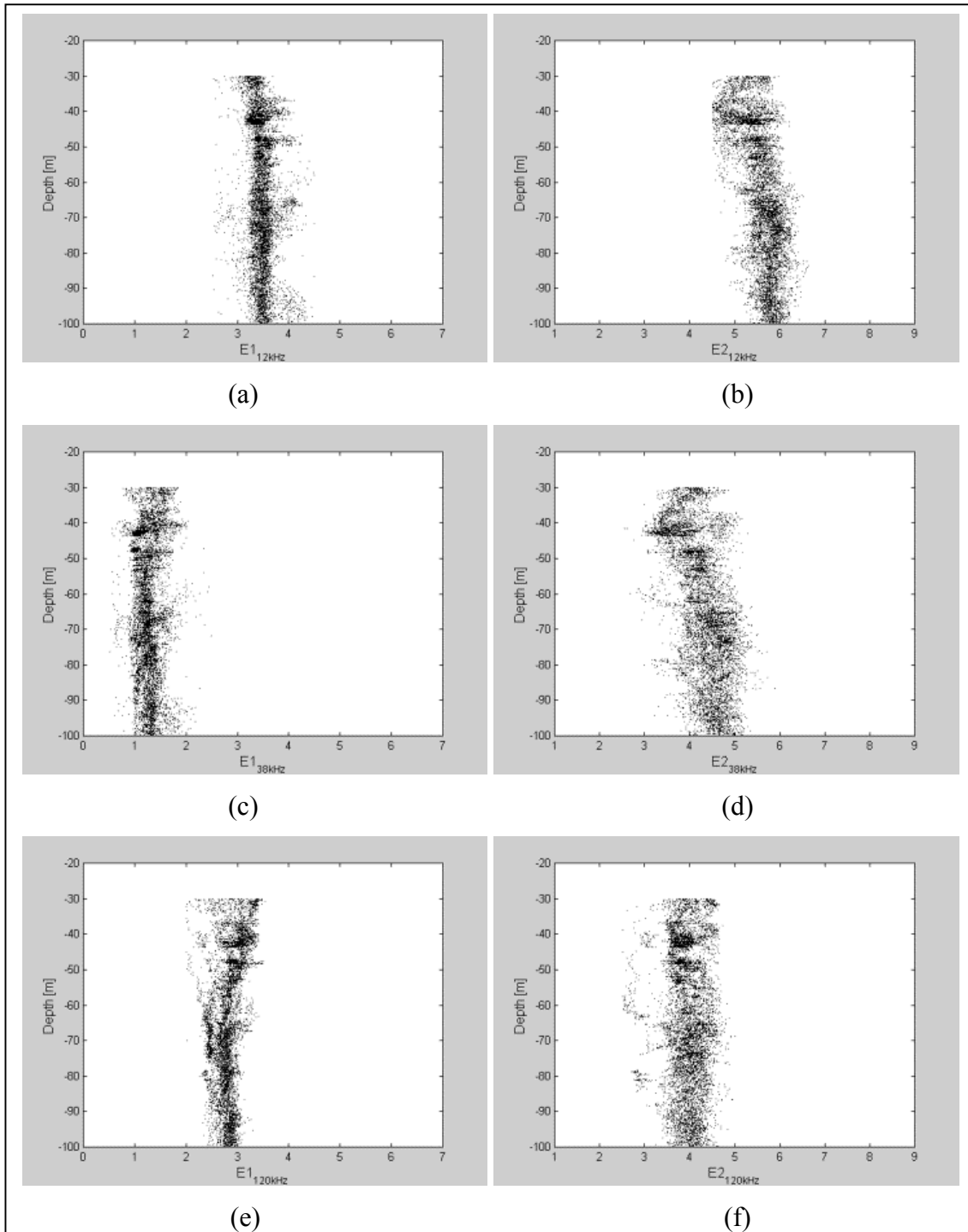
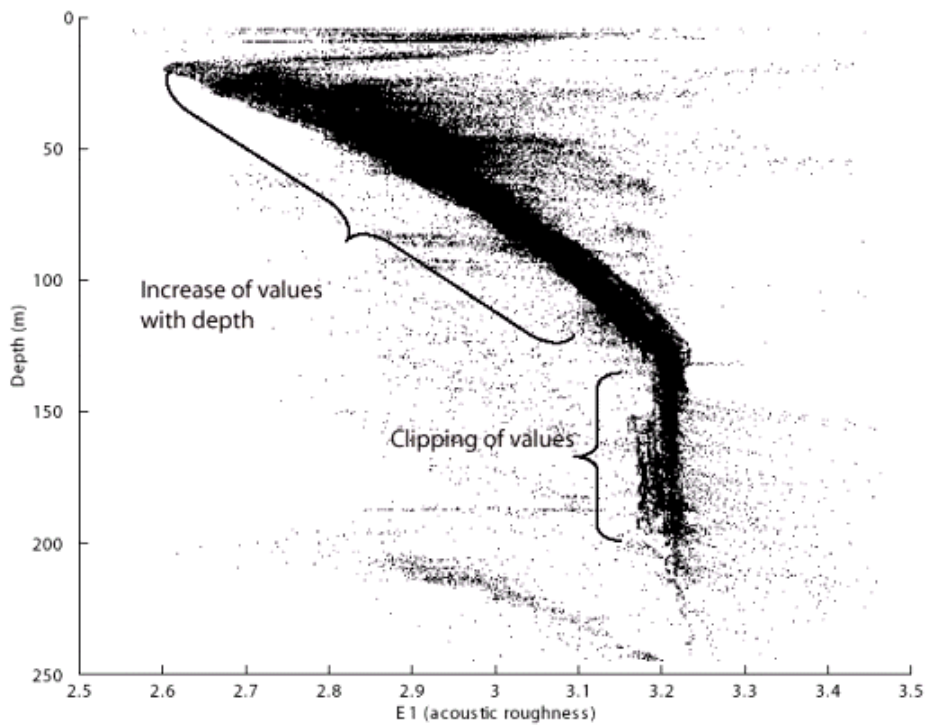


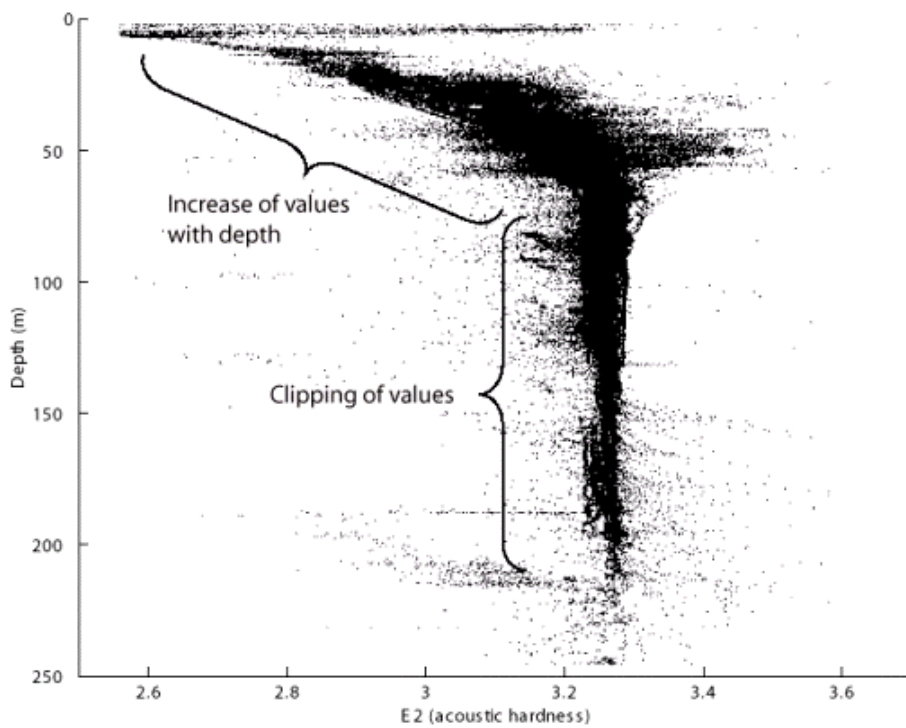
Figure 5.4. Scatterplot of acoustical seabed indices for the SEF study area. (a) roughness index vs depth at 12 kHz. (b) hardness index vs depth at 12 kHz. (c) roughness index vs depth at 38 kHz. (d) hardness index vs depth at 38 kHz. (e) roughness index vs depth at 120 kHz. (f) hardness index vs depth at 120 kHz.



show plots of derived acoustic indices (E1 and E2) versus depth at the three operated frequencies for acoustic data from the NWS and SEF study areas, respectively. It is evident from Figures 5.3 and 5.4 that both roughness and hardness indices have, as earlier discussed, been developed in a largely depth-independent form. In contrast, acoustic data provided by the RoxAnn system attached to the 120 kHz transducer unit operated in the same survey in the SEF region contained major depth biases as shown in Figure 5.5 (Kloser *et al.*, 2001). Prior to the introduction of the constant angular integration interval for E1 index, a constant depth integration interval was adopted to integrate the tail of the first bottom echo to produce E1 index (Ryan *et al.*, 1997; Siwabessy *et al.*, 1999; Kloser *et al.*, 2001). E1 index was found to increase linearly with depth due to spherical spreading of the beam, lengthening the return signal envelopes. However, once the depth trends have been removed from E1 index, the results compared favourably to the E1 index used here, which are derived from a constant angular integration interval. The correlation coefficient between these two sets of E1 index (one was derived using a constant integration interval and the other was derived using a constant angular integration interval) was quite high ( $r = 0.85$ ) but the magnitude was different. The magnitude of the E1 index derived from a constant depth integration interval after the depth trend removal, however, depends on the method of the depth trend removal. The magnitude difference between the two sets of E1 index though might be considered as an offset and is not critical. Similarly, Kloser *et al.* (2001) found that once depth trends from their RoxAnn data (Figure 5.5) have been removed, results compared favourably with their data. Unlike E1 index where a constant angular integration interval was required to produce results independent of depth, E2 index was derived by using a constant depth integration interval. The results remain independent of depth (Figures



(a)

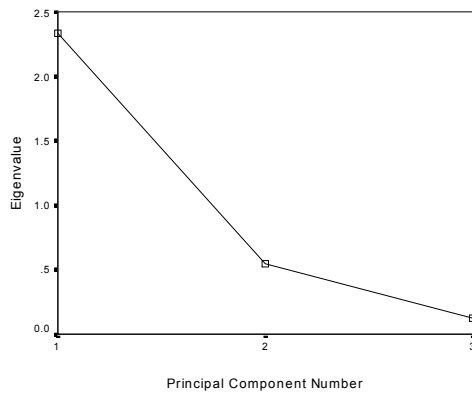


(b)

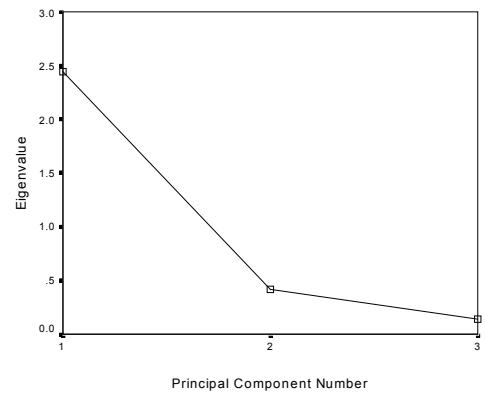
Figure 5.5. Scatterplot of RoxAnn E1 (roughness) and E2 (hardness) indices with depth collected in the SEF study area at the same time and location of the SEF data used here (After Kloser *et al.*, 2001).

5.3 and 5.4). This is understandable since the integration includes the complete second bottom echo that is dominated by coherent components.

Since three frequencies were operated during the surveys in the NWS and SEF regions, there are 3 sets of E1 index and 3 sets of E2 index; each set corresponds to each frequency. To reduce the dimensionality of the E1 and E2 indices, principal component analysis (PCA) was used as outlined in section A.1. The procedure and the criteria concerning the number of principal components (PC) have been described in section A.1. Regardless the magnitude difference, the intercorrelation within E1 index and within E2 index at the three operated frequencies was reasonably high, on average 0.70 and 0.80 for within E1 and within E2 indices, respectively, at the three operated frequencies. The difference in magnitude of E1 index from the three operated frequencies reflects the frequency (and beamwidth) dependency, so does that of E2 index. The frequency dependency however was minimised by using the PCA. The high intercorrelation within E1 index and within E2 index, separately, at the three operated frequencies suggests that E1 (roughness) information at one particular frequency is quite similar to others as is E2 (hardness) information. This in turn suggests that the first PC of E1 and first PC of E2 are sufficient to provide most of the information in the original E1 and E2 data, respectively. Figures 5.6 to 5.7 show the scree plots of principal components of E1 and E2 indices after PCA has been applied to E1 and E2 indices, separately, at the three operated frequencies for the two study areas (Note that PCA was performed on the correlation matrix). It is evident that the eigenvalues of principal components other than the first principal component are less than unity and very small. As such, only the first principal component of E1 (PC1\_E1) and the first principal component of E2 (PC1\_E2) are here retained for further analysis. In addition, this in fact supports the hypothesis mentioned above that

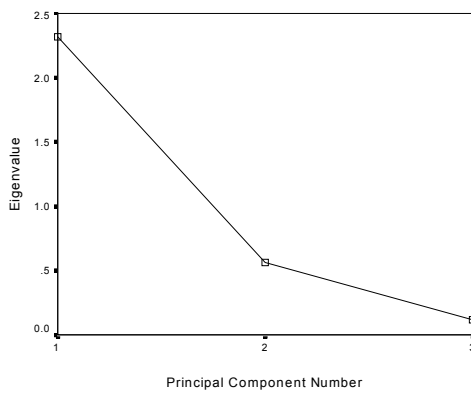


(a)

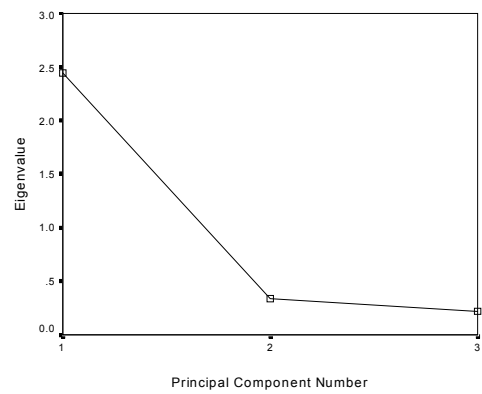


(b)

Figure 5.6. Scree plot of roughness parameter E1 (a) and hardness parameter E2 (b) derived from Table 5.1 for NWS study area. See Table 5.2 for component loadings.



(a)



(b)

Figure 5.7. Scree plot of roughness parameter E1 (a) and hardness parameter E2 (b) derived from Table 5.1 for SEF study area. See Table 5.2 for component loadings.

it is sufficient to retain only one principal component, particularly the first one, due to a high intercorrelation between features within the original data set of either E1 or E2 index. The first principal component accounts for more than 70% of the total variation of the original E1 and more than 80% of the total variation of the original E2 for the two study areas. Details of the total variance of all principal components of E1 and E2 for the two study areas estimated by using equation (A.2) are given in Table 5.1. Component loadings of the first principal component of E1 and E2 for the two study areas were estimated by using equation (A.4) and are given in Table 5.2. Since the PCA was applied to standardised data as discussed in section A.1, the component loadings given in Table 5.2 can also be referred to as component correlations that can be used as a measure of the correlation between the PCs and the original variables. It is evident from Table 5.2 that variables of the first principal component and corresponding variables of the original data are highly correlated. It turns out that the first principal component given as PC1 in Table 5.2 is essentially an average of the three original variables (either roughness indices, E1\_12, E1\_38 and E1\_120, or hardness indices, E2\_12, E2\_38 and E2\_120).

When analysing spatial modelling (section 5.6), it was found that the variance of the first principal component of E1 index in a few straight transects (29% and 18% for the NWS and SEF regions, respectively) was very low. For one particular straight transect, for instance, the first principal component accounted only for 43% of the total variation of the original E1. Three patterns of component loadings were observed from these transects. The first pattern is characterised by high component loadings in E1\_38 and E1\_120 and a low component loading, sometimes with different sign, in E1\_12, in other words the first principal component is influenced most by E1\_38 and E1\_120. The second pattern is identified by E1\_12 and E1\_38

Table 5.1. Eigenvalues and total variances of principal components (PCs) for roughness (E1) and hardness (E2) indices at NWS and SEF study areas. Eigenvalues are calculated by equation (A.3) and are used to plot the scree plot (Figures 5.6 and 5.7). Bold entries are eigenvalues and variances for the first principal component (PC1) and adopted for further analysis.

Roughness Parameter, E1, for NWS study area						
Component	Initial Eigenvalues			Extraction sums of Squared Loadings		
	Eigen value	Variance [%]	Cumulative [%]	Eigen value	Variance [%]	Cumulative [%]
<b>1</b>	<b>2.317</b>	<b>77.249</b>	<b>77.249</b>	<b>2.317</b>	<b>77.249</b>	<b>77.249</b>
2	0.565	18.827	96.076			
3	0.118	3.924	100.000			
Hardness Parameter, E2, for NWS study area						
<b>1</b>	<b>2.443</b>	<b>81.438</b>	<b>81.438</b>	<b>2.443</b>	<b>81.438</b>	<b>81.438</b>
2	0.338	11.281	92.719			
3	0.218	7.281	100.000			
Roughness Parameter, E1, for SEF study area						
<b>1</b>	<b>2.334</b>	<b>77.801</b>	<b>77.801</b>	<b>2.334</b>	<b>77.801</b>	<b>77.801</b>
2	0.544	18.123	95.924			
3	0.122	4.076	100.000			
Hardness Parameter, E2, for SEF study area						
<b>1</b>	<b>2.446</b>	<b>81.528</b>	<b>81.528</b>	<b>2.446</b>	<b>81.528</b>	<b>81.528</b>
2	0.417	13.888	95.416			
3	0.138	4.584	100.000			

Table 5.2. Component loadings of roughness and hardness parameters, estimated from equation (A.4), for NWS and SEF study areas. E1 and E2 are roughness and hardness indices, respectively; 12, 38 and 120 represent 12 kHz, 38 kHz and 120 kHz, respectively.

North West Shelf study area				South East Fisheries study area			
Roughness (E1)		Hardness (E2)		Roughness (E1)		Hardness (E2)	
Variables	PC1	Variables	PC1	Variables	PC1	Variables	PC1
E1_12	0.880	E2_12	0.877	E1_12	0.851	E2_12	0.866
E1_38	0.960	E2_38	0.911	E1_38	0.962	E2_38	0.885
E1_120	0.789	E2_120	0.919	E1_120	0.827	E2_120	0.955

dominating the first principal component where the component loadings of E1\_12 and E1\_38 are high and insignificantly different from one to another whereas the component loading of E1\_120 is low, sometimes with different sign, and significantly different from the component loadings of the other two. The third pattern has E1\_12 and E1\_120 dominate the first principal component, characterised by high component loadings in E1\_12 and E1\_120 and a low component loading, sometimes accompanied with different sign, in E1\_38. As far as surface irregularities of the seabed and beamwidths of the operated transducer units are concerned, the first pattern might be an indication of the scale of surface irregularities that assumes that seabed surface appears rough at 38 and 120 kHz but smooth at 12 kHz and the third pattern, by referring to the beamwidths of the operated transducer units in Table 3.1 in Chapter 3, might indicate a sensitivity of the narrow beamwidth to slopes and sudden drops or rises of the seabed surface as observed by Kloser *et al.* (2001) and Hamilton *et al.* (1999). There seems to be, however, no justification that fits the second pattern.

### **5.3. Training sets**

Training sets for the SEF study area were established from homogeneous areas of the reference sites (training sites) determined from previous surveys in the area. Underwater photographs and grab samples taken in these homogeneous areas confirmed the seabed habitat in the areas (Figure 5.8(a)). Features of acoustic bottom returns within each training site were also visually inspected from the echograms. Locations of the training sites are given as black stars in Figure 5.9(a) and representative examples of echograms from these training sites are shown in Figure 5.10(a). It is important to note that the training sets included only regions of trawlable grounds since underwater photographs used as references came only from trawl

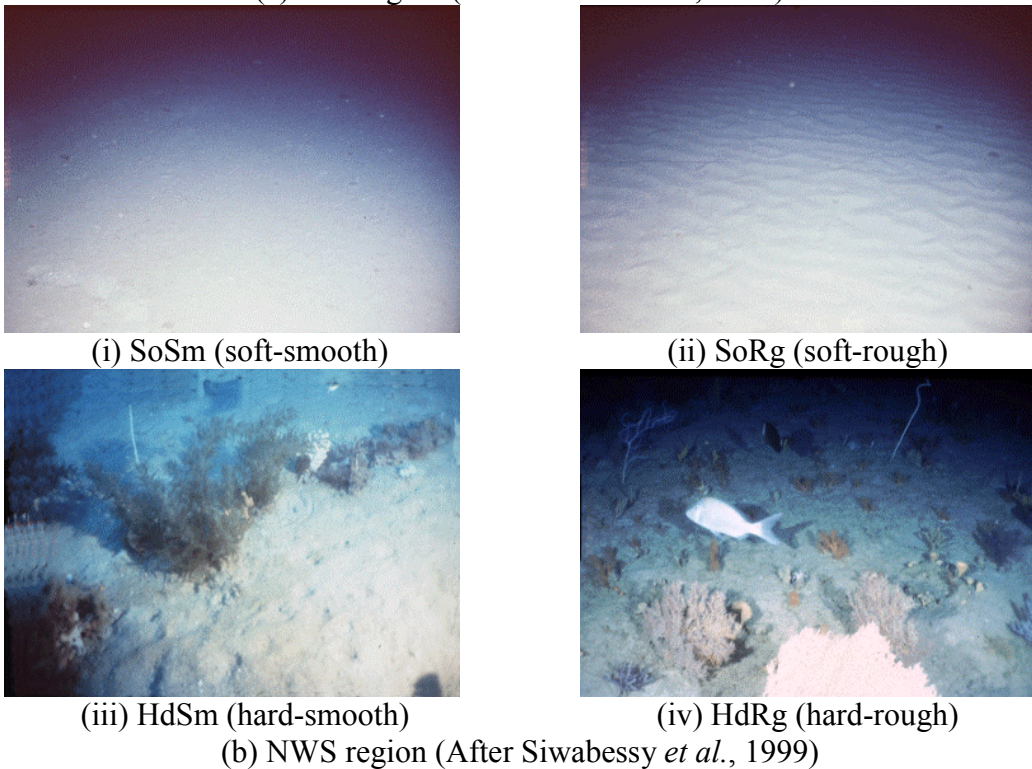
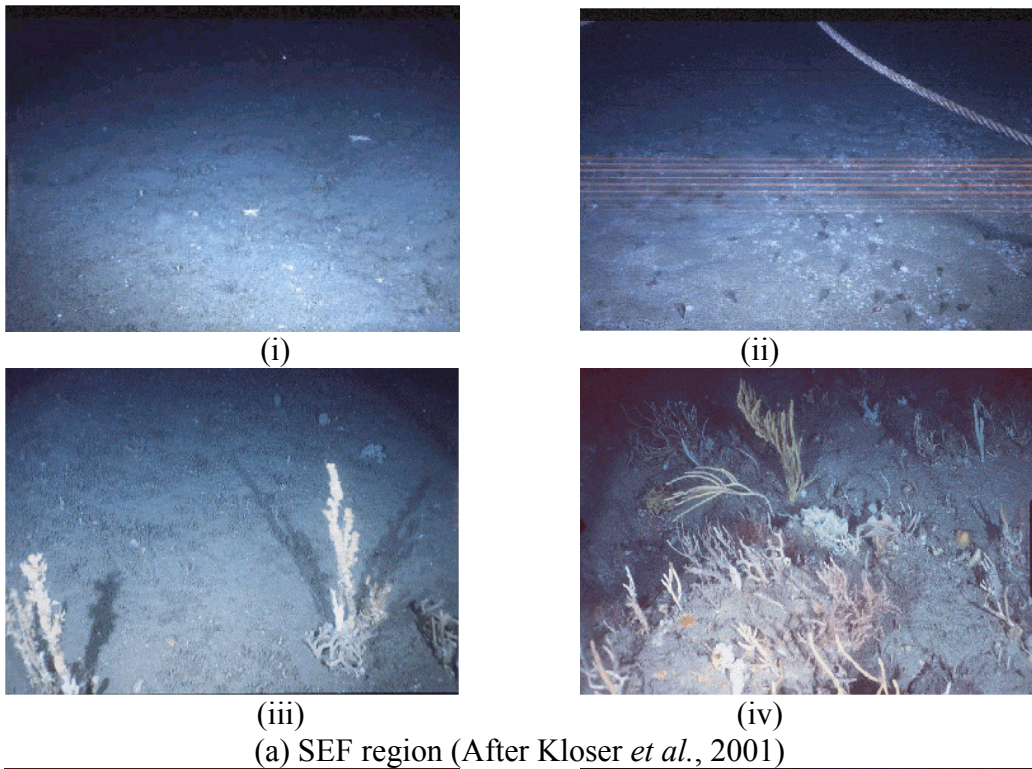
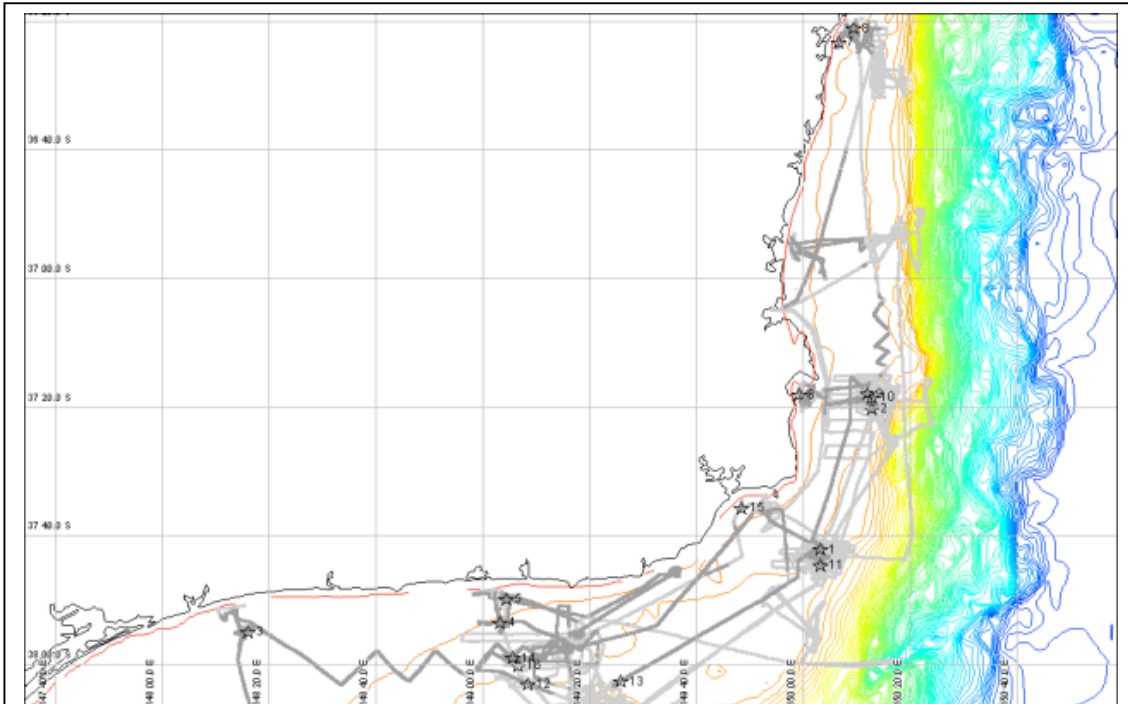
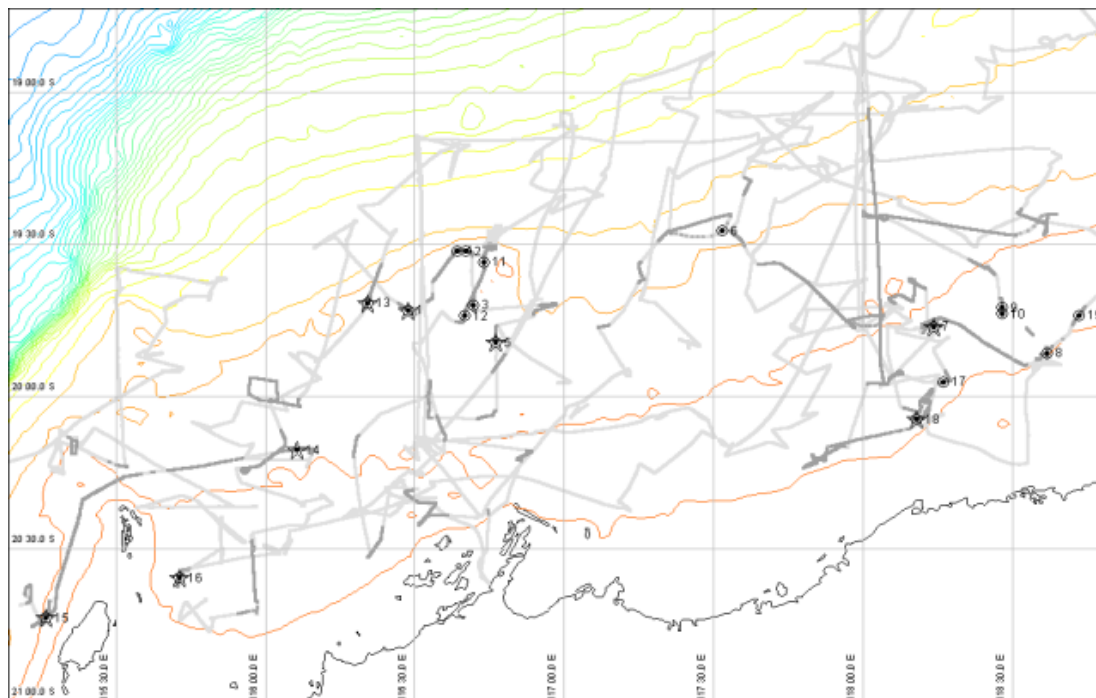


Figure 5.8. Representative examples of seabed images taken by a 35 mm Photosea 1000 camera system in (a) the SEF study area and (b) the NWS study area. Images (i) to (iv) in both study areas (a and b) represent soft-smooth, soft-rough, hard-smooth and hard-rough seabed surfaces and correspond to echograms (i) to (iv) in Figure 5.10.



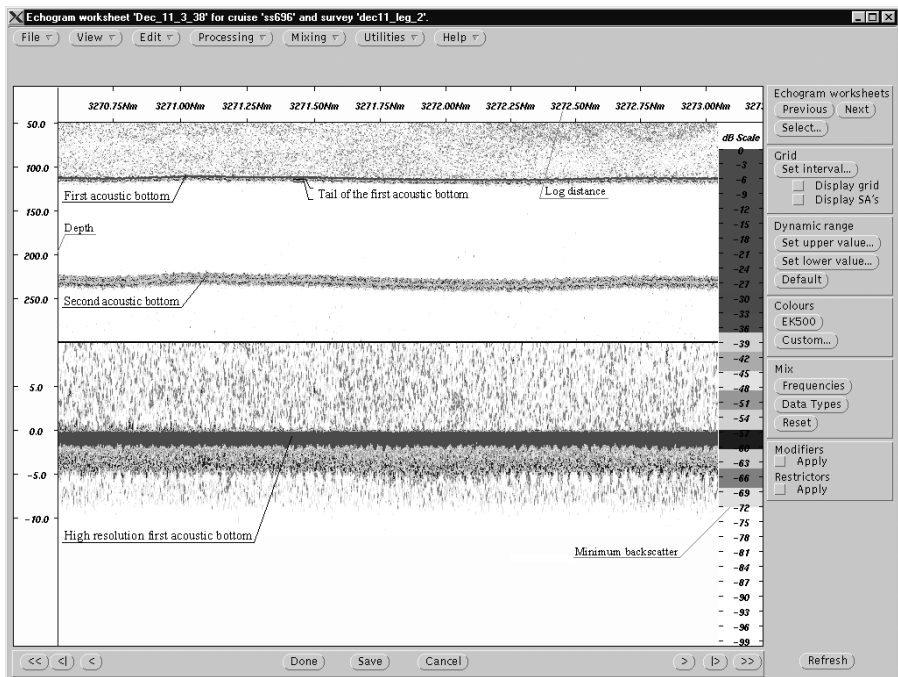


(a) SEF study area

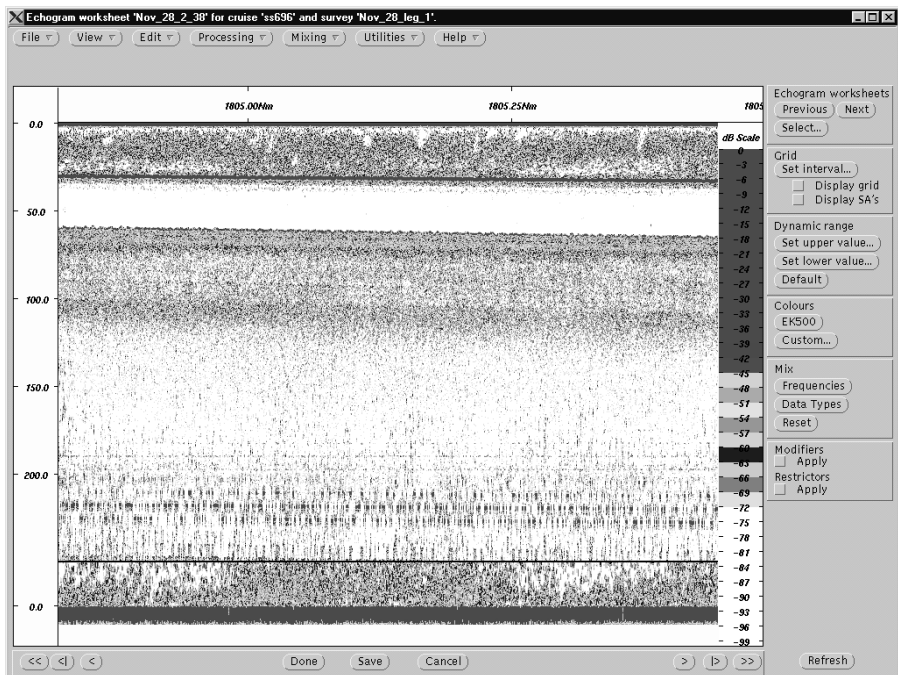


(b) NWS study area

Figure 5.9. Map of the study areas, the complete track (light grey lines), the useful track (dark grey lines) where E1 and E2 indices are available, the coastline and the bathymetry. ☆ = reference sites (for SEF) or training sites (for NWS). ● = available trawl stations. ○ = trawl stations of available photographic records.



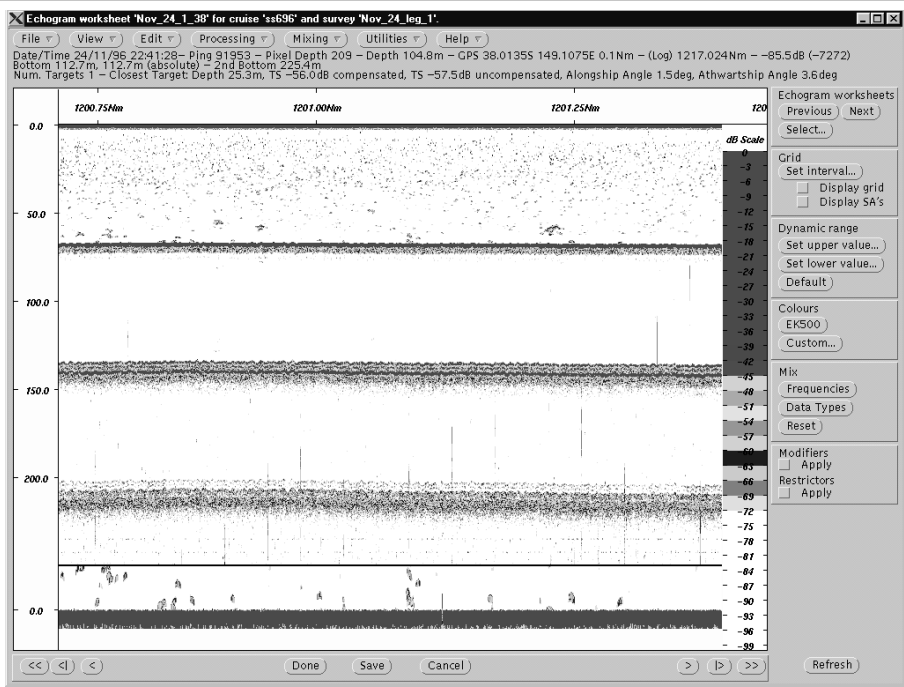
(i) SoSm (12). Frequency=38 kHz. TVG=20logR.



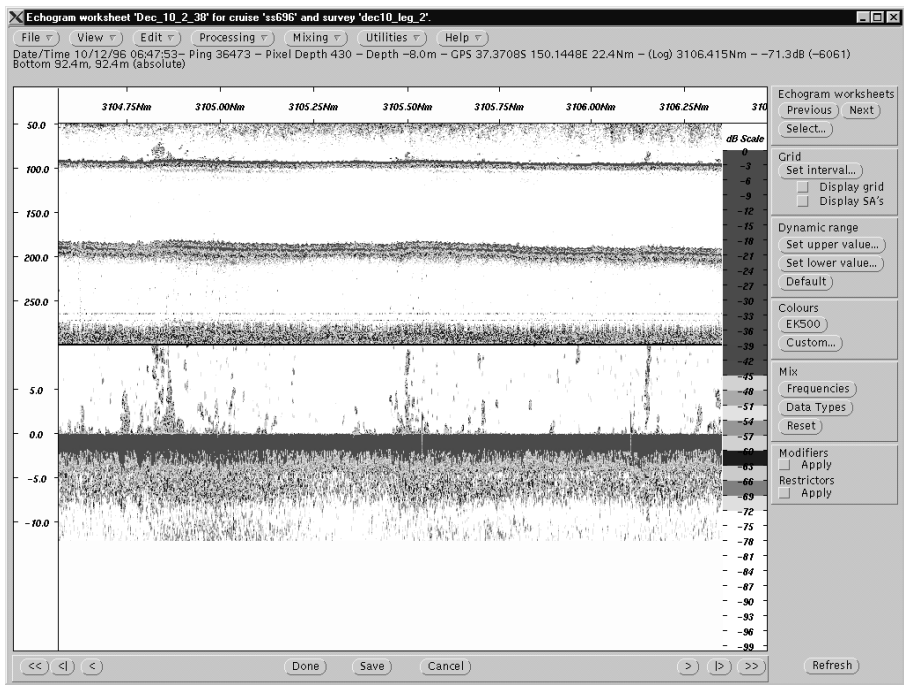
(ii) SoRg (6). Frequency=38 kHz. TVG=20logR.

(a) SEF region

Figure 5.10. Representative examples of echograms of derived seabed class that corresponds to Figure 5.8. Echograms (i) and (ii) correspond to images (i) and (ii) in Figure 5.8(a). Numbers in the bracket correspond to station numbers given in Figure 5.9. Hd = hard; So = soft; Rg = rough; Sm = smooth.



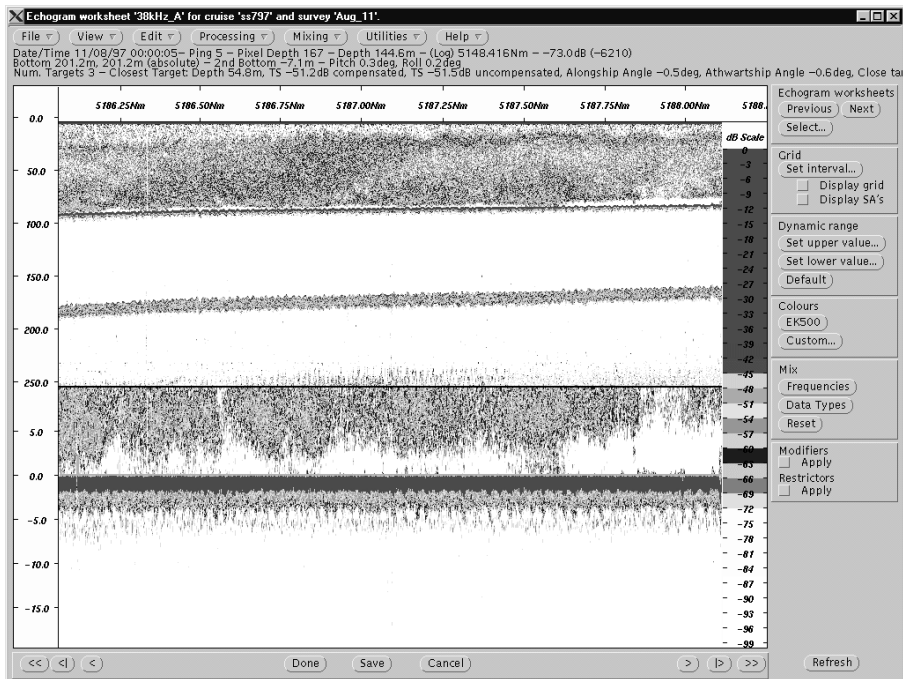
(iii) HdSm (4). Frequency=38 kHz. TVG=20logR.



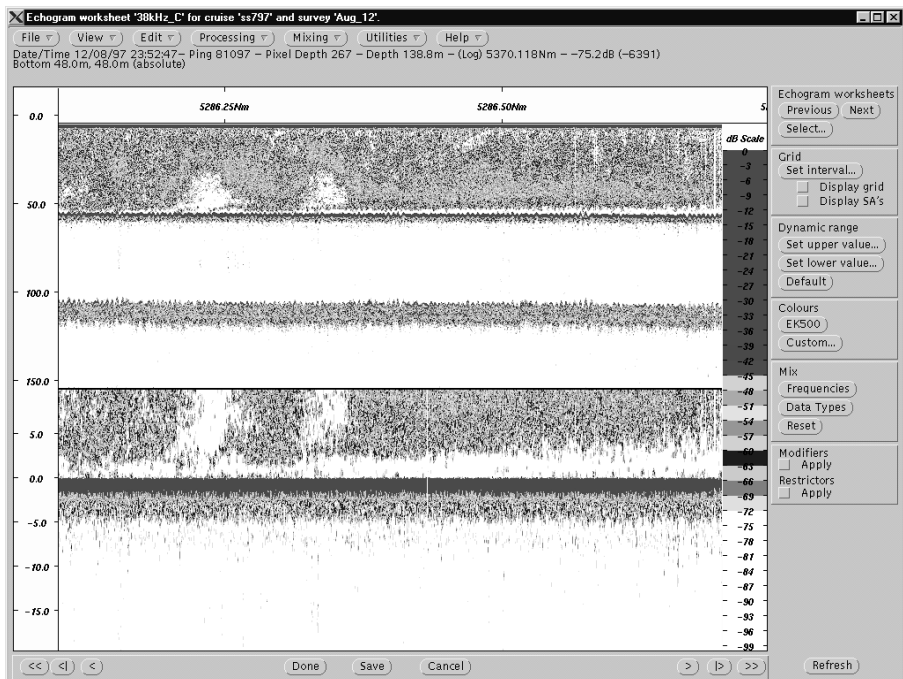
(iv) HdRg (10). Frequency=38 kHz. TVG=20logR.

(a) SEF region

Figure 5.10. *Continued.* Echograms (iii) and (iv) correspond to images (iii) and (iv) in Figure 5.8(a). Numbers in the bracket correspond to station numbers given in Figure 5.9. Hd = hard; So = soft; Rg = rough; Sm = smooth.



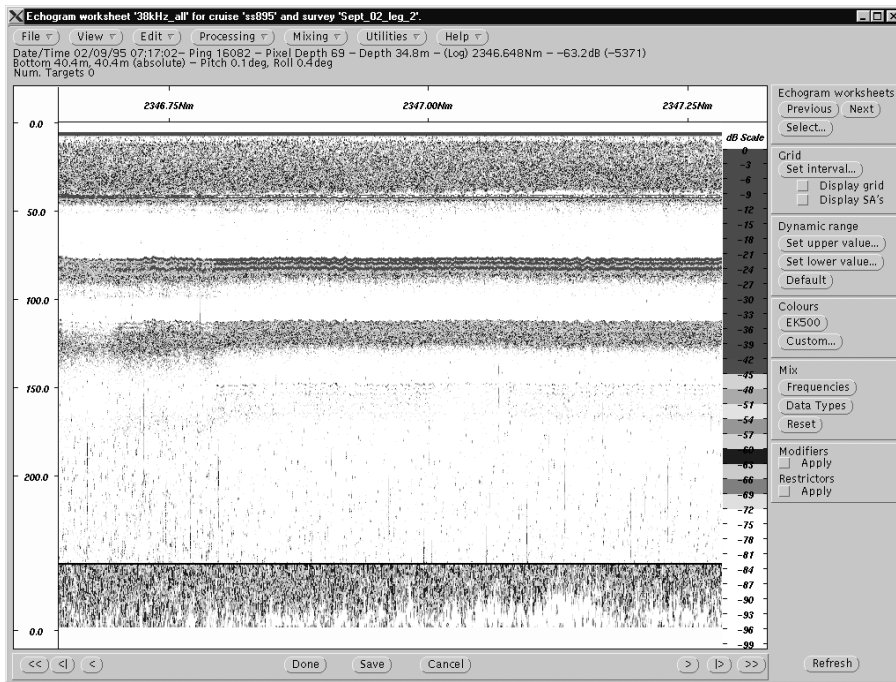
(i) SoSm (13). Frequency=38 kHz. TVG=20logR.



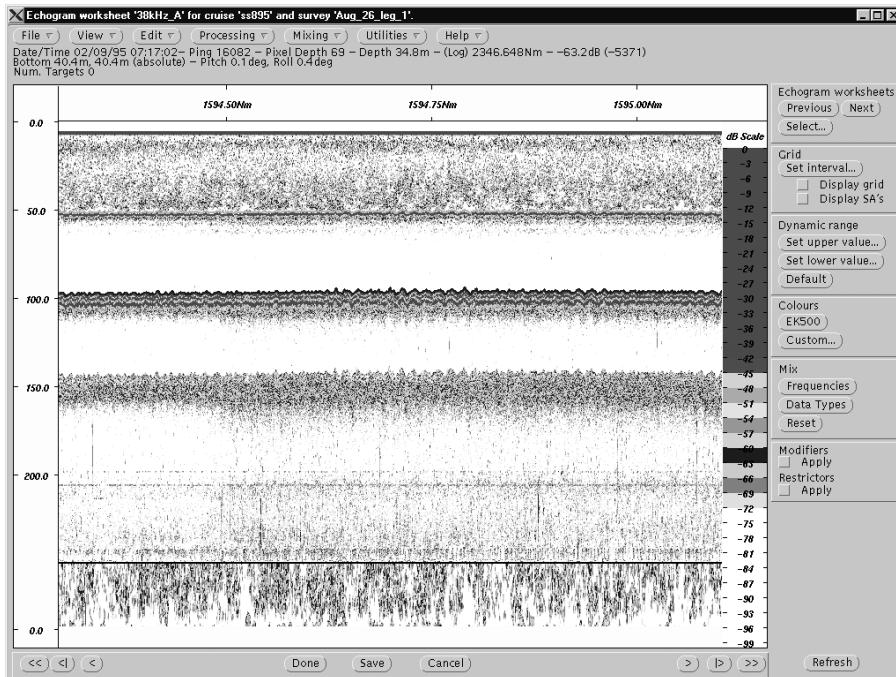
(ii) SoRg (15). Frequency=38 kHz. TVG=20logR.

(b) NWS region

Figure 5.10. *Continued.* Echograms (i) and (ii) correspond to images (i) and (ii) in Figure 5.8(b). Numbers in the bracket correspond to station numbers given in Figure 5.9. Hd = hard; So = soft; Rg = rough; Sm = smooth.



(iii) HdSm (7). Frequency=38 kHz. TVG=20logR.



(iv) HdRg (3). Frequency=38 kHz. TVG=20logR.

(b) NWS region

Figure 5.10. *Continued.* Echograms (iii) and (iv) correspond to images (iii) and (iv) in Figure 5.8(b). Numbers in the bracket correspond to station numbers given in Figure 5.9. Hd = hard; So = soft; Rg = rough; Sm = smooth.

stations even though the acoustic data also included regions of untrawlable grounds as the vessel sailed in regions of trawlable and untrawlable grounds.

The echograms of the soft-smooth seabed type shown in Figures 5.10(a)(i) and (b)(i) are indicated by weak second bottom returns and short tails in the first bottom echoes. For the soft-rough seabed type shown in Figures 5.10(a)(ii) and 5.10(b)(ii), the echograms have long tails in the first bottom echoes and weak second bottom returns. Unlike the echograms of the two seabed types just mentioned, the echograms of the hard-smooth seabed type given in Figures 5.10(a)(iii) and 5.10(b)(iii) show strong second bottom returns. Like the echograms of the soft-smooth seabed types however, the echograms of this seabed types possess short tails in the first bottom echoes. Like the echograms of the hard-smooth seabed type, the echograms of the hard-rough seabed type shown in Figures 5.10(a)(iv) and 5.10(b)(iv) are indicated by strong bottom returns. These echograms, however, have the expected long tails in the first bottom echoes.

For the NWS study area, the training sets were established differently. This was done by firstly selecting seabed images from each trawl station where E1 and E2 parameters, and interpretable seabed images were available and simultaneously noting down any change of general features along the track within each trawl station. These general features included fine sands, coarse sands, muddy sediments, ripples of any kind and size, and outcrops of any kind. The advantages of simultaneously noting down changes of seabed habitat features are twofold. First, it reduces the number of seabed images selected for further analysis since each of them may represent similar features that may exist anywhere within the track of each trawl station. Second, it

provides an initial guesstimate of potential trawl stations or parts of trawl stations that may become training sites.

Of 251 trawl stations from both fisheries surveys (ss895 and ss797) in the NWS region, only 19 given as black dots and stars in Figure 5.9(b) are useable as not only do they provide interpretable seabed images but also have corresponding first and second bottom echoes. Figures 5.8(b) and 5.10(b) show some representative examples of, respectively, interpretable seabed images and echograms common in the 19 trawl stations. Following the procedure outlined in section 3.3.1, the selected seabed images were first scored according to features given in Table 3.2 and the results are given in Table 5.3. Secondly, the hierarchical agglomerative clustering technique was applied to the score data to produce clusters of different seabed habitat types at any association level provided by a dendrogram shown in Figure 5.11. The association level was estimated by using Ward's minimum variance linkage method. The SPSS statistical package that was used here for statistical analysis, however, rescaled the association level to a minimum of 0 and a maximum of 25 units termed rescaled distances (similarity measures). Four different classes at an association level of approximately 13 rescaled distance (see Figure 5.11) are found to be appropriate to represent common seabed habitats in trawl stations in the NWS region. The four classes are hard-rough (HdRg), hard-smooth (HdSm) soft-smooth (SoSm) and soft-rough (SoRg). These four classes correspond respectively to habitats 1 and 2 (H12), habitat 3 (H3), habitat 4 (H4) and habitat 5 (H5) of Sainsbury, *et al.* (in prep). The presence of either outcrops above the seabed, sandwaves or bioturbations or any combination of them, is essentially an indication of seabed roughness. For outcrops of vegetation to grow above the seabed, it is assumed that they require a hard base to support. Therefore they are adopted here to indicate hardness of the seabed. In their

Table 5.3. A set of variables used in the hierarchical agglomerative clustering technique to derive distinguishable seabed classes from photographic records.

Images	Seabed morphology				Sediment type			Benthos type/shape				
	Flat	Ripple	Ridge	Rocky out-cropping	Fine Sand	Coarse sand	Mud	Lump	Cup	Finger	Gor-gonian	Grass
	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
1	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.48	.00	79.95
2	39.63	60.23	.14	.00	100.00	.00	.00	55.10	2.10	.00	2.11	.10
3	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.00	.00	79.95
4	96.11	3.89	.00	.00	.00	.00	100.00	.00	.00	.30	.00	.00
5	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.72	.00	75.00
6	3.86	95.99	.15	.00	.00	100.00	.00	60.12	2.12	.00	2.43	.30
7	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.00	.00	80.15
8	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.48	.00	80.11
9	40.59	59.31	.10	.00	100.00	.00	.00	55.32	2.12	.00	2.13	.00
10	97.21	2.79	.00	.00	100.00	.00	.00	.00	.00	.19	.00	.00
11	39.11	60.66	.11	.12	.00	100.00	.00	59.15	.00	2.15	.00	.25
12	100.00	.00	.00	.00	.00	.00	100.00	.00	.00	.00	.00	.10
13	.00	98.80	1.20	.00	.00	.00	100.00	.00	.00	.91	.00	.20
14	.00	99.00	1.00	.00	.00	.00	100.00	.00	.00	.10	.00	.00
15	95.95	4.05	.00	.00	100.00	.00	.00	.00	.00	.34	.00	.00
16	.00	100.00	.00	.00	100.00	.00	.00	.00	.00	.00	.00	.00
17	95.00	5.00	.00	.00	.00	.00	100.00	.00	.00	.43	.00	.35
18	1.28	98.58	.00	.14	.00	.00	100.00	55.96	1.99	2.14	1.13	.10
19	.00	98.90	1.10	.00	100.00	.00	.00	.00	.00	.89	.00	.19
20	1.24	98.59	.00	.17	100.00	.00	.00	56.12	2.00	2.21	1.11	.11
21	96.10	3.90	.00	.00	100.00	.00	.00	.00	.00	.36	.00	.42
22	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.48	.00	80.10
23	.00	99.50	.50	.00	.00	.00	100.00	.00	.00	.67	.15	.47
24	99.51	.49	.00	.00	.00	.00	100.00	.00	.00	.59	.00	.48
25	.00	100.00	.00	.00	100.00	.00	.00	.00	.00	.00	.00	.00
26	.00	98.80	1.20	.00	.00	.00	100.00	.00	.00	1.11	.20	.39
27	.00	98.88	.14	.98	.00	.00	100.00	50.22	.00	1.59	.00	.00
28	100.00	.00	.00	.00	.00	.00	100.00	.00	.00	.00	.00	.11
29	.00	99.47	.53	.00	100.00	.00	.00	.00	.00	.65	.17	.49
30	95.12	4.88	.00	.00	100.00	.00	.00	.00	.00	.39	.00	.41
31	100.00	.00	.00	.00	100.00	.00	.00	.00	.00	.00	.00	.14
32	.00	98.88	.13	.99	.00	100.00	.00	50.10	.00	1.61	.00	.00
33	.00	98.89	1.11	.00	100.00	.00	.00	.00	.00	1.21	.19	.45
34	98.95	1.05	.00	.00	.00	.00	100.00	.00	.00	.21	.00	.00
35	.00	98.86	1.14	.00	100.00	.00	.00	.00	.00	.13	.00	.00
36	.00	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.11	.00
37	100.00	.00	.00	.00	.00	100.00	.00	.10	.00	.00	.10	75.93
38	.00	100.00	.00	.00	100.00	.00	.00	.00	.00	.00	.00	.00
39	99.15	.85	.00	.00	100.00	.00	.00	.00	.00	.50	.00	.00
40	100.00	.00	.00	.00	100.00	.00	.00	.00	.00	.00	.00	.14



Table 5.3. *Continued.*

Images	Seabed morphology				Sediment type			Benthos type/shape				
	Flat	Ripple	Ridge	Rocky out-cropping	Fine Sand	Coarse sand	Mud	Lump	Cup	Finger	Gor-gonian	Grass
	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
41	.00	98.89	1.11	.00	.00	.00	100.00	.00	.00	.98	.00	.21
42	95.00	5.00	.00	.00	.00	.00	100.00	.00	.00	.41	.00	.39
43	100.00	.00	.00	.00	.00	.00	100.00	.00	.00	.00	.00	.10
44	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.51	.00	79.49
45	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.00	.00	80.20
46	.00	98.90	.10	1.00	.00	100.00	.00	50.21	.00	1.57	.00	.00
47	100.00	.00	.00	.00	100.00	.00	.00	.00	.00	.00	.00	.13
48	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.50	.00	79.51
49	100.00	.00	.00	.00	.00	100.00	.00	.00	.00	.00	.00	80.00
50	40.66	59.22	.12	.00	.00	.00	100.00	55.29	2.14	.00	2.10	.00

Table 5.4. Description of overall seabed types in trawl stations available shown as open circle in Figure 5.9. Rough estimates of the proportion of four seabed types (in % by the total number of images in each trawl station) are given in the bracket.

Cruise ID	Trawl station ID	CSIRO ID	Mean Depth [m]	Description
9508	1	9508042	59	Fine sand and very few mud, ripples (50), outcrops (18), smooth/none (32)
9508	2	9508044	59.5	Coarse sand, few ripples (2), outcrops (60), smooth/none (38)
9508	3	9508047	65	Coarse sand, very few ripples (<1), outcrops (70), smooth/none (29)
9508	4	9508049	56	Coarse sand, very few ripples (<1), outcrops (56), smooth/none (43)
9508	5	9508055	38	Fine sand and very few mud at the end, outcrops (95) including grass-like, smooth/none (5)
9508	6	9508080	75	Coarse sand and very few mud, outcrops (65), smooth/none (35)
9508	7	9508095	24	Coarse sand, outcrops (35), smooth/none (65)
9508	8	9508107	23	Mud, few bioturbation and few ripples at the end (2), outcrops (41), smooth/none (57)
9508	9	9508108	33.5	Coarse sand, ripples and bioturbation (20), outcrops (70), smooth/none (10)
9508	10	9508109	36	Coarse sand, ripples and bioturbation (15), outcrops (65), smooth/none (20)
9707	11	9707013	50	Coarse sand, outcrops (53), smooth/none (47)
9707	12	9707014	48.5	Coarse sand and very few fine sand, outcrops (69), smooth/none (31)
9707	13	9707026	67	Fine sand, few ripples (2), outcrops (13) smooth/none (85)
9707	14	9707032	50	Mud, smooth/none (100)
9707	15	9707033	43	Fine sand, ripples (70), outcrops (20), smooth/none (10)
9707	16	9707055	26	Coarse sand, outcrops (34), smooth/none (34), ripples (32) at the end
9707	17	9707106	27	Fine sand, ripples (14), outcrops (65), smooth/none (21)
9707	18	9707107	25	Fine sand, few ripples (4), outcrops (86), smooth/none (10)
9707	19	9707109	26	Coarse sand, few ripples (4), outcrops (85) including grass-like, smooth/none (11)

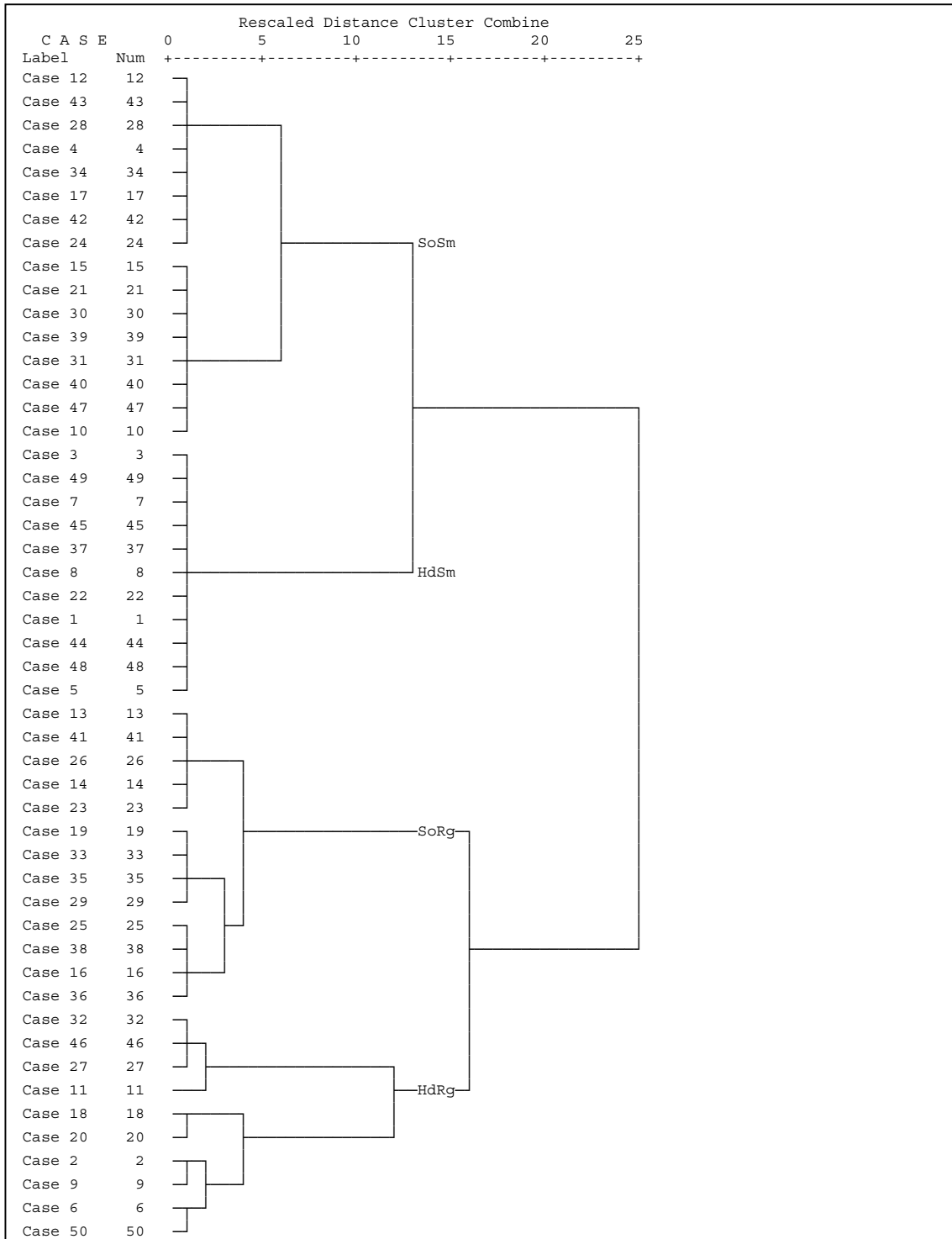


Figure 5.11. Dendrogram showing four derived seabed classes from photographic records produced by applying a set of variables given in Table 5.3 into the agglomerative hierarchical clustering technique. Four different seabed classes occur approximately at an association level of 13 rescaled distance. SoSm = soft-smooth; HdSm = hard-smooth; SoRg = soft-rough; HdRg = hard-rough.

study of benthic habitat types in the NWS region, Sainsbury, *et al.* (in prep) have reviewed the photographic records, a total of 46,419 seabed images, taken from all fisheries surveys between 1983 and 1997 in the region and suggested five different habitat types. Their habitats 1 and 2 (H1 and H2), however, are not acoustically distinguishable here. The four proposed seabed habitat types from this study favour those proposed by Sainsbury, *et al.* (in prep) given their habitats 1 and 2 are acoustically considered similar, hard-rough.

Table 5.4 provides a summary of features of each trawl station along with approximate portions of features along the track of each trawl station that are useful to help in searching for homogeneous trawl stations, in terms of seabed habitat types, to be then taken as training sites. From Table 5.4, fine sand sediments are found to occur in 6 trawl stations (1, 5, 13, 15, 17, 18) whereas coarse sand is dominant in 11 trawl stations (2, 3, 4, 6, 7, 9, 10, 11, 12, 16, 19). Trawl stations 8 and 14 mainly consist of muddy sediments. Trawl station 14 is unique as it has nothing but muddy sediments along the entire track; in other words it is completely (100%) homogeneous. The entire track of trawl station 14 was selected as one of the training sites representing a soft-smooth seabed class. Another training site representing this sort of seabed class was taken from trawl station 13. Since small amplitude ripples and ridges were present in a very small section at the very end of trawl station 13, only the large section at the beginning of trawl station 13 in which flat, fine sands were present was selected. Training sites for a soft-rough seabed class were taken from almost the complete track of trawl stations 1 and 15. It is evident from Table 5.4 that there are many potential trawl stations that represent a hard-rough seabed type. They are trawl stations 2, 3, 4, 5, 9, 11, 12, 17, 18, and 19. Only trawl stations 5 and 18, though, were selected to form training sets of the hard-rough seabed class as they

appeared to have along the track larger sections of this seabed class than others. Although not all trawl stations with the potential to form training sets were selected, the unselected trawl stations were used later for a comparison with the selected ones. Among the derived four seabed types, the hard-smooth class never appear to be the majority at least as far as the trawl stations available to use here, given in Table 5.4, are concerned. Nevertheless, part of the track of trawl stations 7 and 16 are representative enough and were selected to portray the training site of the hard-smooth seabed type. It is important to note that, for those trawl stations where only part of the track is selected to form the training sets, the corresponding acoustic data are extracted using the timestamp of the seabed images of the selected sections and taking into account the lag distance between them. Training sites used here to form training sets are trawl stations 1, 5, 7, 13, 14, 15, 16 and 18. The training sets comprise the first principal component of roughness index (PC1\_E1) and the first principal component of hardness index (PC1\_E2) from the 8 training sites just mentioned.

Initial centroids of the training sets (PC1\_E1 and PC1\_E2) for different classes were estimated by simply taking the arithmetic means of PC1\_E1 and PC1\_E2 values from training sites (Table 5.5). Similarly, centroids of E1 and E2 of each individual were also estimated. Figures 5.12(d) and 5.13(d) show a scatterplot of PC1\_E1 versus PC1\_E2 of the training sets for the SEF and NWS study areas, respectively. Plots of E1 versus E2 of each individual frequency are given in Figures 5.12(a) to 5.12(c) and 5.13(a) to 5.13(c) for the two study areas, respectively. Both Figures 5.12(d) and 5.13(d) suggest four distinct, well separated seabed types in the SEF and NWS study areas with a minor overlapping between classes in comparison to Figures 5.12(a) to 5.12(c) for the SEF region and Figures 5.13(a) to 5.13(c) for the

Table 5.5. Centroids of four seabed types. The initial centroids are simply arithmetic means of the roughness and hardness indices (E1 and E2) for single frequency, and the first principal component of roughness and hardness indices (PC1\_E1 and PC1\_E2) for multi frequencies in the SEF study area and the NWS study area. The final centroids for multi frequencies (12, 38 and 120 kHz) given as bold entries are estimated after applying the iterative relocation clustering (*k*-means) technique using the feedback system and are the centroids adopted for seabed classifications of the remaining data.

Frequency [kHz]	Variable	Class							
		SEF				NWS			
		SoSm	HdSm	HdRg	SoRg	SoSm	HdSm	HdRg	SoRg
12	E1	3.231	3.442	4.106	3.486	3.481	3.645	4.454	3.767
	E2	5.082	5.920	5.851	4.988	5.706	6.278	6.234	5.292
38	E1	0.848	1.407	1.870	1.429	1.067	1.131	1.618	1.197
	E2	4.248	4.693	4.450	3.691	4.197	5.065	4.697	2.659
120	E1	2.161	2.717	3.084	3.066	1.995	2.518	2.806	2.545
	E2	3.272	4.124	3.998	3.745	3.296	4.513	4.188	3.497
12, 38	PC1_E1	-1.655	0.088	2.622	0.237	-0.775	-0.436	1.561	-0.148
	PC1_E2	-0.837	0.855	0.478	-1.609	-0.354	1.094	0.756	-2.086
12, 120	PC1_E1	-2.204	-0.340	2.194	0.611	-1.657	-0.161	1.646	0.072
	PC1_E2	-1.676	0.762	0.484	-1.163	-1.113	1.311	0.819	-1.445
38, 120	PC1_E1	-2.322	0.151	2.000	0.956	-1.364	-0.024	1.568	0.163
	PC1_E2	-1.002	0.715	0.245	-1.031	-0.821	1.431	0.725	-1.738
12, 38, 120	PC1_E1	-2.154	0.013	2.413	0.634	-1.326	-0.219	1.691	0.031
	PC1_E2	-1.175	0.811	0.423	-1.344	-0.769	1.312	0.790	-1.827
	PC1_E1	<b>-2.102</b>	<b>0.566</b>	<b>2.895</b>	<b>0.646</b>	<b>-1.294</b>	<b>-0.305</b>	<b>1.302</b>	<b>0.030</b>
	PC1_E2	<b>-1.162</b>	<b>0.659</b>	<b>0.486</b>	<b>-1.386</b>	<b>-0.595</b>	<b>1.520</b>	<b>0.717</b>	<b>-1.838</b>

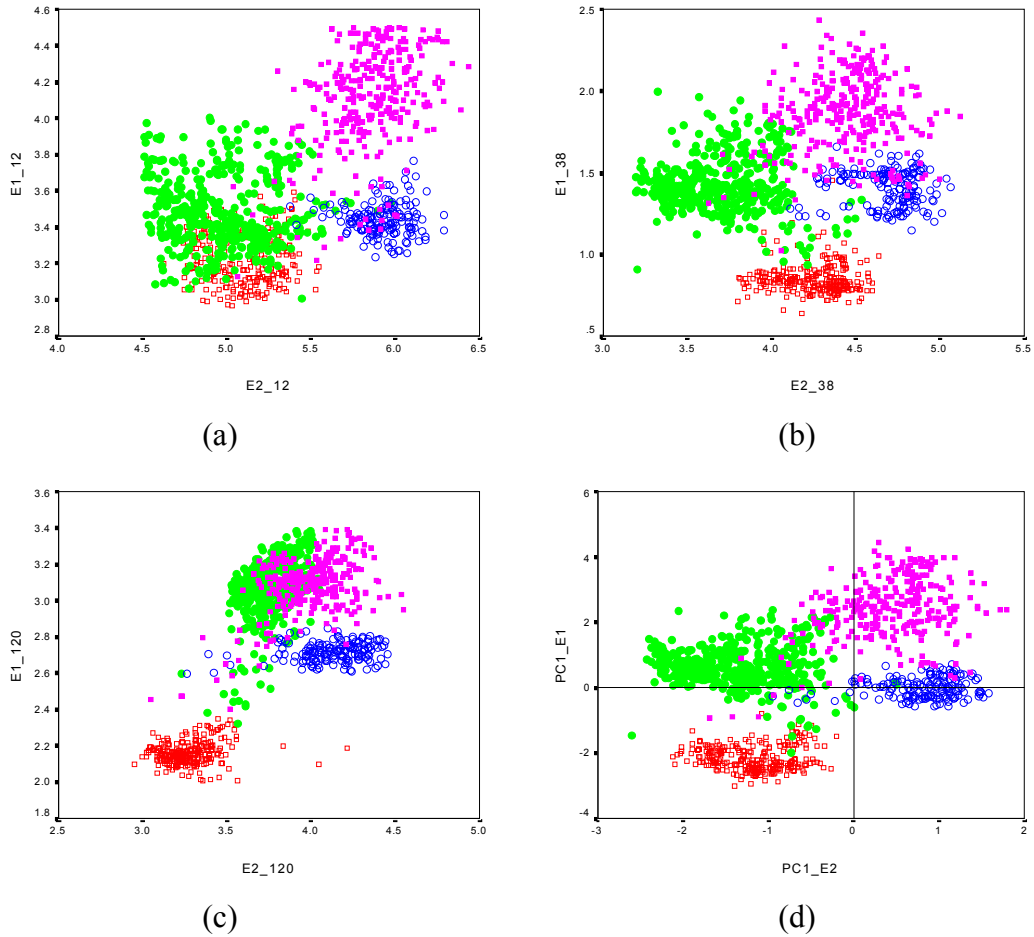


Figure 5.12. Scatterplot of hardness index versus roughness index of the training sets comprising 4 different classes of the seabed for the SEF study area. (a) to (c) roughness index (E1) versus hardness index (E2) for individual 12, 38 and 120 kHz, respectively. (d) first principal component of roughness index (PC1\_E1) versus first principal component of hardness index (PC1\_E2).  $\square$  = soft-smooth;  $\circ$  = hard-smooth;  $\blacksquare$  = hard-rough;  $\bullet$  = soft-rough.

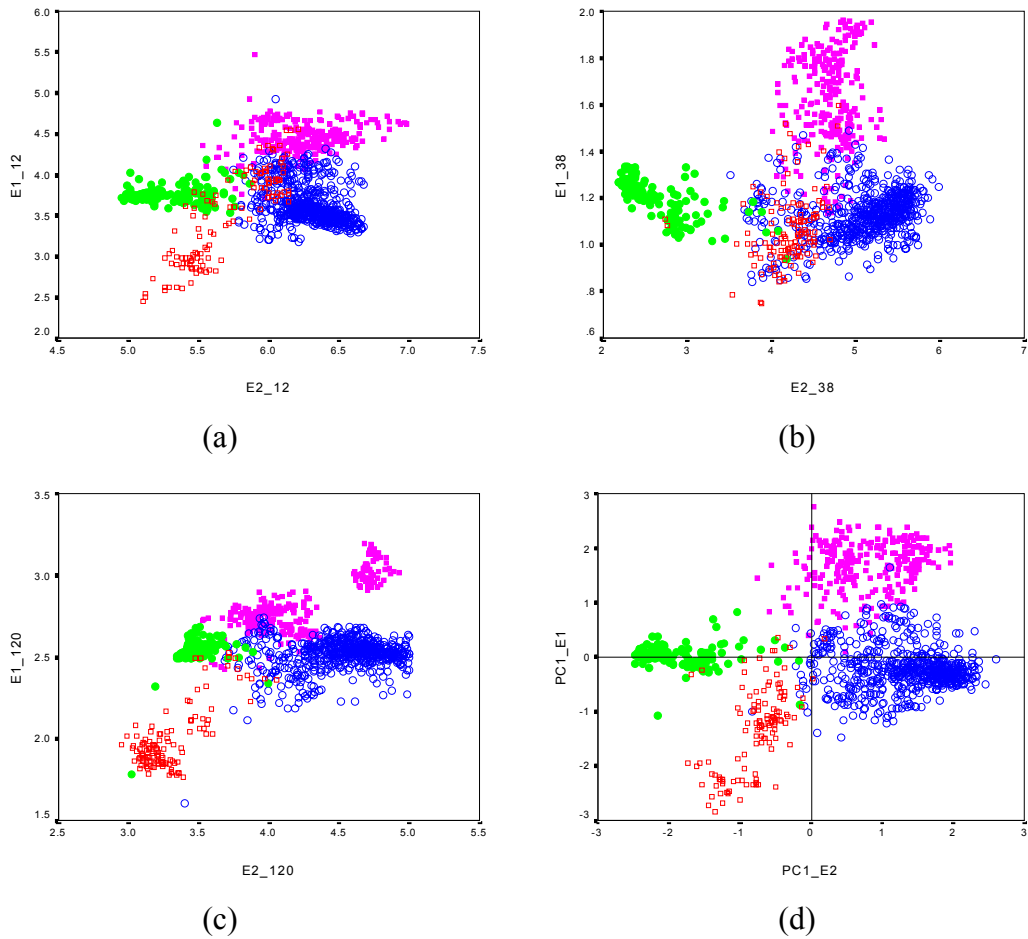


Figure 5.13. Scatterplot of hardness index versus roughness index of the training sets comprising 4 different classes of the seabed for the NWS study area. (a) to (c) roughness index (E1) versus hardness index (E2) for individual 12, 38 and 120 kHz, respectively. (d) first principal component of roughness index (PC1\_E1) versus first principal component of hardness index (PC1\_E2).  $\square$  = soft-smooth;  $\circ$  = hard-smooth;  $\blacksquare$  = hard-rough;  $\bullet$  = soft-rough.

NWS region, in particular classes SoRg and HdRg at 120 kHz from the SEF region and classes SoSm and HdSm at 38 kHz from the NWS region. The Euclidean distances between classes were also estimated from the initial centroids and the results are given in Table 5.6. The general trend is that class separation derived from individual frequencies as indicated by the Euclidean distances is lower than that derived from PCA that includes the three frequencies (multi-frequency). This indicates twofold: (1) there is an increase in class discrimination for the multi frequency technique in combination with PCA adopted here and (2) there is potential overlapping between classes for bottom classification using single frequencies in particular at some of segmented areas (section 5.6) where the correlations between roughness index at the three frequencies and between hardness index at the three frequencies are low (a detailed discussion is covered in section 5.6). The increasing class discrimination indicated by increasing Euclidean distances with the number of frequencies used in the analysis here appears to agree with Kavli *et al.* (1994). In their study of bottom classification, Kavli *et al.* (1994) using the Mahalanobis distance instead of the Euclidean distance adopted here as a measure of class discrimination (separation) found that there is a reduction in class discrimination for each removal of either a frequency or a variable. In this study however, the increasing Euclidean distances between classes with the number of frequencies used from dual frequency to triplet frequency is hardly observed. In addition, some of these Euclidean distances for the dual frequency are even slightly higher than their counterpart for the triplet frequency given as superscripts a and b in Table 5.6. Table 5.7 shows the percentage of the assigned class in each of the four seabed classes based on centroids given in Table 5.5. The percentage of the assigned class increases with the number of frequencies involved, except for entries labelled with superscript a in



Table 5.6. Euclidean distances between four seabed classes estimated for single frequency and multi-frequency (ie. PC1\_E1 and PC1\_E2 after applying principal component analysis) based on initial centroids (simple arithmetic means) shown in Table 5.5. Bold entries are Euclidean distances estimated using final centroids shown in Table 5.5 after the refinement. <sup>a</sup>The Euclidean distance of a particular entry for the dual frequency is higher than its counterpart for the triplet frequency prior to the refinement of the centroids. <sup>b</sup>The Euclidean distance of a particular entry for the dual frequency is higher than its counterpart for the triplet frequency after the refinement of the centroids.

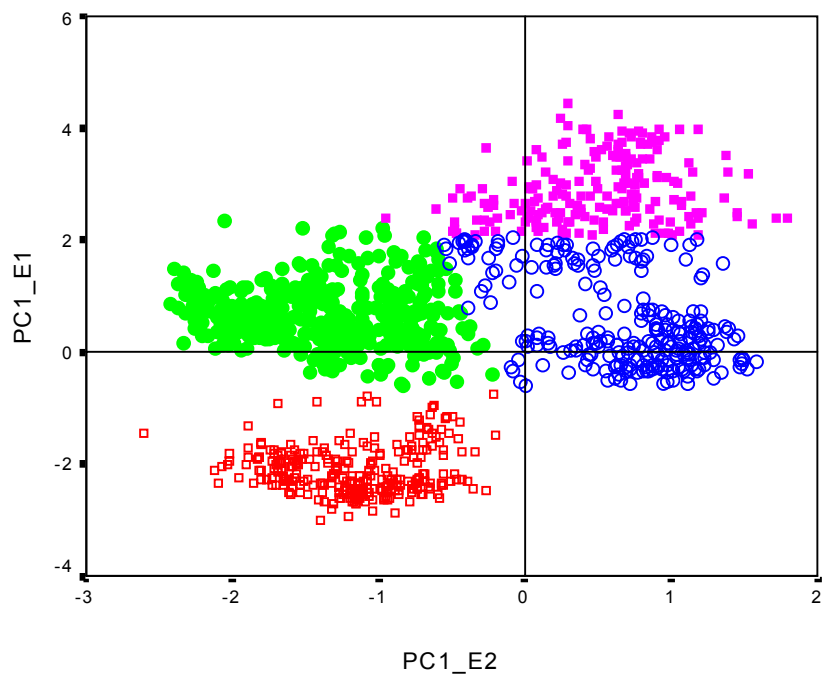
Frequency [kHz]		Class	Euclidean Distance					
			SEF			NWS		
			SoSm	HdSm	HdRg	SoSm	HdSm	HdRg
Single	12	HdSm	0.864			0.595		
		HdRg	1.165	0.667		1.107	0.811	
		SoRg	0.272	0.933	1.062	0.504	0.994	1.166
	38	HdSm	0.714			0.870		
		HdRg	1.041	0.523		0.744	0.610	
		SoRg	0.805	1.002	0.878	1.544	2.407	2.081
	120	HdSm	1.018			1.325		
		HdRg	1.175	0.388		1.206	0.434	
		SoRg	1.022	0.515	0.254	0.585	1.016	0.739
Dual	12, 38	HdSm	2.429			1.488		
		HdRg	4.474	<sup>a,b</sup> 2.561		2.586	<sup>a,b</sup> 2.025	
		SoRg	2.043	<sup>a,b</sup> 2.469	<sup>a,b</sup> 3.169	<sup>a,b</sup> 1.841	<sup>a</sup> 3.193	<sup>a,b</sup> 3.316
	12, 120	HdSm	<sup>a</sup> 3.069			<sup>a,b</sup> 2.849		
		HdRg	<sup>a</sup> 4.900	<sup>a,b</sup> 2.548		<sup>a,b</sup> 3.827	<sup>b</sup> 1.873	
		SoRg	<sup>a,b</sup> 2.862	<sup>b</sup> 2.147	2.284	<sup>a</sup> 1.761	2.766	2.758
	38, 120	HdSm	<sup>a</sup> 3.010			<sup>a,b</sup> 2.620		
		HdRg	4.498	1.908		<sup>b</sup> 3.315	1.742	
		SoRg	<sup>a,b</sup> 3.277	1.923	1.649	<sup>a</sup> 1.782	<sup>a</sup> 3.174	2.836
Triplet	12, 38, 120	HdSm	2.940			2.357		
		HdRg	4.839	2.431		3.395	1.980	
		SoRg	2.794	2.243	2.508	1.721	3.149	3.099
		HdSm	<b>3.230</b>			<b>2.334</b>		
		HdRg	<b>5.261</b>	<b>2.335</b>		<b>2.909</b>	<b>1.796</b>	
		SoRg	<b>2.756</b>	<b>2.046</b>	<b>2.927</b>	<b>1.816</b>	<b>3.374</b>	<b>2.854</b>

Table 5.7. Percentage of the assigned class derived by using the *k*-means technique applied to the training sets of the SEF and NWS study areas, separately, using the centroids given in Table 5.5. Bold entries are the percentage of the assigned class using the refined centroids given as bold entries in Table 5.5. <sup>a</sup>The percentage of the assigned class in a particular seabed class for the dual frequency is higher than its counterpart for the triplet frequency prior to the refinement of the centroids. <sup>b</sup>The percentage of the assigned class in a particular seabed class for the dual frequency is higher than its counterpart for the triplet frequency after the refinement of the centroids.

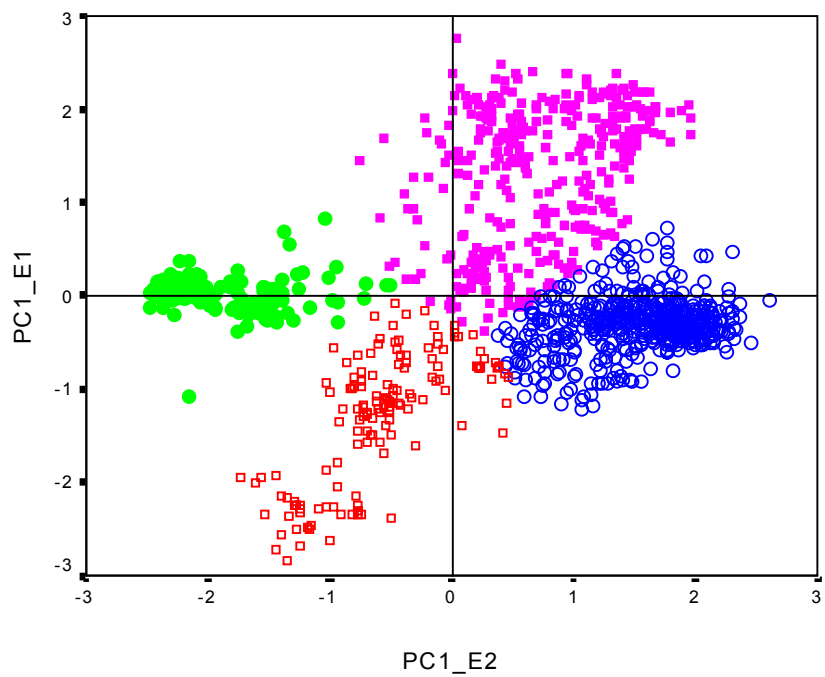
Frequency [kHz]		SEF				NWS			
		SoSm [%]	SoRg [%]	HdSm [%]	HdRg [%]	SoSm [%]	SoRg [%]	HdSm [%]	HdRg [%]
Single	12	75.29	64.75	95.65	83.22	56.2	86.43	83.31	94.19
	38	98.82	89.63	95.03	79.45	82.64	94.29	80.45	84.5
	120	99.22	77.19	94.41	68.84	83.47	97.14	84.58	70.54
Dual	12, 38	99.22	90.55	<sup>a</sup> 98.14	89.04	<sup>a</sup> 92.56	95	84.26	96.9
	12, 120	100	89.86	96.89	<sup>a</sup> 89.73	86.78	97.86	86.17	96.9
	38, 120	100	91.01	96.89	82.53	<sup>a</sup> 92.56	97.86	<sup>a,b</sup> 94.75	95.74
Triplet	12, 38, 120	100	92.63	96.89	89.04	91.74	97.86	93.96	98.06
		<b>100</b>	<b>93.09</b>	<b>98.76</b>	<b>90.41</b>	<b>92.56</b>	<b>98.57</b>	<b>94.28</b>	<b>99.61</b>

Table 5.7. A substantial increase of the percentage of the assigned class occurs when the number of frequencies involved in the analysis increase from 1 frequency (single) to 2 frequencies (dual). For the increasing number of frequencies used in the analysis from 2 frequencies (dual) to 3 frequencies (triplet), however, the percentage of the assigned class only increases slightly and even slightly decreases for 2 seabed classes (HdSm and HdRg for the SEF study area and SoSm and HdSm for the NWS study area) labelled with superscript a (prior to the refinement of the centroids) in Table 5.7.

As discussed above there are several training sites coming from parts of a number of trawl stations. Some of them come from relatively small sections of heterogeneous trawl stations. As such, there are possibilities that seabed types other than that assumed to exist in that particular section may be accidentally extracted. Consequently, the refinement of the centroids of seabed classes is essentially required. The objective is to find sensible centroids for the seabed types in the training sets as well as for the rest of the data sets in order to derive the seabed types along the vessel's track. The refinement of seabed classes was undertaken using a feedback loop as described in section A.2.3. The final centroids after the refinement are given in Table 5.5 and the Euclidean distances between the refined seabed classes are shown in Table 5.6 as bold entries. As shown in Figure 5.14, overlapping between classes disappears after the refinement. In addition, using the refined centroids (given as bold entries in Table 5.5), the percentage of the assigned class increases slightly as shown in Table 5.7 as bold entries, except only for one seabed type (HdSm) in the NWS study area labelled with superscript b in Table 5.7. Class assignment along the vessel's track then uses the final centroids given in Table 5.5 and results along the vessel's track are given in the next section.



(a)



(b)

Figure 5.14. Scatterplot of first principal component of E1 index (PC1\_E1) versus first principal component of E2 index (PC1\_E2) of the refined seabed classes of the training sets for (a) SEF study area and (b) NWS study area.  $\square$  = soft-smooth;  $\circ$  = hard-smooth;  $\blacksquare$  = hard-rough;  $\bullet$  = soft-rough.

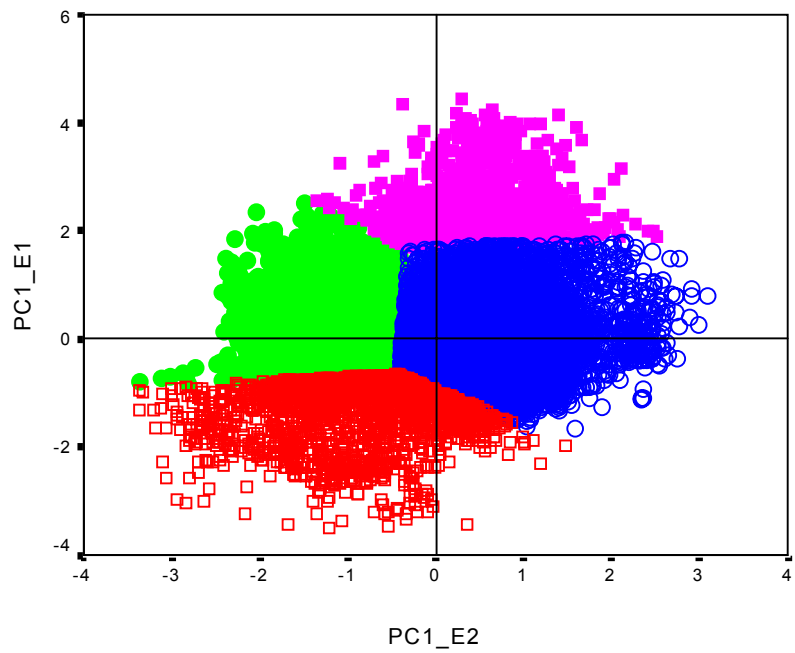
#### **5.4. Along-track seabed classification**

As shown in Figure 5.9, acoustic data available for analysis occupy only part of the entire vessel's track despite the fact that the acoustic instrument was continuously sampling the water column and the seabed. The reasons are threefold. The main reason is that data of the second acoustic bottom returns are available only in part of the total vessel's track. Secondly, malfunction of the acoustic instrument occurred quite often during the operation. Thirdly, performing data quality control as outlined in section 3.1.2 leads to exclusion of all bad acoustic records mainly due to aeration because of bad weather. Acoustic data from cruise ss696 in the SEF region are available for analysis for only 30% of the total vessel's track. For cruise ss895 in the NWS region alone, it is 30% of the total track and for cruise ss797 in the same region, it is only 7% of the total track. Acoustic data available from both cruises in the NWS region in total occupy only 16% of the combined track. The track proportions in which acoustic data are available for analysis are given by dark grey lines in Figure 5.9. It is acknowledged that an alternative use of the leading edge of the first bottom echo (i.e. the specularly reflected part) to the use of the second bottom echo to determine seabed hardness may indeed improve the coverage area. This alternative way, however, is beyond the scope of this current study as it deserves a further, separate study as to how the first and second bottom echoes are related, not to mention the sensitivity of the specularly reflected part of the first bottom echo with pitch and roll of the vessel.

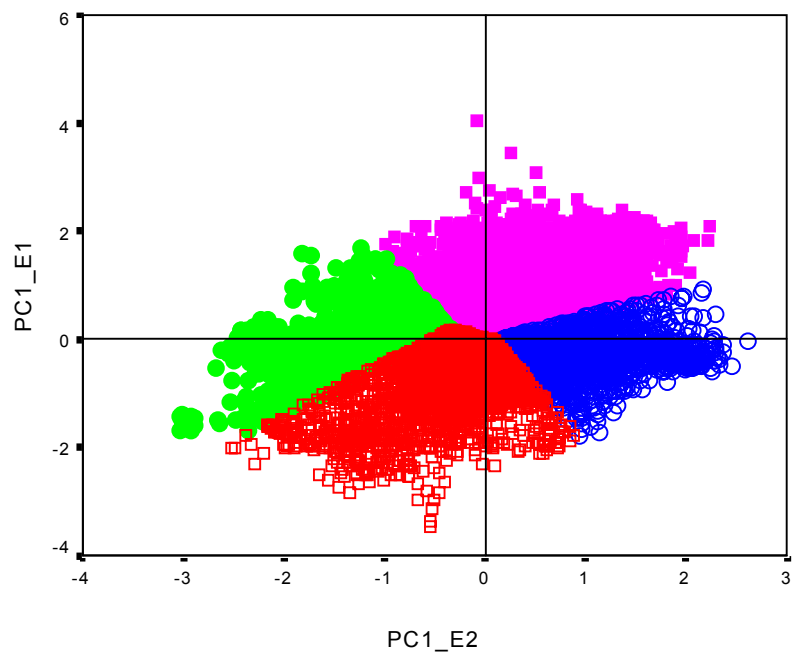
The classification procedure outlined in section A.2.3 was applied to the entire data set for seabed classification along the track using the centroids of the first principal component of roughness index and hardness index (PC1\_E1 and PC1\_E2, respectively) given as bold entries in Table 5.6. Unlike the classification procedure

applied to the training sets, the classification procedure applied to the entire acoustic data involved only class assignments based on the centroids just mentioned without iterative and feedback processes, which are imperative in the training sets to minimise the effect of initial centroids (see section A.2.3). Results of class assignments on the entire data sets are presented as a scatterplot of PC1\_E1 versus PC1\_E2 shown in Figure 5.15. Since new data points are assigned to one particular class they belong to the most, the centroids are subject to change and estimated at the end of the classification process after class assignments have completed. Table 5.8 presents the centroids of the four classes after all data points have been assigned to classes they belong to. It is evident from Table 5.8 that the centroids after all data points have been allocated become smaller than the initial centroids given in Table 5.6. The *Italic* entry inside the bracket shown in Table 5.8 is the absolute value (the absolute PC1\_E1 and PC1\_E2 indices) in a comparable scale to the original roughness and hardness indices at the three frequencies (E1\_12, E1\_38 and E1\_120). The conversion to the absolute PC1 was conducted by using equation (A.1) applied to score coefficients resulting from PCA and roughness and hardness indices of the three frequencies in the original scale. The centroids were then estimated by taking the average of the converted PC1 values for the four derived seabed types. Centroids of the four derived seabed types in both study areas are quite comparable with a maximum difference of 0.3 for the hard-smooth (HdSm) seabed type.

Since the centroids change after class assignments, the Euclidean distances between classes given in Table 5.9 change as well. As the centroids shown in Table 5.8 decrease, so do the Euclidean distances between classes shown in Table 5.9. It is important, however, to note that although the Euclidean distances between classes become smaller, they are, nevertheless, bigger than the Euclidean distances between



(a)



(b)

Figure 5.15. Scatterplot of first principal component of E1 index (PC1\_E1) versus first principal component of E2 index (PC1\_E2) of the entire data together with the four acoustically derived seabed classes for (a) SEF study area and (b) NWS study area.  $\square$  = soft-smooth;  $\circ$  = hard-smooth;  $\blacksquare$  = hard-rough;  $\bullet$  = soft-rough. The boundaries are obtained as a result of the application of the  $k$ -means technique.

Table 5.8. Centroids after all data points have been assigned to the four seabed classes for the SEF study area and the NWS study area separately. The *italic* entry inside the bracket is the absolute value.

Status	Variable	Study Area	Class			
			SoSm	HdSm	HdRg	SoRg
After class assignments of all data points	PC1_E1	SEF	-1.361 <i>(2.612)</i>	.072 <i>(2.969)</i>	2.500 <i>(3.570)</i>	.050 <i>(2.965)</i>
		NWS	-.955 <i>(2.757)</i>	-.284 <i>(2.951)</i>	1.072 <i>(3.411)</i>	-.059 <i>(3.020)</i>
	PC1_E2	SEF	-.898 <i>(4.850)</i>	.622 <i>(5.469)</i>	.515 <i>(5.418)</i>	-1.074 <i>(4.771)</i>
		NWS	-.424 <i>(5.065)</i>	1.073 <i>(5.811)</i>	.389 <i>(5.454)</i>	-1.494 <i>(4.486)</i>

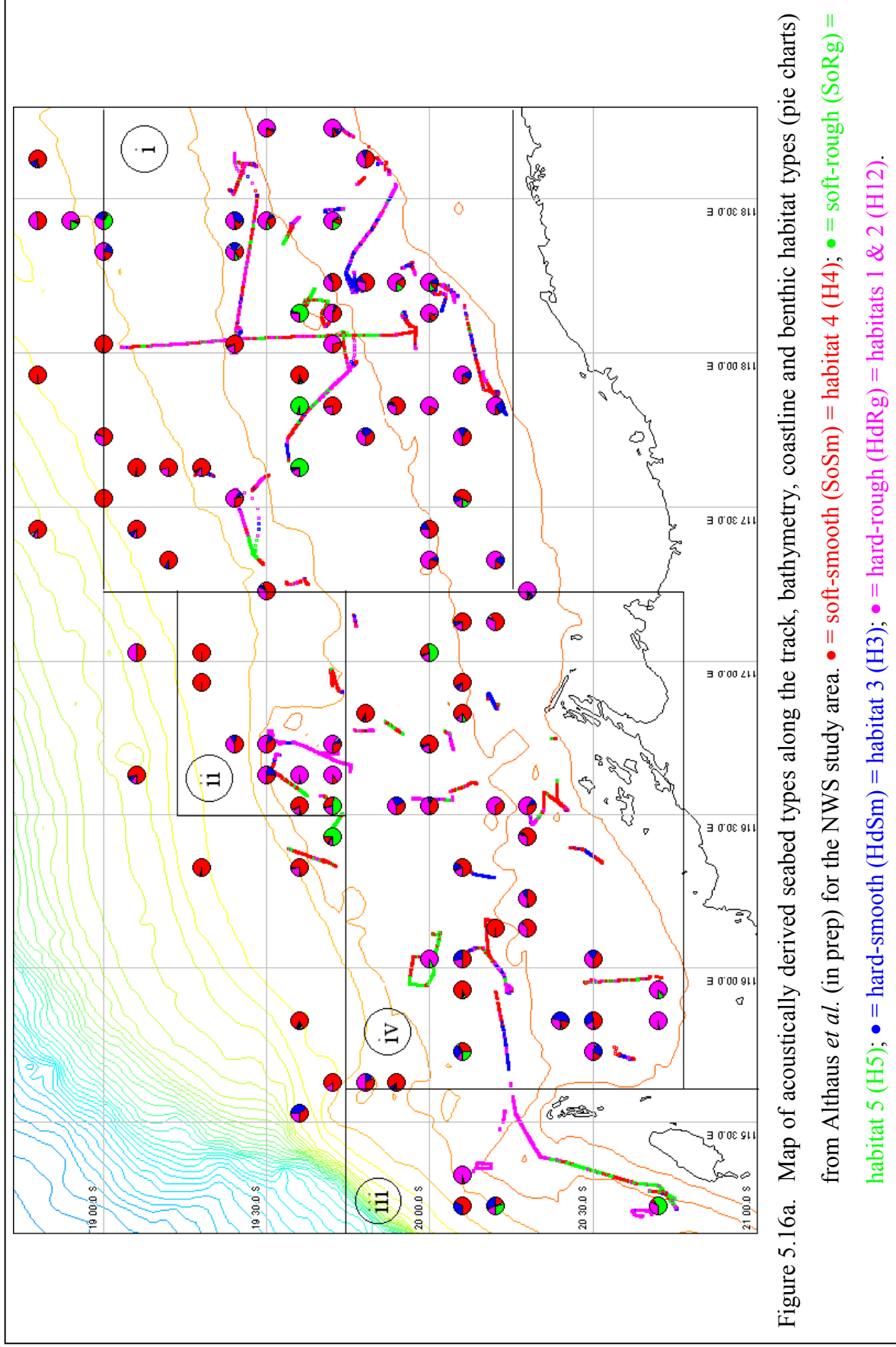
Table 5.9. Euclidean distances between four seabed classes estimated from the centroids after all data points have been assigned to the four seabed classes given in Table 5.8 for the SEF study area and the NWS study area separately.

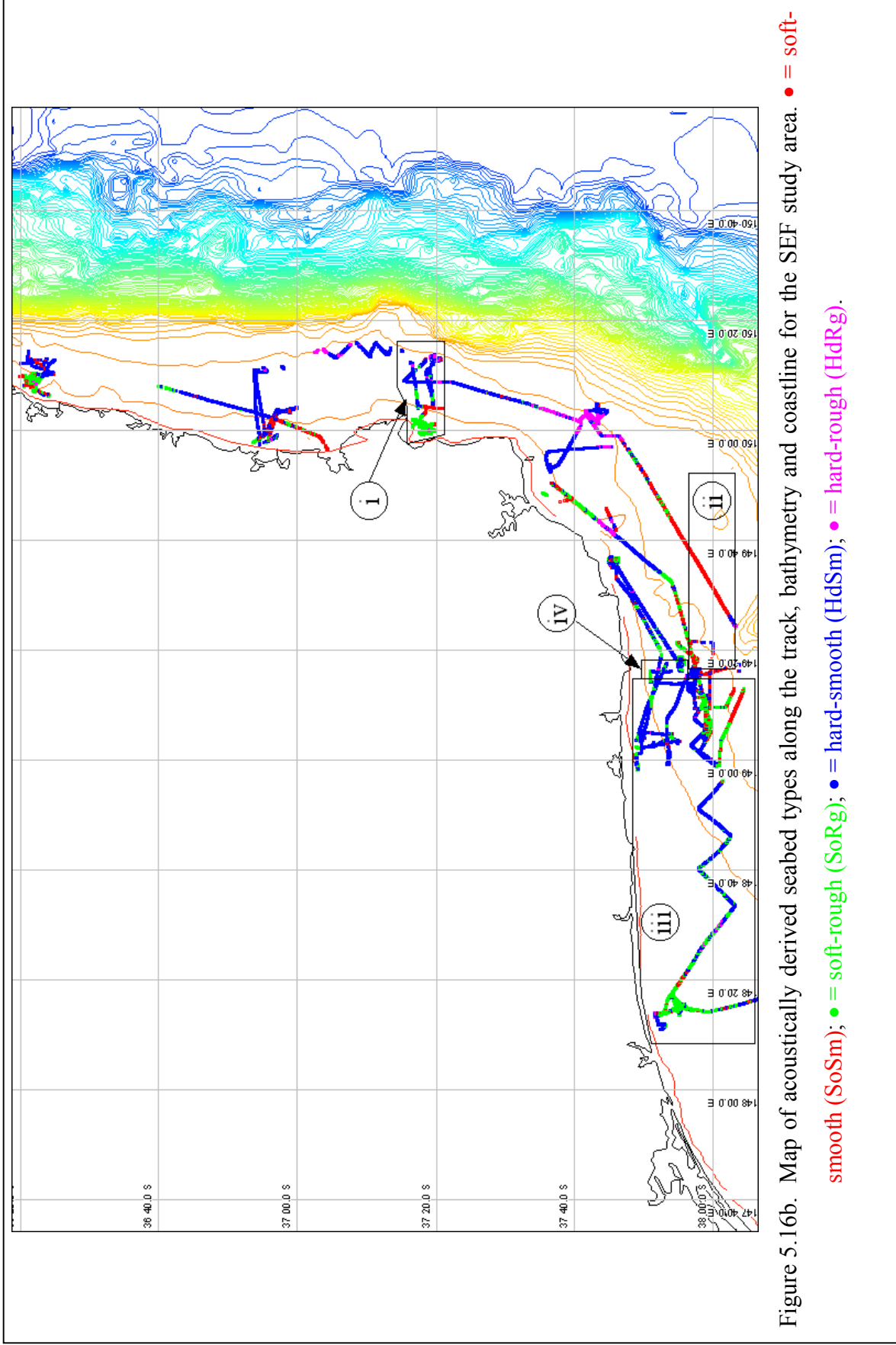
Frequency	Study Area	Class	Euclidean Distance		
			SoSm	HdSm	HdRg
Multi frequencies (PCA)	SEF	HdSm	2.089		
	NWS		1.640		
	SEF	HdRg	4.111	2.430	
	NWS		2.183	1.519	
	SEF	SoRg	1.422	1.696	2.920
	NWS		1.395	2.576	2.196



classes derived from individual frequencies (compare Table 5.9 to Table 5.6 for individual frequencies in particular). This may prove the strength of the multi frequency technique in conjunction with the principal component analysis for seabed classifications over the individual frequency technique and agrees with Kavli *et al.* (1994) who also used the principal component analysis to assess their multi frequency data. Since the principal component analysis involves linear combinations of the original indices from different frequencies and each frequency responds differently for special cases such as those scenarios mentioned in section 5.2, the principal component analysis assigns weights to the original variables that may smear out the ambiguity arising from each individual frequency.

A plot of acoustically derived seabed classes along the vessel's track is shown in Figures 5.16a and 5.16b for the NWS and SEF regions, respectively. There is general agreement between acoustically derived seabed classifications of the NWS region shown in Figure 5.16a, seabed relief of the NWS region as assessed by Jones (1973) shown Figure 5.2a, and benthic habitat types proposed by Sainsbury *et al.* (in prep) in each station as assessed by Althaus *et al.* (in prep) given as pie charts in Figure 5.16a. Sainsbury *et al.* (in prep) extracted 44 features (variables) from 46,419 seabed images in trawl stations taken between 1983 and 1997, and used 46,201 seabed images for further analysis (Sainsbury *et al.*, in prep). Five binary (presence/absence) variables were prepared from the 44 extracted features and introduced to the hierarchical agglomerative clustering technique to produce 5 habitat types (Sainsbury *et al.*, in prep): habitat 1/H1 equivalent to hard-rough/HdRg seabed type (6954 images); habitat 2/H2 equivalent to hard-rough/HdRg seabed type (10138 images); habitat 3/H3 equivalent to hard-smooth/HdSm seabed type (6491 images); habitat 4/H4 equivalent to soft-smooth/SoSm seabed type (18967 images); habitat





5/H5 equivalent to soft-rough/SoRg seabed type (3651 images). Based on these results, Althaus *et al.* (in prep) grouped seabed images into trawl stations they belong to and then produced an appropriate chart for each trawl station. Both acoustically derived seabed classes along the vessel's track and pie charts use the same colour definition: red is soft-smooth (SoSm), green is soft-rough (SoRg), blue is hard-smooth (HdSm) and magenta is hard-rough (HdRg).

To assist the following discussion, the NWS study area was divided into four areas labelled as i, ii, iii and iv in Figures 5.2a and 5.16a. Each area highlights the dominant seabed type and feature based on Figure 5.2a. The total area of these derived areas is shown in Table 5.10. Given also in Table 5.10 are the areas of the acoustic and the photographic coverage within derived areas (i), (ii), (iii) and (iv). As shown in Table 5.10, the acoustics gave better coverage than the photographs in all derived areas even though both acoustic coverage and photographic coverage are very low. The photographs, on the other hand, are better than the acoustic results in terms of spatial distribution in areas other than area (iii) (see Figure 5.16a). Total number of grab samples used to produce Figure 5.2a is not known but it is likely that coverage by such direct sampling would be at a level below either of the figures in Table 5.10.

In area (i) of Figure 5.16a, both acoustically derived bottom classifications and habitat types suggest that the area is mainly hard-rough surfaces whereas Figure 5.2a describes the area as comprising smooth surfaces. Figure 5.2a, however, indicates the presence of rough surfaces in the entire area (i) as well. The discrepancy may be due to different assumptions adopted for the definition of hard-rough surfaces. While in this study, lumps of outcrops of vegetation are considered components of rough-hard surfaces, they, on the other hand, might not be considered rough-hard surfaces in

Table 5.10. The total area, the area of acoustic coverage and the area of photographic coverage in m<sup>2</sup> within four derived areas (i), (ii), (iii) and (iv) shown in Figures 5.2a and 5.16a. Entries inside the bracket are the percentage of coverage relative to the total area.

Category	Area [m <sup>2</sup> ]			
	(i)	(ii)	(iii)	(iv)
Total	22394314000	4057748100	6114108600	18725497000
Acoustic	23877198 (0.1066%)	4427144 (0.1091%)	4853410 (0.0794%)	12465430 (0.0666%)
Photograph	111863.7 (0.0005%)	18892.177 (0.0005%)	6154.5525 (0.0001%)	55291.449 (0.0003%)

Figure 5.2a. However, if hard-smooth and soft-smooth classes from this study are loosely considered smooth surfaces, the along track seabed classification in this area consists mainly of smooth surfaces. Similarly, smooth surfaces become dominant in this area if these two classes in habitat types are combined. In addition, results from the two previous studies identify the presence of sandwaves in this area that may be associated with soft-rough surfaces identified in the similar area here in this study. Results from this study show that hard-rough, smooth (hard-smooth and soft-smooth), soft-rough surfaces that may be attributed to sandwaves account for 33.5%, 57% and 9.5% of the total track in area (i), respectively. For the similar area, habitat types from Althaus *et al.* (in prep) given as pie charts in Figure 5.16a suggest that hard-rough, smooth (hard-smooth and soft-smooth), soft-rough surfaces account for 41%, 50% and 9% of the total area, respectively. A 2×2 contingency table was prepared to test the distribution of acoustically derived seabed types and habitat types just mentioned. The statistical test, however, showed that the acoustically derived seabed types and the habitat types are significantly different at  $\alpha=0.01$  ( $p \ll 0.01$ ).

In area (ii), results from the three studies agree well in terms of the dominant class and the ranked order of the class. Results from this study show that rough surfaces and smooth surfaces account for 66% and 34% of the total track in area (ii), respectively and seem to agree well with the habitat types where rough surfaces account for 52% of area (ii) and smooth surfaces account for 48%. Of 52% rough surfaces from habitat types, only 0.1% were associated with soft-rough surfaces that may be attributed to sandwave. Similarly, results from Jones (1973) given in Figure 5.2a shows that no sandwaves are identified. On the other hand, results from this study indicate the presence of sandwaves in the area. Of 66% rough surfaces in area (ii), 6% are associated with soft-rough, which may be attributed to sandwaves or

muddy, bioturbated surfaces. Although the dominant class and the ranked order of the class from the acoustically derived seabed types and the habitat types may agree, the statistical test similar to the one applied in area (i) reveals that the frequency distribution of them is significantly different at  $\alpha=0.01$  ( $p \ll 0.01$ ).

Unlike areas (i) and (ii), area (iii) is dominated by smooth surfaces. Results from this study reveal that soft-smooth surfaces account for 45% of the total track in area (iii) whereas the habitat types show that these seabed surfaces account for 36% of the area (iii). Hard-rough surfaces, however, occur in area (iii) as well. From this study, hard-rough surfaces account for 32% of the total track in area (iii) whereas from habitat types, they account for 34% of the area (iii). The percentage of soft-rough surfaces in this area is higher than that in area (i). While results from this study show that soft-rough surfaces account for 21% of the total track in area (iii), habitat types show that soft-rough surfaces account for 15% of the area (iii). The statistical test though proved that the frequency distribution of the acoustically derived seabed types and the habitat types was not significantly different at  $\alpha=0.01$  ( $p = 0.157$ ) or in other words the frequency distribution of them are similar. Results from these two studies agree well with results from Jones (1973) given in Figure 5.2a. It is clear from Figure 5.2a that the area, particularly in the bottom left, consists entirely of smooth surfaces. The area of smooth surfaces extends eastward and northward and is bounded by rough surfaces to the east and by sandwaves to the north.

Similar to area (iii), area(iv) consists mainly of smooth surfaces. Seabed relief as assessed by Jones (1973) shown in Figure 5.2a shows that area (iv) comprises mainly smooth surfaces with very few rough surfaces occurring in the west region adjacent to area (iii). Results from this study and habitat types as assessed by Althaus

*et al.* (in prep) support the seabed relief shown in Figure 5.2a. Results from this study show that soft-smooth surfaces occupy 34% of the total track in area (iv) and hard-rough surfaces occur in 28% of the total track in area (iv). In comparison, habitat types suggest that soft-smooth and hard-rough surfaces occupy 47% and 38% of the area (iv), respectively. The statistical test suggested that the frequency distribution of the results from this study and those from Althaus *et al.* (in prep) was similar or was not significantly different at  $\alpha=0.01$  ( $p = 0.8045$ ).

Due to data availability and accessibility, a direct comparison similar to that for the NWS study area is not available for the SEF study area. While previous studies in the NWS study area offer a small-scale view for instance Sainsbury, *et al.* (in prep) and Althaus *et al.* (in prep) for benthic habitat types, those in the SEF study area offer a relatively large-scale view. The comparison method adopted here for the NWS study area is unfortunately not available for the SEF study area. The comparison for the SEF study area is therefore based on a relatively large-scale view.

Unlike the NWS study area, the SEF study area consists mainly of hard-smooth surfaces due to the fact that coarse sediment and shell debris (Jones and Davies, 1983; Kloser *et al.*, 2001) exist in a large proportion as far as the current study area in the SEF region is concerned. They account for 57% of the total track in the entire study area. As shown in Figure 5.16b, soft-rough surfaces also occur in a quite large proportion of the total track (24%). Soft-rough surfaces that may be attributed to sandwaves, in general, occur on the inner shelf, particularly in areas adjacent to the coastline.

Figure 5.16b shows that the outer shelf in area (i) is dominated by hard-smooth surfaces with very few hard-rough surfaces at the outermost shelf whereas the



inner shelf consists mainly of soft-rough surfaces that may be attributed to sandwaves. The presence of soft-smooth surfaces in area (i) is also noticeable. They occur in the central region. Results from this study for this area seem to agree with results from Kloser *et al.* (2001) who found that this area consists mainly of sands with various grainsizes. They observed coarse and medium sands in the outer shelf, very fine sands in the central and fine sands in the inner shelf.

In area (ii), results from this study agree with the surface sediments as assessed by Jones and Davies (1983) as shown in Figure 5.2b. As far as the technique adopted in this study is concerned, the difference between very fine and fine sands, and muddy sediments are not distinguishable but they all are considered soft surfaces. Therefore, soft-smooth surfaces in area (ii) shown in Figure 5.16b might represent the dominant muddy quartzose and calcareous sediments shown in Figure 5.2b.

In area (iii) on the other hand, hard-smooth surfaces are dominant and account for 56% of the total track in the area. As mentioned earlier, this might be due to the presence of the coarse sediment and shell debris in almost all of area (iii) as shown in Figure 5.2b (Jones and Davies, 1983). Although Figure 5.2b does not indicate whether or not sandwaves occur in this area, results from this study, nevertheless, suggest that soft-rough surfaces that may be attributed to sandwaves are present in the area and account for 34% of the total track in the area. Since irregular muddy surfaces due to bioturbation are also considered a sandwave-like surfaces, the presence of soft-rough surfaces in area (iii) might also be an indication of the presence of irregular surfaces of muddy sediments due to bioturbation. This seems to support Jones and Davies (1983) who observed the presence of irregular surfaces of muddy sediments as a result of bioturbation in this area.

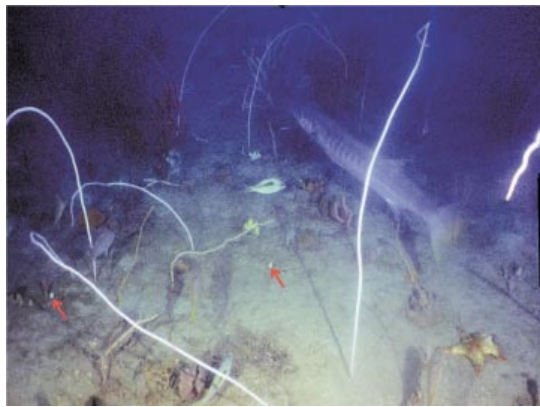
Although area (iv) is dominated by hard-smooth surfaces, hard-rough surfaces are also identified in this area. Hard-rough surfaces account for 6% of the total track in the area. Hard-rough surfaces that might be due to the presence of outcrops in this area are consistent with Jones and Davies (1983) who indicated the presence of reefs and coral reefs in this area as shown in Plate 1 of their report.

From Figures 5.16a and 5.16b, there are two obvious differences between the two study areas. Firstly, distances over which variations of along-track bottom types occurred are shorter in the NWS study area than in the SEF study area. This indicates that the local variations are greater in the NWS study area than in the SEF study area. Secondly, the NWS study area seems to have more hard-rough surfaces than the SEF study area. This might indicate that the NWS study area would have more areas in which reefs and coral reefs exist than the SEF study area would have. These differences might be related to the fact that the NWS study area is located in the tropics whereas the SEF study area is in the temperate and higher latitude regions. A high diversity and a moderate abundance of resources are characteristic of the tropical region. The temperate and higher latitude region is on the other hand characterised by a lower diversity and a higher abundance of resources. A higher local variation of the bottom type in the NWS study area might be an indication of a higher diversity of resources which is expected to occur in the tropics (Clarke and Crame, 1997; Rosenzweig, 1995). In addition, the presence of coral reefs decreases as the latitude moves away from the equator, north or south, or in other words more coral reefs are expected in the tropics than in the temperate and higher latitude regions (Sorokin, 1993; Clarke and Crame, 1997).

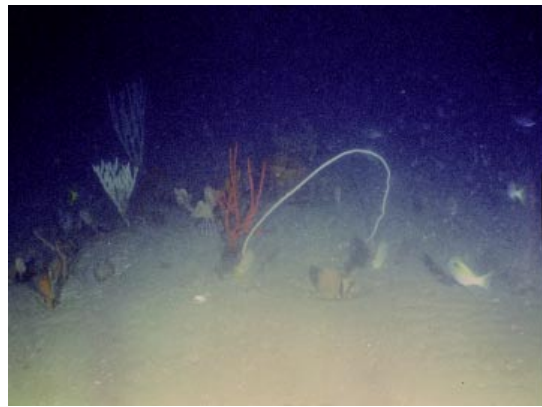
## 5.5. Ground truth versus acoustically derived seabed type

For the NWS study area, the acoustically derived seabed types are compared with benthic habitat types due to Althaus *et al.* (in prep). Representative examples of their benthic habitat types are shown in Figure 5.17. The comparisons are made only at trawl stations where both acoustically derived seabed types and benthic habitat types are both available as shown as stars in Figure 5.9(a). To make the two data sets comparable, benthic habitat types 1 and 2 are combined and are considered as hard-rough seabed surfaces whereas benthic habitat types 3, 4 and 5 are considered equivalent to hard-smooth, soft-smooth and soft-rough seabed surfaces, respectively.

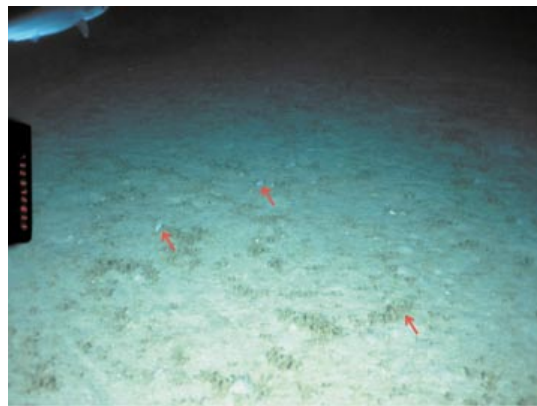
In Figure 5.18, histograms of acoustically derived seabed types and benthic habitat types are presented. The histograms represent frequency of occurrence (in percentage) of acoustically derived seabed types and their equivalent benthic habitat types within each trawl station given as stars in Figure 5.9(a). Classes 1, 2, 3 and 4 shown in both histograms are respectively soft-smooth (=habitat 4), soft-rough (=habitat 5), hard-smooth (=habitat 3) and hard-rough (=habitats 1 and 2). As far as the dominant seabed (or benthic habitat) types shown in Figure 5.18 are concerned, the acoustically derived seabed types agree very well with benthic habitat types (Table 5.11). Despite the agreement between acoustically derived seabed types and benthic habitat types just mentioned, the results also show that they differ in terms of seabed types (or benthic habitat types) other than the dominant ones present in each trawl station. The extreme ones happen in trawl stations 3, 6, and 12 where there is only one acoustically derived seabed type but there are more than 3 significant habitat



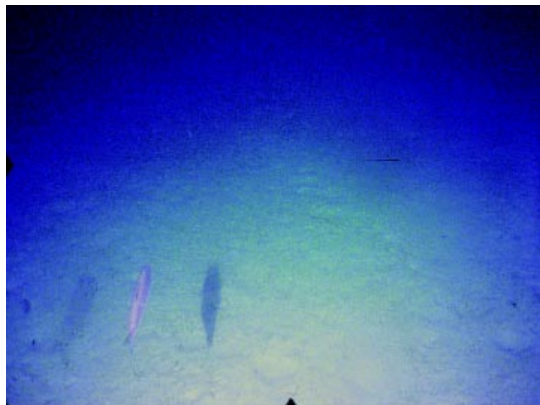
Habitat 1



Habitat 2



Habitat 3



Habitat 4



Habitat 5

Figure 5.17. Seabed images representing five benthic habitat types in the NWS study area proposed by Sainsbury *et al.* (in prep). Habitats 1 and 2 correspond to class HdRg; Habitat 3 corresponds to class HdSm; Habitat 4 corresponds to class SoSm; Habitat 5 corresponds to class SoRg. Images taken after Sainsbury *et al.* (in prep).

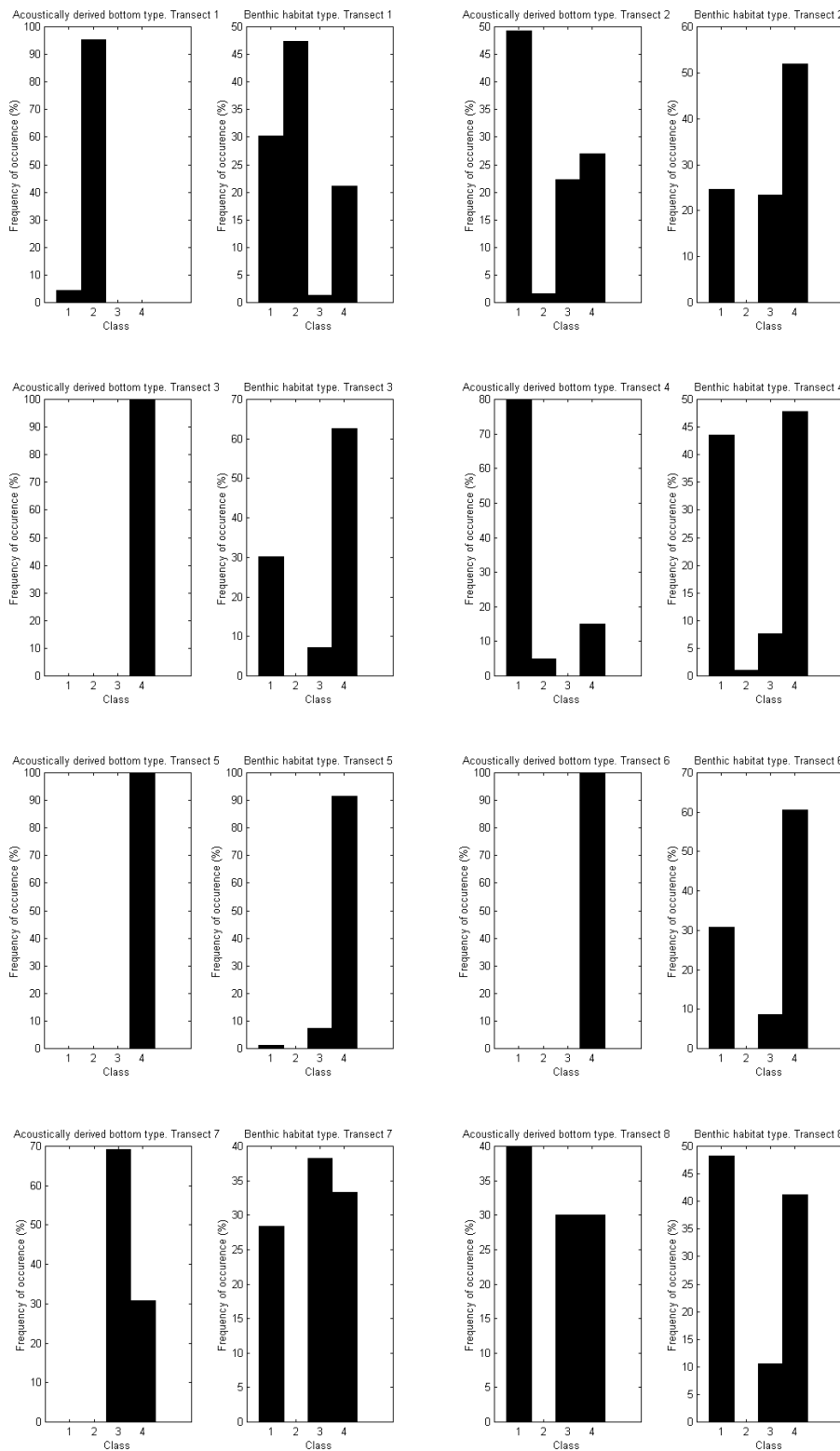
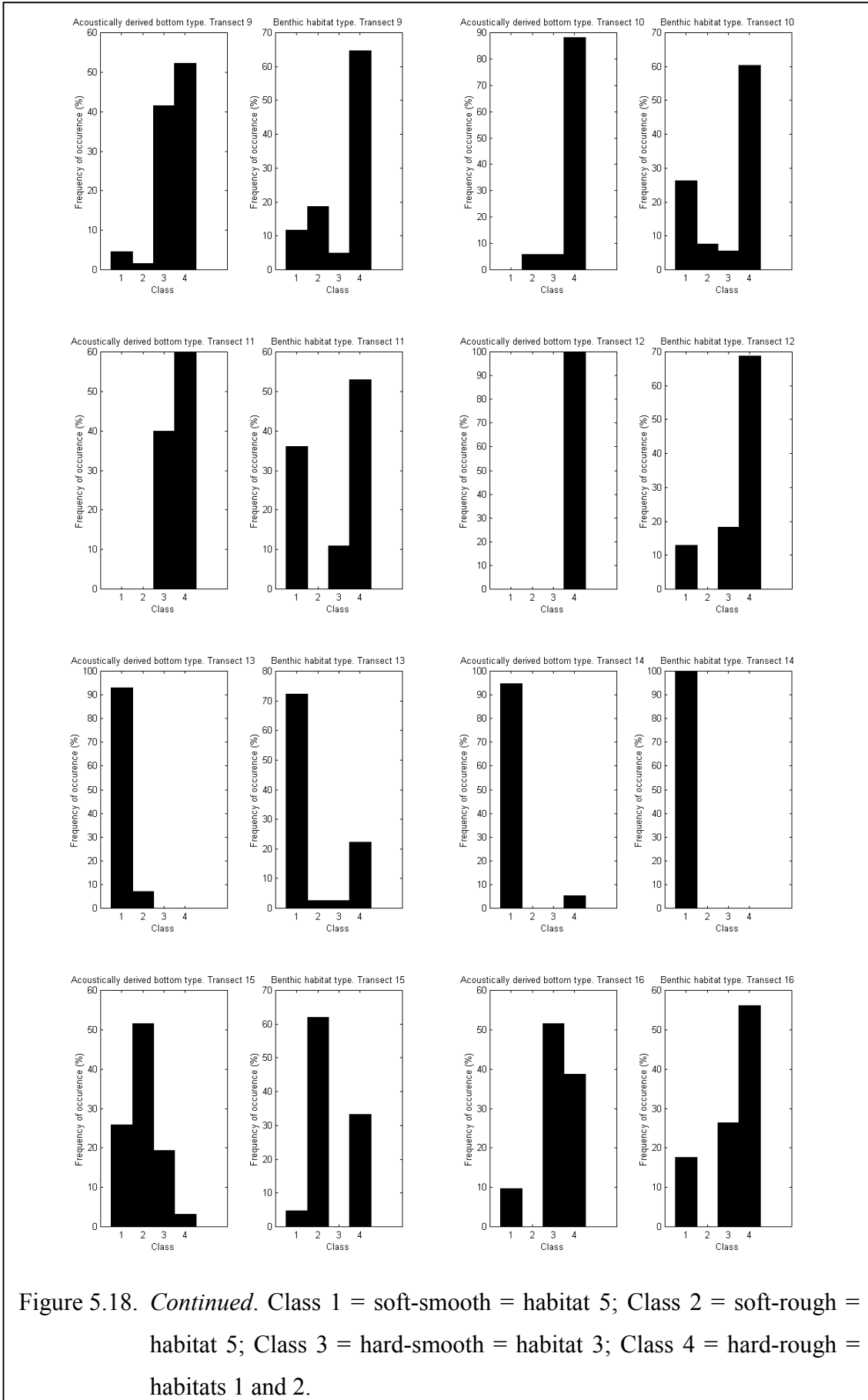


Figure 5.18. Histograms of acoustically derived bottom types and benthic habitat types for the NWS study area. Class 1 = soft-smooth = habitat 5; Class 2 = soft-rough = habitat 5; Class 3 = hard-smooth = habitat 3; Class 4 = hard-rough = habitats 1 and 2.



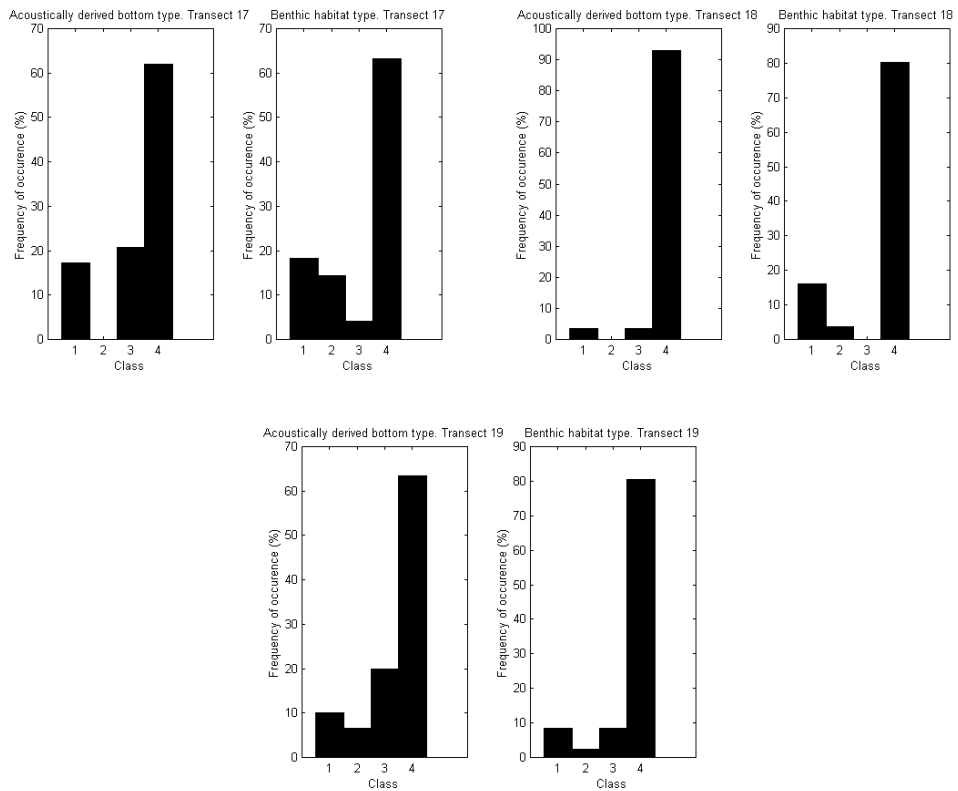


Figure 5.18. *Continued.* Class 1 = soft-smooth = habitat 5; Class 2 = soft-rough = habitat 5; Class 3 = hard-smooth = habitat 3; Class 4 = hard-rough = habitats 1 and 2.

Table 5.11. Confusion matrix between acoustically derived seabed types and benthic habitat types for the NWS study area. I, II, III, IV are classes given in Figure 5.18; H = habitat; SoSm = soft-smooth; SoRg = soft-rough; HdSm = hard-smooth; HdRg = hard-rough.

Acoustically derived seabed types		Benthic habitat types			
		I	II	III	IV
		(H4=SoSm)	(H5=SoRg)	(H3=HdSm)	(H12=HdRg)
I (SoSm=H4)	3			2	
II (SoRg=H5)		2			
III (HdSm=H3)			1	1	
IV (HdRg=H12)				10	

Table 5.12. Confusion matrix between acoustically derived seabed types and seabed types of reference sites for the SEF study area. I, II, III, IV are classes given in Figure 5.19; H = habitat; SoSm = soft-smooth; SoRg = soft-rough; HdSm = hard-smooth; HdRg = hard-rough.

Acoustically derived seabed types		Reference sites			
		I	II	III	IV
		(H4=SoSm)	(H5=SoRg)	(H3=HdSm)	(H12=HdRg)
I (SoSm=H4)	2				
II (SoRg=H5)		5			
III (HdSm=H3)			6	1	
IV (HdRg=H12)				1	



types. There are only 4 trawl stations in which the presence of acoustically derived seabed types exactly match their habitat type counterpart. They are trawl stations 8, 9, 16 and 18.

Table 5.11 shows the confusion matrix between the dominant acoustically derived seabed types and the dominant benthic habitat types from histograms given in Figure 5.18. It is evident that acoustically derived seabed types agree with benthic habitat types in 16 trawl stations out of 19. There are only 3 trawl stations where acoustically derived seabed types disagree with benthic habitat types. They are trawl stations 2 and 4 in which they are acoustically referred to as soft-smooth seabed surfaces but as rough-hard seabed surfaces by benthic habitat category, and trawl station 16 where it acoustically belong to hard-smooth seabed surfaces but to hard-rough seabed surfaces by benthic habitat category.

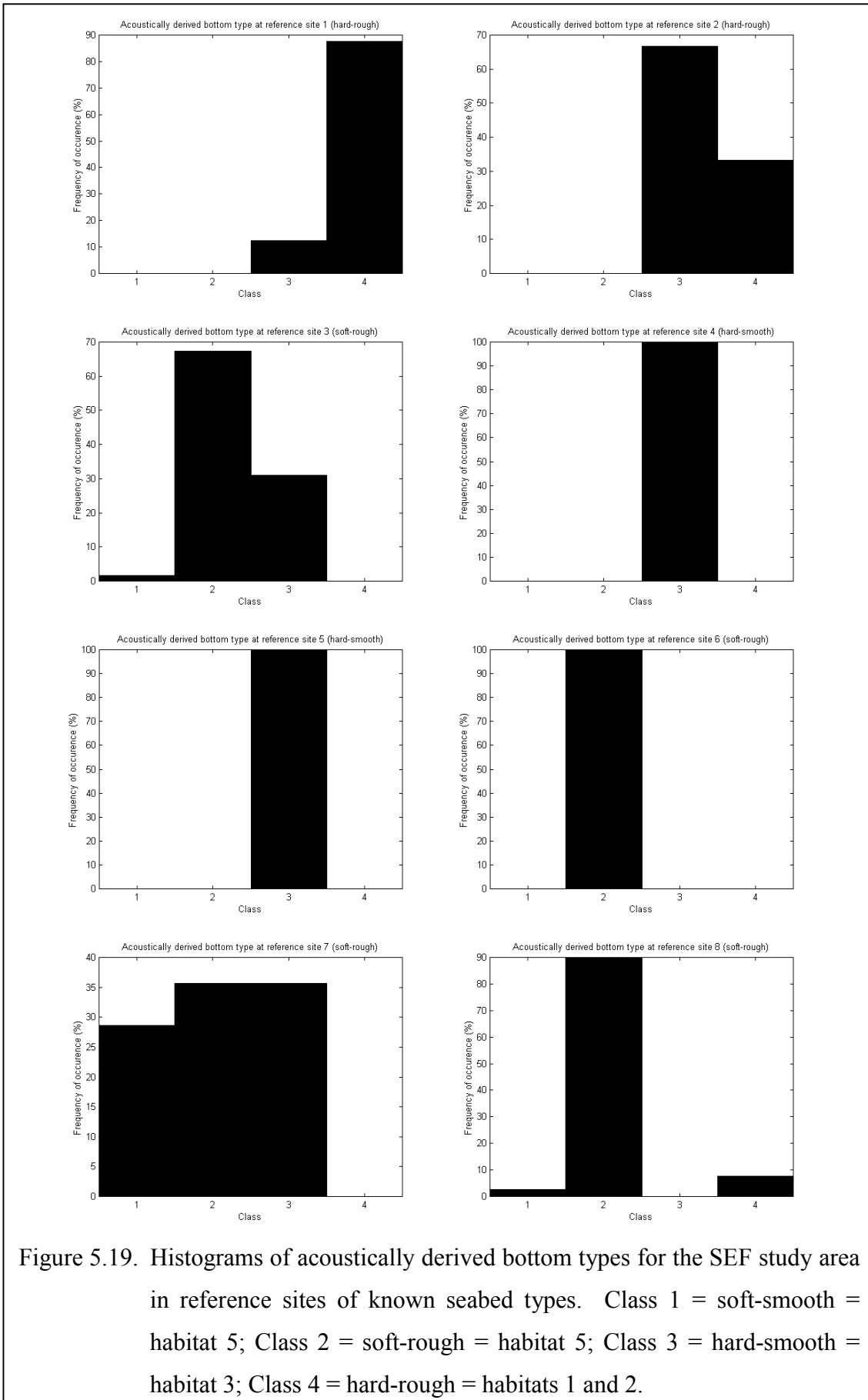
Analysis of the dominant classification was adopted using the hypothesis of marginal homogeneity (Woodward *et al.*, 1990). Wald statistical test was applied to the data in Table 5.11 to test the null hypothesis,  $H_0$ , defined as probability that dominant seabed type is class  $i$  by acoustically derived classification equals probability that dominant seabed type is class  $i$  by benthic habitat classification. The test revealed that the null hypothesis was not rejected ( $p=0.313$ ), indicating that probability of a seabed type derived acoustically and that from benthic habitat classification are similar.

While the seabed types of the training sites in the NWS study area are based on the photographs taken within each training site, the seabed types of the reference sites in the SEF study are well known and the photographs taken in reference sites are used only for confirmation. The seabed types of reference sites given as stars in

Figure 5.9(b) are as follows. Reference sites that belong to hard-rough seabed types are 1 and 2. Reference sites 3, 6, 7, 8 and 15 belong to soft-rough seabed types. Reference sites 4, 5, 9, 10, 11, and 13 consist of hard-smooth seabed types. Reference sites 12 and 14 are mainly soft-smooth seabed types.

Shown in Figure 5.19 is the histogram of the acoustically derived seabed types within each reference site in the SEF study area. Like Figure 5.18, the histogram shows the occurrence frequency of the acoustically derived seabed type in each reference site. Taking the dominant acoustically seabed types from the histograms in Figure 5.19 to indicate the most likely seabed type in each reference site, confusion matrices between acoustically derived seabed types and known reference sites are established and are given in Table 5.12. Both Figure 5.19 and Table 5.12 show that there is good agreement between acoustically derived seabed types and the seabed types of the reference sites. There is only one reference site (reference site 2) where they disagree. Although from previous knowledge reference site 7 belongs to soft-rough seabed types, the acoustically derived seabed type obtained here suggests that this site is not homogeneously soft-rough but rather comprises two other different seabed types namely soft-smooth and hard-smooth in a quite similar frequency of occurrence.

Suppose that one particular site is considered homogeneous if there are at the most 2 different seabed types present; histograms in Figure 5.19 suggest that 9 sites out of 15 are homogeneous whereas histograms in Figure 5.18 indicate only 9 sites out of 19. This again confirms that the local variation of seabed types in the NWS region is higher than that in the SEF region as already mentioned in section 5.4.



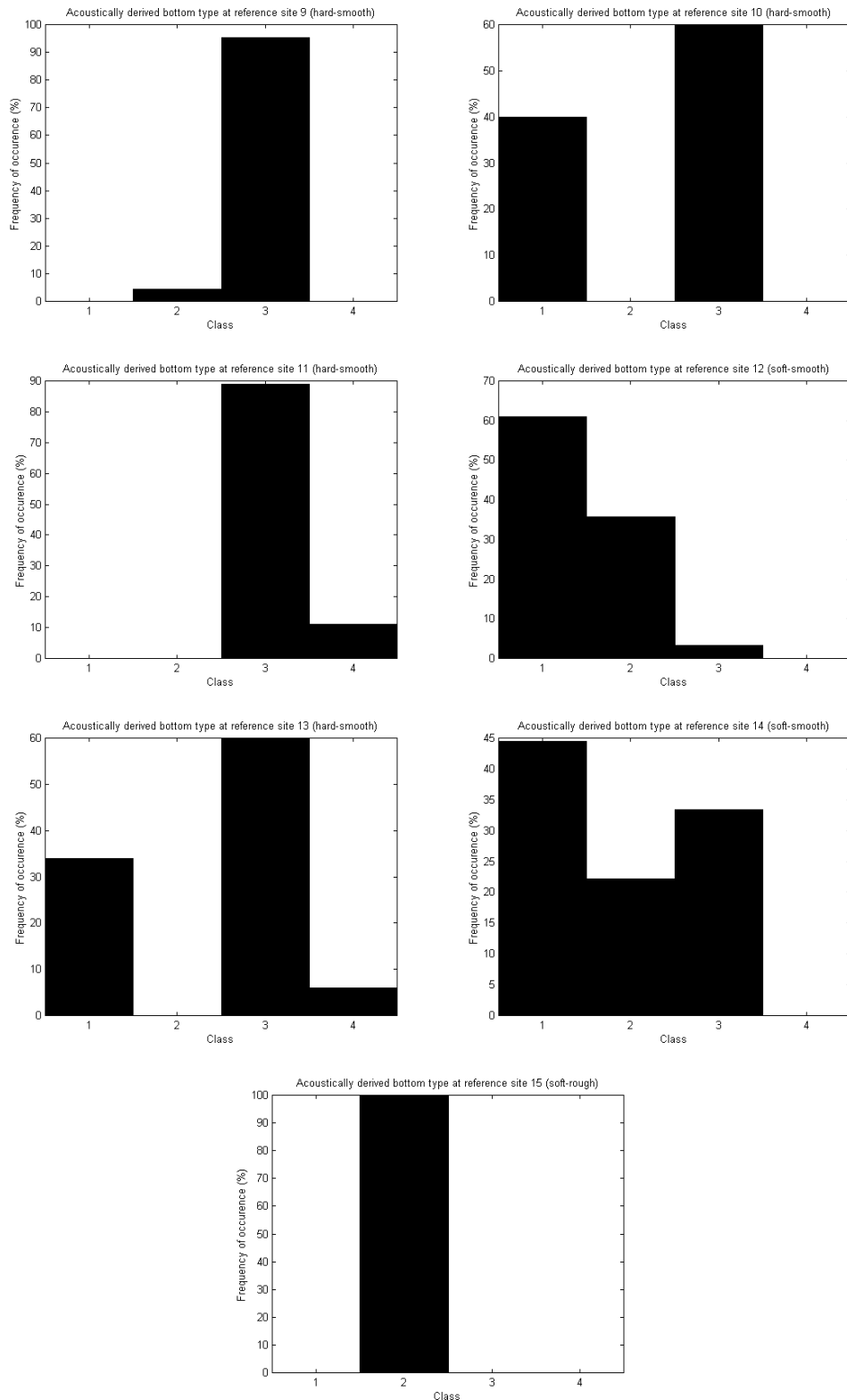
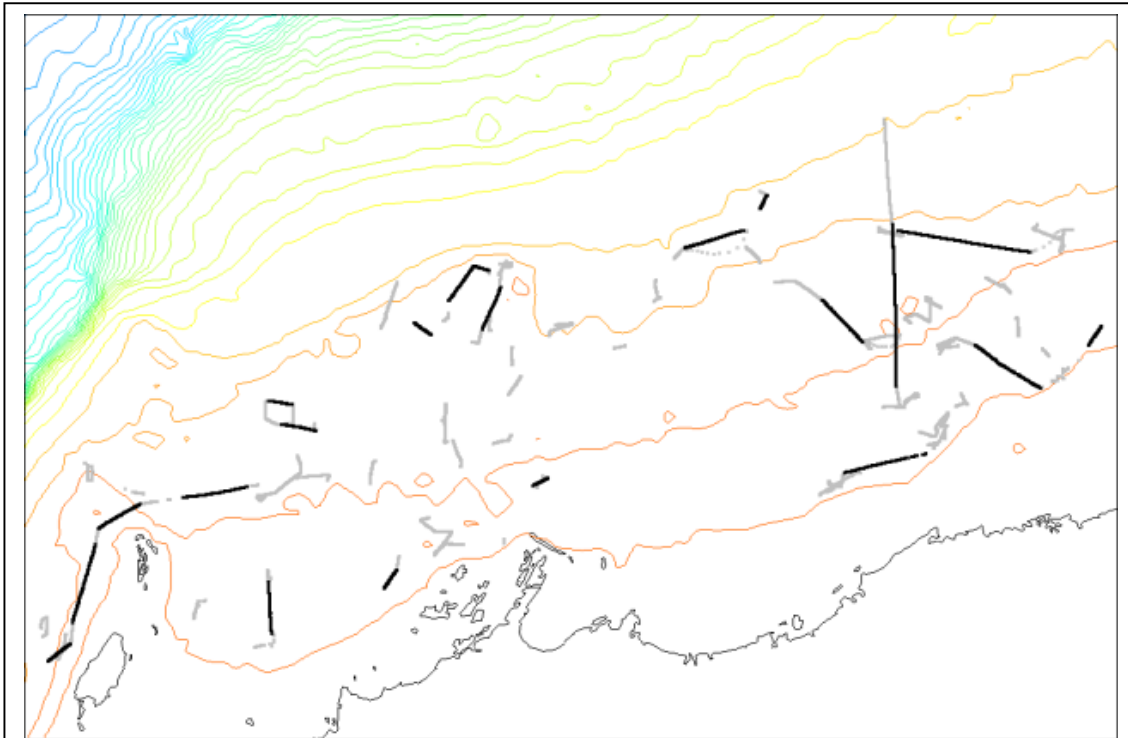


Figure 5.19. *Continued.* Class 1 = soft-smooth = habitat 5; Class 2 = soft-rough = habitat 5; Class 3 = hard-smooth = habitat 3; Class 4 = hard-rough = habitats 1 and 2.

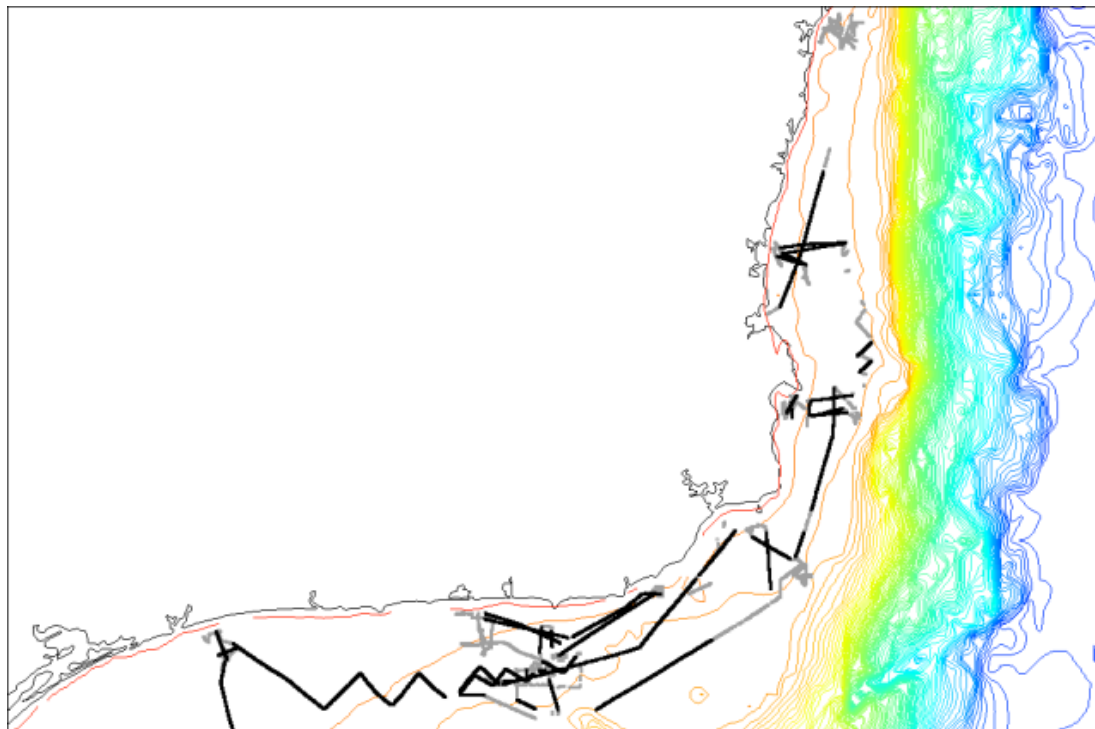
## 5.6. Kriged spatial modelling

In many oceanographic studies, contour maps of various parameters are produced quite often from undersampled data. This dilemma often arises associated with the costs involved in oceanographic sampling. Often too, careful interpolation between measurement sites is called for to minimise the value of relatively sparse measurements. Kriging is one of the interpolation techniques suitable for oceanographic studies since it includes the range of influence (or the autocorrelation characteristic length), i.e. lengths in which attributes of data are self correlated, resulting from a variogram in the interpolation analysis. Autocorrelation analysis was also adopted here to produce the autocorrelation characteristic length as a complementary of the range of influence obtained from the variogram as a matter of assurance. While the Kriging technique does not require either equally spaced or equally directional data, the autocorrelation analysis does require so. To accommodate both the autocorrelation and the Kriging techniques, hence the equally spaced, directional data i.e. linear interpolated straight transects were used.

As described in section 1.6, vessel tracks were first divided into as many straight transects as possible in as many directions as possible. The idea of such division is simply to ease the observation of trend-direction dependency. Figures 5.20(a) and (b) show all straight transects available for the NWS and SEF study areas, respectively. First and second acoustic bottom returns of the three operated frequencies from these straight transects were extracted. Moreover, coordinates in Longitudes and Latitudes were converted into the new ones in Eastings and Northings, respectively. Due to the fact that data of the acoustic bottom returns were not equally spaced, it was necessary to derive equally spaced data in each straight transect. To produce equally spaced data of the first and second acoustic bottom



(a) NWS study area

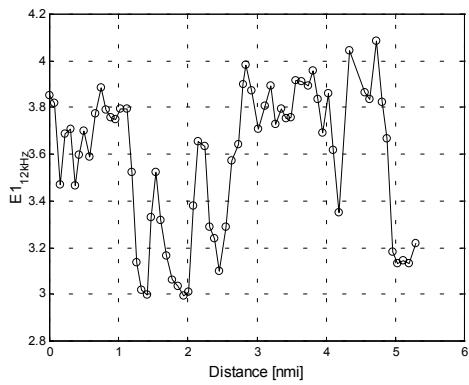


(b) SEF study area

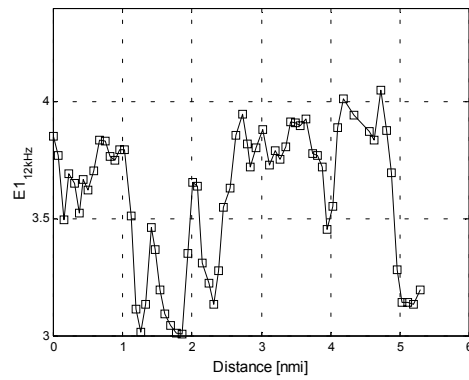
Figure 5.20. Map of all the straight transects (black lines) extracted from the useful track (grey lines) used in the autocorrelation analysis and the interpolation into unsampled areas by Geostatistics (variogram and Kriging) for (a) the NWS study area and (b) the SEF study area.

returns in each transect, a linear interpolation technique was applied to the unequally spaced data of the first acoustic bottom returns and likewise to the unequally spaced data of the second acoustic bottom returns in each transect. Figures 5.21 and 5.22 show representative examples of (a) the original series and (b) the interpolated series of respectively the first and second acoustic bottom returns at three operated frequencies; (i) 12 kHz, (ii) 38kHz and (iii) 120 kHz.

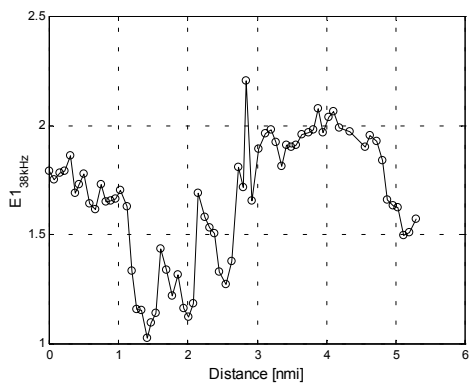
Prior to the analysis of ACF, PCA was applied to roughness index (E1) of the three operated frequencies in each straight transect. Similarly, PCA was applied to hardness index (E2) of the three operated frequencies in each straight transect. Figure 5.23 shows the scree plot of E1 and E2 given in Figures 5.21 and 5.22, respectively. It is evident from Figure 5.23 that the first principal component of (a) roughness index and (b) hardness index (PC1\_E1 and PC1\_E2) is sufficient to represent most of the variation in the original roughness and hardness indices. For this particular example, PC1\_E1 accounts for 89% of the total variation of the original E1 data whereas PC1\_E2 accounts for 80% of the total variation of the original E2 data. The component loadings of PC1\_E1 are 0.92, 0.96 and 0.92 for E1\_12, E1\_38 and E1\_120, respectively and the component loadings of PC1\_E2 are 0.85, 0.98 and 0.84 for E2\_12, E2\_38 and E2\_120, respectively, suggesting that the first principal component (PC1) of roughness and hardness indices is essentially an average. These indicate a high correlation between variables of the PC1\_E1 and their corresponding variables of the original E1 and similarly between variables of the PC1\_E2 and variables of the original E2. Series of PC1\_E1 and PC1\_E2 of this particular transect are shown in Figures 5.24(a) and (b), respectively. It was found that the total variance of PC1\_E2 of all straight transects was never lower than 70%. In contrast, the total variance of PC1\_E1 of some of the straight transects was lower than 50%. The



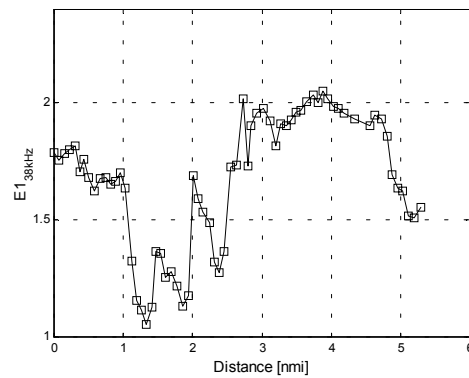
(i) 12 kHz



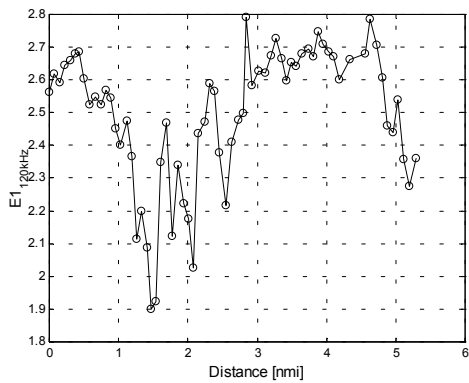
(i) 12 kHz



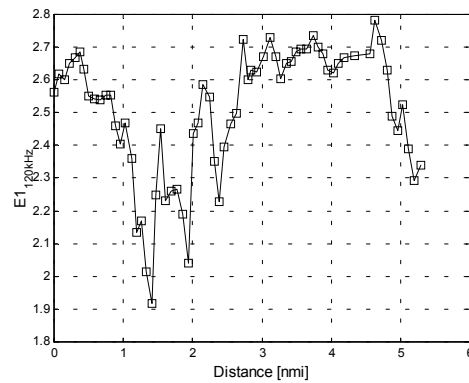
(ii) 38 kHz



(ii) 38 kHz



(iii) 120 kHz



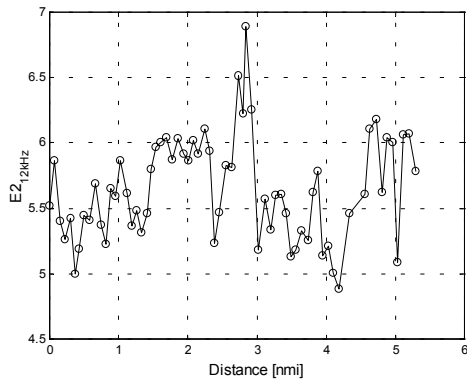
(iii) 120 kHz

(a) Original series

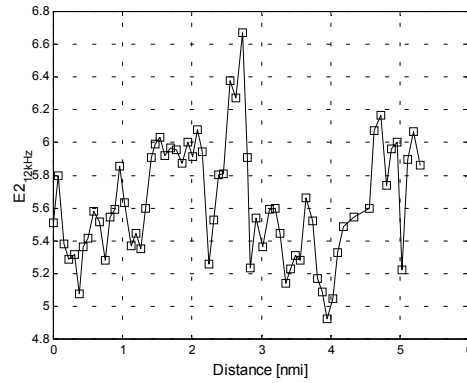
(b) Interpolated series

Figure 5.21. Representative example of series of roughness index ( $E_1$ ) in one of the straight transects at (i) 12, (ii) 38, (iii) 120 kHz. (a) the original series where data points are not equally spaced. (b) the linearly interpolated series with data points being equally spaced that is used for the autocorrelation analysis.

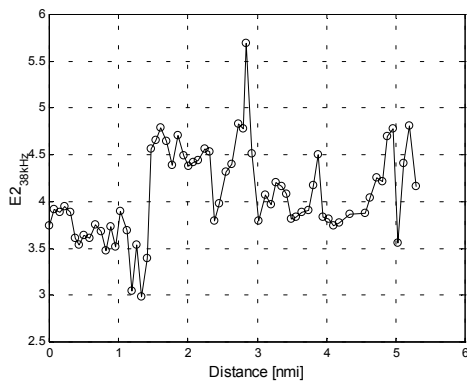




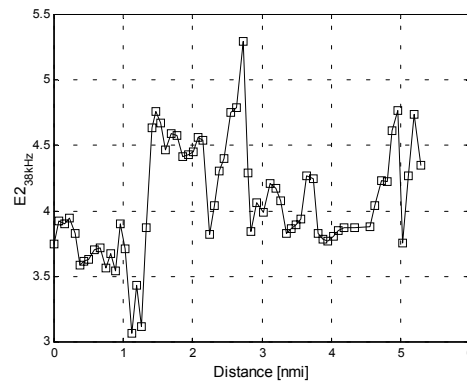
(i) 12 kHz



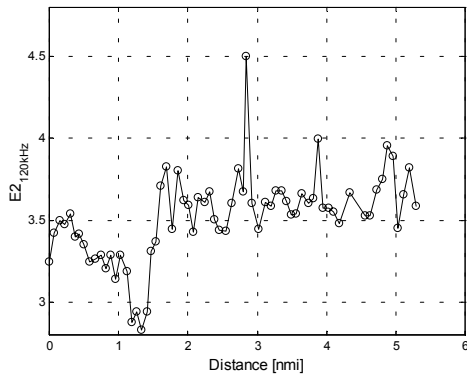
(i) 12 kHz



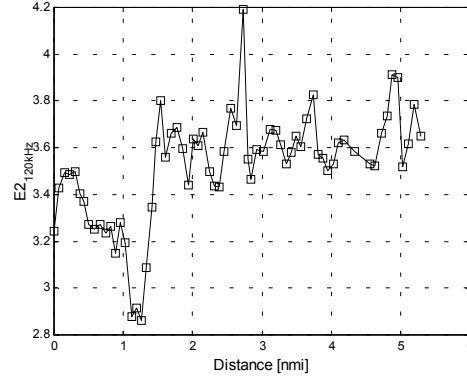
(ii) 38 kHz



(ii) 38 kHz



(iii) 120 kHz

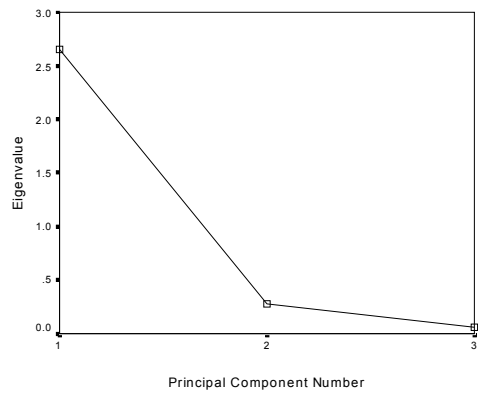


(iii) 120 kHz

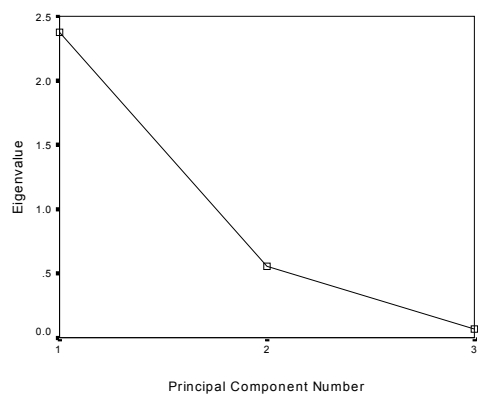
(a) Original series

(b) Interpolated series

Figure 5.22. Representative example of series of hardness index (E2) of the similar straight transect in Figure 5.21 at (i) 12, (ii) 38, (iii) 120 kHz. (a) the original series where data points are not equally spaced. (b) the linearly interpolated series with data points being equally spaced that is used for the autocorrelation analysis.



(a) E1 index corresponding to Figure 5.21



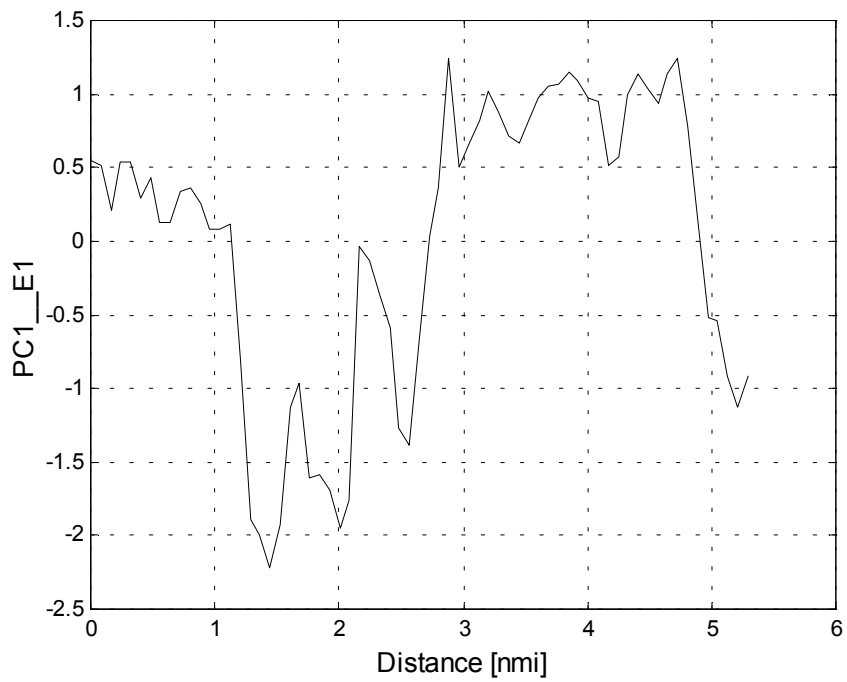
(b) E2 index corresponding to Figure 5.22

Figure 5.23. The scree plot produced by applying the principal component analysis into a set of E1 and E2 given in Figures 5.21 and 5.22, respectively. Series of the first principal is shown in Figure 5.24.

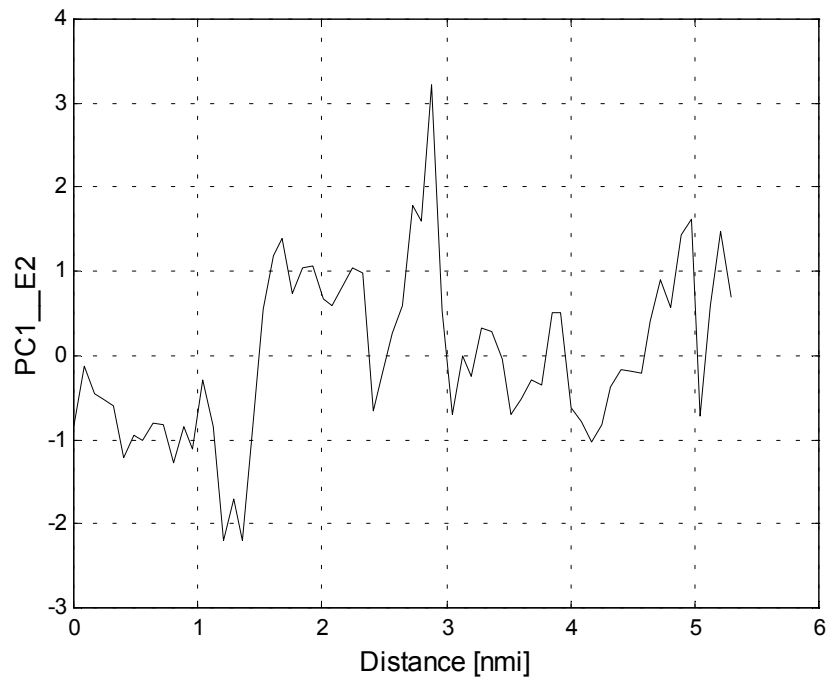
extreme one was as low as 43% as already mentioned in section 5.2. As mentioned in section 5.2, three patterns of component loadings were observed in transects where the total variance of PC1\_E1 was low (less than 70%). This indicates the three scenarios already mentioned in section 5.2. On average, PC1\_E1 and PC1\_E2 account for 70% and 80% of the total variation of the original E1 and E2, respectively.

After PCA had been applied to E1 and E2 in each straight transect separately, ACF was applied to E1 in each transect and similarly to E2 in each transect using the algorithm described in section 1.6. The autocorrelation functions of PC1\_E1 and PC1\_E2 given in Figures 5.24(a) and (b) are shown in Figures 5.25(a) and (b), respectively. The autocorrelation lengths of PC1\_E1 and PC1\_E2 for this particular example estimated by the first zero crossing are 1.5 and 1.1 nmi, respectively. Similar autocorrelation lengths of PC1\_E1 and PC1\_E2 estimated by the two methods were observed in all transects in both study areas. On average, the autocorrelation length of PC1\_E1 for the NWS and SEF study areas is  $3.2 \pm 1.32$  nmi and  $3.28 \pm 1.38$  nmi, respectively and the autocorrelation of PC1\_E2 is  $3.29 \pm 1.5$  nmi and  $3.29 \pm 1.17$  nmi.

The interpolation into unsampled areas was conducted using Kriging in which the dependency of the data on the spatial autocorrelation was taken into account. Variograms of PC1\_E1 and PC1\_E2 were first derived from all transects given in Figure 5.20 for both study areas separately and are shown in Figure 5.26. Similar to the autocorrelation characteristic length derived from the first sign transition in the ACF, the autocorrelation length derived from the variogram (known as range of influence) is, on average, 3 nmi. The range of influence of PC1\_E1 and PC1\_E2 are 3.2 and 3.24 nmi, respectively, for the NWS study area, and 3.58 and 3.24 nmi,

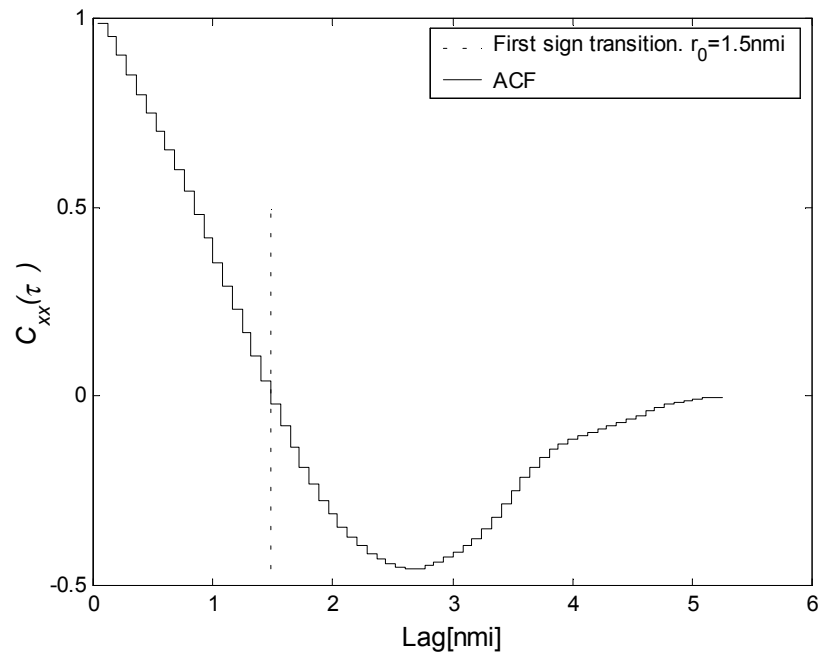


(a) PC1\_E1

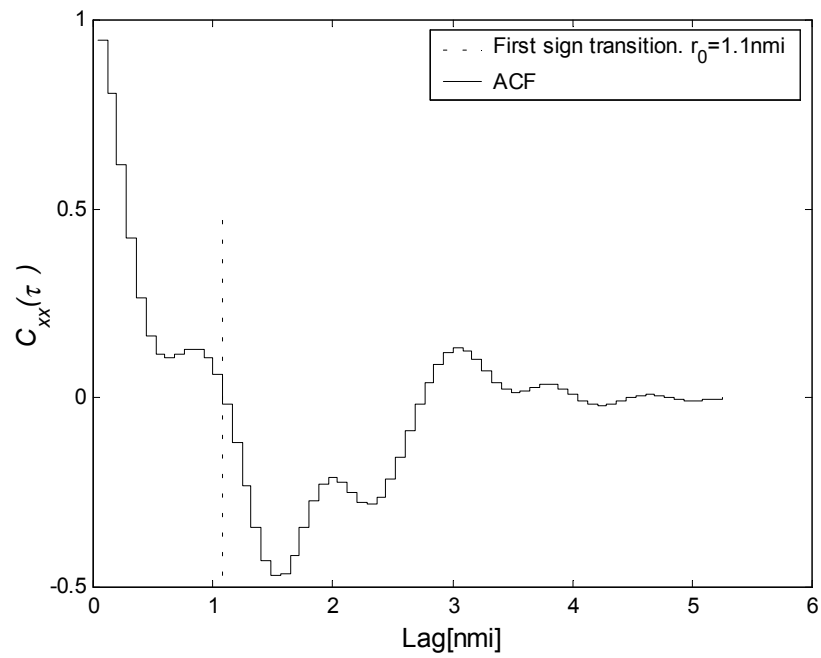


(b) PC1\_E2

Figure 5.24. Series of the first principal component of (a) roughness index (PC1\_E1) and (b) hardness index (PC1\_E2) that correspond to Figures 5.21 to 5.23. The corresponding scree plot presenting the eigenvalues of all principal components is shown in Figure 5.23.

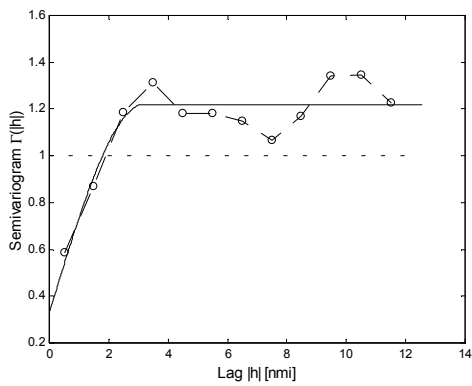


(a) PC1\_E1

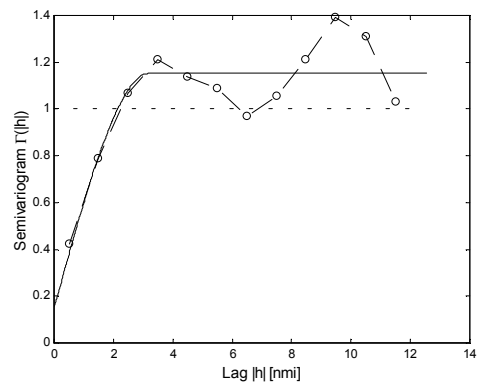


(b) PC1\_E2

Figure 5.25. Plot of autocorrelation function of the first principal component of (a) roughness index (PC1\_E1) of the series given in Figure 5.24(a) and (b) hardness index (PC1\_E2) of the series shown in Figure 5.24(b). The autocorrelation characteristic length of PC1\_E1 and PC1\_E2 estimated by the two methods is 1.5 and 1.1 nmi, respectively.

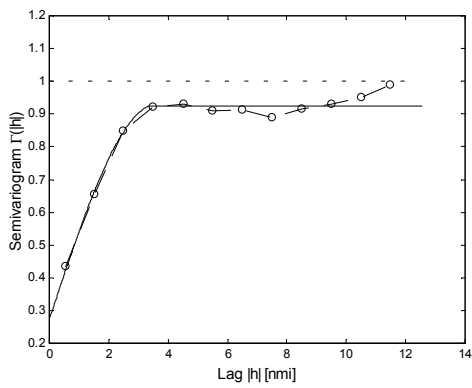


(i) Roughness

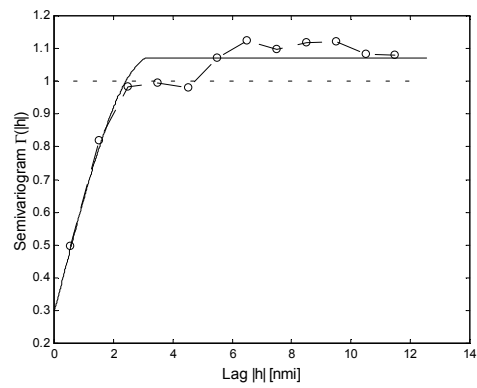


(ii) Hardness

(a) NWS



(i) Roughness



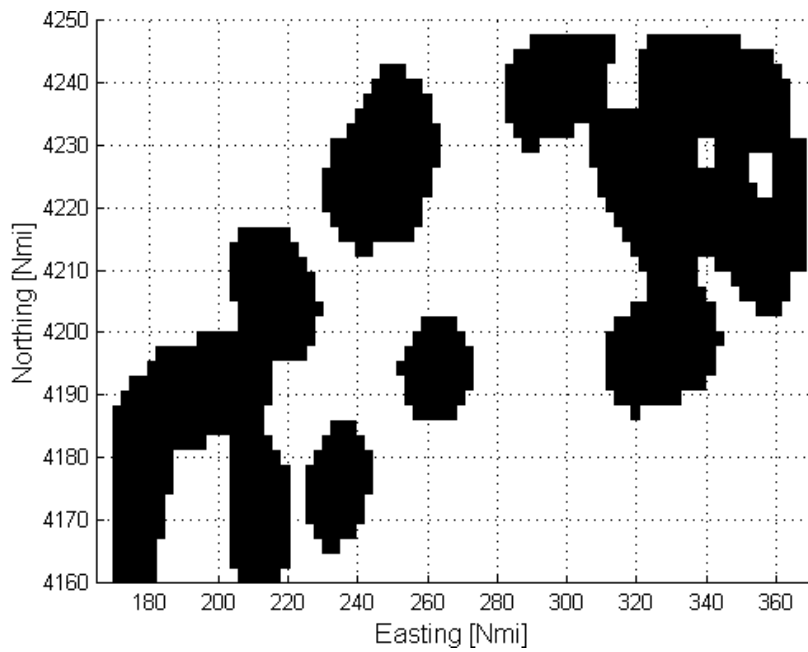
(ii) Hardness

(b) SEF

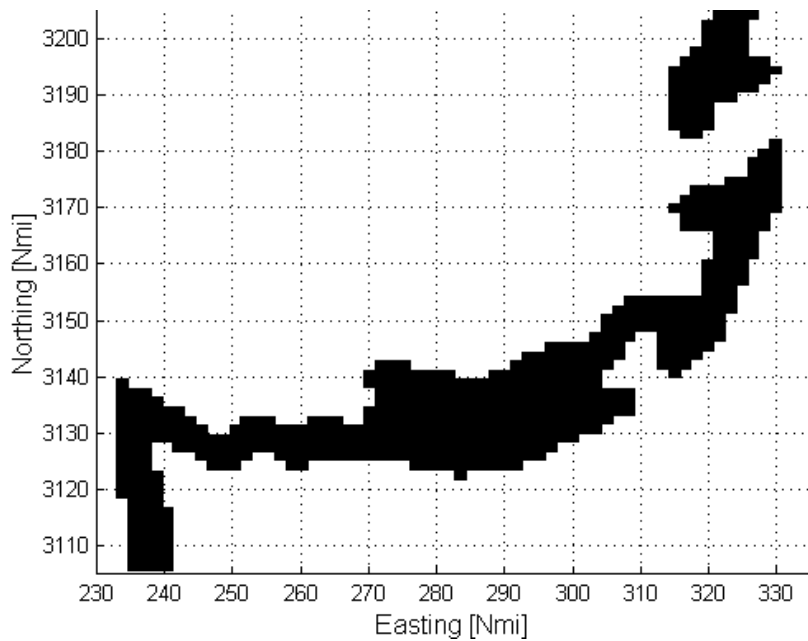
Figure 5.26. Variogram of the first principal component of (i) roughness and (ii) hardness indices (PC1\_E1 and PC1\_E2, respectively) for (a) the NWS study area and (b) the SEF study area. o = sample variogram. Broken lines are intended only to guide the eye. Solid lines are models fitted to the sample variogram.

respectively, for the SEF study area. Sample variograms of PC1\_E1 and PC1\_E2 for the two study areas fit well a spherical model ( $r = 0.92$  and  $0.89$  for respectively PC1\_E1 and PC1\_E2 in the NWS study area, and  $r = 0.99$  and  $0.97$  for respectively PC1\_E1 and PC1\_E2 in the SEF study area). Variograms in all directions in  $10^\circ$  increments for both study areas were derived to observe any direction trends. However, nothing of interest was found in any directions in both study areas. In other words, the variogram models are omnidirectional which suggests that the spatial correlation structure is much the same irrespective of direction in both study areas.

The fitted models was then used to provide interpolated measures of roughness and hardness, using Kriging technique within region of Kriging shown in Figure 5.27. This has, however, the effect of 'smearing' the actual data tracks. The contour plots of Kriged roughness and hardness indices and their associated contour plot of standard errors are shown in Figures 5.28a and 5.28b, respectively, for the NWS region and Figures 5.29a and 5.29b, respectively, for the SEF regions. The plots of Krige estimated roughness and hardness indices versus actual roughness and hardness indices for both regions are shown in Figures 5.28c and 5.29c, respectively. Suppose that all the green contours are considered as rough surfaces for PC1\_E1 and hard surfaces for PC1\_E2, the darkest blue is considered as a transition between rough and smooth surfaces for PC1\_E1 and between hard and soft surfaces for PC1\_E2, and the two other blue contours are considered as smooth surfaces for PC1\_E1 and soft surfaces for PC1\_E2, the NWS study area is mainly dominated by smooth surfaces and a transition between rough and smooth surfaces. Rough surfaces present only in a few regions in particular in regions A, B, C and D shown in Figure 5.28a(i). Similarly, results from Jones (1973) as given in Figure 5.2a indicate rough surfaces in regions just mentioned. From Figure 5.28a(ii), regions of rough surfaces just



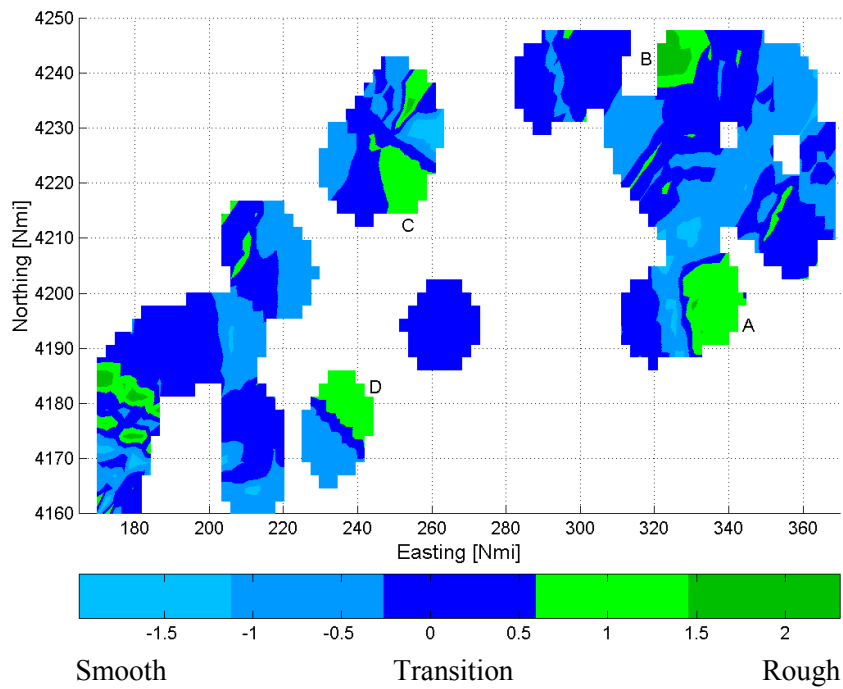
(a) NWS



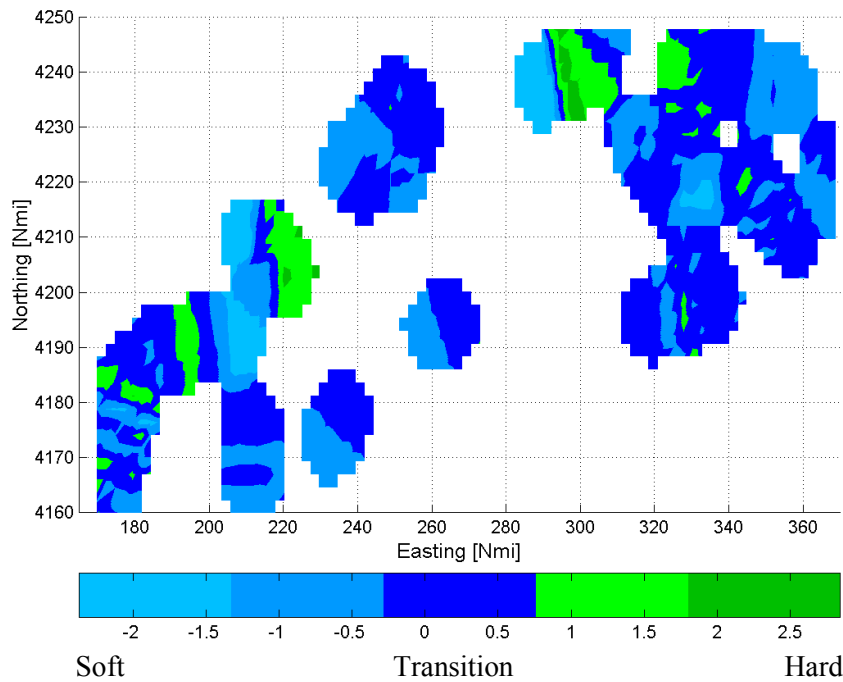
(b) SEF

Figure 5.27. Map of region of Kriging within the autocorrelation characteristic length (the range) of the straight lines shown in Figure 5.20 for (a) the NWS study area and (b) the SEF study area to produce 2-D bottom roughness and hardness structure in the unsampled areas within the range given in Figure 5.26.



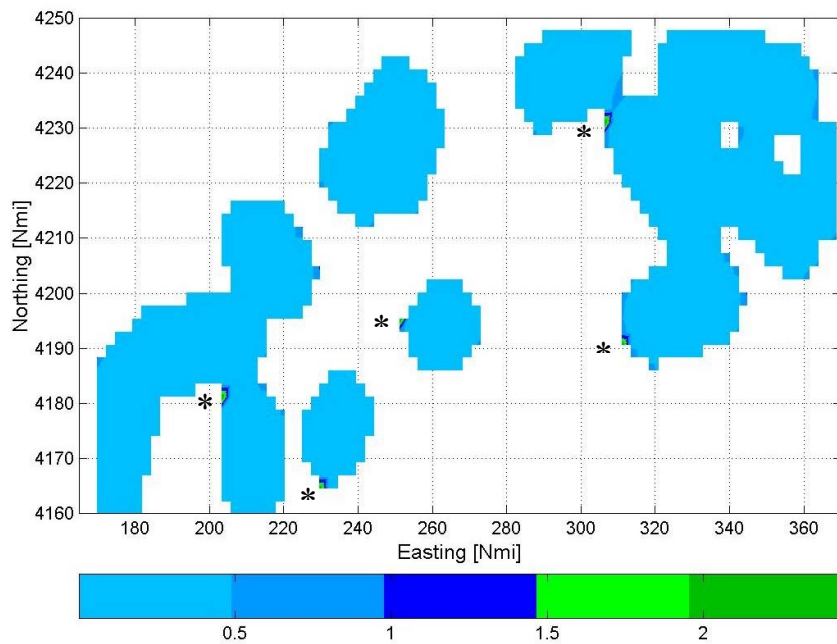


(i)

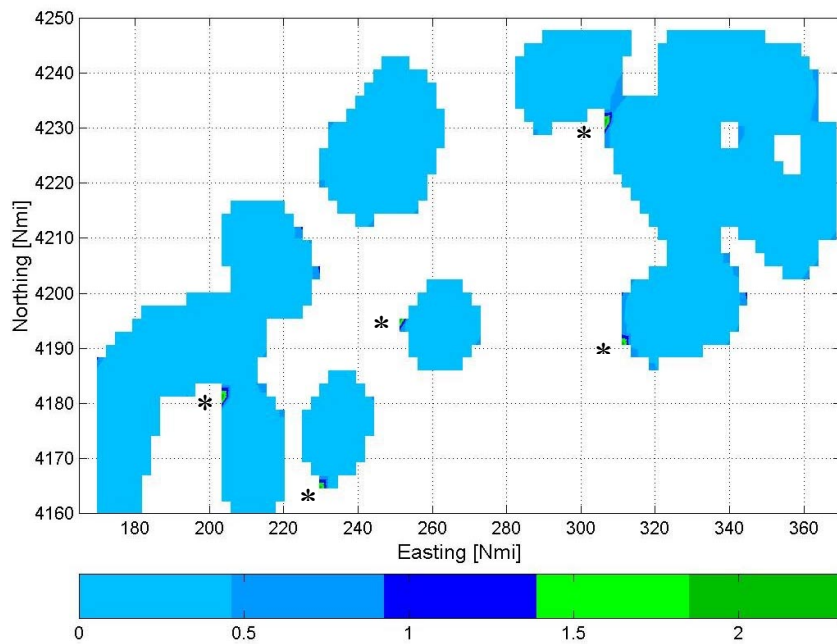


(ii)

Figure 5.28a. Contour plot of (i) bottom roughness structure and (ii) bottom hardness structure based on the region shown in Figure 5.27(a) for the NWS region by using the Kriging technique.

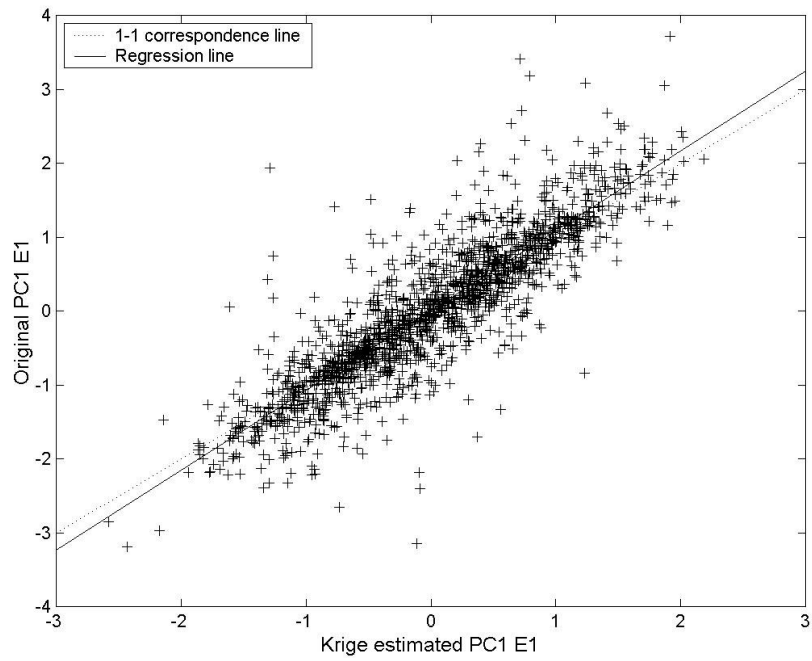


(i)

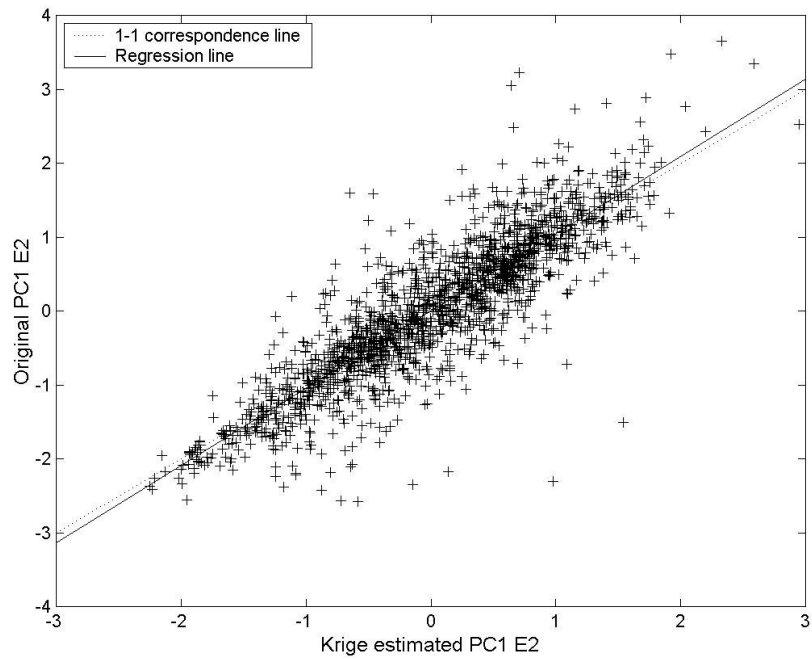


(ii)

Figure 5.28b. Contour plot of standard errors for (i) bottom roughness and (ii) bottom hardness associated with Figure 5.28a for the NWS region by using the Kriging technique. \* = regions of high standard errors.

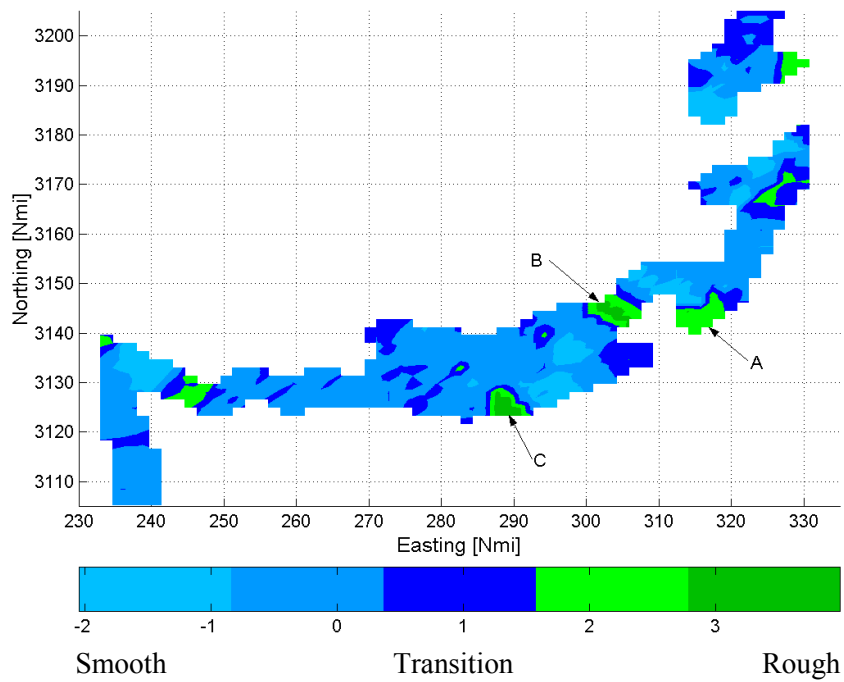


(i)

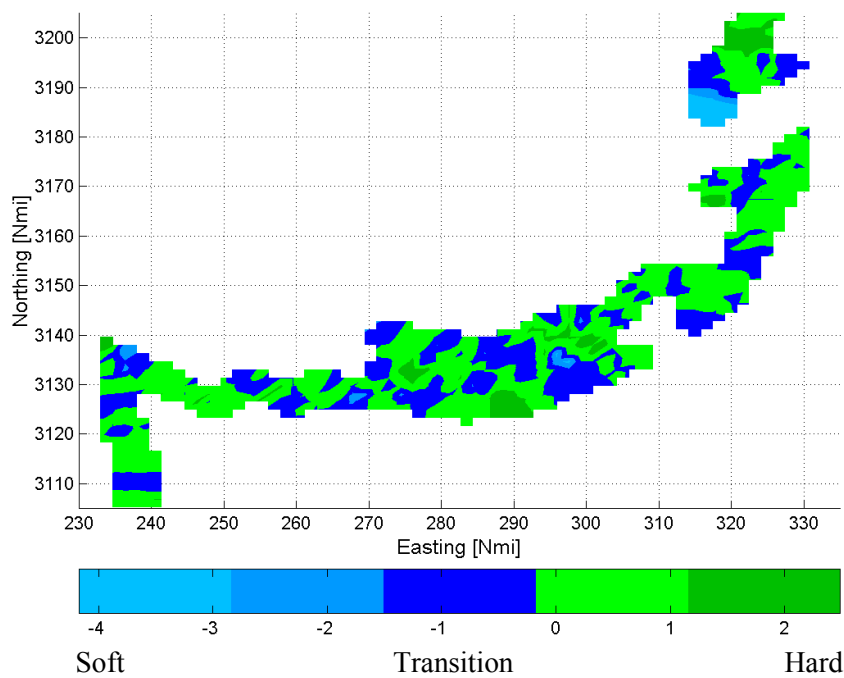


(ii)

Figure 5.28c. Plot of Kriged estimated index versus actual index for (i) bottom roughness and (ii) bottom hardness for the NWS region using the Kriging technique. — = regression line with the equation for (i)  $y = 1.080x + 0.001$  ( $r = 0.88$ ) and (ii)  $y = 1.047x - 0.001$  ( $r = 0.88$ ).

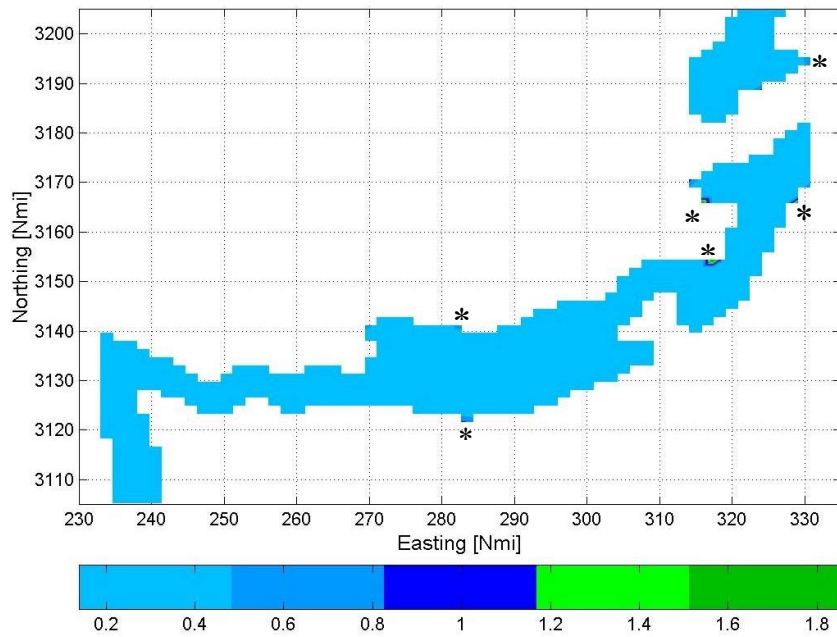


(i)

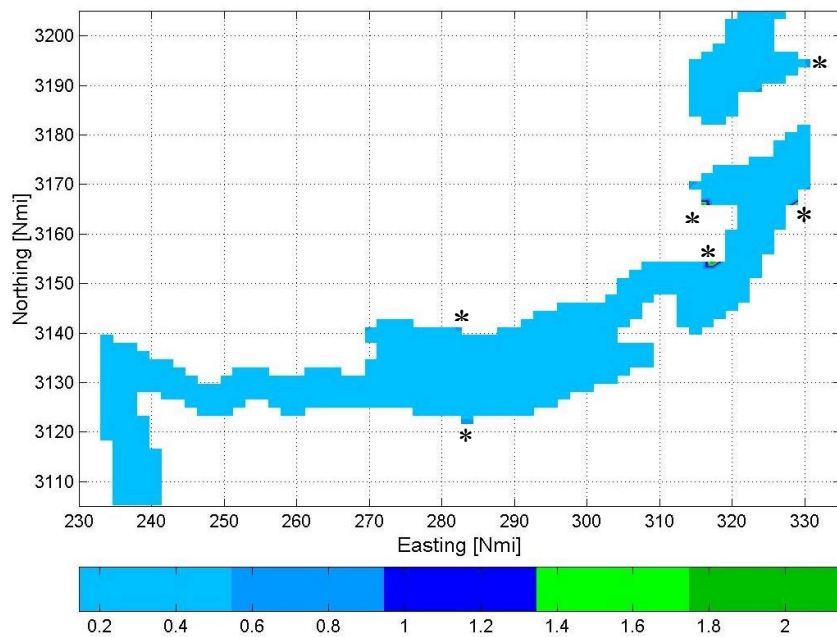


(ii)

Figure 5.29a. Contour plot of (i) bottom roughness structure and (ii) bottom hardness structure based on the region shown in Figure 5.27(b) for the SEF region by using the Kriging technique.

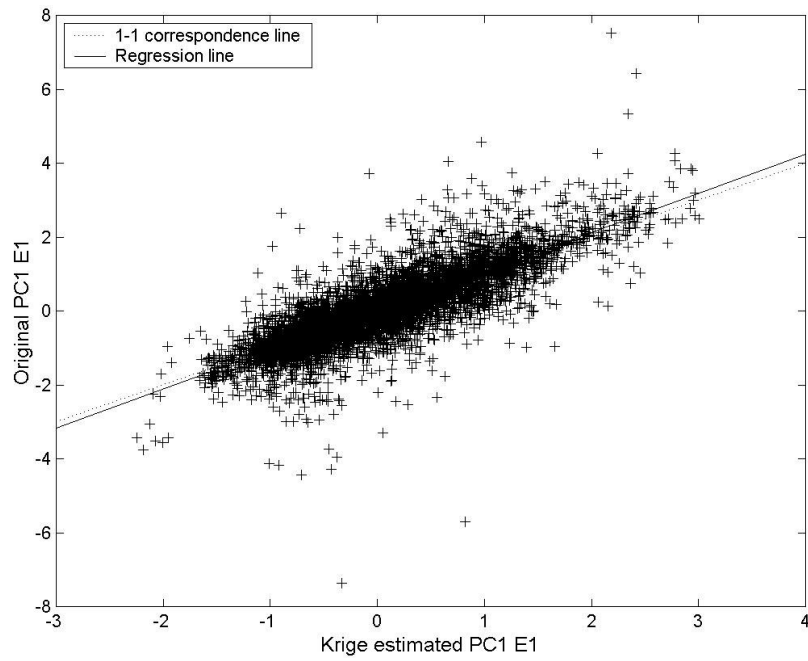


(i)

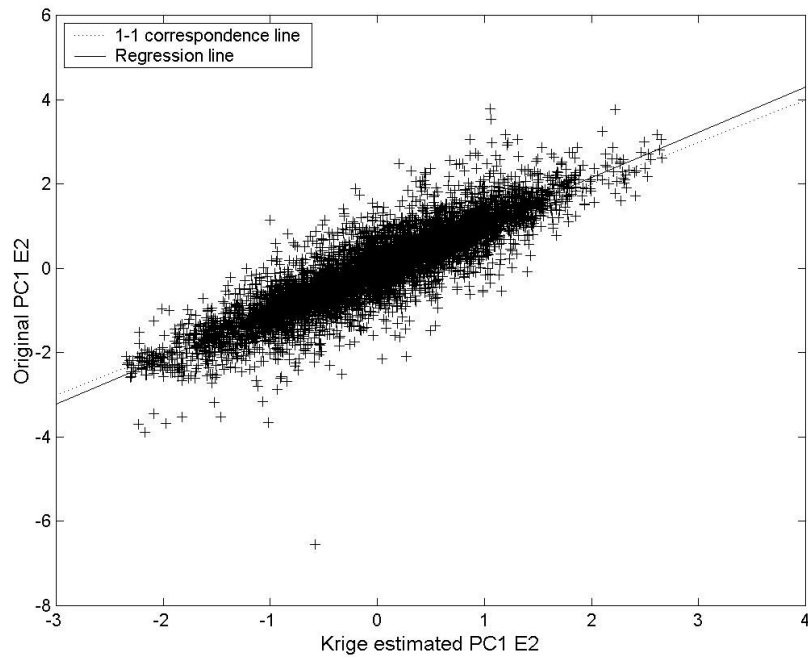


(ii)

Figure 5.29b. Contour plot of standard errors for (i) bottom roughness and (ii) bottom hardness associated with Figure 5.29a for the SEF region by using the Kriging technique. \* = regions of high standard errors.



(i)



(ii)

Figure 5.29c. Plot of Kriging estimated index versus actual index for (a) bottom roughness and (b) bottom hardness for the SEF region using the Kriging technique. — = regression line with the equation for (i)  $y = 1.080x + 0.001$  ( $r = 0.88$ ) and (ii)  $y = 1.047x - 0.001$  ( $r = 0.88$ ).

mentioned have different associated hardness. Rough surfaces in regions A, C, and D for instance are transitional surfaces (between hard and soft surfaces) whereas the other rough surface (region B) consist mainly of hard surfaces.

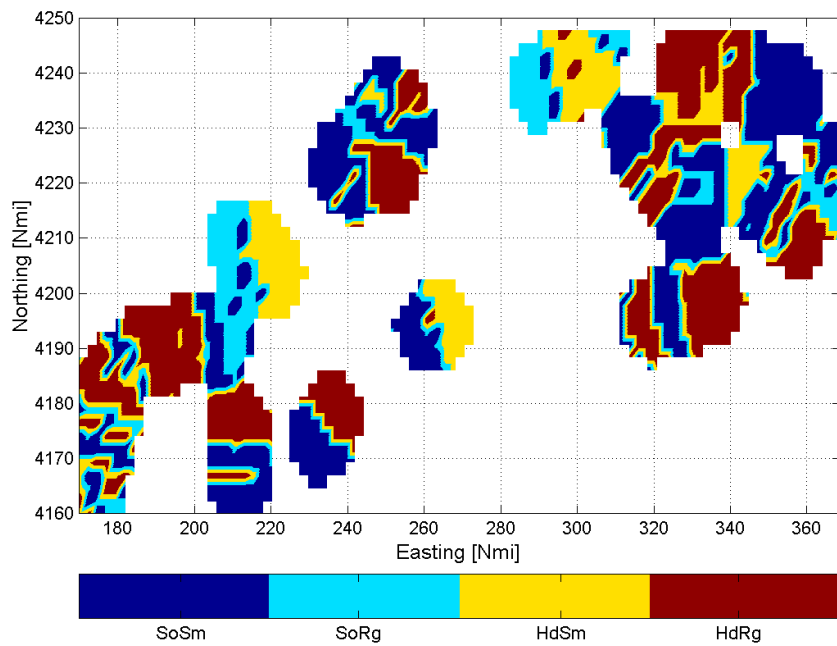
For the SEF study area, the seabed consists mainly of hard-smooth surfaces, except in regions A, B and C shown in Figure 5.29a(i) where rough surfaces are present. Dominated hard-smooth surfaces in the SEF region are associated with the presence of coarse sediment and shell debris in a large proportion (Jones and Davies, 1983; Kloser *et al.*, 2001). The presence of different grainsizes of sands and sandwaves in region B in Figure 5.29a were also observed by Kloser *et al.* (2001).

Unlike other interpolation techniques, Kriging also provides standard errors for every interpolated value (see equation (B.8) in Appendix B). The contour plots of standard errors for roughness and hardness indices in both regions possess the same characteristic. Each of them suggests that the Krige estimated standard error is approximately the same over the entire interpolated region except in areas labelled as \*; all located at the edge of interpolated region. On average, the standard error of prediction is 0.48 for both Krige estimated roughness and hardness indices for the NWS region and 0.55 and 0.45 for Krige estimated roughness and hardness indices, respectively, for the SEF region. The cross-validation was also performed to compare between the actual roughness and hardness indices and their associated Krige estimator (Figures 5.28c and 5.29c). Figures 5.28c and 5.29c show a high correlation between the actual and the predicted roughness and hardness indices. The correlation coefficient is 0.88 for both roughness and hardness indices in the NWS region and 0.83 and 0.89 for roughness and hardness indices, respectively, for the SEF region. The slope of the regression line is close to one for both indices and regions. The slope

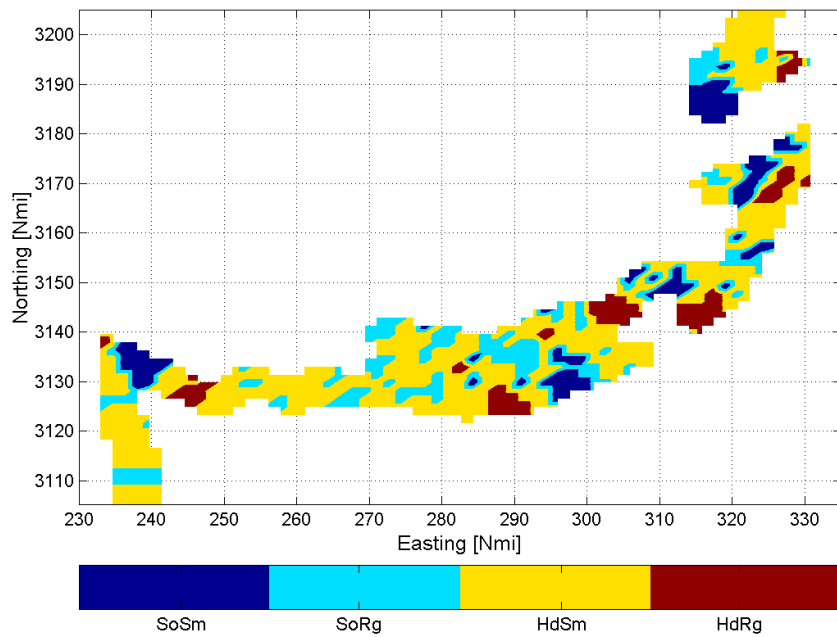
of the regression line is  $1.080 \pm 0.014$  and  $1.047 \pm 0.014$  for roughness and hardness indices, respectively, in the NWS region and  $1.057 \pm 0.008$  and  $1.075 \pm 0.006$  for roughness and hardness indices, respectively, in the SEF region.

To produce contour plots of seabed types for the NWS and SEF study areas within the Kriging regions given in Figures 5.27(a) and (b) (note that Kriging regions for PC1\_E1 and PC1\_E2 of each study area are the same), the classification procedure outlined in section A.2.3 was applied to Kriged PC1\_E1 and PC1\_E2 indices shown in Figures 5.28a(i) and (ii) for the NWS study area and Figures 5.29a(i) and (ii) for the SEF study area. Like the classification procedure prepared for the along-track seabed classification (section 5.4), the classification procedure applied to Kriged roughness and hardness indices shown in Figures 5.28a and 5.29a involved only class assignments based on the centroids given in Table 5.6 as bold entries without iterative and feedback process. Results of the contour plots of seabed types are shown in Figures 5.30(a) and (b) for the NWS and SEF study areas, respectively. As far as Figure 5.30(a) is concerned, the NWS study area consists mainly of soft-smooth (SoSm) surfaces (36%) and hard-rough (HdRg) surfaces (32%). The two other seabed types namely soft-rough (SoRg) and hard-smooth (HdSm) surfaces account for 14% and 18% of the Kriging region, respectively. Although the NWS study area is dominated by soft-smooth and hard-rough surfaces, the distribution of the four derived seabed types are quite comparable. Unlike the NWS study area, more than half of the SEF study area consists mainly of hard-smooth (HdSm) surfaces. Hard-smooth surfaces make 62% of the Kriging region of the SEF study area. Soft-smooth (SoSm), soft-rough (SoRg) and hard-rough (HdRg) surfaces account for respectively 12%, 8% and 18% of the Kriging region of the SEF study area. While presence of the four derived seabed types in the NWS study area is quite comparable, it is not so in





(a)



(b)

Figure 5.30. Contour plot of seabed type derived from Kriged PC1\_E1 and Kriged PC1\_E2 indices given in Figures 5.28a and 5.29a for (a) the NWS study area (b) the SEF study area.

the SEF study area. In addition, distances over which variations of Kriged seabed types occurred are longer in the SEF study area than in the SEF study area. Together with the percentage of the seabed types just mentioned for both study areas, this again might indicate that the local variations are greater in the NWS study area than in the SEF study area.

## Chapter 6

# Relationship between common species and acoustically derived seabed type

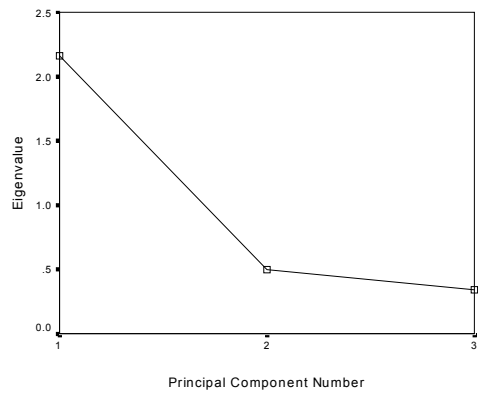
### 6.1. Introduction

As already discussed in Chapter 4, the difference between  $s_A$  estimates from nets and acoustics is substantial, despite the fact that there is an overall linear relationship. The fact that there is a substantial variation in detail between the two  $s_A$  estimates from results in Chapter 4 leads to a conclusion that it is convenient here in this section to adopt only the  $s_A$  estimates from acoustics rather than from nets or both to explore the relationship between common species and acoustically derived seabed types. Nevertheless, species compositions from nets are useful indeed to provide positive identification of recovered organisms that cannot be accomplished here in this study by acoustic techniques. Data used to explore the relationship between common species and acoustically derived seabed types are the average values of PC1\_E1, PC1\_E2 in each trawl station, centroids of PC1\_E1 and PC1\_E2 representing the mode of seabed types in each trawl station, the average values of PC1\_SA, and compositions of the common species and weights of the common species per area trawled. The common species to form adopted groupings of fish species came from areas of trawlable grounds. To extend the study to include also species from areas of untrawlable grounds, gill nets and traps may serve as an alternative to the trawl. This, however, is not available in this current study.

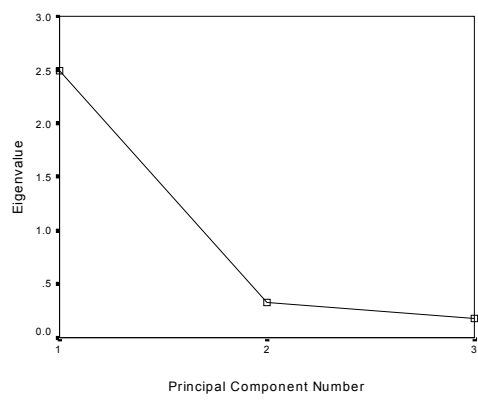
Assemblages such as fish species, fish communities or groups of fish prefer certain spatial configurations and make use of them in particular ways such that the overall space is not occupied randomly (Gerlotto, 1993). Assemblages of fish species into coherent communities or groups is subject to physical factors such as, among others, seabed types (Gerlotto, 1993), as well as chemical ones. As far as acoustics is concerned, Gerlotto (1993) insists that the acoustic returns offer a synthesis of the physical and biological characteristics of the assemblage. This has led to studies of spatial structures being used as the basis for the identification techniques despite the fact that fish may not remain in the space always in the same manner (Fréo *et al.*, 1990; Scalabrin and Massé, 1993 cited in Gerlotto, 1993).

## **6.2. PCA of the area backscattering coefficient $s_A$**

Before the hierarchical agglomerative clustering technique is applied to the data, it is necessary to reduce the dimensionality of the area backscattering coefficient,  $s_A$ , from the three operated frequencies. In order to eliminate the sensitivity of PCA to different scales of variables, the area backscattering coefficients of the three frequencies were first transformed in a logarithmic form. Since  $s_A$  values from the three operated frequencies are intercorrelated, the first principal component is sufficient to represent most of the variation in the original  $s_A$  values. This is confirmed by the scree plot as shown in Figure 6.1. It is clear in Figure 6.1 that eigenvalues of the second and third principal components are quite similar and low, and are very different from the eigenvalue of the first principal component. Since the eigenvalues of the second and the third principal components are very low, they contribute little to the total variation of the original  $s_A$  values and therefore can be neglected. For the NWS study area, the first principal component of  $s_A$  values (PC1\_SA) accounts for 72% of the total variation of the original  $s_A$  values whereas for



(a) NWS

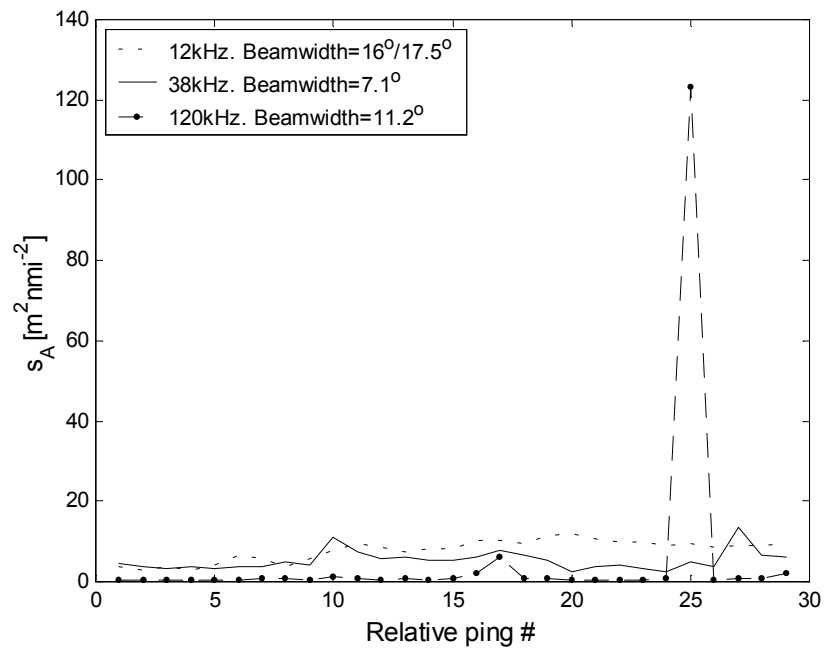


(b) SEF

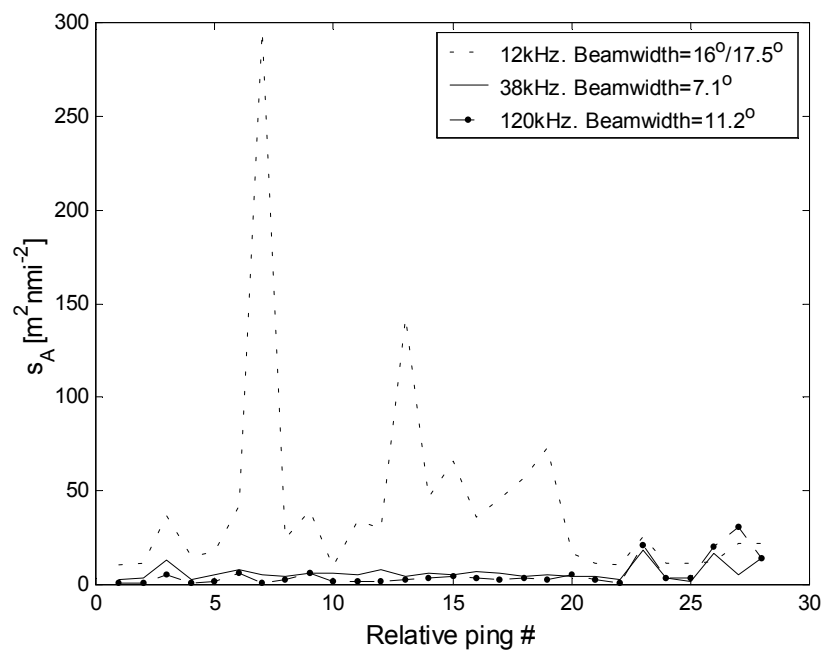
Figure 6.1. Scree plot of area backscattering coefficient,  $s_A$ , formed from the three operated frequencies in (a) the NWS region and (b) the SEF region.

the SEF study area, it accounts for 83%. This PC1\_SA, together with PC1\_E1, PC1\_E2 and other variables mentioned earlier then form the data matrix for classification purposes described in section A.2.2. Since the variance of PC1\_SA in the two study areas differs by 10%, it is of interest to apply PCA to  $s_A$  values in trawl stations individually to get a clear picture of the characteristic of acoustic data from the three operated frequencies at each individual trawl station.

For the NWS study area, the variance of PC1\_SA in trawl stations individually varies from 52% (lowest) to 93% (highest) and for the SEF study area, it varies from 51% (lowest) to 97% (highest). Two patterns in component loadings of the first principal component for low variances are observed. In the first pattern, the component loadings of SA\_12 and SA\_120 ( $s_A$  estimates at 12 kHz and 120 kHz, respectively) in the first principal component are similar but are very different from the component loading of SA\_38 ( $s_A$  estimates at 38 kHz). In the second one, the component loadings of SA\_38 and SA\_120 are similar but very different from the component loading of SA\_12. Series of  $s_A$  estimates at the three operated frequencies given in Figure 6.2 reveal the effects of the differing beamwidth of the three operated transducer units on a highly patchy distribution of fish. It is important to note that the beamwidth of 12 kHz, 38kHz and 120 kHz transducer units is  $16^\circ/17.5^\circ$ ,  $7.1^\circ$  and  $11.2^\circ$ , respectively. Figure 6.2(a) shows a representative example of series of the first pattern mentioned earlier whereas Figure 6.2(b) shows a representative example of series of the second one. In comparison, Figure 6.3 shows representative examples of the series of  $s_A$  estimates at the three operated frequencies from those PC1\_SA that have very high variance. Although both study areas have PC1\_SA as low as 50%, the low PC1\_SA occurs more frequently in the NWS study area, which is located in the tropics, than in the SEF study area, which is located in the temperate, higher latitude



(a) First pattern



(b) Second pattern

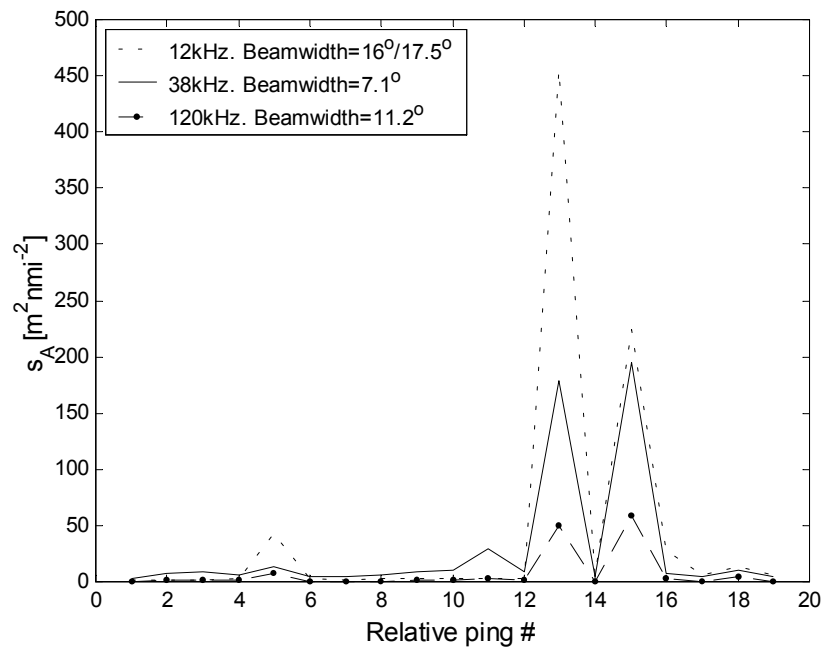
Figure 6.2. Series of acoustical  $s_A$  estimates at three operated frequencies, -- = 12 kHz, — = 38 kHz and —•— = 120 kHz, showing the beamwidth effect on a highly patchy distribution of fish for the two observed patterns in the component loadings of variables in the PC1 having low variance accounted for most of the variation in the original variables.

region. This indeed might be used explicitly to confirm the highly local variation of seabed types and benthic types in the tropic NWS study area discussed in section 5.4.

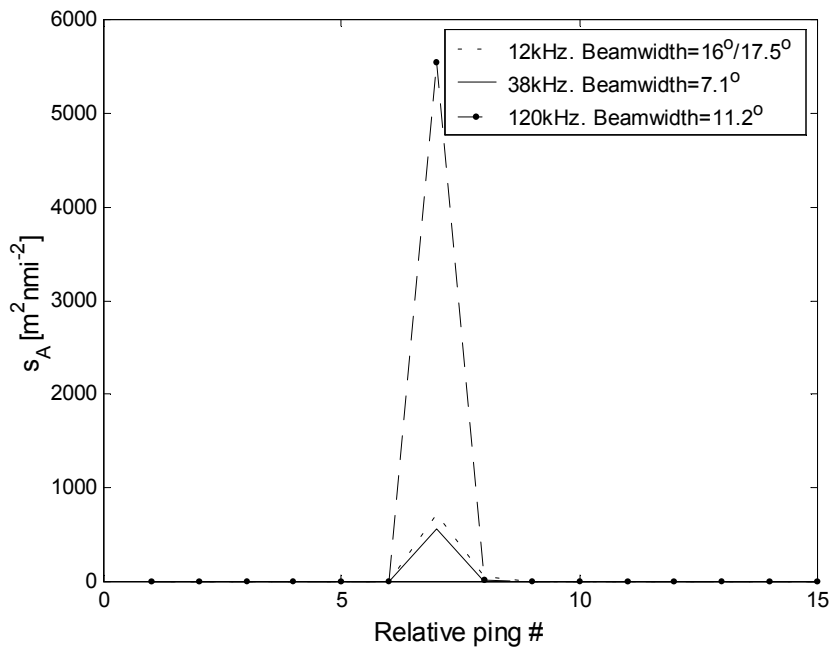
### **6.3. The "quasi acoustic population"**

In each individual trawl station, the average values of PC1\_E1, PC1\_E2 and PC1\_SA are calculated. It is important to note that the average is taken from the first principal component of  $s_A$  estimated from all the trawl stations available at once rather than that estimated individually at each trawl station as it gives the average value of zero because the calculated principal components are based on standardised  $s_A$  values having zero mean and unit variance. Likewise, this is applied to the first principal components of the E1 and E2 indices. These average values together with catch compositions in each trawl station form the so-called "acoustic population", a concept developed by Gerlotto (1993) in his study to identify and spatially stratify tropical fish concentrations. He defined the acoustic population as a set of discriminate characteristics of the acoustic data and other supportive information logically related to main biological characteristics or behaviours of the group of fish species. This concept was adopted here but it involved not only the acoustic parameters but also the biological parameters in particular the relative abundances of derived groupings of fish species. As such, it was rather called the "quasi acoustic population." The notion underlying the inclusion of the biological parameters is that they provide the obvious, positive identification of the recovered organisms in specific seabed types that the current acoustic technology adopted here fails to fully offer. As far as studies of tropical multispecies identification are concerned, the acoustic technique is able to discriminate not more than 5 or 6 fish species (Rose, 1992 cited in Gerlotto, 1993). It was also expected that the outcomes (groups of the quasi acoustic population) were equally weighted between the physical and biological parameters at once instead of





(a)



(b)

Figure 6.3. Series of acoustical  $s_A$  estimates at three operated frequencies, -- = 12 kHz, — = 38 kHz and -●- = 120 kHz, for those first principal components having the highest variance accounted for most of the variation in the original variables.

either one as an excuse for the other one. A set of parameters used in the quasi acoustic population for the NWS and SEF study areas are presented in Tables 6.1 and 6.2, respectively. The first four variables shown in Tables 6.1 and 6.2 represent seabed characteristics in each trawl station as was suggested by Gerlotto (1993). While the first two are simply the average values of the first principal component of E1 and E2 indices (PC1\_E1 and PC1\_E2) in each trawl station, the last two are the centroids of PC1\_E1 and PC1\_E2 of the four derived seabed types adopted for seabed classification in sections 5.3 and 5.4. The latter provides the information of the most common seabed type that exists in each trawl station. The rationale underlying the inclusion of the centroids and the average of PC1\_E1 and PC1\_E2 is as follows. The centroids are required to take into account the contribution of the common seabed type in each trawl station for the purpose of the clustering analysis. The averages on the other hand become a measure as to whether or not the centroids best represent the common seabed type in each trawl station. The more the averages differ from the centroids in a particular trawl station, the more the seabed type in that particular trawl station differs from the centroid-represented seabed type and the closer the seabed type in that particular seabed type is to centroids of the adjacent seabed type. The fifth variable in Table 6.1 and 6.2 provides information on the relative biomass of fish in each trawl station and the rationale for the inclusion is as follows. The relative biomass of fish is an indicator of the concentration of fish species or their assemblage into communities or groups on particular areas identified by the first four variables. While this variable provides no information on recovered fish identities, catch compositions in both study areas do provide that information and hence are adopted here as shown in Tables 6.1 and 6.2 in the last eight variables. These eight variables include four different derived groups of fish for each study area. Those groups of fish

Table 6.1. Parameters adopted to form quasi acoustic population in the NWS study area. \*The most occurrence acoustically derived seabed type in each trawl station. SoSm = soft-smooth. SoRg = soft-rough. HdSm = hard-smooth. HdRg = hard-rough. PC1\_E1 = first principal component of roughness indices. PC1\_E2 = first principal component of hardness indices. Centroids are related to the most occurrence acoustically derived seabed type in each trawl station. PC1\_SA = first principal component of acoustical  $s_A$  estimates. Group 1 = fish from genera *Lethrinus* and *Lutjanus*. Group 2 = fish from genera *Saurida* and *Nemipterus*. Group 3 = fish from genus *Epinephelus*. Group 4 = fish from genus *Parupeneus*.

Trawl Station ID	CSIRO ID	Mode seabed type*	Mean PC1_E1	Mean PC1_E2	Centroids		Mean PC1_SA	Fish group composition							
					PC1_E1	PC1_E2		Group 1		Group 2		Group 3		Group 4	
								%	kg/nmi <sup>2</sup>	%	kg/nmi <sup>2</sup>	%	kg/nmi <sup>2</sup>	%	kg/nmi <sup>2</sup>
1	9508042	SoRg	.0108	-1.422	.030	-1.838	.991	19.88	9260.00	2.66	1240.84	10.37	4827.55	.00	.00
2	9508044	SoSm	-.191	.328	-1.294	-.595	-.881	2.87	101.02	13.19	464.12	6.98	245.81	.00	.00
3	9508047	HdRg	1.854	1.460	1.302	.717	-.533	5.26	333.36	10.54	667.95	1.07	67.91	.71	45.07
4	9508049	SoSm	-.600	-.296	-1.294	-.595	.066	21.69	2610.70	51.51	6199.26	1.28	154.33	1.03	123.47
7	9508095	HdRg	.803	.931	1.302	.717	.439	4.07	234.59	1.25	72.23	.00	.00	.00	.00
9	9508108	HdRg	.523	.484	1.302	.717	.426	25.11	3376.81	4.85	651.90	.09	12.35	2.54	342.00
11	9707013	HdRg	.785	.453	1.302	.717	.227	15.98	1961.90	13.45	1651.02	1.98	243.66	.26	32.35
12	9707014	HdRg	1.915	.864	1.302	.717	-.103	10.02	711.31	39.03	2770.82	.93	66.17	2.33	165.42
13	9707026	SoSm	-.969	-.555	-1.294	-.595	-.329	11.72	494.81	18.37	775.08	4.18	176.47	2.71	114.19
14	9707032	SoSm	-1.550	-.905	-1.294	-.595	-.631	25.29	7078.03	8.41	2354.95	.05	12.86	2.84	794.88
15	9707033	SoRg	-.186	-.534	.030	-1.838	-.136	.00	.00	66.38	4754.55	.00	.00	2.60	186.24
16	9707055	HdSm	-.018	.601	-.305	1.520	1.301	8.17	260.67	.07	2.09	.00	.00	.72	22.94
17	9707106	HdRg	.267	.590	1.302	.717	.153	8.63	429.62	1.09	54.47	.20	9.87	.19	9.53
18	9707107	HdRg	.928	1.124	1.302	.717	-.218	4.93	356.27	.38	27.55	.24	17.36	1.10	79.25
19	9707109	HdRg	.423	.578	1.302	.717	.495	18.87	350.60	.91	16.82	.00	.00	.91	16.82
20	9707017	SoSm	-1.350	-.813	-1.294	-.595	1.255	17.56	815.21	22.99	1067.44	.00	.00	4.58	212.73
21	9707034	HdRg	1.116	.234	1.302	.717	-1.000	15.15	1043.59	20.50	1412.18	.00	.00	9.93	684.37

Table 6.2. Parameters adopted to form quasi acoustic population in the SEF study area. \*The most occurrence acoustically derived seabed type in each trawl station. SoSm = soft-smooth. SoRg = soft-rough. HdSm = hard-smooth. HdRg = hard-rough. PC1\_E1 = first principal component of roughness indices. PC1\_E2 = first principal component of hardness indices. Centroids are related to the most occurrence acoustically derived seabed type in each trawl station. PC1\_SA = first principal component of acoustical  $s_A$  estimates. Group 1 = sharks. Group 2 = cods. Group 3 = flatheads. Group 4 = pelagic fish (mainly mackerel and trevally).

Trawl Station ID	CSIRO ID	Mode seabed type*	PC1_E1		PC1_E2		Centroids		PC1_SA	Fish group composition							
							PC1_E1	PC1_E2		Group 1		Group 2		Group 3		Group 4	
										%	kg/nmi <sup>2</sup>	%	kg/nmi <sup>2</sup>	%	kg/nmi <sup>2</sup>	%	kg/nmi <sup>2</sup>
1	9606010	HdRg	2.628	.655	.486	-.008	35.46	2989.64	4.54	382.97	1.47	123.54	0.00	0.00			
3	9606034	SoRg	-.182	-.645	-.386	.520	1.68	194.34	0.00	0.00	3.68	425.96	0.11	12.96			
4	9606045	HdSm	-.131	.537	.659	-1.432	1.30	203.72	0.00	0.00	1.78	277.80	1.89	296.32			
5	9606069	HdSm	1.023	.849	.659	.229	3.62	460.78	0.00	0.00	0.92	117.29	0.00	0.00			
6	9606096	SoRg	.071	-.617	-.386	2.555	0.00	0.00	0.00	0.00	0.56	24.01	0.00	0.00			
7	9606106	SoRg	-.119	-.292	-.386	.287	1.64	82.11	0.00	0.00	1.26	63.15	11.69	586.40			
8	9606108	SoRg	1.366	-1.461	-.386	.399	0.00	0.00	0.00	0.00	7.96	259.28	1.19	38.83			
9	9606129	HdSm	.147	.766	.659	-1.657	0.14	40.74	0.00	0.00	1.24	364.23	56.67	16608.12			
10	9606158	HdSm	-.272	-.057	.659	-1.450	0.45	47.62	0.00	0.00	2.68	280.45	11.47	1201.15			
11	9606168	HdSm	.401	1.238	.659	.593	0.00	0.00	0.00	0.00	1.26	95.25	5.97	449.77			
12	9606220	SoSm	-.824	-1.038	-1.162	-.045	0.51	154.33	0.20	61.73	0.16	49.39	3.45	1043.29			
13	9606221	HdSm	-.254	.441	.659	.473	0.00	0.00	1.80	555.60	0.00	0.00	18.91	5852.32			
14	9606224	SoSm	-.855	-.365	-1.162	-.712	0.22	220.48	0.00	0.00	0.44	440.95	2.91	2927.92			
15	9606226	SoRg	1.346	-.559	-.386	1.315	0.00	0.00	0.00	0.00	0.73	86.43	0.00	0.49			

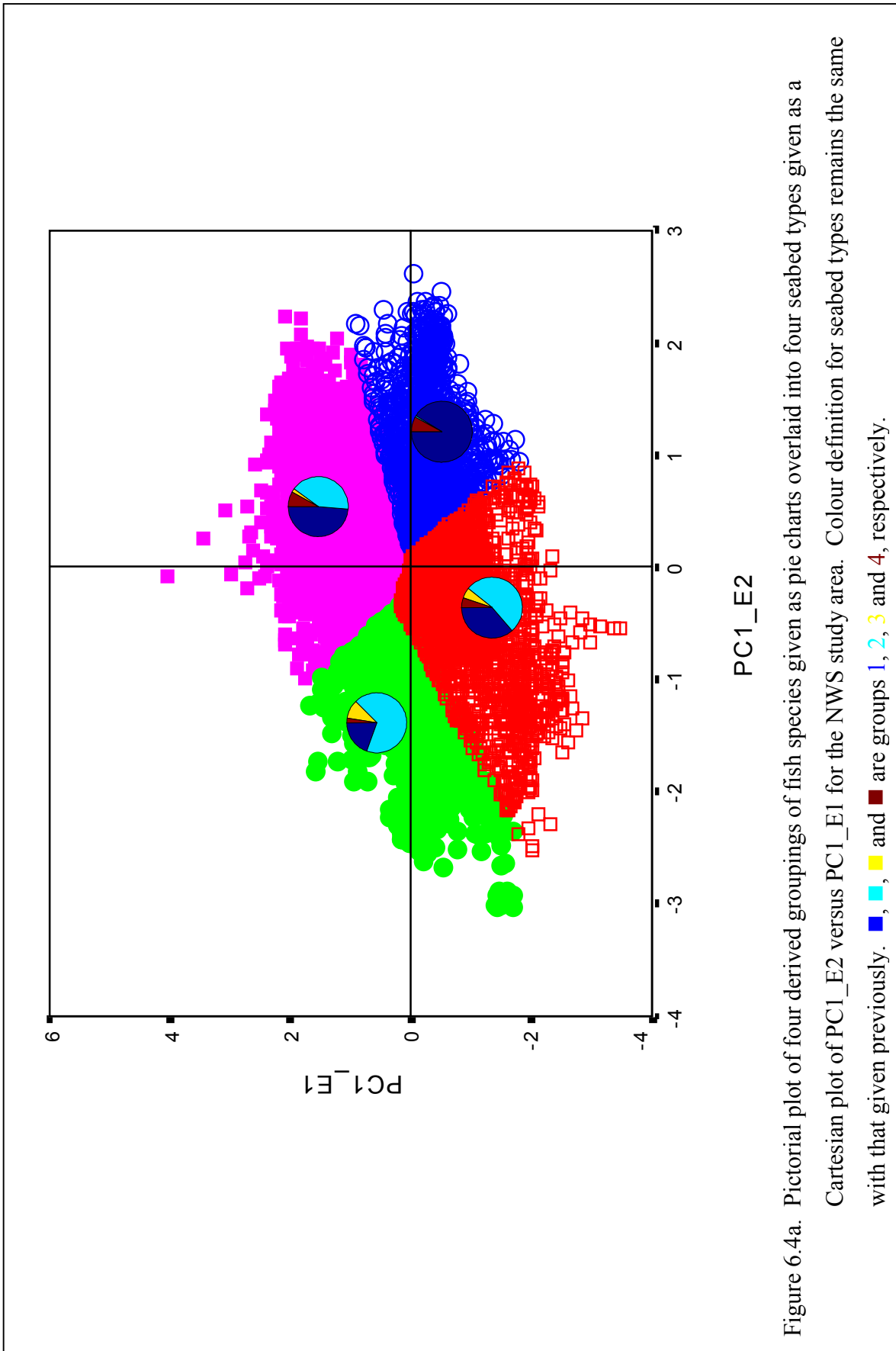
established for each study area are based on common species suggested by Sainsbury (1991) for the NWS region and by Bax *et al.* (1999) for the SEF region. The idea for having catch compositions in particular those of these groups of fish species in the clustering analysis is to put a certain weight of the contribution of the biological factors on the derived groupings of the quasi acoustic population since they have shown some sorts of association in any way with the benthic habitat of the seabed (Sainsbury, 1991; Bax *et al.*, 1999). For the NWS study area, the four fish groups are those from genera (1) *Lethrinus* and *Lutjanus*, (2) *Saurida* and *Nemipterus*, (3) *Epinephelus*, (4) *Parupeneus*. For the SEF study area, the fish groups are (1) Sharks, (2) Cods, (3) Flatheads, (4) Pelagic species mainly mackerel and trevally. The inclusion of the flatheads in the analysis for the SEF study area is rather ambiguous since they are acoustically difficult to sample as they usually reside on the bed. Nevertheless, they are included in the analysis due to the fact that they are abundant in the SEF study area and therefore may be associated in any way to a particular seabed type. To compensate for this dilemma, catch compositions provided from nets are included in the analysis. For each group of fish species, two variables are introduced, namely fish group compositions from the total catch in each trawl station (% by weight) and total weight of fish group normalised to the area trawled. While the former offers a contribution of the composition of derived groupings of fish species within their natural communities in each trawl station on the clustering analysis to derive groups of the quasi acoustic population, the latter provides the outcomes (groups of the quasi acoustic population) from the clustering analysis with a contribution of the concentration of derived groupings of fish species within their natural communities in each trawl station.

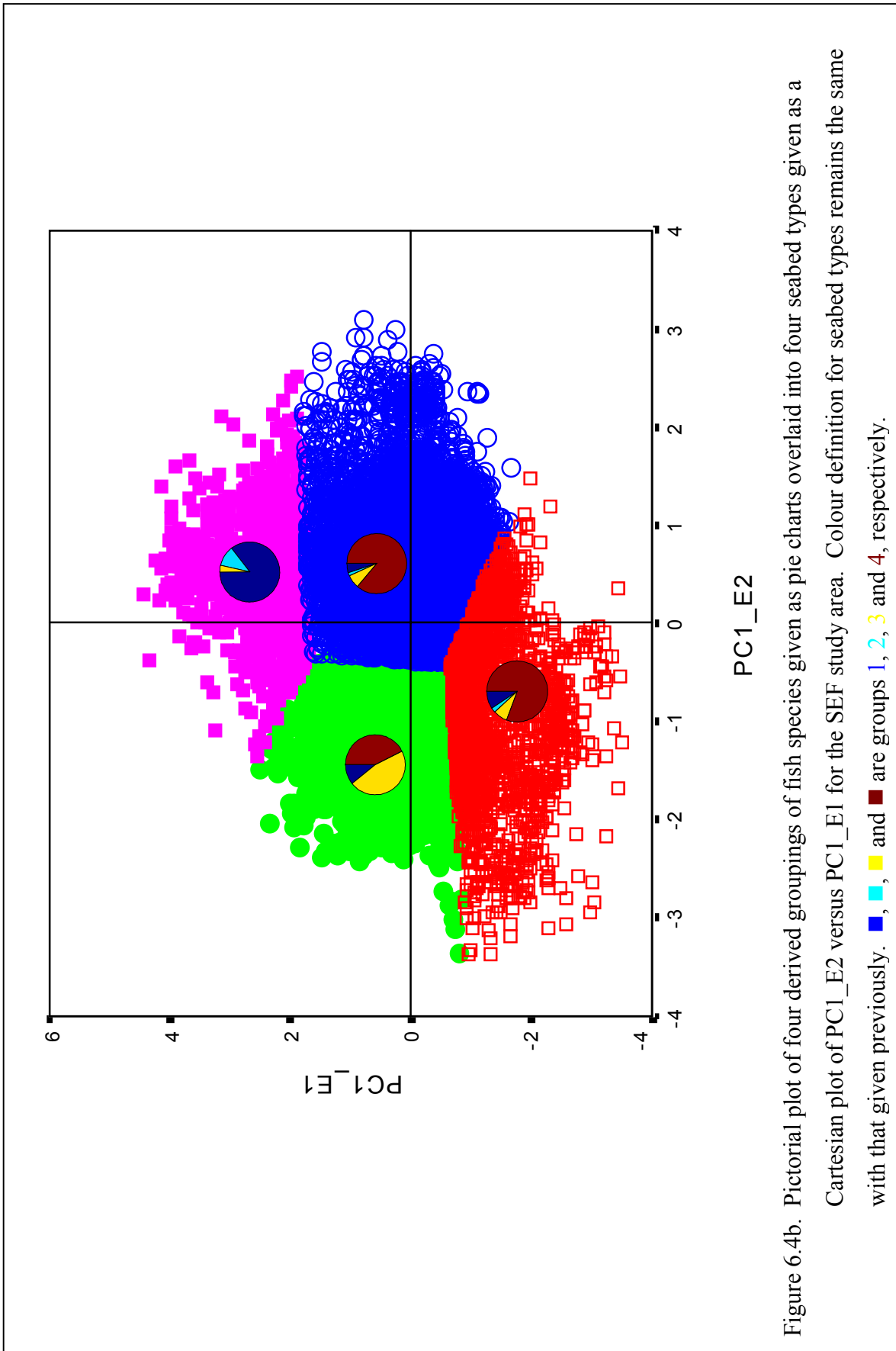
In Figures 6.4a and 6.4b, the relative abundances of four derived groupings of fish species shown as pie charts (based on figures given in Tables 6.1 and 6.2) is overlaid into a plot of PC1\_E2 versus PC1\_E1 of the four derived seabed types for the entire NWS and SEF regions. For the NWS study area as shown in Figure 6.4a, it appears that fish group 1 is dominant in hard seabeds whereas fish group 2 is dominant in soft seabeds. It is interesting to note that although fish group 1 appears dominant in soft seabeds, fish group 2 exists in a substantial proportion in HdRg seabeds as well whereas it completely disappears in HdSm seabeds. For soft seabeds, the presence of fish group 1 is also noticeable in less proportion than the dominant fish group 1.

For the SEF study area as shown in Figure 6.4b, fish group 4 appears to be dominant in smooth seabeds. Other fish groups are also present in smooth seabeds but in an insubstantial proportion in comparison to the dominant fish group 4. Although fish group 3 is dominant in SoRg seabeds, the presence of fish group 4 is noticeable in a substantial proportion as well. For HdRg seabeds, fish group 1 is dominant. Like smooth seabeds, the presence of fish groups other than the dominant one is insubstantial.

### 6.3.1. The first principal component of roughness and hardness indices versus the first principal component of area backscattering coefficient

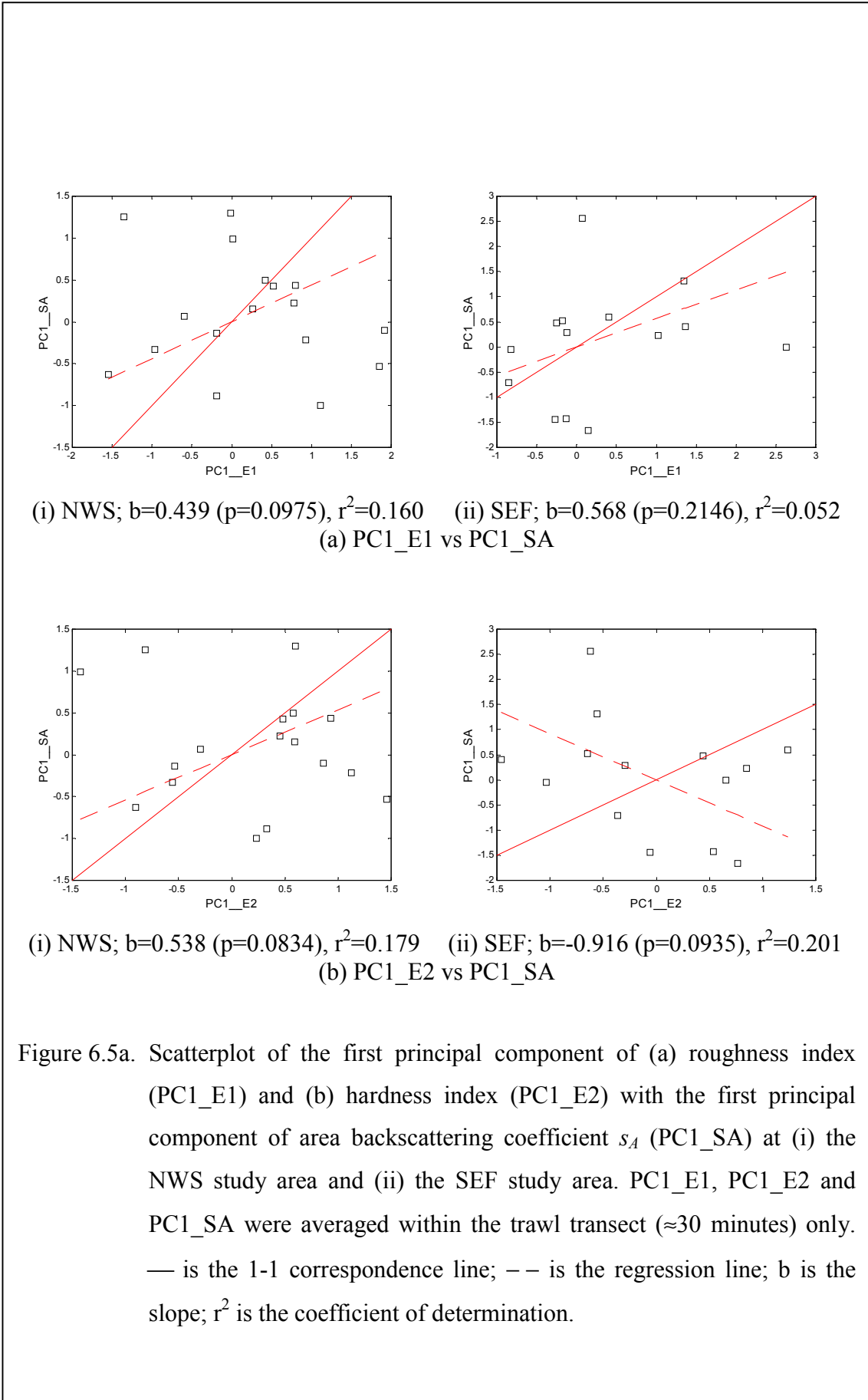
Figure 6.5a shows scatterplots of (a) PC1\_E1 versus PC1\_SA and (b) PC1\_E2 versus PC1\_SA in both study areas. Although Figure 6.5a(a) suggests that there might be an overall linear trend between PC1\_E1 and PC1\_SA in both study areas, there is extensive scatter. This might suggest that  $s_A$  estimates that represent relative measures of biomass increase with roughness index E1, the correlation however is







insignificant. The linear regression through the origin produces slopes ( $b=0.439$  and  $b=0.568$ ) significantly different from zero at  $\alpha=0.01$  ( $p=0.0975$  and  $p=0.2146$ ) but results in poor fits of the regression line ( $r^2=0.160$  and  $r^2=0.052$ ) shown as dashed lines for the NWS and SEF study areas, respectively. For hardness index versus relative measures of biomass, there is again an overall linear trend between PC1\_E2 and PC1\_SA for the NWS study area. In contrast, PC1\_SA seems to be independent of PC1\_E2 or perhaps to have an inverse relationship with PC1\_E2 for the SEF study area. The correlation is again insignificant. The linear regression through the origin produces slopes ( $b=0.538$  and  $b=-0.916$ ) significantly different from zero at  $\alpha=0.01$  ( $p=0.0834$  and  $p=0.0935$ ) but again gives poor fits of the regression line shown as dashed lines ( $r^2=0.179$  and  $r^2=0.201$ ) for the NWS and SEF study areas, respectively. The similar trend was observed when all the records of PC1\_E1, PC1\_E2 and PC1\_SA for both areas were averaged within 30-minute (equivalent to duration of trawl deployment), 25-minute, 20-minute, 15-minute, 10-minute and 5-minute intervals, but the scatter still persists. Representative examples for PC1\_E1 versus PC1\_SA and PC1\_E2 versus PC1\_SA averaged within 15-minute and 5-minute intervals from all the acoustic records are shown in Figures 6.5b and 6.5c, respectively. The similar trend showing in Figure 6.5a is evident in Figures 6.5b and 6.5c but again the correlations are insignificant. The coefficient of determination between PC1\_E1 and PC1\_SA and between PC1\_E2 and PC1\_SA for these time intervals was also produced. The objective was to observe the trend of the coefficient of determination with varying time intervals. No particular trend of the coefficient of determination was observed. It was also expected that varying time intervals substantially affected the range of variation of PC1\_E1, PC1\_E2 and PC1\_SA, separately. The analysis of variance (ANOVA) however revealed that the variances



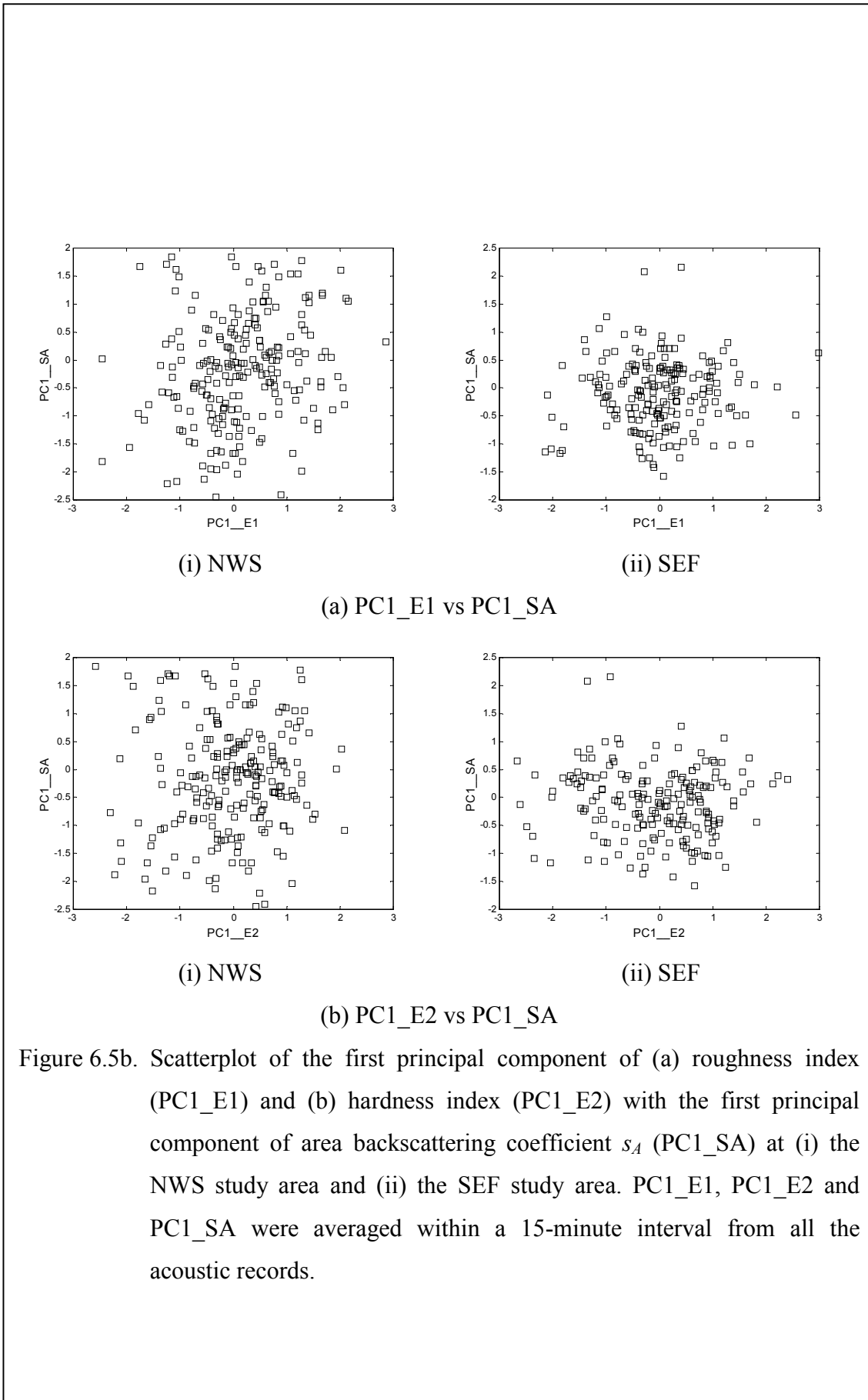
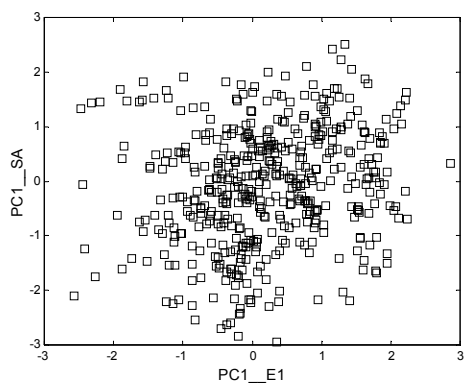
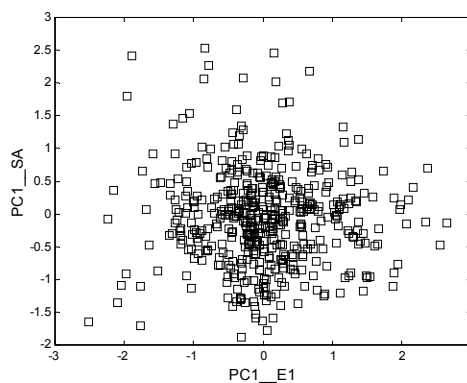


Figure 6.5b. Scatterplot of the first principal component of (a) roughness index (PC1\_E1) and (b) hardness index (PC1\_E2) with the first principal component of area backscattering coefficient  $s_A$  (PC1\_SA) at (i) the NWS study area and (ii) the SEF study area. PC1\_E1, PC1\_E2 and PC1\_SA were averaged within a 15-minute interval from all the acoustic records.

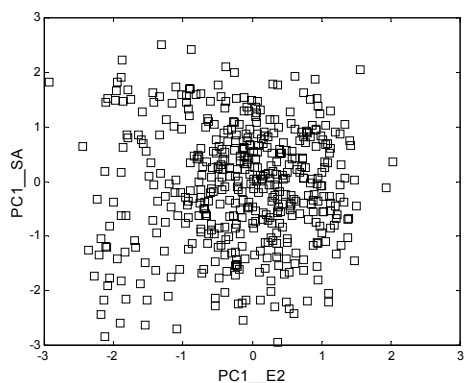


(i) NWS

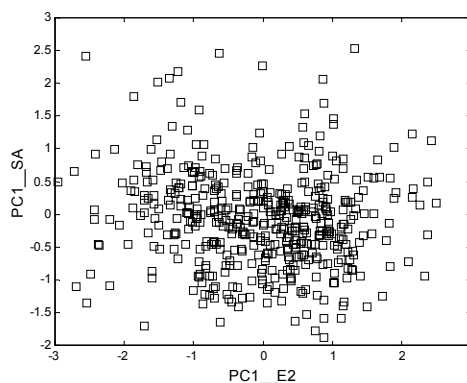


(ii) SEF

(a) PC1\_E1 vs PC1\_SA



(i) NWS



(ii) SEF

(b) PC1\_E2 vs PC1\_SA

Figure 6.5c. Scatterplot of the first principal component of (a) roughness index (PC1\_E1) and (b) hardness index (PC1\_E2) with the first principal component of area backscattering coefficient  $s_A$  (PC1\_SA) at (i) the NWS study area and (ii) the SEF study area. PC1\_E1, PC1\_E2 and PC1\_SA were averaged within a 5-minute interval from all the acoustic records.

of PC1\_E1 in the three different time intervals were not significantly different at  $\alpha=0.01$  ( $p=0.8661$  and  $0.3753$  for the NWS and SEF study areas, respectively), neither were the variances of PC1\_E2 ( $p=0.1679$  and  $p=0.9077$  for the NWS and SEF study areas, respectively) and the variances of PC1\_SA ( $p=0.4817$  and  $p=0.6701$  for the NWS and SEF study areas, respectively).

### 6.3.2. Fish group versus seabed type

The hierarchical agglomerative classification is calculated on a set of parameters shown in Tables 6.1 and 6.2 to produce a dendrogram for each study area as shown in Figure 6.6. The dendrogram given in Figure 6.6(a) suggests that trawl stations in the NWS study area may form two main groups of quasi acoustic population labelled as A and B in the dendrogram at an association level of 16 rescaled distance. By looking only at the acoustically derived seabed types as well as the benthic habitat types given in Table 6.1, group A belongs to the hard seabed type and group B to the soft seabed type. The dendrogram given in Figure 6.6(a) suggests that no further divisions are available for group A. Approximately at an association level of 16 rescaled distance however, group B seems to be divisible into two subgroups namely soft-rough and mostly soft-smooth seabed types, respectively. Nonetheless the division of group B into the two subgroups just mentioned is rather ambivalent due to the fact that only a small number of the soft-rough seabed type are involved in the analysis. In addition, trawl station 15, one of the soft-rough seabed types, is clustered into the soft-smooth seabed type rather than to its counterpart, trawl station 1. Trawl stations forming the two main groups of quasi acoustic population produced by the dendrogram shown in Figure 6.6(a) are overlaid into the acoustically derived seabed type along the vessel's track as shown in Figure 6.7a. The abundance of fish groups in percentage is given by the pie charts. Colour definition for the

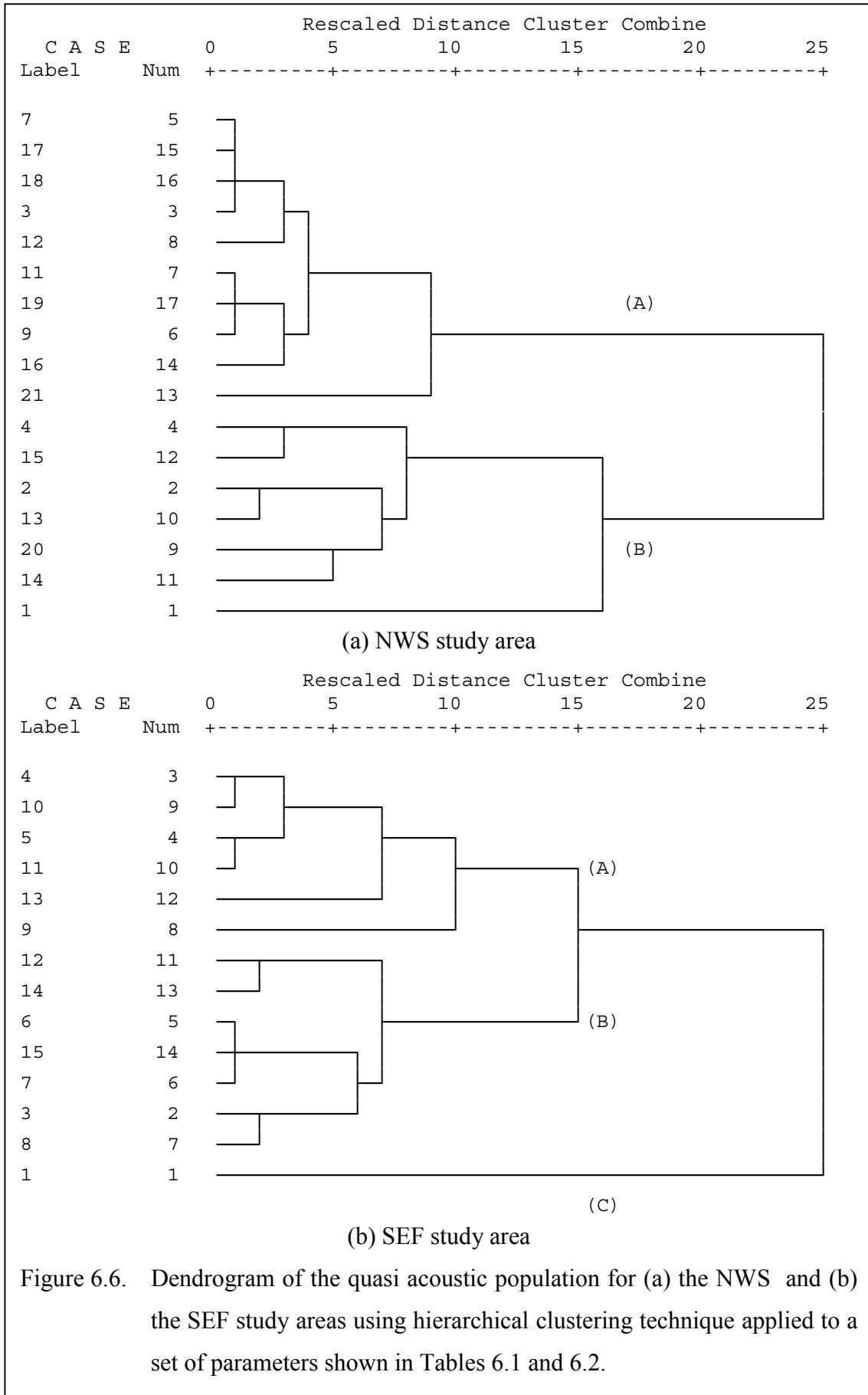


Figure 6.6. Dendrogram of the quasi acoustic population for (a) the NWS and (b) the SEF study areas using hierarchical clustering technique applied to a set of parameters shown in Tables 6.1 and 6.2.

acoustically derived seabed types remains the same. The magenta circles in Figure 6.7a are trawl stations that belong to group A and the red circles are those that belong to the other group (B). For pie charts in Figure 6.7a, fish groups 1 (from genera *Lethrinus* and *Lutjanus*), 2 (from genera *Saurida* and *Nemipterus*), 3 (from genus *Epinephelus*) and 4 (from genus *Parupeneus*) are shown as magenta, red, green and blue, respectively. From Table 6.1 and Figure 6.7a, it is clear that although fish groups 1 and 2 exist in the two main groups of quasi acoustic population shown in the dendrogram in Figure 6.6(a), the abundance of each of them in these two groups are however unique. In group A, fish group 1 is found to be the dominant in comparison to fish group 2. Only in trawl stations 3, 11 and 12, the abundance of fish group 2 is greater than that of fish group 1. In group B on the other hand, the abundance of fish group 2 is greater than that of fish group 1. Only in trawl stations 1 and 14, the abundance of fish group 2 is less than that of fish group 1. The two patterns of association between acoustically derived seabed types and fish groups seem to support the results from Sainsbury (1991). Genera *Lethrinus* and *Lutjanus* that form fish group 1 inhabit coral reefs or the vicinity of coral reefs, coralline lagoons, seagrass beds, rock areas, and also areas with flat bottoms and occasional coral outcrops, sponges, and sea whips (Froese and Pauly, 1998). Genera *Saurida* and *Nemipterus* that form fish group 2 are found over sandy or muddy bottoms of coastal waters (Froese and Pauly, 1998).

The SIMPER (Similarity Percentages) routine in the PRIMER (Plymouth Routines In Multivariate Ecological Research) developed by Plymouth Marine Laboratory, UK (Clarke and Gorley, 2001) was used to assess variables, listed in Table 6.1 used for the clustering analysis, that account primarily for the observed assemblage different between derived groups identified a posteriori from the

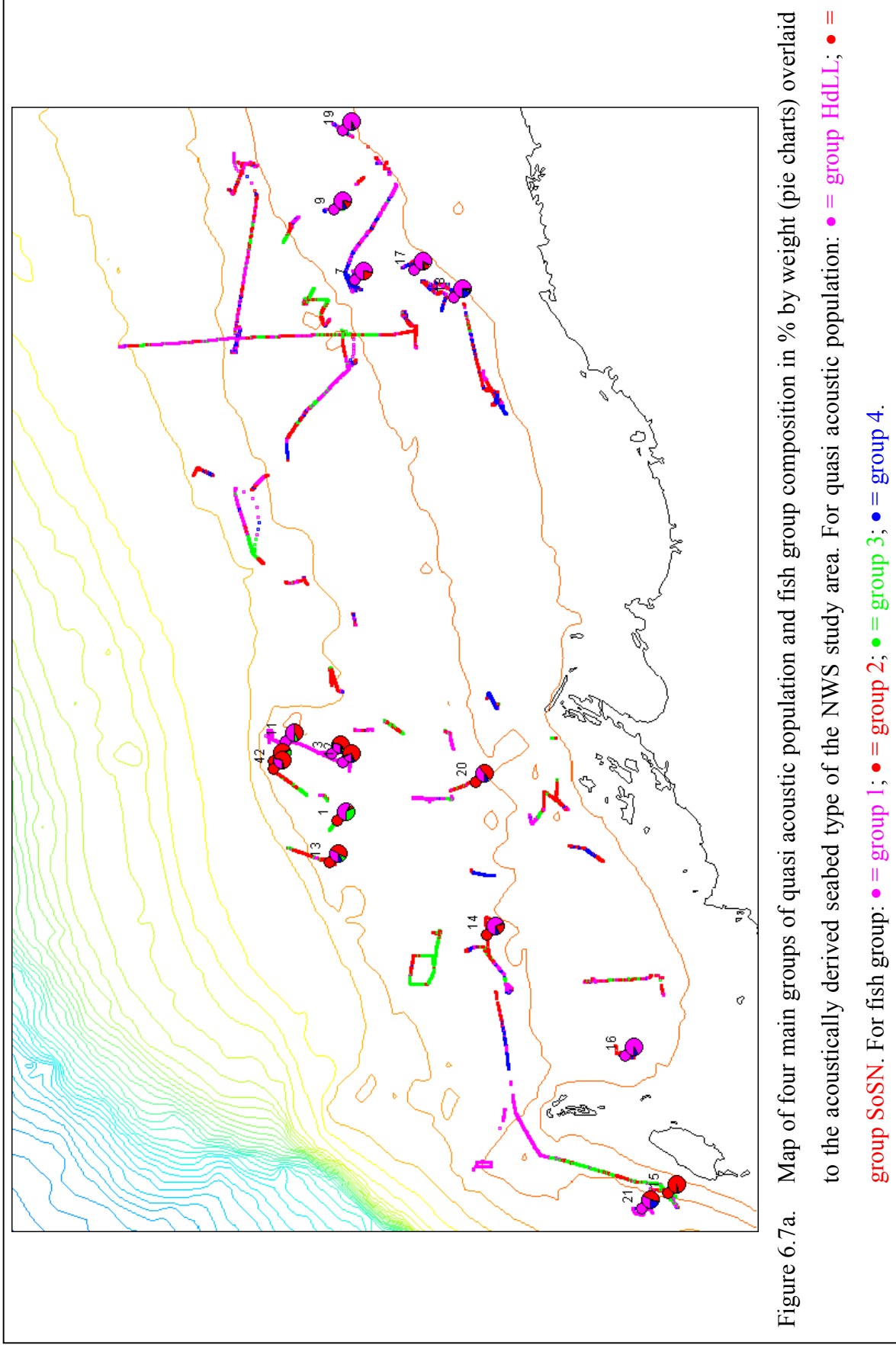


Figure 6.7a. Map of four main groups of quasi acoustic population and fish group composition in % by weight (pie charts) overlaid to the acoustically derived seabed type of the NWS study area. For quasi acoustic population: ● = group HdLL; ● = group SoSN. For fish group: ● = group 1; ● = group 2; ● = group 3; ● = group 4.



dendrogram shown in Figure 6.6(a) and that contribute to the similarity within a group. The SIMPER routine uses the Bray-Curtis similarity index in the analysis. Since this index is very sensitive to the magnitude difference in the variables being analysed, the relative abundances given in the tenth, twelfth, fourteenth and sixteenth column of Table 6.1 are logarithmically transformed ( $\log_{10}(x+1)$  where  $x$  is the relative abundance) and then standardised to have zero mean and unit variance. Note that a similar transformation and normalisation has already been applied to the acoustic parameters (see the derivation of PC1\_E1, PC1\_E2 and PC1\_SA in the previous sections and chapters). Since the SIMPER routine is workable only on the positive values, rescale is required to all variables but the catch compositions so that the minimum is equal to zero and the maximum is equal to unity. Since the catch compositions given in the ninth, eleventh, thirteenth and fifteenth column of Table 6.1 are in percentages, they are only converted to fractional form. The outcomes from the SIMPER routine are as follows:

The average dissimilarity of all pairs of inter-group parameters in groups A and B is 41.80. The seabed type parameters, both the averages and centroids of PC1\_E1 and PC1\_E2, account for most of dissimilarity between these two groups (50.50% of the overall dissimilarity of 41.80). Further 9.56% and 9.24% are contributed by fish groups 3 (from genus *Epinephelus*) and 4 (from genus *Parupeneus*), respectively, that cumulate to 69.29% of dissimilarity between these two groups.

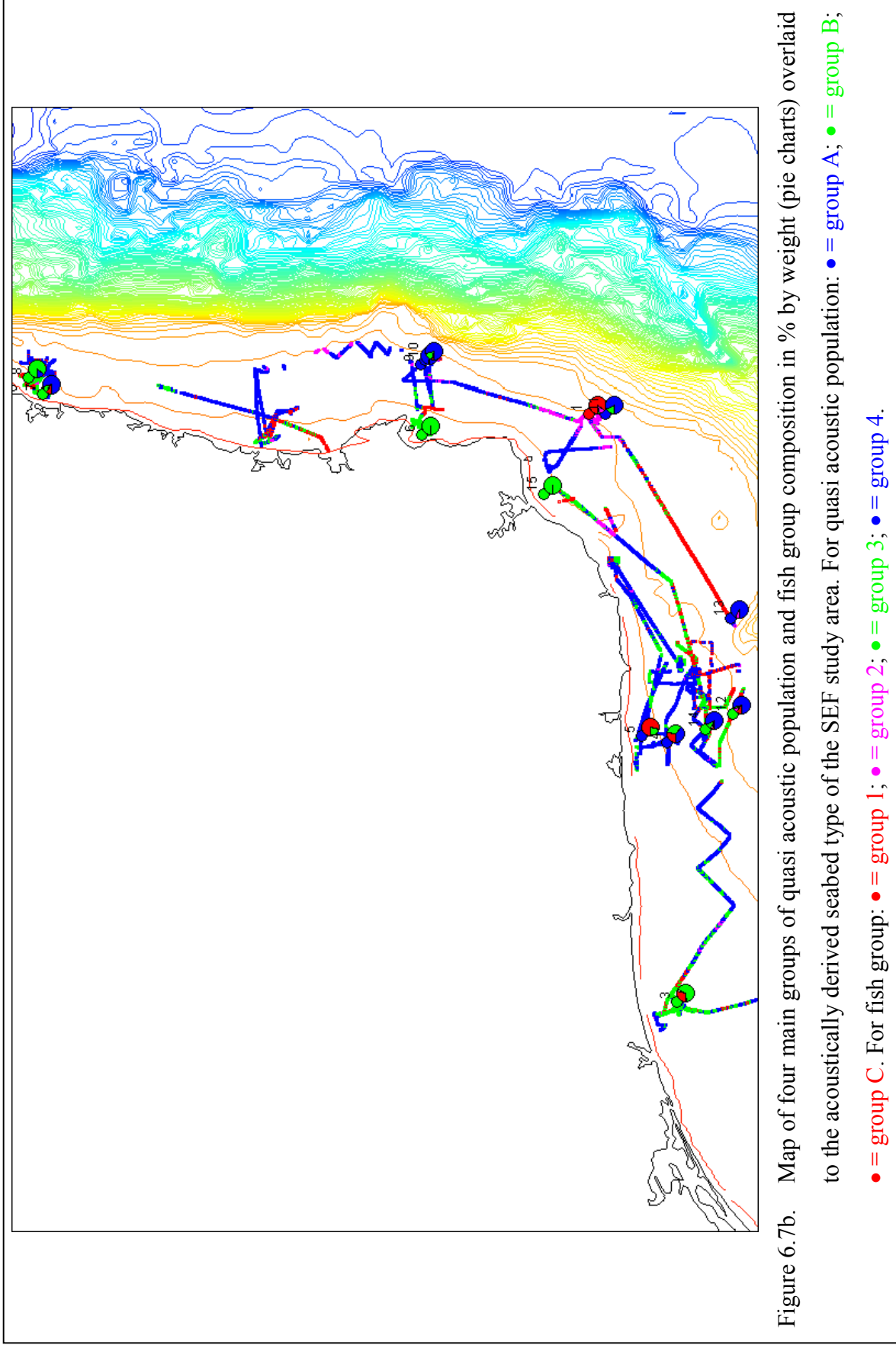
The average similarity of all within-group parameters is 76.29 and 67.58 for groups A and B, respectively. The seabed type parameters, both the averages and the centroids of PC1\_E1 and PC1\_E2, are typical of group A and account for 58.35% of

the overall similarity of 76.29. Fish group 1 (from genera *Lethrinus* and *Lutjanus*) is typical of group A and together with the seabed type parameters cumulate to 70.05% of the overall similarity. For group B however, all fish groups are common and account for most of similarity (67.16). Even though all fish groups are common in group B, fish group 2 (from genera *Saurida* and *Nemipterus*) contributed the most; one-half of the contribution of all fish groups.

The dendrogram shown in Figure 6.6(b) suggests that trawl stations in the SEF study area might form 3 main groups of quasi acoustic population at an association level of 15 rescaled distance or 2 main groups of quasi acoustic population at an association level of 16 rescaled distance (similar to that of two main groups of quasi acoustic population in the NWS study area). Having 2 main groups, however, does not provide a unique structure associated either with seabed types or fish groups. The former on the other hand reveals a unique structure associated with seabed types and fish groups in three derived groups of quasi acoustic population. As far as seabed types are concerned, group A belong to the hard-smooth seabed type, group B to the soft seabed type and group C to the hard-rough seabed type. Unlike group A and C, group B can be further divided into two subgroups approximately at an association level of 6 rescaled distance. The two subgroups of group B are the soft-smooth and soft-rough seabed types, respectively. Trawl stations clustered into the three main groups of quasi acoustic population produced by the dendrogram given in Figure 6.6(b) are overlaid into the acoustically derived seabed type along the vessel's track as shown in Figure 6.7b. As in Figure 6.7a, pie charts show the abundance of fish groups in percentage at each trawl station in the SEF study area. Colour definition for the acoustically derived seabed types remains the same. The three groups of quasi acoustic population produced by the dendrogram shown in Figure 6.6(b) are presented

in Figure 6.7b as blue, green and red circles, respectively. For pie charts in Figure 6.7b, fish groups 1 (Sharks), 2 (Cods), 3 (Flatheads) and 4 (Pelagic species) are shown as red, magenta, green and blue, respectively. Table 6.2 and Figure 6.7b clearly show that fish groups 1, 3 and 4 are common in the three main groups of quasi acoustic population. Bax *et al.* (1999) found that species, forming the three fish groups adopted here, were common to both Black Head and adjacent Disaster Bay macrohabitats but occurred at markedly different abundance. Fish group 2, on the other hand, exists only in an area associated with group C. This seems to contradict Bax *et al.* (1999) who found that cods occurred only on one bottom type (soft). This might be due to the fact that they included only two macrohabitats in their analysis whereas more than two macrohabitats were included in the analysis in this current study. In addition, Scott *et al.* (1974) found that cods, which form fish group 2, are mostly found on hard bottom over exposed rocky reefs. Two fish groups of interest are flatheads (fish group 3) and pelagic species (fish group 4). These two fish groups are common to groups A and B but occurred at different abundance in each group. As far as the abundance is concerned, the flatheads, which form fish group 3, are more dominant in group B (soft seabed type) than in group A (hard-smooth seabed type). Bax *et al.* (1999) also found that flatheads prefer to live in the soft seabed types. In contrast, pelagic species, which are mainly the mackerel and trevally that form fish group 4, seems to be more abundant in group A than in group B.

Prior to the introduction of the analysis of similarity percentages, all parameters adopted in the analysis as listed in Table 6.2, the similar transformation procedure applied to the NWS data (Table 6.1) was applied to adopted parameters of quasi acoustic population for the SEF study area presented in Table 6.2. The outcomes from the SIMPER routine for the SEF data are as follows:



The average similarities of all within-group parameters in groups A, B and C as shown in the dendrogram (Figure 6.6(b)) are 75.00, 62.08 and 100, respectively. Note that the high average similarity for group C is due to the fact that it has only a single member. The seabed type parameters, in particular the hardness ones (the average and the centroid of PC1\_E2), are typical of group A and account for 44.06% of the overall within-group similarity. The typical parameter accounting for most of the within-group similarity in group B is the relative abundance of fish group 3 (Flatheads) responsible for more than one-third of the within-group similarity. Although fish group 3 are common to groups A and B, they are more abundant in latter group than in the former one, accounting for only 14.42%. Since group C possesses only a single member, the breakdown similarity is not available.

The average between-group dissimilarity of all pairs in groups A and B is 45.55. The discriminator between these two groups is the seabed hardness which accounts for 23.13% of the between-group dissimilarity. Together with the acoustically derived relative abundance (PC1\_SA) of fish groups 3 (Flatheads) and 1 (Sharks), they make 56.00% of the overall between-group dissimilarity.

Cumulating to one-third of the overall between-group dissimilarity of 37.50, fish groups 4 and 2, respectively, are the discriminator between groups A and C. The seabed type parameters, both the averages and centroids of PC1\_E1 and PC1\_E2, are the other discriminator, cumulating to three-fourths of the overall between-group dissimilarity.

The average dissimilarity between groups B and C is 47.71. Of this, 14.02% and 12.87% are contributed by fish groups 3 (Flatheads) and 4 (Pelagic species), respectively, which cumulate to 26.90% of the overall between-group dissimilarity.

The contribution of the seabed type parameters is also evident and cumulates to 73.83% of the between-group dissimilarity.

## Chapter 7

### Conclusions

This chapter draws conclusions from the outcomes of this research. The results of all the experiments, assessments and analyses performed in this work are reviewed.

This thesis deals with the investigation of the use of acoustic data, including acoustic bottom returns and benthic-pelagic returns, together with other supportive data such as catch compositions. The motivation was to explore post classification techniques that effectively relate acoustically derived seabed types with fish groups above the seabed.

The study areas of this research comprised parts of north west continental shelf of Western Australia, known as the North West Shelf (NWS) region, and the south east continental shelf of Australia, known as the South East Fisheries (SEF) region. While the former is located in the Australian tropics, the latter is situated in a temperate, higher latitude region of Australia. Data collected in the NWS study area comprised 2 different surveys conducted between August and September 1995 and 1997 and data collected in the SEF study area comprised a single survey conducted between November and December 1996. From the FRV *Southern Surveyor*, acoustic data were collected along the vessel's track using a calibrated SIMRAD EK500 scientific echosounder with hull-mounted transducers of three different frequencies, 12 (single beam), 38 (split beam) and 120 kHz (split beam), and net data and

photographic data were collected in trawl stations selected following a stratified random technique using a McKenna demersal trawl and a 35 mm Photosea 1000 camera system.

*ECHO*, a software package developed by CSIRO Marine Research (Waring *et al.*, 1994; Kloser *et al.*, 1998), was the main software used for acoustic data quality control and post-processing (sections 3.1.2 and 3.13). This research has contributed to the enhancement of the *ECHO* software; one contribution in particular is an algorithm which corrects bottom spikes in the echo records at several frequencies arising from a time jitter that appears before and after the time of the first seabed return (section 3.1.2). Manipulation and integration of acoustic data were also conducted using the *ECHO* software. The main advantage of post-processing the acoustic data using the *ECHO* software over commercial bottom classifiers was that the quality of the acoustic data used for further analysis was guaranteed. Acoustic data of poor quality may cause a misclassification of the bottom type. Performing data quality control then becomes imperative. In contrast, had the available commercial bottom classifiers operated directly on the original data, they would have provided very significant misclassification (section 5.2).

As mentioned earlier, Kloser *et al.* (2001) in sections 1.4.2 and 5.2, and Voulgaris and Collins (1990) in section 1.4.2 experienced a depth dependence in their RoxAnn data. The depth dependence of RoxAnn data was an artefact that was not numerically meaningful and could not be explained by differences in bottom types (Kloser, *et al.*, 2001; Hamilton *et al.*, 1999). In this study, the assumption adopted was that roughness and hardness indices (E1 and E2) were independent of depth (section 2.4). For roughness index (E1), this was achieved by introducing a constant



angular integration interval to the tail of the first bottom returns (section 2.4.1). A constant angular integration interval was selected such that a similar tail sector was integrated regardless of the beamwidth of the three operated transducer units. It was found that the angular integration interval between  $14^{\circ}$  and  $40^{\circ}$  plus an offset of 3 m for 38 and 120 kHz produced roughness index (E1) independent of depth (sections 2.4.1 and 5.2). For 12 kHz, the angular integration interval between  $0^{\circ}$  and  $35^{\circ}$  plus an offset of 9 m produced roughness index (E1) independent of depth (sections 2.4.1 and 5.2). Unlike roughness index (E1) where a constant angular integration interval was required to produce results independent of depth, hardness index (E2) was derived by using a constant depth integration interval. The results remained independent of depth (section 5.2). This is understandable since the integration includes the complete second bottom echo that is dominated by coherent components (particularly in the initial sector of acoustic bottom returns) which are merely depth-dependent.

For the NWS study area, training sites were selected based on information provided by underwater photographs. From the photographic data, four different seabed types were established using the hierarchical agglomerative clustering technique (section 5.3). Twenty one trawl stations were available for bottom classifications and assessments of the relationship between acoustically derived seabed types and fish groups. Homogeneous seabed types were selected from the 21 trawl stations available as training sites. Unlike the NWS study area, the SEF study area involved a collection of acoustic data in reference sites.

To reduce the dimensionality of acoustic data formed by data from three different frequencies, principal component analysis was used. A simple iterative

relocation clustering technique (*k*-means) was applied to the dimensionally reduced acoustic data resulting from the principal component analysis to form four seabed classes as suggested by the photographs (section 5.3). A feedback system was found to be effective to minimise the sensitivity of the initial centroids of PC1\_E1 and PC1\_E2 of the four seabed classes in the *k*-means clustering technique. The centroids adopted were those employed as the initial centroids that did not change after all class assignments had finished (section 5.3). As far as the ground truth is concerned, the technique adopted here works well (section 5.5). The accuracy of the technique determined from the confusion matrices (acoustically derived seabed types versus ground truth) was 84% for the NWS study area and 93% for the SEF study area. It is important to note however that these estimates involve a very small number of ground truth sites.

Results from training sites for the NWS study area and reference sites for the SEF study area were then used to classify the remaining data in both study areas independently (section 5.4). Results along the vessel's track were in good agreement with previous studies by Jones (1973) and Sainsbury, *et al.* (in prep) for the NWS study area and by Jones and Davies (1983), Bax *et al.* (1999) and Kloser *et al.* (2001) for the SEF study area. Distances in which variations of along-track bottom types occurred were shorter in the NWS study area than they were in the SEF study area. This might be related to the fact that the NWS study area is located in the tropics whereas the SEF study area is in the temperate, higher latitude region. A high diversity and a moderate abundance of resources are characteristic of tropical regions.

A comparative analysis of the area backscattering coefficients derived from acoustics and nets from the North West shelf data set at 38 kHz showed that the two

estimates were equivalent in that over long times and track lengths they both clustered around the 1-1 correspondence line (section 4.5). Substantial short term variations were observed however, which might involve uncertainties in target strength, net efficiency and spatial distribution of targets. In respect to the latter, comparison of the series of the acoustic area backscattering coefficients at the three operated frequencies, with different beamwidths, confirmed the influence of spatial distribution of targets on the acoustically derived area backscattering coefficients at the varying cross track scale lengths involved (section 4.5). Results from this comparison led to a decision to use acoustic  $s_A$  estimates rather than net derived  $s_A$  estimates to investigate the relationship between acoustically derived seabed types and fish groups using the concept of acoustic population proposed by Gerlotto (1993) applied to the hierarchical agglomerative clustering technique. The decision to use the acoustic data seems to favour Gerlotto (1993) although he, himself, did not provide reasons to select only the acoustic data into his concept of acoustic population. Unlike his acoustic population, the acoustic population established here involved also catch composition data derived from the analysis of data from net trawls undertaken at the time of acoustic data acquisition and was called the quasi acoustic population.

Examinations of series of the roughness and hardness indices at the three operated frequencies indicated that the autocorrelation characteristic lengths derived from along-track measurements were much less than average transect spacing (section 5.6). It would appear that the spatial variability of these two indices involved distances smaller than the average transect separation necessary on the survey. This suggested that interpolation into unsampled area was possible only within the autocorrelation characteristic length, i.e. within lengths of 3 nmi only, using variograms and Kriging techniques. Sample variograms of PC1\_E1 and PC1\_E2 for

the two study areas fit well a spherical model ( $r = 0.92$  and  $0.89$  for respectively PC1\_E1 and PC1\_E2 in the NWS study area, and  $r = 0.99$  and  $0.97$  for respectively PC1\_E1 and PC1\_E2 in the SEF study area). Variograms in all directions in  $10^{\circ}$  increments for both study areas were derived to observe any direction trends. However, nothing of interest was found in any directions in both study areas. In other words, the variogram models are omnidirectional which suggests that the spatial correlation structure is much the same irrespective of direction in both study areas. In order to produce a full 2-D bottom structure in the entire study area, a closer transect separation is required.

Following the concept of acoustic population introduced by Gerlotto (1993), the relationship between seabed type and common species was investigated by using the hierarchical agglomerative clustering technique (section 5.7). Two main groups of quasi acoustic population were found in the NWS study area whereas three main groups were evident in the SEF study area (section 5.7). Those main groups of quasi acoustic population in both study areas had a strong association with the acoustically derived seabed types, i.e. hard and soft seabed types for the NWS study area and soft, hard-smooth and hard-rough seabed types for the SEF study area. While the two main groups of the quasi acoustic population in the NWS study area were not divisible into small, meaningful subgroups, one of the three main groups of the quasi acoustic population in the SEF study area provided two small, meaningful subgroups in terms of seabed types; soft into soft-smooth and soft-rough seabed types. These main groups of quasi acoustic population in both study areas were also highly associated with fish groups adopted in this study. Species contributing most to the dissimilarity between groups of acoustically derived bottom types in the NWS study area were from genera *Lethrinus* and *Lutjanus* that form fish group 1 and from genera

*Nemipterus* and *Saurida* that form fish group 2. Despite the fact that they were common to all acoustically derived bottom types, they occurred at different abundances in different bottom types. Fish group 1 prefers to live in hard-rough seabed types whereas fish group 2 seems to exist mainly in soft-smooth and soft-rough seabed types. Likewise, few species from shark, cod, flathead and pelagic species (mainly mackerel and trevally) in the SEF study area were common to all seabed types but occurred at different abundances. Two fish groups of interest in the SEF study area were flatheads (fish group 3) and pelagic species (fish group 4). The flatheads (fish group 3) dominantly occupy group B (soft seabed type). In contrast, pelagic species (fish group 4), mainly the mackerel and trevally, prefer to abundantly exist in group A (hard-smooth seabed type).

Apart from its use to reduce the dimensionality of the acoustic data formed from three operated frequencies, Principal Component Analysis has shown its strength to also detect anomalies that might occur at acoustic data from individual frequencies due to the ratio between the wavelength and the scale of roughness and the effect of the beamwidth of the transducer units employed on the slope of the seabed in the case of acoustic bottom returns and on the highly patchy distribution of fish in the case of acoustic pelagic returns (sections 5.2 and 5.7). In the case of acoustic bottom returns, three patterns of component loadings were observed in those first principal components with low variances (section 5.2). Two of them, i.e. the first and third patterns, may be associated with surface irregularities of the seabed and beamwidths of the operated transducer units. The first pattern denoted by high component loadings in E1\_38 and E1\_120 and a low component loading, sometimes with different sign, in E1\_12 might be an indication of the scale of surface irregularities. The third pattern verified by high component loadings in E1\_12 and

E1\_120 and a low component loading, sometimes accompanied with different sign, in E1\_38 might indicate the sensitivity of the narrow beamwidth to slopes and sudden drops or rises of the seabed surface. In the case of benthic-pelagic returns, two patterns in component loadings of the first principal component with low variances were observed (section 5.7). The first pattern was indicated by the first principal component having the component loadings in SA\_12 and SA\_120 ( $s_A$  estimates at 12 kHz and 120 kHz, respectively) very similar but very different from the component loading in SA\_38 ( $s_A$  estimates at 38 kHz). The second one was indicated by the first principal component having the component loadings in SA\_38 and SA\_120 very similar but very different from the component loading in SA\_12. Series of  $s_A$  estimates at the three operated frequencies confirmed that these two patterns arose from the effect of the differing beamwidth of the three operated transducer units on a highly patchy distribution of fish. This was indeed another indication to prove substantial variation between acoustically derived  $s_A$  estimates and net derived  $s_A$  estimates. In addition, it was found that these two patterns occurred more frequently in the NWS study area than in the SEF study area.

## Appendix A

### Principal Component Analysis and Cluster Analysis

#### A.1. Principal Component Analysis (PCA)

PCA is a dimension-reduction technique. It attempts to reduce the dimensionality of a data set comprised of a large number of interrelated variables with minimal loss of sample variation (information). The process includes an orthogonal transformation from the axes representing the original variables into a new set of axes called principal components (PCs). The new axes or the PCs are uncorrelated one to another and are ordered in such a way that the first few PCs hold as much of the total variation present in the original data set as possible. While PCs, from a geometric point of view, are orthogonal projections of all the original variables, PCs algebraically are linear combinations of the original variables. In addition, a linear combination of variables is an essential concept in multivariate analysis and is indeed fundamental to PCA.

Let  $\mathbf{X}$  be the  $p \times n$  data matrix where  $p$  is the number of variables and  $n$  is the number of objects. Modifying the equation presented in Harris (1975) into a matrix notation, the linear combinations to derive PCs are defined by

$$\mathbf{PC}_j = \mathbf{v}_j^T \mathbf{X} \quad (\text{A.1})$$

where  $\mathbf{PC}_j$  is the  $j^{\text{th}}$  principal component (new variable or dimension) and  $\mathbf{v}_j^T$  is the transposed column vector ( $1 \times n$  matrix) of constants  $\mathbf{v}_j$ . The normalisation constraint

$\mathbf{v}_j^T \mathbf{v}_j = \mathbf{1}$  (Harris, 1975) on the components of  $\mathbf{v}_j$  is usually adopted. The variances of PCs are then defined by (Johnson and Wichern, 1998; Timm, 1975)

$$\text{Var}(\text{PC}_j) = \mathbf{v}_j^T \mathbf{C} \mathbf{v}_j \quad (\text{A.2})$$

where  $\mathbf{C}$  is the sample covariance matrix of the variables in  $\mathbf{X}$ . The solution of  $\mathbf{v}_j$  involves the eigen equation, which is defined by (Harris, 1975; Timm, 1975)

$$(\mathbf{C} - \lambda_j \mathbf{I}) \mathbf{v}_j = 0 \quad (\text{A.3})$$

where  $\lambda_j$  is the eigenvalue and  $\mathbf{v}_j$  is the vector of coefficients in equation (A.1) and is also the corresponding eigenvector. The values of new variables can be found simply by substituting the eigenvector  $\mathbf{v}_j$  into equation (A.1). These transformed values are called Principal Component scores (PC scores). The covariance matrix  $\mathbf{C}$  of variables in data matrix  $\mathbf{X}$  becomes the correlation matrix  $\mathbf{R}$  of matrix  $\mathbf{X}$  if  $\mathbf{X}$  has zero mean and unit variance. In other words, variables in  $\mathbf{X}$  have first been standardised. In practice, it is also common to scale the eigenvectors such that their sum of squares is equal to the corresponding eigenvalues (i.e.  $\mathbf{e}_j^T \mathbf{e}_j = \lambda_j$  where  $\mathbf{e}_j$  is the scaled eigenvectors). The scaled eigenvector is then defined by

$$\mathbf{e}_j = \lambda_j^{-\frac{1}{2}} \mathbf{v}_j \quad (\text{A.4})$$

These scaled coefficients are called component loadings. The component loading  $\mathbf{e}_j$  can also be a measure of the correlation between the PCs and the original variables in  $\mathbf{X}$  if  $\mathbf{X}$  is a matrix of standardised variables (unit variance). In such cases, they are referred to as component correlations.



Let  $\mathbf{E}$  be the matrix of the series of eigenvectors ( $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p$ ) and  $\mathbf{L}$  be the diagonal matrix of series of the corresponding eigenvalues ( $\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_p$ ), the general form of the eigen equation becomes

$$(\mathbf{C} - \mathbf{L}\mathbf{I})\mathbf{V} = 0 \quad (\text{A.5})$$

and the eigenvectors satisfy  $\mathbf{V}^T\mathbf{V} = \mathbf{V}\mathbf{V}^T = \mathbf{I}_p$ . It is important to note that  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$  and  $\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_p$  correspond to the first, second,  $\dots, p^{\text{th}}$  principal components ( $\text{PC}_1, \text{PC}_2, \dots, \text{PC}_p$ ), and that the eigenvalues are the respective variances of the different PCs. In addition, the result, which follows the solution of equation (A.5), is that the sum of the variances of the original variables is equal to the sum of the variances of PCs (i.e. the eigenvalues).

The PCA of data is critically scale-dependent. The scale of variables, which are in particular not comparable one to another, plays a significant role in influencing the outcomes of the PCA. When variables are not comparable from one to another, either in terms of magnitude or unit, their variances are not comparable and are different in magnitude. Those variables with large variances will consequently dominate the first few PCs of the covariance matrix whatever the correlation structure might be. A widely acceptable solution to this problem is to initially standardise all the variables such that they have unit variance before introducing PCA. In other words, the PCA is performed on the correlation matrix of the data instead of the covariance.

The results of PCA arising from the correlation matrices for different sets of random variables are more directly comparable than those derived from the covariance matrices. This is understandable since the PCs based on the covariance

matrices are very sensitive to the units or scales used to measure the variables. This then becomes a primary argument for using the correlation matrices rather than covariance matrices. To avoid the scale-dependent problem, E1 and E2 parameters, and area backscattering coefficient,  $s_A$  (defined in section 3.1.3.2 and equations 3.4 to 3.6), used in this study were first standardised separately to have unit variance.

Since one objective of using PCA in this study was to reduce the number of variables needed in the cluster analysis and intercorrelation within individual parameters was high, only the first principal component (PC1) of each individual set of parameters was taken and used in the cluster analysis. Nonetheless, the total variation, in percentage, of PC1 (> 70 %) and the scree plot, a plot of eigenvalues with sequential numbers of principal components, were also used to monitor whether or not PC1 alone was representative enough (sections 5.2 to 5.5 and 5.7).

The first principal component of E1 parameters (PC1\_E1) and the first principal component of E2 parameters (PC1\_E2) were then used in cluster analysis, the iterative relocation (*k*-means) technique (see discussion in section A.2.3). When investigating the relationship between seabed types and demersal fish communities (section 5.5), PCA was applied also to the area backscattering coefficient,  $s_A$ , to produce the first principal component of  $s_A$  estimates (PC1\_SA). Like E1 and E2 parameters,  $s_A$  estimates of the three operated frequencies were highly intercorrelated so that the first principal component was sufficient to account for most of the variation in the original  $s_A$  estimates. Results (PC1\_SA), together with the mean of PC1\_E1, PC1\_E2, and key species compositions in those trawl stations, were introduced to the hierarchical agglomerative method.

## A.2. Cluster Analysis (CA)

The clustering problem is defined as follows: The set of  $n$  objects  $X = \{X_1, X_2, \dots, X_n\}$  is to be clustered. Each  $X_i \in \mathfrak{R}^p$  is an attribute vector consisting of  $p$  real measurements describing the object. The objects are to be clustered into non-overlapping groups  $C = \{C_1, C_2, \dots, C_k\}$  ( $C$  is known as a clustering), where  $k$  is the number of clusters,  $C_1 \cup C_2 \cup \dots \cup C_k = X$ ,  $C_i \neq \emptyset$ , and  $C_i \cap C_j = \emptyset$  for  $i \neq j$ . The objects within each group should be more similar to each other than to objects in any other group, and the value of  $k$  may be unknown. If  $k$  is known, the problem is referred to as the  $k$ -clustering problem.

### A.2.1. Measure of similarity

Most efforts to produce a rather simple group structure from a data set involve a process of measurements of similarity. A small distance between objects should indicate a high similarity. Thus a distance measure can be used to quantify dissimilarity.

Let  $d(P, Q)$  be a distance between two points  $P$  and  $Q$  and  $R$  be an intermediate point. The distance is a metric measure only if it holds the following conditions (Anderberg, 1973; Everitt, 1974; Gordon, 1981).

$$\begin{aligned} d(P, Q) &= d(Q, P) \quad \text{for all } P \text{ and } Q \text{ in } S \\ d(P, Q) &> 0 \quad \text{if } P \neq Q \\ d(P, Q) &= 0 \quad \text{if } P = Q \\ d(P, Q) &\leq d(P, R) + d(R, Q) \quad \text{for all } P, Q \text{ and } R \text{ in } S \end{aligned} \tag{A.6}$$

The first property imposes symmetry. It requires that the distance between  $P$  and  $Q$  be the same as the distance between  $Q$  and  $P$ . The second property indicates that positive distances are required. The third property implies that point  $P$  is zero

distance apart from Q, i.e. any points with zero distance apart must be identical. The fourth property is referred to as the triangle inequality. This simply states that the length of one line of the triangle is equal to or less than the sum of the lengths of the other two lines. This property has also been called the metric inequality. Measures that are not metric may not be jointly monotonic, i.e. the value of different measures used with the same data will not necessarily vary conjointly. This raises the disturbing issue that these measures could suggest quite different relationship among the entities (data points). The Pearson product-moment correlation coefficient is an example of a non-metric measure. It fails to meet the second criterion and in many applications may well not meet the fourth criterion.

There are several distance measures available in the literature for clustering. Perhaps the most popular and familiar distance measure however is the Euclidean metric and was adopted here in this current study (sections 5.3 and 5.4). The Euclidean distance is actually a special case of the more general Minkowski distance. Let  $\mathbf{x}=[x_1, x_2, \dots, x_p]^T$  and  $\mathbf{y}=[y_1, y_2, \dots, y_p]^T$  be two  $p$ -dimensional objects. The Euclidean distance between these two objects is defined as

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} = \sqrt{(\mathbf{x} - \mathbf{y})^T (\mathbf{x} - \mathbf{y})} \quad (\text{A.7})$$

$d(\mathbf{x}, \mathbf{y})$  is the straight line distance between the two points representing the objects. The Euclidean distance becomes unsatisfactory if variables are of different scale. Like in PCA, solution to this problem in this study was to first standardise all the variables such that they had zero mean and unit variance before evaluating the Euclidean distance. The Euclidean distance of standardised variables will preserve relative distances. An alternative distance measure is the city-block distance

(Manhattan distance). This measure however is problematic if the attributes are correlated. The Mahalanobis distance is a standardised form of the Euclidean distance. This measure scales data in terms of standard deviations and adjusts for intercorrelations between the variables. It is however not as commonly used as the Euclidean as prior knowledge of the clusters are required to compute the correlation matrix.

#### A.2.2. Hierarchical agglomerative methods

The hierarchical clustering technique is a clustering technique that produces a solution in which some groups are nested within other groups. The technique may be subdivided into two methods namely agglomerative and divisive methods. The former proceeds by a series of successive fusions of  $n$  entities (objects). The latter, as its name implies, works in the opposite direction of the former. It divides an initial set of objects into finer groupings. Both methods operate on a matrix of similarity or distance measures between objects. Results from both methods may be presented in the form of a dendrogram, a 2-dimensional diagram illustrating the fusions (for agglomerative methods) or divisions (for divisive methods) which have been made at each stage of the procedure. The basic procedure with all these methods is similar and is as follows (Anderberg, 1973; Johnson and Wichern, 1998).

1. Begin with  $n$  groups as many as the number of objects and the symmetrical matrix of similarity or distance measures between objects  $\mathbf{D} = d_{ij}$ .
2. Find from the symmetrical distance matrix  $\mathbf{D}$  the most similar pair of groups. Let  $k$  and  $l$  be the most similar pair, and  $d_{kl}$  be the distance between them.

3. Fuse the most similar pair  $k$  and  $l$ . Label the merger of the most similar and update the distance matrix to take into account the similarity measures between the new the new group and all other existing groups. Delete rows and columns corresponding to the two merging groups  $k$  and  $l$ .
4. Repeat steps 2 and 3 a total of  $n-1$  times until all objects merge into a single, big group. Record merging groups and the level of similarity between them at each stage when mergers take place.

Results at each stage of the fusions are then described completely by the dendrogram.

Hierarchical methods are distinguished primarily by ways of defining similarity or distance measures either between an observation and a group of several objects or between two groups of objects. There are four important agglomerative clustering algorithms: single-linkage, complete-linkage, average-linkage and Ward's minimum variance method. For single-linkage (nearest neighbour), the distance between two groups is the distance between the two nearest objects in those groups. Problems occur however when groups are poorly delineated; this method can result in long chains with dissimilar objects at the ends. Complete-linkage (furthest neighbour) joins the two groups with the minimum distance between their two furthest objects. This thus eliminates the chaining problem experienced with single-linkage clustering. Average-linkage (group-average clustering) defines the distance between two groups to be the average distance from all objects in one group to all objects in the other groups. This approach however tends to combine groups with small variances. In addition, this method is biased towards producing groups with approximately equal variance. Centroid clustering, a variant of the group average clustering, defines the distance between two groups as being the distance between their centroids. This

approach however can produce messy and confusing results, since the centroids move as groups are combined. Furthermore, the distance between two groups may be less than the distance between the centroids of groups merged at earlier stage.

Unlike the above clustering techniques, Ward's minimum variance technique optimises an objective statistic, the sum of the squared distance between each object and its group centre. At each step, the algorithm merges the groups that will minimise the increase of this statistic. This technique differs from the centroid clustering method and its variant in which it involves weighting of the distance centroids when distances are computed. Consequently, the error sum of squares never decreases as the clustering process proceeds. Hence, this method is not subject to reversals, which happen in the centroid method and its variant. Since comparison studies show that this technique clusters relatively accurately over a wide range of data types, it was used in this study (section 5.7). Since the hierarchical agglomerative clustering method is restricted to smaller data sets due to the need to store similarity matrices, it was used in conjunction with Ward's minimum variance linkage method only when investigating the relationship between demersal fish communities and seabed types (section 5.7). The reasons were (1) the number of data sets was small, and (2) levels of mergers were easily monitored.

#### A.2.3. Iterative relocation (*k*-means) technique

The iterative relocation technique in this study for seabed classification (sections 5.3 and 5.4) was used due to the following considerations. Unlike hierarchical methods, which require the calculation and storage of a matrix of similarities or distances between objects, iterative relocation methods are implemented directly upon the raw data. They therefore offer the opportunity of

handling larger data sets than do hierarchical methods. Moreover, the iterative relocation method makes more than one passage through the data and can compensate for a poor initial partition of the data. This thereby avoids one of the major drawbacks of the hierarchical method. This method also produces single-rank clusters that are not nested and are therefore not part of a hierarchy. Most iterative relocation techniques do not permit overlapping groups.

Unlike hierarchical clustering techniques that attempt to converge to an empirical number of groups which satisfy the implied group definition, i.e. the exact number of groups being a consequence of data characteristic relative to the clustering criteria employed, the iterative relocation technique best clusters the objects into a fixed (prespecified) number of groups. As the iterative relocation technique is designed to cluster objects, rather than variables, into a collection of  $k$  groups, it starts with a prespecified set of group centres (or centroids). In this technique, a modification of the initial partition into  $k$  groups is implemented by moving objects from one group to another such that the sum of squares of the modified groups due to the movements is minimised (or reduced). Movements could involve allocation or re-allocation of objects to the group whose centroid is closest to them. Once all of the objects have been allocated, the respective group centroids are updated. Relocation of objects and revision of group centroids keep continuing until a minimum sum of squares has been reached. The relocation and the revision will terminate if none of the movements being considered will further reduce the sum of squares. This method differs with respect to the initial partitions, the objective function, the reassignment processes and the terminating criteria.



The initial group partition could be randomly selected, systematically selected or be the partition resulting from other technique. Examples of prespecified partition might include (Jensen, 1977)

1. random coordinates within the range of the data in the multidimensional space,
2. random data points or objects selected from the actual aggregate of data,
3. equally spaced coordinates along either the positive or negative transversal or diagonal of the multidimensional space,
4. symmetrical coordinates as mutually exclusive over the range of the data as is possible,
5. actual data points (objects) as mutually exclusive as is possible selected from a practical sample of the data, or
6. input user determined coordinates, which may or may not correspond to actual data points (objects).

Objective functions commonly used as clustering criteria include: (1) minimisation of trace ( $\mathbf{W}$ ), (2) minimisation of the determinant of ( $\mathbf{W}$ ), and (3) maximisation of trace ( $\mathbf{B}\mathbf{W}^{-1}$ ), where  $\mathbf{W}$  is the pooled within-cluster covariance matrix and  $\mathbf{B}$  is the between cluster covariance matrix.

Two popular algorithms of this clustering technique are the  $k$ -means and hill climbing techniques. However, the  $k$ -means technique performs well in comparison to the hill-climbing technique, although it is sensitive to its initial partition. The  $k$ -means technique is also less affected by outliers, the choice of distance and the presence of irrelevant attributes or dimensions. As such, the  $k$ -means technique was

adopted for seabed classification in this study (sections 5.3 and 5.4). In addition, initial partitions were based on information provided by ground truth (photographs and reference sites). To minimise sensitivity of the  $k$ -means technique to the initial partitions, the procedure used in this study (section 5.3) is as follows:

1. A training set is established by extracting PC1\_E1 and PC1\_E2 from ground truth area.
2. Centroids of PC1\_E1 and PC1\_E2 for each seabed type are calculated and then become initial centroids.
3. PC1\_E1 and PC1\_E2 are applied to the  $k$ -means technique.
4. Final centroids are calculated when (3) has been terminated.
5. Final centroids are compared to the initial centroids, if they are different, step 3 is repeated by taking final centroids as the initial centroids. If they are not different, the procedure ends and the final centroids become initial centroids to cluster all the remaining data.

## Appendix B

### Variogram and Kriging techniques

Kriging is a spatial interpolation technique that is optimal under fairly general conditions. The rationale underlying the construction of the variogram is as follows. Suppose  $Y$  is a regionalised variable of either PC1\_E1 or PC1\_E2 at a point  $x$ , then the mean of  $Y(x)$  is  $E[Y(x)] = m$ , which is constant for all points. The covariance at a lag  $h$ ,  $C(h)$ , between points  $x$  and  $x+h$  is then lag-dependent only. The second-order stationarity conditions are then (Simmonds *et al.*, 1992)

$$E[Y(x+h) - Y(x)] = 0 \quad (\text{B.1})$$

and

$$\text{VAR}[Y(x+h) - Y(x)] = 2\Gamma(h) \quad (\text{B.2})$$

where the function  $\Gamma(h)$  is the semi-variogram, which is very often loosely called the variogram. The variogram has also been used to refer to a plot of the variance of  $[Y(x+h) - Y(x)]$  against the lag distance  $h$ . Bearing in mind the above rationale, the semi-variogram is the mean square value of the difference between points,  $Y(x)$  and  $Y(x+h)$  as follows (Simmonds *et al.*, 1992):

$$\Gamma(h) = \frac{1}{2} E[Y(x+h) - Y(x)]^2 \quad (\text{B.3})$$

Suppose that the number of any possible pairs of data points is  $N(h)$ , then the robust estimator of the variogram proposed by Cressie (1991) is

$$2\Gamma(h) = \left( \frac{1}{|N(h)|} \sum_{i=1}^{N(h)} |Y(x+h) - Y(x)|^{1/2} \right)^4 / \left( 0.457 + \frac{0.494}{|N(h)|} \right) \quad (\text{B.4})$$

Ideally, it is expected to have  $\Gamma(h)=0$  at  $h=0$ . However, this is often not the case. This can be due to measurement error or discontinuity in the data on a very fine scale. A mathematical model may then be fitted to the calculated semi-variogram. Several common models are spherical, exponential, gaussian and power. If the plot does not have a structure, then the data set is considered random and the spatial distribution of the data set does not exist. At least 4 features can be extracted from the plot of the variogram. Firstly, from variograms derived from a number of different  $h$  orientations, a description of the anisotropy of the spatial distribution of the data can be obtained. Secondly, the plot of the variogram provides a maximum lag distance called the range in which the autocorrelation of the data exists. Together with the previous feature, if the ranges of the variograms from different  $h$  orientations are very different from one to another, the data set is considered anisotropic, otherwise it is isotropic. Thirdly, the sill (plateau), the maximum  $\Gamma(h)$ , offers the variance beyond the local autocorrelation. Typically, when the anisotropy happens, it is expected the sill of variograms from different  $h$  orientations remains the same or very close one to another even though the associated ranges might differ. Lastly, the nugget, the intercept on the y-axis where  $\Gamma(h) \neq 0$  at  $h=0$ , gives the sampling variance.

From the above features, the usefulness of the variogram as a descriptive tool is as follows. Several mathematical functions can be applied to the variogram if required and this helps describing different structures which exist in the area for

instance a small scale structure because of schools and a large scale structure because of the overall density distribution. The variogram is also very useful when appropriate dimensions of strata are desired. Lastly, if it happens that a different mathematical function is best descriptive in different parts of the area, this suggests that different probability distributions may exist.

The model that has best fitted the calculated semi-variogram can then be used for the interpolation (section 5.6) known as Kriging technique. The estimated Krige value  $Y^k(x_o)$  at an unsampled point  $x_o$  around sampled points  $Y(x_i)$  is (Simmonds *et al.*, 1992; Petitgas, 1996)

$$Y^k(x_o) = \sum_{i=1}^n w_i Y(x_i) \quad (B.5)$$

where  $w_i$  are weights and the sum of them is 1. The weights  $w_i$  are estimated in such a way that the Kriging variance is minimised. The Kriging variance is estimated as (Simmonds *et al.*, 1992):

$$\sigma_k^2 = E[Y_o - Y_o^k]^2 = E[Y_o^2] - 2E[Y_o Y_o^k] + E[(Y_o^k)^2] \quad (B.6)$$

To minimise the above Kriging variance taking into account that the weights  $w_i$  sum to 1, the Lagrange method is applied and the solution in matrix notation is

$$\mathbf{C} \bullet \mathbf{W}^k = \mathbf{U}$$

$$\begin{bmatrix} C_{11} & \dots & C_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C_{n1} & \dots & C_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \bullet \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} C_{10} \\ \vdots \\ C_{n0} \\ 1 \end{bmatrix} \quad (B.7)$$

Following matrix notation, the Kriging variance, equation (B.6), becomes

$$\sigma_k^2 = \sigma^2 - \mathbf{W}^k \cdot \mathbf{U} \quad (\text{B.8})$$

## Appendix C

### Listing of source code of the employed MATLAB functions

This appendix gives the source code of two MATLAB functions employed in this work. The first function *autocorrelation.m* given in C.1 was used to produce the autocorrelation function (ACF) using power spectral density (PDS) technique and the autocorrelation characteristic length. The second function *shift\_bottom.m* given in C.2 was prepared to correct the spike noise above the seabed due to a time jitter.

## C.1. *autocorrelation.m*

```
function [tlin,r,first0,R0]=autocorrelation(data);
% AUTOCORRELATION computes the autocorrelation using power spectral density
% technique.
%
% INPUTS: data containing equally spaced distance in the first column (real)
% and time series of indices at the second column (real).
% OUTPUTS: tlin = lag distance (real).
% r = autocorrelation (real).
% first0 = autocorrelation length derived from the first sign
% transition (real).
% R0 = autocorrelation length derived from Gaussian model (real).
% CHANGES: none.
%
% Created by Justy Siwabessy 21/02/00.
%

% extract equally spaced distances and indices from matrix data
tlin=data(:,1); % equally spaced distance
ztlin=data(:,2); % bottom indices (E1, E2, PC1_E1 or PC1_E2)

% remove mean and linear trend from the time series
ztlin=detrend(ztlin,mean(ztlin)); % remove mean
ztlin=detrend(ztlin); % remove linear trend

% bandpass mean-removed, linearly detrended time series
% using Butterworth filter design
% 5th-order bandpass Butterworth filter with a passband from
% 0.01% to 0.25% of the Nyquist frequency
[b,a]=butter(5,[0.01 0.25]);
ztlin=filtfilt(b,a,ztlin); % filter out unwanted components

% compute PSD
Zt=[ztlin; zeros(length(ztlin),1)]; % add zeros at the end of the time series
N=length(Zt);
Zt=fft(Zt); % compute DFT
Zt=Zt(1:length(Zt)/2)/N; % retain the first half of Zt only
delta=tlin(2)-tlin(1);
s=Zt.*conj(Zt)*2*N*delta;

% compute ACF by taking the inverse Cosine transform of Zt
r=idct(s); % compute inverse Cosine transform
r=r/max(r); % normalise r

% find the first 0
i=find(r<0);
first0=mean(tlin(i(1)-1:i(1))); % autocorrelation length from first transition

% fit Gaussian model
k=find(tlin<=first0);
lnc=log(r(k));
ltlin=tlin(k);
lnC=[ones(size(lnc)) ltlin.^2];
a=lnC\lnc;
lnCE=lnC*a;
CE=exp(lnCE);
R0=sqrt(-1/a(2)); % autocorrelation length from Gaussian model

% find the max correlation but excluding the correlation at lag 0
maxr=max(r(2:end));

% plot ACF derived from PSD.
plot([first0 first0],[min(r(2:end)) max(r(2:end))/2],'k','linewidth',1);
hold on;
plot(ltlin,CE,'k-.');
stairs(tlin(2:end)-(tlin(2))/2,r(2:end),'k');hold on
plot([min(tlin(2:end)) max(tlin(2:end))/8],[maxr maxr],'r-','linewidth',1);
text(max(tlin(2:end))/8,maxr,['\it' num2str(maxr)],'fontsize',9);
xlabel('Lag[nmi]','fontsize',12);
ylabel('\itR_{xx}(\tau)','fontsize',12);
legend(['First zero crossing. r_{o}=' num2str(first0)],['Gaussian model. r_{o}='
num2str(R0)],'ACF');
hold off
```



## C.2. *shift\_bottom.m*

```
function [X,Y,Z]=shift_bottom(X,Y,Z,ll,ul);
%SHIFT_BOTTOM finds the corrupted, shifted pings and relocate these pings
%                to the their "true" location estimated from good pings.
%INPUTS:  Z is the corrupted MxN matrix of Sv values.
%         ll is the allowable lower limit of the maximum Sv values within each ping to
%         occur.
%         ul is the allowable upper limit of the maximum Sv values within each ping to
%         occur.
%OUTPUTS: Z is the corrected MxN matrix of Sv values.
%CHANGES: Z is the input MxN matrix data as well as the output MxN matrix data.
%
%SYNTAX:  [output_matrix]=shift_bottom(input_matrix,min_limit,max_limit).
%
flg1=0;flg2=0;
while (~isempty(flgl) & ~isempty(flg2))
    [imaxz,imaxz]=max(Z);
    iimaxz=find(imaxz<ll);
    flg1=iimaxz;
    % If you like to use the first +ve value of the bottom returns
    % as a reference for shifting, include the below procedures
    % as far as the 'Stop here' sign.
    % [i,j]=find(Z(68:imaxz,iimaxz)>-45 & Z(68:imaxz,iimaxz)<-20);
    [i,j]=find(Z(70:max(imaxz(iimaxz)),iimaxz)>-55 & ...
        Z(70:max(imaxz(iimaxz)),iimaxz)<-10);
    unj=unique(j);
    for ii=1:length(unj)
        dummy=i(j==unj(ii));
        imaxz(iimaxz(ii))=dummy(1);
        iimaxz(iimaxz(ii));
    end
    nsll=abs(113-(imaxz(iimaxz)+69)); %must plus 69 to get back
    iiimaxz=find(imaxz>ul);
    iimaxz=iiimaxz(maxz(iimaxz)>0);
    flg2=iiimaxz;
    % If you like to use the first +ve value of the bottom returns
    % as a reference for shifting, include the below procedures
    % as far as the 'Stop here' sign.
    % [i,j]=find(Z(120:imaxz,iiimaxz)>-45 & Z(120:imaxz,iiimaxz)<-20);
    % [iiimaxz' imaxz(iiimaxz)' maxz(iiimaxz)']
    [i,j]=find(Z(120:max(imaxz(iiimaxz)),iiimaxz)>-45 & ...
        Z(120:max(imaxz(iiimaxz)),iiimaxz)<-10);
    unj=unique(j);
    for ii=1:length(unj)
        dummy=i(j==unj(ii));
        imaxz(iiimaxz(ii))=dummy(1);
        iimaxz(iiimaxz(ii));
    end
    nsul=abs(113-(imaxz(iiimaxz)+119)); %must plus 119 to get back
    % [iiimaxz' imaxz(iiimaxz)' nsul' maxz(iiimaxz)']
    % iiimaxz=iiimaxz(nsul<40);
    % nsul=nsul(nsul<40);
    % [iiimaxz' imaxz(iiimaxz)' nsul' maxz(iiimaxz)']
    [m,n]=size(Z);
    if length(nsll)>=length(nsul)
        n_loop=length(nsll);
    else
        n_loop=length(nsul);
    end
    for i=1:n_loop
        if i<=length(nsll)
            Z(:,imaxz(i))=[zeros(nsll(i),1)+NaN; ...
                Z(1:m-nsll(i),imaxz(i))];
        end
        if i<=length(nsul)
            Z(:,iiimaxz(i))=[Z(nsul(i)+1:m,iiimaxz(i)); ...
                zeros(nsul(i),1)+NaN];
        end
    end
    end
    %Z(isnan(Z))=min(min(Z)); %For printing purposes only. Canon BJC 4500 gets trouble
    %when printing "NaN".
end
```

## Appendix D

### Shell and gawk scripts

The "shell" script *extract\_indices* (D.1) extracts parameters of interest from the *ECHO* output files. It calls two "gawk" scripts, *extract\_indices\_awk* (D.2) and *average\_indices\_awk* (D.3). The first gawk script extracts all parameters of interest and, if required, rounds off the time to the nearest minute depending on whether or not the *r* flag is specified. If the *r* flag has been specified, the second gawk script will be launched to calculate the average of parameters of interest within each one-minute interval. In principle, all information from different indices can be stored in a single, big file. If that is the case, it is important to specify a unique ID for each index. This script requires 1 input argument and or without an output argument and the syntax is:

```
extract_indices <list_of_input> [output_file]
```

where [output\_file] is the name of the output file to which results are directed and <list\_of\_input> contains 2 fields separated by WHITE SPACE. The first field contains numbers and a flag separated by COMMA and the second field contains the filename and path. e.g.:

```
400,r,2 Aug_11/38kHz_all/e1_loangle\(\AL,27,40\)
```

400 is the low-res echogram range, *r* is the flag to have the time rounded to the nearest minute, 2 is the assigned ID, Aug\_11/38kHz\_all/ is the path and e1\_loangle\(\AL,27,40\ is the input file.

## D.1. *extract\_indices*

```
#!/bin/sh
# extract_indices
#
# It extracts respectively time, date, the name of survey and worksheet, vessel log,
# GPS log, latitude, longitude, SA* and Std Sa*, assigns numbers provided by the
# database and if required rounds the time to the nearest minute. The later is
# required when latitudes and longitudes are missing and it needs to patch up the lats
# and lons from underway gps data which have a one-minute resolution. As the time is
# rounding to the nearest minute, vessel log, depth, SA* and Std Sa* are averaging
# accordingly.
#
# Created by Justy Siwabessy 21/02/00.
#

#HOME=/mnt/NWS_Study/script
HOME=/data/NWS_Study/script
ECHO=/bin/echo
GAWK=/bin/gawk
CAT=/bin/cat
MORE=/bin/more
SORT=/usr/bin/sort
READ=read
MKDIR=/bin/mkdir
EGREP=/bin/egrep
RM=/bin/rm
SED=/bin/sed
DUMMY_FILE=$HOME/$$.dummy
EXTRACT=$HOME/extract_indices_awk
AVERAGE=$HOME/average_indices_awk

if [ $# -lt 1 ]; then
    $ECHO
    $ECHO "Usage: $0 <list_of_input> [output_file]"
    $ECHO
    $ECHO "<list_of_input> contains 2 fields separated by BLANK SPACE."
    $ECHO "The 1st field contains numbers and a flag separated by COMMA."
    $ECHO "The 2nd field contains the filename and path."
    $ECHO "e.g.:"
    $ECHO "400,r,2 Aug_11/38kHz_all/e1_loangle\ (AL,27,40\)"
    $ECHO "400 is the low-res echogram range"
    $ECHO "r is the flag to have the time rounded to the nearest minute"
    $ECHO "2 is the background_id produced by MS Access Database"
    $ECHO "Aug_11/38kHz_all/ is path."
    $ECHO "e1_loangle\ (AL,27,40\)" is the input file."
    $ECHO
    $ECHO "The 1st field options:"
    $ECHO "r perform time rounding to the nearest"
    $ECHO " minute but those near 00:00 are rounded"
    $ECHO " down to 23:59."
    $ECHO "v1 limit extraction based on echogram"
    $ECHO " range v1."
    $ECHO "v1,v2 limit extraction based on v1 and write"
    $ECHO " background_id v2 provided by the MS"
    $ECHO " Access Database. Note: v2 must always"
    $ECHO " come after v1."
    $ECHO "v1,r or r,v1 limit extraction based on v1 and perform"
    $ECHO " time rounding."
    $ECHO "r,v1,v2 or v1,r,v2 or v1,v2,r limit the extraction, write the"
    $ECHO " background_id and perform time rounding."
    $ECHO " Note: v2 must always come after v1"
    $ECHO " wherever the location of r is."
    $ECHO "Where v1 is echogram range either hires or lowres."
    $ECHO " v2 is background_id provided by the MS Access database."
    $ECHO " r is the flag of time rounding."
    $ECHO "Note: v1 and v2 are always numeric."
    $ECHO "Default: no rounding, no limit and no background_id."
    $ECHO
    exit 1
fi
```

fi

INPLIST=\$1  
OUTFILE=\$2

```
if [ $# -eq 1 ]; then
  $SED -n '/./p' $INPLIST |
  (while true; do
    $READ LINE || exit
    FLAG="-v inval="`$ECHO $LINE | $GAWK '{print $1}'`
    INPFILE=`$ECHO $LINE | $GAWK '{print $2}'`
    INPFILE=`pwd`"/"$INPFILE
    $ECHO $FLAG | $EGREP 'r' > /dev/null 2>&1
    if [ $? -eq 0 ]; then
      $EXTRACT $FLAG $INPFILE > $DUMMY_FILE
      $AVERAGE $DUMMY_FILE
      $RM $DUMMY_FILE
    else
      $EXTRACT $FLAG $INPFILE
    fi
  done)
elif [ $# -eq 2 ]; then
  CTR=0
  $SED -n '/./p' $INPLIST |
  (while true; do
    $READ LINE || exit
    FLAG="-v inval="`$ECHO $LINE | $GAWK '{print $1}'`
    INPFILE=`$ECHO $LINE | $GAWK '{print $2}'`
    INPFILE=`pwd`"/"$INPFILE
    $ECHO $FLAG | $EGREP 'r' > /dev/null 2>&1
    if [ $? -eq 0 ]; then
      $EXTRACT $FLAG $INPFILE > $DUMMY_FILE
      if [ $CTR -eq 0 ]; then
        $AVERAGE $DUMMY_FILE > $OUTFILE
      else
        $AVERAGE $DUMMY_FILE >> $OUTFILE
      fi
      $RM $DUMMY_FILE
    else
      if [ $CTR -eq 0 ]; then
        $EXTRACT $FLAG $INPFILE > $OUTFILE
      else
        $EXTRACT $FLAG $INPFILE >> $OUTFILE
      fi
    fi
    let CTR=$CTR+1
  done)
fi
```

## D.2. *extract\_indices\_awk*

```
#!/bin/gawk -f
# extract_indices_awk
#
# This gawk script extracts and rounds the time to the nearest minute and
# The averaging process by using average_indices_awk is then conducted
# immediately after extracting and rounding process and then stores the final
# result to a file.
#
# Removes much of the information that is unnecessary for analysis
# Extracts information of interest and rounds off time stamp, if flag is on, for
# later use in matching up with underway GPS data for lats and lons.
#
# Created by Justy Siwabessy 02/02/00.
#

function iprint(value1, value2)
{
  printf("%lf%s", value1, value2);
}
```

```

}

function sprintf(value1, value2)
{
    if (value1 == "")
        printf("NaN%s", value2);
    else
        printf("%s%s", value1, value2);
}

BEGIN \
{
phi=3.141592654;
rounding_flag=0;
i_rounding_flag=0;
for (i=1; i<11; ++i)
{
    getline l;
    if (i==1)
    {
        k=split(l, l_data, ",");
        cruise=l_data[1];
        survey=l_data[2];
        worksheet=l_data[3];
    }
    if (i==7)
    {
        k=split(l, l_data, " ");
        l=l_data[1];
        k=split(l, l_data, "(");
        l=l_data[2];
        k=split(l, l_data, ")");
        l=l_data[1];
        k=split(l, mask_info, ",");
        reference=mask_info[1];
        start_depth_angle=mask_info[2];
        stop_depth_angle=mask_info[3];
    }
}
if (insval=="")
{
    background_id=" ";
    reference="default";
    rounding_flag=0;
}
else
{
    k=split(insval, insval_i, ",");
    for (i=1; i<k+1; ++i)
    {
        if (insval_i[i]=="r" || insval_i[i]=="R")
        {
            rounding_flag=1;
            i_rounding_flag=i;
        }
    }
    if (k==1)
    {
        if (rounding_flag && i_rounding_flag)
        {
            background_id=" ";
            insval=10000;
        }
        else
        {
            background_id=" ";
        }
    }
    if (k==2)
    {
        if (rounding_flag && i_rounding_flag)
        {
            if (i_rounding_flag==1)
            {
                background_id=" ";
            }
        }
    }
}

```

```

        inval=inval_i[2];
    }
    else
    {
        background_id=" ";
        inval=inval_i[1];
    }
}
else
{
    background_id=inval_i[2];
    inval=inval_i[1];
}
}
if (k==3)
{
    if (i_rounding_flag==1)
    {
        background_id=inval_i[3];
        inval=inval_i[2];
    }
    if (i_rounding_flag==2)
    {
        background_id=inval_i[3];
        inval=inval_i[1];
    }
    if (i_rounding_flag==3)
    {
        background_id=inval_i[2];
        inval=inval_i[1];
    }
    if (i_rounding_flag!=1 && i_rounding_flag!=2 && i_rounding_flag!=3)
    {
        background_id=" ";
        reference="default";
        rounding_flag=0;
    }
}
if (k>3)
{
    if (rounding_flag)
    {
        background_id=" ";
        reference="default";
    }
    else
    {
        background_id=" ";
        reference="default";
        rounding_flag=0;
    }
}
}
if (reference=="BL" || reference=="BH" || reference=="B2")
{
    min_depth=-stop_depth_angle;
    if (reference=="B2")
    {
        max_depth=(inval-min_depth)/2;
    }
    else if (reference=="BL")
    {
        max_depth=inval-min_depth;
    }
    else
    {
        max_depth=10000;
    }
    max_depth_hires=max_depth;
    cos_angle=1;
}
else if (reference=="AL")
{
    min_depth=5;
    max_depth=inval*cos(stop_depth_angle/180*phi);
    max_depth_hires=max_depth;
}

```

```

        cos_angle=1;
    }
else if (reference=="AH")
    {
    min_depth=5;
    max_depth=10000;
    max_depth_hires=insval-10; # remember to subtract 10m above seabed.
    cos_angle=cos(stop_depth_angle/180*phi);
    }
else
    {
    min_depth=5;
    max_depth=10000;
    max_depth_hires=max_depth;
    cos_angle=1;
    }
while (getline cl)
    {
    m=split(cl, cl_data, ",");
    n=split(cl_data[3], cl_time_data, ":");
    o=split(cl_data[2], cl_date_data, "/");
    if (cl_data[9]>min_depth && cl_data[9]<max_depth && (cl_data[9]/cos_angle-
cl_data[9])<max_depth_hires && cl_data[12]>0 && cl_data[17]>0)
        {
        if (rounding_flag)
            {
            if (cl_time_data[3] > 30)
                {
                ++cl_time_data[2];
                if (cl_time_data[2] > 59)
                    {
                    ++cl_time_data[1];
                    cl_time_data[2] = cl_time_data[2] - 60;
                    if (cl_time_data[1] > 23)
                        {
                        # This below formation is of use when extraction and averaging are
                        # conducted one after another for each day data separately. The other
                        # formation is to comment #2 and #3 and to uncomment #1 and #4 when
                        # extraction is first conducted for all day data and then average is
                        # conducted later at once for all day data in a big, single block of
                        # data (the whole cruise).
                        #
                        cl_time_data[1] = 0;
                        cl_time_data[1] = 23;
                        cl_time_data[2] = 59;
                        #
                        #
                        cl_date_data[1] = cl_date_data[1] + 1;
                        date_flag = 1;
                        }
                    }
                }
            }

        }
    sprintf(background_id, ",");
    sprintf(survey, ",");
    sprintf(cl_data[1], ",");
    if (date_flag==1)
        {
        if (length(cl_date_data[1]) < 2)
            {
            sprintf(" 0", "");
            sprintf(cl_date_data[1], "/");
            }
        else
            {
            sprintf(" ", "");
            sprintf(cl_date_data[1], "/");
            }
        if (length(cl_date_data[2]) < 2)
            {
            sprintf("0", "");
            sprintf(cl_date_data[2], "/");
            }
        else
            {
            sprintf(cl_date_data[2], "/");
            }
        if (length(cl_date_data[3]) < 2)
            {
            sprintf("0", "");
            }
        }
    }
}

```

```

        sprint(cl_date_data[3], "");
    }
    else
    {
        sprint(cl_date_data[3], "");
    }
    date_flag==0;
}
else
{
    sprint(cl_data[2], "");
}
if (length(cl_time_data[1]) < 2)
{
    sprint(" 0", "");
    sprint(cl_time_data[1], ":");
}
else if (length(cl_time_data[1]) == 2)
{
    sprint(" ", "");
    sprint(cl_time_data[1], ":");
}
else
{
    #sprint(" ", "");
    sprint(cl_time_data[1], ":");
}
if (length(cl_time_data[2]) < 2)
{
    sprint("0", "");
    sprint(cl_time_data[2], ",");
}
else
{
    sprint(cl_time_data[2], ",");
}
}
else
{
    sprint(background_id, ",");
    sprint(cl_data[2], "");
    sprint(cl_data[3], ",");
    sprint(survey, ",");
    sprint(cl_data[1], ",");
}
for (j=4; j<8; ++j)
{
    sprint(cl_data[j], ",");
}
sprint(cl_data[9], ",");
sprint(cl_data[12], ",");
sprint(cl_data[13], "\n");
} # end of the very first if
} # end of while
} # end line (end of begin)

```

### D.3. *average\_indices\_awk*

```

#!/bin/gawk -f
# average_indices_awk
#
# Averages values of interest, removes much of the information that is unnecessary
# for analysis and extracts after rounding process by extract_indices.awk.
#
# Created by Justy Siwabessy 02/02/00
#
function iprint(value1, value2)
{
    printf("%lf%s", value1, value2);
}

```



```

function sprint(value1, value2)
{
    if (value1 == "")
        printf("NaN%s", value2);
    else
        printf("%s%s", value1, value2);
}

BEGIN \
{
    n = 0;
    while (getline cl)
    {
        m=split(cl, cl_data, ",");
        if (counter == 0)
        {
            background_id = cl_data[1];
            time_flag = cl_data[4];
            survey_flag = cl_data[2];
            worksheet_flag = cl_data[3];
        }
        if (cl_data[4] == time_flag)
        {
            acc_vessel_log = acc_vessel_log + cl_data[5];
            acc_depth = acc_depth + cl_data[9];
            acc_sa = acc_sa + cl_data[10];
            acc_std = acc_std + cl_data[11];
            if (cl_data[m] != 0 || cl_data[m] != "NaN")
            {
                ++n;
            }
            background_id = cl_data[1];
            time_flag = cl_data[4];
            survey_flag = cl_data[2];
            worksheet_flag = cl_data[3];
        }
        if (cl_data[4] != time_flag)
        {
            if (n != 0)
            {
                ave_vessel_log = acc_vessel_log / n;
                ave_depth = acc_depth / n;
                ave_sa = acc_sa / n;
                ave_std = acc_std / n;
                if (ave_sa !=0)
                {
                    sprint(background_id, ",");
                    sprint(time_flag, ",");
                    sprint(survey_flag, ",");
                    sprint(worksheet_flag, ", ");
                    printf("%7.3f", ave_vessel_log);
                    sprint(" ", ",");
                    sprint(" ", ",");
                    sprint(" ", ",");
                    printf("%6.1f", ave_depth);
                    printf("%1.6e", ave_sa); #floating point xx.xxExx
#                    printf("%11.3f", ave_sa); # floating point xxx.xxx
#                    printf("%1.6e\n", ave_std); #floating point xx.xxExx
#                    printf("%11.3f\n", ave_std); # floating point xxx.xxx
                }
                acc_vessel_log = cl_data[5];
                acc_depth = cl_data[9];
                acc_sa = cl_data[10];
                acc_std = cl_data[11];
                n = 1;
            }
            background_id = cl_data[1];
            time_flag = cl_data[4];
            survey_flag = cl_data[2];
            worksheet_flag = cl_data[3];
        }
        counter += 1
    }
} # end line

```

## Appendix E

### List of publications

Siwabessy, P.J.W., J.D. Penrose, R.J. Kloser and D.R. Fox. 1999. "Seabed habitat classification." *Proc. International Conference on High Resolution Surveys in Shallow Waters* DSTO, 18-20 October 1999, Sydney, Australia.

Siwabessy, P.J.W., J.D. Penrose and R.J. Kloser. 2000. "A comparison between acoustically derived  $s_A$  estimates and those derived from nets." *ECUA 2000, Proc. 5<sup>th</sup> European Conference on Underwater Acoustics*, 10-13 July 2000, Lyon, France. M. Zakharia, P. Chevret and P. Dubail (eds), 1431-1436.

Siwabessy, P.J.W., J.D. Penrose, D.R. Fox and R.J. Kloser. 2000. "Bottom Classification in the Continental Shelf: A Case Study for the North-west and South-east Shelf of Australia." *Acoustic 2000, Proc. Aust. Acous. Soc.*, 15-17 November 2000, Joondalup, Perth, Western Australia, 265-270.

Kloser, R.J., P.J.W. Siwabessy and M. Lewis. 2000. "*In situ* target strength measurements of black oreo (*Allocyttus niger*)." CSIRO Rep., 33pp.

## References

- Aglen, A. 1996. "Impact of fish distribution and species composition on the relationship between acoustic and swept-area estimates of fish density." *ICES J. Mar. Sci.* 53, 501-505.
- Althaus, F., X. He, K. Woolley, S. Condie and K.J. Sainsbury. in prep. "Spatial distribution of benthic habitats and their response to management implementation." CSIRO Marine Research, Hobart, Australia.
- Anderberg, M.R. 1973. *Cluster Analysis for Applications*. Academic Press. New York. 359pp.
- Anon. 1993. *Simrad EK500 Scientific Echo Sounder Instruction Manual*. Simrad Subsea P2170E.
- Anon. 1995. "Echo sounder seabed classification results." *Quester Tangent Corporation*. Document SR-25-1-April 1995. Sidney, BC, Canada.
- Arnaya, I.N., N. Sano and K. Iida. 1989. "Studies on acoustic target strength of squid III: measurement of the mean target strength of small live squid." *Bull. Fac. Fish. Hokkaido Univ.* 40(2), 100-115.
- Bailey, T.C. and A.C. Gatrell. 1995. *Interactive Spatial Data Analysis*. Longman Scientific & Technical. New York. 413pp.
- Bartlett, W.M. 1981. *Western Australian Year Book*. No. 19. Australian Bureau of Statistics, Western Australia Office. 586pp.
- Bax, N.J. and A. Williams (Eds). 2000. "Habitat and fisheries productivity in the South East Fishery." Final Report to FRDC Project 94/040. (CSIRO Marine Research: Hobart, Australia).
- Bax, N.J., R.J. Kloser, A. Williams, K. Gowlett-Holmes and T. Ryan. 1999. "Seafloor habitat definition for spatial management in fisheries: a case study on the continental shelf of southeast Australia using acoustic and biotic assemblages." *Oceanologica Acta* 22(6), 705-719.

- Benedito-Cecilio, E., A.A. Agostinho and R.C.C. Velho. 1997. "Length-weight relationship of fishes caught in the Itaipu reservoir, Parana, Brazil." *Naga, ICLARM Q.* 20(3-4), 57-61.
- Blaber, S.J.M., D.T. Brewer and A.N. Harris. 1994. "Distribution, biomass and community structure of demersal fishes of the Gulf of Carpentaria, Australia." *Aust. J. Mar. Freshwater Res.* 45, 375-396.
- Brekhovskikh, L.M. and Yu.P. Lysanov. 1982, *Fundamentals of Ocean Acoustics* 2<sup>nd</sup> Edition. Springer Series on Wave Phenomena. Springer-Verlag, Berlin Heidelberg, Fed. Rep. of Germany, 270pp.
- Brook, D. and R.J. Wynne. 1988. *Signal Processing. Principles and Applications.* Edward Arnold. London. 308pp.
- Burns, D.R., C.B. Queen, H. Sisk, W. Mullarkey and R.C. Chivers. 1989. "Rapid and convenient acoustic seabed discrimination." *Proceedings of the Institute of Acoustics* 11, 169-178.
- Caton, A., K. McLoughlin and D. Staples. 1997. "Fishery status reports 1997." *Bureau Resour. Sci.*, 139pp.
- Chivers, R.C., N. Emerson and D.R. Burns. 1990. "New acoustic processing for underway surveying." *Hydro. J.* 56, 9-17.
- Clarke, A. and J.A. Crame. 1997. "Diversity, latitude and time: Patterns and shallow seas." *Marine Biodiversity: Patterns and Processes.* R.F.G. Ormond, J.D. Gage and M.V. Angel (eds). Cambridge University Press, New York. 449pp.
- Clarke, K.R. and R.N. Gorley. 2001. *PRIMER v5: User Manual/Tutorial.* PRIMER-E Ltd. Plymouth, UK. 91pp.
- Clay, C.S. 1966. "Coherent reflection of sound from the ocean bottom." *J. Geophys. Res.* 71(8), 2037-2046.
- Clay, C.S. and H. Medwin. 1977. *Acoustical Oceanography. Principles and Applications.* Ocean Engineering: A Wiley Series. John Wiley and Sons, Inc. New York. 544pp.
- Collins, W., R. Gregory and J. Anderson. 1996. "A digital approach to seabed classification." *Sea Tech.* August, 83-87.
- Collins, W.T. and R.A. McConnaughey. 1998. "Acoustic classification of the sea floor to address essential fish habitat and marine protected area requirements." *Proc. Canadian Hydrographic Conference*, March 1998.

- Cressie, N.A.C. 1991. *Statistics for Spatial Data*. Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, Inc. New York. 900pp.
- Diday, E. and J.C. Simon. 1976. "Clustering analysis." *Digital Pattern Recognition*. K.S. Fu, W.D. Keidel and H. Wolter (eds). Springer-Verlag. Berlin. 47-94.
- Edwards, J.I., F. Armstrong, A.E. Magurran and T.J. Pitcher. 1984. "Herring, mackerel and sprat target strength experiments with behavioural observations." *Int. Coun. Explor. Sea CM* 1984/B:34, 21pp.
- Edwards, R.J. 1990. "Upwelling could hold clues to fish patterns." *Aust. Fish.* 49(1), 18-20.
- Elliot, N.G. and R.J. Kloser. 1994. "Comparison of acoustic and trawl results of the 1989 CSIRO deepwater survey in southeastern Australia waters." CSIRO Report 216, 19pp.
- Everitt, B. 1974. *Cluster Analysis*. Heinemann Educational Books Ltd. London. 122pp.
- Everson, I. and D.G.M. Bone. 1986. "Effectiveness of the RMT8 system for sampling krill (*Euphausia superba*) swarms." *Polar Biology* 6, 83-91.
- Everson, I. 1987. "Some aspects of the small scale distribution of *Euphausia crystallorophias*." *Polar Biology* 8, 9-15.
- Foote, K.G. 1982. "Optimizing copper spheres for precision calibration of hydroacoustic equipment." *J. Acoust. Soc. Am.* 71(3), 742-747.
- Foote, K.G. 1983. "Maintaining precision calibrations with optimal copper spheres." *J. Acoust. Soc. Am.* 73(3), 1054-1063.
- Foote, K.G. 1987. "Fish target strengths for use in echo integrator surveys." *J. Acoust. Soc. Am.* 82(3), 981-987.
- Foote, K.G. 1996. "Quantitative fisheries research surveys, with special reference to computers." *Computer in Fisheries Research*. B.A. Megrey and E. Moksness (eds). Chapman and Hall. London. 254pp.
- Fréo, P., F. Gerlotto and M. Soria. 1990. "Evaluation of the influence of vessel noise on fish distribution as observed using alternatively motor and sails aboard a survey vessel." ICES statutory meeting, Bergen, Oct. 1990, CM/B, 55, 15pp.
- Froese, R and D. Pauly. 1998, *FishBase 98. Concepts, Design and Data Sources*. ICLARM. Manila. 293pp+CD-ROM.

- Garcia-Arteaga, J.P., R. Claro and S. Valle. 1997. "Length-weight relationships of Cuban marine fishes." *Naga, ICLARM Q.* 20(1), 38-43.
- Gerlotto, F. 1993. "Identification and spatial stratification of tropical fish concentrations using acoustic populations." *Aquat. Living Resour.* 6, 243-254.
- Gordon, A.D. 1981. *Classification*. Chapman and Hall. London. 193pp.
- Greenstreet, S.P.R., I.D. Tuck, G.N. Grewar, E. Armstrong, D.G. Reid and P.J. Wright. 1997. "An assessment of the acoustic survey technique, RoxAnn, as a means of mapping seabed habitat." *ICES J. Mar. Sci.* 54, 939-959.
- Hamilton, E.L. 1972. "Compressional-wave attenuation in marine sediments." *Geophysics* 37(4), 620-646.
- Hamilton, E.L., H.P. Bucker, D.L. Keir and J.A. Whitney. 1970. "Velocities of compressional and shear waves in marine sediments determined *in-situ* from a research submersible." *J. Geophys. Res.* 75(20), 4035-4049.
- Hamilton, L.J., P.J. Mulhearn and R. Poeckert. 1999. "Comparison of RoxAnn and QTC-View acoustic bottom classification system performance for the Cairns area, Great Barrier Reef, Australia." *Cont. Shelf. Res.* 19, 1577-1597.
- Harris, R.J. 1975. *A Primer of Multivariate Statistics*. Academic Press. New York. 332pp.
- Heald, G.J. and N.G. Pace. 1996. "An analysis of the 1st and 2nd backscatter for seabed classification." *Proc. 3<sup>rd</sup> European Conference on Underwater Acoustics*, 24-28 June 1996 vol. II, 649-654.
- Holloway, P.E. 1983. "Tides on the Australian North-west shelf." *Aust. J. Mar. Freshw. Res.* 34, 213-230.
- Holloway, P.E. 1987. "Internal hydraulic jumps and solitons at a shelf break region on the Australian North West Shelf." *J. Geophys. Res.* 92(C5), 5405-5416.
- Holloway, P.E. and H.C. Nye. 1985. "Leeuwin current and wind distributions on the southern part of the Australian North West Shelf between January 1982 and July 1983." *Aust. J. Mar. Freshw. Res.* 36, 123-137.
- ICNAF. 1963, *The selectivity of fishing gear. Spec. Publ. int. Comm. NW Atlant. Fish.* 5, 1-225.

- Jackson, D.R. and K.B. Briggs. 1992. "High-frequency bottom backscattering: Roughness versus sediment volume scattering." *J. Acoust. Soc. Am.* 92(2), 962-977.
- Jensen, D.L. 1977. *The Role of Cluster Analysis in Computer Assisted Mass Appraisal*. The Lincoln Institute of Land Policy. Cambridge, Massachusetts. 75pp.
- Jensen, F.B., W.A. Kuperman, M.B. Porter and H. Schmidt. 1994. *Computational Ocean Acoustics*. AIP Press. Woodbury, New York. 612pp.
- Jin, X. 1990. "Acoustic estimation of near seabed fish abundance." M. Phil. Thesis. University of Bergen, Norway, 81pp.
- Johnson, R.A. and D.W. Wichern. 1998. *Applied Multivariate Statistical Analysis* 4<sup>th</sup> Edition. Prentice-Hall International, Inc. New Jersey. 816pp.
- Jones, H.A. 1973. "Marine geology of the Northwest Australian continental shelf." *B.M.R. Bull.* No. 136, 102pp.
- Jones, H.A. and P.J. Davies. 1983. "Superficial sediments of the Tasmanian continental shelf and part of Bass Strait." *B.M.R. Bull.* No. 218, 25pp.
- Kaiser, M.J., P.J. Armstrong, P.J. Dare and R.P. Flatt. 1998. "Benthic communities associated with a heavily fished scallop ground in the English Channel." *J. Mar. Biol.* 78(4), 1045-1059.
- Kavli, T., E. Weyer and M. Carlin. 1994. "Real time seabed classification using multi frequency echo sounders." *Proc. Oceanology International 94 vol. 4*. Brighton, UK, March 1994, 1-9.
- Kloser, R.J. 1996. "Deeply-towed transducer improves precision of acoustic surveys of benthopelagic fish." *ICES J. Mar. Sci.* 53, 405-413.
- Kloser, R.J., J.A. Koslow and A. Williams. 1996. "Acoustic assessment of the biomass of a spawning aggregation of orange roughy (*Hoplostethus atlanticus*, Collet) off South-eastern Australia, 1990-93." *Mar. Freshwater Res.* 47, 1015-1024.
- Kloser, R.J., P.V. Sakov, J.R. Waring, T.E. Ryan and S.R. Gordon. 1998. "Development of software for use in multi-frequency acoustic biomass assessments and ecological studies." CSIRO Report to FRDC project T93/237, 74pp.
- Kloser, R.J., P.J.W. Siwabessy and M. Lewis. 2000. "In situ target strength measurements of black oreo (*Allocyttus niger*)." CSIRO Rep., 33pp.

- Kloser, R.J., N.J. Bax, T. Ryan, A. Williams and B.A. Baker. 2001. "Remote sensing of seabed types in the Australian South East Fishery – development and application of normal incident acoustic techniques and associated "ground truthing"." *Mar. Freshwater Res.* 52, 475-489.
- Kochzius, M. 1997. "Length-weight relationship of fishes from a seagrass meadow in Negros Oriental, Philippines." *Naga, ICLARM Q.* 20(3-4), 64-65.
- Koslow, J.A., R.J. Kloser and A. Williams. 1997. "Pelagic biomass and community structure over the mid-continental slope off southern Australia based upon acoustic and midwater trawl sampling." *Mar. Ecol. Prog. Ser.* 146, 21-35.
- Lurton, X. and E. Pouliquen. 1992. "Automated seabed classification system for echosounders." *Oceans 92* Vol. 1, New York, IEEE, 317-321.
- MacLennan, D.N. and E.J. Simmonds. 1992. *Fisheries Acoustics*. Fish and Fisheries Series. Chapman and Hall. London. 325pp.
- Magorrian, B.H., M. Service and W. Clarke. 1995. "An acoustic bottom classification survey of Strangford Lough, Northern Ireland." *J. Mar. Biol. Ass. UK.* 75, 987-992.
- McKinney, C.M. and C.D. Anderson. 1964. "Measurements of backscattering of sound from the ocean bottom." *J. Acoust. Soc. Am.* 36(1), 158-163.
- McLoughlin, R.J. and P.C. Young. 1985. "Sedimentary provinces of the fishing grounds of the North West Shelf of Australia: Grain-size frequency analysis of surficial sediments." *Aust. J. Mar. Freshw. Res.* 36, 671-681.
- Medwin, H. and C.S. Clay. 1998. *Fundamentals of Acoustical Oceanography*. Academic Press. Boston. 712pp.
- Merella, P., A. Quetglas, F. Alemany and A. Carbonell. 1997. "Length-weight relationship of fishes and cephalopods from the Balearic Islands (Western Mediterranean)." *Naga, ICLARM Q.* 20(3-4), 66-68.
- Ona, E. and O.R. Godo. 1990. "Fish reaction to trawling noise: the significance for trawl sampling." *Rapp. P.-v. Reun. Cons. int. Explor. Mer.* 189, 159-166.
- Ona, E. and R.B. Mitson. 1996. "Acoustic sampling and signal processing near the seabed: the deadzone revisited" *ICES J. Mar. Sci.* 53, 677-690.
- Orlowski, A. 1984. "Application of multiple echoes energy measurements for evaluation of sea bottom type." *Oceanologia* 19, 61-78.



- Pace, N.G. and Z.K.S. Al-Hamdani. 1985. "The range dependence of normal incidence acoustic backscatter from a rough surface." *J. Acoust. Soc. Am.* 77(1), 101-112.
- Pace, N.G. and R.V. Ceen. 1982. "Seabed classification using the backscattering of normally incident broadband acoustic pulses." *Hydro. J.* 26, 9-16.
- Pariwono, J.I., J.A.T. Bye and G.W. Lennon. 1986. "Long period variations of sea level in Australasia." *J. Roy. Astro. Soc.* 87, 43-54.
- Pauly, T., S. Nicol, W.K. de la Mare, I. Higginbottom and G. Hosie. 1997. "A comparison between the estimated density of krill from an acoustic survey with that obtained by scientific nets on the same survey." WG-EMM-97/43, 15pp.
- Pauly, T. and J.D. Penrose. 1998. "Laboratory target strength measurements of free swimming Antarctic krill (*Euphausia superba*)." *J. Acoust. Soc. Am.* 103(6), 3268-3280.
- Penrose, J.D. and G.T. Kaye. 1979. "Acoustic target strengths of marine organisms." *J. Acoust. Soc. Am.* 65(2), 374-380.
- Petitgas, P. 1996. "Geostatistics and their applications to fisheries survey data." *Computer in Fisheries Research*. B.A. Megrey and E. Moksness (eds). Chapman and Hall. London. 254pp.
- Prager, B.T., D.A. Caughey and R.H. Poeckert. 1995. "Bottom classification: operational results from QTC view." *OCEANS '95 - Challenges of our changing global environment conference*, October 1995, San Diego, California, USA.
- Rencher, A.C. 1995. *Methods of Multivariate Analysis*. John Wiley and Sons, Inc. New York. 627pp.
- Richardson, M.D. and D.K. Young. 1980. "Geoacoustical models and bioturbation." *Mar. Geo.* 38, 205-218.
- Ruiz-Ramirez, S., G. Lucano-Ramirez and J. Mariscal-Romero. 1997. "Length-weight relationships of soft-bottom demersal fishes from Jalisco and Colima States, Mexico." *Naga, ICLARM Q.* 20(3-4), 62-63.
- Rose, G.A. 1992. "A review of problems and new directions in the application of fisheries acoustics on the Canadian East Coast." *Coast. Fish. Res.* 14, 105-128.

- Rosenzweig, M.L. 1995. *Species Diversity in Space and Time*. Cambridge University Press, New York. 436pp.
- Ryan, T., R.J. Kloser and P. Sakov. 1997. "Data management, analysis and mapping of acoustic seafloor indices." CSIRO Unpublished Internal Report. 38pp.
- Sainsbury, K.J., P.J. Kailola and G.G. Leyland. 1985. *Continental Shelf Fishes of Northern and North-western Australia*. Clouston & Hall, Canberra, Australia, 375pp.
- Sainsbury, K.J. 1991. "Application of an experimental approach to management of a tropical multispecies fishery with highly uncertain dynamics." *ICES Marine Science Symposium* 193, 301-320.
- Sainsbury, K.J., R.A. Campbell and A.W. Whitelaw. 1993. "Effects of trawling on the marine habitat on the North West Shelf of Australia and implications for sustainable fisheries management." *Bureau of Resources Sciences Proceedings*, 137-145.
- Sainsbury, K.J., R.A. Campbell, R. Lindholm and A.W. Whitelaw. 1997. "Experimental management of an Australian multispecies fishery: Examining the possibility of trawl-induced habitat modification." *Global Trends: Fisheries Management*. E.K. Pikitch, D.D. Huppert and M.P. Sissenwine, editors. American Fisheries Society. Bethesda. Maryland, 107-112.
- Sainsbury, K.J., X. He, F. Althaus, C. Stanley, R. Campbell and K. Woolley. in prep.. "Benthic habitat types and their associations on the Northwest Shelf of Australia." CSIRO Marine Research, Hobart, Australia.
- Scalabrin, C. and J. Massé. 1993. "Acoustic detection of the spatial and temporal distribution of fish shoals in the Bay of Biscay." *Aquat. Living Resour.* 6, 269-283.
- Schlagintweit, G.E.O. 1993. "Real-time acoustic bottom classification: a field evaluation of RoxAnn." *Oceans* 93, pIII-214-pIII-219.
- Scott, J.S. 1982. "Selection of bottom type by groundfishes of the Scotian Shelf." *Can. J. Fish. Aqua. Sci.* 39, 943-947.
- Scott, T.D., C.J.M. Glover and R.V. Southcott. 1974, *The marine and freshwater fishes of South Australia*. A.B. James (ed), Government Printer, South Australia. 2<sup>nd</sup> edition, 392pp.

- Sigurdsson, T. 1993. "Application of acoustic information in bottom trawl surveys." Cand. Scient. Thesis. University of Bergen, Norway, 135pp.
- Simmonds, E.J., N.J. Williamson, F. Gerlotto and A. Aglen. 1992. "Acoustic survey design and analysis procedure: A comprehensive review of current practice." *ICES Cooperative Research Report No. 187*. 131pp.
- Siwabessy, P.J.W. 1995. "A laboratory apparatus for the measurement of sea-water acoustic absorption." MSc Thesis. University College, The University of New South Wales, Australian Defence Force Academy, Canberra. 157pp.
- Siwabessy, P.J.W., J.D. Penrose, R.J. Kloser and D.R. Fox. 1999. "Seabed habitat classification." *Proc. International Conference on High Resolution Surveys in Shallow Waters* DSTO, 18-20 October 1999, Sydney, Australia.
- Siwabessy, P.J.W., J.D. Penrose, D.R. Fox and R.J. Kloser. 2000a. "Bottom Classification in the Continental Shelf: A Case Study for the North-west and South-east Shelf of Australia." *Acoustic 2000, Proc. Aust. Acous. Soc.*, 15-17 November 2000, Joondalup, Perth, Western Australia, 265-270.
- Siwabessy, P.J.W., J.D. Penrose and R.J. Kloser. 2000b. "A comparison between acoustically derived  $s_A$  estimates and those derived from nets." *ECUA 2000, Proc. 5<sup>th</sup> European Conference on Underwater Acoustics*, 10-13 July 2000, Lyon, France. M. Zakharia, P. Chevret and P. Dubail (eds), 1431-1436.
- Sorensen, P.S., K.N. Madsen, A.A. Nielsen, N. Schultz, K. Conradsen and O. Oskarsson. 1998. "Mapping of the benthic communities common mussel and neptune grass by use of hydroacoustic measurements." *Proc. 3<sup>rd</sup> European Marine Science and Technology Conference*, 26 May 1998, Lisbon, Portugal.
- Sorokin, Yu.I. 1993. *Coral Reef Ecology*. Springer-Verlag, Berlin. 465pp.
- Sullivan, B.K., P.H. Doering, C.A. Oviatt, A.A. Keller and J.B. Frithsen. 1991. "Interactions with the benthos alter pelagic food web structure in coastal waters." *Can. J. Fish. Aquat. Sci.* 48, 2276-2284.
- Tilzey, R.D.J. 1994. "The South East Fishery - a scientific review with particular reference to quota management." *Bureau Resour. Sci.*, Canberra, 360pp.
- Timm, N.H. 1975. *Multivariate Analysis with Applications in Education and Psychology*. Wadsworth Publishing Company, Inc. Monterey. 689pp.

- Tolstoy, I. and C.S. Clay. 1966. *Ocean Acoustics. Theory and Experiment in Underwater Sound*. Advanced Physics Monograph Series. McGraw-Hill. New York. 293pp.
- Tranter, D.J. 1962. "Zooplankton abundance in Australian waters." *Aust. J. Mar. Freshw Res.* 13, 106-129.
- Tucker, D.G. and B.K. Gazey. 1966. *Applied Underwater Acoustics*. Pergamon Press, Oxford. 244pp.
- Van Andel, T.H. and J.J. Veevers. 1967. "Morphology and sediments of the Timor Sea." *B.M.R. Bull.* No. 83.
- Urick, R.J. 1954. "The backscattering of sound from a harbour bottom." *J. Acoust. Soc. Am.* 26, 231-235.
- Voulgaris, G. and M.B. Collins. 1990. "USP RoxAnn ground discrimination system: a preliminary evaluation." ARE Portland UTH Tech Memo 36/90. RE005314. University of Southampton, Dept. of Oceanography, Marine Consultancy Services. *Tech. Rep.* No. SUDO/TEC/90/5C, 75pp.
- Walsh, S.J. 1992. "Factors influencing distribution of juvenile yellowtail flounder (*Liminda ferruginea*) on the Grand Bank of Newfoundland. Netherlands J. Sea Res. 29, 193-203.
- Waring, J.R., R.J. Kloser and T. Pauly. 1994. "ECHO -Managing fisheries acoustic data." *Proc. International Conference on Underwater Acoustics* University of New South Wales, Dec. 1994, 22-24.
- Whitelaw, A.W. 1998. "Cruise report ss 07/97." CSIRO Division Marine Research Marine Laboratories internal report, 9pp.
- Wickens, T.D. 1995. *The Geometry of Multivariate Statistics*. Lawrence Erlbaum Associates, Publishers. New Jersey. 165pp.
- Wong, H.K. and W.D. Chesterman. 1968. "Bottom backscattering near grazing incidence in shallow water." *J. Acoust. Soc. Am.* 44(6), 1713-1718.
- Woodward, J.A., D.G. Bonett and M. Brecht. 1990. *Introduction to linear models and experimental design*. Harcourt Brace Jovanovich Publishers, Academic Press, San Diego. 626pp.
- Wyrtki, K. 1961. "Scientific results of marine investigation of the South China Sea and the Gulf of Thailand 1959-1961." *NAGA Rep.*, Vol. 2, 195pp.