

BROAD SCALE REMOTE SURVEY AND MAPPING OF SUBLITTORAL HABITATS AND BIOTA

1999

R.L. Foster-Smith, J. Davies & I. Sotheran

This is the final Technical Report on sublittoral mapping methodology of the BMAP project, a project supported by the European Commission under the *Life* programme.

SeaMap Research Group
Department of Marine Sciences and Coastal Management,
University of Newcastle-upon-Tyne,
Newcastle upon Tyne,
U.K. NE1 7RU

BROAD SCALE REMOTE SURVEY AND MAPPING OF SUBLITTORAL HABITATS AND BIOTA

Preface

The BROADSCALE Mapping Project (BMP) was initiated by the statutory country conservation agencies (the Countryside Council for Wales, English Nature and Scottish Natural Heritage) and the Crown Estates. The objective was to develop methodologies for broad scale mapping of sublittoral habitats and biota (biotopes) based on acoustic remote sensing. The BMP was a three-year project funded by a consortium consisting of the Crown Estate, the Countryside Council for Wales (CCW), English Nature (EN), Scottish Natural Heritage (SNH) and the *SeaMap* Research Group (Newcastle University) and has been supported by the European Commission's *Life* programme.

The views expressed in this report are those of the authors and cannot be taken to represent those of the country conservation agencies or the Crown Estate.

Acknowledgements

The authors would like to thank Mandy McMath (CCW), Dan Laffoley (EN), John Baxter (SNH) and Tony Murray (Crown Estate) for their help and support throughout the project. The authors also received useful assistance on surveys and for analysis by other members of *SeaMap* including Robert Walton, Clare Wormald and Emily Strong. Many other people assisted with the field work and the authors would like to acknowledge the skill and support of the skippers and crew of the various boats used including Bill Watson of the Eastern Joint Sea Fisheries Committee, Alistair Simpson of Newcastle University's research boat *Bernicia*, Mark Burton from the Dale Fort Field Studies Centre, Steve Hartley skipper of MV *Sulair*, the skipper and crew of the RV *Prince Madog* of the University of North Wales, David Ainsley skipper of MV *Porpoise* and the skipper and crew of RV *Calanus* from the Dunstaffnage Marine Laboratory. Dr Peter Garwood of *Identichaete* provided valuable expertise with the identification of infaunal samples.

Advice and comments on the BMP outputs was also gratefully received from Alistair Davison (SNH), Helen Vine, Paul Gilliland and Ian Patterson (EN), Blaise Bullimore and Mike Camplin (CCW), Ivor Rees (University of North Wales, Bangor) and Bill Cooke (North Wales and North West Sea Fisheries Committee).

CONTENTS

1	Executive summary	3
1.1	The Broadscale Mapping Project: Mapping the habitats and biota of large areas of the seabed	3
1.2	The methodology	3
1.3	Survey of the three trial areas	4
1.4	Map accuracy and confidence	5
1.5	The future	6
2	The BMP Project	7
2.1	Aim of the Broadscale Mapping Project	7
2.2	Development of the project	7
2.3	Purpose of Technical Report	8
2.4	Structure of the Technical Report	9
3	Broad scale mapping in a management-support role	11
3.1	Introduction: The role of broad scale mapping in management	11
3.2	Defining broad scale mapping of habitats and biota based on remote sensing	11
3.3	Maps in integrated data management systems	12
3.4	Summary of the role of habitat and biotope maps in a management information system	13
3.5	Acoustic images and their analysis	15
3.6	Error and accuracy assessment	16
4	Technology of remote sensing in the marine environment	18
4.1	An overview of the technology available for broad scale survey in the marine environment	18
4.2	Choice of systems	20
4.3	An introduction to AGDS	20
5	An introduction to survey using acoustic ground discrimination systems (AGDS)	28
5.1	Acoustic survey and real-time AGDS	28
5.2	Generating a continuous coverage from track data	31
5.3	Interpreting AGDS data	32
6	Mapping units	35
6.1	Introduction	35
6.2	Biotope mapping units suitable for broad scale survey	35
6.3	Substrata	38
6.4	Limitations to habitat or biotope discrimination using acoustic data	39
6.5	The output from broad scale survey: what maps show	43
7	Assessing the performance of a map	52
7.1	Confidence and accuracy	52
7.2	Visualising and describing the data	52
7.3	Measures of accuracy and uncertainty	53
7.4	An assessment of the map from expert local opinion	54
7.5	Ease of use	54
8	Planning a broad scale mapping survey	56
8.1	Key stages in planning and executing a broad scale survey	56
8.2	Survey types based on AGDS	57
8.3	Integrated survey strategy	63
8.4	Design of field survey	65
9	Data analysis and display	68
9.1	Assembling an analysis and display software system for biotope mapping	68
9.2	Software used for data processing and map preparation	69
9.3	Data analysis: types of operation	69
9.4	Critical stages in data analysis	70
10	Preliminary AGDS data editing and treatment	72
10.1	Additional labelling and identification	72
10.2	Correcting depths to chart datum	72
10.3	Editing track data	72
10.4	Standardisation of E1 and E2	73
10.5	Deriving secondary variables	74
10.6	Summary of data-file structure	74
11	Exploratory data analysis and data treatment	75

11.1	Exploring relationships between E1 and E2	75
11.2	Classifying acoustic ground types	75
11.3	Spatial patterns in the AGDS data: The variogram	77
12	Interpolation and resolution	81
12.1	The purpose of interpolation	81
12.2	Interpolation of AGDS data	81
12.3	Grid spacing	85
12.4	Using bathymetric data to assist interpolation	85
13	Classification	87
13.1	Introduction	87
13.2	Regular arrangement of E1/E2 classes	87
13.3	Classification using calibration	87
13.4	Unsupervised classification	88
13.5	Supervised classification	88
13.6	Probability images and maximum likelihood classification	92
13.7	Displaying levels of probability	94
13.8	Uncertainty measurement	95
13.9	Using probabilities to classify track data and amalgamate the data from different surveys	96
14	Error matrices and crosstabulation	98
14.1	Introduction to matrices	98
14.2	Uses of matrices	99
15	Resolution of maps	102
15.1	Introduction	102
15.2	Consistency in the performance of the AGDS	103
15.3	Limits to spatial resolution	105
15.4	Survey intensity	107
15.5	Consistent use of habitat and biotope categories	108
15.6	Information from the AGDS used for discrimination	110
15.7	Intensity of ground truthing on consistency of interpretation	110
15.8	Variability in interpretation	111
15.9	Cartographic scale and resolution	111
16	The use of 3-D bathymetric models and drapes for interpretation.	114
16.1	Topography of the Firth of Lorn	114
16.2	Draping life forms over DEMs in Pembroke	115
16.3	Draping the probability distribution of <i>Sabellaria spinulosa</i> for an aggregate dredging area off Lincolnshire	117
17	Other AGDS systems.	119
17.1	Multi frequency echo sounders	119
17.2	QTC View	120
18	Integrating sidescan into broad scale mapping.	122
18.1	Northumberland	122
18.2	Firth of Lorn	124
19	Guidelines for broad scale survey	127
19.1	Preliminary planning	127
19.2	Expectations of a broad scale survey	127
19.3	Field survey design	129
19.4	Analysis	130
20	Conclusions	134
20.1	The broad scale survey and mapping methodology	134
20.2	Future developments of broad scale survey and mapping	135
21	Glossary	138
22	References	143
	Appendix 1. Life forms and colours used in the biotope maps with equivalent MNCR biotope complexes and MNCR colours	147
	Appendix 2: Modified Folks triangle used in sediment analysis	155

1 Executive summary

1.1 The Broadscale Mapping Project: Mapping the habitats and biota of large areas of the seabed.

Within the last decade our view of the marine environment has undergone a radical transformation as man's impact on the sea has become clearer and public awareness has been heightened. New and increasing uses of marine resources, legislative requirements and the range of duties placed on many bodies that manage the seas as well as the economic significance of activities offshore have transformed marine conservation into an essential element in considering any marine activity or development. Knowledge of the extent and distribution of marine habitats and the variety of life they support is a key requirement for the designation of marine nature reserves, special areas of conservation and marine protected areas. This information is also important when considering the possible impacts upon the marine environment from industries such as aggregate dredging, fish farming, fisheries, seaweed harvesting and oil and gas exploration and extraction.

Maps showing habitats and marine life (biotopes) on the seabed are needed to provide the spatial information so that decisions on the best way to manage these resources can be made. However, such map-based information is lacking for most of the waters around the United Kingdom. This is undoubtedly due to the difficulties in the past of acquiring an overview of the seabed. New techniques based on acoustic remote sensing make it possible to create images of large areas of the sea floor and they will provide broad scale maps of seabed habitats and marine life.

The Broadscale Mapping Project (BMP) was formed to develop appropriate methods for the survey and mapping of habitats and their associated life of the shallow seas. Many different organisations will need to become involved with mapping the seabed if large areas of the U.K. waters are to be covered and the objective of the Project was to devise a methodology that is readily accessible and cost-effective. In addition, distribution maps of habitats and biota should be easy to use by scientist and non-scientist alike.

The Project was devised and funded by a consortium consisting of the Crown Estate, the Countryside Council for Wales, English Nature, Scottish Natural Heritage and Newcastle University through the *SeaMap* Research Group. The project has also been supported by the European Commission's *Life* programme. The techniques have been trialed in three selected areas as a pilot project in order to provide the necessary experience and information for a possible future large scale operation to undertake mapping of the distribution of biotopes within European waters under UK jurisdiction. These areas are (1) Milford Haven and the Pembrokeshire coast and islands; (2) the Wash, Lincolnshire coast and north Norfolk coast, and; (3) the Firth of Lorn.

1.2 The methodology

The methodology developed relied on the relatively inexpensive acoustic ground discrimination system (AGDS) based on a single beam echo sounder. This system analysed the return signal from an echo sounder and gave a measure of the hardness and roughness of the sea floor, as well as depth. These records were converted to digital images that enabled powerful image processing software, designed primarily for the processing of satellite images, to be applied to the acoustic data.

Image processing required ground truth samples of the biota and habitats from the seabed. Remote video systems were especially useful since they were easy to deploy and provided a vista of the sea floor. Image processing was a three-stage process. In the first stage ground truth sites were used to calculate the acoustic characteristics (acoustic 'signature') of each habitat or biotope class. In the second stage each element of the digital image (pixel) was matched to the signatures and the

likelihood of the pixel belonging to each habitat or biotope calculated. Lastly, an automated process then selected the most likely category for the pixel.

However, image processing is only a part of the analysis that can be applied to broad scale mapping. The project incorporated other knowledge into the analysis, an approach termed Cartographic Modelling that has been adopted in the BMP to form a rule-based system for biotope mapping.

1.3 Survey of the three trial areas

A variety of areas were chosen to develop a single methodology that could be applied to a wide range of circumstances. Each of the areas has its own special characteristics: The habitats of the Wash survey area are largely comprised of sediment ranging from muddy sand to cobbles with smaller areas of soft mud and silty boulders. Both the Firth of Lorn and the Pembrokeshire survey areas have a much wider range of rock substrata. In particular, there are very deep rocky outcrops in the Firth of Lorn and this area probably supported the widest range of biotopes found during the BMP. The trial area in Wales also gave the BMP the opportunity to include other surveys and amalgamate different data sets. This extend the mapped area to include south Cardigan Bay.

If the methodologies for broad scale mapping are to be used to cover very large sea areas, then it is inevitable that the data from many different AGDS and operators will need to be amalgamated. This will require training and exchange of expertise between operators to ensure consistency and that the most up-to-date analysis is used. The BMP has given training in a number of instances. In particular, training in the use of *RoxAnn* and videography has been given to the Eastern Sea Fisheries Joint Committee and the North Wales and North West Sea Fisheries Committee.

1.3.1 Wash, Lincolnshire coast and north Norfolk coast

Much of the biota is buried in soft sediment and was sampled using the traditional biologists' grab. However, the area also has substantial Ross worm (*Sabellaria spinulosa*) reefs. This worm constructs tubes out of sand grains and the reefs were more effectively sampled using remote video. A complete description of the area required the combination of the two sampling methods and this has lead to a significant advance in knowledge of the biology of the region - the graphic description of substantial *Sabellaria spinulosa* reefs with their associated high diversity and large numbers of crustacean species.

These reefs have a great significance for the management of the biological resources and conservation interest of the area since the Ross worms are found in areas with strong tidal currents and a heavy load of suspended sand, an environment that many other species find harsh. In an environment that might otherwise be species-poor and not very productive, Ross worms create a 'worm-made' (biogenic) habitat that is an oasis of life. These reefs support high numbers of the commercially exploited pink prawn *Pandalus montagui*. A predictive map of their distribution has contributed to the formulation of policy towards licensed dredging sites and the monitoring of the favourable status of the reefs within the Wash candidate European Marine Site.

The Wash, Lincolnshire and north Norfolk coast was the largest sea area surveyed in the BMP and this gave valuable experience in the use of other previously published data to augment the picture of the area built up from the survey data. Three-dimensional models have been constructed and these show how predicted sediment distribution follows the ridges, banks and channels of the region. Visualisation will be important in making and explaining management decisions in the future.

1.3.2 Milford Haven and the Pembrokeshire coast and islands

The trial area in Wales gave the BMP an excellent opportunity to develop ways in which data gathered under different conditions can be analysed together. Altogether seven different AGDS were used in an area that stretched from Pembrokeshire to Aberystwyth.

Amalgamation was successfully carried out by applying image processing techniques to each data set and then amalgamating the probability plots for a particular biotope from each survey.

Another issue that becomes apparent with large sea areas is that there are broad scale trends in the distribution of biotopes. Unless these trends are taken into consideration biotopes can be 'found' in areas where they are unknown and extremely unlikely to occur. The use of Bayesian statistics allows broad scale trends determined from sample data to be incorporated into the image classification process. This process was successfully used in Wales and in other trial areas.

1.3.3 The Firth of Lorn

A very wide range of physical conditions was encountered in the Firth of Lorn. West facing coastlines and off-lying rocks are exposed to the full force of the prevailing Atlantic waves in contrast to the extremely sheltered conditions at the head of the mainland sea lochs. Similarly, powerful tidal currents sweep through the great whirlpool of the Coryvreckan and narrow sounds between islands whereas little water movement occurs in the deeper sediment basins of sea lochs. A huge range of depths was encountered within the study area. To the west of the Garvellachs, depths of over 200m were recorded within 1km of the shore, and vertical submarine cliffs over 90m high extended for many kilometres along the edge of a deep trough.

Mapping these extreme physical environmental regimes posed a considerable logistical and technological challenge. It was only possible to survey the offshore areas in calm conditions and the tidal channels at slack tide. Two AGDS systems operating at low frequency (28kHz) and high frequency (200 kHz) were required to survey the wide depth range and a special video camera system was required to record pictures from over 200m deep in the bottom of the troughs. Over fifty biotopes were recorded from the study area. Dense fields of featherstars were encountered on mud in the deep troughs - a rarely recorded biotope previously only observed by submarine.

Processing the acoustic and video data to create maps relied heavily on mathematical models and probability statistics. Correction factors were developed to offset the effect of the depth on the acoustic signals, and to combine data from different acoustic systems used on different boats. Extending the probable distribution of biotopes from known locations helped account for the spatial heterogeneity of the physical environmental factors.

1.4 Map accuracy and confidence

All maps are predictive and, ultimately, confidence will only be established through their use. However, the BMP has applied accuracy and uncertainty assessment procedures based on error matrices and spatial modelling that are designed to indicate the level of confidence that can be placed in the maps.

The smallest scale (1:250 000) shows broad trends in biotope distribution whilst the larger scales (1:100 000 and 1:50 000) give a much better idea of how different biotopes interrelate spatially.

However, there are limits to the accuracy of the technology and broad scale survey is not designed to give highly accurate boundaries to biotopes at a small scale. Indeed, one of the justifications for broad scale maps is their use in identifying areas where a greater level of detail is required to address specific questions, particularly those related to monitoring change and the effects of human intervention on habitats. Thus, broad scale mapping is part of an integrated survey strategy and links desk-top studies to detailed, specific-issue surveys.

1.5 The future

The methodology to emerge from the BROADSCALE Mapping Project is based on acoustic remote sensing and it is capable of mapping habitats and the marine life of large areas of the seabed. The methodology has been tried by other organisations and the BMP has demonstrated that results from different surveys can be amalgamated.

1.5.1 Mapping wider sea areas

There is no doubt that broad scale habitat and biotope maps are a useful resource in their own right for supporting management of large sea areas, as can be judged by their use to date in areas of high conservation interest. If these broad scale surveys are extended over very large sections of the U.K. coastline, this new knowledge of the distribution of biotopes will help in the management of the wider seas.

1.5.2 Maps in data management systems

Biotope resource maps are just one layer of information of an electronic integrated data management system. A single layer showing the most likely biotope distribution is easiest to view, but probability images for each biotope are even more versatile and can readily be incorporated into risk assessment models.

1.5.3 Maps in the study of dynamic processes

Biotope maps are essentially a snap-shot in time. How will distribution change over time? How can our knowledge of the dynamics of habitats and biotopes be incorporated into dynamic distribution models? Again, probability maps could form the basis of dynamic models by providing predictions on changing patterns of biotope distribution at different scales that can be tested through a highly targeted, stratified sampling program.

1.5.4 Maps as part of a wider integrated survey strategy

The lessons learnt from the use of mapping in conjunction with more traditional grab sample surveys could have enormous implications for the future design of monitoring surveys. Knowledge of the distribution of biota could structure future surveys leading to a cost-effective strategy that is based on refining our knowledge of the processes that determine biological community structure.

Biotope maps will undoubtedly fulfil a central role in the support of scientifically-based decision making in the marine environment. The next step is to provide this spatial information for larger sea areas and in a format that can be used in the new generation of data management and risk assessment systems that are being developed for integrated marine management.

A new generation of data management and risk assessment systems are being developed for integrated marine management and these require, above all else, resource maps which can be used as a context for other data, a basis for modelling and prediction as well as simply a visual interface for querying the database for information held. These resource maps will help in the management of specific sites as well as whole regions and will help to formulate future strategies for monitoring the health of the seas around the U.K.

2 The BMP Project

2.1 Aim of the Broadscale Mapping Project

The aim of the Broadscale Mapping Project (BMP) was to develop appropriate techniques for the survey and mapping of habitats and their associated biota of the shallow seas. These techniques were trialed in selected areas as a pilot project in order to provide the necessary experience and information for a possible future large scale operation to undertake mapping of the distribution of biotopes within European waters (out to the continental shelf) under UK jurisdiction.

2.2 Development of the project

Within the last decade our view of the marine environment has undergone a radical transformation as man's impact on the sea has become clearer, reflected in a marked increase in public awareness. Seabed habitat and biota information in the form of maps provides the spatial context within which scientifically based decisions may be taken. Maps showing the distribution of marine habitats and biota are key requirements for the designation of marine nature reserves, special areas for conservation and marine protected areas. The maps could be equally used by those who give marine conservation advice, those who wish to exploit the seas renewable (e.g., fisheries) or non-renewable resources (e.g., aggregates, oil and gas) and to those who need to formulate policies and set the balance between potentially conflicting interests. The key requirement is for basic spatial information on habitat and biota, particularly in the subtidal zone out to territorial limits.

Until recently obtaining information on seabed habitats and biota was largely dependent upon slow and expensive observations by divers or extensive infaunal grab sampling surveys. However, seabed mapping based on acoustic techniques coupled with rapid techniques for ground truthing can provide a means of identifying the location and extent of habitats and their associated biota. A methodology for mapping habitats and biota using acoustic remote sensing techniques was developed by the LIFE funded *BioMar* team at Newcastle University and this provided a cost-effective and rapid system for acquiring data. The Broadscale Mapping Project (BMP) was initiated by the statutory country conservation agencies (the Countryside Council for Wales, English Nature and Scottish Natural Heritage) and the Crown Estates with the aim of continuing the development of the mapping techniques to the point where they could be used on a larger scale for the mapping of the distribution of biotopes within European waters under UK jurisdiction. This pilot project was designed to provide the necessary experience and information for a possible large scale operation.

At a general level this project is a significant contribution towards the overall goal for the UK of integrated marine resource management. This can be viewed as having seven principle information needs which have to be considered together:-

- (a) Seabed topography
- (b) Geology and sea bed sediments
- (c) Seabed habitats and biota (biotopes)
- (d) Seabed dynamics (processes and change)
- (e) Visualisation of management practices
- (f) Visualisation of the extent of voluntary and non-statutory plans
- (g) Visualisation of patterns of exploitation of resources

The aim for the BMP was to present spatial information on seabed habitats and biota as one layer of information (e.g., in a geographic information system) which, when combined with other layers of data would provide a comprehensive picture of both the living and non-living attributes of the seabed. At a more detailed level, this project was also designed to produce information to support a wide range of local initiatives and programs, particularly those concerned with marine survey and monitoring. The habitat and biotope maps complement information already available from navigation charts and seabed sediment maps.

The project has been funded by a consortium consisting of the Crown Estate, the Countryside Council for Wales, English Nature, Scottish Natural Heritage and the Newcastle University (*SeaMap* Research Group). The European Commission under the Life programme has also supported it. During this project specific technical development work undertaken was as follows:-

- (a) the development of surveying and mapping methodologies suitable for large geographic areas, including offshore seas out to territorial waters;
- (b) the application of acoustic remote sensing and remote sampling systems for the recognition of specific habitats and biota;
- (c) the further development of acoustic systems based on echo sounders that had previously been utilised by the *BioMar* project and the use/appraisal of other cost-effective acoustic technology;
- (d) the incorporation of data analysis and presentation into a PC-based management support system.

Three large sea areas were selected as the main focus areas for the survey work. These were Milford Haven and the Pembrokeshire Islands, Lincolnshire and north Norfolk (including the Wash) and the Firth of Lorn (Figure 2.1). The experiences from other surveys carried out by *SeaMap* were also incorporated where appropriate. The main outputs from the BMP are this technical report giving details of the methodology, a report for each of the three trial areas and maps of each of the trial areas at a scale of 1:250 000 or, for smaller areas within the main survey areas, 1:100 000 - 1:50 000 showing the distribution of habitats and biota. There was also the requirement that, where possible, the methodology was to be shared with other potential users through training and collaboration.

The BMP built upon the work already undertaken by the *BioMar* project to provide a cost-effective method for broad scale mapping. Since the methodology developed is to be recommended for widespread use by organisations wishing to undertake broad scale mapping, the techniques adopted are affordable and, also of importance, the data produced are relatively easy to process and interpret. The team at Newcastle University reviewed acoustic and remote sampling techniques that were available and selected the most appropriate techniques for the BMP.

The availability of computer software, in particular for image processing and geographic information systems (GIS), has advanced very rapidly over the last decade. The BMP has made extensive use of PC-based GIS and image processing systems for the analysis of acoustic images. It is one of the aims of the technical report and training given to potential users to promote the use of these software programs for marine survey and to link outputs to existing GIS.

2.3 Purpose of Technical Report

The purpose of the Technical Report is to give guidance on the use of acoustic techniques for broad scale mapping. There are two intended readerships:

- Firstly, it is hoped that those commissioning mapping surveys or using the resulting habitat or biotope maps will benefit from an appreciation of the principles underlying broad scale survey and mapping using acoustic remote sensing. Part 1 of this Technical Report is designed as a guide through these stages.
- Secondly, the document gives details of survey and analytical techniques so that other surveyors can follow the procedures developed for the BMP and adapt them to meet particular circumstances. Acoustic mapping is a rapidly evolving science, but it is hoped that the generalities presented will keep pace with changing technologies and advances in analytical software even though details of procedures will be overtaken by developments.

2.4 Structure of the Technical Report

Part 1. An introduction to broad scale mapping.

The readers are introduced to the basic concepts of broad scale mapping and how it integrates with other types of survey in Section 3. There follows a description of the remote sensing technology in Section 4 and an introduction to the use of acoustic ground discrimination systems (AGDS) for habitat and biotope surveys in Section 5. This lays down the principles involved with acoustic ground discrimination and the interpretation of the data. A basic understanding of the acoustic ground discrimination system and the remote sampling techniques is required to appreciate the issues connected with the choice of biotope mapping units (Section 6). This Section also gives examples of what maps can show and their information content using examples from each of the three trial areas. Part 1 ends with a short discussion on sources of error and uncertainty and how they may be assessed (Section 7).

Part 2. Planning a broad scale survey and analysis of the results.

Purpose of this part will be to take the reader through the complete process of planning and carrying out a broad scale survey, including data analysis, assessment and map preparation. Section 8 discusses the aims, objectives and appropriate designs for different types of broad scale survey. Section 9 outlines analytical procedures which are then discussed in more detail in the following Sections. Editing the data is explained in Section 10 and this is followed by a description of various exploratory data analysis techniques in Section 11. Interpolation of point data to create a continuous coverage is the first stage in the interpretation of the data (Section 12) and the resulting digital images are then amenable to various classification techniques available in image processing/GIS packages, as described in Section 13. Error measurement is a vital for assessing the quality of the output of image interpretation (Section 14) and this is developed further in Section 15 with a discussion on scale, resolution, detail and their relation to cartography.

Plan maps are just one way of displaying habitat and biotope information and three dimensional views where the information is draped over bathymetric models have also been used extensively in the BMP to help interpret habitat and biotope distributions (Section 16). Section 17 discusses other AGDS and Section 18 describes how the BMP has used sidescan sonar in the context of broad scale mapping and particularly how it has been used in conjunction with AGDS.

Section 19 outlines the main points to emerge from the BMP which can be used as guidelines for broad scale survey and, lastly, the overall achievements of the BMP are summarised and future directions for broad scale survey described in Section 20.

Notes:-

1. The main technical terms are in **bold** when first introduced and defined or explained in the Glossary.
2. A two digit system is used to number tables and figures: The first number is taken from the major section in which they are located and the second number follows the sequence in which they appear.
3. The data are primarily taken from the BMP surveys of the three trial areas, but additional data and examples are drawn from other *SeaMap* surveys. Other surveys in which AGDS have been used for habitat or biotope mapping that have been reported (either in the scientific or technical press or in grey literature reports) have been reviewed.

Part 1. An introduction to broad scale mapping.

The purpose of Part 1 of the Technical Report is to introduce the basic concepts of broad scale mapping. Readers who are involved with planning and commissioning broad scale surveys or who are reading survey reports may wish to appreciate these basic concepts without the need for a detailed knowledge of the methodology. There are many forward references in the text to Part 2 (a more detailed account of the processes) so that readers can easily look up this information should they wish to do so.

This division of the Technical Report into two parts inevitably results in some repetition as many of the issues introduced in Part 1 are taken up again in Part 2.

3 Broad scale mapping in a management-support role

Purpose: The aim of this Section is to explain the management requirement for map-based information and, in particular, the requirement for broad scale maps; to introduce the underlying concepts and expected outputs of broad scale mapping surveys in order to appreciate the role of mapping in an integrated survey strategy.

3.1 Introduction: The role of broad scale mapping in management

Maps which show the distribution of habitats and biota, together with accompanying data and statistics, are central to many aspects of environmental appraisal, but particularly for the assessment of the natural heritage (conservation) interest and the impacts of human activities on biological resources. Broad scale survey and mapping can provide a basic level of knowledge of the biological resources of an area where there is little existing information. The purpose of broad scale mapping is to identify the basic resources of an area in terms of major physical habitat characteristics and community types and to map their distribution and extent. The justification for this level of survey is given as providing a knowledge of the location, extent and distribution of wildlife and associated habitats which forms an objective baseline for (1) determining which sites might deserve consideration for protection and warrant a more detailed survey; (2) determining the strategic importance of sites to nature conservation, including those sites of relatively limited conservation interest; and (3) making an assessment of the environmental impacts of human activities. This knowledge enables planners to respond quickly to proposals for development, accidents and to save time and money through anticipation and avoidance of controversy (Nature Conservancy Council 1990). Broad scale survey also provides data for developing management plans for marine protected areas and have a key role in defining site boundaries and boundaries, at a coarse level, of specific features of conservation importance (Downie *et al.* 1999).

Underlying this statement is the importance of dialogue between all those concerned with environmental management and involved with environmental appraisal, impact assessment and monitoring. A common understanding of biological resources is necessary for dialogue and the resolution of conflicting demands upon the environment. Broad scale maps can provide an elegant way to display complex information for this purpose.

Broad scale surveys conducted over large areas are useful in alerting coastal managers to the potential conservation interest of a site, perhaps identifying the site for the first time. Such data, especially if accessed through map-based enquiry systems, could serve as the primary route for retrieving information about specific locations when reacting to issues with potential implications for conservation management. Perhaps the most valuable use of broad scale maps is to highlight where more detailed information is required and what type of survey would be most suitable. Mapping and associated surveys should not, therefore, be seen in isolation but as part of a larger integrated survey strategy.

3.2 Defining broad scale mapping of habitats and biota based on remote sensing

A map is normally defined as a representation to scale of some feature of the earth's surface (Lawrence 1971). However, this is too general for the purposes of this report and it is worth taking some time to define what is meant by broad scale mapping before the purpose, strategy and techniques for mapping of subtidal habitats and biota can be discussed.

The issue of scale is central to mapping and some explanation of the terms ‘broad scale’ (as opposed to ‘fine scale’) is required. It should be stressed that there is no hard and fast distinction between broad and fine scale mapping, but the following account contrasts important elements.

Broad scale implies that large areas are mapped and show the approximate disposition of broadly defined classes of habitats or biotopes (small scale). Fine scale implies that small areas are mapped to a higher level of detail and accuracy (large scale). The distinction between broad and fine scale mapping can be linked to survey techniques, particularly in the subtidal environment. An area is mapped to a fine scale if a complete coverage of biotopes is obtained through direct observation and boundaries are known with some precision throughout the mapped area. Clearly, such intensive survey is expensive if conducted over large areas.

An alternative approach adopted for broad scale mapping is based on remote sensing. For a general account of remote sensing see Cracknell and Hayes (1993) or Wilkie and Finn (1996). An image of the area is obtained using remote sensing and information about certain attributes of the seabed is collected by direct or remote sampling at selected sites. The distribution of these attributes is inferred from the relationship between the sample data and associated characteristics of the image - a process called **ground truthing**. There are two important points here: Firstly, the object of this process is to give an attribute to every possible location within the surveyed area (in other words, to generate a **continuous coverage**). Secondly, the objects or attributes to be mapped are not directly measured at all locations, but are estimated from the remotely sensed image. The word ‘estimated’ is emphasised since it is important to realise that there will be a level of uncertainty about any estimate and that the maps are predictive. Further sampling will be needed based on these predictions to indicate just how valid the map is (how accurate the predictions are).

Attention must also be drawn to the distinction between maps that are produced from point samples of biotopes alone and maps derived from remote sensing. In the former case boundaries can only be drawn around sample points with the position of the boundary lying approximately half way between it and adjacent dissimilar samples. Boundaries drawn on such a basis are of little use when there is a large variation between one sample point and the next. Remote sensing, on the other hand, provides a basis for drawing boundaries between point samples and, by extrapolation, for interpreting parts of the image where little or no ground truthing exists.

3.3 Maps in integrated data management systems

Whilst there will always be a demand for maps in paper form, electronic maps and underlying databases are increasingly being used in integrated data management systems. Electronic data storage and manipulation allows for very flexible queries of large and diverse datasets. In general, these sophisticated data management systems include a comprehensive database on existing data and combine this with remotely sensed images, geographic query, modelling and information technology. The purpose that these systems have is to make best use of existing information to address a wide range of issues, including ‘what-if’ questions, without the need for additional data gathering. Clearly this makes sense if the result is cost-effective and avoids duplication of effort, especially of expensive survey work (Doody 1999). Two examples of such systems are the Environmental Risk Assessment package developed by Associated British Ports (Hannan 1999) and *SimCoast*TM which is designed to support coastal governance (Hogarth 1999).

However, these systems can only work with the data that are available and, being essentially geographic-based systems, they require basic resource distribution maps above all other forms of data. In electronic form, these maps can be thought of as an **atlas of thematic maps**, each showing a selected subject (such as habitat or biotope type). The process of data analysis, map preparation and map-based queries can be done electronically using the basic tools common to most geographic information systems (GIS). At an elementary level, maps can be used for geographic query of a database. Two or more thematic maps (e.g., biotope maps and jurisdictional boundaries) may be overlain to produce a third map showing biotopes within a particular area of interest. These are not particularly sophisticated operations. Maps can be reclassified, perhaps using a biotope map to create a sensitivity map if biotopes could also be classified as to their sensitivity to a particular impact. More

complicated mathematical or logical operations can be performed on a single biotope map or with map overlays. This could include inputs from dynamic models. Thus, quite complex environmental modelling can be based on resource maps.

Categorical maps which show only one category at any location can give a misleadingly definitive picture of habitat or biotope distribution. They portray the categories as being mutually exclusive and with sharply delineated boundaries (often termed '**hard**' or '**crisp**' boundaries). However, the reality is, as will be discussed in Section 5, that these categories are not mutually exclusive nor are boundaries necessarily clearly defined on the ground. Many of the problems associated with categorical data on maps can be overcome through the use of biotope **probability maps**. These show the likelihood of a biotope being present. These probability maps extend the usefulness of resource mapping in integrated data management systems by facilitating mathematical operations that can be performed on probabilities. This topic will be discussed in more detail in Section 13.

However, some caution must be exercised with all data management systems since they cannot replace individual expertise in particular subject areas and cartographic knowledge systems are no different from any other 'expert system' in this respect (Kraak and Ormeling 1996). The application of 'rules' in cartographic modelling is excellent discipline for an analyst. It forces the analyst to think about the procedures adopted and to justify each decision and choice of parameters used in any calculation. But ultimately, accuracy and confidence of the underlying data (and this includes the resource maps) underpins the success of any analysis or modelling.

3.4 Summary of the role of habitat and biotope maps in a management information system

Maps have a wide range of applications to management and form an important and integral part of management information systems. The following is a synthesis of some of these applications.

1. Baseline surveys

- **Inventory and proportions of different biotopes:** Map-based surveys help create a balanced inventory of the major biotopes (in terms of representation in an area). Furthermore, a comparison of areas occupied by different biotopes is valuable in its own right for assessing the importance of each biotope in terms of extent: Statistics on the extent of biotopes can be used to quantify statements about the rarity or typicalness of a biotope at the regional, national and international scale.
- **The wider seas:** Broad scale, rapid survey permits a wide area to be covered so that areas with benthic biotopes of high, moderate and low conservation status can be described and delineated. This is particularly important in the marine environment where activities in one area can have a high impact on neighbouring sites. Broad scale information is needed to establish and justify zones for different activities not just in sites of special conservation importance, but in the wider seas in which they lie.
- **Description of areas:** Broad scale biotope maps show the general distribution patterns of the habitats and biological communities in an area. This is vital for building up an overview of the area that is needed to help explain the significance of unusual features, variations in specific biotope attributes (such as rare species) or the concentration of features of interest in parts of the survey area. Although maps are essentially a 'snap-shot' in time, they can lay the foundation for an understanding of dynamic processes and their spatial implications.

2. Bringing together data, Graphic User Interface (GUI) as front end for an integrated data management system

- **Base map for showing other biotope data:** Maps can place in context available detailed biotope information, much of which are point source data. The significance of these data can be more fully appreciated if the likely spatial extent can be estimated from the broad scale biotope maps. The implications of studies of single biotope types or species can be more

clearly assessed and quantified if biotope extent and distribution is known. Spatial information can be built into ecosystem models.

- **Relation to land and sea use:** The position of biotopes relative to patterns of ownership and land use will have implications for management. In this context, maps are a suitable way of summarising the interaction between different types of information.
- **Biotope maps as base maps for geographic query:** Perhaps one of the most versatile ways to search for and use data collected by different surveys is through geographic query. If data have an associated geographic location (**georeferenced**) then they can be queried on the basis of position. Maps, especially if in a GIS linked to electronic databases or spreadsheets, form the natural front end for geographic query and for displaying the results. This is particularly useful in the sublittoral environment where different data sets may be used to help interpret an acoustic image.

3. Hypothesis generating and survey planning

- **Maps as predictors:** Maps based on remote sensing predict the distribution of habitats or biotopes. These predictions should be based on 'rules' (sometimes termed **cartographic models**) which can be tested and modified. Maps, in this predictive capacity, can help to refine our knowledge of the processes that structure biotopes.
- **Planning more detailed survey:** Planning a survey based on prior knowledge of the area is an extension of the use of maps in their predictive role. Broad scale maps can be used to plan more detailed survey by ensuring an adequate and equitable coverage of biotopes. This will lead to a greater return value from detailed survey for the resources committed to it.

4. Monitoring change/effectiveness of management

- **Establishing a monitoring programme:** Broad scale biotope maps are useful for developing a meaningful monitoring programme. Monitoring may require the selection of a limited number of sites for regular detailed sampling (for example, to monitor the population of specific species of interest, general species diversity, biomass and productivity). These sites should be chosen so that the data collected are not susceptible to small fluctuations in biotope boundaries or to poor positioning. In other words, biotope maps would indicate where suitable large homogenous areas are located. Monitoring might also require the repeat survey of transects. Again, biotope maps might be used to select suitable locations for transects.
- **Remote survey as a monitoring tool:** Broad scale survey techniques might be used in conjunction with other more detailed survey techniques to indicate if any gross changes in biotope distribution take place. Awareness of such changes could trigger targeted, more detailed survey.

In summary, maps should be seen as part of an **integrated survey strategy** and not as a stand-alone product of one survey. Surveys in general can range from the geographically extensive to the site specific and provide a general broad brush account of biotopes or more exhaustive, detailed descriptions and a knowledge of dynamic processes. Viewing the marine environment from the broad perspective to the more detailed is often termed the '**top down**' approach, whilst piecing together detail into the broader picture is termed '**bottom up**'. They are complementary: the top down approach can use broad scale maps to provide the basis for the design of more detailed survey and the bottom up approach can use maps to provide a broad context for the interpretation of the data from detailed survey.

Maps derived from broad scale survey can be used to design a subsequent **stratified sampling** programme where **prediction** based on broad scale survey can form a hypothesis for testing. A stratified sampling programme is more cost-effective than 'blind' sampling. Detailed sampling programmes are usually stratified to some extent, but the justification for the choice of locations for sampling may be based on very little available information. Broad scale survey provides much more information on the spatial distribution of biotopes than is available from other sources. The demand for stratified sampling based on acoustic remote sensing has been recognised by management groups as a way of focussing limited resources (Curran 1995).

3.5 Acoustic images and their analysis

Acoustic techniques can be used in the subtidal environment to produce an image which is continuous (or approximately so in the case of AGDS) over the whole of a survey area. The characteristics of the sound that is reflected from the seabed depends primarily on the physical nature of the sediments and the spatial patterns of acoustic reflectance are used to generate an image of the sea floor. These images, although nominally representing continuous data (in the sense that every possible location on the image has its own attribute value), are actually area data. The images are made up of rows and columns of small tiles that abut onto each other. These tiles, or **pixels**, have defined dimensions and area. Each pixel will have an attribute value that applies to any point within the pixel. Thus, an image can be represented by rows and columns of numbers that lie in the centre (or **centroid**) of each pixel. This is a **digital image**. Once in this format, the images are amenable to a wide variety of mathematical and logical operations collectively termed **image processing** (Wilkie and Finn 1996). This approach was developed for AGDS data in the *BioMar* project (Sotheran *et al.* 1997) and has been adopted as the standard methodology of the BMP.

As with all forms of remote sensing there are limits to resolution and what can be seen directly from the images. It is important to understand exactly what the stages are in the interpretation of the image:

- (a) What is actually sensed? In the case of sonar this is the reflective properties of the seabed (the **acoustic ground type**).
- (b) What can be directly interpreted from this using links between ground truth data and acoustic characteristics? These acoustic images need to be interpreted using point sample data. This involves establishing links between the acoustic ground type and those habitat features of the seabed that are responsible for its reflective properties.
- (c) Many important attributes of the ground (particularly biological attributes) will not affect the acoustic response and so cannot be directly interpreted. How can the distribution of these attributes be inferred from the acoustic images?

Links between acoustic data and habitat or biotopes are rarely sharply defined. The links are more statistical and, since many acoustic variables may be used to discriminate between habitats, the statistics are likely to be multivariate. Thus, the analyst will need to consider a series of images that can be superimposed on each other, each one representing a separate (and, ideally, independent) variable. Each point on the earth's surface within the image is represented by values from all of the superimposed images.

The last point raised above (point 'c') is the most intractable and the one that most concerns biological resource mapping. In principle these point attribute data are linked to the acoustic data by a **set of conditions** which are, in effect, a series of hypotheses that can be tested and subsequently modified in the light of experience. Examples of conditions that link acoustic data to biotopes are the statements (1) 'kelp forests are found on hard substrata' and, (2) 'kelp forests are found in shallow water'. Clearly, a whole series of conditional statements (perhaps more quantitative than the examples given above) might be required to define the areas where kelp can be predicted with confidence.

Interpretation that proceeds through these stages, building upon what is sensed and incorporating more and more information to refine the habitat or biotope maps, is termed **cartographic modelling** (Eastman, 1997) and provides a powerful and disciplined route for resource mapping.

The above is the basis for all remote sensing and mapping through image processing irrespective of the technology. Many acoustic techniques of the present day have poor spatial resolution and the information may be limited to a few variables. Indeed, as will be seen with the single beam acoustic ground discrimination systems (Section 4.3), the data may not be a complete coverage. Although advances in sonar technology will provide greater resolution and information, it is unlikely that any remote sensing system will 'see' the biota directly in sufficient detail to be able to identify the biotope. There will always be a need for extensive ground truthing and careful interpretation. Thus, the principles behind mapping biota on the sea floor as laid out in this document are likely to remain

valid even though technological advances may overtake details of sonar applications as used in the BMP.

This being the case, it is worthwhile discussing at this stage some of the general issues regarding sources of error and uncertainty, limits to resolution as well as cartographic limitations.

3.6 Error and accuracy assessment

Accuracy of maps and the level of confidence that should be invested in them is extremely hard to convey visually or to quantify since there are so many different ways for evaluating and expressing accuracy and uncertainty. Nevertheless, some indication of confidence must be given for the maps to be useful to environmental managers and scientists alike.

3.6.1 Sources of error

The key steps in the construction of a map should be explained and illustrated. Those commissioning or using maps of the seabed (or any other survey based on remote sensing) should have a knowledge of the way in which the maps are constructed. The limitations to the information content of maps due to both survey methodology and cartography should be appreciated by the user. These are examined in more detail in Section 15, but the potential sources of error are:-

- (a) **The inherent variability within and between the remote sensing system:** All remote sensing techniques have inherent variability that degrades their ability to discriminate features on the ground.
- (b) **Positional error of the remote sensing system:** Satellite positioning systems have revolutionised position fixing on survey. However, it is unrealistic to expect real-time positions to have a greater accuracy than 5m (Gilbert 1999) for the survey vessel.
- (c) **Positional error of the ground truth system:** Sampling devices deployed from a vessel will have positional uncertainty over and above that of the vessel itself. Its position on the seabed relative to the vessel can only be estimated from the length of umbilical or wire paid out and the direction of drift of the vessel and sampling device. The combined positional error can be quite considerable.
- (d) **Combined errors for images and ground truth data:** Image processing requires the location of the ground truth samples on the image so that image characteristics can be associated with the ground truth classes. The combined positional errors will give rise to a locus (or 'envelope' of uncertainty).
- (e) **Measurement error of the ground truth data:** Image processing can be undermined by poor attribute measurement of the ground truth samples. This is particularly likely when the attributes are habitat or biotope classes and the analyst must decide how best to match the sample data to a classification system.
- (f) **Variability due to analysis:** There can be many stages involved in image processing from data editing through to statistical analysis and modelling. However, the route followed by an analyst may be hard to replicate by another person since there are many possible pathways, each with different parameters that must be set. It is hoped that analysis is robust, but there is always the possibility that the interpretation is sensitive to apparently trivial parameter settings.
- (g) **Cartographic simplification:** Cartography is a mixture of aesthetics and science. There is generally a compromise between scale, map detail and ease of use of the map. Too much detail will render a map confusing and hard to read.

3.6.2 Map confidence

Maps showing track data, ground validation points, acoustic properties and interpreted habitat maps will all help people assess the confidence of biotope maps. These maps can be presented along with the main habitat and biotopes map. This is common practice both for the British Geological Survey quaternary sediment maps and the Soil Survey soil maps which both suffer from the similar problems of synthesising different types of data into one interpretation.

The strengths of the relationships between the remotely sensed variables and the ground attributes can be tested statistically and the results of the interpretation of the remotely sensed images based on these relationships assessed for accuracy. It is likely that this assessment will, in the first instance, be a measure of the consistency with which the ground truth data used to classify the acoustic data match the resulting maps. This is a measure of **internal accuracy**. The ground truth data are overlain onto the biotope map and the proportion of correctly matched pixels counted.

A more satisfactory measure is the use of an external data source instead of the ground truth data. The method of calculation of the accuracy measure is exactly the same as above although in this case the map's **predictive** capability is being tested and, since these samples are separate from the ground truth samples, this gives a measure of **external accuracy**. Ultimately, the value of the biotope maps will be established through their use in a management support capacity.

4 Technology of remote sensing in the marine environment

Purpose: To review current technologies as a background necessary to understand use in mapping surveys.

4.1 An overview of the technology available for broad scale survey in the marine environment

The high absorption of electromagnetic radiation in all but shallow, clear water means that satellite and airborne remote sensing have limited application to the remote sensing of subtidal seabeds. Sound, however, does travel readily through water and acoustic techniques can be used to obtain images of the seabed. Sonar, therefore, replaces traditional remote sensing technology using electromagnetic spectra (EMS) for sea floor imaging. There are many different acoustic systems available for survey and Riddy and Masson (1996) give a very accessible account of a range of sonar techniques.

That is not to say that EMS remote sensing techniques have not been used successfully in the subtidal environment and airborne electromagnetic systems are reviewed briefly to balance the emphasis placed on acoustic systems in this report. Their inclusion also forms a link to the image processing software used in the BMP that is designed primarily for the digital images EMS systems produce. These image processing systems are extremely sophisticated and there is a wealth of experience in applying remote sensing to scientific and management issues. The ultimate goal would be, therefore, to emulate these remote sensing technologies using sonar. Acoustic technologies are emerging that can produce continuous digital images with many layers of information (equivalent to EMS multispectral images) and will become less expensive. It is worth stressing the point that the general principles established in this report on the use of image processing as applied to acoustic mapping of biotopes will apply equally to these emerging technologies as to those used in the BMP.

4.1.1 Airborne remote sensing based on electromagnetic radiation

Aerial photographs using ordinary colour film have been used for very shallow water and compact airborne spectrographic imagers (CASI) in tropical waters (Green *et al.* 1996). They have high resolution, a good range of spectral information and good coverage. Additionally, images are often available through archives but, if not available, airborne systems are cost-effective to deploy. Their disadvantages are that they have poor and locally patchy penetration through water. Light spectra are variably absorbed by water so that quite complicated corrections are required to compensate for depth. Since depth is not measured, bathymetric information must come from an alternative source.

Lidar (Laser incidence direction and ranging) has been used for comprehensive bathymetric surveys (Sinclair 1998). Lidar is an airborne active scanning system. A pencil of laser light is directed over the ground in a grid configuration (typically one point per 4 sq metres). The time difference between the first weak reflection (from the water surface) and the second reflection (from the sea floor) is measured from which depth can be calculated (Muirhead and Cracknell 1986). Resolution is good and penetration is about 3 x Sechi disc distance (visible light penetration). The bathymetric data are of high quality, meeting the International Hydrographic Organisation standards, and a swath width of between 200 and 500m means that a considerable area can be covered rapidly. However, Lidar only measures bathymetry and the system is, at present, expensive.

4.1.2 Single beam acoustic ground discrimination systems (AGDS)

These use a typical echo sounder to generate a signal and they analyse the echo. They not only measure depth under the vessel but also reflective properties of the seabed. The complex shape of the echo is determined by characteristics of the seabed. Fishermen have long since used the picture of their echo sounders to give them information on the seabed and this stimulated the development of

*RoxAnn*TM, one of the first AGDS (Burns *et al.* 1985). *RoxAnn* measures two parameters of the echo that are taken to represent hardness and roughness. Two parameters are not many and hardly measure up to the **multispectral** information from airborne remote sensing systems. However, it is relatively inexpensive and easy to operate, measures depth and can be set to work in any depth. Another proprietary system, *QTC View*TM measures many more parameters than *RoxAnn* and offers the potential for detailed analysis of echoes for the discrimination of different ground types (Collins *et al.* 1996).

All AGDS have the drawbacks that coverage is poor since the data are essentially point information along tracks with no information available between tracks. A continuous coverage needs to be mathematically interpolated. It is also time consuming to build up a comprehensive coverage from running parallel tracks.

4.1.3 Sidescan sonar

Blondel and Murton (1997) give a general account of the principles of the sidescan sonar. Two transducers (one port one starboard) are usually housed in a 'fish' which is towed behind a vessel or mounted in a bracket in front of the bow of a boat. Each transducer emits pulses of sound which are 'fan-shaped' in that they are very wide in the port/starboard plane but very narrow fore/aft. A single pulse results in an echo from the seabed that is more distant as the slant angle (the angle at which the sound meets the seabed) increases. The first part of the echo to be picked up by the transducer is the reflection from the sea floor directly under the fish and time taken for the rest of the echo to return to the transducer is a measure of the distance from the fish. The surface unit displays the vessel's position centre-screen and the return signal to the right or left (starboard/port) at a distance from centre screen which is in proportion to the time taken for the return signal to be picked up by the transducer. The trace on the screen (or paper trace) advances with each new pulse until an image of the sea floor is built up line by line.

The intensity of the echo is also measured so that an image can be produced that looks remarkably like a strongly side-illuminated black and white photograph. The intensity of the echo depends on reflectance which in turn depends upon topography. Surfaces angled towards the fish produces a strong echo whilst surface hidden from the 'line of sight' of the fish results in a sound shadow. Intensity is also dependant upon absorption of sound on the sea floor, so that some idea of sea floor sediments can be visualised but it is not a straightforward measurement. Sidescan sonars are relatively cheap, have a good resolution and produce an image that covers a swath of up to 500m (or more with specialist sidescan systems). However, bathymetric data are limited to an estimate directly under the vessel and not for the full swath. The image, whilst giving excellent information on topographic features has little measurable point data on sediment characteristics. The images require careful interpretation in much the same way as a black and white photograph would. Although automated pattern recognition algorithms are used to analyse the image, analysis usually requires skilled interpretation by eye.

Research is being carried out into how the signal can be processed to give information on sediment characteristics (similar to AGDS) and this will enable detailed images of the sea floor with information on sediment properties as well as topographic information.

4.1.4 Swath multibeam systems

de Moustier (1988) gives a detailed description of swath systems. A series of sound beams are formed from a limited number of transducers in a fish or hull-mounted pod. These produce a series of pulses at known angles of elevation to port and starboard of the array and from this (and the time of the echo) the depth at the point where the sound meets the seabed can be calculated. A large swath of accurate bathymetric measurements can be made with a spatial resolution of about 10cm². Again, intensity can also be measured and the resulting image can look very like a sidescan image with the advantage of a continuous coverage of bathymetric data, albeit at a slightly reduced resolution. As with sidescan, swath multibeam systems produce little detailed analysis of reflectance properties at

any one point on the seabed, although research is being carried out in this field. Multibeam systems have been used to map the topography of very large areas of the seabed (Gardener *et al.* 1998). The system is very dependent upon boat movement and excellent pitch, roll and heave information is required. Multibeam systems are expensive but, especially with recent developments on signal processing, they promise to have great potential for biotope mapping (Scherzinger *et al.* 1997; Safer 1997). In particular, classification using pattern recognition methods might provide a way of describing features at different scales (Kavli *et al.* 1993).

4.2 Choice of systems

The choice of systems to use as the primary tool for remote sensing by the BMP was, of course, constrained by cost both in terms of capital outlay and staff time in processing data. However, these restrictions were not considered to be unrealistic since one of the primary aims of the project was to devise a methodology that could be adopted by many other organisations so that large areas of the seabed could be surveyed. Clearly, the use of very expensive and/or experimental technology would exclude many organisations from participating in such survey work.

AGDS was considered to be inexpensive, easy to operate and produce data in a format that could be processed easily. Of the various proprietary systems available *RoxAnn*TM had the advantage of being well established and is already installed on many survey vessels in the U.K. *QTC View*TM had potential advantages over *RoxAnn* in that the range of variables measured was greater. However, the system was somewhat experimental when the BMP project started and, although trialed thoroughly by *SeaMap* a system was not, in the end, purchased.

AGDS systems gave BMP the ability to cover large areas of sea, but at a very low resolution. Sidescan sonar, on the other hand, gives images of the seabed of very high resolution but with high survey costs for covering large sea areas. It was decided that this system would complement the AGDS and could be deployed either to cover small areas comprehensively or as a way of sampling representative areas of the AGDS map, much in the same way as a remote viewing device, but covering a greater area in its sample.

SeaMap operated a *Geo Acoustics*TM SS490 side scan sonar which can switch between 100kHz and 500kHz and was linked to an *EOSCAN* digital acquisition system (Polaris Imaging Inc) which provided full geo-referenced data capture and post-processing capability. *SeaMap* also used an *EOMAP* system (Polaris Imaging Inc) for combining individual sonar lines into a mosaic to create a map of the survey area.

Swath multibeam systems were beyond the budget of the project, although *SeaMap* staff attended demonstrations of various swath systems (enough to convince them of the potential and the need to ensure the BMP methodology was generic and could absorb these emerging systems).

4.3 An introduction to AGDS

How do acoustic ground discrimination systems work? What is the nature of the data they produce and what are the principles involved with the interpretation of these data? This Section gives an introductory account of AGDS, concentrating particularly on the *RoxAnn* signal processor. The following account is intended as a general introduction for the understanding of the way in which AGDS can be used for habitat and biotope mapping, stressing the potential sources of variability and error. It is not a technical account of the way in which echo sounders and acoustic ground discrimination systems work. The reader is referred to Mitson (1983) for more technical information on echo sounders and a technical description of *RoxAnn* is given by Chivers *et al.* (1990) and Collins and Voulgaris (1993). The account illustrates the principles underlying interpretation using the basic editing features to be found in the real-time logging and display software. The limitations for interpretation are explained and the foundation is laid for more powerful post-survey processing using image processing, geographic information systems and statistical software.

4.3.1 Echo sounders

4.3.1.1 Basic operation of echo sounders

Echo sounders consist of a shipborne surface unit that generates an electrical signal that is transmitted to a transducer mounted under a boat in contact with the water. The electrical energy is transformed into mechanical energy by the transducer that converts the electrical signal into a compression wave signal. The compression wave travels through the water and rebounds off the seabed and this echo is detected by the transducer which converts the mechanical signal back into an electrical signal that is transmitted back to the surface unit. The signal is amplified and displayed on a screen. The time delay between the signal being generated by the sounder and the detection of its echo is a function of depth. The time delay determines how far down the screen the echo is displayed and this vertical distance is calibrated as a depth measurement. See Urick (1983) and Mitson (1983) for a description of echo sounders and the theory of sound propagation in water.

Additionally, the intensity of the display of the echo is a function of echo strength. The echo will be stronger if it rebounds off a hard object than a soft object. However, since signal strength decreases with increased depth, most echo sounders apply a depth-compensating amplification to the echo before it is displayed.

Echo-sounders send a pulse of sound that has a particular frequency. The duration of the pulse and the frequency is fixed by the sounder: The most common frequencies used lie between 30kHz and 200kHz. Transponders shape the pulse of sound into an approximate cone directed towards the sea floor. The area **ensonified** (analogous to the term ‘illuminated’) by the echo sounder directly under the vessel depends upon the beam angle (angle of the apex of the cone of sound) and depth of the sea floor. For example, an echo sounder with a beam angle of 15° with a depth under the boat of 5m and 60m would ensonify an area with diameter of about 1.25m and 15m respectively.

Sound waves travelling down through the centre of this cone hit the seabed first and depth is measured from time taken for this returning sound energy to be detected by the transponder. The sound energy that spreads away from the centre of the cone produces a weaker echo. This wave energy takes slightly longer to reach the seabed because of the extra distance travelled, and this time lag increases the further away from vertical axis of the cone the sound travels. These weaker echoes are not used for depth measurement, but contain useful information of the nature of the sea floor and are used by AGDS.

The strength of the echo from a pulse of sound produced by a transponder and the way it decays with time produces a complex signal whose shape depends to a large extent on the nature of the sea floor. Hard, flat surfaces result in very strong echoes that decay rapidly, although there may be one or more multiple echoes as the returning sound energy bounces off the water surface and rebounds from the sea floor a second (or third) time. Hard, rough surfaces will produce a strong echo that decays more slowly since sound spreading some distance from the vertical may reflect off an inclined surface angled towards the vessel and its transponder. Soft surfaces (and this may include rock substrata that are acoustically softened by overgrowth of biota) result in a weak signal. AGDS work on the principle that particular sea floor types can be distinguished by their reflectance properties.

4.3.1.2 Differences between echo sounders

The output from echo sounders and their transponders differ quite markedly from each other and these characteristics will introduce variability between data collected by different systems. Clearly, beam width, signal frequency and pulse length are fixed with a particular echo sounder system.

Frequency: Echo sounders are commonly designed to operate at a particular frequency of between 150-200 kHz. This range is quite high and will respond to superficial sediments only since penetration through sediment is poor. Penetration increases as frequency decreases; Riddy and Mason (1996) give approximate penetration for a 200kHz echo sounder as 0-0.5m and a 35kHz sounder as 0-2m depending on the substratum. The potential resolution of objects on the seabed increases as frequency increases which is why high resolution sidescan sonar utilise high frequencies (e.g., 200-400kHz). However, beam width also affects resolution (see below).

Power: The signal attenuates with depth, AGDS operating in deep water will require a more powerful signal than for one operating in shallow water. Frequency may also be considered since low frequency sound attenuates less than high frequency sound and is more suitable for sounders operating in deep water. Signal strength varies between echo sounders of the same frequency. Many systems will automatically adjust the power output of the signal according to the depth range set and some systems may offer additional power output settings controllable by the operator.

Beam width and shape: Transponders also shape the sound in different ways and the description of the signal as a ‘cone’ of sound is an oversimplification since transponders will produce extensions of the cone in certain directions, termed ‘side lobes’. Some transponders accentuate the side lobes, particularly those designed as fish-finders. Every transponder will have its own characteristic ‘footprint’ on the seabed, even if they are nominally identical. Thus, it should be expected that the responses of two sounder systems to the same sea floor types would not be the same.

4.3.1.3 Echo sounder variability

Additionally, there is the possibility that the same system might have slightly different characteristics over time. This might arise when portable systems are set up on different vessels or also caused by sea conditions. It is possible that careful standardisation of the acoustic data could mitigate variation between different systems and between surveys using the same system conducted at different times.

The above account has stressed the variability that may be encountered between different echo sounder systems. However, it is important to appreciate that this variability between echo sounders does not necessarily mean that the *interpretations* of the acoustic data will be significantly different if, irrespective of the sounder system, the same sea floor types return distinctive echoes. However, it may be expected that some sounders are better suited to discriminate between sediments than others because of the shape of their footprint, signal strength and the frequency at which they operate.

4.3.2 How AGDS work

The way information is extracted from the echo is very different with each proprietary AGDS and is the subject of commercial confidentiality and copyright.

4.3.2.1 The single frequency *RoxAnn*TM AGDS

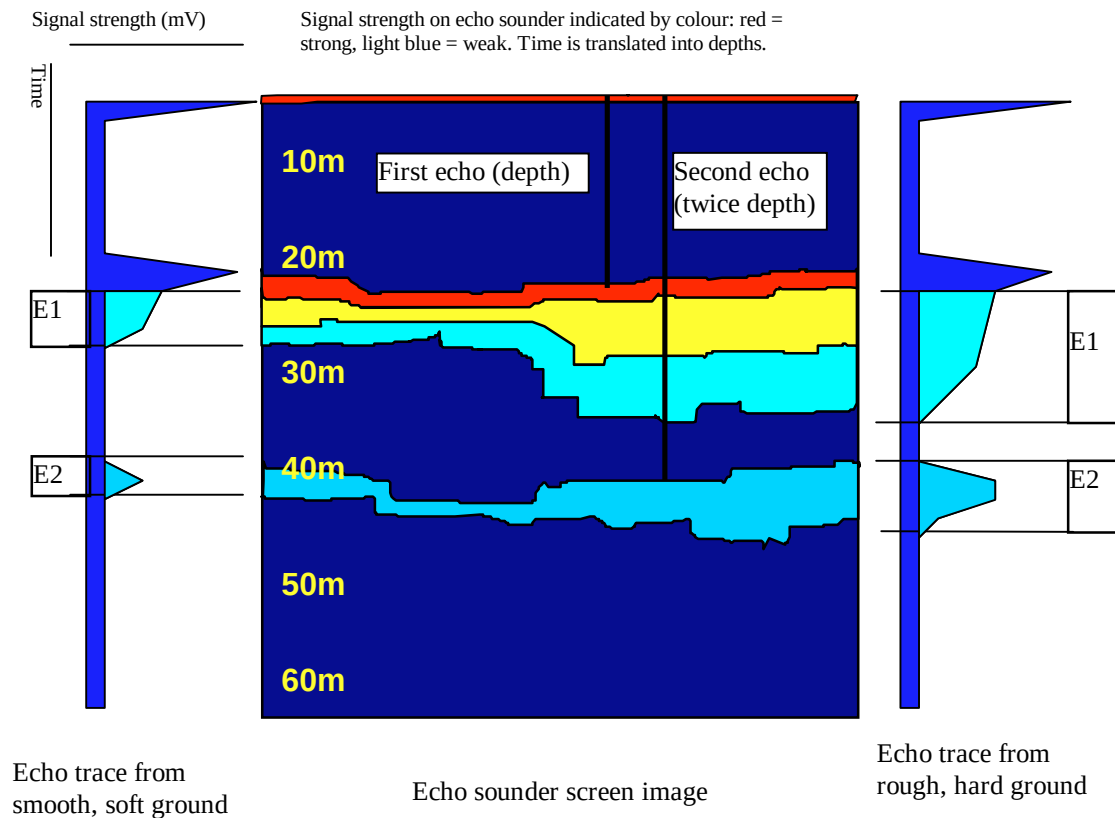
The *RoxAnn*TM system uses signal processing hardware to select two elements from the analogue echo and measure a value that is an integration of echo signal strength (in millivolts) and time. The first selected segment of the echo is the decaying echo after the initial peak in strength. This measure of time/strength of the decaying echo is termed ‘Echo 1’ (or ‘E1’) and is taken to be a measure of roughness of the ground. The beam width of the sounder is important for E1 since a wide beam will give greater scope for measuring signal decay away from the perpendicular than a narrow beam. For this reason it is recommended that AGDS operate with a sounder of moderate beam width (15° – 25°).

The second segment is the whole of the first multiple echo. Multiple echoes are produced when the first echo is sufficiently strong to rebound off the water/air interface and reflect for a second time from the seabed. This is taken by the designers of *RoxAnn* to be a more sensitive measure of ground hardness than the strength of the first return. It is termed ‘Echo 2’ (or ‘E2’). The sounder should emit a signal which is sufficiently powerful to create a second echo but not so powerful as to saturate the AGDS. There is a trade off in terms of maximum and minimum operating depths since a signal powerful enough to create a second echo (E2) in deep water may be too powerful for shallow water and saturate the return signal. Thus, it is important to ensure that the sounder is properly set up to operate within the required working depth range.

The two paired variables (E1 and E2) can be displayed quite simply on a Cartesian XY plot, and this is the basis of the *RoxAnn* real-time display as used in the data logging and display systems *Microplot*TM and *RoxMap*TM. Rectangular areas on the Cartesian plot can be marked out so that records lying within that section of the plot can be colour-coded and displayed on the track plot (see Figure 4.4). This forms the basis of real-time ground discrimination that will be discussed in the following Section.

BOX

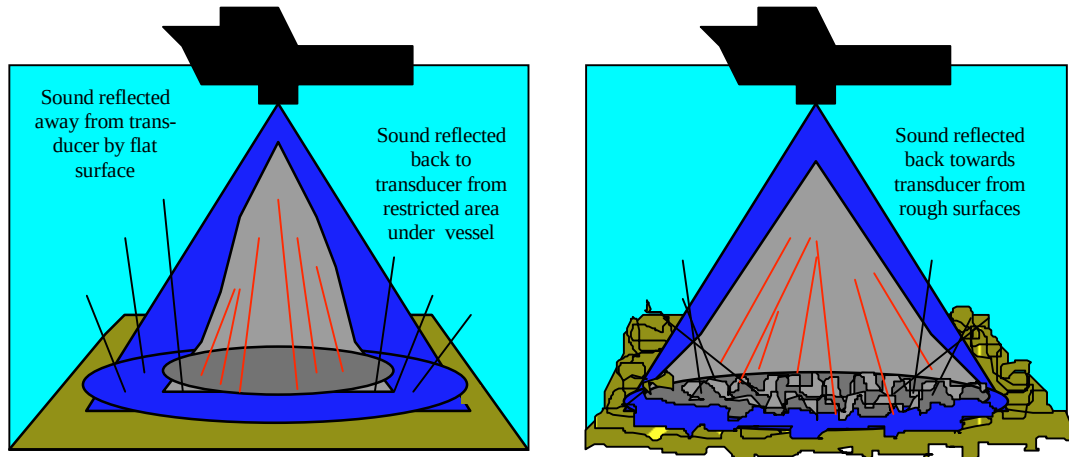
Figure 4.1. The relationship between an echo sounder screen display, the echo trace and the parameters measured by E1 and E2 by RoxAnn.



Echo sounders show characteristics of the return signal as a visual display from which qualitative information about ground hardness, roughness and other ground characteristics can be obtained. Acoustic ground discrimination systems analyse the signal and quantify selected parameters. Figure 4.1 above shows what an echo sounder screen might look like as a vessel tracks over smooth, soft ground on the left of the screen to rough, hard ground on the right. The echo traces shown either side of the central display illustrate a single pulse and its return signal corresponding to each of these two ground types. The strength of the peak of the first echo (not the same as E1; see below) depends on ground hardness. However, it is claimed that the strength of the second echo is more sensitive to hardness (Chivers *et al.* 1990). The second echo results from sound reflected from the seabed rebounding off the water/air interface and reflecting off the seabed for a second time. The RoxAnn signal processor integrates time and signal strength of the second echo (E2) to derive a value that is taken as a measure of hardness. Other AGDSs, however, do use the first echo to measure hardness properties of the sea floor.

Other parts of the signal more susceptible to scattering of the sound are used to derive a measure of roughness (Figure 4.2). The echo sounder emits a cone of sound (dark blue) which ensonifies an area of the seabed (equivalent to light from a torch illuminating a circle). The energy of the reflected sound that is picked up by the transducer and converted into electrical energy (grey/blue) comes mostly from the region that is directly underneath the vessel.

Figure 4.2. The behaviour of sound reflected from smooth and rough ground.



Sound diverging from the point on the sea floor directly under the vessel that would reflect away from the transducer on smooth ground may, on rough ground, find surfaces that bounce the sound back towards the transducer. The greater distance this sound travels compared with sound hitting the sea floor directly under the vessel is translated into a longer time interval. Thus rough ground extends the time base of the first echo and can increase the amount of energy it contains. *RoxAnn* measures the energy in the 'tail' of the first echo after the peak of the signal has passed and this is labelled E1.

4.3.3 The nature of AGDS track data

It follows from the above description of the way an AGDS works in conjunction with a single beam echo sounder that a set of measurements (E1, E2 and depth in the case of *RoxAnn*) is made for each echo. The E1 and E2 values are taken from an area of sea floor covered by the footprint of the signal, but the area will increase with depth. Not all measurements are logged, however. In the case of *Microplot*, if data are saved every 2 seconds, then average values for E1, E2 and depth are calculated over a 2 second interval. These average values are logged together with the ships' position at the moment when the values are saved. In the case of *RoxMap*, only the values from the echo preceding the save are logged. No average is calculated and the majority of the data are not saved nor do they contribute to the saved values. Thus, with *Microplot* the recorded values can be regarded as an average (and hence smoothed) value returned at the trailing edge of a variably shaped oblong of sea floor. With *RoxMap* the recorded values are not smoothed but there will be gaps on the sea floor between successive save points for which there are no data.

These differences between the two logging systems could result in greater variability in the data set recorded in *Roxmap* than for *Microplot*. Similarly, different save rates selected for *Microplot* would lead to different degrees of smoothing.

The data are, strictly speaking, in the form of single sets of values recorded from an area of sea floor of variable shape and size whose location corresponds with the ship's last calculated position (taken from the DGPS) prior to saving the data. The combined ellipse of DGPS error, footprint size and/or gap between saved data gives the maximum resolution of the AGDS in the direction of travel of the vessel. Thus, for a vessel working in about 10m of water at 10 km/h with a beam angle of 15° and a DGPS error of 10m, one might expect a set of E1, E2 and depth values saved every 2 seconds to represent an area of approximately 13m wide by 20m long. Any increase in depth, speed or save rate would lower the resolution further. However, for most purposes, the recorded values can be regarded

as centred on a series of points following the ship's course. It is not until the track data are displayed at a very fine scale that this approximation becomes apparent.

Acoustic data for a survey area are built up as the survey vessel tracks backwards and forwards across it. Since it is unlikely that parallel tracks will lie closer together than the maximum along-track resolution (15m in the example above), it follows that the between-track distance imposes a far greater limit to the resolution of an image than along-track resolution.

4.3.4 Simultaneous dual frequency AGDS

Two variables (three, if depth is included) would not seem to offer great potential for discrimination between sediment types. On the other hand, it might be said that the simplicity and robustness of the system compensate for the potentially low powers of discrimination. However, *SeaMap* have been experimenting with the simultaneous deployment of two systems working on two different frequencies to see if information content and powers of discrimination can be improved.

High frequency sounders respond to seabed surface sediment characteristics. Rock covered with a thin film of sediment, for example, might give a very different response than bare rock. Lower frequencies with their greater penetration might respond to coarser sediment (and bedrock) underlying a thin layer of surface sediment. Thus, high and low frequency systems might respond differently to sediments and combinations of sediments and, potentially, give complementary information. Studies on a limited range of seabed types in calibration sites suggested that a better classification was obtained using a low (either 38 or 18 kHz) and a high frequency (200 kHz) echo sounder than was obtained using just one frequency (Kavli *et al.* 1994). These workers did not, however, find any improvement in discrimination between sediments if all three frequencies were used.

It was on the strength of this potential increase in sediment discrimination that a simultaneous dual frequency *RoxAnn* system was purchased for the BMP. It was important that the two frequencies were recorded simultaneously to avoid positional errors that would render a pair-wise comparison between high and low frequency data problematic. The dual frequency system comprised of two *RoxAnn* units whose outputs were logged simultaneously. The outputs from the two systems could be displayed simultaneously (but separately) on Cartesian E1/E2 plots in real time. However, any combined analysis of the data was only possible using post-processing.

4.3.5 The *QTC View*TM AGDS

*QTC View*TM is an alternative proprietary AGDS that works in a very different manner to *RoxAnn*. The *QTC View* system was taken on trial in the Firth of Lorn and Pembroke and, although not purchased for the BMP, *SeaMap* have some knowledge of this system.

The echo is converted from analogue to digital form and is then subjected to analysis using a large number of algorithms for wave-form analysis. The *QTC* choice of algorithms and the way they are applied to the echo is considered commercially sensitive. However, the second echo (so important to the *RoxAnn* AGDS) is not used.

At the start of a survey, a selection of known ground types are chosen to calibrate the AGDS. The *QTC* system analyses the echo using its full complement of algorithms and then performs a principal component analysis (PCA) on the variables and selects those that contribute most to the discrimination between seabed types. These then form the variables which will be recorded for the rest of the survey. The principal components analysis is run in real time and the data displayed on a three-dimensional plot of the first three principal components, termed '*Q space*'. This is equivalent to the Cartesian plot of *RoxAnn* system except that many more than two variables are involved. Again, as with *RoxAnn*, three dimensional boxes can be marked out on the display and the records colour coded and displayed on the track plot.

The three principal component values are recorded together with a confidence measure (how well new records match to one or other of the original ground types). If the confidences are low, this suggests that another ground type might be present and that a new box might be required and the PCA run again.

Both *RoxAnn* and *QTC View* have been designed to be real-time AGDS and require calibration against known ground types. It is expected that the arrangement of the boxes are finalised (but can be refined) early in a survey and that all new data are categorised into one or other of these ground types according to where they are located on the Cartesian or *Q space* plots. The way this operates is explained for *RoxAnn* in the next Section.

However, this real-time classification is very limiting and *SeaMap* have explored ways of analysing the raw E1 and E2 values post-survey. Although it is not possible to utilise the raw variables in the standard *QTC* system, it is possible that the principal component values (eigenvalues) might fulfil the same function. More recent versions of *QTC View* do allow the export of the raw variables, but this system has not been tested in the BMP. *Simrad* also report on their development of single beam AGDS which appeared to be similar to the *QTC* system in that principal components analysis to reduce the volume of data extracted from the echo (Kavli *et al.* 1993; Kavli *et al.* 1994). This developmental system does not seem to have been marketed.

4.3.6 Equipment for acoustic ground discrimination

The equipment routinely used in the *SeaMap* project for acoustic survey was based on *RoxAnn* signal processors. Two *RoxAnn* systems were used and are referred to in this report: (a) a single frequency *RoxAnn*TM system operating at 200kHz with a Koden CVS-8112 echo sounder (referred to hereafter as the 'single AGDS') and (b) a dual frequency *RoxAnn*TM system operating at both 200kHz and 38kHz with a JMC V122 echo sounder (referred to hereafter as the 'dual AGDS'). Geographic co-ordinates were provided by a differential global positioning system (DGPS). The track data were collected and saved at time intervals which can be set by the operator (usually 2 or 4 seconds) on a laptop computer, which also supplied time and date for each point, utilising *Microplot* navigation software (Figure 4.3) for the single AGDS and *Roxplot* for the dual AGDS.

Both *Microplot* and *Roxplot* display track data (superimposed, if required, on a map or chart of the coast) as they are being collected. The track is coloured according to combinations of E1 and E2 or by depth. The combinations of E1 and E2 were displayed graphically on the screen and boxes drawn to encompass ranges of values as desired (see lower right of screen display, Figure 4.4).

Both the single and the dual systems were entirely self-contained and portable so that they could be set up on a wide variety of craft. The transducers were attached to scaffolding poles that could be rigged to the side of a boat.

In addition to the *RoxAnn* AGDS owned by *SeaMap*, other systems were also used as the opportunity arose. Thus, the unit fitted to the Eastern Sea Fisheries Joint Committee's research vessel *Surveyor* was used extensively in the Wash. The *R.V. Prince Madog* from the University College of North Wales and the North West Sea Fisheries Committee vessel *Aegis* are also each equipped with a *RoxAnn* units and data from these sources have been used for the survey in Wales.

Figure 4.3. Schematic diagram of the single frequency *RoxAnn* AGDS and survey equipment

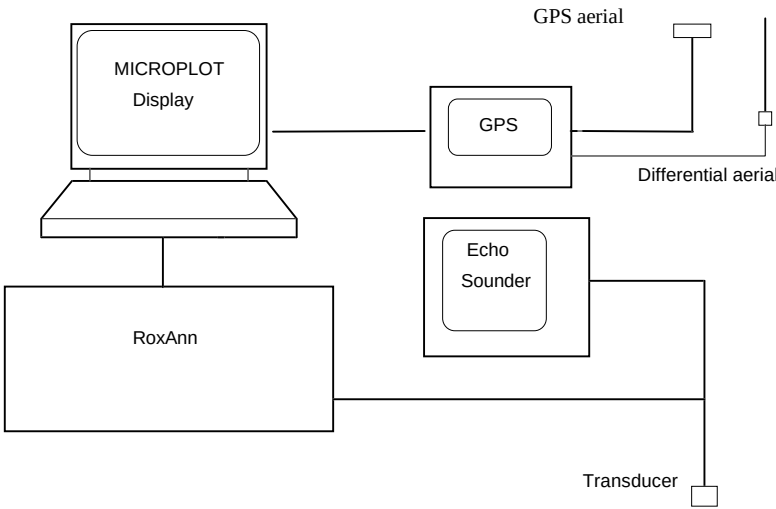
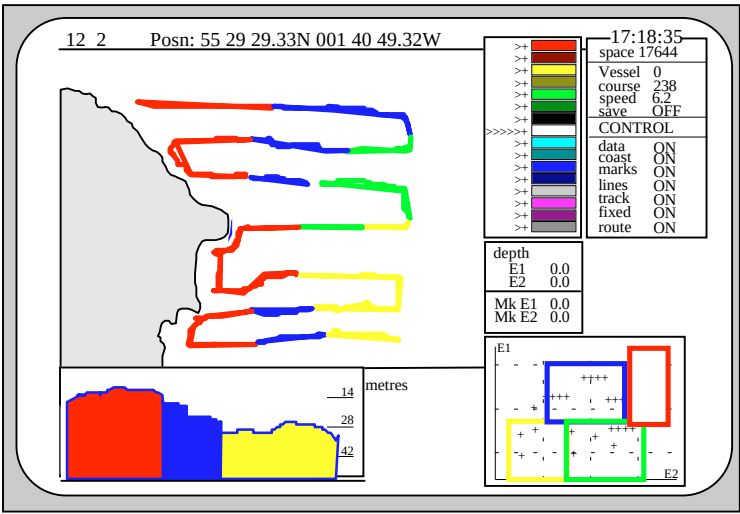


Figure 4.4. Screen display of track data in *Microplot*



5 An introduction to survey using acoustic ground discrimination systems (AGDS)

Purpose: To give the reader an overview of the methodology adopted by the BMP, the basic operation of the AGDS and the use of the data for the construction of a habitat or biotope map. The account stresses sources of error and uncertainty.

5.1 Acoustic survey and real-time AGDS

The basic principles of mapping using an AGDS can best be explained using the editing facilities within *Microplot*, which forms a very useful tool for real-time ground discrimination and was used extensively during field survey to visualise different ground types and select sites for ground truthing. However, it was little used for the more advanced data analysis for map production.

There are two basic approaches to the use of AGDS for real-time ground classification. The method intended by the designers of AGDS is classification through calibration. However, this assumes that the ground types in the survey areas are known and, at least in some representative plots, their spatial distribution has been established. The second method uses an iterative process of analysis of the acoustic data to detect clusters of data and identify what these clusters represent through ground sampling.

5.1.1 Ground discrimination through the use of calibration sites.

Areas of known seabed type are first located and described in detail using divers, remotely operated cameras or intensive grab sampling. Once a representative selection of seabed types has been identified, the acoustic characteristics of these ground types is measured and defined by upper and lower E1 and E2 values (i.e., the box dimensions on the *Microplot* display). These boxes, once defined, are then used to classify all further AGDS data. There is an assumption that the arrangement of the boxes can be exported from one site to another. Indeed, the basic arrangement of the boxes is regarded by Chivers *et al.* (1990) to have a standard configuration with regard to sediment type and that local calibration can be achieved by setting the maximum and the minimum values for the area.

This approach, whilst perhaps suitable if the use of the system is to be restricted to one area (the operators' 'home patch'), is unlikely to be the most appropriate use for very broad surveys and when the system is used in many different locations where the range of ground types is unknown and the response of the system is likely to vary with subtle changes in seabed characteristics throughout the region.

5.1.2 Ground discrimination through iterative AGDS survey and ground truth sampling.

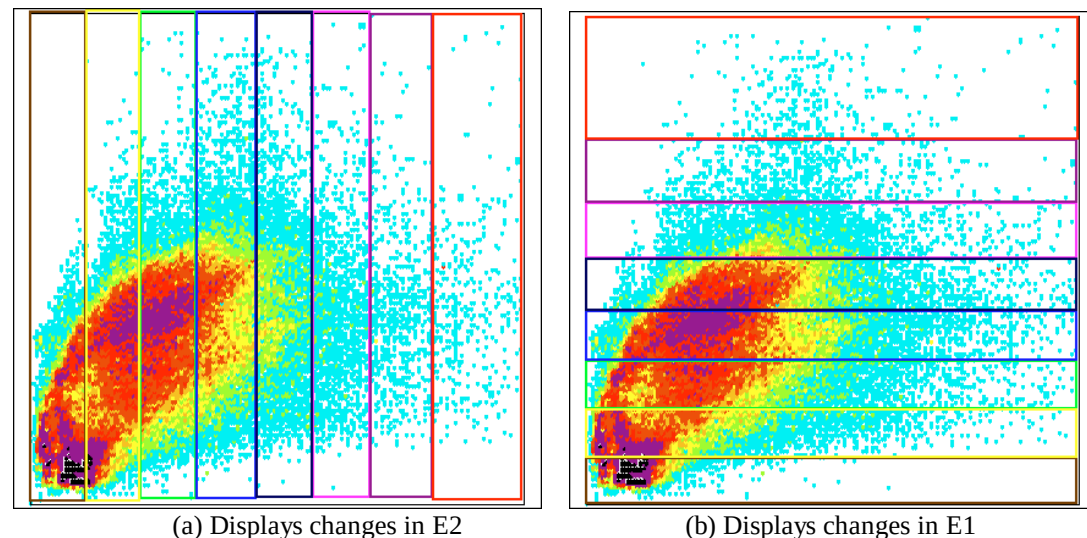
An iterative process can be used for survey based on acoustic remote sensing where initial characteristics of E1 and E2 suggest suitable locations for ground truthing. The information ground truthing gives regarding habitats and biotopes are used to refine the initial *Microplot* display boxes to better reflect the different ground types found. Further ground truth sites are then selected to test the adjustments made to the boxes. A good intuitive feel is gained for the acoustic characteristics of the major habitat and biotope types by the end of a field survey. Subsequent post-survey analysis tests these ideas more rigorously.

The following account is somewhat idealised since time constraints on a survey usually mean that some of the stages might be compromised, particularly through the lack of opportunity to revisit locations during the survey due to logistical constraints.

Stage 1. Preliminary acoustic survey:

The survey commences with the systematic collection of acoustic data over one section of the survey area. The initial layout of the boxes in the real-time *RoxAnn* display is designed to show, separately, small increments in the values of E1 (roughness), E2 (hardness) or depth. The track data can be edited and re-displayed showing each parameter in turn during the field survey so that a first impression of the distribution of acoustic characteristics can be gained. Figure 5.1 shows the arrangement of boxes that display E1 and E2. These boxes are stored in the library of the display software and can be used at any time.

Figure 5.1. Arrangement of boxes to show increments of E2 and E1 respectively. The boxes are shown superimposed on a scatterplot of a complete data set which is coloured according to point density.



Stage 2. Range and distribution of acoustic characteristic:

After the conclusion of the first stage of field survey, three hard copy print outs are made of the track data displaying E1, E2 and depth respectively. These maps can be printed out to the same scale on transparent film so that they can be superimposed (a process called **acetate overlay**). A composite map of E1/E2 is produced indicating areas with similar acoustic characteristics. Depth remains an acetate overlay. This process is easier if contour lines enclosing equal-value points are drawn onto each of the acetates. Contouring can be done adequately by eye at this stage.

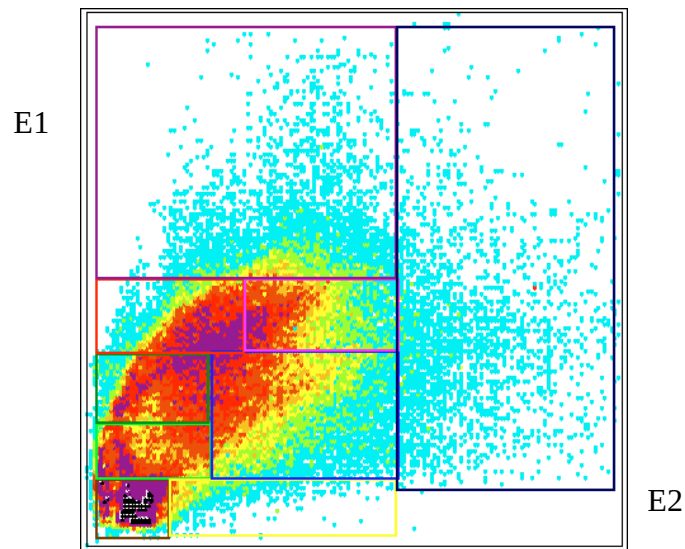
Since there are no data about habitats or biotopes at this stage, the discrimination between ground types can only be made on acoustic characteristics. The resulting distribution map, therefore, shows **acoustic ground types**.

Boxes can now be drawn in the display system for E1/E2 combinations that are commonly found on the composite map. The track can be re-displayed according to this new arrangement of the boxes (Figure 5.2) to test if the map shows acoustically homogeneous regions. These boxes then represent the acoustic ground types to be sampled in the next stage of the survey.

Stage 3. Ground truthing acoustic data:

Locations representing the different acoustic ground types are sampled. Careful selection of ground-truth stations within an area of uniform acoustic ground type increases the confidence with which acoustic values can be attributed to the ground truth data. The boxes can be expanded, contracted, subdivided or amalgamated as experience from ground truth sampling more clearly defines the habitat or biotope that the boxes represent. This editing can be done during the survey so that the operator can test and modify ideas as the survey progresses.

Figure 5.2. Initial arrangement of E1/E2 boxes in *Microplot* that correspond to acoustic ground types.

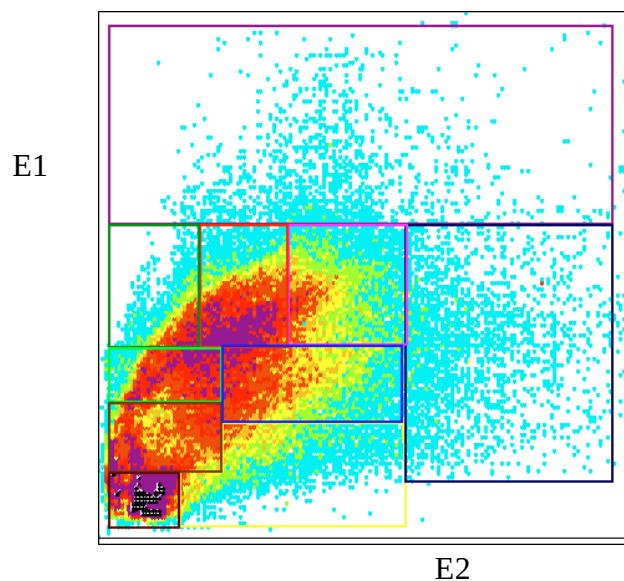


The aim is to be able to predict which habitat or biotope is most likely to be found at a location from the acoustic characteristics. This process of box modification (Figure 5.3) to predict a biotope followed by ground truthing is repeated through the survey.

Stage 4. Completing the field survey:

Other parts of the survey area can now be used to test and refine the relationship between acoustic ground characteristics and habitats or biotopes. Ground truth locations will be selected either to test predictions or to investigate new acoustic ground types. Predictive capability will vary depending upon the ground type. For example, hard/rough ground in shallow water might be found to support a kelp biotope in all of 5 ground truth samples.

Figure 5.3. Modified arrangement of boxes within *Microplot*



Areas of similar acoustic characteristics would require little or no extra ground truthing to predict the presence of kelp. On the other hand, moderately hard and rough ground in deep water might support brittle star beds or a dense turf of the soft coral *Alcyonium digitatum*. It would be more worthwhile to

direct sampling effort towards the latter ground than the former. The surveyor should aim to establish if acoustic characteristics can be refined to discriminate between these two biotopes or (if the two biotopes share the same box) to discover if they can be separated geographically (i.e., there is a broad scale trend in their distribution).

This process works well for small areas, but overlaps in the acoustic characteristics of different biotopes become more common as the survey area and range of habitat or biotope type increases. This is to be expected since two dimensions (E1/E2), together with depth, do not provide much scope for the discrimination of a large range of habitat or biotope types. It should be expected, considering the importance of depth in structuring benthic communities, that bathymetric data would need to be taken into account when mapping biotopes, but not necessarily habitat type.

5.2 Generating a continuous coverage from track data

5.2.1 Why generate a continuous coverage?

Satellite or aerial remote images are comprehensive in that there are directly measured data for each pixel of the image and the pixels are contiguous. The AGDS track data, in contrast, do not constitute a comprehensive coverage since the echo sounder footprints from which acoustic data are derived will rarely be contiguous between tracks: there will be large gaps over the survey area for which there are no remotely sensed data.

Whilst it is possible to interpret track point data and display the results on a map, it is far easier to see spatial patterns in data if these are displayed as a continuous picture. This is particularly true for broad scale survey where a user-friendly, general account of the distribution of habitats and biotopes is required. It must be stressed that producing a continuous coverage from point data does not improve the accuracy of the data. The track data cannot be improved upon in this way. This may seem an obvious point, but one that can cause confusion. Indeed it should be appreciated that much of the estimated data introduced through point-to-area transformation may be of dubious validity whilst important but small features seen on the track data may be lost. This begs the question as to whether a continuous surface is required for all types of survey. For example, monitoring surveys may be better served through an interpretation of the track data rather than continuous coverages generated from the original data.

However, there are more profound reasons for transforming point data to a continuous coverage. The spatial nature of the point data must be modelled to produce a satisfactory continuous coverage, and this exercise is useful in exploring and understanding the processes that determine habitat and biotope distribution. This approach ‘pushes’ the interpretation of the data towards prediction and hypothesis-testing and is the underlying philosophy of using maps as a predictive tool that the BMP have adopted. It should be said that this approach might not be suitable for all purposes, such as monitoring studies of small areas where precision and accuracy are required.

5.2.2 Introduction to point-to-area conversion

The ground between tracks can be filled with what appears to be the most appropriate acoustic value by eye. In fact, the human eye can do this quite successfully with a single parameter and can apply subtle changes to the data using intuitive ‘rules’ although this approach suffers from a lack of rigour and repeatability. Burroughs and McDonnell (1998) review the process of generating an areal coverage from point data. The areal coverage is generated through a mathematical process termed ‘**interpolation**’ in which values are predicted at locations for which there are no direct measurements using known values at multiple locations around the unknown value. In essence, interpolation places a grid of ‘n’ rows and ‘m’ columns over the track point data and calculates a value for each grid node. Since the real data are replaced by the new grid values, it is clear that this step is crucial in the use of AGDS data for biotope mapping.

Many **geostatistical** and mapping software packages offer a range of different interpolation methods to choose from (e.g., *Surfer*TM and *Vertical Mapper*TM). In addition, each interpolation method requires a number of parameters to be set by the investigator. The picture to emerge from the interpolation process could potentially be very dependent upon the choices made by the investigator. However, most interpolation procedures will calculate values that are faithful to real data if these are spatially close to an estimated value. Thus, unless there are large gaps in the track data, interpolation is a robust procedure and unlikely to be unduly sensitive to the interpolation method used. Unfortunately, large gaps inevitably do occur in survey data. In such cases, the emphasis of interpolation changes from a process of filling in missing data between tracks with 'obvious' values to one of mathematical modelling in which values are predicted. These predictions should, ideally, be tested by further survey.

In order to decide how best to proceed with interpolation in situations where the track spacing is other than close, some knowledge of the spatial correlation in the acoustic data set is required. The primary method for investigating spatial correlation is the use of the **variogram** (Burroughs and McDonnell 1998). The reasoning underlying the use and production of the variogram is as follows: It is a basic assumption of interpolation that the values of points that lie close together are more likely to be similar than points which lie further apart. The correspondence decreases with increasing distance until a distance is reached where the values of two points have no systematic relationship to each other that is dependant upon the distance that separates them. The way in which the relationship varies with distance, and in particular the range of the relationship, is important to interpolation. If tracks lie further apart than twice this range, then it is clear that there will be ground between the tracks where values cannot be estimated from the spatial patterns in the data of the tracks that lie either side. The simple average of all points in a locality (i.e., without any form of distance-weighting) is the best estimate in such cases. Indeed, since interpolation procedures can calculate values way beyond the range of spatial correlation if a large search radius is chosen, it is important to realise that estimated values might simply be the local average of the data 'found' within the search radius. The variogram is explained more fully in Section 11.

As track spacing increases, interpolation introduces an increasing degree of uncertainty and the resulting biotope maps will prove to be increasingly inaccurate. Such inaccuracies, whilst acceptable to scientists wishing to generate and test hypotheses, may be unacceptably high for a useful biotope map. At what track spacing do the inaccuracies inherent in interpolation become unacceptably large? A value for inter-track distance cannot be universally applied since many factors must be considered. For example, homogeneous areas require less intensive tracking than heterogeneous areas. If biotope maps are only acceptable if all patches of biotopes above a certain minimum size are plotted, then the size of these features and their frequency of occurrence will determine inter-track distance. In other words, scale must be considered. Clearly, this subject is of importance to survey design and has implications for survey effort and costs (see Section 15).

5.3 Interpreting AGDS data

5.3.1 Habitat and biotope identification and classification

A grid of interpolated values for the variables measured by *RoxAnn* (E1, E2 and depth or other variables derived from the data) can be treated as a digital image. This opens up the possibility of using powerful image processing software designed primarily for use with **multispectral** satellite or airborne remote images and modelling within geographic information systems (GIS). However, as with the use of any automated procedures, there is the potential to create an interpretation that is sensitive to the choices made by the investigator (e.g., choice of interpolation algorithm and the various parameters to be set in the calculation of estimated values). Interpretation of the digital images, no matter how sophisticated the process used, must be matched against what might be expected from expert biological opinion. The following is a general discussion of the principles of image processing here, although this topic will be covered in more detail in Part 2. See also Cracknell and Hayes (1993) and Wilkie and Finn (1996) for general accounts of image processing.

Classification of the acoustic data assigns a pixel to a category, such as a habitat or biotope class, on the basis of its spectral values (in the case of AGDS this is equivalent to E1, E2 and depth). There are two approaches to classification that differ primarily in the order in which the acoustic data and the ground-truth data are used to classify the acoustic images.

Unsupervised classification: The acoustic data are examined to detect discrete clusters of values when plotted in an E1/E2/depth co-ordinate system and then each of these clusters is assigned as unambiguously as possible to a habitat or biotope type. This is essentially the procedure used when using the editing facilities in *Microplot* to construct boxes on the Cartesian plot of E1 and E2. Modification of the shape of the boxes is an attempt to identify these clusters with habitat or biotope type more precisely. Re-displaying the track data in geographic space according to the modified boxes tests how well the classification has performed in describing the spatial distribution of habitats or biotopes. If the clusters are discrete, then the map and the ground truth sample data will match well.

Supervised classification: An alternative approach is to select **training sites** for each habitat or biotope class and to derive a template from the E1, E2 and depth values within the training area. This is termed the acoustic **signature** of the category. The track data (or pixels in a digital image) are then compared to these signatures and every pixel is given a likelihood value representing the degree of its membership to each of the habitat or biotope categories. The likelihood values are then compared in an automated process called **maximum likelihood classification** and each pixel is assigned to the category which has the highest probability and, therefore, to which the pixel is most likely to belong.

Whilst it might seem that the two processes would arrive at similar interpretations, the differences between them are more profound than might at first appear. Unsupervised classification makes the assumption that clusters of data occur within the acoustic data set. However, do such clusters exist? Can they be said to be discrete? Whilst clusters may be apparent when only a few, distinct habitat or biotope classes are involved, the picture is less clear-cut as the number of categories increases. In most BMP surveys a moderately wide range of habitats or biotopes were encountered. Scattergrams of E1/E2 from these surveys indicate that the swarm of data points is continuous with no clear dividing lines between clusters, even if there are areas within the plot where the density of records are higher than other areas (see Figure 11.1).

More fundamentally a large degree of overlap might be expected between the acoustic signatures of different habitat or biotope classes. This is because the classes themselves are not clearly defined but merge into each other. Thus, it is not realistic to draw boundaries between the acoustic characteristics of the classes. Overlapping signatures are, however, allowed in the development and application of acoustic signatures in supervised classification.

For these reasons, unsupervised classification is undertaken most often as a guide to the collection of ground truth data. Once this information is available, supervised classification is the more usual approach since there are stronger theoretical grounds for its use and the outputs offer greater scope for subsequent analysis and modelling (Wilkie and Finn 1996).

5.3.2 Supervised classification of acoustic data

Interpretation of a digital image is a three-stage process of (a) deriving an acoustic profile or 'signature' for each biotope type, (b) using these signatures to calculate the likelihood of pixel membership to habitat or biotope classes and (c) assigning each pixel to the category to which it is most likely to belong. *SeaMap* routinely use *Idrisi* (Eastman 1997) for image processing and classification.

5.3.2.1 Signature development

Signature development is crucial to data interpretation and it is important to appreciate why these signatures may not be sufficiently closely defined to be used for discrimination between habitat or biotope categories. The technical limitations of the system have been discussed in Section 4.3.

However, there are other sources of variability that are due to poor data collection which can degrade the quality of the signatures and result in poor image interpretation.

There are two main sources of error in signature development: (1) poor categorisation of the samples resulting in similar biotopes being inconsistently labelled and, (2) poor positioning of the samples so that acoustic data not properly associated with a biotope class are included in the calculation of a signature.

The descriptions of the categories should be distinctive and easily recognisable. Poor and inconsistent categorisation of the ground truth sites will undermine the most careful analysis of the AGDS data. Biotope description and identification is discussed in Section 6 of this report. Assuming that the sample data are of good quality, then the next stage is to extract acoustic values that are spatially associated with these sites so that an acoustic profile can be built up of each biotope or habitat type. This signature is constructed from the mean, maximum and minimum values and standard deviations for each of the variables taken from the subset of the data extracted around the ground truth samples.

However, selecting the AGDS data to associate with each biotope class depends, once again, on the analyst. There must be a minimum of data before a signature can be created. Using only those acoustic values nearest to a ground truth site may not result in sufficient data. It is reasonable, bearing in mind positional inaccuracies of the vessel and the sampling device, to accept all values within a small radius (or ‘buffer’) around the samples. However, the resulting acoustic signature may include pixels which are unlikely to be associated with a particular biotope or habitat (perhaps because the sample lies close to a boundary between two biotopes). This is exacerbated when (1) the ground is very heterogeneous over small areas; (2) different habitat or biotope types lie close together on the seabed (e.g., either side of a real biotope boundary); (3) samples are not located near any real acoustic data (so that uncertainty exists in the interpolated data), and; (4) buffers are too large. Editing signatures based on the knowledge of environmental limits to certain biotopes (e.g., depth limits to algal growth) can mitigate some of these problems.

5.3.2.2 Likelihood calculation

The values of the variables in each pixel are then compared with those of the signatures for each of the habitat or biotope categories and an estimate of the likelihood (expressed as probabilities) that the pixel belongs to each category calculated (for details of the calculation, see Eastman (1997)). Most pixels will have a probability of belonging to more than one category. Assuming that all possible categories have been described during the ground truth sampling program, then the probabilities for all the categories will add up to 1 for each pixel.

5.3.2.3 Maximum likelihood classification

The next stage in the classification process automates assigning each pixel to a single habitat or biotope class. This is done on the basis of the class with the highest likelihood membership. However, uncertainty of this classification will be low when a pixel has similar likelihood values for two or more habitat or biotope categories.

The use of likelihood values offers ways to (a) assess uncertainty and (b) show the probability of finding a particular biotope (even if not the first choice). These give valuable insights into habitat and biotope distribution patterns in their own right and will be discussed in Section 13.

Image processing and GIS software provide a range of options for both signature development and supervised classification and, since image processing is a rapidly developing field, choice will increase in future years. The analysis chosen for interpreting the acoustic data could potentially have far reaching consequences for the outcome of the biotope map. However, as with interpolation, variability between treatments need not be as great as the above account suggests if there are sufficient good quality ground truth data and the biotope classes chosen for interpretation are not too similar in terms of their acoustic characteristics.

6 Mapping units

Purpose: To discuss the practical application of the biotope concept to broad scale mapping. The issues addressed are the ability to recognise biotopes using remote viewing and sampling techniques; the ability to discriminate between biotopes acoustically; scale and biotope detail.

6.1 Introduction

Broad scale maps show the distribution of categories of habitats or biotopes. That is to say, habitat or biotope type is either identified and recorded in the field or, more usually, sufficient information is recorded to allow a category to be assigned to each ground truth record at some later date. This process is critical to the interpretation of the acoustic data.

Analysis of the ground truth data can be undertaken to devise a list of categories relevant to that data set. However, if mapping is to be undertaken on a national level, then ground truth data will need to be compared to a national classification. Analysis of the ground truth data then has the aim of finding a match between a sample and a biotope or habitat category taken from a list provided to the surveyors. Local variations or new biotopes might be described based on the degree of divergence between the ground truth samples and descriptions in the national classification scheme.

This matching process is difficult enough when the same recording techniques are used on survey as were employed in the original data gathering exercise for the construction of the classification scheme. However, the process is even more problematic if different recording techniques are used.

6.2 Biotope mapping units suitable for broad scale survey

Broad scale survey techniques place particular technical demands on any mapping units. Partly this arises because the units must be approximately the same size as the maximum resolution of the acoustic system. In other words, the sampling unit should give a good idea of the area of the seabed covered by the footprint of the AGDS. However, the maximum resolution of the AGDS is not the same as the maximum resolution of the interpolated acoustic values or the interpreted digital image. Also, the remote sampling techniques of choice will be those that are rapid and inexpensive. These might not give the level of detail needed for the positive identification of habitat or biotope category at the finest level of detail.

Thus, the technical limitations of the remote sensing/remote sampling system may impose quite severe limitations on (1) the spatial representation of biotopes or habitat types and also (2) the level of detail of their description. The first limitation requires conventions about the representation of the minimum mappable area by the predominant (or most likely) category whilst the second may mean that biotopes will need to be amalgamated into more broadly defined categories. These limitations need not necessarily be viewed as detracting from the value of broad scale mapping since the interpretation of the biology of an area may be better served by using a more general, broad brush approach to mapping.

6.2.1 The biotope as a mapping unit

Biotopes are loosely defined in CORINE documentation as the flora and fauna interacting with each other and with abiotic factors (Devillers *et al.* 1991). The biological communities are described by their species composition and characterised by their predominant, conspicuous or 'key' species. Likewise, habitats are characterised by the predominant features and further qualified by other 'minor' features. Biotopes, therefore, provide a summary description of the biological community together with physical features likely to influence the distribution of community types. Marine biotopes for the north-east Atlantic have been described and classified as part of the Marine Nature Conservation Review (MNCR) by Connor *et al.* (1997).

Although no definitive size has been ascribed to biotopes in the Corine classification system (Devillers *et al.* 1991) it is useful to have some guide for the development of subtidal broad scale survey and mapping methodology. A minimum biotope size has been given as 5m x 5m (Foster-Smith 1995; Bunker and Foster-Smith 1996; Connor *et al.* 1997). This is convenient for remote viewing since it equates approximately to a single view from a video and is also approximately equivalent to the area that might be ensonified by an echo sounder.

6.2.2 Minimum size of map units

A minimum size for a mapping unit is taken to be approximately 25m x 25m. There are technical and cartographic reasons for the choice of this size: Firstly, it is an area that represents the maximum resolution of the acoustic remote sensing system, bearing in mind combined positional error. Secondly, it is probably the smallest area that can be represented on a broad scale map (giving a square of 2.5mm x 2.5mm on a map drawn at a scale of 1:10,000). Most sublittoral biotope maps will be drawn to a smaller scale (1:50,000 - 1:250,000) and the minimum area that can be represented would be correspondingly larger.

6.2.3 Acoustic discrimination and detail of biotope mapping unit

A more fundamental issue regarding mapping based on remote sensing is ensuring that the mapping unit matches the discrimination capabilities of acoustic remote sensing. It might be possible to recognise many biotopes down to quite a detailed level. For example, *Laminaria hyperborea* kelp forest with dense foliose red algae (MNCR biotope code LhypR.Ft: Connor *et al.* 1997) might be identified from the ground truth sample record although the acoustic signature may be indistinguishable from other kelp biotopes. In a situation where ground truthing showed a variety of kelp biotopes to be found within the survey area it might not be possible to predict which one would be found at a particular location on the basis of acoustic signatures. In this situation it is not possible to use the acoustic properties to draw boundaries around biotopes. The greatest level of detail that could be mapped is, in this example, 'kelp'. This suggests that categories that are more broadly defined than 'biotope' might be useful for broad scale mapping.

6.2.4 Arranging biotopes into broad categories for mapping

The problems associated with biotope size and map scale may be surmounted by the use of broader biotope categories although it must be stressed that detail and size do not necessarily go hand-in-hand. Thus, a biotope identified to a fine level of detail may nevertheless cover a very large area. Conversely, very broadly defined biotopes might only occupy a small area. (The issue of biotope size and patchiness, and representation of biotope mixes on maps is discussed in Section 6.3)

Biotopes have been arranged into a hierarchical classification system to provide a consistent framework for the description of the biological character of an area and for the objective assessment of site quality and sensitivity necessary for the management of the marine environment (Connor *et al.* 1997). Higher biotope categories also can assist the interpretation of the biological character of an area by emphasising general trends in the distribution of biotopes. The nature of broad biotope categories will dictate what this emphasis will be. Thus, it is important to decide what aspects of the biotope will be selected for the derivation of broadly defined categories.

6.2.4.1 The biotope complexes and habitat complexes as a basis for broad scale mapping

The biotope is a hybrid unit incorporating very different types of information: biota and physical features. There are many possible arrangements of biotopes into a hierarchical (nested) system using either or both types of features. The system adopted for the MNCR classification uses both biota and physical habitat features, but the emphasis changes with the level of classification. The lower levels of the MNCR biotope classification system (biotopes and sub-biotopes) are described primarily according to community (species) composition. These have been arranged into categories of similar biological character (biotope complexes) which are the broadest biologically defined biotope groupings. (Connor *et al.* 1997).

Above this level in the hierarchy there are two upper tiers comprising habitat complexes (defined variously on the basis of zonation, substratum and exposure to wave action) which are arranged into major habitats (e.g. infralittoral rock and sublittoral sediment). The habitat categories have been defined to reflect important biological break points between communities. For example, shallow infralittoral rock supports macroalgae whilst deeper circalittoral rock is characterised by predominantly faunal communities.

Biotope complexes and habitat complexes may provide the higher level units suitable for broad scale mapping. One of the advantages of stressing the relationship between biotopes and habitat parameters is for the prediction of biotopes based on physical features (Connor *et al.* 1997). It is important that a survey of biotopes in a previously unsurveyed area samples across the full range of possible biotopes. The sample programme is likely to be based on a desk-top study of geological and hydrological information. The MNCR classification system provides a biological justification for dividing up the physical environment into representative units that can be sampled for biotope composition.

However, there are disadvantages in the use of biotope complexes that are nested within biologically defined physical habitat categories for surveys based on remote sensing.

- There is necessarily some overlap between biotope complexes that straddle two habitat categories. For example, kelp biotopes span all categories from exposed infralittoral rock to sheltered infralittoral rock and the lack of clear acoustic and visual discrimination between the various kelp biotopes may lead to misclassification/misidentification. Again, using kelp as an example, it might not be possible to discriminate between the biotope complex 'kelp with faunal cushion and red seaweeds' from 'kelp and red seaweeds'. The former lies within the exposed infralittoral rock habitat complex, whilst the latter is included in the moderately exposed infralittoral rock habitat complex. Thus, the highest common category that unites the two biotope complexes would be sublittoral rock. Clearly, broad scale mapping can deliver more information than the biotope classification system permits to be shown.
- The effect of attributing a doubtful biotope tag to a sample becomes very important when colour is the main visual cue for map interpretation. For example, one surveyor might class a kelp biotope as belonging to an exposed infralittoral rock habitat complex (colour code red in the MNCR habitat manual) whilst another worker might assign it to a moderately exposed infralittoral rock habitat complex (colour code blue). If alternately both kelp biotopes were to be given the same colour (separated by different shades or hatch patterns) then any confusion between them would not be so visually conspicuous.

6.2.4.2 Life forms

An alternative arrangement of biotopes into categories described by their overall appearance has been used for broad scale biotope surveys (Bunker and Foster-Smith, 1997; Foster-Smith *et al.* 1997). The emphasis is placed where possible on the appearance of the biota and for this reason the categories have been termed **life forms**. This term has a long history: Remane, in the 1940's (quoted in Schafer (1972)) called species with similar life habits 'behavioural types' and, if species with similar habits and functions were also morphologically similar they could be categorised into 'life forms'. A similar definition has been used to define life forms for broad scale mapping survey:

Life forms are the conspicuous, morphological adaptations of organisms to their environment, particularly adaptations to trophic requirements, maintenance of their position in the physical milieu in which they live and avoidance of predation.

Life form terminology can be applied to whole communities by labelling them by the predominant, conspicuous or 'key' life form. This definition of life form communities is clearly very similar to that used for biotopes although there is an emphasis on the overall biological appearance of community. It should be stressed that life forms are simply a convenient arrangement of defined biotopes suitable for broad scale mapping. Thus, in the example given in Section 6.1.3 'Kelp' is a life form. A list of life forms used for broad scale mapping in the *BioMar* project and refined for use in the BMP is given in Appendix 1, which includes the relationship between the life forms and the MNCR biotope complexes.

The life form arrangement of biotopes is not a rigid classification system. Although biotopes and sub-biotopes are attributed uniquely to one life form (and thus coded by one colour), there is scope for intermediate levels of life forms between the biotope and the major life form. These intermediate levels, whilst maintaining a standard colour, may be distinguished by different shades and hatch patterns chosen by the designer of the map.

This flexibility is required to take account of different survey objectives, map scales and survey conditions. The function of the life form system is to draw together biotopes which are most similar in terms of appearance and acoustic characteristics. These two properties do not go hand-in-hand and the ability to view biotopes and discriminate between them acoustically will vary depending upon survey design, techniques used and the nature of the survey area. For example, it may be possible in an open coast situation to discriminate between (a) sublittoral fringe kelp, (b) shallow water dense kelp forest and (c) deeper water sparse kelp park. These (a, b and c) could form useful subsidiary life forms for 'kelp', although not corresponding with biotope complexes. However, such an arrangement might not work in a sheltered sea loch where (d) kelp on bedrock can be discriminated from (e) silty kelp communities on mixed substrata. These latter groups (d and e) could form an alternative subsidiary life form arrangement.

6.2.4.3 Cartography and biotope mapping units.

The mapping units used in the BMP are (a) the biotope and (b) the life form. The biotope complex is not used although in many cases there is a good correspondence between life form and biotope complex (see Appendix 1).

The use of colours for mapping is extremely important and colour coding biotopes by their life forms means that maps show patterns in the distribution of biological communities. The clarity with which these broad trends will be shown on a map will, in turn, depend on the colours and symbols selected to represent biotope categories. **It is fundamental that maps convey visually (by the careful use of colour, shading, hatching, line styles and symbols) an overview of biotope distribution.**

Whether biotopes or life forms are used for mapping, the constant and unifying feature is the use of the life form colour scheme as set out in Appendix 1. The shades and hatch patterns for a biotope might change from one set of maps to another depending upon cartographic requirements, but the main colour representing the life forms will remain constant.

Because the life form system is poorly developed for sediment habitats, the primary feature that is amenable to broad scale mapping is sediment type. Thus, the life form colour coding system is based on the colours used in the MNCR classification system, although utilising stronger shades for cartographic reason.

Note that the life form colour scheme has been used in all biotope maps in the BMP. The term 'biotope' is used even though they are represented on the map coloured according to the life form colour scheme. The term 'life form' is used to denote broader biotope groupings.

6.3 Substrata

Mapping substrata is an important stage in the interpretation of the acoustic data and forms one of the outputs from remote survey. There is a well-established procedure for the analysis of sediments based on the Folks triangle (Buchanan 1984) and a standard colour scheme to represent these sediments on maps is used by the British Geological Survey (BGS). However, sampling and sediment analytical techniques suited for soft sediments are not applicable to seabed habitats with a substantial rock component. In these situations sampling is better carried out by divers and remote video and the habitat described by a visual assessment.

Likewise the BGS colour scheme is inadequate for the representation of rocky seabeds and a modified form of the Folks triangle has been devised to include larger rock elements such as cobbles, boulders and bedrock (see Appendix 2). The inner triangle is the standard Folks diagram but there are three additional triangles representing cobble, boulder and bedrock. Each of these is larger than the preceding triangle but underlying it so that the four triangles form a nested series.

The utilisation of this extended Folks triangle is as follows:- The first stage is to describe the fine sediments. If sediment samples are available or the visual records are sufficiently detailed, then these are given a position in the inner standard Folks triangle. If larger rock elements are present, then first the appropriate outer triangle representing cobbles, boulders or bedrock is selected and secondly the sub-unit in this triangle that represents the combination of fine and rock sediment is located.

The modified Folks triangle cannot be used in a quantitative way, as can the standard Folks triangle, but is used to select a structured description that is felt to best describe the sediment.

Video and grab sampling are, however, quite different and can result in very different impressions of the nature of the sea floor, particularly with regard to the rock elements. This creates difficulties when using the samples for ground truthing (i.e., the creation of acoustic signatures for interpretation of the AGDS data). The samples used for the development of signatures must be carefully selected from the complete data set using the following rules:-

- Where two types of samples are taken at the same station, the data are checked to see if they are compatible. If they are, then the modified Folks triangle is used as the common denominator between the samples. If there is obvious conflict in the description of the small particle fraction, then neither sample is used in the development of the signatures.
- Samples are also rejected if two samples are taken within a distance of about 25m of each other (in other words, where positional error may result in the two samples actually being at the same location – see Section 15.3.3) and are not of the same category.

6.4 Limitations to habitat or biotope discrimination using acoustic data

Despite the use of higher habitat or biotope categories (whether biotope complexes, life forms or habitat complexes) there will be many problems encountered when trying to assign a category to samples. Difficulties in deciding how to tag ground truth records stem from:-

- inherent patchiness of the ground;
- the occurrence of intermediate biotopes between two or more biotope types;
- biotopes forming a series of nodes with overlapping characteristics. These issues can be overcome to some extent by applying certain conventions when tagging records and displaying biotopes on a map.

Small areas may contain more than one biotope and resorting to broad biotope categories may not provide the answer since very different biotope types might be involved. This is one of the hardest issues to resolve in broad scale biotope mapping. The expedient of labelling the sample with the predominant biotope may be entirely misleading since there is no way of knowing what the acoustic response of the mixture will be nor how this relates to the acoustic properties of the components.

6.4.1 Scale and heterogeneity

Figure 6.1 illustrates how these scales relate to ground truthing and tracking with an acoustic ground discrimination system and Table 6.1 summarises the conventions that might be used when mapping biotope heterogeneity.

6.4.1.1 Heterogeneity at the biotope scale (large scale)

The finest scale (Figure 6.1a) corresponds approximately to the field of view of a video and the area of sea floor ensonified by an echo sounder. The acoustic reflectance values are an average over the mixture of the components present and an acoustic signature can separately identify none of the components. A mixed biotope at this scale can be treated like a single biotope for mapping purposes and given a new label to denote the particular mixture. The mixed biotope may give a distinctive signature or a very variable signature. These samples should be avoided for the purposes of signature development (unless the mixture is very commonly found).

6.4.1.2 Heterogeneity at the video tow scale (intermediate scale)

The next, intermediate scale, relates to the dimensions of the minimum area that can be represented on a map. Two samples may be located close to each other or different biotopes are recorded at discrete intervals from a video tow (Figure 6.1b). The biotopes cannot be given separate positions on the acoustic map with confidence. In other words, they are within the same minimal mapable area.

Each biotope forms patches that produce an echo that might be distinctive. However, the patches are small and the track will be composed of a sequence of points each with different values for E1 and E2 from the previous record.

There will be many instances where the acoustic footprint will straddle two biotopes and return intermediate values for E1 and E2. For this reason it would be dangerous to rely on applying signatures developed on homogeneous seabeds to situations where the biotopes are very mixed at this scale. One solution is, therefore, to map mixtures. At this scale it is quite likely that the biotopes are related to each other through their membership to a physiographic feature. For example, a shallow water reef might encompass kelp forests, faunal turf on reef faces, faunal crusts in surge gullies, scoured urchin grazed rock and barren mobile sand. As another example, 'algal turf on cobble plains' might be an acceptable way of referring in short-hand to extensive areas of cobble and shell with patches of sand and bedrock outcrops in shallow water that are predominantly covered by red foliose algae. The mixtures can, therefore, form an association that can be given an appropriate label.

6.4.1.3 Heterogeneity of AGDS patches (small scale)

The smallest scale in the example equates to an area which is amenable to tracking and ground truthing at selected sites (Figure 6.1c). Short lengths of track might be quite distinctive, but the patchy ground results in a poor correspondence between two parallel tracks. Interpolation between tracks will not be successful, resulting in small, apparently disconnected patches which may in fact be linked. More intensive tracking might resolve the finer structure of the distribution of the ground types, but at a substantial cost to the survey. However, is it necessary to know the distribution of all the patches? If not, one solution would be to treat the patches as an aggregation of biotopes. The patches as interpolated could then be regarded as indicative of the nature of the patchiness of the ground without being definitive as to their location.

6.4.2 Distinct habitats or biotopes with overlapping acoustic characteristics

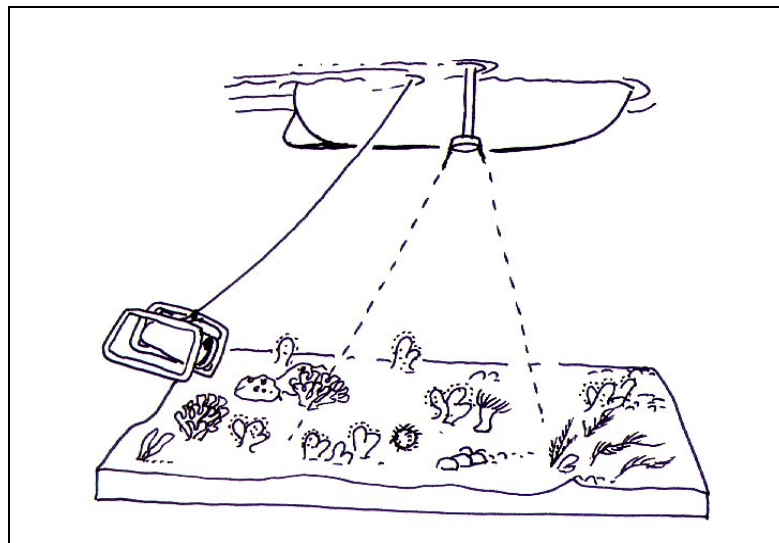
In the ideal situation, each biotope would have a distinctive combination of acoustic characteristics and depth range not shared by other biotopes (**exclusivity**). Whilst some biotopes can be distinguished by their distinctive acoustic properties, this will not be possible where two or more biotopes have similar acoustic characteristics (unless they are separated by some distance and their integrity can be verified by ground-truthing). Biotopes will have to be combined into larger categories where two or more biotopes are indistinguishable acoustically and are found in the same general locality. For example, some kelp biotopes are defined by their key species (*Laminaria hyperborea* versus *Laminaria saccharina*). These two species cannot be separated acoustically and, if they occur in close proximity, they will need to be combined into a single kelp biotope for mapping purposes.

6.4.3 Biotopes with a wide tolerance of habitat

Another source of variability stems from biotopes that are loosely defined in terms of habitat type and, therefore, are likely to encompass a wide range of acoustic ground types. In other words, some biotopes have a low **fidelity** to ground type.

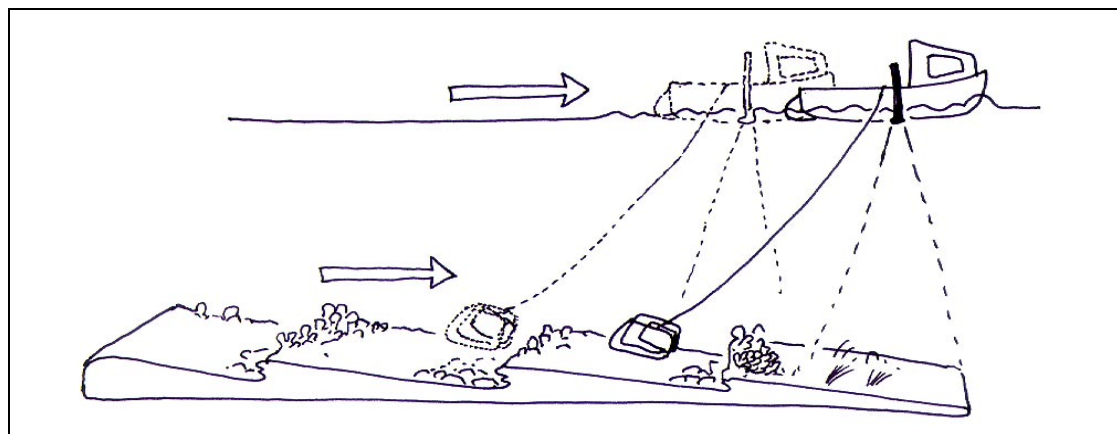
In summary, there are a number of limitations to habitat or biotope discrimination based on remote sensing and these will affect the way in which habitats or biotopes are shown on a map.

Figure 6.1. Different scales of heterogeneity relate to the techniques for acoustic survey and ground truthing by remote video. Three different scales are illustrated.



(a). The biotope scale: Mixtures of biotopes (either as intermediates between two or more biotopes or as small patches forming mosaics). The scale equates to an area ensonified by an echo sounder and to a field of view of a video.

(b). The video tow scale: associations of biotopes over a distance that equates to a video tow and short length of AGDS track.



(c). Small scale heterogeneity: Patches of different acoustic ground types detected by AGDS tracking, each of which can be ground truthed.

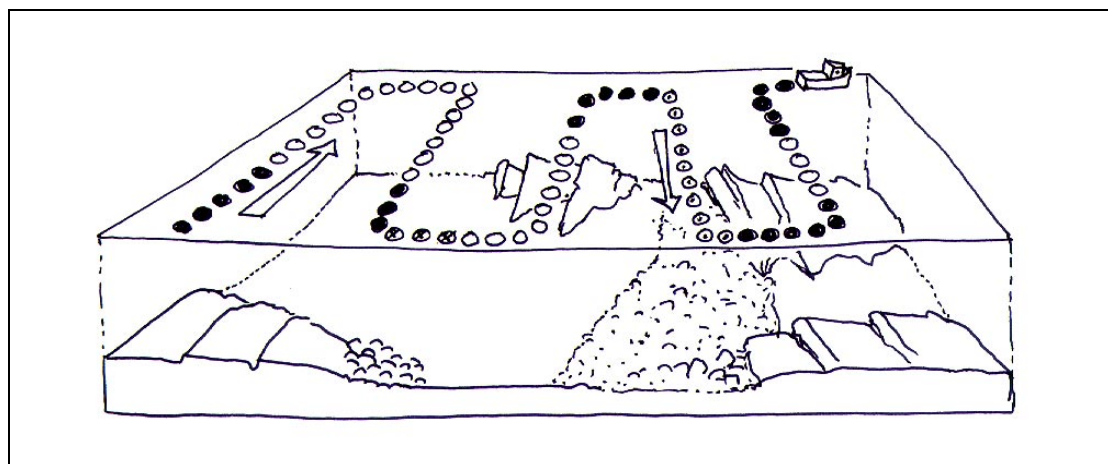


Table 6.1. Matching biotopes to acoustic parameters and conventions for map presentation

BIOTOPE	AREA*	ACOUSTIC PARAMETERS (Depth, roughness, hardness)	MATCH BETWEEN BIOTOPE & PARAMETERS	REPRESENTATION ON MAP
One biotope with good match to description in national classification	Large	Distinct throughout survey area	Simple biotope/parameter match	Use unique colour / hatch / letter code to represent biotope
One biotope with good match, possibly divisible by habitat or biotic attributes	Large	Apparently 2 sets of distinct properties	Can biotope be divided on easily recognisable features? Is it worthwhile?	If 'yes'; split biotope and give separate hatch/letter codes for each division. Add supplementary letters to biotope code and note these as provisional
One biotope that matches poorly with existing national classification	Large	Distinct throughout survey area	Simple biotope/parameter match	(1) Make do with a poor match if only of very local interest, use life form and note description (2) If of wider interest, create new biotope and give it a provisional code
Two biotopes	Many large polygons within survey area	Not distinct from one another	(1) Can other features be used to distinguish biotope, such as slope, distance from shore? (2) Cannot be distinguished: Both biotopes likely to be found, but not necessarily in close proximity to each other	(1) Use one code for each biotope (2) Are the biotopes similar at a higher level? If 'yes' use colour for life form. If 'no' use most likely biotope and note chances of finding other biotope to denote uncertainty. (3) If uncertainty is great between very different biotopes, highlight area as being poorly surveyed
Two biotopes with a common boundary	Large	There are very distinct acoustic characteristics either side of the boundary	The acoustic boundary marks a distinct biotope boundary	If desired, use a thick line to show degree of confidence of the boundary between the two biotopes
Two biotopes with a common boundary	Large	There is a gradual change in acoustic characteristics between biotopes	Biotopes may be distinct but grade into each other or biotopes are 'nodes' along a continuum	If desired, use boundary drawn with a thin line to represent either low confidence or that the boundary is somewhat artificial
Two or more biotopes	Small	Very variable characteristics indicative of the patchy nature of the ground	(1) One biotope predominant; or more than 2 biotopes in equal proportions:- (2) Biotopes of similar life form (3) Biotopes very different	(1) Use code of predominant biotope (2) Use common life form (3) Create an association of all biotopes and use special colour code by mixing predominant 2 life form colours in a hatch pattern.
Two or more biotopes	Very small	Very variable characteristics indicative of the patchy nature of the ground	Biotopes not found at a transition but are commonly found together	Create a new mosaic biotope. Use colour of one life form with unique hatch pattern and new letter code

*Large = area large enough so that ground truth position can be confidently placed well within uniform acoustic ground type.

Small = too small for ground truth point to be located with confidence within ground of a particular acoustic characteristic but where video shows only one biotope per field of view.

Very small = video indicates biotopes mixed at a scale smaller than the field of view.

6.5 The output from broad scale survey: what maps show

Does a habitat or biotope map show the information required at the expected level of detail? Section 6.5 illustrates what may be expected from broad scale mapping, the strengths and weaknesses of different styles of presentation. Examples are taken from each of the survey areas (Wash, Firth of Lorn and Pembroke) and each shows different approaches taken to mapping biotopes which are largely a function of the nature of the seabed.

6.5.1 Mapping biotopes in the Wash

The Wash does not have a large range of life forms or substrata and most biotopes are either infaunal or biogenic sand reefs on mixed substrata with sparse attached epifauna (Foster-Smith and Sotheran 1999). Grab samples were the primary tool for ground truthing and the infauna were analysed in some detail by Dr Peter Garwood of *Identichaete*. These were supplemented by video and dredge samples. Hence, the level of information and detail of the infauna was commensurate with the level of detail of the data used in the construction of the MNCR biotope classification. Thus, it was possible to describe the ground truth samples to the biotope level. Indeed, it was often felt that biotopes could be further subdivided into local variations on the basis of the detail collected and their distinctiveness from the nearest MNCR biotope. For example, a large number of samples could be classified as the *Sabellaria spinulosa* biotope (MNCR code SspiMx) since this appears to be a key species for biotope identification. However, it would be pointless to attempt image processing of the acoustic data if the majority of samples were of a single biotope category and much valuable information on the distribution of biota would be lost. Consequently, the *Sabellaria spinulosa* biotope has been subdivided into a number of more detailed biotopes suitable for the local data set.

Some samples did not match biotope categories sufficiently for a confident identification. The nearest biotope has been used, but other indentifications could also be valid as indicated in Table 6.2.

Table 6.2. The biotopes used for the classification of infaunal samples from the BMP survey with provisional local codes where appropriate.

Community categories	National biotope code equivalents (alternative interpretations in brackets)	Possible local code justified
Sabellaria	CMX.SspiMx	CMX.SspiMx.reef(Wash)
Sabellaria/Lanice	CMX.SspiMx	
Sabella discifera/Sabellaria	CMX.AbrNucCor (CMX.SspiMx)	CMS.AbrNucCor.Sdisc(Wash)
Sabella pavonina	CMS.AbrNucCor	CMS.AbrNucCor.Spav(Wash)
Abra	CMS.AbrNucCor	
Ophiura albida	CMS.AbrNucCor	CMS.AbrNucCor.Ophalb(Wash)
Scoloplos/Spiophanes	IMS.FaMS.SpiSpi	CMS.ScoSpi(Wash)
Nephtys/Scoloplos	IMS.FaMS.SpiSpi	CMS.ScoNep(Wash)
Nephtys/Bathyporeia	IGS.Fas.NcirBat	
Sparse polychaetes/nemertean	IMS.FaMS.SpiSpi	CMS.Nem(Wash)
Lanice	IGS.Fas.Lcon (IGS.Fas.Mob)	
Ensis	IGS.FaMS.EcorEns	
Modiolus	CMX. ModMx	

Three maps are presented (Figures 6.2, 6.3 & 6.4) which show how the biotopes can be displayed according to different levels of the classification and using two colour palettes. Figure 6.2 displays the biotopes labelled according to their most characteristic or conspicuous species and uses the life form

colour palette. These biotopes can be directly translated into MNCR biotopes but giving the codes a suffix to represent to local variants (Figure 6.3). In Figure 6.4 the biotopes displayed have been restricted to those in the MNCR classification (Connor *et al.* 1999). These maps illustrate the following two issues:-

Biotope detail: There is no reason why biotope distribution cannot be mapped to the lowest level of the MNCR biotope classification or even to the level of subdivisions if (a) this level can be recognised from the ground truth samples and (b) there is a reasonable separation of these biotopes using acoustic techniques. However, the Wash example illustrates that a direct comparison between the sample data and the biotope descriptions is not always easy and, although the ground truth samples have been consistently identified within the survey itself, their direct translation into MNCR biotope classes may be open to different interpretations by other people.

Colour palettes: Where a direct translation of the mapped classes and the MNCR biotope code is possible, then the maps can be displayed using either the life form or the MNCR colour palette. The life form and MNCR colour palettes for the sediment biotopes are, in any case, fairly similar. The life form colour scheme in the Wash (Figure 6.2) emphasises the following five components: (a) biotopes associated with mobile, medium fine – coarse sand (yellow); (b) muddy sand (olive); (c) biogenic sand reefs and accretions (orange); (d) muddy biotopes (brown), and; (e) shelly biotopes (grey). Similar trends are harder to distinguish from the colours in Figures 6.3 and 6.4. Many apparent trends in the distribution of infralittoral and circalittoral biotopes are spurious. These maps do, however, show major trends of substrata ranging from coarse sand in many shallow areas, through muddy sands in shallow and intermediate depths to mixed sediments in deeper water.

6.5.2 Firth of Lorn

The ground truth samples taken in the Firth of Lorn were attributed to the most detailed level in the MNCR classification that it was possible to identify from the video data. This was normally at the biotope level, although occasionally it was not possible to classify beyond a biotope complex. It should be noted however that the information which could be extracted from drop-down or towed video recordings was less detailed than often required to confidently attribute a biotope to the sample. In particular the following problems frustrated the classification process:

- Difficulty was sometimes experienced in identifying the key species which define a biotope or assessing abundance where this information was critical;
- Particular difficulty was experienced in clearly identifying the composition of a faunal turf which were often only identified to the level of hydroids and bryozoans;
- It was very difficult to separate closely related sediment types;
- Samples with a low species diversity yielded insufficient information to attribute a biotope;
- Mixed habitats where the visually dominant biotope was not recorded from the most common habitat. For instance where isolated boulders with large kelp plants were embedded in sediment, the associated acoustic data would reflect the more common sediment, rather than the less common boulders.

Notwithstanding these problems, stations were assigned to a total of 57 classes (47 biotopes and 10 biotope complexes). The large number of biotopes identified posed a number of problems when attempting to map their spatial distribution. Firstly, there were simply too many biotopes to legibly display on a map at a broad geographical scale ($>1:100,000$). Secondly, some of the biotopes were small units within a larger biotope; for instance faunal turf and seaweeds on shallow vertical rock (IR.CorMetAlc) occurred in kelp forests and kelp parks. Thirdly, some of the biotopes identified had identical acoustic signatures simply because they occurred on the same substrata; this was particularly true for the *Alcyonium* dominated biotopes on circalittoral bedrock. For these reasons it was necessary to group similar biotopes into more generic classes which had the additional benefit of reducing the likelihood of misclassification between samples. Two generic classification schemes were used in the present report: the biotope complexes from Connor *et al.* (1997), and the *life form classification* proposed by Foster-Smith *et al.* (1997). Table 6.3 describes how the biotopes were amalgamated into the more generic classifications. Note that the arrangement is hierarchical with biotope complexes being nesting uniquely within a life form.

Inspection of the acoustic signatures for the groups listed in Table 6.3 revealed a number of inconsistencies. Whilst the signatures were edited to remove obvious irregularities such as incorrect depth limits for algal classes and the incorporation of broad geographic trends for certain biotopes into the classification process (see Section 13.5.4), some issues could not be completely resolved:

- Some signatures were very general because the stations assigned to that class were themselves variable. For instance, brittlestar beds and bryozoan/hydroid turfs were recorded on habitats ranging from bedrock through mixed stones/sediment to sediment, with a varying density of animals. These stations were split into tightly defined classes based on the habitat to improve the correlation with the acoustic data (for example, MCR.Bri [rock] and MCR.Bri [stones]).
- Other biotope complexes recorded on mixed habitats could not be sub-divided and therefore their signatures remained general.

ID	Life form	Biotope complex	Biotopes recorded
1	Kelp on infralittoral rock	EIR.KFaR MIR.KR MIR.GzK SIR.K	EIR.LhypR; EIR.LhypR.Ft; EIR.LhypR.Pk; EIR.LsacSac; EIR.FoR; MIR.Lhyp.Ft; MIR.Lhyp.Pk; MIR.Lhyp.TFt; MIR.Lhyp.TPk MIR.LhypGz.Ft; MIR.LhypGz.Pk SIR.LhypLsac.Ft IR.CorMetAlc; IR.AlcByH
2	Kelp on infralittoral mixed substrata	MIR.SedK	MIR.XKScrR; MIR.EphR; IMX.LsacX
3	Faunal turf/faunal crust on circalittoral stones	ECR.Efa MCR.ByH *	ECR.PomByC MCR.SnemAdia; MCR.Flu; MCR.Flu.HbyS; MCR.Flu.SerHyd; MCR.Urt
4	Faunal cushions and faunal turf on circalittoral rock	ECR.Alc ECR.BS MCR.Xfa MCR.GzFa SCR.BrAs	ECR.AlcTub; ECR.AlcMaS; ECR.AlcSec; ECR.AlcC; ECR.AlcBri ECR.BalTub MCR.PhaAxi; MCR.ErSSwi MCR.FaAlc; MCR.FaAlc.Abi SCR.AntAsH; SCR.NeoPro CR.Ant
5	Brittlestar beds	MCR.Bri *	MCR.Oph
6	Mixed coarse sediments	IGS.FaS CGS	IGS.Mob; IGS.ScupHyd IMX.FaMX (A biotope complex but stations included in IGS.FaS) No biotope CMX.ModHo (A biotope complex but stations included in CGS)
7	Tideswept gravel	CGS	CGS.Ven.Neo
8	Tideswept sand	CGS	No biotope
9	Muddy sand	IMS.FaMS CMS	No biotope CMS.VirOph; CMS.VirOph.Has; COS.AmpPar
10	Circalittoral mud	CMU	CMU.SpMeg; CMU.SpMeg.Fun; CMU.Lept

Table 6.3 The relationship between the different classifications used for biological mapping. Biotope complexes marked * were sub-divided into two classes for mapping.

Figure 6.5 shows an example of part of the Firth of Lorn where both the life form and the biotope complex classification has been used. The map showing the biotope complexes provides considerable biological information, for instance on the spatial heterogeneity in this area. It should be noted however, that this map will be influenced by misclassified ground truth samples. Also, the colour scheme makes it difficult to formulate an overview of any broad patterns or trends throughout the area. Using the life form colour scheme, it is possible to extract broad patterns and the map will be less influenced by the misclassification of ground samples. Perhaps the best solution in this case would have been to use the life form colour scheme at the biotope complex level.

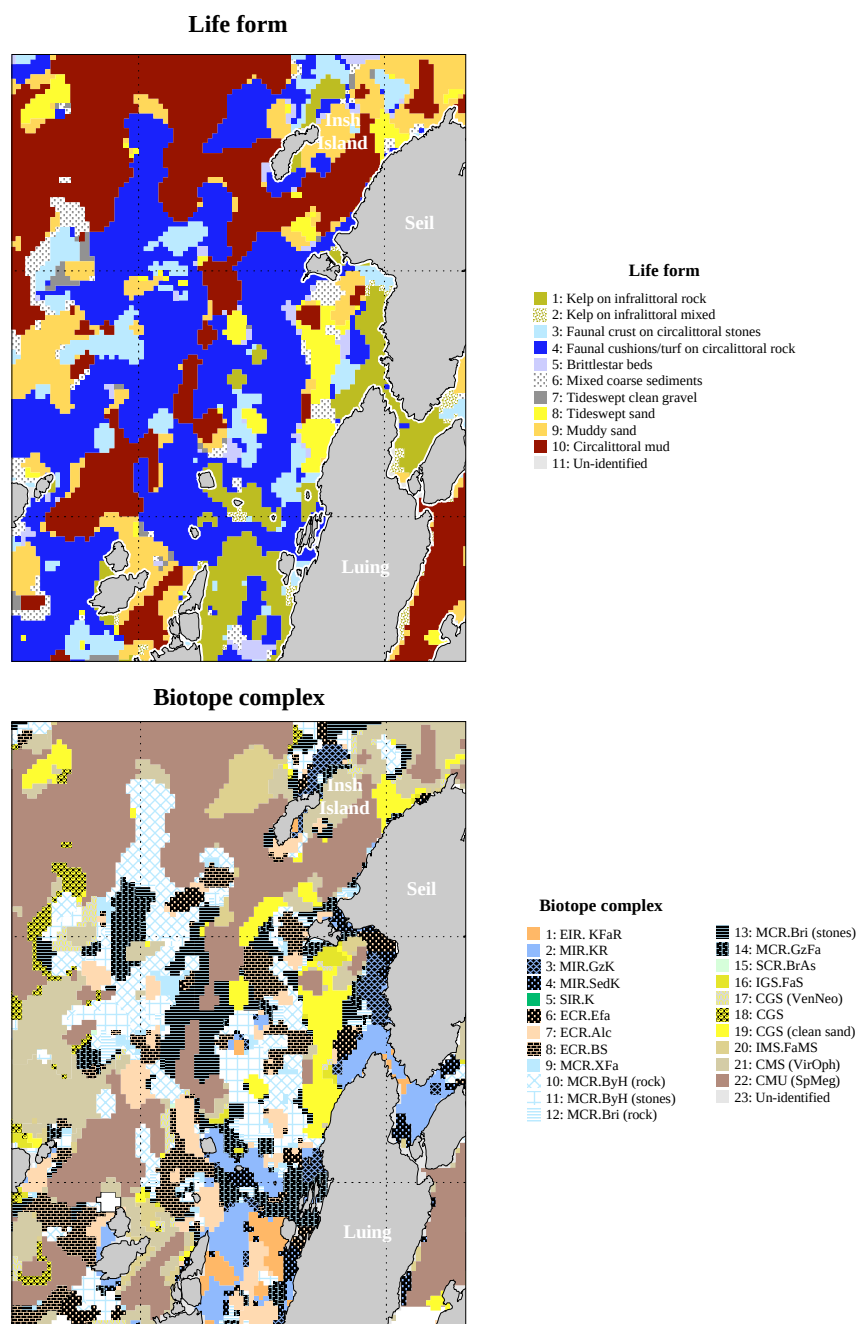


Figure 6.5 An example of showing the predicted distribution of life forms (upper map) and biotope complexes (lower map) in Easdale Bay, Firth of Lorn.

6.5.3 Mapping life forms and substrata in Pembroke

A situation in which the biotopes overlap considerably both in terms of species composition and substratum is not ideal for image processing since discrimination is compromised by this high degree of overlap. This is well illustrated by the Pembrokeshire survey. Many of the ground truth samples were not identified with confidence because (a) the species composition of the biotopes were not clear cut but graded into each other (leading to possible misclassification) and (b) the biotopes were found across a wide range of substrata so that there was considerable overlap in sediment characteristics between many biotopes. Table 6.4 summarises those infaunal biotopes used in the image processing of the acoustic data. The biotope, habitat and biotope complex codes are followed by the life form. The representation of these biotopes across the range of substrata is given as number of occurrences in the columns to the right of the life form code.

It can be seen from Table 6.3 that many of the biotopes were found on many different types of substrata with few, especially amongst the epifaunal communities, showing a marked preference for one particular substratum. For example, there was a large number of faunal turf life form biotopes (8 biotope classes which comprised 94 samples out of a total of 273) with few obvious substratum preferences. Indeed, many of the species that numerically dominate the faunal turf biotopes were common to them all and this made identification difficult and, in many cases, somewhat arbitrary.

Likewise, brittlestar beds were catholic in their substratum preferences, being found on rugged bedrock to gravel, although not on sand. Of the biotopes characterised by epibiota, only the *Pomatoceros*, *Balanus crenatus* and algal crust biotope (MNCR code PomByC) was found almost exclusively on one substratum, that of cobbles.

Classification of the acoustic images according to biotopes is made more problematic by overlapping characteristics between biotopes and classification using the broader life form categories was more successful (Figure 6.6). By contrast, the substratum classes used to categorise the samples appeared to be easily recognisable and distinct. Mapping at the life form level would not seem to offer a great deal of information (considering the effort expended in survey). However, there is scope for refining the description of the biotopes through the use of information on predicted habitat distribution. The approach adopted was to cross-classify the life form and substratum distribution maps to give more detail on the nature of the biotopes (see Figure 6.7). In theory, cross-tabulation is possible whenever biotopes (or life forms) and substrata are mapped separately. In practise, this is a complex task that is prone to the creation of many small 'slivers' of improbable combinations where the two overlays do not match properly. For example, in boundary conditions between rock and sand the biotope map might indicate a rich faunal turf biotope where the substratum map indicates sand. These slivers need to be edited. Nevertheless, cross-tabulation may provide a useful means of combining biota and habitat information to increase the information content of a map.

Colour palette: The MNCR colour palette was used to colour a version of the life form map (see Figure 6.8). Since the life forms encompassed a number of biotopes from different exposure classes, hatching was required to represent these mixtures of biotopes. The two colour palettes used in Figures 6.6 and 6.8 emphasise different aspects of the biotopes. The life form colours show shallow water kelp biotopes fringing rocky faunal turf biotopes close to rocky headlands with extensive faunal and algal crusts on cobble. The map does not distinguish between the predominantly inshore bedrock/large boulder faunal turfs and the offshore boulder/cobble faunal turfs since their biotope composition is not distinctive enough. The MNCR colour system shows a fringe of infralittoral biotopes (stronger colours) close to the rocky headlands that are a mixture of exposed and moderately exposed biotopes and hence the fringe is represented by two hatch patterns denoting kelp and algal turf respectively. The majority of the area, however, has been classed as circalittoral (pale colours). The faunal turf biotopes are, once again, mixtures of exposed and moderately exposed biotopes and are represented by cross hatching.

Clearly, since cross hatching has already been used with the MNCR colour scheme, the results of cross-tabulation with substrata would be difficult to display and has not been attempted.

6.5.4 Summary

The mapping units: Clearly, mapping at the most detailed level should be attempted since biotopes can always be amalgamated into larger units post-survey, but broad categories cannot be split post-survey. Thus, the mapping units used for classifying the ground truth samples into categories suitable for image processing (i.e., resulting in the best discrimination) are, in order of detail and preference:-

- The biotope (possibly subdivided when the biotope is too broad a category for discrimination).
- The biotope complex
- The life form (incorporating habitat complexes for sediment types as a last resort where sediments biotopes are poorly sampled: Possibly sub-divided on the basis of sediment characteristics alone. For example, the MNCR habitat category 'gravels & sands' was often subdivided into separate substratum categories 'gravel', 'coarse sand' and 'medium fine sand').

By and large these units nest into each other with very few cases where the biotopes within a biotope complex required allocation to different life forms (see Appendix 1). A mixture of levels could be mapped depending upon the varied levels of discrimination possible. In particular, poorly sampled sediments classified only to sediment type (equivalent to the sediment categories of the habitat complex level of the MNCR classification system) were often mapped whilst rocky habitats were identified and represented on the same map at the biotope level. However, the level of detail of the maps varied widely from area to area depending upon the nature of the substrata and the spatial heterogeneity of the biotopes.

In the case of the Wash, biotopes or sub-divisions of biotopes were consistently mapped whilst in Pembroke and the Firth of Lorn the mapped units used on a single map ranged from biotopes (such as brittlestar beds) to life forms (faunal turf) and habitat complexes (circalittoral muddy sand). The use of mixed levels of biotope/life form categories is purely pragmatic to make best use of the data and show the maximum amount of information possible.

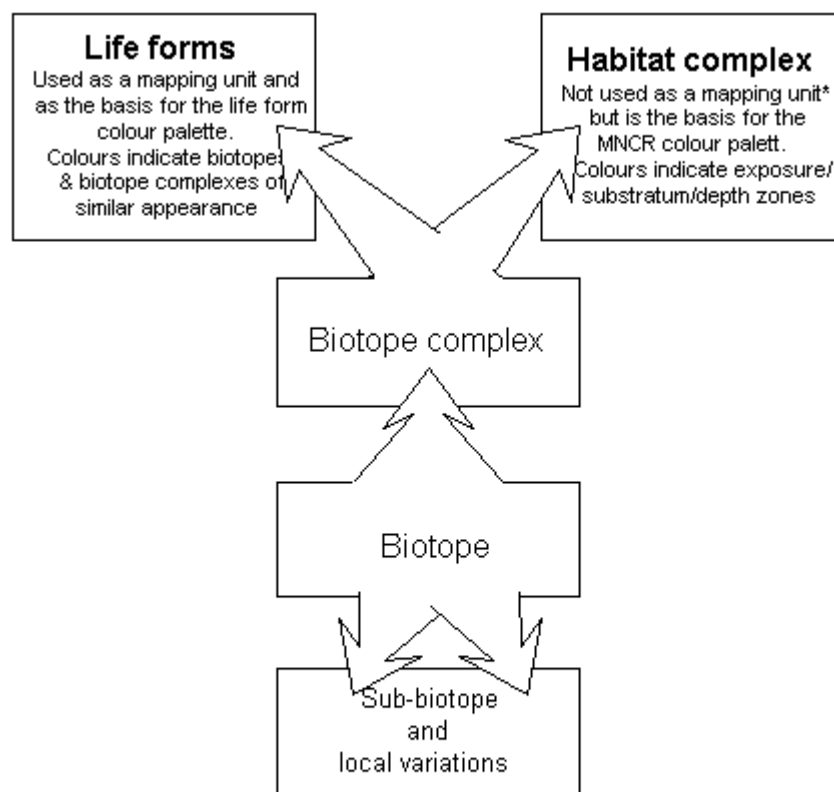
The colour palettes: Figure 6.9 shows the nested hierarchical arrangement of biotopes (and their subdivisions) into either life form categories or habitat complexes. The arrangements diverge above the biotope complex; i.e., biotope complexes can either be arranged by their overall appearance (life form) or by substrata/exposure/depth. Thus, if a map shows biotopes or biotope complexes the colour palettes can be switched quite simply (as can be seen from the Wash maps). The Firth of Lorn examples illustrate maps that compare detailed biotopes/biotope complexes using the MNCR palette with these units amalgamated into life form units for the use of the life form palette. However, there is no reason why the maps could not have shown the biotopes/biotope complexes using the life form colour scheme (in which case the level of detail would have been the same).

In this situation it is a matter of choice as to which palette is used since the level of detail does not change. However, in both the Wash and the Firth of Lorn surveys, it is the view of the authors of this report that more biological information is conveyed with less inconsistencies using the life form palette.

When higher level life form units are used for mapping it is very likely that these will be composed of biotopes/biotope complexes that have a similar appearance but are classified in different habitat complexes. These mapping units can only be displayed using the MNCR colour palette (based on habitat complexes) if different colours are mixed with hatch patterns, although the scope for this is very limited.

Maps showing biotopes/substrata: Biotopes are defined by their biota and physical habitat. However, in reality most biotopes are narrowly defined by their biota and the range of habitats on which the communities are found can be quite wide. This dichotomy between the community and the physical habitat (Figure 6.9) becomes even more obvious with biotope complexes and life forms (in that the importance of the biota to the definition of the unit becomes greater whilst the habitat ranges become wider). Thus, mapping units (biotopes/biotope complexes/life forms) can be usefully subdivided on the basis of the substrata on which they are found to add further information to a map. This has been with the Pembroke data (Figure 6.7) by cross-tabulating the life form maps with the habitat (substrata) maps. These are displayed using a two-colour hatch pattern. Clearly, this scheme would not be possible with the MNCR palette.

These examples show the level of detail to which biotope mapping can be taken and this varies between the different trial areas. However, the lowest common denominator between the three surveys is the life form (incorporating modified habitat complex categories for sediments). Additionally, the use of a palette based on life forms emphasised biological trends in the distribution of biotopes. For these reasons the life form colour palette was adopted as standard for the BMP.



* but sediment types were incorporated into the life form system (with subdivision where appropriate) for poorly sampled sediments

Figure 6.9. The relationship between biotopes, biotope complexes, life forms and habitat complexes. The biotopes could be subdivided where finer levels were needed for image interpretation.

Table 6.3. Biotopes (also with life form descriptions) from the Pembrokeshire survey identified from the ground truth samples. The representation of each biotope across the spectrum of substrata are given in the columns to the right of the life forms.

Biotopes				Substratum								Total biotopes
Biotopes	Habitat	Complex	Life form	Bedrock	Rugged bedrock	Sandy bedrock	Boulder	Cobble	Gravel	Sand	Fine sand	
FoR.Dic	EIR	KFaR	Algal turf	2	1		2	9				14
XKScrR	MIR	SedK	Algal turf	3	2	2	1	2				10
Oph	MCR	Bri	Brittlestar bed	2	1		1	5	1			10
PomByC	ECR	EFa	Faunal & algal crusts	1		3	1	34	3			42
AlcC	ECR	Alc	Faunal turf	1				1				2
AlcMaS	ECR	Alc	Faunal turf	7	5			3				15
AlcTub	ECR	Alc	Faunal turf	1								1
ErSPbolSH	MCR	Xfa	Faunal turf	8	8		1	5				22
FaAlC	MCR	GzFa	Faunal turf					5				5
Flu.HByS	MCR	ByH	Faunal turf	2	2	1		7				12
Flu.SerHyd	MCR	ByH	Faunal turf	1		5		3		1	2	12
SNemAdia	MCR	ByH	Faunal turf	5		2		17	1	1		26
TubS	ECR	BS	Faunal turf	5				1				6
FabMag	IGS	FaS	Gravel					2		1		3
Mob	IGS	FaS	Sand							2		2
NcirBat	IGS	FaS	Sand							4		4
Ven	CGS		Gravel						3	1		4
VenNeo	CGS		Gravel						1			1
Lhyp.Ft	EIR	KFaR	Kelp	7	1	1	3					12
Lhyp.Pk	EIR	KFaR	Kelp	4	5			4				13
LsacChoR	MIR	SedK	Kelp				4	1				5
SpMeg	CMU		Mud								2	2
AbrNucCor	CMS		Silty sand					2	3	8		13
AfilEcor	CMS		Silty sand								14	14
EcorEns	IMS.	FaMS	Silty sand			4				7		11
MacAbr	IMS	FaMS	Silty sand						1			1
SpiSpi	IMS	FaMS	Silty sand					4		7		11
SspiMx	CMX		Sabellaria gravel							2		2
MolPol.Sab	MCR	As	Sand reefs			1						1
Total substrata				48	25	19	13	105	13	25	18	273

7 Assessing the performance of a map

Purpose: To illustrate what habitat and biotope maps can show. To summarise ways in which information can be presented to map-users to enable them to make a judgement about map confidence

7.1 Confidence and accuracy

Accuracy and confidence are extremely hard to assess and there is no single measure of accuracy that can be applied to a map. Nevertheless, some assessment of accuracy is required so that potential users of maps can make their own judgement as to the degree of confidence they can place in a map. This report has already stressed the potential for error in broad scale mapping: The sources of error in remote sensing in general have been introduced in Section 3.4. Also, errors inherent in the AGDS technology was discussed in Section 4.3 and variability due to analytical techniques in Section 5. Lastly, uncertainty in identifying biotope classes from sample records was discussed in Section 6.

This Section summarises these various sources of error and suggests ways in which accuracy and confidence can be objectively assessed. The issues are not dealt with in detail but refer the reader to other Sections in Part 2 of the report.

7.2 Visualising and describing the data

No map should be taken on trust without at least having knowledge of the data upon which it is based, how the data were collected and analysed and what are the potential sources of error. There are many ways in which the track data and the various stages of the interpretation process can be visualised which familiarise the viewer with the data.

1. **Visualising AGDS coverage:** Track spacing over the survey area will give a good idea of coverage and point up areas of where there will be uncertainty due to a low intensity of tracking (Section 11.2).
2. **Visualising ground truth coverage:** Likewise, a distribution map of the ground truth samples will show where interpreted habitat or biotope distribution is well supported by ground truth data and where areas are predictive with few data to reinforce the interpretation of the AGDS.
3. **Visualising the acoustic variables:** The ranges of variables measured by *RoxAnn* are few and this limits the scope for discrimination. These variables can be shown as separate thematic maps and these are useful as a visual guide as to the general patterns of the variables upon which all interpretation is based. If there are numerous anomalies, they will be obvious at this stage. However, distribution maps of the variables will only be possible where there is a single data set or where two or more data sets can be standardised. (Sections 10.4 & 13.8).
4. **Visualising acoustic ground types:** The echo sounder measures reflectivity properties of the sea floor and not the habitat or biota directly. Acoustic ground types might be shown, including an unsupervised classification of the acoustic data. Again, this is only possible where data sets are comparable (Section 13.2).
5. **Visualising the pattern of spatial variability for the data set:** Interpolated values are used to create continuous-coverage digital images suitable for processing. These are calculated values that may be poorly supported by real data with increasing distance from tracks. Variograms can be presented and the appropriate ranges displayed as buffer zones around the acoustic tracks (Section 11.3).
6. **Demonstrating relationships between biotopes:** It is assumed for the purposes of map production, that the list of biotopes and accompanying descriptions used for the categorisation of samples are complete and the biotope classes are discrete entities. Neither of these assumptions need necessarily hold true. Intermediates between two or more biotopes will occur and there may even be samples that cannot be matched to any existing biotope. A description of the biotopes used should be given indicating where there may be confusion between classes.
7. **Habitat and biotope probability maps:** Biotope probability maps may provide an alternative to maps showing maximum likelihood. Each class of habitat or biotope will have its own probability

distribution throughout the mapped area. These probabilities range from zero (a nil probability of it being found at a particular location) to one (a certainty of its occurrence). There will be one map per class, but this does overcome the problem of over-generalisation that is an inevitable consequence of using maximum likelihood classification (see Section 13.6).

8. **Showing the range of habitats or biotopes that might be found:** Maximum likelihood classifiers assigns each pixel to the habitat or biotope class to which it most likely belongs. This requires the classifier to select between biotopes that may have very similar probabilities. Thus, there is the possibility that other classes with slightly lower levels of probability may be under-represented on the map. Although there are ways of conveying this uncertainty (see Section 13.7), it is likely that many users will assume that the classes are mutually exclusive. If a single biotope is used as a short-hand convention to indicate a suite of biotopes, then the other biotopes that might be expected to be present should also be listed in the legend of the map or even as codes on the map itself.

7.3 Measures of accuracy and uncertainty

Whilst visualisation of confidence are important and effective, there are other methods that measure and demonstrate accuracy and uncertainty more directly.

7.3.1 Accuracy

Accuracy can be measured by finding the proportion of ground truth samples that match the habitat or biotope map compared to the proportion that do not match. These accuracy assessments take the form of an error matrix which is formed by matching the predicted biotopes or habitats from the map with the actual categories (the ground truth samples or external sample data set). The diagonal of the matrix indicates how well each biotope class has been predicted. Matrices can be used to assess **internal accuracy** and **external accuracy**. Accuracy can be measured as a simple percent correct or the agreement between the sample data and the map above and beyond chance (see Section 14). These measures can be applied to the map as a whole or to individual habitat or biotope classes.

An internal accuracy measure assesses how well the map matches with the ground truth data that have been used to classify the AGDS data. It is really a measure of the success of the analysis. It should be stressed that the ground truth data are used both for classifying the acoustic data and for accuracy assessment and internal accuracy is, therefore, not a good measure of the predictive power of a map. A better measure of confidence and predictive capability of a map is obtained if the map is tested against an external sample data set.

These accuracy assessments must be interpreted with caution. Small areas with only a few habitat or biotope classes will always return higher accuracy values than will be the case for large diverse areas. Thus, accuracy can be increased if biotope or habitat classes are amalgamated into a smaller number of broadly defined classes. Success must be judged not only by the number of correct and incorrect predictions ('hits and misses') but against the degree of closeness of the predicted community to that actually found and also the level of detail predicted. For example, if a hydroid turf is predicted, but a hydroid/bryozoan turf found instead, this might still represent a close prediction. On the other hand, sand and rock might be correctly predicted but this might be considered a rather coarse level of discrimination. There is a trade-off between accuracy and detail: accuracy increases as the level of detail decreases.

A one-to-one match between a map and an external sample data set is a very strict use of the external error matrix and makes no allowance for near misses that might be expected with the combined positional errors of the track data and the sample locations. Some tolerance might be acceptable, but this must be used with caution since there is a risk that the 'correct' biotope will be found by chance if the search distance is set too high.

7.3.2 Uncertainty

Uncertainty of the classification of an image can be calculated and displayed in map-form. In essence, the calculation is the same as that used for the maximum likelihood classification. The probability that a pixel belongs to each of the biotope classes is first calculated. Maximum likelihood classifies a pixel according to the most likely class. If the probability for one particular biotope is much higher than for any other, then the uncertainty about this classification will be low. However, if a pixel could be classified as a number of biotope classes with more or less equal probability, then the uncertainty of classifying the pixel as any one of them will be high. The process results in an overall map expressing uncertainty and a series of probability maps for each biotope class (see Section 13.7).

7.4 An assessment of the map from expert local opinion

However well the analyst regards the match between the map and the survey data, a map is unlikely to carry conviction if it conflicts with the view of local biological expert opinion. It is important, therefore, that maps are circulated for comment and any concerns expressed explored by reanalysis of the data. All the maps in the BMP were circulated for comment. The predicted distribution

In many cases these comments reflect broad scale trends in the distribution of habitats or biotopes which may not have been considered in the analysis of the data. For example, some biotopes may be restricted to sheltered inlets but the automated image classification process searches for (and possibly 'finds') this biotope throughout the survey area. If these broad trends are pointed out, then they can be incorporated into the analysis (discussed in more detail in Section 13.5.4).

Ultimately, there is only one way to test the usefulness of a map and that is by testing its predictive capability. Ideally, a user should plan further sampling based on the map to prove its accuracy.

7.5 Ease of use

Maps should be easy to read and use. This may entail a compromise between the level of detail available and simplicity. This is especially important for digital image processing where it is tempting to represent the interpretation of every pixel no matter how isolated it may be. Thus, filtering is usually required to simplify the image for ease of viewing. The user should know what level of filtering has been applied.

Part 2. Planning a broad scale survey and the analysis of results.

The purpose of Part 2 of the Technical Report is to give guidance, based on the experience of *SeaMap* and the Broadscale Mapping Project in particular, to those readers who are involved with the detailed planning of broad scale mapping. There are many options open to surveyors and no single ‘best’ pathway for survey and data analysis can be put forward. Thus, there are number of Sections each targeting particular parts of the survey methodology, some of which are general in nature and others that are more specific. This approach to Part 2 is somewhat modular and Section 19 attempts to redress this by summarising the methodology.

Many of the topics have been introduced in Part 1 and there is, inevitably, some repetition of the material in Part 2.

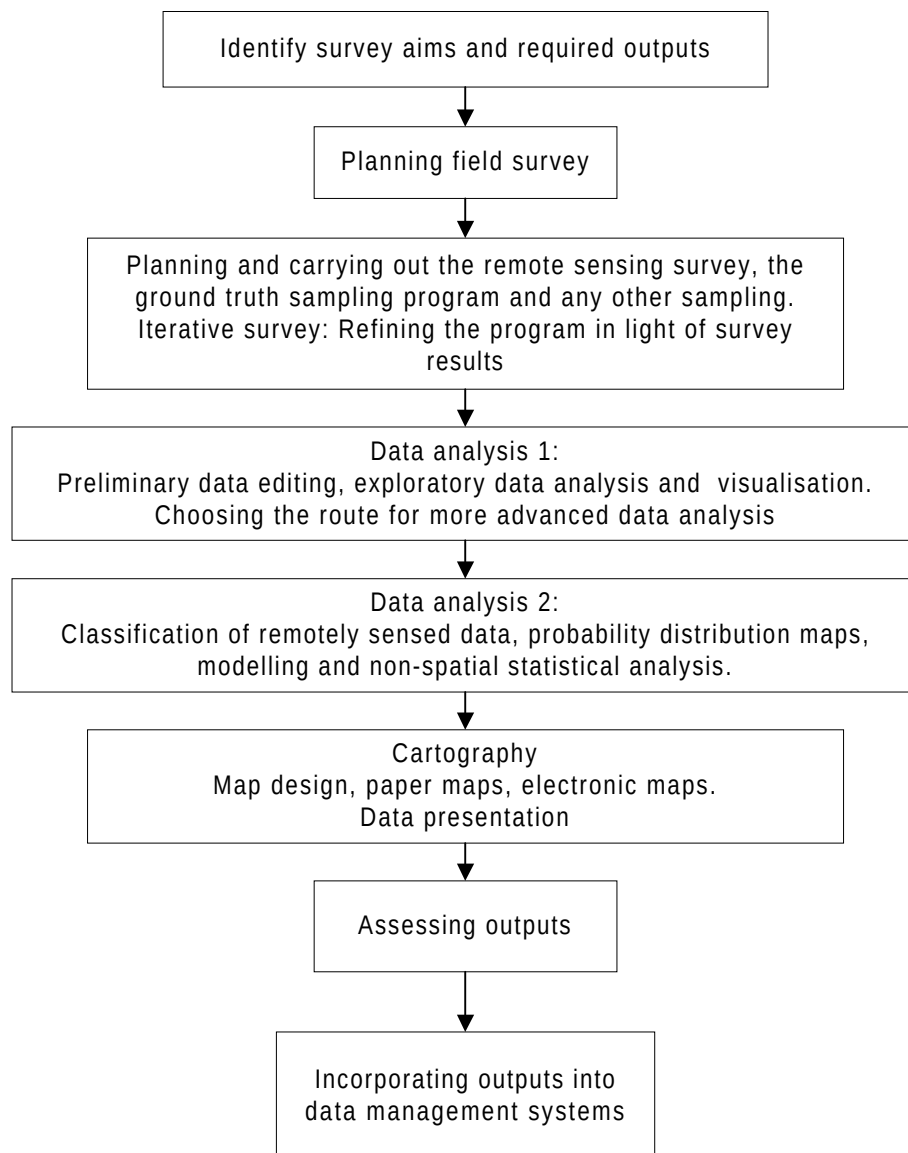
8 Planning a broad scale mapping survey

Purposes: (a) to identify and if necessary prioritise the aims of the survey (b) to decide upon the required outputs from the survey and the format in which they are presented and stored; (c) to match survey design to aims and required outputs

8.1 Key stages in planning and executing a broad scale survey

Broad scale mapping surveys should be planned with clear aims and a realistic expectation of the outputs. Although detailed planning might be left to the surveyors, those commissioning surveys should be aware of the key decisions to be made at each stage and the implications for time and costs. The key stages are set out in Figure 8.1.

Figure 8.1. Key stages in planning broad scale survey



8.2 Survey types based on AGDS

Broad scale survey is defined by the application of remote sensing to the survey. Broad scale mapping surveys should be planned with clear aims and a realistic expectation of the outputs. There are many different types of survey where remote sensing and the spatial interpretation of data are important (see below). If it is decided that a broad scale survey is required at least as part of a survey strategy, then there are many subsidiary issues which need to be considered. Figure 8.2 shows a decision tree that can be used to help decide what type of survey is most appropriate to a particular situation.

This Technical Report is mainly concerned with the application of remote sensing in broad scale surveys, but there are many different types of survey where remote sensing and the spatial interpretation of data are important. Many of the procedures used routinely for broad scale mapping may not be appropriate for the detailed survey and monitoring of small sites. This particularly applies to resolution, intensity of survey and interpolation (point-to-area conversion).

It is important to establish the anticipated resolution of a habitat or biotope map early in the planning stage so that expectations of positional accuracy and detail are not over optimistic (see Section 15). The resolution of an acoustic survey will not only depend on survey intensity, but also on the way the AGDS is used to discriminate between features. For example, the precision of boundaries will be set by the maximum resolution of the AGDS where changes in the acoustic values alone can be confidently used to map biotopes. AGDS can be used at the maximum resolution if used intensively in small areas where there are boundaries between a small number of distinct biotopes that form relatively large homogeneous areas. The ground truth samples are used in this situation as a 'check' to confirm what is obvious from the AGDS data on their own.

On the other hand, if the full classification process is required to identify and establish the position of a wide range of biotopes, then statistical procedures are required to link ground truth samples to acoustic data and the resolution is limited by the combined uncertainty of the acoustic and the ground truth surveys. These two ways of using the AGDS must not be confused in the survey planning stage.

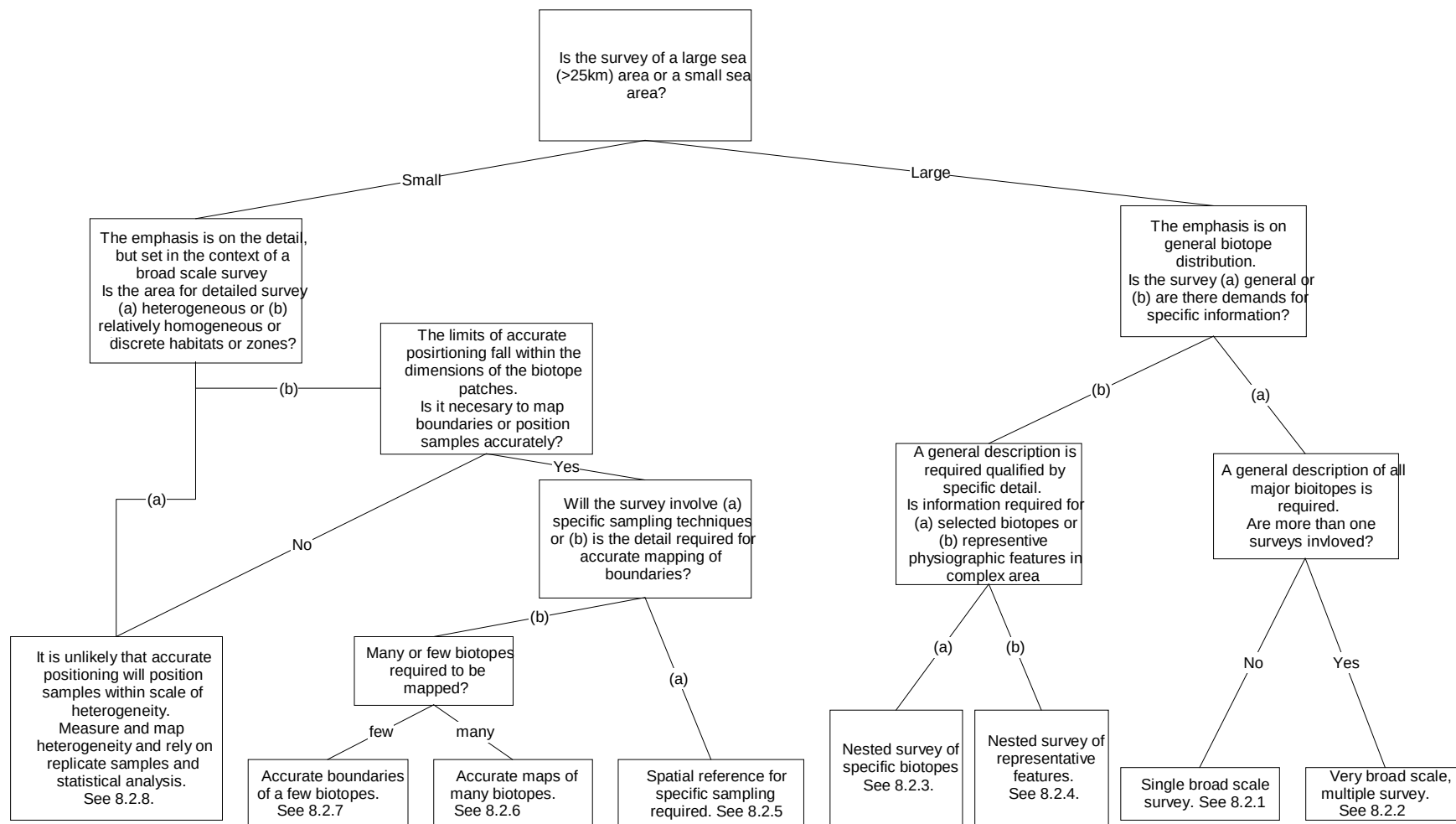
The reasons for interpolating point data were discussed in Section 5.2.1 where it was pointed out that the requirement for habitat and biotope maps as a continuous coverage applies particularly to broad scale survey and mapping. For small areas that have been intensively surveyed, as might be the case in monitoring studies, the interpretation of the track point data will be more accurate and still convey spatial trends in the distribution of habitats and biotopes.

As a useful generalisation, broad surveys of large areas can be compared with detailed surveys of small areas in the following way:-

Broad surveys of large areas	Detailed surveys of small areas
Wide track spacing	Narrow track spacing
Low density of ground truth samples	High density of ground truth samples
Interpolation required to model spatial distribution patterns	Spatial patterns may be obvious from the track point data without interpolation
The emphasis of AGDS interpretation is on prediction since many areas of the map are sparsely sampled	The emphasis of AGDS interpretation is on refining boundaries around habitat and biotope samples

Figure 8.2 gives a decision tree suggesting choices between survey types that may be made depending upon the objectives, area to be surveyed and costs. The survey types are described in the following pages.

Figure 8.2. Survey design



8.2.1 Broad scale survey

Purpose: To provide a baseline of knowledge of the extent and distribution of a broad range of biotopes or habitat types where little information exists.

AGDS track spacing: Broad track spacing of about 250m and an approximately even coverage throughout survey area. Alternatively, some sectors within the survey area may be surveyed at a much lower intensity where the ground is relatively homogeneous and/or the ground is of less interest than other sectors. The track spacing may be as much as 1-2km.

Ground truth sampling: Rapid sampling techniques such as video. Sidescan sonar would be recommended deployed for ground truthing of some representative features. Grab samples may be used where sediment habitats are predominant; an even coverage throughout area with emphasis on sampling representative acoustic ground types; a minimum number of ground truth samples will be needed per major biotope (about 5) spread throughout the survey area, but no replicate samples at each sampling station.

Data analysis and outputs: The basic procedure for classifying image processing can be applied to the whole image. The results can be presented as a small scale map covering the whole area. Since interpolation introduces detail which is more apparent than real, there may be an advantage in producing a version which is a very general account of the distribution of major biotopes or suites of biotopes.

Costs: Approximately 20-30 km square per 5 day week/2 people which would generate 2 weeks of analysis. Total of 4 person/weeks.

Advantages: Large areas covered; even coverage simplifies data analysis and the whole area can be represented to a uniform level of detail.

Suitable for: Apart from main purpose, also suitable for locating possible features of interest that may have special management requirements. Broad scale surveys are also suitable for planning more detailed, targeted survey.

Disadvantages: Distribution patterns only approximate and interpolation between tracks will only be able to represent broad biotope trends and not fine scale patterning. Repeat surveys will need to follow previous tracks closely to avoid differences in interpretation due to positional changes.

Not suitable for: This type of survey is not suitable for an accurate estimate of areas of biotopes nor will it establish boundaries with confidence. It follows that it is also not suitable for monitoring change over time except for gross change.

8.2.2 Joining two or more broad scale surveys

Purpose: To provide a baseline of knowledge of the extent and distribution of a broad range of biotopes or habitat types over very large areas of the seabed by joining two or more broad scale surveys.

AGDS track spacing: Broad track spacing and, preferably, an equal treatment for each of the component survey areas. Different surveys may be conducted using different AGDS systems.

Ground truth sampling: Rapid sampling techniques; an even coverage throughout area with emphasis on sampling representative acoustic ground types; the emphasis may change between surveys, but should be standardised if possible. Sidescan sonar would be recommended for some representative features deployed for ground truthing.

Data analysis and outputs: The basic procedure for classifying image processing cannot be applied to the whole image because of regional trends in the data and differences in the AGDS systems. There may also be differences in resolution between surveys. Analysis will need to (1) take account of regional trends in biotope or habitat distribution; (2) make allowances for different responses of the AGDS and; (3) use collateral data from other sources in an attempt to get an even coverage throughout the area.

Costs: An allowance of approximately 4 weeks will be needed for the analysis required for combining different surveys

Advantages: Large areas covered; use is made of various data sets without the need for re-survey.

Suitable for: Providing a regional context for the assessment of site specific issues.

Disadvantages: Distribution patterns only approximate and combining different surveys introduces an additional level of uncertainty. It is unlikely that surveys will be repeated on such a large scale so that there is a risk that inadequate surveys would remain as part of the baseline information for the region by default.

Not suitable for: This type of survey is not suitable for an accurate estimate of areas of biotopes nor will it establish boundaries with confidence. It follows that it is also not suitable for monitoring change over time except for gross change.

8.2.3 Nested survey targeted on selected biotopes

Purpose: Broad scale survey with nested finer scale survey to provide a baseline of knowledge of the extent and distribution of a broad range of biotopes or habitat types but with a requirement for increased detail on specific biotopes or habitats where these are thought to occur within the survey area.

AGDS track spacing: Wide spacing and an even coverage throughout survey area combined with nested areas of narrow track spacing (50m) dependant upon locating smaller areas of interest during the broad scale survey phase.

Ground truth sampling: Rapid sampling techniques (possibly including sidescan sonar) with an emphasis on sampling representative acoustic ground types; focused intensive sampling on areas of interest, perhaps using techniques to extract detailed information.

Data analysis and outputs: Areas of detail will need additional analysis. The results can be presented as a small scale map covering the whole area together with separate maps of targeted areas mapped to a greater level of detail. Alternatively, these areas may be included within the main map clearly marked as being at a finer spatial resolution (smaller pixel size).

Costs: Approximately 15 km square per 5 day week/2 people which would generate 3 weeks of analysis. Total of 5 person/weeks.

Advantages: Large areas covered to rapidly locate areas of interest and provide context for evaluating probability of having located the majority of targeted biotopes. Cost-effective way of locating biotopes of interest and obtaining reasonable level of detail of finer structure of such areas.

Suitable for: As for broad scale survey; additionally for drawing boundaries around features of interest with a reasonable accuracy and confidence in that most features of a size larger than the inter-track spacing will have been located. Thus, will be able to provide useful statistics on extent of features of interest.

Disadvantages: Distribution patterns only approximate and uncertainty will exist as how well the survey has performed in detecting all biotopes of interest. Nesting maps requires special explanation as to the effect on scale/level of detail. Repeat surveys will need to follow previous tracks closely to avoid differences in interpretation due to positional changes.

Not suitable for: Accurate location of all features of interest or accurate monitoring for change.

8.2.4 Nested survey of representative features within a broad scale survey

Purpose: Broad scale survey with nested finer scale survey to provide a baseline of knowledge of the extent and distribution of a broad range of biotopes or habitat types but with a requirement for increased detail on representative physiographic features which can be identified from maps and charts.

AGDS track spacing: Broad track spacing and an even coverage throughout survey area combined with nested areas of finer track spacing (50m) on selected sites within the survey area.

Sidescan sonar: Mosaic sidescan may be used for a comprehensive description of representative samples of features with the AGDS data supplementing the image with point information on sediment characteristics.

Ground truth sampling: Rapid sampling techniques (possibly including sidescan sonar deployed as a method of ground truthing the AGDS) with an emphasis on sampling representative acoustic ground types; focused sampling within selected areas

Data analysis and outputs: Areas of detail will need additional analysis. The results can be presented as a small scale map covering the whole area together with separate maps of targeted areas mapped to a greater level of detail. However, there may also be a requirement for a version that is a general account of the distribution of major biotopes or suites of biotopes based on an extrapolation from the representative areas to similar physiographic features. This would be based on extending the distribution patterns obtained from the representative areas to other similar areas as ascertained through a combination of the very broad coverage, hydrographic charts and other map-based assessment of physiographic features. These general maps may be presented as a digitised vector/polygon map rather than a raster map.

Costs: Approximately 5 km square nested within a 20 km square per 5 day week/2 people which would generate 4 weeks of analysis. Total of 6 person/weeks.

Advantages: A cost effective way of rapidly surveying large topographically complex areas. Will produce an inventory of biotopes that may be expected to occur within the survey area.

Suitable for: A broad scale survey but where it is unlikely (because of topographic complexity) that a regular, even coverage will adequately represent the range of features present in the area.

Disadvantages: Distribution patterns only approximate and uncertainty will exist as how well the survey has performed in detecting all biotopes of interest. Nesting maps requires special explanation as to the effect on scale/level of detail. Repeat surveys will need to follow previous tracks closely to avoid differences in interpretation due to positional changes.

Not suitable for: Accurate location of all features of interest or accurate monitoring for change.

8.2.5 Special sampling techniques within broader scale context

Purpose: A small sampling area selected to represent an important habitat or biotope typical of the survey area is surveyed in detail for monitoring potential change or study of dynamics. If nested within a broad scale survey, then the significance of any change can be assessed in a wider spatial context.

AGDS track spacing: Fine track spacing (25-50m) over the test area is nested in a broader survey (100-250m). The broad scale survey should extend over the whole habitat or biotope feature that the test area represents and overlap with neighbouring habitats or biotopes.

Sidescan sonar: Mosaic sidescan images may be required for test area with the AGDS data supplementing the image with point information on sediment characteristics.

Ground truth sampling: The test area sampled intensively using specialist sampling procedures to assess diversity, biomass etc. The broad scale survey would use rapid sampling techniques (possibly including sidescan sonar deployed as a method of ground truthing the AGDS).

Data analysis and outputs: The maps provide a backdrop for the display of non-spatial statistics. If a relationship between a measured variable (e.g., biomass) and acoustic images can be established, then the distribution of these variables can be mapped.

Costs: Approximately 2 km square nested within a 5 km square per 5 day week/2 people which would generate 4 weeks of analysis. Total of 6 person/weeks.

Advantages: Spatial context to special sampling program.

Suitable for: Environmental impact assessment, monitoring and special studies on dynamic processes.

Disadvantages: Limited geographic extent and specialist sampling may lead to high costs of survey and analysis.

Not suitable for: Broad scale coverage.

8.2.6 Fine scale survey with intensive ground truth

Purpose: Detailed survey of a representative transect that lies across important habitat or biotope boundaries or follows an environmental gradient for the purpose of monitoring the change of boundary conditions. This transect is nested in a broad scale survey to assess the likely significance of any change in a broader context.

AGDS track spacing: Track spacing at the minimum that is allowable with the DGPS over the transect with wider track spacing over the whole of the physiographic feature that the transect represents.

Sidescan sonar: Mosaic sidescan images may be required for transect with the AGDS data supplementing the image with point information on sediment characteristics.

Ground truth sampling: Intensive sampling on areas of interest, perhaps using techniques to extract detailed information. Rapid sampling techniques elsewhere (possibly including sidescan sonar deployed as a method of ground truthing the AGDS) if these can identify the biotopes or habitats of interest.

Data analysis and outputs: Analysis may be carried out on track point data rather than interpolated data. Transect areas will need additional analysis, particularly with respect to biotope detail.

Costs: Approximately 2 km square nested within a 5 km square per 5 day week/2 people which would generate 4 weeks of analysis. Total of 6 person/weeks.

Advantages: Gives a wider vista to boundaries not possible without acoustic remote sensing and is an insurance against falsely concluding die-back/encroachment when the reality is a spatially fluctuating boundary.

Suitable for: Monitoring the dynamics of boundaries between biotopes.

Disadvantages: Distribution patterns are limited in resolution by positional errors. Small boundary changes may not be measured with confidence.

Not suitable for: Very precise boundary changes.

8.2.7 Fine scale survey of a few biotopes to map boundaries without intensive ground truthing

Purpose: Detailed survey of a small area, perhaps a representative transect, for the purpose of monitoring the change of boundaries between a limited number of biotopes. This area may be nested in a broad scale survey to assess the likely significance of any change in a broader context.

AGDS track spacing: Track spacing at the minimum that is allowable with the DGPS over the target area with wider track spacing over the larger context area.

Sidescan sonar: Mosaic sidescan images may be required over the target area if topographic features are diagnostic of the biotopes.

Ground truth sampling: A small number of samples will be taken to check the acoustic data/biotope linkage used to determine boundaries.

Data analysis and outputs: Analysis is unlikely to involve training areas or buffers. Instead, a direct link is proved which can be used to discriminate between two or a limited number of acoustically distinct biotopes. The analysis should be carried out on track point data rather than interpolated data to maintain the integrity of the acoustic data and avoid introduction of position errors.

Costs: Approximately 2 km square nested within a 5 km square per 5 day week/2 people which would generate 2 weeks of analysis. Total of 4 person/weeks.

Advantages: Gives the best possible boundary accuracy since it avoids compounding AGDS with ground truth sample positional errors inherent in other forms of analyses.

Suitable for: Monitoring the dynamics of boundaries between a small number of biotopes.

Disadvantages: Small patches will not be mapped accurately due to the limitations of the maximum resolution of the AGDS. Will only be able to work with a small number of acoustically distinct biotopes.

Not suitable for: Mapping a wide range of biotopes.

8.2.8 Replicate sampling in heterogeneous areas

Purpose: To describe the heterogeneity of a survey area and monitor the status of component biotopes of selected test areas.

AGDS track spacing: Track spacing at the minimum that is allowable with the DGPS over the test areas with wider track spacing over the whole of the survey area.

Sidescan sonar: Mosaic sidescan images may be required for test area.

Ground truth sampling: Intensive sampling on areas of interest, perhaps using techniques to extract detailed information. The component biotopes would be described using random sampling within each test area. Rapid sampling techniques elsewhere (possibly including sidescan sonar) if these can identify the biotopes or habitats of interest.

Data analysis and outputs: Heterogeneity measured and modelled for different parts of the survey area and for the selected test areas in particular. The heterogeneity would be estimated from probability measures for each component biotope. If the areas are small, the track point data should be interpreted rather than interpolated data to preserve fine scale variation in the along-track data that would be 'lost' with the averaging involved with interpolation.

Costs: Approximately 5 km square nested within a 20 km square per 5 day week/2 people which would generate 4 weeks of analysis. Total of 6 person/weeks.

Advantages: Allows monitoring programs to be designed for heterogeneous areas where there may be a risk of falsely detecting change when, in fact, sampling at 'fixed points' is very position sensitive. Gives a wider vista to distribution of biotopes not possible without acoustic remote sensing.

Suitable for: Monitoring heterogeneous areas.

Disadvantages: This random stratified sampling survey would not describe fine scale patch dynamics between component biotopes.

Not suitable for: Very precise boundary changes.

8.3 Integrated survey strategy

It is clear from the above that, whilst all the surveys are based on broad scale survey methodology, many of them use other survey techniques as well. In many cases the broad scale survey provides a spatial context for more detailed ecological survey whilst in others detailed sampling 'fleshes out' the broad scale survey by providing a more thorough description of representative locations. Broad scale survey should be regarded as an integral part of a larger survey strategy that spans a range of spatial scales and ecological knowledge.

A generalised scheme for an integrated survey strategy is illustrated in Table 8.1.

Table 8.1: A strategy for integrating broad scale mapping with other types of survey

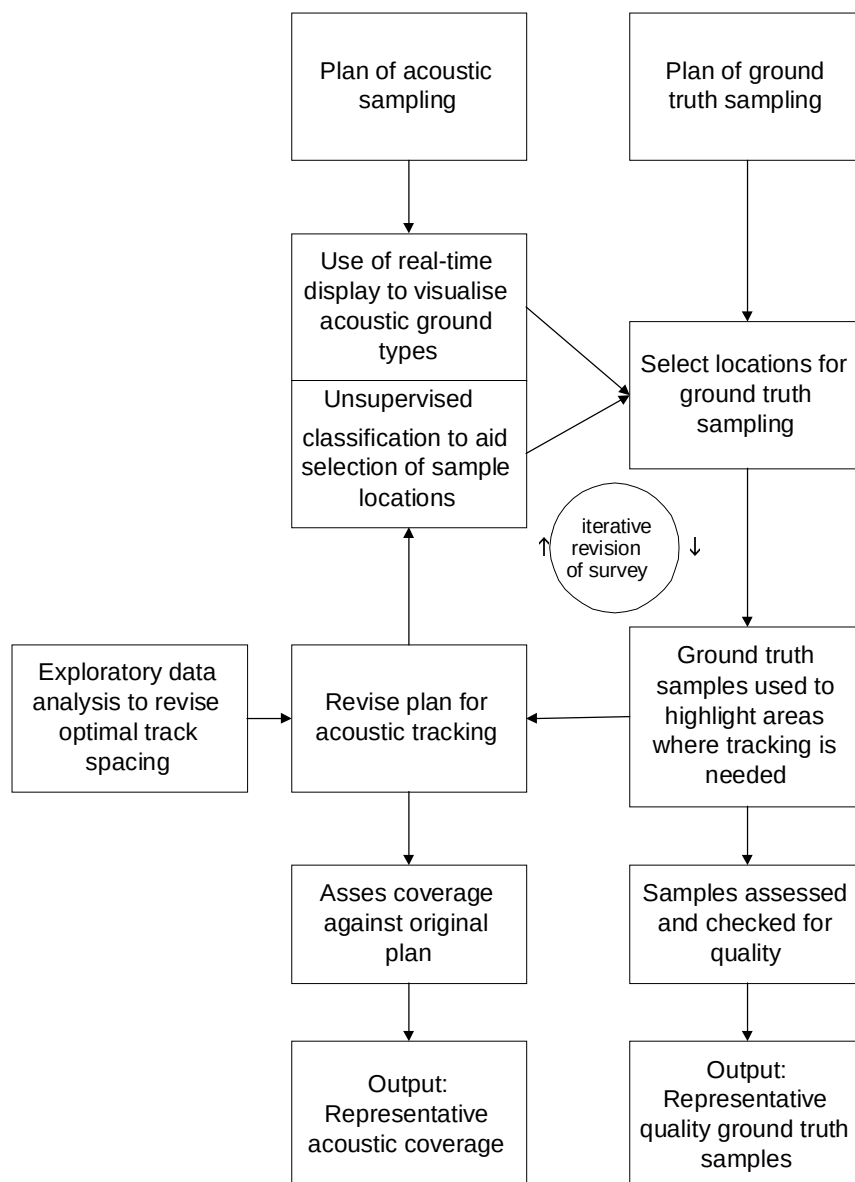
STAGE	PURPOSE	METHODS	OUTPUTS	LINKS TO NEXT STAGE
Defining objectives	1. Define area of interest, management aims and objectives. 2. Review available information to assess the need for further information gathering in the light of management requirements.	1. Review literature, survey reports, maps and databases. 2. Make inventory of sources of data.	1. Description of area of interest in regional/national context. 2. Description of components of the area. 3. An assessment of the adequacy of the data and an identification of the information required.	1. Decide what information will need to be gathered and how this data will be incorporated into a database/display system. 2. Decide if a broad scale survey is appropriate for gathering the required information.
Survey planning	Prepare background material that will be needed for planning the field work of the Phase 1 survey.	1. Consult Hydrographic charts, geological maps, Ordnance survey and any published biotope maps. 2. Establish boundaries for areas of interest. 3. Define any socio-economic features/boundaries of relevance. 4. Georeference any relevant biological data.	1. Thematic maps and linked database 2. An assessment of the likely variety and distribution of broad habitat types; highlight those of potential conservation interest.	1. Define the most cost-effective use of any survey effort that may be required. 2. Decide how survey data will link in with existing database/display system. 3. Prepare documentation for any survey that might be put out for tender.
Broad scale survey	Prepare maps of biotopes based on the analysis of remote imaging linked to ground ground truthing.	1. Collect and analyse acoustic data from the sublittoral. Possibly use satellite, airborne remote sensing (ARS) or aerial photographs for shallow sublittoral in clear water. 2. Use remote viewing or sampling of sublittoral sites for ground truthing. 3. Use the ground ground truthing data or existing biotope information to interpret images to obtain a continuous coverage of biotopes	Maps showing the distribution of biotopes at the appropriate level of detail and scale. 1. Continuous coverage of biotopes or suites of biotopes. 2. A broad-brush assessment of biotope distribution and diversity. Habitat maps may be produced, especially where biotope information is insufficient. 3. A holistic appraisal of the area of interest using biotope maps. This should involve the re-interpretation of biotopes at a broader scale to show benthic landscape features.	Identify where data is required to quantify diversity and species richness through the use of summary and comparative statistics.
Detailed survey	1. Obtain detailed descriptions of sites (including quantitative and semiquantitative data) and the derivation of summary and comparative statistics. 2. Accurate and detailed biotope maps where a need has been identified. 3. Surveys designed to answer specific queries (e.g. estimation of biomass/productivity, distribution of subtle biotope subdivisions, measurement of environmental and biological variables). 4. Monitoring change: assessing the requirement for management or the effectiveness of management strategy.	1. Direct observations using diver-recorders or detailed analysis of faunal composition of grab samples and trawls, sediment analysis etc. 2. More exhaustive sampling to ensure adequate coverage of the area and biotope types indicated in the Phase 1 survey. 3. Quantitative sampling methods, particularly stratified random sampling.	1. Detailed species lists with quantitative or semi-quantitative data. 2. Detailed physical habitat descriptions. 3. Data analysed to produce statistical information. 4. Data georeferenced for inclusion on map-based display systems. 5. Reports on specific surveys.	Data to feed back into database and display system.

8.4 Design of field survey

Planning a survey involves deciding what survey intensity is suitable and whether this should vary over the whole survey area to gain most benefit from the acoustic survey in terms of coverage and to optimise the value of the samples in the interpretation of the acoustic ground types.

The range of ground types expected and complexity of the coastline and submarine features will greatly influence survey intensity and an initial assessment can be made from existing information, hydrographic charts and published maps of sea bed sediments. However, it is to be expected that the survey will need to be modified as it proceeds depending upon findings. The process of revising the field work program is outlined in Figure 8.3.

Figure 8.3. Key tasks in carrying out field survey



8.4.1 Calibration

Surveyors need to be aware of the possibility of variability in the performance of the AGDS during and between surveys. It is the experience of many users that the performance can change over time, although the reasons for this are often obscure. The use of test areas for calibration at the beginning and end of each day's survey is recommended for AGDS surveys by Rukavina and Cadell (1997). This enables a check to be kept on the variability between data sets collected on different days. However, the test areas must be homogeneous in terms of depth, E1 and E2 in order that they can be used as calibration sites since small positional changes in a heterogeneous area would undermine comparisons.

Calibration sites might be impossible to establish in surveys of large areas where the survey vessel may use a variety of harbours so that returning to the calibration sites is impractical. There are, however, other ways in which consistency of performance can be checked. It is reasonable to assume that, whilst the magnitude of the values returned by a particular AGDS might vary between survey or even between days, the AGDS will respond to ground types in a systematic way. This is certainly the experience of *SeaMap*. The general relationship between E1 and E2 can be checked using a scatterplot and, if there are no obvious changes in pattern, different data sets can be standardised (see Section 10).

A variation on this procedure is useful where there are overlapping surveys. The data are checked and standardised in the overlapping region first before the standardisation parameters are applied to the complete data sets.

It should also be pointed out that directly comparable values between surveys of the same area conducted at different times and/or the use of different AGDS are not essential, although desirable. If the data sets are processed separately, then it is the interpretations that are being compared and not the E1 or E2 values (see also Section 13.8 where this important point has been used as the basis for amalgamation of different data sets that could not be standardised).

8.4.2 Position of tracks and track spacing

The intensity of tracking required for an adequate coverage will vary with the size of the features that are to be mapped. Small features will need more intensive tracking than larger more uniform areas. However, some economies can be made when surveying large and complex areas (such as highly indented coasts) by dividing the area into physiographic types and surveying representative examples of each. In effect, this is **nesting** large scale patterns and variability within the overall, small scale picture.

It is difficult to make hard and fast rules for tracking a survey area: so much depends upon specific circumstances and the ideal is often compromised by constraints of time and tide. However, the following points provide some guidelines:

1. **Tracks can be planned to follow expected major environmental trends:** For example, most inshore open coast sites show a major trend from shallow to deep water with increasing distance from the shore. Tracks should run at right angles to the shore in these cases since tracks parallel to the shore will over-emphasise ground associated with particular depth contours and intermediate contours will be under represented.
2. **Track spacing should reflect the scale of feature on the sea floor:** Pinnacles and other irregular topographic features will need closely spaced tracks whilst uniform sediment might be adequately covered with widely spaced tracks. Intensive survey may achieve a track spacing of about 25m as a realistic minimum but would demand a good DGPS. Track spacing of about 250m would be suitable for broad scale survey whilst gaps between tracks greater than about 500m are generally unsatisfactory. Some ground is very variable over small distances (the E1/E2 values differ markedly from one save point to the next) and yet this pattern is found consistently over large areas. In other words, the ground is uniform if viewed at a broad scale. In these circumstances it is unlikely that even closely spaced tracks will reveal fine detail.

3. **Shorelines, islands and narrows should be tracked as close to the shore as possible:** Interpolation between widely spaced tracks does not estimate values very well where some of the tracks end close to the shore. Estimated values for depth, E1 and E2 at locations along the coastline will be calculated from real data that will inevitably include those from deeper water. Interpolation will also 'cross' islands and narrow headlands to incorporate data that are, in reality, separated by land. The resulting interpolated surfaces are obviously unsatisfactory along the shoreline. Whilst it is possible to mitigate these effects through the incorporation of bathymetric data from Admiralty charts (see, for example, Section 12), the process is time consuming and somewhat unsatisfactory. It is far better to plan to track around the coastline and, as far as is possible, islands and narrows.
4. **Representative features can be surveyed:** The idea of broad scale mapping is to cover a lot of ground relatively rapidly. Covering a large area in great detail would be counter to the aims of broad scale mapping. However, it may be desirable to investigate representative features or systems in some detail and **nest** them in a broader picture. For example, an inlet or reef might be selected for detailed survey as representative of many such similar features. Subsequently, it may only be necessary to plot where other such features occur and to briefly check that they are similar to the 'type' feature selected.

8.4.3 Ground truth sampling program

8.4.3.1 Sampling techniques

Biotope mapping requires information on the physical nature of the sea floor and the infauna, epifauna and flora. Remote viewing using a video sledge might be sufficient for sampling habitats and biotopes that can be identified from surface features and epibiota. However, other methods that retrieve samples from the sea floor will be needed for infaunal biotopes and when more detailed data are needed. Boat operated grabs, corers, trawls and dredges can be used and divers might be employed to make observations and take samples. The expense and time involved with using these techniques will place upper limits on the number of samples that can be taken and processed. In addition, analysis of the sample data becomes difficult if too many techniques are used in a survey.

Sidescan sonar can be regarded as a ground truth sampling device if the images are used to construct training sites for the development of acoustic signatures (see Section 19).

8.4.3.2 Selection of sampling locations

The careful selection of ground truth sites, as well as the number of samples taken, will influence map accuracy. It is important that the sampling covers the full range of acoustic ground types found in an area and that a sufficient number of samples are taken from the same ground type.

The intensity of ground truthing will vary with the predictability of the habitat or biotope. If repeated sampling within one ground type indicates that the biotope can be predicted with confidence, then further sampling will add little extra information for the interpretation of the whole acoustic picture. It would be more appropriate to direct effort towards other ground types where prediction is less certain. However, it would be difficult to develop links between acoustic characteristics and biotope type if there were less than 5 samples for each acoustic ground type. However, some ground types may be very variable and prediction based on acoustic characteristics will be weak irrespective of sample number. The use of broader biotope categories or mixed categories might be considered.

Uniform areas of E1/E2/depth characteristics should be chosen where possible so that the sample can be confidently attributed to a set of ground characteristics. However, it is often useful to investigate boundaries between ground types. Some boundaries are quite distinct in terms of changes in E1/E2 and sampling across boundaries will establish what habitat or biotope changes take place.

9 Data analysis and display

Purpose: To introduce the concept of cartographic modelling through the use of a suite of analytical techniques.

9.1 Assembling an analysis and display software system for biotope mapping

There are many stages of analysis between collecting the data and the final production of biotope or habitat maps and this requires a suite of software packages to be assembled in which data can be passed from one to the other as demanded by the scheme for analysis that is being followed. The scheme or flow chart constructed to guide analysis is termed the cartographic model and is valuable in the laying down of guidelines for a step-by-step approach to complex data processing. This assembly of software is termed a 'ring' of modules (Eastman 1997) and these modules may be found within one software package. However, it is more likely that a number of packages will need to be assembled since most have strengths and weaknesses. Compatibility between different programs becomes less of an issue since, with every new version of software published, the ability to exchange data between different formats increases. There are five main components of an analytical system

1. *Database/spreadsheet*: data input, editing, storage and retrieval
2. *Statistical/geostatistical component*: multivariate analysis, interpolation
3. *Image processing: image enhancement*: identification of features and classification
4. *Geographical Information System*: overlay, spatial query, modelling
5. *Cartography*: map production and display

Figure 9.1 shows the main elements to the ring of modules that might be used in a comprehensive computer-assisted data analysis and cartographic system.

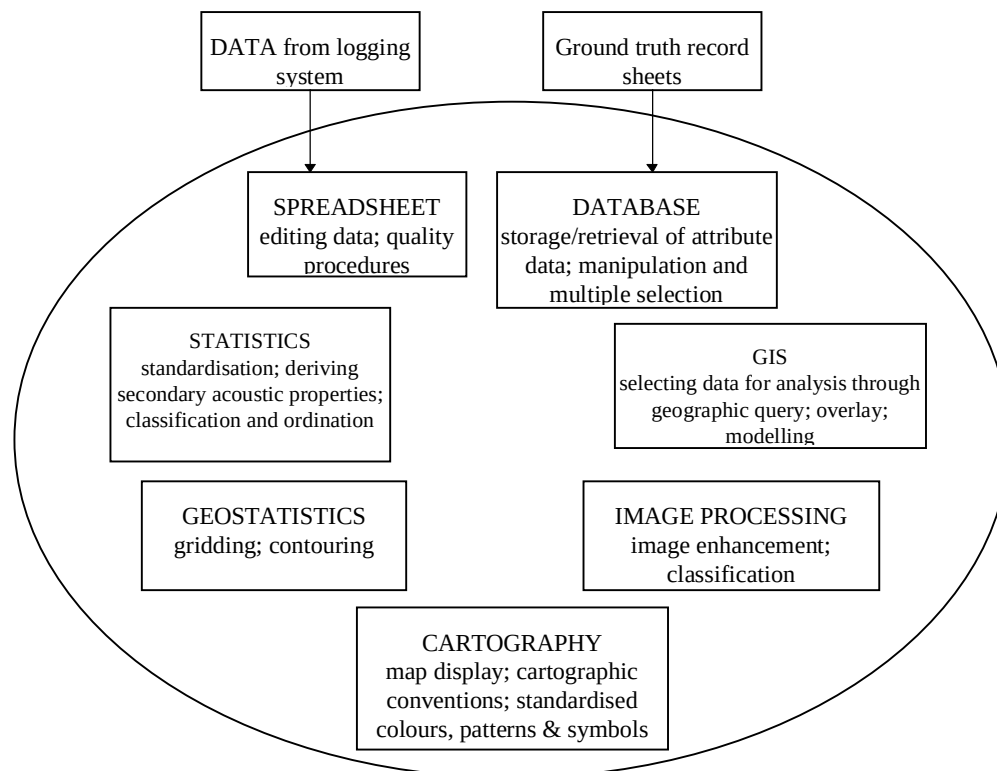


Figure 9.1. The ring of modules in a computer-assisted data analysis and cartographic system.

9.2 Software used for data processing and map preparation

The data are exported from the logging software as ASCII text files and imported into a spreadsheet (*Excel*TM) where various routines are employed to edit out spurious data, derive other variables and generally prepare the data for subsequent analysis and image processing. *Excel* is also used for much of the exploratory data analysis and some statistical treatment and graphical display of data resulting from analyses performed in other software.

The data are imported into software specifically designed for interpolation. *Surfer for Windows*TM is used primarily although *Vertical Mapper*TM is also used in some instances and, specifically, for the creation of three-dimensional models based on bathymetry.

Classification of the acoustic data is performed using *Idrisi for Windows*TM. Other operations are also performed in *Idrisi*, such as image enhancement, cross-tabulation and the calculation of error matrices. The conversion of the *Idrisi* raster images to vector polygons suitable for import into the GIS involved the use of modules in *ArcInfo*TM.

MapInfo Professional is the geographic information system (GIS) in which most of the spatial display (including map design), the creation of vector layers (such as coastlines), the digitising of points and lines from paper charts, spatial editing, spatial query and the creation of vector buffers are performed.

9.3 Data analysis: types of operation

There are many stages in the analysis of the data that involves editing, transformation and enhancement of the data, exploratory and statistical treatment, mathematical and logical overlay of two or more spatial layers of data and the inclusion of parameters into processes that have been determined elsewhere. It is important that analysis does not degenerate into a series of options taken independently of each other, but follows some established guidelines. Setting out the stages in the analytical process is vital so that the procedures can be explained and justified. A schematic diagram of the process is termed a **cartographic model**. This is defined as a graphic representation of the data and analytical procedures used, particularly as related to the spatial nature of the data. This formal approach to analysis helps focus on the 'rules' that have been established to guide the whole process.

There are two different, but interrelated, approaches to data analysis: non-spatial and spatial (geographic). Spatial analysis is defined as any process where the spatial arrangement of the data is explicitly important (Bailey and Gatrell 1995). The two approaches are linked since results of non-spatial relationships within the data can be displayed geographically and, conversely data can be searched for and extracted on the basis of geographic location (geographic query) and then treated non-spatially. Figure 9.2 shows some of the methods used in the analysis of AGDS data either by *SeaMap* or other workers. Most books on statistics treat the subject from a non-spatial perspective. However, texts on GIS explain the fundamentals of spatial analysis (e.g. DeMers (1997); Burroughs and McDonnell (1998) and there are a number of authors who deal specifically with mathematical spatial data analysis (e.g. Cressie (1990); Bailey and Gatrell (1995)).

Data analysis is often separated into three levels: data exploration (including visualisation), data analysis and modelling (Bailey and Gatrell, 1995). However, there is no clear distinction between exploratory and more advanced data analysis and the same software might be used for both types of analysis. For example, contouring of the variables will be carried out during the exploratory phase using the same software involved with more advanced analysis. The distinction lies in the extent to which the full power of the analysis and display tools within software packages are used and the degree of critical scrutiny to which the analysis is subjected.

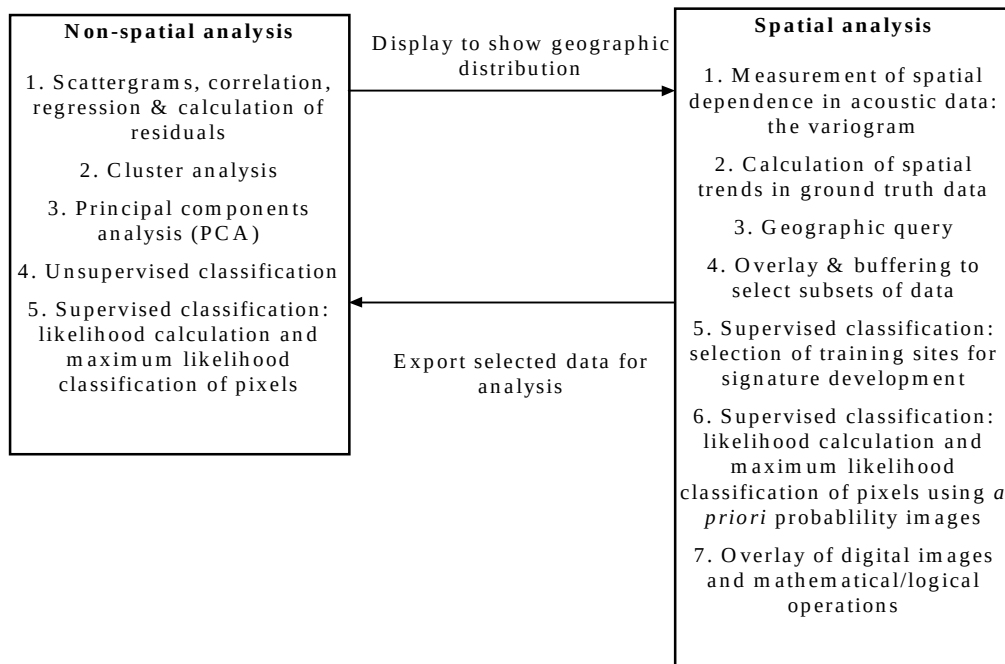


Figure 9.2. Many of the non-spatial and spatial analytical processes used in the interpretation of the acoustic and ground truth data.

9.4 Critical stages in data analysis

Flow diagrams, such as those shown in the Figure 19.1, are vital in establishing the stages required for building up habitat and biotope maps. Constructing the flow diagrams highlights those stages where there are choices to be made between different routes. These stages are critical to the analysis and selecting the most appropriate route requires a good understanding of the nature of the data to be input into a particular stage, the transformation performed by analysis and the nature of the data output from that stage. The key stages in analysis (Figure 8.1) are outlined below and are described in more detail in subsequent Sections.

9.4.1 Preliminary data editing, storage and export

Before the acoustic data can be analysed in detail, the data need careful scrutiny and editing to eliminate spurious data points. This editing can be done within *Microplot* but more advanced editing, including automated searches for dubious records, is more easily performed once the track data have been exported into a spreadsheet, such as *Excel*. Here, depths are corrected to chart datum (see Section 10.2) dubious data are highlighted for possible removal, data are standardised and measures of variability are derived which can be used in subsequent analysis.

9.4.2 Exploratory data analysis and data treatment

The primary purpose of exploratory data analysis is to visualise (a) obvious trends, clusters and features in the acoustic data; (b) the geographic distribution of major acoustic ground types; (c) the links between the ground truth sample characteristics (particularly habitat and biotope type) and the acoustic ground type.

Complex analysis is not needed to show major trends in the distribution of ground types nor how they relate to the ground truth data. This can be done visually with acetate overlays (see Section 5.1.2) and this may be the only analysis that many users of AGDS will require for their purposes, especially if the emphasis is on rapid analysis for day-to-day use.

Visualising the various parameters, such as E1, E2 and depth, is important when justifying and explaining more complex interpretations of the data. As a minimum requirement separate maps are prepared of E1, E2 and depth together with plots of the key features of the ground truth data (for

example; habitat types, sediment grades, biotope types). Overlaying these maps show relationships between the data.

Plots of E1 against E2 show the general relationship between these two variables and are useful in highlighting any obvious clusters. Often these clusters arise because of slightly different responses of *RoxAnn* systems to the same ground types and point up the need for special standardisation or even the removal of sections of tracks from the data set (see Section 10.4). The spatial variability of the acoustic data must also be established through the construction of the variogram so that the ranges for interpolation can be calculated (see Section 11.3).

9.4.3 Interpolation and the creating of the digital images

Interpolation is required to convert the AGDS data into a continuous surface in digital format (Section 12). Various algorithms are available for interpolation and within each there are many parameters to be set. Thus, this vital stage in the image processing procedure needs to be carried out with care.

9.4.4 Image classification

Image classification can take the form of simple grouping E1 and E2 values into 'boxes'. However, more powerful classification processes include unsupervised classification (in which the data are explored for clusters which are then investigated to establish their connection with ground truth data) and supervised classification (in which **training sites** are created from the ground truth data and these used to create acoustic signatures of the habitat or biotope types) (see Section 13).

9.4.5 Modelling

Classification, and in particular the calculation of probability images for different classes of habitat or biotope, is the start of many other procedures designed to incorporate general biological knowledge and local distribution data into the image processing system (see Section 13.5.4).

9.4.6 Cartography

The results of analysis will need editing to produce a map that is easy to read and understand. This requires cartographic skills as well as analytical skills. For example, the results of analysis may have too much fine scale variability or detail to be successfully displayed at the desired scale. Filtering the image may be required to simplify the display. It may be that this detail cannot really be supported by the data and so editing the maps gives the analyst the opportunity to eliminate areas for which there can be little confidence.

10 Preliminary AGDS data editing and treatment

Purpose: To edit, treat and store data in a form which can be imported into software programs necessary for further analysis.

Before the acoustic data can be analysed in detail, the recorded depths are corrected to chart datum, and the acoustic data are scrutinised carefully and edited to eliminate spurious data points. Although this editing can be done within the data-logging software (*Microplot* or *Roxmap*) it is recommended that all editing should be carried out on a copy of the data after they have been exported into a spreadsheet. Additionally, further treatment of the data may be required depending upon the type of analysis anticipated, including standardisation of E1 and E2. The following operations are routinely used by *SeaMap*.

10.1 Additional labelling and identification

Although data are labelled with time and date, it is useful to give each record a unique ID number. If files are combined these ID numbers are run sequentially throughout the whole data set for a survey area. The data are labelled with ID numbers before any editing so that gaps in numbers can be used to indicate where data have subsequently been deleted. Indeed, the deleted data are stored in a separate table and can be restored if necessary.

10.2 Correcting depths to chart datum

Depths are corrected to chart datum by applying corrections calculated from the tidal prediction program using the simplified harmonic method produced by the UK Hydrographic Office (Anon 1991)). The corrections are applied at time intervals between 10-30 minutes. It should be stressed that the resulting depths may not be very accurate and may conflict with the soundings on Admiralty hydrographic charts. This must be expected since atmospheric conditions affect tides and, additionally, the nearest reference port may be some distance away. However, using 10 minute intervals eliminates steps in the depth track records that can be apparent if longer time intervals are used. This is particularly important for the construction of digital elevation models (see Section 16).

10.3 Editing track data

All editing is performed on a copy of the original data and no records are irretrievably lost through the editing process. Although detection of dubious data can be automated, deletion of data should only be undertaken with great caution. Automated processes are used to highlight possible errors in the data for visual inspection, rather than for automatic deletion.

10.3.1 Editing data on the basis of positional uncertainty or boat speed

Acoustic data are inspected to eliminate dubious points related either to large skips in position caused by GPS error or to spurious depth, E1 and E2 values. Positional errors occur even when using DGPS since the differential signal may be lost on occasions. Small positional errors (10m) are difficult to identify but larger skips (of up to 500m) are usually transitory and cause obvious spikes on the track. Additionally, values of E1 and E2 can increase when the speed of the vessel is very low or when the vessel comes to a stop, although this effect is more apparent on some vessels than others. Track data associated with low vessel speed (below 2 knots) are, therefore, deleted.

An automatic procedure is applied to the eastings and northings to highlight sections of track, (1) where there are large skips in position and (2) where there is little change in the position of the survey vessel. The distance between two consecutive points is calculated and positions highlighted where there is a jump in distance much greater than expected from the speed of the vessel or where the vessel slows to a speed below the acceptable minimum for recording AGDS (about 1 metre per second). Note that if a record is selected for deletion, then the whole record, including depth, E1 and E2, is deleted.

10.3.2 Editing data on the basis of erratic changes in depth

An automatic procedure is applied to the depth data to highlight those sections of track where changes in depth are erratic. A record is compared to the average value of the two previous track points together with the two following points. Track points where a large difference ($>5\text{m}$) is calculated are highlighted and inspected. If a point appeared to be out of step with its neighbours, then it is deleted. Note that the whole record, including E1 and E2, is deleted.

10.3.3 Erratic changes in E1 and E2 values

Spurious changes in E1 and E2 are harder to detect than spurious position and depth records and their causes are numerous. For example, values for E1 and E2 can change when the vessel alters course. Certain sea states might affect the values differently depending upon the ship's course. The apparatus has also been known to return lower values than expected on some tracks, often after many hours of recording acoustic data possibly due to variability in the transducer (Collins and Voulgaris 1993). A second echo sounder operating on a similar frequency to the *RoxAnn* unit on the vessel may cause interference. Sometimes this interference is obvious but more insidious interference may only come to light during post-survey analysis.

There is no automatic way of highlighting erratic E1 and E2 values. A careful watch is kept on the track during field recording with reference to the echo sounder screen for reassurance. This alerts the operator to any potential problems so that measures can be taken to rectify the situation. Plotting E1 against E2 may reveal clusters of potentially dubious data that lie outside the expected distribution of the data. These selected points are labelled and then displayed geographically to see if their distribution reveals a likely cause. For example, a cluster might be associated with changes in the ship's course or any other of the potential factors listed above. Plotting the track data showing ranges of E1 or E2 values can reveal single anomalous points which should be investigated and removed if necessary. It must be stressed that a small number of track points which are outstandingly different from others in the neighbouring region can introduce serious errors in future analysis, particularly when interpolating between data and, therefore, it is important that they are detected.

The use of a test area for day-to-day calibration of the AGDS would also alert operators to potential problems with stochastic differences in E1 and E2 values. This is discussed in Section 8.4.1.

10.4 Standardisation of E1 and E2

Standardisation is recommended when data from surveys of the same area obtained on separate dates and/or using different vessels are combined for processing or compared. E1 and E2 cannot be standardised using the maximum value since every data set has a small number of high values that would unduly influence the standardisation. Instead, the data can be standardised by dividing all the records by the 95th percentile value (allowing for a wide scatter of values in the upper 5% band). This value is found by sorting the records for E1 (the process being repeated for E2) in descending order of magnitude and finding the record that separates the first 5% of the records from the subsequent 95%. The value for E1 for this record is used for standardisation. This has been found to produce a good match between surveys and works best when tracks over the same area can be used for inter-calibration. These standardised values are used in place of the original values.

Note that data from consecutive days on one survey are amalgamated prior to standardisation, having first ascertained, using a scatter plot, that there is no obvious day-to-day variability. This ensures that the data encompass the widest possible range of habitats prior to standardisation.

In some cases, data obtained from systems with very similar ground-response characteristics can be amalgamated even though the magnitude of the values varies between systems. In other cases, the response characteristics are different and standardisation will not make the data sets directly comparable. Other methods for amalgamating surveys under these conditions have been devised (see Section 13.8).

10.5 Deriving secondary variables

10.5.1 Principal components

E1 and E2 are not independent variables. In theory, the maximum information from E1 and E2 is obtained by performing a simple principal components transformation on the data. The new axes are $E1 + E2$ and $E1 - E2$ assuming there to be a linear relationship between E1 and E2. However, experience has shown that image classification using the transformed data are very little different from using the untransformed E1 and E2 values. For this reason PCA transformation has been little used.

10.5.2 Variability

Other information can be extracted from the track data that is used to help ground discrimination. For example, some ground is more variable than others and this can be measured by the variability between sequential track data points (i.e. $E1_n - E1_{n+1}$; $E2_n - E2_{n+1}$). The spatial distribution in variability in E1 and E2 are very similar and a double square root transformation of variability of $E1 + E2$ produces a useful index for summarising acoustic variability. Variability over larger sections of track, using variance or standard deviation, will give a measure of heterogeneity over different spatial scales and is the basis of the variogram (see Section 11.3).

10.6 Summary of data-file structure

The following are the fields (columns) created and retained for further analysis:-

Longitude	Decimal degrees, OSGB36 datum
Latitude	Decimal degrees, OSGB36 datum
Eastings	Metres, OSGB36 datum
Northings	Metres, OSGB36 datum
Time	24hr clock, hr/min/sec
Date	Month/day/year
E1 raw	Unstandardised
E2 raw	Unstandardised
Depth	Uncorrected
Depth	Corrected to chart datum
E1 standardised	95 th percentile
E2 standardised	95 th percentile
E1/E2 variability	Calculated from standardised data

11 Exploratory data analysis and data treatment

Purpose: To visualise (a) obvious trends, clusters and features in the acoustic data; (b) the geographic distribution of major acoustic ground types; (c) the links between the ground truth sample characteristics (biotope type) and the acoustic ground type.

11.1 Exploring relationships between E1 and E2

Scatter plots of E1 against E2 show that they are correlated with a broad trend along a gradient of ‘smooth/soft’ ground to ‘rough/hard’ ground (low to high values of both E1 and E2). For the majority of paired data points, E1 is a good predictor of E2 and the strong correlation between E1 and E2 is the reason why PCA transformed data may be preferred to untransformed E1 and E2 data.

It is unusual to find any obvious and distinct clusters in large data sets. Instead, scatter plots show a cloud of points with the highest density lying along the regression line with indistinct clusters lying above and below this line (Figures 11.1a & b). This suggests that the acoustic ground types grade into each other along a continuum without any clear-cut divisions between them. However, clusters may be apparent when a limited range of ground types are sampled (e.g. sand and cobble) but become obscured when further data from a wider range of ground types are incorporated into the full data set. A more critical examination of clusters and breaks in the data can be performed using cluster analysis and underpins unsupervised classification (see Section 13.4).

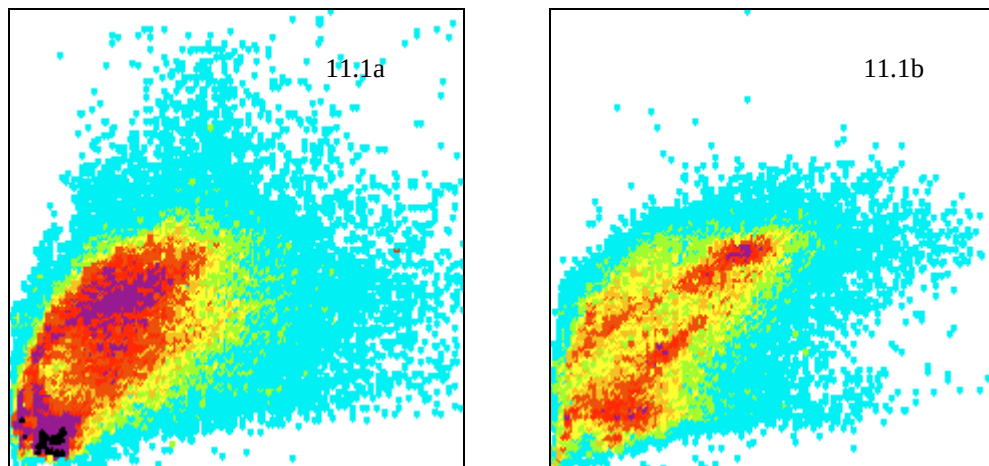


Figure 11.1. Plots of E1 against E2 showing density of the point data for the Northumberland coast (11.1a) and the Wash, Lincolnshire and north Norfolk coast (11.1b).

11.2 Classifying acoustic ground types

It is worthwhile classifying the track data into ranges of E1/E2 to keep a check on subsequent analysis performed on interpolated data. Classifying E1/E2 can be performed in *Microplot* (see Section 4). However, classification is most easily performed in image processing software packages such as *Idrisi*. The number of classes of E1 and E2 and the ranges of values used is a matter of convenience. Five classes of both E1 and E2 is considered to be a workable number (corresponding to ‘very smooth, smooth, moderately rough, rough and very rough’ and ‘very soft, soft, moderately hard, hard and very hard’ respectively). The range is defined by dividing the whole range from zero to the 95th percentile value into 5 equal classes (allowing the top 5% to ‘spill over’ into some extra classes). Thus, there is a

The track records are coded according to the box in which they lay and these coded values can be interpolated (using nearest neighbour interpolation – see Section 13.2) to give a continuous image – a rough but useful form of unsupervised classification .

11.3 Spatial patterns in the AGDS data: The variogram

It is important that some knowledge of the spatial patterns in the AGDS data is obtained so that various treatments of the data can be justified. In particular the track data need to be converted into areal coverages, and the optimal way to interpolate depends upon the nature of the spatial pattern in the data.

Patterns can be divided into three components operating at different scales-

1. **Large scale variability:** At the finest scale (largest scale) it would be expected that there will be variation between one AGDS record and the next irrespective of how little distance separates the points. This is a stochastic variability that is often referred to as 'noise'. If enough records were made at a point, then the distribution would be normally distributed around a mean value. (In fact, there may be spatial patterns, but they are at a finer scale than the remote sensing technology can determine with the result that small positional changes give a apparently random changes in values.)
2. **Medium scale spatial patterns:** The closer two records are the more likely they are to be similar. The greater the separation between points, the less likely they are to be similar. The way in which this correlation declines with increasing distance can be generalised for a whole data set: The pattern provides a 'model' for the area as a whole, although there may well be local differences in the way the data behave over large survey areas. Nevertheless, spatial pattern at this scale is not so much dependent on exactly where the two points are in the survey area, but upon the distance that separates them.
3. **Small scale trends:** Eventually, a distance will be reached where there is no longer any correlation between two points that depends on separation distance ('lag'). At this scale the broad trends may be apparent when many local values are averaged (e.g., inshore/offshore, north/south, shelter/exposure). In a sense, the medium scale spatial pattern becomes 'noise' for the broad scale trends just as the hidden pattern in the large scale variability created noise for the medium scale patterns.

A graph showing the variance between points at increasing separation illustrates the nature of the spatial correlation within the data. These graphs are termed 'variograms'. The variance (average of the sum of squares of the differences) between values at different 'lags' (the term used to denote the distance that separates two points) are plotted against the lag distances (Burroughs and McDonnell, 1998). The differences in, for example, E1 values are calculated between records as follows:- $Z_i - Z_{i+2}$; $Z_i - Z_{i+3}$; $Z_i - Z_{i+4}$ $Z_i - Z_{i+n}$. These calculations are performed on the whole data set and the average variance found for each lag distance. The lag distances can be estimated from the average distance between consecutive points (Pythagorean distance).

A general form of the variogram shown in Figure 11.3 which illustrates the points described above. The shape of the relationship between variance and lag distance illustrated is only one of many models that might be found in practice (see Burroughs and Donnell 1998), but is the general shape found with AGDS data. The graph starts at a minimum distance representing the maximum spatial resolution of the AGDS (e.g., 25m) and the variance below this distance is 'noise' (see '1' above). Variance may then rise quite steeply initially over the range of spatial correlation indicating that the strongest links between data are over short distances and correlation dwindles rapidly as distance increases. The point where the graph levels off (the 'sill') marks the maximum distance over which spatial correlation can be detected and is called the 'range'. However, the correlation is very weak at distances approaching the maximum range and a smaller working range for interpolation may be chosen .

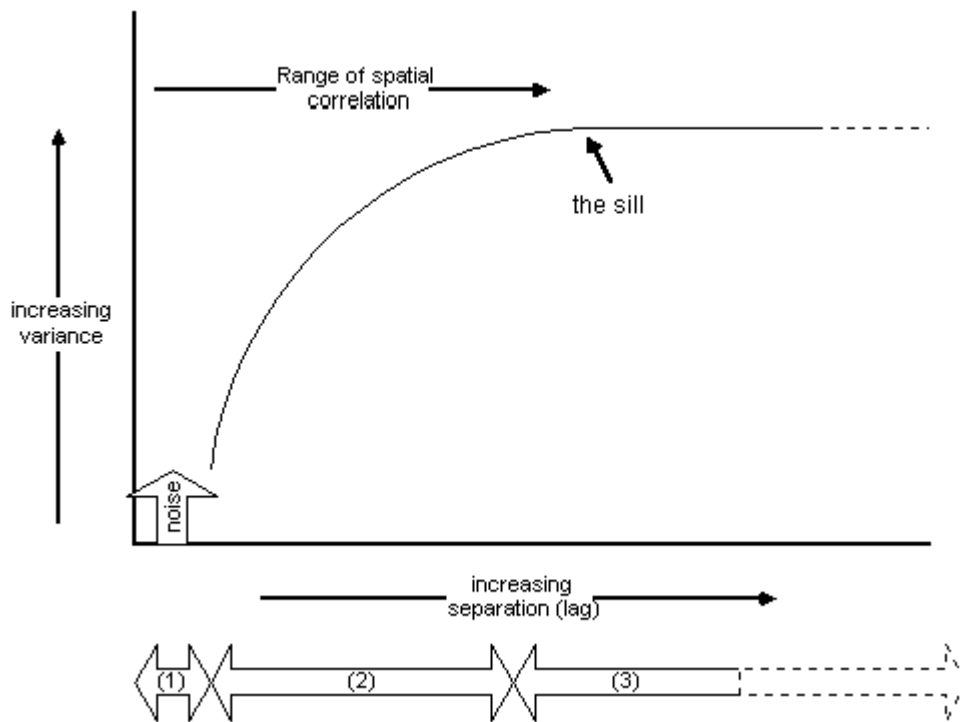


Figure 11.3. A typical variogram showing the major features and terminology. The numbers in the block-arrows at the foot of the figure refer to the scales of spatial pattern described in the text.

Why is an exploration of these spatial patterns important? Firstly, it is reassuring to demonstrate that the variance due to 'noise' is small in comparison to the overall variance of the data set. If this was not the case, then any attempt to interpret the AGDS data would be doomed to failure. Secondly, it is fundamental to broad scale mapping that some generalisation about the distribution of acoustic characteristics and hence habitats and biotopes can be made at a local, large scale throughout a survey area. It is assumed that some spatial dependence operates. If this were not the case, then each location on the map would be entirely independent of its neighbours and broad scale mapping would be impossible. However, whilst it may be safe to make an assumption of a degree of medium scale spatial correlation, it is only by modelling the nature of this spatial pattern interpolation, statements about map resolution and cartographic detail can be justified. Measurements taken from the variogram are used to set the parameters for interpolation and/or indicate those areas where it is likely that the estimated values are simply the local average (see Sections 5.2.2 and the discussion that follows in Section 12).

The following two graphs (Figures 11.4a & 11.4b) are based on data from the Wash (where the ground can be variable over short distances) and the Lincolnshire coast (where the ground is more homogeneous). The first graph (Figure 11.4a) presents the standard layout of the variogram (Burroughs and McDonnell 1998) with the lag shown on a linear scale of distance. In the second graph (Figure 11.4b) the lag is on a log scale which expands the pattern of spatial variability over the smaller distance ranges. These graphs indicate that the spatial correlation is quite high over small distances and increases exponentially over moderate distances. The variance then levels off towards a maximum (the 'sill') and this point on the graph indicates the maximum range over which interpolation can be justified. It is probably reasonable to assume that this along-track variogram would also hold true for points between tracks and can be used to justify distances over which interpolation can be taken to give more information than simply the local mean.

Figure 11.4 indicates that a good degree of spatial correlation might be assumed for distances of about 500m for both the Wash and the Lincolnshire coast (giving a variance of about 0.05 in both cases) and an upper range of 2.5km, a point at which the variance is about one half the maximum in both data

sets. This assumes that there would be some advantage to interpolation over estimation using the local average if the acceptable variance was about one half that of the maximum. The corresponding inter-track distances are twice these distances.

Figure 11.4a. Variogram for the Wash and Lincolnshire E1 track data using a linear scale for the lag.

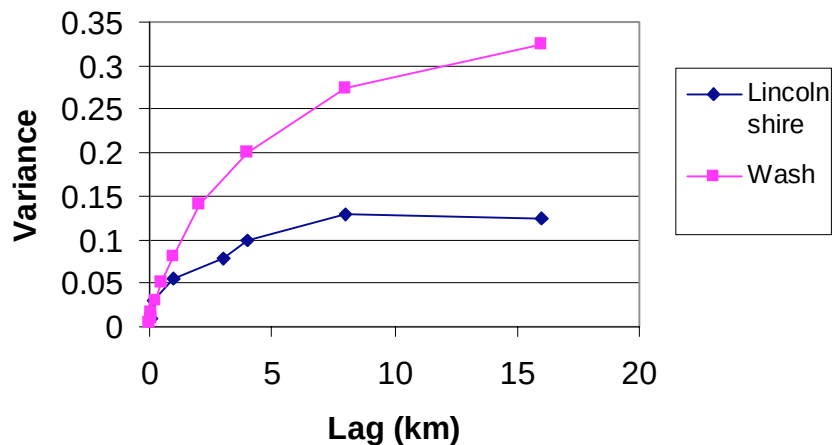
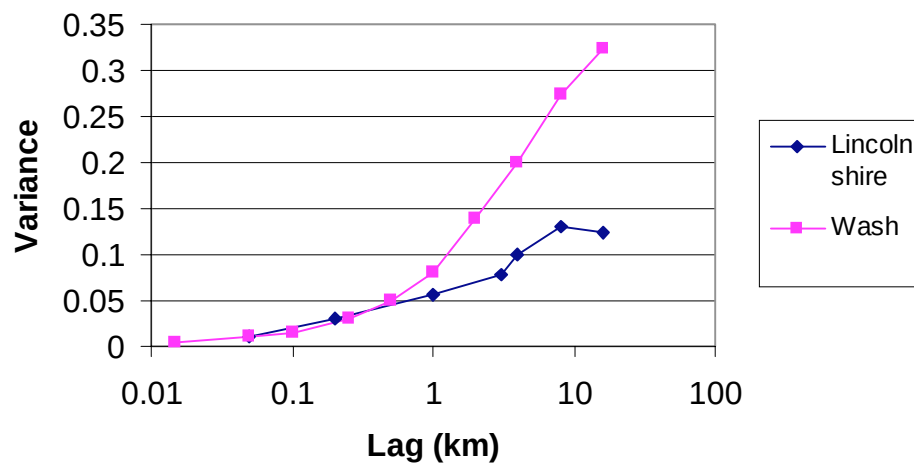


Figure 11.4b. Variogram for the Wash and Lincolnshire E1 track data using a logarithmic scale for the lag.



These ranges are guidelines for the survey area and it is to be expected that there are finer scale differences in the variogram throughout the survey area. Nevertheless, these ranges are useful as a means of determining areas where the uncertainty in the interpolated values might be expected to be high. A zone of 750m and 2.5km either side of the track has been created and is shown as dark and light shaded areas respectively in Figure 11.5. This indicates that the Wash and southern section of the offshore survey area have been covered adequately, but that the inter-track distance has exceeded the guideline distance for the north Lincolnshire coast.



Figure 11.5. Zones around track data indicating areas where interpolation could be expected to return reliable estimated values for E1 and E2 (dark shade) and less reliable estimates (pale shade). Estimates for areas outside of these zones would be the local non-spatial average (see text).

A second example is taken from an intensive survey of a small area from Loch Maddy, North Uist in the Outer Hebrides (Foster-Smith *et al.* 1999). The variograms follow a similar pattern, although the ranges are far smaller due to the finer scale heterogeneity of the area and compactness of the site.

Figure 11.6: Variogram for Sponnish Harbour and Outer part of Loch Maddy

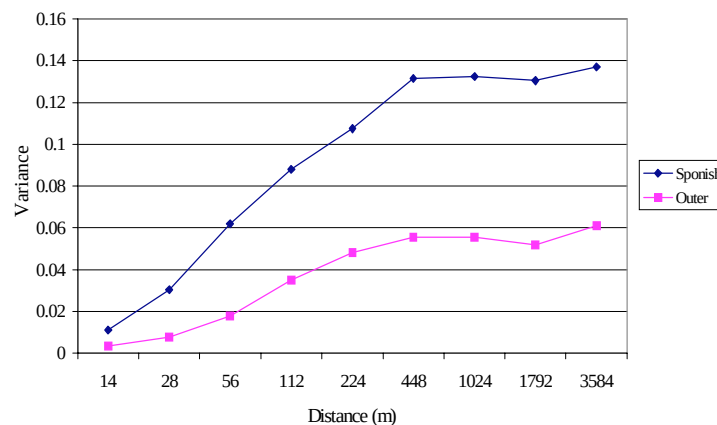


Figure 11.6 is based on data from (a) Sponish Harbour (where the ground is presumed to be variable over short distances) and (b) the outer loch (where the ground is more homogeneous). If very reliable estimated values are required (e.g., 0.2 x maximum variance), then a range of about 25m would be required for Sponish Harbour but 60m for the outer Loch. These ranges are useful as a means of expressing areas where the uncertainty in the interpolated values might be expected to be high because of distance away from the track. The differences in the variances also gives an indication on the variability in the acoustic reflectivity of the ground, being greater in Sponish Harbour than for the outer loch.

12 Interpolation and resolution

Purpose: To derive a continuous coverage of the survey area from irregularly spaced point data

12.1 The purpose of interpolation

Interpolation is required to view the variables as **contours** or **surfaces** (Cressie 1990; Keckler 1994; Burroughs and McDonnell 1998). Contours are lines joining equal values and are a familiar way of representing continuous variables. Contour maps are very useful for showing how individual variables are distributed over the survey area and lines enclose areas of defined ranges bounded by upper and lower contours (DeMers 1997).

However, a more accurate and mathematically versatile way of representing interpolated values is by a continuous surface. Every point has its own associated value rather than sharing the attribute with all other points enclosed by contour lines. These surfaces are actually composed of a grid of rows and columns of values and this grid can be regarded as a digital image.

Interpolation, therefore, is common to both contouring and the creation of digital images. Computer programs, such as *Surfer* and *Vertical Mapper*, generate a regular grid of point values from the original track data and offer many different algorithms for interpolation. In essence, a grid is placed over the geographic display of track data and values computed for a point at every intersection of rows and columns. These values are computed from a search of real data around each grid intersection. The analyst must select the most appropriate interpolation algorithm, set the grid spacing, search pattern and other parameters.

Contours are derived from the grid by drawing lines between nodal points that separate values that lie above and below the required contour values. The resulting contour lines will be composed of numerous straight lines that can then be 'smoothed' by the computer program to produce more pleasing lines.

In the creation of a digital **raster image** each grid node becomes a **centroid** of a small tile-shaped region. The tiles are contiguous so that a complete coverage is composed of rows and columns of tiles corresponding to the grid nodes. During exploratory data analysis contoured maps and raster images are complementary to each other and emphasise different aspects to the data. Whilst it is easier to read values off a contour map, raster images convey the spatial variability of the values which is often closer to the true nature of the ground.

12.2 Interpolation of AGDS data

The track data from the acoustic ground discrimination system are somewhat unusual in the context of grid generation since there are numerous more or less regularly spaced data points along the track and relatively large gaps between tracks. Values computed for grid nodes that lie near to tracks are extremely well supported by the data whilst nodes between tracks may be poorly supported. The estimated value at a location depends upon (a) the distance of the location from any real data and (b) the pattern of spatial variability of the real data. The variogram is used to indicate how spatial correlation varies with distance and the appropriate interpolation algorithm and parameters are selected accordingly. Figure 12.1 shows the main parameters that must be set when estimating a value at a grid point. These are :-

- Interpolation algorithm (model of distance-weighting to be used)
- Search radius (distance over which interpolation is to be performed)
- The number of points to be used in the calculation (maximum and minimum)

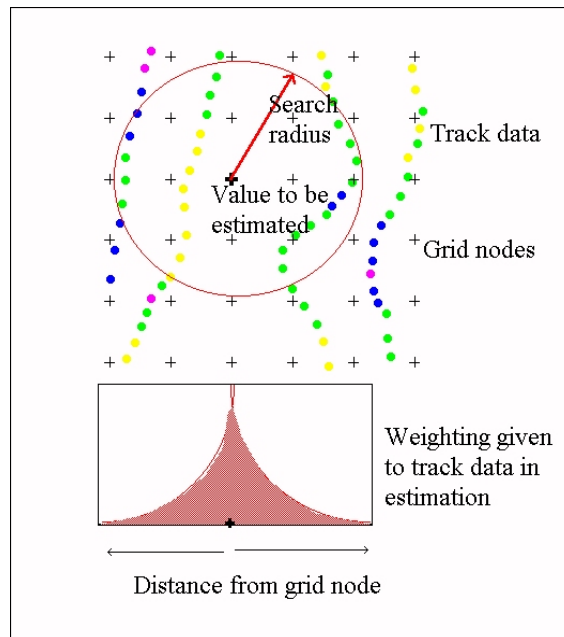


Figure 12.1. A small section of track with grid superimposed. The bold cross indicates the grid node for which an estimated value is to be calculated. A search radius is set by the range of the variogram and the distance weighting used depends upon the interpolation algorithm chosen. An example of inverse distance square weighting is shown in the box below the track data. However, not all data within that radius need be included if the maximum number of data points is set below the 26 points (in the example illustrated) that lie within the circle.

12.2.1 Choice of algorithm

Of the algorithms that are commonly available, two are routinely used in the interpolation of AGDS data. The simplest of these is inverse distance-weighted interpolation. The assumption is that the estimated value is a distance-weighted average of the real data that occur within a search radius set by the analyst. The interpolation can be set to be linear or, more usually, an inverse distance square function is selected.

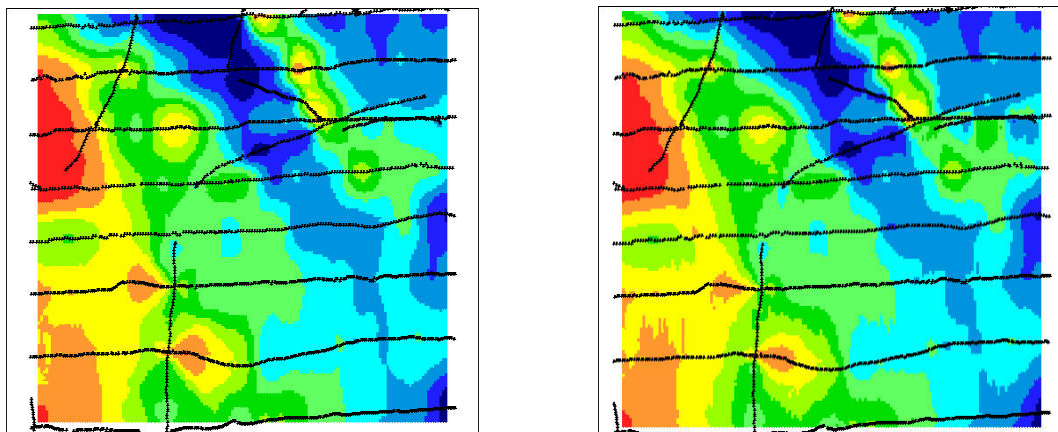


Figure 12.2 a & b. Interpolated depth data, with tracks superimposed, for an area of approximately 5km by 5km off the north Norfolk coast. Inverse distance square was used for Figure 12.2a and Kriging in 12.2b.

Kriging is an alternative distance-weighted interpolator that optimises the use of the model of spatial correlation as shown by the variogram (Burroughs and McDonnell 1998). Because of this Kriging is thought to give a more accurate estimation of interpolated values.

In practice, as the example in Figure 12.2 illustrates, the difference is not marked where the inter-track spacing is well within the range of the data. Since Kriging makes use of many assumptions about the spatial nature of the data that are hard to verify, it is more prudent to use the simpler method of inverse distance.

12.2.2 Search pattern

It should be pointed out that the range as defined by the variogram does not preclude the estimation of point grid values at greater distances away from real data. However, it does mean that these values do not depend upon the spatial arrangement of the real data points in the neighbourhood. The best estimate is, in these situations, the **local mean** of all points within the search radius. In fact, there is a gradual shift from estimates based on spatial correlation to those on local means as distance from real data increases in a manner that depends on the variogram model.

Where the track spacing is wide and the interpolation is approaching the limits of the range of the variogram model, the interpolation algorithm and parameters can make a great difference to the estimated grid values. The analyst must adjust the search pattern to meet the average conditions of the whole survey or part of the area if it has been divided into sections for interpolation. The main parameters to be set are the algorithm and distance weighting function (or variogram model if Kriging is the interpolator used), search radius, the number of records to include in the calculation and whether the search should require data from all points of the compass or simply the nearest data.

It is tempting, since the spatial discrimination along the track is high, to use a closely spaced grid with a small search radius. Estimates close to track data will reflect the fine scale variability along the track although there will, inevitably, be some smoothing since the estimates are an average of a number of records. However, as distance increases away from the track two effects may become apparent, depending upon the algorithm used, the search radius and minimum number of points set.

- If linear distance is selected and the number of records set high, then the resulting estimated values will smooth variability and this will be more apparent with increasing distance from real data. The search circles for each point estimate will overlap so that some data will be included into the estimates of many grid points. In effect, the interpolation is acting as though it was a **'moving window'** filter which returns a smoothed image.
- If an inverse distance square interpolation algorithm is used and a small number of records accepted then the estimates will more closely reflect the values of the nearest real data and the result will be very similar to the use of **'nearest neighbour interpolation'** in which averaging is underplayed.

In practice, a compromise needs to be reached for the choice of search radius and the number of real data accepted such that valid variability in the track data is not too heavily smoothed, but not so restricted that any local spatial pattern is, in effect, ignored. The choice will clearly affect the interpretation of the track data (see example in Figure 12.3).

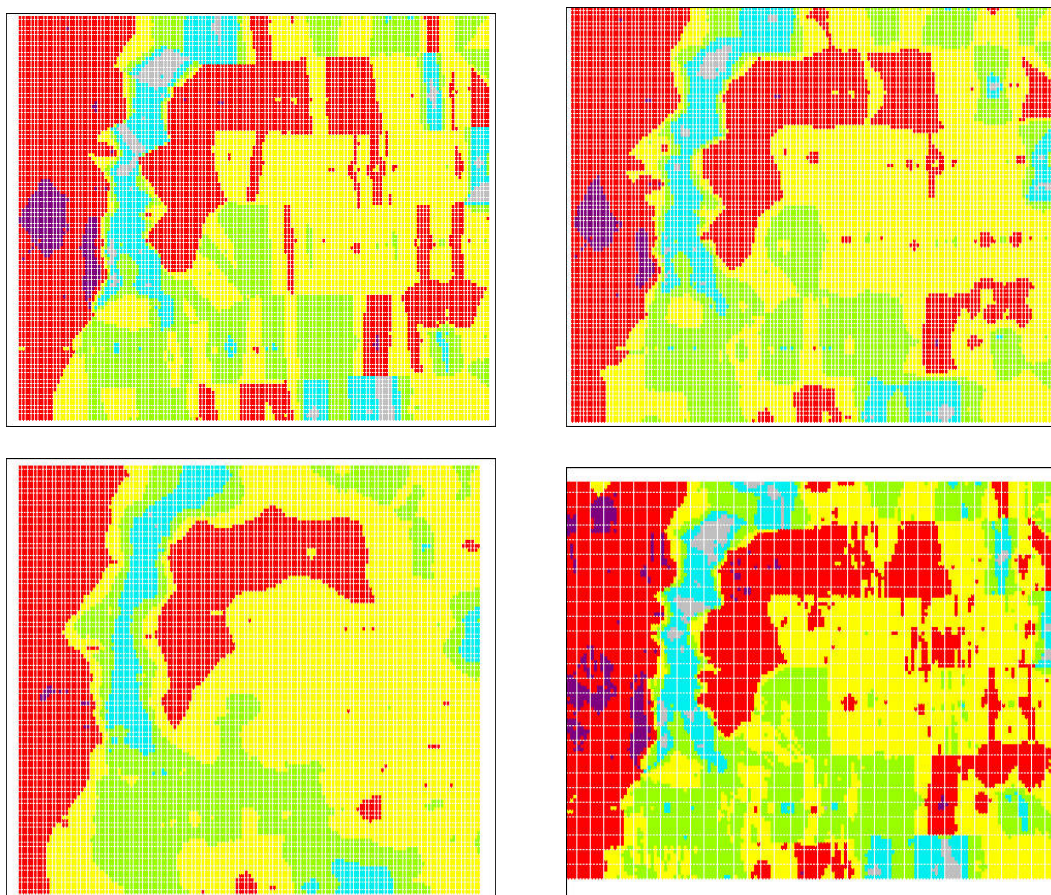


Figure 12.3. The effect of changing the parameters is demonstrated using this example of interpolation of E1 data from widely spaced track data taken from the north Lincolnshire coast in the locality of Silver Pit. Purple and red colours represent high values of E1 and blue and pale blue represent low values. The area is approximately 15km by 15km. There is no single correct solution and the final choice might depend upon experimenting with different parameters and selecting those that smooth some of the obvious jumps in values without losing real features from the image. In the first example (top left) inverse distance square has been used with a search radius of 3km, but accepting the first 20 points in the estimation of interpolated values. The result is very 'blocky'. Increasing the number of points in the calculation to 100 produces a smoother effect (top right) and using a linear function rather than inverse distance square smooths the data further, although many of the more extreme low E1 values are poorly represented (bottom left). The last example (bottom right) uses Kriging as the interpolator with the same search radius of 3km and using 100 points in the estimation.

A working compromise in situations where the tracks are variably, but sometimes widely, spaced might be to select inverse distance to the power of 1.5 or 2, a search radius sufficient to bridge all tracks (but no larger). The number of records accepted and search pattern might be set so that there is a reasonable expectation that a representative selection of records are included from the surrounding tracks and not just the nearest tracks.

12.2.3 Variability of AGDS data

Estimating values between tracks is not likely to be a serious problem for depth data, where depths vary gradually along the track and intermediate depths can be safely assumed to occur between two adjacent points. However, E1 and E2 values can vary markedly along a track so that much information may be lost by the averaging involved with interpolation. Adjacent tracks may differ markedly and intermediate values between them cannot be assumed. It may be decided that the

information lost due to averaging is not great and that interpolation should proceed in the same way as for depth. However, a measure of variability can be used to compensate for smoothing. A measure of heterogeneity can be derived from the track data (see Section 10.5.2) which can be mapped in the same way as E1 and E2 and used as a third acoustic parameter.

12.2.4 Categorical data

Interpolation works well with continuous variables, but interpolation of classed (categorical) track data is problematic because intermediate values between two classes are not permissible. There are occasions where categorical data are required to be interpolated, such as the interpolation of the colour box codes assigned to records (see Section 13.2). However, there are interpolation procedures that are suitable for classed data, such as **nearest neighbour** methods (Keckler 1997) where the value at a grid point is taken from the nearest real value.

12.3 Grid spacing

Resolution and spatial detail is primarily determined by track spacing and is also a function of the heterogeneity of the sea floor. It is all too easy to introduce a spurious level of detail into a map that is not substantiated by the data since interpolation procedures will generate values even for locations well away from any track data.

12.3.1 Track spacing and resolution

Closely spaced tracks will support a more detailed map than widely spaced tracks and more intensive ground truthing will directly confirm the existence of many more of the biotope or habitat polygons on the map. There is, however, no simple direct link between track spacing and resolution. Heterogeneity must be considered and this may well vary throughout the survey area. One way of measuring heterogeneity of the acoustic data is through the use of the variogram. This shows graphically how the correlation between records changes with increasing distance between points and can be used to set a range beyond which interpolation will not give better estimates than the local mean. This can be used to judge if the track spacing is adequate for interpolation. The use of the variogram is discussed in more detail in Section 15.4.

Clearly the maximum resolution is determined by the remote sensing technology. But track spacing and the variogram model will, in most surveys, set the maximum resolution at a higher level than the technical maximum even when track spacing is very narrow (25m). The resolution of a map is unlikely to be the maximum, especially with broad scale surveys, because of the time and cost involved with closely spaced tracks. Often, track spacing will vary over the map with some areas being intensively surveyed compared to others and this creates additional problems. If a single map is to be produced for all the data, then the minimum resolution should be used. However, separate maps may be prepared for the smaller areas that are more intensively surveyed.

12.3.2 Grid size, accuracy and resolution

Grid spacing in itself does not affect the accuracy of the estimates. Choosing a fine grid spacing simply adds more estimated grid values to an image. However, grid spacing should reflect the resolution of an image. A fine grid spacing will produce an image with smoother boundaries to the features identified, but there is a cost in terms of the size of the images and the processing times required for any operation. In addition, it is pointless to fill an image up with numerous grid points that have been estimated either from local averages at one extreme (where the estimated values are some distance from any track data) or duplicating real data (where the estimated values lie very close to track that interpolation simply reproduces the track data in a slightly different position). However, valuable information is lost when the grid spacing is very wide relative to the track data so that only a meagre selection of estimated values are calculated.

12.4 Using bathymetric data to assist interpolation

There may be situations where the track data are too widely separated to map large topographic features that are shown on hydrographic charts. For example, a ridge might lie north west/south east across a small number of AGDS data tracks that run east/west. The tracks will show this feature as

short lengths of track of elevated bathymetry that are offset from track to track. Interpolation does not cope well with this type of data and the result will look like a series of offset blocks.

Bathymetric data from charts will improve the depth data from the AGDS and these data are easy to incorporate into the data set. However, can bathymetric data also help to improve the interpolation of E1 and E2 values? In some environments, such as the Wash, the survey area is very hydrodynamically active and there are a number of very well defined topographic features with associated sediment types (Cameron *et al.* 1992). Under these circumstances it is reasonable to assume that the sediment characteristics (and hence E1 and E2 values) are related to depth at a very local level. If additional depth data are available from charts, then it is possible to interpolate E1 and E2 data for a particular depth range into areas that are bounded by those same depth ranges. In other words, interpolation of E1 and E2 follows the depth contours taken from the charts. The following procedure carried out for the Lincolnshire data shown in Figure 12.4:-

1. Depth contours were digitised (in the form of closed polygons) for the ranges 0-5m, 5-10m, 10-20m, 20-30m, 30-40m, 40-50m, 50m+.
2. The acoustic data were divided into subsets for depth ranges of 0-7m, 3-15m, 7-25m, 15-35m, 25-45m and 35-55m and 45m+.
3. The data in the subsets were then interpolated but only the values inside their respective contour ranges of 0-5m, 5-10m, 10-20m, 20-30m, 30-40m, 40-50m and 50m+ were removed. Note that depth ranges for the acoustic data subsets are wider than their corresponding contour ranges. This overlap is necessary to smooth interpolated values at contour boundaries. A wide grid spacing (250m) was selected so that a relatively small number of values are created. This was done so that subsequent interpolation (see step 5) is weighted towards real values.
4. These interpolated values were then added to the original track data.
5. A second interpolation was carried out on this data set, but using a smaller grid spacing (100m).

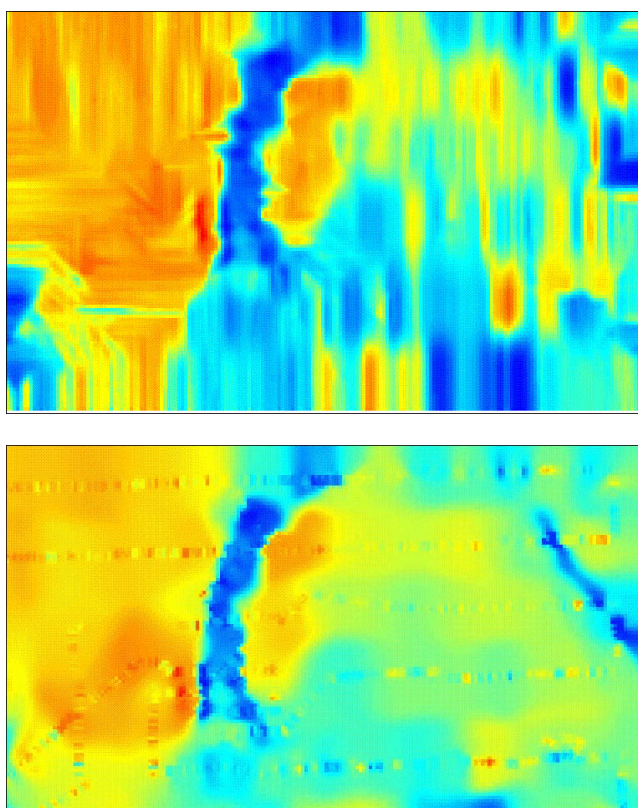


Figure 12.4. E1 values interpolated for an area of approximately 15km by 25km off the north Lincolnshire coast in the locality of Silver Pit. In Figure 12.3a the track data have been interpolated without reference to data from the hydrographic charts whilst in Figure 12.3b the interpolation has been constrained within depth ranges as explained in the text.

13 Classification

Purpose: To interpret the acoustic data according to acoustic ground type, habitat or biotope category.

13.1 Introduction

Raster images are the starting point for analysis using powerful image processing packages. Both continuous variables and categories can be converted into raster coverages. Variables will be converted into a number of classes each with consecutive ranges of values. Raster images can simply be a colour picture displaying the variable or classes. However, the term **smart** raster applies where the centroid ID for each tile is associated with data contained in a spreadsheet or database. Smart raster images are amenable to manipulation and analysis using sophisticated GIS and image processing software.

13.2 Regular arrangement of E1/E2 classes

The simplest form of unsupervised classification (Figure 13.1) of the digital images is an extension of the classification using E1/E2 boxes introduced in Section 11.2 and using nearest neighbour interpolation (Section 12.4).

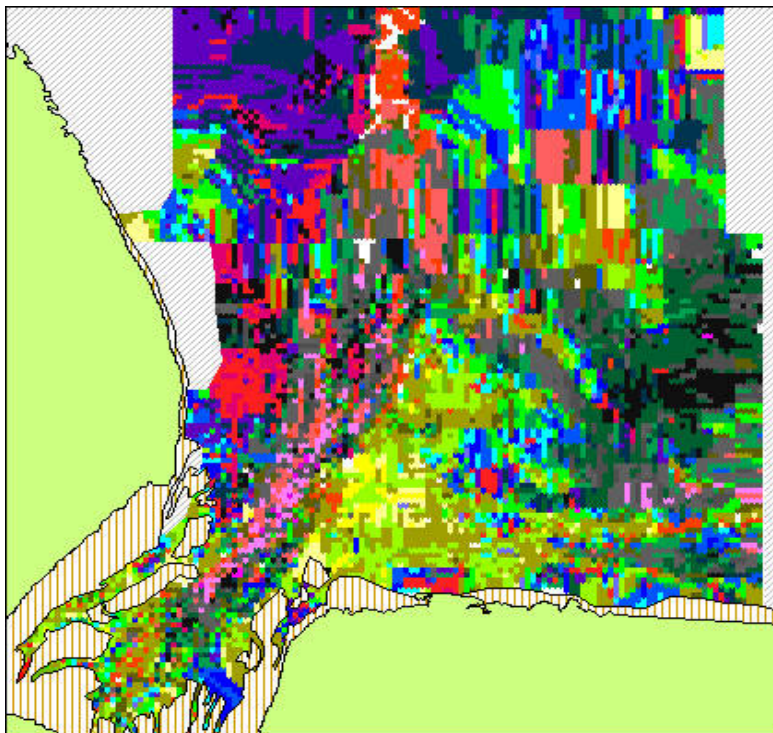


Figure 13.1. The track data have been coded according to the E1/E2 boxes in Figure 11.2 and then interpolated using nearest neighbour. In essence, this tracks the tracks as displayed in Figure 11.2 and extends the colours away from the track using the nearest real colour code record to label the interpolated grid point.

13.3 Classification using calibration

The use of areas of known ground type for designing boxes on the Cartesian plot of E1 and E2 was introduced in Section 4. Ground discrimination by calibration has been used to test the efficacy of *RoxAnn* (Schlagintweit 1993) for discrimination between sediment types and he reports that the system gave a reasonable degree of correlation between ground type predicted by *RoxAnn* and that found by divers. However, the maps presented in that work indicate significant areas of discrepancy.

Collins and Voulgaris (1993) used *RoxAnn* and the standard *Microplot* box display to discriminate between 8 different seabed types off the Kent coast and found (a) that there was overlap between the

boxes used to define each type and (b) that even over small distances (15km) the position of the boxes for the various ground types was not constant. (Magorrian *et al.* 1995) also used the basic editing facilities in *Microplot* to survey the seabed of Strangford Lough, Northern Ireland and distinguished a small number of bottom types (*Modiolus*, *Nephrops*, brittlestar beds and soft ground).

The use of calibration for survey has been used by a number of workers. Rukavina and Caddell (1997) used *RoxAnn* to survey contaminated sediment in nearshore sediments in Lake Ontario, Canada. The arrangement of the boxes was determined through calibration and used, with minor alterations, for many different areas over a number of years. Mapping was done using very closely spaced tracks and no interpolation was needed to obtain a visual representation of a continuous coverage. They found discrimination to range from good to poor largely dependant upon the heterogeneity of the ground.

An alternative way of calibrating *RoxAnn* has been to use a contouring routine to interpret the E1/E2 plot of the AGDS data from calibration sites in terms of sediment characteristics (Hull and Nunny 1999). In this case the contoured E1/E2 plots were used to identify areas within the plot that could be associated with particular sediment qualities (e.g., percent gravel or mud) and the authors used this information to help in the construction of the arrangement of boxes on the E1/E2 plot.

13.4 Unsupervised classification

There are serious disadvantages to using a simple box arrangement drawn around clusters of points on a Cartesian plot of E1 against E2. The boxes are bounded by upper and lower E1 and E2 values forming rectangular boxes. This is termed **paralleliped** classification and has been used to map limited numbers of biological communities (Murphy *et al.* 1995; Magorrian *et al.* 1995). However, the rectangular nature of the boxes may not be the most appropriate shapes for the clusters and this limitation can be overcome by drawing irregular shaped polygons (Greenstreet *et al.* 1997).

However, the definition of these clusters is subjective and dimensions other than E1 and E2, such as depth and variability, cannot be accommodated. *Idrisi* provides a simple clustering technique that can use three variables (it is designed to be used with red, green and blue images from colour aerial photographs). A three dimensional frequency histogram is formed from three variables (e.g., E1, E2 and depth) and clusters are defined from the peaks and troughs in the histogram. These clusters are then associated with particular ground types.

This technique was used by Greenstreet *et al.* (1997) for the discrimination of sediment types in the Moray Firth. The clusters were defined and sediment samples collected on the basis of the spatial distribution of these AGDS data clusters. Multidimensional scaling ordination (MDS) was then used to see if the sediment samples were associated in the same groupings as the AGDS clusters. They used analysis of similarity (ANOSIM) to test the significance of the sediment clusters and found that there was significant overlap between many of the clusters, although a small number were quite distinct. A similar approach was used by Kaiser and Spencer (1996) and Kaiser *et al.* (1998) in the study of the effect of fishing gear on seabed communities. Again, they found considerable overlap in the acoustic characteristics of different bottom types although *RoxAnn* could be used to distinguish between trawled and untrawled areas of the sea floor.

The objective of clustering, however, is to find non-overlapping groups and for this reason may not be most suitable for environmental variables that merge into one another with no clear division between categories.

13.5 Supervised classification

The principles of supervised classification have been introduced in previous Sections. This Section gives more details as to its use for the interpretation of AGDS data. The starting point for successful image classification is to have good digital images of the variables to be used in the classification and a ground truth data set in which the samples have been confidently identified.

13.5.1 Importing digital images

The digital images created in *Surfer* are imported into *Idrisi* as ASCII text files. These images need to be converted into the format for spectral images and the data are rescaled so that values between the minimum and maximum range from 0-255. Interpolation may create values above and below this range and these are ignored in the transformation.

13.5.2 Signature development

An acoustic signature is formed by associating a set of values for the various acoustic parameters (E1, E2, variability and depth) with particular habitat or biotope types. The acoustic signature will be composed of a mean, standard deviation and range for all parameters.

13.5.2.1 Selection of training sites

The approach to signature development most often described is where polygons are digitised directly into the GIS using the digital image or other land class map to identify suitable training sites. These polygons are then used to extract spectral information from the digital images. This is impractical for subtidal acoustic surveys using point ground truth samples since there is no direct observational evidence to support the construction of extensive training polygons. (However, this approach is possible using sidescan sonar images – see Section 19).

Instead the ground truth data are usually in the form of point or line data and it is assumed that a small area can be drawn around the location of these samples to form training sites. This is done by creating a small **buffer** around the ground truth sample position. The buffer is used to extract acoustic data, which is assumed to be associated with the habitat or biotope of the ground truth sample. A small buffer is preferable although the minimum extent of the buffer is determined by the combined positional errors of the AGDS and the remote sampling device. Thus, there will usually be some uncertainty about the association of the AGDS records used for signature development and ground type (unless the ground is very homogeneous). This will inevitably result in some degree of smoothing of the mean of the signature (i.e., the mean values for the signatures will tend towards the mean of the complete data set) whilst the range of values and standard deviations will be large. The data to be included in a signature can be edited to reduce the chances of including records not properly associated with the ground truth sample. At a coarse level a ground truth sample might be excluded if it comes from a very heterogeneous area. Other editing can be performed on the data from the buffer by inspecting the frequency histogram of E1 and E2 for discrete and widely separated peaks in the data and excluding those values that are obviously not associated with the ground type (e.g., low values of E1 and E2 where the ground type was of rock). This has been used cautiously in the BMP since it assumes that there is already some knowledge of what the ‘correct’ values should be. The use of data editing in this way is also implied in the approach adopted by other users of *RoxAnn* (Hull and Nunny 1999).

Signature development requires a minimum of data for each class which can be met by (1) increasing the number of ground truth data for each habitat or biotope type (not an option after the field survey), (2) amalgamating ground truth types into a smaller number of classes (thereby increasing the amount of data for the remaining classes), (3) increasing the buffer size to include more acoustic data (only acceptable in homogeneous ground), and (4) by increasing the number of estimated values (decreasing grid spacing) of the image which will increase the number of estimated values that fall within each buffer. The latter is preferable since small buffers reduce the chance of including acoustic data that are unrelated to the ground type and also because large pixels will reflect the average acoustic values over large areas. However, as has already been stated, images with a large number of grid nodes create processing penalties due to file size.

13.5.2.2 Selection of variables to include in the classification process

The variables that the *RoxAnn* AGDS produce directly are E1, E2 and depth. However, a variability index can be derived for E1 and E2 (see Section 10.5.2). The analyst must decide which variables to include in the analysis.

- A degree of correlation between variables is expected in multispectral data (Wilkie and Finn, 1996). Likewise, E1 and E2 are not independent variables. Although E1 and E2 can be subject to PCA transformation, there is no obvious improvement in discrimination and transformation does not appear to be necessary.
- Depth is very important in structuring biological communities and may be directly related to sediment type through hydrography and seabed. It is doubtful if depth should be regarded as equivalent to E1 and E2 in image processing. It could be argued that E1 and E2 on their own should, in the first instance, be used for classification and that depth should subsequently be overlain to reclassify the habitats and biotopes. However, this would appear to be unnecessarily complicated and depth is routinely included with E1 and E2 in signature development by *SeaMap*.
- Variability is not a point measurement as are E1, E2 and depth but rather a measure of the acoustic properties of more than one acoustic record. This would seem to make variability substantially different in kind from E1, E2 and depth. However, the latter are themselves averaged through the process of interpolation. Indeed, as has been explained in Section 10.5.2, variability is taken to be a measure of acoustic ground characteristic that compensates for the smoothing effect of interpolation.

13.5.3 Maximum likelihood

Although there are a number of classifiers, maximum likelihood is universally acclaimed as the most satisfactory (Bailey and Gatrell 1995; Wilkie and Finn 1996; Eastman 1997). The classification is based on a fundamental theorem of probability put forward by Thomas Bayes in the eighteenth century. The probability is calculated that a pixel with particular acoustic values (expressed as a 'value vector' composed of n elements) is a member of a particular habitat or biotope class. This membership probability is calculated from a variance-covariance matrix and the pixel to class distance.

The estimation of the likelihood of a pixel belonging to any of the habitat or biotope classes for which there are acoustic signatures is a key stage in the classification process. A separate image for each habitat or biotope class is created. Once the probability is calculated for each class, then the maximum likelihood classifier takes each pixel in turn, selects the most likely class from these images for a particular pixel and assigns this class to it.

However, these probability images can be viewed independently of the maximum likelihood classification process and Figure 13.2 is an example of a probability image for one biotope, *Sabellaria* reef. It is inevitable that many pixels will have no class that stands out clearly from the others. In these circumstances the choice of the maximally likely class will be somewhat arbitrary. This situation is most likely to arise when there are a number of habitat or biotope classes being considered that are very similar to each other. It may be preferable to amalgamate similar classes into a new, larger class.

Choice of habitat or biotope classes is, therefore, very important to the classification process. The choice of class can be made before the classification process in anticipation of confusion that is likely to arise between similar classes. However, amalgamations between similar classes can be performed at various stages in the classification process through the use of the individual class probability images. The probabilities for two similar classes can be added and the maximum likelihood classification re-run.

13.5.4 Prior probabilities

Confusion between biotopes with similar acoustic signatures is very common with AGDS. However, there are broad scale trends in distribution of these biotopes which can help to separate biotopes. This is done through the use of prior probabilities.

When ground truth data are plotted on a map of the survey area it becomes clear that there are broad scale trends in the distribution of biotope and habitat type. For example, bedrock was commonly found close in shore on the Pembrokeshire coast, but never found more than a few kilometres from the coastline. Likewise, *Modiolus* communities were found some distance off the Norfolk and Lincolnshire coasts, but never in the Wash. However, it is possible that the process of matching a pixel to acoustic signatures will give a positive probability value for a habitat or biotope class which is believed not to occur in that part of the image.

Maximum likelihood classification can take broad distribution trends into account through the application of prior probabilities although this facility is underused in classification (Eastman 1997). It is often assumed in image classification that each class is equally likely to be represented in an area since there is often little information to know *a priori* what the proportional composition actually is. However, Bayes' theorem makes use of prior knowledge to estimate the probability of an hypothesis being true (the membership of a pixel belonging to a biotope class is an hypothesis in this context). These prior probabilities can be estimated from the distribution of the ground truth data and other data or knowledge and expressed as a **prior probability image**. These images can be incorporated into the maximum likelihood classification and modify the posterior (output) likelihoods.

Creating prior probability images from ground truth and other sample data: A map showing the distribution of the ground truth data are divided into polygons each containing between 5 and 10 samples. These polygons are drawn to coincide with important physiographic boundaries to enclose areas with a fairly similar representation of habitats or biotopes. The ground truth samples are counted and the proportion of each biotope or habitat class calculated for that polygon. This is repeated for all polygons which results in a probability density distribution for each habitat or biotope class separately.

However, prior probability images based on these polygons will have sharp transitions at polygon boundaries. This problem can be overcome by substituting the polygons by points lying well within polygon boundaries and interpolating the probability density values from these points. This results in a prior probability image with more gradual transitions between values. These substitute points can be created in a number of ways: for example, it is convenient to use the positions of the original ground truth locations for the positions of the new points. Alternatively, a regular arrangement of widely separated points could be generated. These points will have the probability density values of the polygon in which they are situated.

The same geographic boundaries and grid spacing are set for interpolation that were used for the creation of the acoustic images so that the prior probability image will overlay the acoustic images exactly. With so few points, it is important not to use an interpolation algorithm that emphasises individual point values (producing what is called a 'bulls-eye effect'). Minimum curvature is used because it gives a very smooth interpolation (Keckler 1994). As with all interpolation, the calculation may result in impossible probability values less than zero and greater than one and for this reason the prior probability images are re-scaled forcing a minimum value of zero and a maximum value of one when they are imported into *Idrisi*. At the conclusion of this process there is a prior probability image for each of the habitat or biotope classes. When maximum likelihood is run, these prior probability images are selected.

The following maps (Figure 13.2) show an example of prior probabilities applied to the distribution of *Sabellaria* reefs. Figure 13.2a is the prior probability distribution of the reefs (suggesting where reefs are most likely to be found on the evidence of ground truth data). Figure 13.2b shows the **posterior probability** after the pixels have been matched to the acoustic signature of the reef biotope. As another example, basic acoustic signatures did not provide discrimination between the various kelp biotope complexes in the Firth of Lorn that differed only in their wave exposure. These classes were amalgamated into a more generic class for the basic life form maps, and the actual stations where samples were recorded were used to generate maps of prior probability to help the classification process.

Prior probabilities reflect what the analyst believes about the likely distribution of a habitat or biotope before looking at the acoustic evidence. This belief should be justifiable and arrived at in a systematic

way. However, the use of ground truth data is only one way to derive a prior probability image. Prior probability images might be based on knowledge of the physical limits to the extent of some biotopes. For example, if kelp forests are known to have a lower depth limit in a survey area of 20m, then an image of the 0-20m depth range can be created and given a high prior probability for kelp forest biotopes. If pixels that match kelp forest biotopes signatures are found below 20m, then the prior probability image would depress the likelihood that, in fact, kelp would be found at that location.

Note: It is important to appreciate that this process will not ‘find’ a particular biotope where the pixels do not match the biotope’s signature. But where a pixel might be classified as either of two biotopes (a) because signatures overlap considerably or (b) because the characteristics of a pixel lie equidistant between two signatures, then the choice is influenced by the prior probability by increasing the posterior probability of one biotope and reducing the probability of the other.

In the case where a biotope is given a probability of zero (i.e., it is considered impossible for a biotope to be in a particular location; for example an exposed kelp biotope in a very sheltered sea loch) then even though the pixel characteristics match the signature much better than any other a second choice will be forced since the posterior probability will also be zero. Extensive use of prior probability images has been made in the analysis of the data from all survey areas.

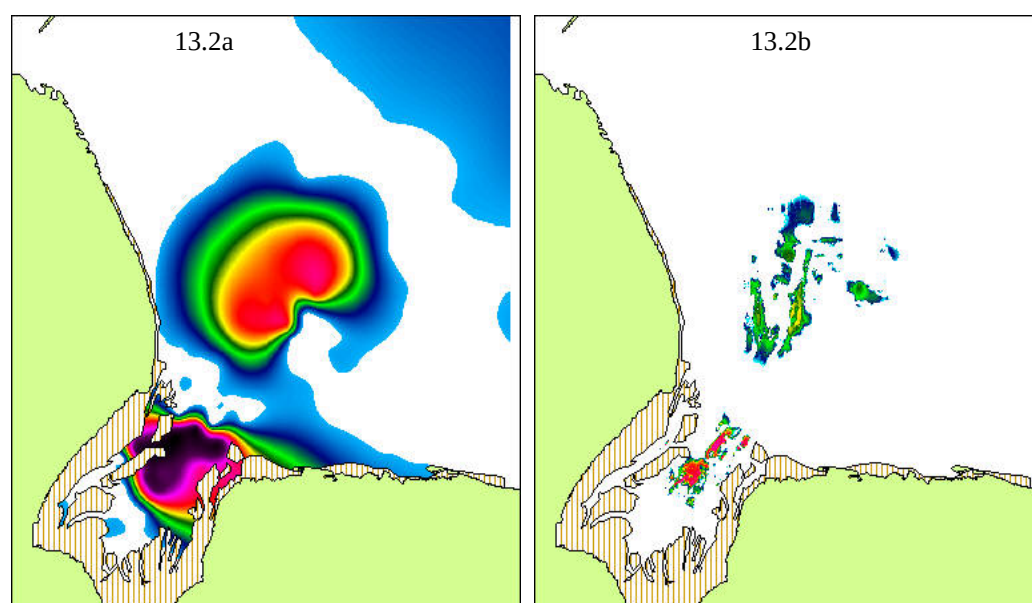


Figure 13.2. A prior probability image (13.2a) has been used to in the calculation of the posterior probability distribution of *Sabellaria* reef (13.2b).

13.6 Probability images and maximum likelihood classification

Probability maps for all biotopes can be displayed and compared with the distribution of the biotopes as shown by maximum likelihood alone. This comparison shows where there is a high degree of certainty in the maximum likelihood map and where the certainty is low (although still higher than any other biotope). Figure 13.3 shows a selection of four biotopes and for each a comparison can be made between the boundary defined by maximum likelihood and the probability values. This example is taken from the Wash, north Norfolk and Lincolnshire coasts. The maps show probabilities greater than 0.1 (colour-graded from dark purple to blue) and the boundaries for maximum likelihood.

The probabilities for super-abundant *Sabellaria* and *Nephtys/Bathyporeia* biotopes agree well with the distribution of these biotopes as determined by maximum likelihood. There is also reasonably good agreement between probabilities >0.1 and the maximum likelihood for *Abra* in the Wash although *Abra* may be more widely distributed than the maximum likelihood classification suggests in the north of the survey area.

Sabellaria/Lanice, on the other hand, has been selected as the most likely biotope for large areas offshore although the likelihood is low (<0.1). This would indicate that there is a high level of uncertainty over much of this area, despite the fact that *Sabellaria/Lanice* was the only biotope found within the appropriate biotope boundaries. Thus, although likelihood is low, *Sabellaria/Lanice* is still the best prediction and this is confirmed by the ground truth data.

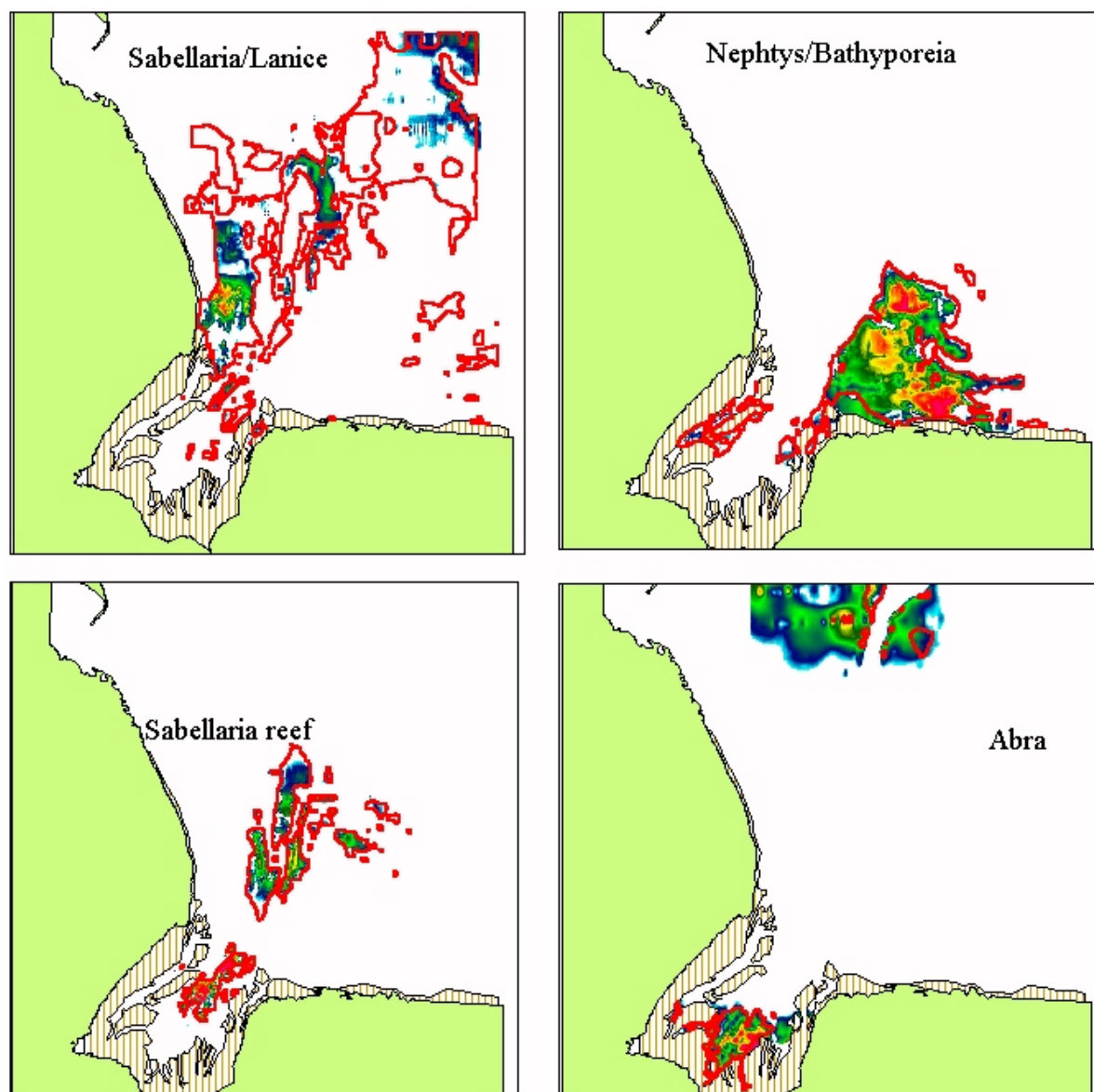


Figure 13.3. Examples of likelihood maps for four different biotopes. Blue = likelihood of 0.1; red = likelihood of 0.9. Likelihoods <0.1 have been excluded to simplify the maps. The red lines enclose areas that have been classified by the maximum likelihood classification procedure (i.e., the areas enclosed by the red lines have been classified as being the corresponding biotope whilst all areas outside of the red line have been classified as some other biotope).

13.7 Displaying levels of probability

The probabilities can be re-classified to give not only the maximum likelihood, but also the second, third and fourth next likely biotope. The first level image indicates the class with the highest probability, the second level image shows the class with the next highest probability, the third level image shows the class with the third highest probability, and so on. These levels can be very useful in determining the character of a pixel since it allows one to break down the dominant and successively subordinate characteristics. The following examples, illustrated in Figure 13.4 are taken from the Firth of Lorn.

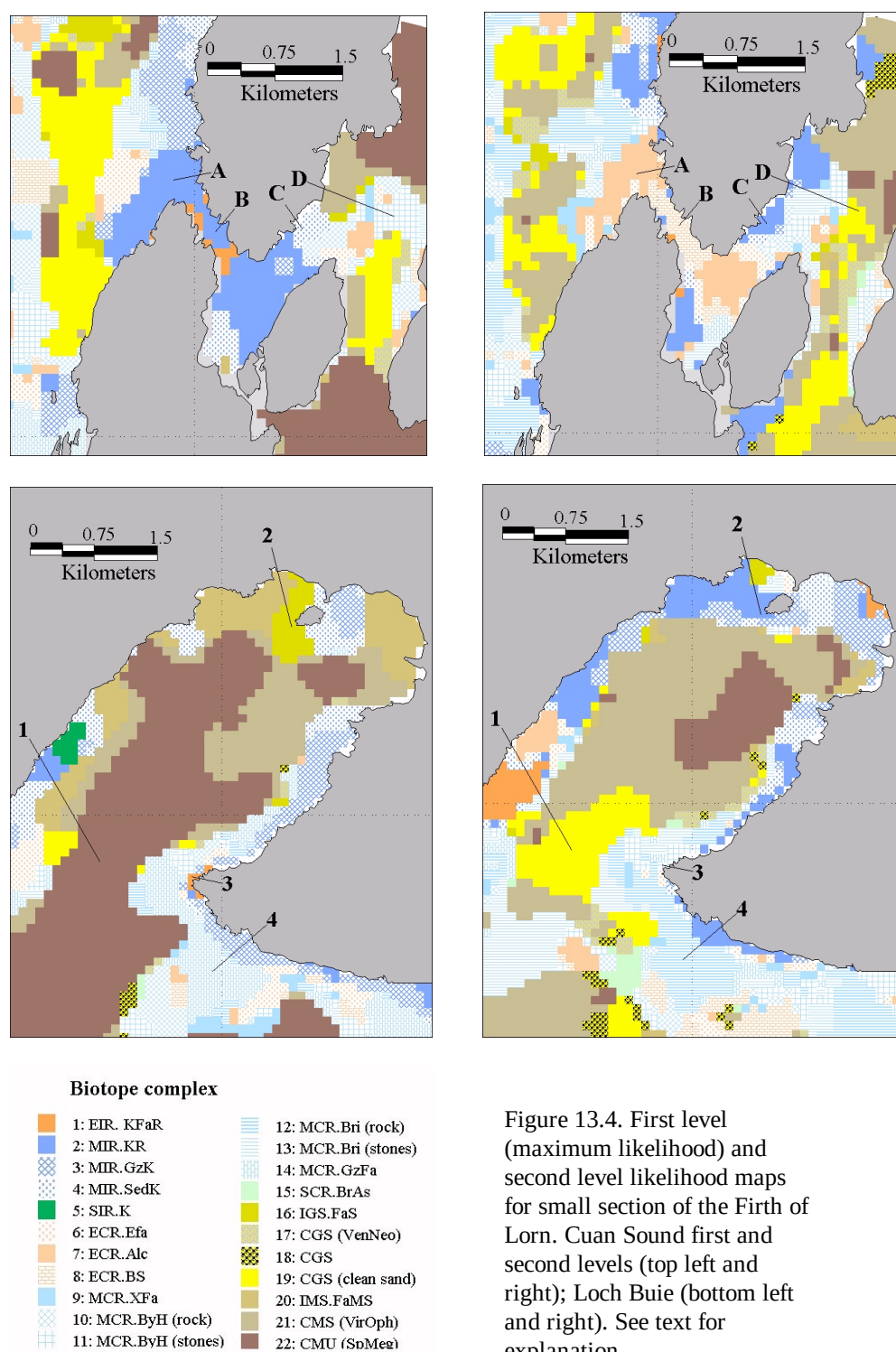


Figure 13.4. First level (maximum likelihood) and second level likelihood maps for small section of the Firth of Lorn. Cuan Sound first and second levels (top left and right); Loch Buie (bottom left and right). See text for explanation.

Cuan Sound: The entrance to Cuan Sound (A) is of rugged bedrock of depths ranging irregularly from infralittoral to circalittoral over short distances. The acoustic characteristics of the pixels in this area match the signature of infralittoral kelp (MIR.KR) more closely than any other biotope signature. However, the next most likely match is for circalittoral *Alcyonium* (soft coral) turf (ECR.Alc) which was, in fact, present at the lower margins of the infralittoral zone in this area. Similarly, Cuan Sound (B) has been classified as being of kelp, but there is also a strong likelihood that circalittoral barnacle/sponge (ECR.BS) biotope complex was also present. This latter biotope was indeed found, especially in the deeper, U-shaped channel in the centre of the Sound.

Sheltered infralittoral habitats were often a mixture of bedrock, boulders and patches of sand with a narrow strip of bedrock and boulders adjacent to the shore. The spatial resolution of the maps cannot represent these transitions. Whilst the overall area at locations such as at 'C' are classified at the first level as kelp on sediment (MIR.SedK) and kelp on bedrock (MIR.KR) at the second level, the reality might be patches or zones of both. Location 'D' illustrates two apparently very contrasting biotopes at the first and second level. Thus, faunal turf on stones (MCR.ByH –stones) is the first choice whilst clean sand (CGS-clean sand) is the second. However, the area was typical of a transition zone between these two different biotopes.

Loch Buie: Sediments sampled at location '1' appeared to reflect a gradual transition from exposure to extreme shelter with increasing distance into the loch. There was, however, considerable spatial heterogeneity at a fine scale. The first level biotope reflects the more sheltered conditions (CMU.SpMeg) whilst the second level raises the possibility that more exposed sediments might also be present towards the outer part of the loch. The contrasting biotopes at '2' (the coarse sand IGS.FaS and the kelp biotopes MIR.KR and MIR.GzK) is a further illustration of patchy biotope distribution.

At small but topographically significant locations, such as the exposed headland at point '3' where the rock falls away steeply into deep water, the mixture between infralittoral biotopes and circalittoral biotopes as shown by the first and second level biotopes reflects the range of conditions likely to be found from the extremely high wave energy shallow zone to the more benign deeper water which will have biological significance. In the last example (point '4') the biotope present was transitional between at least two biotope classes and the identification was difficult to make. These two biotopes are the first and second level choices (MCR.Bri(rock) and MCR.GzFa) and either might be considered to be present.

13.8 Uncertainty measurement

Probability distribution maps can also be used to determine uncertainty of image classification. *Idrisi* allows the calculation of the extent to which one class is more likely to be found than any other class on a pixel-by-pixel basis. If the margin is great, then the classification of that pixel has a high level of certainty. If, on the other hand, the margin is not great, then the classification is uncertain. The resulting value for this margin can be plotted for each pixel giving an uncertainty map for the survey area.

The following example (Figure 13.5) is of a section of the Pembrokeshire coast in which the pixels have been given an uncertainty measure. Uncertainty can result from poor signature development or from heterogeneity of the ground (to be discussed further in Section 15.5.3). In the case illustrated in Figure 13.5 the highest levels of uncertainty coincide with the rocky habitats and these are quite variable, with bedrock, boulder and patches of sand and gravel. This might indicate that heterogeneity has caused high uncertainty.

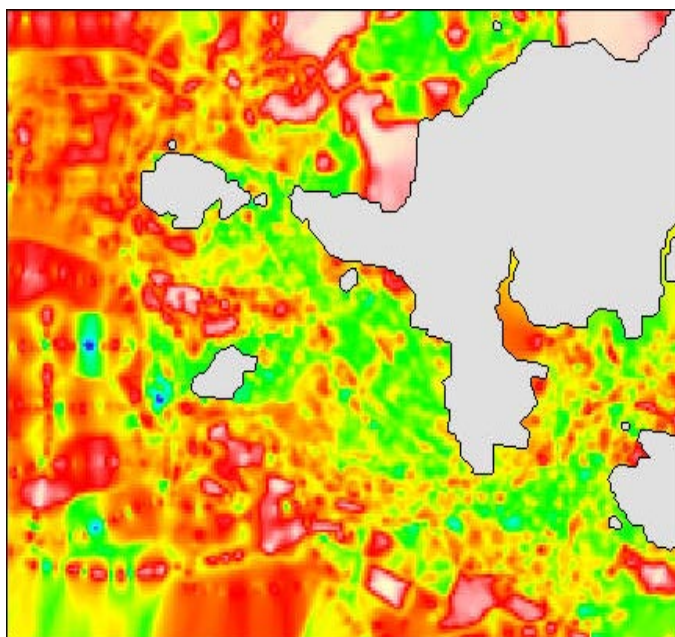


Figure 13.5. The pixels in this image of Skomer and the entrance to Milford Haven have been coloured according to the level of uncertainty of the maximum likelihood classification. Blue and green indicate high levels of uncertainty whilst red and white indicate low levels of uncertainty.

13.9 Using probabilities to classify track data and amalgamate the data from different surveys

The probability images prove to be a very versatile way of handling data. Probability images can be used to join data from many different and overlapping surveys that have used different AGDS systems. This is a specialised use of probability images that has the potential for opening the way for bringing together many different surveys into one single interpretation.

It cannot be assumed that different *RoxAnn* AGDS surveys will give compatible data. If the same system is used (but on different vessels and at different times), standardisation can render the different data sets compatible. This has been done to amalgamate data derived from separate surveys using the *SeaMap* single frequency *RoxAnn* system. It was also found to be satisfactory for combining the data from this system and the *RoxAnn* unit on the ESFJC vessel *Surveyor*. However, other systems, even if operating at the same frequency, have very different characteristics and transformation and standardisation will not render them compatible. This was the case with data sets from the single and the dual 200kHz AGDS operated by *SeaMap* and the data obtained from the *Prince Madog* and the NW & NWSFC vessel *Aegis*. Data from the dual frequency *RoxAnn* system was also found to respond differently in the deep and shallow water of the Firth of Lorn.

One solution is to classify the data separately and then amalgamate the interpretations. However, one must bear in mind the sensitivity of the maximum likelihood classification process to small differences in probability values when selecting the maximally likely class. The results from different interpretations will often show a high level of unconformity where they overlap. Since the classified data are categorical, the classes cannot be merged through interpolation as could a continuously variable attribute, such as depth.

Probabilities, however, can be treated as continuous variables. This offers an approach for joining data sets together. The procedure is as follows:-

1. Digital images of the track data are created using a small grid spacing and small search radius so that estimated values are only calculated for points that lie very close to tracks. The rest of the image is left blank. In other words, the images for E1, E2 and depth are a very close reflection of the track values.

2. Signatures for habitats or biotopes are developed in the usual way. Probability images are then calculated for each of the habitat or biotope classes, again as previously described. The results, however, are images showing track probabilities.
3. The positive values for each class are then exported from *Idrisi* as point values as an ASCII XYZ text file (the background zero values are not exported).
4. The procedure is repeated for other data sets.
5. The values for each habitat or biotope class from all data sets are then combined and imported into *Surfer* for interpolation.
6. The interpolated images for each habitat or biotope class can now be re-imported into *Idrisi* and these images classified using the maximum likelihood classifier.

The procedure is illustrated using a section of the Pembrokeshire coast that includes Skomer and the entrance to Milford Haven. Figure 13.6a shows the amalgamated tracks coloured according to the probability of finding faunal turf. These probabilities have been interpolated and the resulting continuous surface is shown in Figure 13.6b. In both illustrations blue indicates low probabilities and red indicates high probabilities.

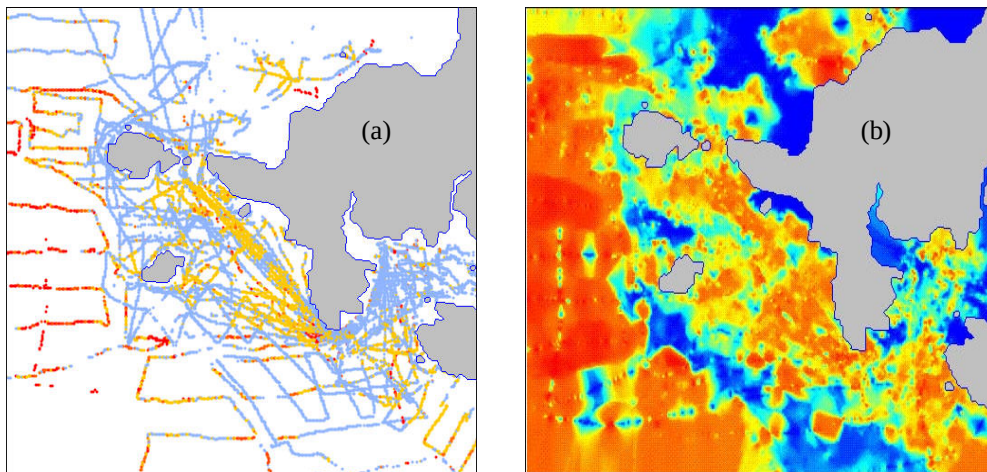


Figure 13.6. The probability of faunal turf has been calculated for the track data (13.6a) for a number of surveys and are shown superimposed. The probability values have been amalgamated and interpolated. The resulting continuous surface is shown in 13.6b.

14 Error matrices and crosstabulation

Purpose: To compare two images for accuracy assessment and determination of change between two images.

14.1 Introduction to matrices

Two digital images that are co-registered and have the same pixel size can be overlain and the pixels from one image compared with those of the other. This is an extremely useful facility for the comparison of two images. For example, accuracy assessment requires a comparison between the ground truth data and the interpreted map. One image shows ground truth information: Pixels in the buffer zone around ground truth samples are given the habitat or biotope code of the sample. The second image is the interpreted acoustic map classified by the same codes. When the two images are overlain, each pixel with a code in the ground truth image can be compared with the code of the corresponding pixel of the map. Each pixel-to-pixel comparison has two possible interpretations. Firstly, if the map incorrectly predicts the class, then the error is termed one of **commission** in that the map has included classes that were not, in fact, present. These classes have been over-represented on the map. The second interpretation is that the correct class has been omitted, so the comparisons can be framed as errors of **omission**. These classes have been under-represented on the map.

These comparisons are made through the construction of a matrix (see Table 14.1) of ground truth data (columns) and mapped data (rows). The row totals gives the number of pixels in the map image that have been classified as particular habitat or biotope classes (but only counting those pixels for which there are corresponding ground truth pixels). The figures in the various columns of that row give the break down of this total in terms of the classes as indicated by the ground truth data.

Table 14.1. An example of an error matrix of the infaunal biotopes of the Wash. The figures are pixel numbers. (the biotope references are, where not obvious:- PolyNem = polychaetes/nemertean; NeoBat = *Nephtys/Bathyporeia*; ScoSpi = *Scoloplos/Spiophanes*; Sabdisc = *Sabella discifera*; SabLan = *Sabellaria/Lanice*; reef = super-abundant *Sabellaria spinulosa*; Sabpav = *Sabella pavonina*).

		Columns: As classified in ground truth image													Total pixels classified
		Lanice	PolyNem	NepBat	NepSco	ScoSpi	Abra	Sabdisc	SabLan	Reef	Sabpav	Ensis	Ophiura	Modiolus	
Rows: As classified on map	Lanice	20	4	7	0	0	0	1	0	0	0	0	0	0	32
	PolyNem	0	28	0	0	14	11	0	11	0	0	0	0	0	64
	NepBat	7	0	92	16	23	1	0	12	0	7	0	0	0	158
	NepSco	0	11	2	41	0	5	4	14	0	0	0	0	0	77
	ScoSpi	0	7	0	0	0	0	0	0	0	0	0	0	0	7
	Abra	0	0	0	0	7	25	0	7	0	0	0	10	0	49
	Sabdisc	0	0	12	0	0	0	17	7	0	0	0	0	0	36
	SabLan	0	14	11	7	18	0	0	125	16	0	8	7	0	206
	Reef	7	4	5	1	1	0	0	38	53	0	0	6	0	115
	Sabpav	0	0	1	0	0	0	0	0	0	8	0	0	0	9
	Ensis	0	0	0	0	0	0	0	0	0	0	8	0	0	8
	Ophiura	0	6	0	0	12	11	0	13	0	0	0	22	0	64
	Modiolus	0	0	0	0	0	0	0	0	0	0	0	0	21	21
Total ground truth pixels		34	74	130	65	75	53	22	227	69	15	16	45	21	846

The column totals give the number of pixels that have been assigned to a particular class from the ground truth data whilst the figures in the rows give the break down of this total in terms of how these pixels were interpreted by image processing. Note that the biotope *Scoloplos/Spiophanes* was rarely predicted. This point is discussed in more detail in Section 14.2.2.

The diagonal gives the tally of correct predictions and the percent correct calculated as an overall figure for the accuracy measure (total correct matches/total pixels used in the comparison). In the above example the overall figure is 55%. Percentage correct can also be calculated for each individual biotope and these are shown in Table 14.2 below.

Table 14.2. Percentage correct matches for individual infaunal biotopes.

Lanice	PolyNem	NepBat	NepSco	ScoSpi	Abra	Sabdisc	SabLan	Reef	Sabpav	Ensis	Ophiura	Modiolus
61	41	64	58	0	49	58	59	58	66	66	40	100

The percent correct does not distinguish between errors of commission and omission and an alternative measure, the Kappa coefficient, gives the measure of agreement between two images beyond chance (see Congleton (1991) and Wilkie and Finn (1996) for a full explanation). The measure for the Wash infauna is 30%, which is quite low. However, a wide range of biotopes are involved in the classification process and much of the confusion is between similar biotopes. This will be discussed in more detail in Section 14.2.2 below.

14.2 Uses of matrices

14.2.1 Assessing internal consistency of the classification and predictive power

The use of matrices as described above indicates the level of agreement between the ground truth data used for classification and the map. It is an internal measure of agreement (**internal accuracy**) and does not test the predictive power of the map. An external data source is required for the latter (**external accuracy**), although positional errors, possible changes over time and poor correspondence between techniques used to identify samples can make external accuracy measures particularly prone to uncertainty.

An illustration of the difference between internal and external accuracy is provided by a study on redundancy of ground truth sampling in Loch Maddy, North Uist (Foster-Smith *et al.* 1999). A small area was sampled intensively and this provided sufficient ground truth data to be able to remove a proportion of the data from the classification process and reserve it for external accuracy assessment. The ground truth data used in the classification process were used for internal accuracy measurement. The ground truth data were divided into subsets as follows: (a) one quarter of the data were set aside for external accuracy assessment, three quarters used for classification; (b) the data were split half and half; (c) three quarters of the data were set aside and one quarter used for classification. These ground truth samples were collected by drop down video deployed at intervals during the collection of the acoustic data. Thus, the positional error was kept to the minimum. An independent set of samples were collected by divers and ROV on the same survey, but at a different time and the data were also used for external accuracy against a classification using all the drop down video. Table 14.3 summarises the internal and external accuracy measures for the four treatments. The figures are overall percentage accuracy.

Table 14.3. A comparison between internal and external accuracy measures taken from a study in Loch Maddy. See text for explanation.

Treatment (data used for classification)	Internal accuracy	External accuracy
Three quarters	52%	51%
One half	56%	38%
One quarter	86%	31%
Diver/ROV	N/A	21%

Internal accuracy was variable but, although there are insufficient data for confirmation, it would seem that smaller data sets gave a higher internal accuracy. This is perhaps not surprising since the signatures generated by a small data set may not have the spread of values of a larger data set. Thus, whilst the resulting interpretation may be internally accurate, the signatures do not fully represent the biotopes.

The accuracy measured against an independent data set gives a very different view of the relative performance of the different data sub sets used for classification. They show that there is a steady decrease in accuracy as the sample numbers used for classification are reduced. The independent data set provided by the diver/ROV samples gives the poorest accuracy measure possibly because of increased positional error.

Internal accuracy assessments are only useful as a check on the performance of the analyses and not the relative accuracy of different data sets. If it is important to assess how accurate the interpretation is, then additional ground truth samples should be collected in order to calculate an external accuracy error matrix. In addition, the extra samples could allow for re-interpretation using various combinations of the total ground truth sample, a process analogous to statistical ‘bootstrapping’ (Manly 1997). However, external accuracy again must be used with caution, especially in areas where the seabed is heterogeneous (as is the case in the Loch Maddy trial area). A direct pixel-on-pixel match ignores near misses that would be expected with the positional errors involved with survey.

14.2.2 Matrices as indicators of overlap between habitat or biotope classes

Internal accuracy matrices serve another very useful function: They give a good indication as to where confusion between habitat or biotope classes and their acoustic signatures is most likely to arise. An inspection of the error matrix for the Wash infauna (Table 14.1) shows that *Sabellaria* reef was infrequently omitted from the map (there was a 16/69 or 23% chance of reef being omitted) and where it was omitted, it was always found in the closely related *Sabellaria/Lanice* biotope. However, the extent of reef was more likely to be over-estimated (a 53/115 or 46% chance of other biotopes being found in areas labelled as reefs). Of these other biotopes, the related *Sabellaria/Lanice* biotope was the most likely to be found (a 33% chance), with a lesser chance of others.

Other biotopes are also confused, often quite markedly. For example, the biotope *Scoloplos/Spiophanes* was never correctly predicted and it is likely that this biotope is ‘shared’ between many other biotopes in the area. The data from the error matrix suggests that, because of poor acoustic discrimination, this biotope could be usefully discarded and the records amalgamated into other biotope categories. In fact, its component fauna and sediment characteristics could almost be regarded as a common denominator between a wide range of biotopes.

Ophiura biotopes are also poorly mapped and, again, these brittlestars are quite catholic in their habitat requirements. It might be better to regard these species as being present in a wide variety of biotopes rather than as key species for the identification of biotopes in their own right.

14.2.3 Matrices to compare change

Matrices can also be used to assess change between classified two images, such as the results of surveys conducted over the same area at different times. This process is termed **crosstabulation** but the principles are the same. This is used in the following section on repeatability (Section 15).

15 Resolution of maps

Purpose: To summarise the technical limitations to survey and discuss cartographic detail.

15.1 Introduction

What can broad scale mapping techniques show? Or, put another way, what can they resolve? Resolution of maps is a complex issue which relates to the technical limitations of the survey system on one hand and cost and survey effort on the other hand. The term ‘resolution’ can refer to the level of detail to which habitats or biotopes are discriminated or refer to the spatial detail that has been mapped. The following questions about resolution and accuracy are pertinent to mapping surveys:-

- What is the maximum level of habitat or biotope detail that it is possible to discriminate? How good is the AGDS in discriminating between different ground types? What is the accuracy of discrimination?
- What is the maximum spatial resolution that is technically possible? How accurately have features been mapped?
- What is the actual spatial resolution of a map bearing in mind survey intensity? Does a map show the level of habitat or biotope detail and spatial detail required? Are there cartographic ways in which maps can maximise information content?
- Can all the detail it is possible to resolve be shown on a map drawn to a particular scale? If not, what cartographic procedures have been used to generalise the detail?

The answers to these questions are necessarily hedged about by conditions and they often appear unsatisfactory to the enquirer. The answers are also very dependent upon costs: Resolution, confidence and accuracy can be increased, but at a cost. If very large areas are to be surveyed (which was the case in most of the BMP surveys) then the costs may be too high and there may not be the requirement for this quality of output. However, for other purposes, such as monitoring of small areas, then the costs might be justified.

This Section attempts to provide some clarification of the above questions, if not definitive answers, by summarising BMP experience of a number of issues that are involved with resolution. The corner stone of interpreting acoustic data is linking the data to ground truth samples. Good discrimination and mapping of habitats and biotopes in broad scale survey depends upon:-

- (a) consistent performance of the AGDS system;
- (b) good position fixing of both AGDS and the ground truth samples;
- (c) intensity of AGDS survey;
- (d) categorisation of the ground truth data;
- (e) the ability of the AGDS to give sufficient acoustic information for discrimination;
- (f) intensity of ground truth survey;
- (g) variability in interpretation.

Because fine scale monitoring did not form part of the major remit of the BMP program, a monitoring study conducted by *SeaMap* for Scottish Natural Heritage in Loch Maddy, North Uist, has been used to illustrate many of the points in this Section (Foster-Smith *et al.* 1999).

15.2 Consistency in the performance of the AGDS

15.2.1 Consistency in values for E1 and E2 on a day-to-day basis

All remote sensing techniques have inherent variability that degrades their ability to discriminate features on the ground. In sonar systems this may be due to instability of the electronics and, in particular, the transponder. Two systems with the same nominal performance will nevertheless have slightly different characteristics. Variability may, however, be due to external sources that interfere with the quality of the image of the seabed. The sources are numerous and include sea water temperature and thermoclines, salinity, turbidity and sea conditions. Calibration to achieve a high level of consistency is notoriously difficult and there will always be a residual variability over different spatio-temporal scales which may have an unpredictable effect on the interpretation of the data.

However, if data are to be combined into a single data set for interpretation, then there must be some assurance that the AGDS has been functioning reasonably consistently throughout the survey. Clearly, if this were not the case, then any analysis of the data would be meaningless. Collins and Voulgaris (1993) trialed *RoxAnn* using artificial ponds with controlled sediments on the pond floor and other users of *RoxAnn* have operated the system under very controlled conditions using electronic test equipment to monitor the response to a known calibration signal before, during and at the end of a field survey (Rukavina and Caddell 1997). Both studies have shown that the system is capable of acceptable levels of repeatability although there is always ‘noise’ (inherent variability) in the data.

The use of calibration sites during a survey gives some assurance about the consistency of the performance of the AGDS, although they can never be strictly controlled (see Section 8.4.1).

An example of a calibration exercise is given in Table 15.1 where two sites showing a hard and a soft bottom type respectively were tracked on three consecutive days (taken from the Loch Maddy survey).

Table 15.1. Calibration of a 200kHz *RoxAnn* AGDS on three consecutive days.

Hard ground	Average E1	Stdev	Average E2	Stdev
Day 1	0.439	0.05	0.257	0.07
Day 2	0.438	0.07	0.25	0.09
Day 3	0.445	0.12	0.36	0.13
Soft ground	Average E1	Stdev	Average E2	Stdev
Day 1	0.212	0.05	0.043	0.02
Day2	0.221	0.06	0.047	0.03
Day3	0.23	0.06	0.077	0.04

There was no significant difference between the mean values for each of the ground types over the three days. This gave reassurance that the data from the complete data sets for each of the three days could be combined into a single data set for the three-day survey.

15.2.2 Consistency in values between surveys

If the same system is used on different surveys carried out in the same area, then it is useful to be able to directly compare and possibly combine the different data sets (see Section 10.4). However, it is not strictly necessary that the AGDS give directly comparable values between surveys since it is unlikely (and certainly not recommended) that a survey would be conducted without an adequate concurrent ground truth sampling program. The acoustic signatures developed in one area should not be used in

another, for example. If two surveys are to be compared, it is safer to do an independent interpretation of each and compare the results.

15.2.3 Consistency of interpretation

Repeat surveys may have been conducted using different AGDS. It is unlikely that different systems will have a directly similar response to the seabed. If the different data sets are classified separately, then the lesser assumption that the systems are responding to the same seabed features would be enough to ensure that the surveys should be repeatable.

The more pertinent question for repeat surveys is “can AGDS data be consistently interpreted from one survey to another (despite the possibility that the E1 and E2 values are not directly comparable)?” This is a difficult question to answer since positional errors (however seemingly small) and heterogeneous ground are beyond the control of the surveyor and introduce a level of uncertainty that is hard to quantify. Even with the greatest care surveys carried out at two different times will never be strictly comparable since real changes, however apparently trivial, may have taken place in the intervening time between two calibrations that might significantly affect the reflectance properties of the sea floor (e.g., siltation).

Bearing this in mind, how comparable are two surveys conducted over the same area? Figure 15.1 shows the data for the Wash from 1996, 1997 and 1996 & 7 classified (maximum likelihood) using the same ground truth sample data in each case.

There is a similarity between the maps in that the broad distribution patterns of the biotopes are comparable. However, there are quite substantial differences on a pixel-by-pixel comparison. Table 15.2 gives the overall similarities between the images and it can be seen that the similarity is at about the 37-57% level, although this varies between biotopes.

Table 15.2. Overall percentage similarity between the images shown in Figure 15.1.

Data used in comparison	Overall percent similarity
1996 compared with 1997	37.5%
1996 compared with combined 1996/7	56.5%
1997 compared with combined 1996/7	47.5%

There are many reasons why the pixel-to-pixel comparison returns low percent similarities. The 1997 survey was not designed to repeat the 1996 tracks so that there was a disparity in track position and survey intensity. Even small differences in track position can lead to very different acoustic values, especially if the ground is heterogeneous, and this has the potential for influencing biotope signatures. Uncertainty of boat position is inevitable, even with DGPS (see Section 4.3.1.1) and this may be sufficient for different values to be recorded from the same recorded location. Thus, whilst re-running tracks may reduce variability between surveys, there is no guarantee that differences may not be due to positional uncertainty between data sets.

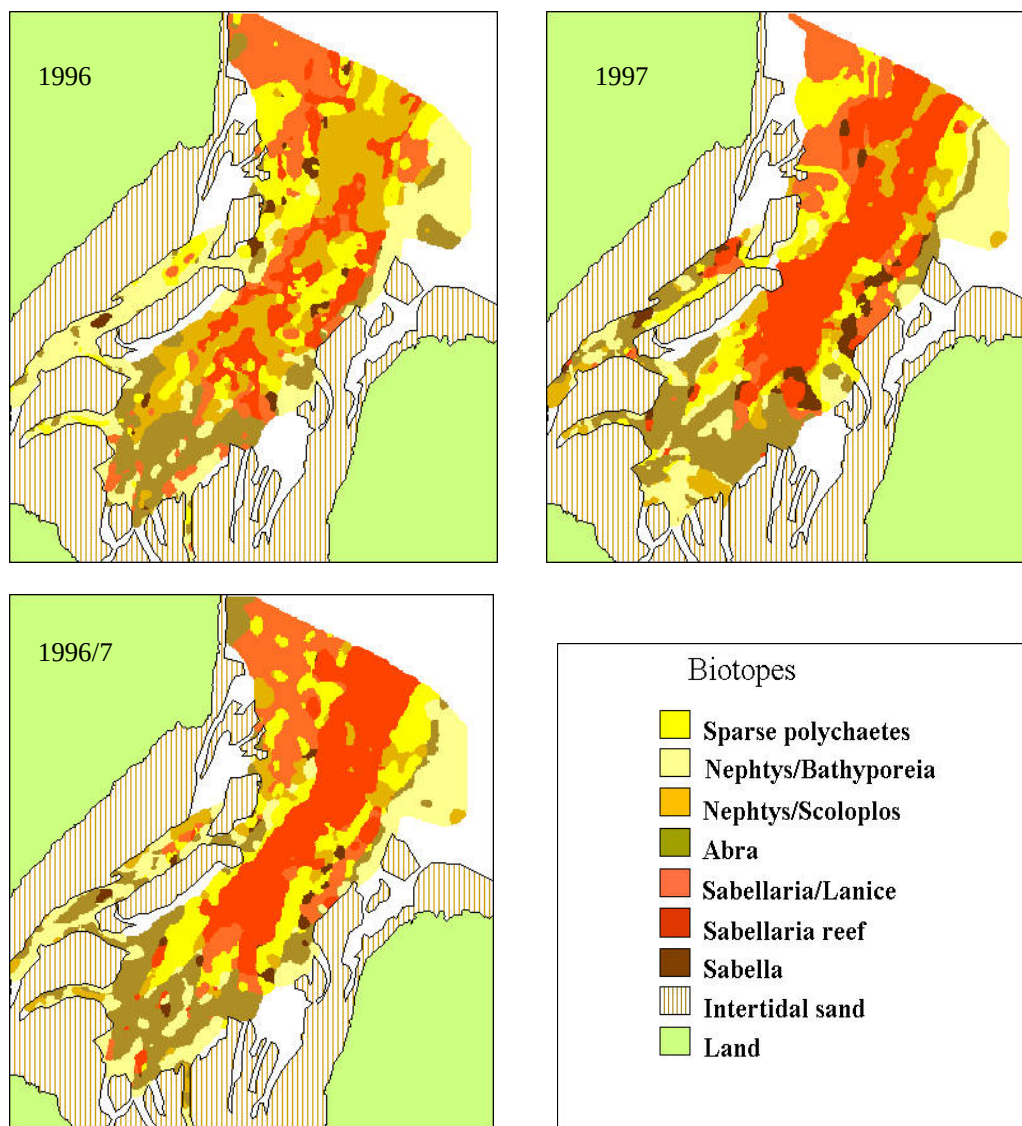


Figure 15.1. Classification of two acoustic data sets (1996 & 1997) using the same ground truth data. The combined 1996/1997 has also been classified for comparison.

15.3 Limits to spatial resolution

Resolution of a remote sensing technology is the minimum sized feature that can be detected and mapped (given a unique spatial position) by the acoustic system. This is not a straightforward issue when applied to AGDS data since the resolution along a track consisting of numerous individual records may be quite high whilst inter-track spacing may be such that resolution after interpolation will be lower.

15.3.1 The acoustic footprint

A single beam echo sounder produces an approximate cone of sound with a specified beam angle that ensonifies an area of the sea floor. This area is approximately circular, although sounders produce many side-lobes that make the footprint a more complex shape in practice. This area ultimately is the limit to the resolution of the AGDS system. However, the area ensonified increases with depth so that a single value for the maximum resolution cannot be given. Thus an echo sounder with a beam angle

of 15° with a depth under the boat of 5m and 60m would encompass an area with diameter of about 1.25m and 15m respectively (approximately 0.25 x water depth; see Section 4.3.1.1). However, the AGDS logging system may average a number of signals (see Section 4.3.3) so that the effective acoustic footprint may be much larger. AGDS cannot detect features which are smaller than this footprint size. Thus, the way in which the position of a record relates to the ground covered also varies with the data logging system used (Section 4.3.3).

15.3.2 AGDS positional error

Detecting a feature is not the same as giving it an accurate position. Consecutive records along a track may indicate changes in ground type over small distances (typically 15m between points), but the position of the whole track may have an error associated with it that depends on the GPS/DGPS system. These errors may not be appreciated by looking at a track in isolation. However, discrepancies might become apparent when positions can be compared with a feature, such as a pier, which gives an independent reference position or when different tracks, perhaps from two consecutive days, overlap and the position of a feature from the two tracks do not coincide.

Positional error for the track will range from as much as 50m with GPS to 5m with public broadcast DGPS (Gilbert 1999), although sub-metre accuracy can be achieved with specialist systems. This means that a position of a feature (such as a change in ground type) will, if recorded often enough over a period of time, range between 5-50m (depending on the system) in any direction from the true position, but with the mean position lying close to the true position. It must be stressed that a small feature will probably be detected on all tracks but its recorded position will vary. In other words, the feature can be detected, but cannot be mapped accurately. However, if the feature (a boundary, for example) is detected on numerous occasions, then the average position will be a good estimate of the true position.

The effect of this positional uncertainty will make itself felt in two ways: Firstly, if a feature is identified from the track records there is a likelihood that it will not be found in that exact position on subsequent re-surveys. Secondly, if data from numerous tracks are interpolated, then the averaging involved in the calculation may result in small features being 'lost' and the sharpness of boundary conditions being blurred.

In conclusion, it is unlikely that features (including intricacies of boundaries between two habitats or biotopes) can be mapped if the detail falls below the positional error of the GPS/DGPS.

15.3.3 Positional error of ground truth samples

Acoustic ground types, which are determined directly from AGDS data, must not be confused with mapping habitats or biotopes since these are interpreted from the acoustic ground types and not measured directly. Thus, the detection and mapping of habitats and biotopes depends not only on AGDS data, but also on the ground truth samples. There is an issue both with positional error of the samples and misclassification. This Section is concerned with the former issue.

The position of a sampling device deployed from a boat is very prone to uncertainty partly due to positional error of the boat (as above) and partly due to the uncertainty of the position of the sampling device relative to the boat. Unless sonar positioning devices are attached to the sampling device the position can only be estimated from the amount of umbilical or wire deployed relative to water depth and the direction of drift.

The uncertainty is much greater if the device is towed rather than simply 'dropped down' since it is usual to have a much greater length of umbilical paid out than the water depth. This lay-back depends upon the speed of the tow (whether active tow or passive drift) but may be as much as two to three times the water depth. Estimation of the position of the video is confounded by tidal currents on the umbilical. In favourable conditions, however, a video will sink directly beneath the boat with minimal drift. However, even if a video is dropped with a length of umbilical 1.25 times water depth, the video might have landed up to a distance equivalent to half the water depth away from the boat (although

the area might be reduced by observing the direction in which the umbilical has set). Thus, very approximately, the true position of the video might lie anywhere within a circle with a radius equivalent to DGPS error + $\frac{1}{2}$ water depth. For example, if a video is dropped in 20m of water, and position error is estimated to be 5m, then the true position of the video might be anywhere within a 20m locus from the recorded position of the boat (assuming that the video could drift in any direction). Clearly, in conditions where the boat is drifting in strong winds or tides the positional error may greatly exceed this. Thus, any data within an extended locus of twice DGPS error + $\frac{1}{2}$ water depth might be considered to be associated with the sample. To this must be added the footprint of the echo sounder which, as has been discussed, is also dependant upon water depth. Thus, using the 20m example above, the ground truth sample might be associated with the seabed up 20m from the recorded position. To this must be added the radius of the acoustic footprint to assess what parts of the seabed could possibly have an influence on the acoustic records associated with the ground truth sample.

These calculations are very approximate and resolution may be bettered in practice and positional error can be reduced by better quality DGPS (Gilbert 1999). However, the above example shows how positional errors grow through addition and serves as a caution against assuming that broad scale mapping through interpretation of acoustic data can achieve a fine resolution with high confidence. It should be stressed that overlaying ground truth records on a map of acoustic data can give an impression of positional accuracy and resolution which is misleading.

15.4 Survey intensity

15.4.1 Track spacing and resolution

The previous Section discusses the maximum spatial resolution that is technically feasible. However, it is unlikely that a survey will be carried out so intensively that the maximum technical resolution is achieved evenly throughout the whole of the survey area. What determines the actual resolution of a map based on broad scale AGDS survey?

Adjacent tracks will inevitably be further apart than the spacing between consecutive records along a track. Clearly, it is possible that features smaller than the intertrack spacing may go undetected. This is a coarse indication of resolution. But how likely is this given the pattern of spatial variability in the acoustic data?

Variability along tracks is a good indication as to the scale of features that might be detected by adjacent tracks and this along-track variability can be conveniently measured for variables such as E1, E2 and depth by the variogram (Section 11.3). For an example where (a) point-to-point variability along a track is low, (b) adjacent tracks are close (in relation to along-track variability) and (c) adjacent tracks have similar acoustic values it is safe to assume that these tracks will have detected most of the features present on the seabed.

On the other hand, if (a) point-to-point variability is high along a track, (b) adjacent tracks are wide (in relation to along-track variability) and (c) adjacent tracks have a similar variability it is very likely that features will not be large enough to straddle tracks and many small features lying between tracks will go undetected.

From these two examples it is clear that there is a relationship between along-track variability, track spacing and resolution. If along-track variability is low and track spacing is narrow, then resolution will be good. The variogram gives a good indication of the distance over which points are spatially correlated. However, what is the threshold for the difference between AGDS values that needs to be exceeded before a change in habitat or biotope is indicated? There is no easy answer to this that could be applied to all habitats or biotopes since acoustic signatures are based on multivariate statistics and the separation between signatures will vary between different habitat or biotope classes.

However, suppose that variance and resolution are directly related, then distance away from tracks corresponding to increase in variance (read from the variogram) would indicate decreasing resolution. This is the basis of Figure 11.5 which displays buffer zones either side of tracks corresponding to increasing variance. A variance of 0.05 was taken as the threshold for substantial change and this corresponded to a distance of about 500m. Thus, within the areas 500m either side of the track (giving a track spacing of 1 km) it is likely that most features distinguished by a change of E1 or E2 values greater than the threshold will have been detected.

A variance of 0.05 equates to a standard deviation of 0.22 and this might seem rather a large change in E1 or E2 values. A lesser variance of 0.02 (standard deviation 0.14) for the threshold would mean that the track spacing would need to be reduced to about 200m to make it likely that most features have been detected.

However, the assumption with these examples is that the pattern of spatial variability is the same over the whole data set. This is unlikely to be the case since spatial variability will change with patterns of homogeneity/ heterogeneity.

15.4.2 Track spacing and consistency of interpretation

Studies designed to test the effect of track spacing on interpretation were carried out by *SeaMap* (Foster-Smith *et al.* 1999) in Loch Maddy, North Uist. A trial area was intensively surveyed using a single frequency *RoxAnn* AGDS operating at 200kHz. In addition to the full data set, other subsets of the data were formed by removing either one half or two thirds of the track data. The data sets were subjected to the standard image processing procedure (i.e., the track data were interpolated and classified using ground truth samples and the maximum likelihood classifier). The differences in (a) the interpolated E1 and E2 values and (b) the classified pixels were assessed through cross tabulation and expressed as percentages. Table 15.3 summarises the results and it can be seen that changes in track intensity can affect the interpolated acoustic values and have an even more marked effect on the interpreted images. Presumably, the differences on E1, E2 and depth will produce cumulative errors in the interpretation of the AGDS data.

Table 15.3. Effect of increasing track spacing on interpolated E1 and E2 values and the percentage similarity of the resulting interpreted AGDS data.

Difference compared with the full data set	Half track data	One third track data
E1	9.4%	15.1%
E2	12.7%	20.72%
Interpreted AGDS data	22%	44%

15.5 Consistent use of habitat and biotope categories

Image processing can be undermined by poor attribute measurement of the ground truth samples. This is particularly likely when the attributes are habitat or biotope classes and the analyst must decide how best to match the sample data to a classification system. Biological classes are rarely distinctive and many samples will not fit clearly into one class. They may lie between two or more classes or fit no description at all. How does the analyst cope with these situations? The temptation will be to force the sample into a class, but this runs the risk of degrading image processing.

15.5.1 Misclassification

It is often assumed when viewing a map that the classes depicted are discrete, mutually exclusive and easy to identify. However, as any field surveyor knows, this is not the case. Samples may be attributed

to different classes by different surveyors and intermediates between types seem to occur all-too frequently. These problems in labelling ground truth records have been discussed in Section 6 and their effect on interpretation of the AGDS data cannot be overstated. A sensible grouping of habitats or biotopes into broader groupings of similar appearance reduces the severity of misclassification. Validation of the way in which ground truth samples have been classified by an independent expert would help to ensure a level of consistency between surveys, but there will still be an element of subjectivity.

15.5.2 Heterogeneity of the seabed within a sample

An allied issue to that discussed above is where mixtures of two or more types occur together in the vista of a video frame (equivalent in scale to the acoustic footprint – see Section 6.4.1.1) and classification of samples will depend upon the perception of heterogeneity.

The habitats and biotopes may be heterogeneous at any scale, including areas of ground covered by the acoustic footprint. The acoustic values recorded at each point will be an integration of the reflectance properties of all the various ground types within the footprint and this will mean that the range of habitats and biotopes that are to be discriminated must encompass commonly occurring mixtures as well as ‘pure’ (homogeneous) types.

15.5.3 Heterogeneity of the seabed within the limits of spatial resolution

In practise it is unlikely that one acoustic record (of E1, E2 and depth) can be linked spatially to a particular location on the seabed and its habitats or biotopes because of positional uncertainty. Instead, acoustic records within an ‘envelope’ of combined positional uncertainty (Section 15.2.3) will be used for the derivation of an acoustic signature for a habitat or biotope type. Thus, an assessment made of spatial heterogeneity of the seabed from a visual assessment of a short video tow is more appropriate for signature development than attempting to link one record with one small patch of the seabed (see Section 6.4.1.2).

The issue of heterogeneity of the sea bed at this scale has been raised throughout the report. The nature of the sea bed and the biotopes it supports varies considerably over quite small distances in some situations (e.g., close inshore in rocky habitats) whilst in others the sea floor is relatively homogeneous (e.g., offshore sediment plains). The primary effect of fine scale heterogeneity and positional uncertainty in the data is to compromise any interpretation of the acoustic data since it is inevitable that some of the acoustic data will be attributed to the wrong habitat or biotope.

Some authors have calibrated their AGDS on homogeneous patches of ground and then used the boxes constructed for E1/E2 to detect the ground types in other areas (Chivers *et al.* 1990). Likewise it is possible to develop signatures for homogeneous areas and apply these to the whole image (Collins *et al.* 1996). However, these signatures cannot be applied with confidence to heterogeneous areas since the footprint of the sounder system and interpolation will result in average reflectance values that do not match to any ‘pure’ signature. Thus, it is unlikely that such procedures will overcome the problem of heterogeneity in practice.

Perhaps the most workable approach to heterogeneity is to accept that the component habitats or biotopes cannot be mapped accurately and find ways to represent heterogeneous areas. A simple way around the problem of heterogeneity is to map heterogeneous areas and annotate them according to their biotope composition. The use of probabilities may open up other ways of assessing heterogeneity. Fine scale heterogeneity will result in low probabilities for a wide range of habitats or biotopes rather than the hoped-for high probability for one class (Figure 15.2). This, however, gives one possible way of measuring heterogeneity. Images that show the degree of uncertainty are a good measure of heterogeneity given that the area in question has been surveyed fairly intensively.

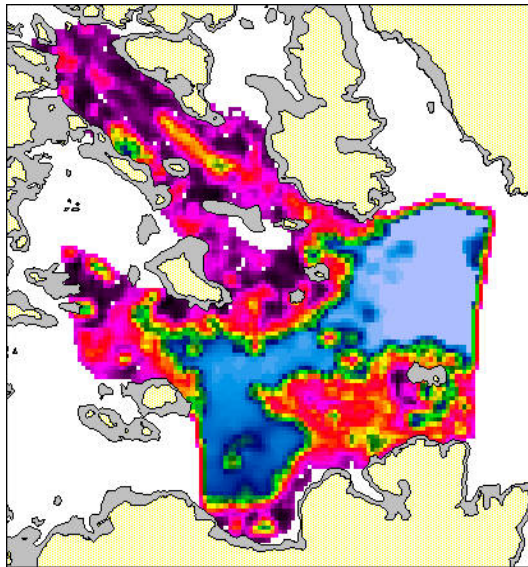


Figure 15.2. Uncertainty in the classification of the Loch Maddy study area. The blue represents low levels of uncertainty whilst dark purple and black represents high uncertainty. Since the uncertainty has been calculated using the full data set, it is presumed that the uncertainty is a reflection of heterogeneity rather than poor signature development.

This approach has a practical application to the design of monitoring surveys. It has been recommended that areas of fine scale heterogeneity are identified and this knowledge used to formulate an appropriate survey strategy (Foster-Smith *et al.* 1999). There is little point in concentrating more and more survey effort in such areas to try to improve the overall performance of biotope maps when this will be undermined by positional uncertainty (the maximum resolution) and sea floor heterogeneity.

A second way of representing heterogeneous areas is to map the probability for each biotope separately (see Section 13.6).

15.6 Information from the AGDS used for discrimination

It must be remembered that the two variables E1 and E2 (four variables if depth and a variability index are included) will limit the scope for discrimination between biotopes and the powers of AGDS should not be over-inflated. Positional uncertainty (as explained above) will inevitably mean that records are incorrectly associated with a habitat or biotope class. This will 'dilute' the acoustic signature and reduce the power for discrimination resulting in many habitat or biotope classes having similar or overlapping signatures. Very often this confusion arises between closely associated classes so that misinterpretation of the AGDS data is not entirely misleading.

Discrimination between a very small number of habitats or biotopes known to occur in a restricted area is usually good. However, as the range of habitats or biotopes increases, the possibility of overlapping signatures also increases, so reducing discrimination. The use of more broadly defined habitat or biotope classes may be considered as one way of overcoming problems with discrimination if this lower level of detail is acceptable.

15.7 Intensity of ground truthing on consistency of interpretation

Ground truth samples for large survey areas are hard-won and it is difficult to avoid using all the data for classification of the acoustic images. There is little scope for redundancy of ground truth data. This is, admittedly, unsatisfactory since the implication is that the interpretation will be sensitive to the position and number of the samples. The Loch Maddy study allowed the effect of ground truth sample intensity on AGDS interpretation to be measured since a small area was intensively surveyed and sampled.

In this study an acoustic data set was interpreted using all available ground truth samples and subsets of three quarters, one half and one quarter of the samples. The subsets were generated by grouping samples by biotope and eliminating the requisite proportion of records from each biotope category.

This was done so that the subsets had approximately the same proportions of biotope types as the full ground truth sample data set. The exercise was duplicated using a different choice of samples. The results are shown in Table 15.4.

Table 15.4. The effect of reducing the number of ground truth samples used in the classification of the acoustic data. Data from the Loch Maddy study were used.

	Three quarters of the ground truth samples	One half of the ground truth samples	One quarter of the ground truth samples
Number of samples used	66	42	22
Modal number of samples per biotope	5	3	2
Percent similarity to the full data set	78%	81%	32%

The results indicate that consistency of interpretation remained at about 80% until the number of training sites (ground truth samples) dropped below about 3 per biotope.

An alternative view of the effect of reduction in the number of ground truth samples can be gained by looking at the interpreted images (Figure 15.3). There seems to be a convergence between interpretations in the overall patterns in the biotope distributions as more samples are incorporated into the analysis. However, with small numbers of ground truth points some obvious discrepancies between the interpretation of the acoustic data emerge. For example, kelp was predicted in the outer loch where, with the benefit of larger numbers of ground truth points, sea pens might have been expected. This illustrates the unpredictable effects that low ground truth numbers can have on signature development and subsequent image classification.

15.8 Variability in interpretation

There can be many stages involved in image processing from data editing through to statistical analysis and modelling. Of course it is the intention of the analyst to improve interpretation through analysis. However, the route followed by an analyst may be hard to replicate by another person since there are many possible pathways, each with different parameters that must be set. Although each decision taken during the whole process should be taken for a sound reason, in practice many decisions are somewhat arbitrary and taken without a full understanding of underlying theory. It is the experience of the BMP that interpretation is quite robust with perhaps two exceptions. (a) Habitat or biotope classes that have not been well sampled but which have a large range of data values can sometimes swamp other classes and may need to be removed and (b) boundary conditions appear to be sensitive to differences in interpretation and variation will occur around the edges of biotopes.

It may be that the data should be independently analysed and in critical situations the extra cost may be justified.

15.9 Cartographic scale and resolution

Cartography is a mixture of aesthetics and science. There is generally a compromise between scale, map detail and ease of use of the map. Too much detail will render a map confusing and hard to read. Too little detail will reduce the value of a map. However, broad scale maps may need to be simplified. Simplification can introduce errors when an analyst has to decide how best to combine classes that occur in mixtures (of small clusters or isolated pixels) at a scale that cannot be represented on a map.

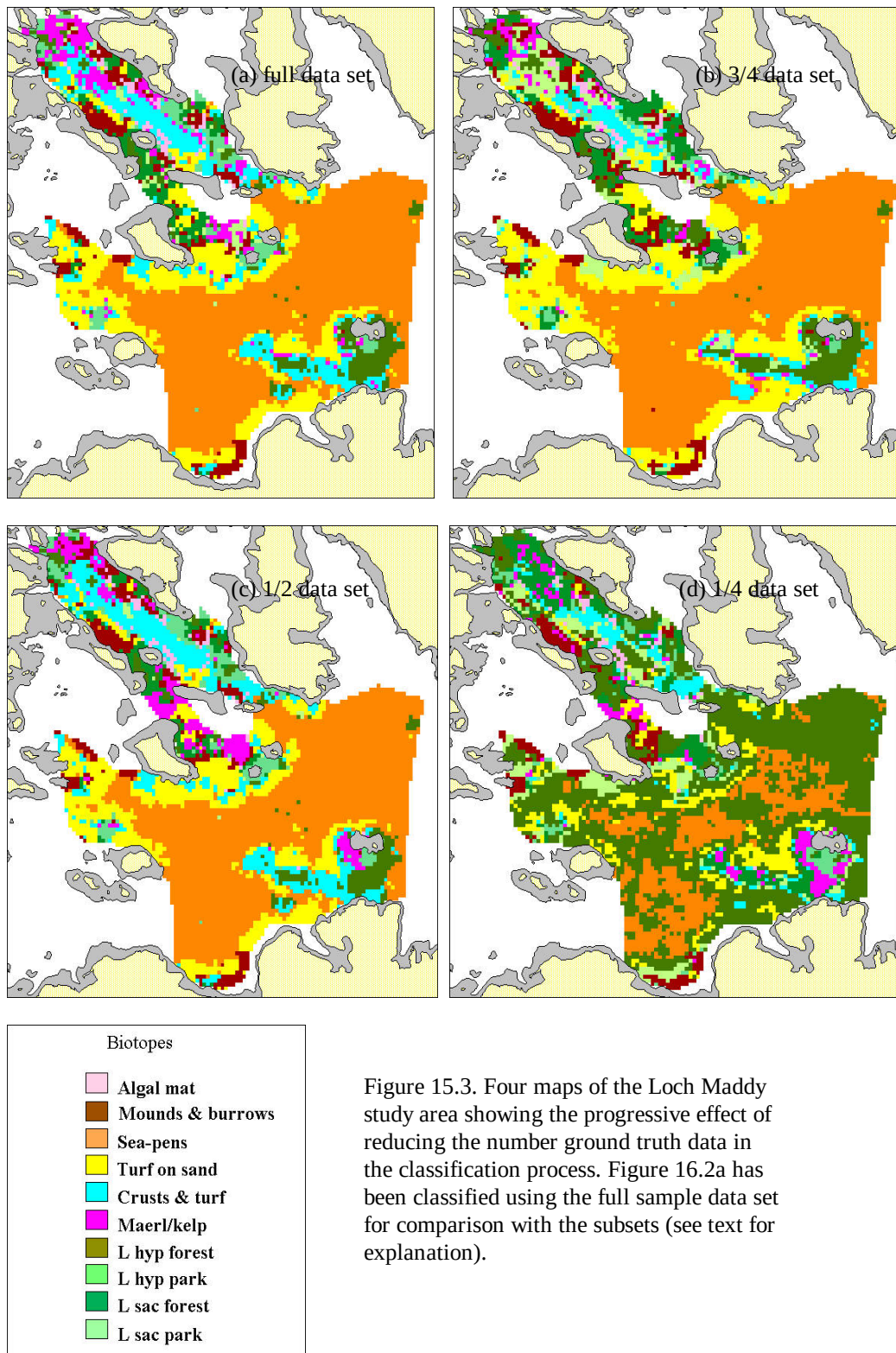


Figure 15.3. Four maps of the Loch Maddy study area showing the progressive effect of reducing the number ground truth data in the classification process. Figure 16.2a has been classified using the full sample data set for comparison with the subsets (see text for explanation).

This may involve a reduction in the number of isolated pixels in a digital image through filtering and the combination of small classes (e.g., biotope classes) into more broadly defined classes. Grid spacing chosen for interpolation of estimated values (and the pixel size of the images derived from the grid) might reflect resolution, but is not its equivalent. As discussed in Section 12, grid interpolation is

based on the spatial pattern of the track data and the intensity of the track spacing. It must be stressed that the additional grid values over those which are strictly required to reflect the resolution of the data will be just as valid as any other of the estimated values; there will simply be more of them. Thus, as a general rule, grid spacing should be as fine as possible without incurring heavy penalties in terms of file size and data processing.

It is possible to show fine detail in the case of a survey employing very fine track spacing. However, a small pixel size will create cartographic problems since isolated areas less than about 2mm x 2mm are difficult to show on a map without creating a cluttered appearance. Filtering can be used to reduce the number of isolated pixels. This will result in the loss of detail for the sake of clarity. If more detail is required, then a finer scale map will need to be produced.

On the other hand, a fine grid spacing may be used even where track spacing is wide as a way of avoiding the 'tiled' effect often found with wide grid spacing. A line ceases to have an obtrusively 'stepped' appearance when the pixel size on a map is less than about 0.05mm in size. Thus, a grid size could be chosen to ensure that this size is not exceeded. Filtering can also be used in this situation in conjunction with the fine grid spacing to reduce the number of isolated pixels. Since these are unlikely to be well supported by the track data little loss of information will result.

For aesthetic reasons, the tiled appearance of a map can be softened by sub-dividing pixels (into quarters or 16ths) and then filtering the resulting image. This process smooths lines by rounding corners off the original tiles.

16 The use of 3-D bathymetric models and drapes for interpretation.

Bathymetric data can be used to create three dimensional models of the seabed, termed **digital elevation models** (DEM) and many contouring and GIS packages provide modules for their construction.

Not only do these models provide a visual medium for explaining topography, but other information can be draped over them, such as biotope or probability maps, which displays the information in a very striking way. In this capacity DEMs bring out connections between seabed features and biotope distribution that would be extremely difficult using flat map (plan) displays.

Editing and interpolation of the depth data are critical stages in the construction of a DEM since any spurious data that survive editing or are introduced through interpolation will result in small but visually intrusive bumps and dips in the DEM. Draping the track data over the DEM will show if some of these 'features' are associated with singleton records that should be removed. Interpolation necessarily involves a degree of smoothing (see Section 12) and this may be emphasised in the process of DEM construction by using a linear interpolator (or inverse distance to the power 1.5) and setting the minimum number of records for the calculation of estimated values to be larger than might be used for image processing. The degree of smoothing is a matter of choice, but the model cannot support small scale topographic features which cannot be reliably detected by the track records.

The following are examples of how three dimensional models have been used in the Broadscale Mapping Project.

16.1 Topography of the Firth of Lorn

The underlying geological formations and the retreat of ice following the last ice age are responsible for most of the present landscape features of the Firth of Lorn. The solid geology of the seabed is complex and has had a significant influence on the subtidal landscape. For example, differential erosion of the bedrock by glacial activity created a number of deep troughs within the Firth of Lorn and the adjacent sea lochs. Offshore the seabed is covered with extensive deposits of Pleistocene sediments, mostly comprised of mud.

Central Firth of Lorn was the most topographically complex part of the whole study area: the wave exposure varied from exposed to sheltered, tidal streams ranged from slow to very strong, depth varied from 0 to greater than 200m, and a full range of seabed types were recorded. This high habitat diversity was reflected in a very high biotope diversity throughout the area, where most of the biologically richest examples of the biotopes recorded during the whole study were found in the central Firth of Lorn. These conditions were typified by the seabed around the Garvellachs (Figure 16.1).

To the south-east of the Garvellachs the seabed slopes steadily down into a basin at a depth of greater than 230m. Nearshore habitats were predominantly rocky giving way to muddy sand with stones and shells and then to mud in the deepest regions. In contrast, to the west of the Garvellachs there were two deep (> 100m), near vertical-sided glacial troughs separated by rock outcrops where the shallowest depths were < 20m. The western side of the outermost trough comprised a massive wall, overhanging in places (personal observation from scuba diving) extending for more than 10km in a north/south direction. Further west there was an extensive reef of rugged bedrock.

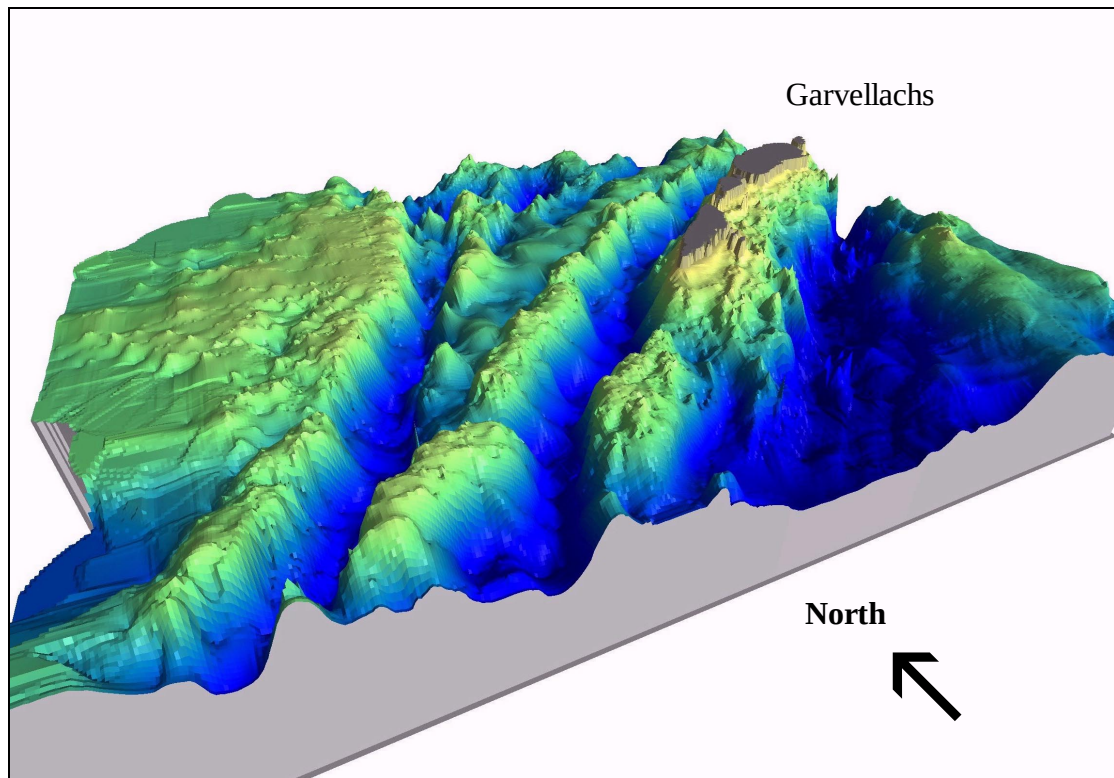


Figure 16.1. A three dimensional representation of the seabed around the Garvellachs, central Firth of Lorn

16.2 Draping life forms over DEMs in Pembroke

Draping biotope information over DEMs helps to explain their distribution, but can also be misleading in certain situations. The example from Pembroke illustrates both of these aspects to the use of three dimensional visualisation. Figure 16.2 is a view of Skomer and Skokholm looking towards Marloes Peninsula from the west. The map of the life forms/habitats as shown in Figure 6.8 has been used in this example. Note that the land has been given the notional depth of zero for interpolation in the construction of the DEM.

The vertical axis has been exaggerated in order to accentuate the topography of the seabed and this demonstrates how the bedrock areas (dark blue) drops into deep water close inshore and thereafter the reefs are rather uneven. The coarse sand bank between Skokholm and Skomer is also very conspicuous. The drape model also illustrates the level gravel/cobble plain that is so extensive west of Skomer and Skokholm.

However, the model gives the impression that submerged cliffs of Skomer are also of cobble. In fact the area occupied by the bedrock cliffs in plan view is very narrow and was poorly sampled by boat due to access. The three dimensional view has greatly exaggerated the cliffs and makes this poor match between substratum and topography very obvious. The plan map would need to be edited subjectively to correct for this and this step might be justified in this case. Other cases where habitat and topography are in conflict are harder to detect.

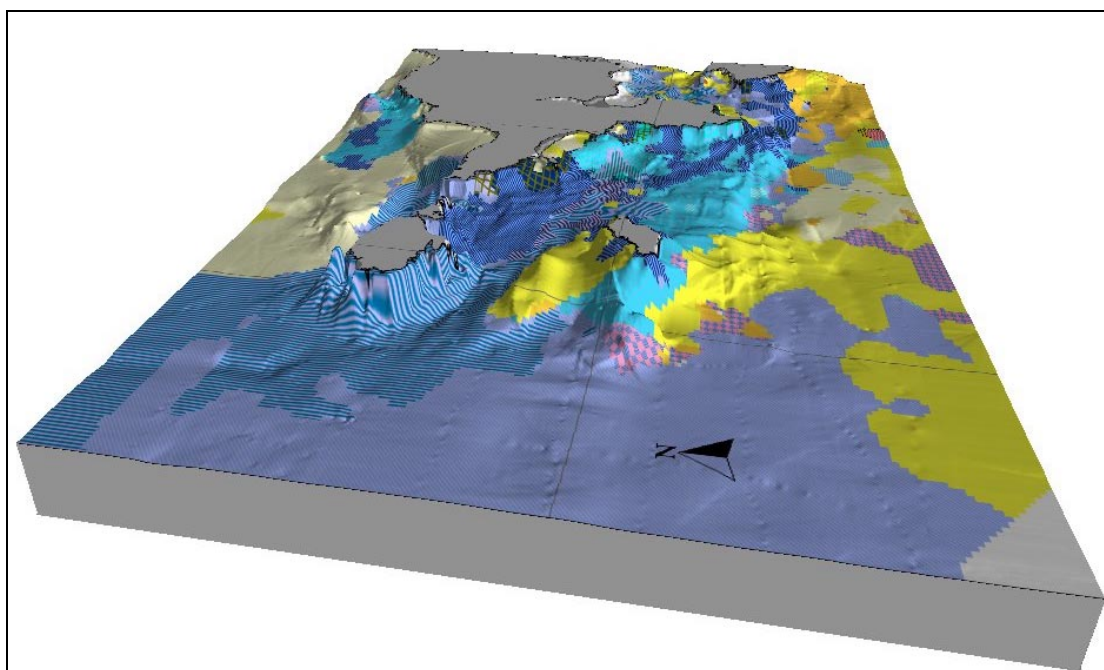


Figure 16.2. The habitats/life form categories (see Figure 6.8) have been draped over the DEM for Skomer, Skokholm and Marloes Peninsula as viewed from the west.

16.3 Draping the probability distribution of *Sabellaria spinulosa* for an aggregate dredging area off Lincolnshire

Probability maps (see Section 13) give much more information on the distribution of individual biotopes than maps showing maximum likelihood. Figure 16.3 shows the likelihood of *Sabellaria* reefs being present off the south Lincolnshire coast. The black outline shows the boundary as defined by maximum likelihood (i.e. it shows the area where *Sabellaria* reef is the most likely biotope). This shows the central part of one of the maps in Figure 13.3 at a smaller scale. The red polygon is a licensed area for aggregate extraction and the indications are that reefs run north/south through the aggregate extraction area. Does this distribution, and particularly the area of greatest likelihood, coincide with any particular topographic feature? This probability map has been draped over the DEM in Figure 16.4. The *Sabellaria* reef coincides with an east-facing slope bordering a channel which is an extension of the Lynn deeps in the Wash. The slope is one side of a ridge of coarse shelly sand. This visual display of the distribution of the *Sabellaria* in relation to topography may not be conclusive but suggests associations which might be the subject of specific studies.

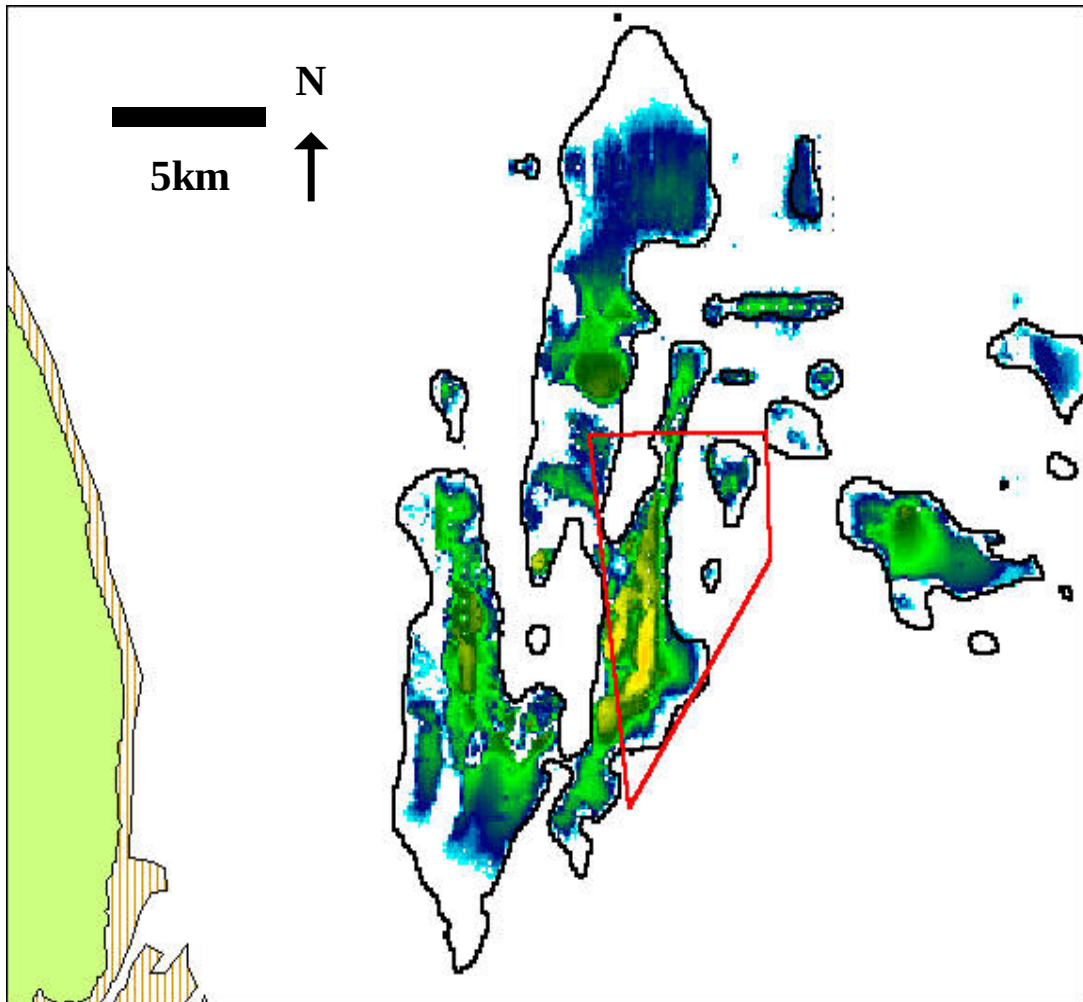


Figure 16.3. The probability distribution of *Sabellaria* reefs within the areas where they have been predicted to be the most likely biotope (area marked by the black outline). The red polygon marks an aggregate extraction area.

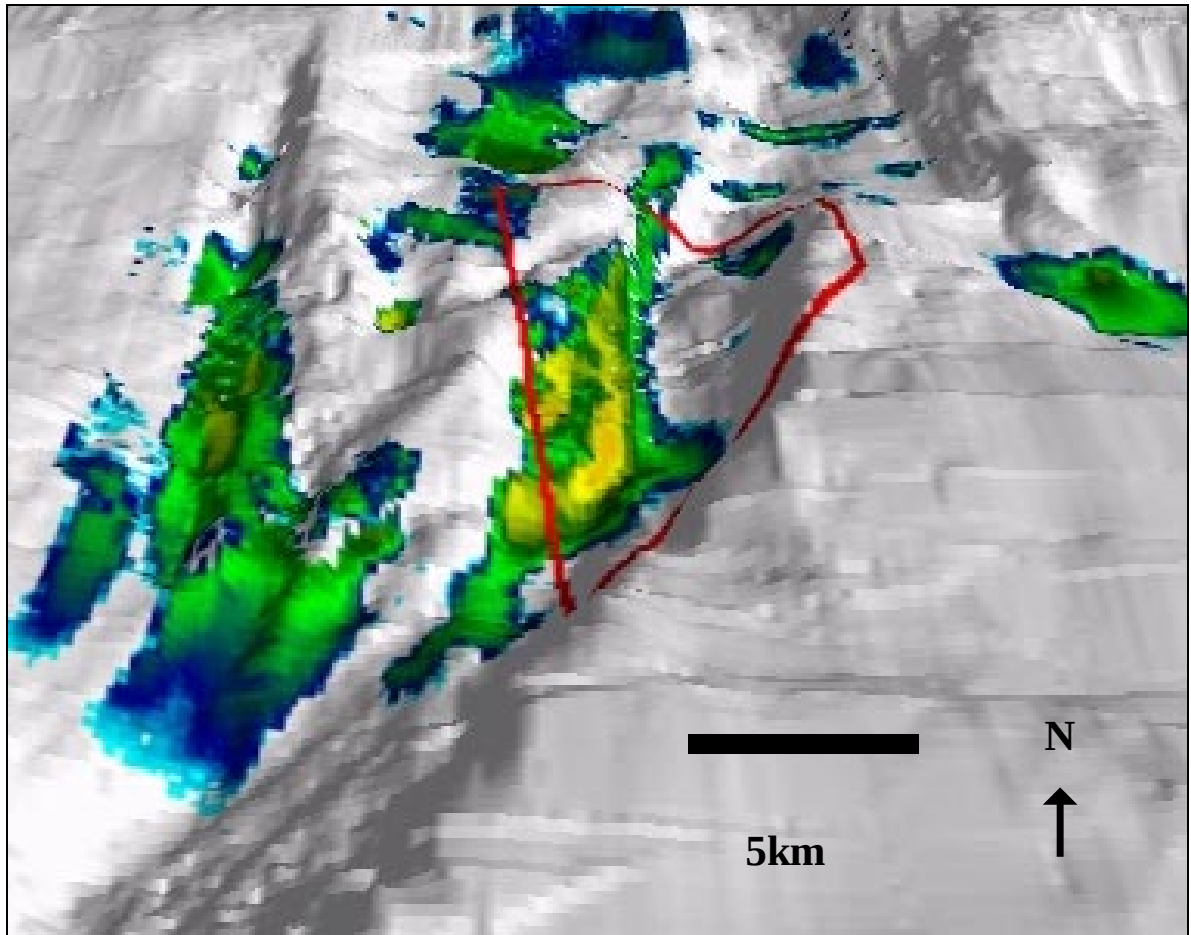


Figure 6.4. The probability of the occurrence of *Sabellaria spinulosa* reefs in the vicinity of the aggregate extraction area 107 (marked with a red outline). The probability map for *Sabellaria* has been draped over a DEM for the area.

17 Other AGDS systems.

17.1 Multi frequency echo sounders

Simultaneous dual frequency was used extensively in the Firth or Lorn survey and in the much smaller Loch Maddy trial area. The results of the discrimination attained using the dual frequency as compared with the single frequency were somewhat equivocal. Because of these results, and given the patchy coverage of dual frequency data for the other trial areas, analysis of the dual frequency data was not attempted for Pembrokeshire and the Wash.

17.1.1 Loch Maddy

The most encouraging study was conducted in Loch Maddy. The data from the dual frequency system were classified using the data from each of the frequencies separately and then the combined data. The match between the ground truth data and the images were assessed for internal accuracy and the results are shown in Table 17.1 below. However, it must be remembered that a higher internal accuracy does not necessarily mean that the interpretations are necessarily better.

Table 17.1. Error matrices for each treatment

Frequency of Dual system	Accuracy
200kHz	0.5928
38kHz	0.6242
200 & 38kHz	0.7692

17.1.2 Firth of Lorn

Prior to using data from both frequencies in the classification process, it was important to assess whether there was any benefit over classifying each frequency independently. Initially data from

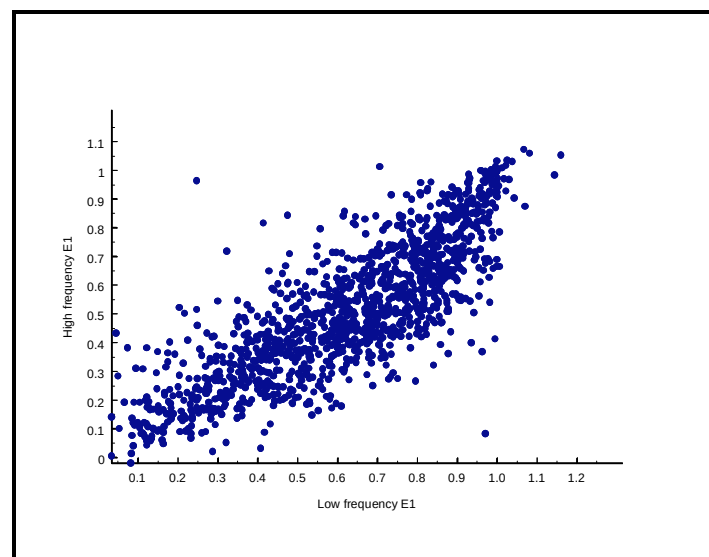


Figure 17.1. Relationship between E1 values recorded simultaneously by the low and high frequency RoxAnn systems. A similar relationship was observed for E2 values.

each frequency were plotted against each other to determine whether they were highly correlated and therefore not actually contributing any additional information to a multi-frequency classification. Although there was a broad relationship between the two frequencies (Figure 17.1) there was sufficient variability in this relationship to suggest that both frequencies might yield different information about the sea floor. These results were considered sufficient justification to use both low and high frequency data together in the classification process to predict biotope distribution. E1 and E2 are correlated for each frequency, and there was a broad correlation between frequencies for each variable. All four *RoxAnn*TM variables were submitted to a principle components analysis (PCA) and the first three axes explained the majority of the variance in the original data.

The data from the dual frequency system were classified using the data from each of the frequencies separately (E1, E2 and depth), the combined data (depth, high E1 and E2, low E1 and E2) and the results of the PCA analysis (depth combined with PCA 1, 2 and 3). The match between the ground truth data and the classified images was assessed for internal accuracy and the results for the classification of life forms in the Firth of Lorn are shown in Table 17.2 below.

Table 17.2. The results of internal accuracy for each habitat/life form and the overall accuracy for data from the dual frequency system for the Firth of Lorn.

Life form	All	Low	High	PCA
Infralittoral rock	59	61.5	56.5	53
Mixed infralittoral	46.5	28.5	11.5	41
Circalittoral stones	30.5	11	26.5	20
Circalittoral rock	43.5	59	31.5	46.5
Brittlestar beds	32	9.5	28	24.5
Silty stones	49	4.5	18	41.5
Gravel	39.5	8.5	14	27
Sand	59	46.5	41	47
Muddy sand	43.5	28.5	11.5	28
Circalittoral mud	62.5	70	59	66.5
Overall	43	40	33	38

It should be noted that these observations were not fully tested with independent ground truth samples. However, it would appear that there is a benefit to using the dual frequency system over a single frequency system, although the difference in the internal accuracy was marginal and not be reflected in predictive accuracy of the maps.

17.2 QTC View

At the outset of the BMP *SeaMap* sought to broaden its experience of AGDS through the purchase of a *QTC View*TM system and extensive trials were conducted in Northumberland, Pembroke and the Firth of Lorn. However, the system was not considered to be sufficiently reliable at that time and, whilst the potential for its use was appreciated, there was not sufficient time available to the BMP to persevere with further trials. However, the ways in which *QTC View* might be incorporated into the BMP protocol for survey described below.

The benefits of the system is that the acoustic echo is much more completely analysed than is the case for *RoxAnn*, as explained in Section 4.3.5. This, potentially, gives the system more scope for acoustic discrimination between ground types. However, the standard *QTC* system (at the time the trials were conducted) reduced the output to three principal components which explain the majority of the variability of the acoustic data. These were designed to be displayed in a real-time as three dimensional parallelepiped blocks, akin to the boxes in the *RoxAnn* real time display (see Section 13.4),

although the extent to which a record belongs to a defined block is measured and displayed as a confidence measure.

However, unlike E1 and E2, these PCA values were not measurements as are E1 and E2. Thus, the scope for exporting these values into an image processing system for supervised classification was unclear. It is understood that recent versions of *QTC View* allow researchers to record many of the measurements of the echo and thus there may be the potential for multispectral (or even hyperspectral) classification.

In its standard form, *QTC View* had another drawback, especially for portable AGDS. If the same calibration were to be used, there would be no certainty that the system would be responding to the seabed in the same way and the effect on the PCA would be unpredictable. The system would need to be calibrated for each set-up and it is uncertain whether the same variables are recorded and used in the PCA. Each survey would need to be interpreted separately and the various interpretations amalgamated.

Whilst this is not impossible (and was done to incorporate the survey of the Smalls into the main Pembrokeshire map) the drawbacks are that interpretations of the different data sets would not match exactly (and may not match even roughly) which would almost certainly create problems in editing the various interpretations so that a consistent final map is produced.

18 Integrating sidescan into broad scale mapping.

Sidescan sonar is a technique for recording images in a broad swath of the surface morphology of the seafloor. An important attribute of side scan sonar is its ability to achieve high resolution of seafloor features: high frequency systems operating at short ranges (500kHz at 50m range) can resolve features such as cobbles, small boulders, and sand waves/ripples. Side scan sonar data are recorded as a vessel moves along a track and therefore the images show a strip of the seafloor. Computer technology can join these strips by their respective geographic co-ordinates to form accurate maps of seafloor structure. Overlapping sidescan images can be joined together in a process termed mosaicing. It is, therefore, possible to create maps of topographic features over areas of the seabed that will be accurate within the positional accuracy limits of the system and the ability to distinguish between the different features from the images. Images can also be transferred to geographic information systems (GIS) to link with other data to further improve image interpretation and map production.

The images can be interpreted visually into categories of features and video can be used to check the feature identification. The images can also be geo-registered so that polygons can be drawn around features to create a vector map of the categories.

The BMP used sidescan in conjunction with AGDS. There are two approaches that are possible using the two systems in combination:-

1. **Feature location:** In the first approach the sidescan is used more as a ground truthing technique. Training sites are selected from the images which are then used to classify the acoustic data.
2. **Mapping through mosaics:** In the second approach the seabed is surveyed intensively by sidescan and the AGDS used to give more information, particularly on sediment composition, to supplement what can be ascertained from the sidescan images alone (Curran 1995).

The first of these approaches was used in the BMP much more than the second since building up maps for large areas with sidescan sonar is very time consuming both in terms of survey and analysis. Optimum image quality was achieved when the vessel was travelling at very low speeds (1-2 kn). Whilst side scan sonar provided very detailed images, this slow rate of coverage did not fit with a broad scale survey strategy and mapping using sidescan was therefore restricted to a small number of sites. Nevertheless the pictures and mosaic images extracted gave a clear indication of the potential for this technique to make a significant contribution to an overall broad scale mapping strategy. Pictures of a large area of complex seabed can help interpret video recordings by providing a broad spatial context. Similarly, a mosaic of a complex area can assist the interpretation of adjacent *RoxAnn*TM tracks.

18.1 Northumberland

Sidescan sonar was used as a ground truthing technique extensively off the north Northumberland coast to describe and map the reef system. Seven different seabed types were distinguished, although some were more distinctive than others. These were soft sand, coarse sand, gravel, sand/gravel/cobble/boulder mix, sand with reef outcrops, reefs with sand, rugged bedrock reefs. Figure 18.1 shows an example of a sidescan image of a transition from sand to reef with the AGDS data superimposed. This illustrates how the AGDS data correlate with the obvious features from the images.

Figure 18.2 is of the classified AGDS data for north Northumberland using training sites drawn from the topographic features identified from the sidescan images. The classification process has only used E1, E2 and variability. Depth was not used since the distribution of the topographic features should be independent of depth. The distribution of the reefs and other substrata are very distinct and, although no external accuracy has been performed, the results are a more convincing representation of reefs in the area than have been produced using the video samples.

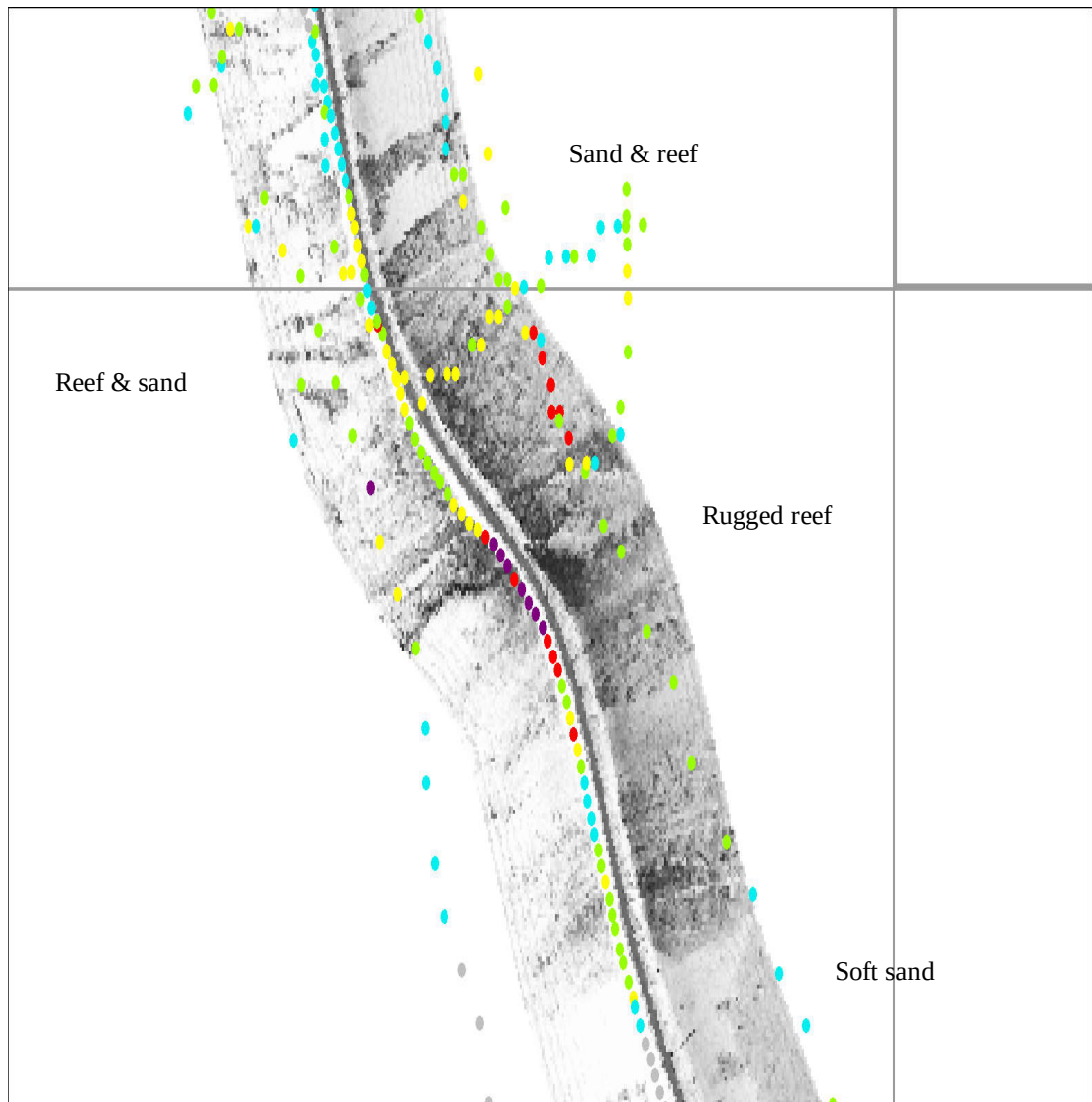


Figure 18.1. An example of a sidescan image of reefs and sand off the north Northumberland coast with AGDS data superimposed, coloured according to E2 (hardness). Purple represents high values, light blue and grey represent low values. The width of the swath was approximately 400m. The right and left hand sides can be joined together by removing the space immediately either side of the centre line that represents the water column beneath the boat. This has not been done in this case so that the depth profile can be seen.

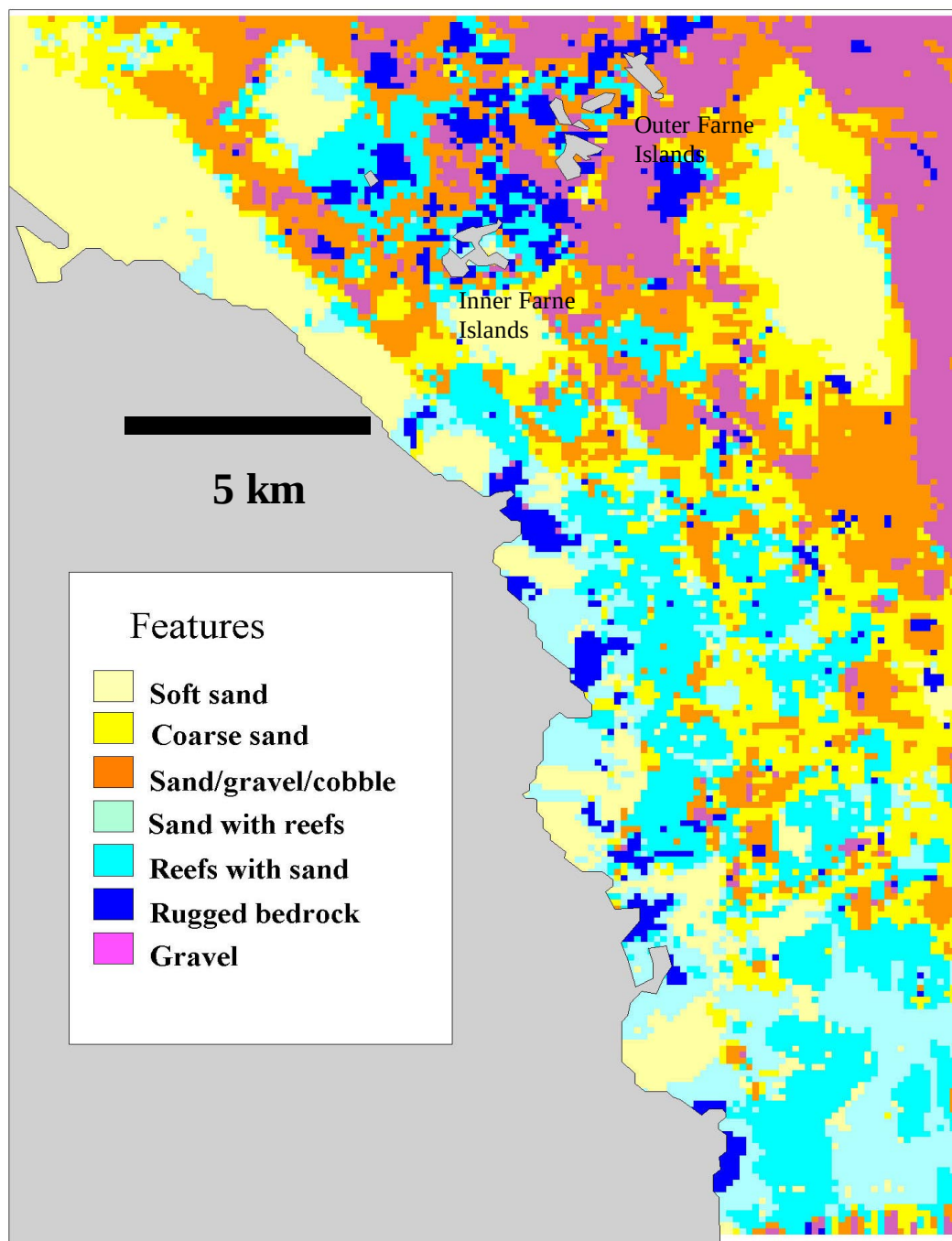


Figure 18.2. Main seabed features of north Northumberland identified from AGDS interpreted using training sites selected from sidescan images.

18.2 Firth of Lorn

During data acquisition and in replay mode, the EOSCAN software provided a *waterfall* display of the seabed passing under the tow fish. At any point in time, the display could be frozen and the image saved as a picture file (Figure 18.3). The dimensions of these pictures will depend on the sonar range (width of the image) and the speed of boat over the ground. These digital pictures can be manipulated

using appropriate software to, for instance, improve the sharpness, or to reduce the size of the image by discarding unnecessary material. Geographic co-ordinates can be derived for these pictures and so they can be transferred to a geographic information system (GIS) to link with other data.

Figure 18.3. is an image of the left half of the screen display (the tow fish travelled along the right edge of the image) showing an area of seabed at the Torren Rocks approximately 190m long by 160m wide, at a depth of 18m. It reveals a number of interesting seabed features: there are rugged bedrock reefs in the top and bottom right separated by waves of sediment in the centre. The reef in the upper right is criss-crossed by small gullies/fissures and becomes more broken towards its left edge. The reef in the lower right appears in three parts separated by sediment. It is possible to estimate the height of the reef from the length of its shadow – approximately 8m high. Similarly it is possible to estimate the wavelength of the sand waves by measuring the distance between the troughs (dark lines) – approximately 3m in the centre of the image, 1.5m towards the top. In the left half of the image there are some isolated reefs surrounded by sediment although it is more difficult to distinguish any features.

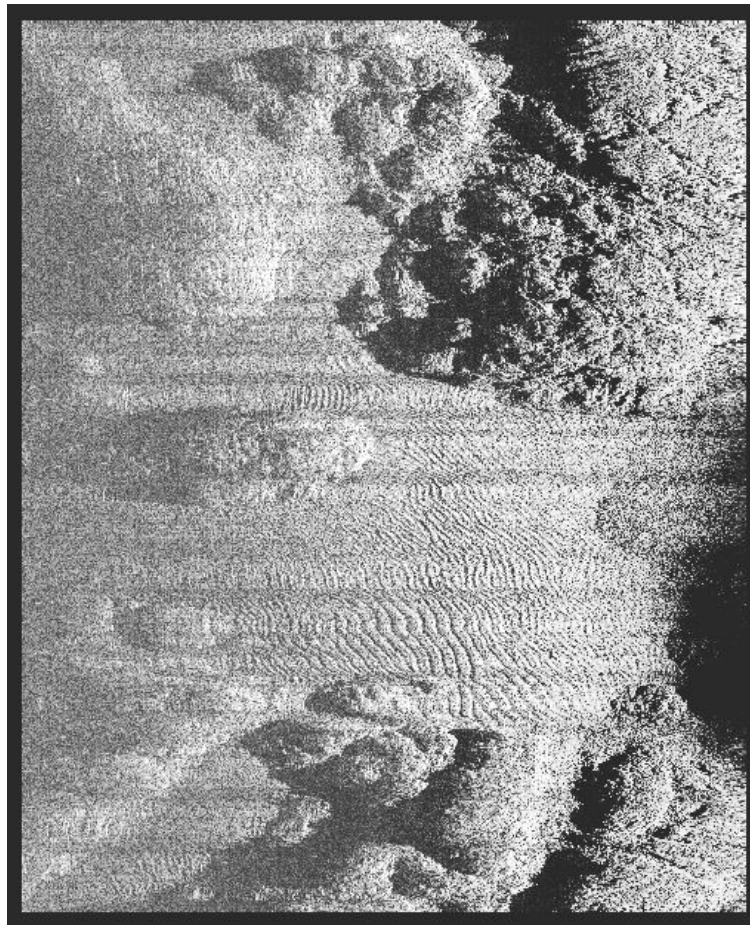


Figure 18.3. A sidescan sonar image showing details of rock reefs and sediment waves west of the Torren Rocks, the Firth of Lorn.

Figure 18.4 shows a portion of a side scan sonar mosaic around the west Torren Rocks showing an area of seabed approximately 2km long by 2km wide, at depths between 20-30m. It provides an interesting insight into the spatial arrangements of the major seabed habitats at a broad scale: there are extensive rocky reefs with large areas of sediment. Within these broad areas however, it is possible to discriminate sand-filled gullies within the reef and rocky outcrops on the sediment.

Much of the complex small-scale pattern revealed by these sidescan images were not apparent from the *RoxAnn* AGDS data. The rugged nature of the reef formation shows many gullies, cracks and fissures which were filled with sediment; ripples and waves were visible in the sediment indicating rapid water movement within these gullies. Video records showed changes in the biotic composition

associated with edges and vertical steps although it was not possible to gain a clear understanding of their spatial arrangement from the video. Mosaics illustrate these features and shows how side scan sonar can contribute to the understanding of spatial patterns on these complex reef formations.

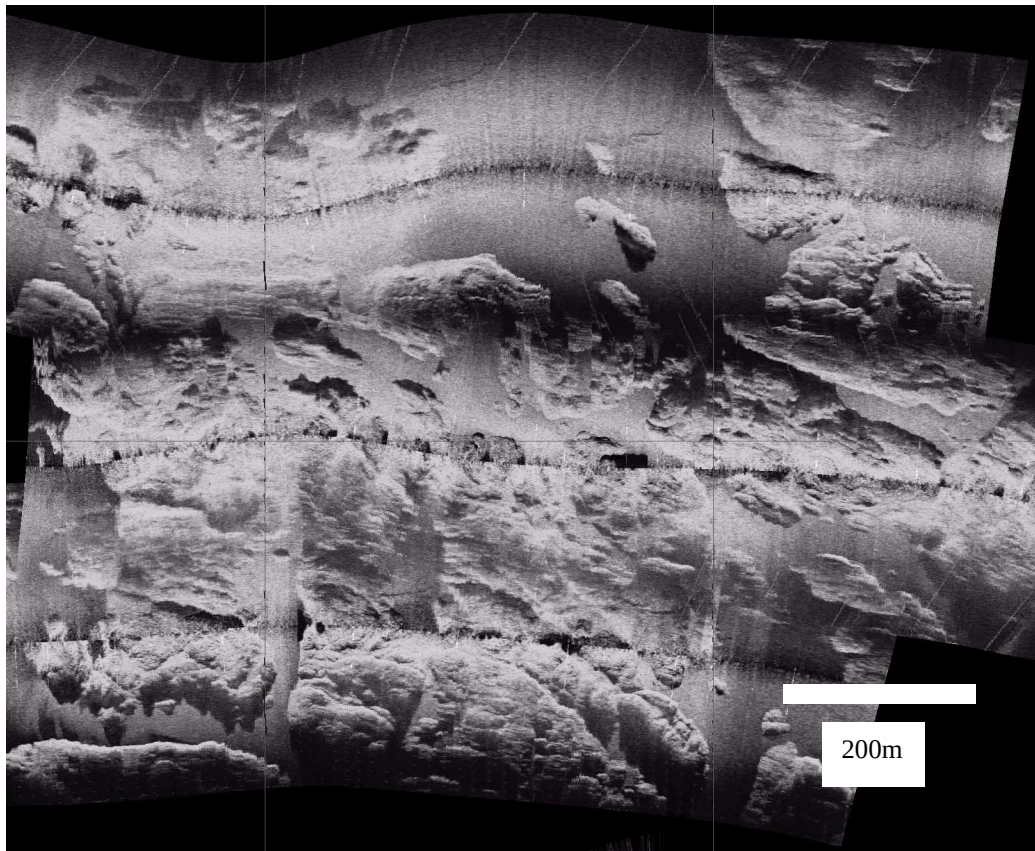


Figure 18.4. An small example of the mosaic of side scan sonar lines at the west Torren Rocks.

19 Guidelines for broad scale survey

Purpose: To give summary guidelines on planning, carrying out and interpreting broad scale mapping surveys as developed during the BMP program.

19.1 Preliminary planning

1. Is a broad scale survey needed (at least as part of the information requirements) and is it feasible?

Broad scale survey is defined by the application of remote sensing to the survey. There are many applications of broad scale survey and mapping to management of the coastal zone (Section 3) and the following is a series of questions that need to be addressed in the preliminary planning stages of any proposed survey where broad scale mapping is considered:-

- Are maps showing the distribution and extent of habitats, biotopes or some other attribute as a continuous coverage required?
- What are maps expected to show in terms of (a) the level of descriptive detail of habitats, biotopes or other attributes, (b) spatial resolution of mapped entities, (c) accuracy of boundaries and, (d) confidence and repeatability of the interpretation?
- What are the technical limitations? Is it possible to provide an output that matches the expectations?
- What are the limitations due to cost and time? Are there any ways in which an effective compromise can be achieved that provides an acceptable output for the resources available?

2. What is the most appropriate survey strategy?

If it is decided that a broad scale survey is required, at least as part of a survey strategy, then there are many subsidiary issues which need to be considered when deciding what type of survey best suits the aims. Figure 8.2 shows a decision tree that can be used to help decide what type of survey is most appropriate to the situation and the reader is referred to Section 8 for more details, including an estimation of costs.

3. Will the survey be integrated into a wider strategy for gathering information?

Broad scale survey should be seen as part of an integrated strategy for collecting information for management. Will the output of broad scale survey set the context for more detailed survey (see Table 8.1 in Section 8)?

19.2 Expectations of a broad scale survey

It is important to have realistic expectations of broad scale survey and this Technical Report has stressed the limitations of broad scale survey and the potential sources of error. These can be grouped under the following headings:-

1. Positional error and maximum technical resolution of the survey methodology

The corner stone of interpreting acoustic data is the use of ground truth data located within the acoustic data. Position error reduces the ability of the system to discriminate the ground type and the combined positional error of the boat and the sampling device error can be quite large (see Section 15.3). However, AGDS can be used to its maximum technical resolution in surveys designed to determine boundaries in small areas where there are a small number of distinct acoustic ground types in large homogeneous patches.

Although positional error can be reduced by careful use of DGPS and survey design, it is unrealistic to expect high positional accuracy for most broad scale surveys.

2. Survey intensity and acceptable resolution

Intensive survey is expensive and survey of all but the smallest areas is likely to involve a compromise between survey effort and resolution. This involves both track spacing and ground truth sampling intensity. Wide track spacing results in gaps in the data and it must be decided if interpolation is required to 'fill' these gaps. Broad scale surveys will inevitably miss many small features and boundaries cannot be expected to be accurate. Nesting survey intensity is suggested as a compromise for maintaining a high level of detail for small, selected locations within a wider survey area .

Economies in the number of ground truth samples will shift the emphasis of an acoustic survey from the use of the acoustic data to substantiate boundaries around the ground truth samples to the interpretation of AGDS data by extrapolation from ground truth locations to other parts of the survey area where there is no ground truth data. This extrapolation may be justifiable for very broad scale survey of areas where there is little existing information on the distribution of biotopes, but might result in poor predictive maps for other purposes.

3. Ground truthing and information requirement

Poor ground truth sample identification, misclassification and inconsistent identification will seriously undermine any survey, and broad scale mapping is no exception. This is especially the case where samples are tagged with the predominant biotope. If biotope identification is critical then an independent check on identification might be built into the survey. This may involve re-analysis of the original data where consistency is crucial and will have implications for survey costs.

Alternatively, identification to a more general level (biotope complex or life form) may be adequate and independent verification may be unnecessary.

Ground truth data labelled according to the predominant biotope might be sufficient for the interpretation of general biotope distribution maps. However, there are other ways in which biotopes can be shown, such as through the use of probabilities. Probability maps may be a much more realistic way of displaying the distribution of selected biotopes.

Other specific attributes of the ground truth samples (e.g., species, abundance and biomass) can be superimposed on the biotope distribution map and this can be an effective way of showing possible connections between attributes of interest and habitat/biotope distribution. This mixture of 'hard' biotope categories for setting general context with attribute data overlain might provide the ideal mapping solution for many monitoring aims.

4. Acoustic discrimination

It must be remembered that the two variables E1 and E2 (four variables if depth and a variability index are included) will limit the scope for discrimination between biotopes and the powers of AGDS should not be over-inflated.

Discrimination between a very small number of habitats or biotopes known to occur in a restricted area is usually good. However, as the range of habitats or biotopes increases, the possibility of overlapping signatures also increases, so reducing discrimination. The use of more broadly defined habitat or biotope classes (such as biotope complexes or life forms) may be considered as one way of overcoming problems with discrimination if this lower level of detail is acceptable. The distribution maps need not be restricted to the use of one consistent level of discrimination. For example, biotopes and life forms can be mixed to make the best use of acoustic discrimination.

There might be broad scale trends in a large survey area encompassing very different physiographic features. The wide range of ground types and biotopes might compromise the development of signatures for classification in such large areas. The result may be overlapping signatures and poor discrimination. It may be necessary to separate the survey area into two or more parts on the basis of their physiographic features rather than attempt to treat the survey area as one unit.

5. Heterogeneity of the sea floor

Classification of the acoustic data is compromised by positional uncertainty and heterogeneity of the ground. It will be impossible to map the patches of biotopes accurately if they are smaller than the maximum resolution of the survey. These heterogeneous areas can be labelled according to the major biotope or by the suite of biotopes likely to be found.

If it is important to sample the biotope composition more quantitatively, then it is recommended that areas of fine scale heterogeneity are identified early in a survey and this knowledge used to formulate an appropriate survey strategy. For example, one component of a monitoring survey might be the statistical description of selected heterogeneous sampling areas within using stratified random sampling. The analysis of the acoustic data in terms of probabilities suits this approach to survey. It also does not place undue emphasis on positional changes in biotope distributions that, as demonstrated in many of the tasks in this survey, are extremely hard to explain with any certainty.

6. Classification

It is important to realise that there is uncertainty in the classification process and that the automated choice might have been made between two or more biotopes with very similar likelihoods. For this reason mapping the probability distribution of individual habitats or biotopes may be preferable to maximum likelihood classification for many purposes.

Additionally, different analysts are likely to produce different interpretations of the same data (through choice of interpolation parameters, training sites etc.). If it is important to assess how accurate the interpretation is, then additional ground truth samples should be collected in order to calculate an external accuracy error matrix. In addition, the extra samples could allow for re-interpretation using various combinations of the total ground truth sample, a process analogous to statistical 'bootstrapping' (Manly 1997).

7. Scale and cartography

Many of the above issues are related to the scale at which the maps are required. Surveying to a fine level of detail (both spatially and in terms of habitat or biotope description) would be pointless, for example, if the results are too small to be shown on a map produced at the required scale.

However, it is always possible to generalise data and it is safer to survey in more detail than is strictly necessary (if costs permit) than to survey to the required scale.

19.3 Field survey design

The purpose of field survey is to obtain the best representative coverage of the survey area in terms of acoustic data and ground truth sample data within the constraints of cost and time.

The preliminary plan for survey will establish in broad terms what intensity of survey will be required, where tracking and ground truth sampling may be targeted and the expected range of habitats or biotopes. This plan will be based on survey requirements and available knowledge of the area.

The survey plan will need revision as the data are collected and assessed and this process is iterative in that the ground truth locations will be determined by a preliminary assessment of the acoustic ground types and, in turn, tracking will be targeted in areas of uncertainty as determined from the ground truth samples. The output from the field survey should be assessed in terms the original plan and to ensure that the acoustic data and the ground truth samples are representative of the area. Field survey design are discussed in Section 8 and is summarised in Figure 8.3.

19.3.1 The survey plan

Track spacing and intensity, which depends on survey type, will need to be established and a plan of suggested tracks drawn for the area. This should take account of priority areas and make provision for

alternatives in case of adverse survey conditions, equipment breakdown and other unforeseen constraints.

Track spacing should be near the limits of positional accuracy for intensive survey (approximately 25m spacing). A track spacing between 50-250m (depending on heterogeneity) would give reasonable broad scale coverage whilst wider track spacing greater than 250m is only acceptable if a very general broad scale picture is acceptable.

The extra costs involved with surveying using narrow track spacing are considered to be minimal for small areas, but they will be significant for large areas.

Ground truthing will require a minimum of 5 samples for every major habitat or biotope type to enable interpretation to be carried out with a reasonable level of accuracy. However, this will vary with the predictability of the biotopes as determined through the course of the survey and more samples may be needed where predictability is low. More intensive ground truthing may be required to meet specific objectives.

19.3.2 Field survey

Real-time display and editing of track data will, in the first instance, help guide the specific location of the ground truth samples. This might be augmented by unsupervised classification carried out off-ship during the field survey. As soon as the ground truth data become available, these will be used to refine the acoustic survey.

19.3.3 Preliminary assessment of survey data

At the end of the field survey there should be a good understanding, albeit without advanced analysis for support, of the distribution of acoustic ground types and what they mean in terms of habitat or biotope distribution. The outputs should be assessed in terms of the original plan and any significant departure should be justified. External assessment of the data, particularly the ground truth samples, might be desirable.

A field report should give details of these three stages together with a preliminary assessment of the results of the survey. A visual presentation of the track data (e.g. see Section 11.2) is all that is required at this stage. The data should be in a format that can be easily taken forward for interpretation and it should be possible to pass the data and the field report on to a different team for analysis.

19.4 Analysis

Analysis of the data can take many forms and the choice will depend on the nature of the data (track spacing, intensity of ground truth sampling) and the size and range of physiographic features present in the survey area. Analysis is relatively straightforward for a one-off small intensive survey and becomes progressively more complicated as the survey area becomes larger and more varied, especially if many different surveys are to be amalgamated. However, the key stages are set out in Figure 19.1 and these are likely to be common to many broad scale surveys.

19.4.1 Preparation of acoustic data

The data are edited to ensure that they are of sufficiently high quality. The data should be subjected to data-cleaning procedures and the resulting data visually displayed to highlight any potential problem areas.

19.4.2 Preparation of ground truth data

The data are analysed to identify the habitats and biotopes present. If possible, these identifications should be checked by an independent expert. The records should be tagged by the predominant habitat

or biotope type (or higher level category such as life form or biotope complex). Additional attribute data may also be required, such as the proportions of subsidiary habitats or biotopes, species abundances or quantitative data.

19.4.3 Exploratory data analysis

This is a vital stage in the interpretation of the data and, if done thoroughly, will help to prevent possible classification errors in subsequent analyses.

1. The nature of the acoustic data need to be explored through the use of scattergrams and, in particular, variograms showing the spatial variation in the data. It may be necessary to break the survey area into different physiographic regions if it is likely that the spatial variation will vary substantially between regions. The variogram will determine safe interpolation distances and these should be shown as buffers around the track data to highlight areas where values may be confidently interpolated and areas where such estimates will be doubtful.
2. Track data could be classified by E1/E2 categories as a check on subsequent interpolation. The variables will need to be interpolated and the results displayed as single raster or contour images to highlight possible areas where there may be a problem with interpolation. The details of interpolation method and parameters used should be given.
3. It may be useful to further explore the data using unsupervised classification techniques to identify clusters in the acoustic data and non-spatial modelling and statistics to find correlations between the acoustic variables and the sample attribute data. These methods are appropriate for (a) selecting sites for further investigation on the strength of the acoustic data and (b) to find non-spatial explanations of acoustic/ground truth attribute relationships.
4. The ground truth data (and any available external data source) should be inspected for broad trends in distribution.

19.4.4 Supervised classification

The primary outputs from supervised classification are maximum likelihood classification of habitats and biotopes. However, the classification process also produces individual probability distribution maps for habitat or biotope classes. They are important and should be available if the survey requires this information.

Prior probability images are useful for incorporating a variety of information on constraints to the distribution of habitats and, in particular, biotopes and life forms. If these are used in supervised classification, then the basis for their construction should be explained. It is likely that their use will be needed for large and complex areas.

19.4.5 Accuracy assessment

1. **Internal accuracy measures** are extremely useful for indicating which habitat or biotope classes are likely to be confused. This usually occurs between closely related classes.
2. **External accuracy measures** are the best test of the predictive capability of the maps. This is often not possible within the constraints of a survey. However, if possible, the cost of extra sampling should be included in any proposed survey.
3. **Uncertainty maps** are useful in showing areas where there is no one habitat or biotope class with high probabilities of occurrence. This is likely to be due to heterogeneity of the ground where the area has been intensively surveyed, but could also mean poor signature development where the survey has been carried out at a low intensity.

19.4.6 Cartographic display

Maps need to be designed for the scale at which they are to be displayed. Very small pixel size can introduce a level of detail that is visually confusing if shown on small scale maps. On the other hand, large pixels give a map a tiled look which is unattractive even though this may be the representation which best matches the data and interpretation.

In general, small pixel sizes should be used to reduce the tile-effect, but filtered to remove isolated pixels when the map is to be shown at a small scale. If maps are to be shown at a large scale, then the tile effect can be reduced by quartering the pixels (or even further subdivision) followed by filtering. This smoothes tiles by removing sharp corners.

Paper plan maps (single theme maps) of the predominant biotopes will be the main output of broad scale survey. Three dimensional (digital elevation models) are also very useful methods of displaying habitat and biotope information. Single theme maps may also be required showing the distribution of specific attributes.

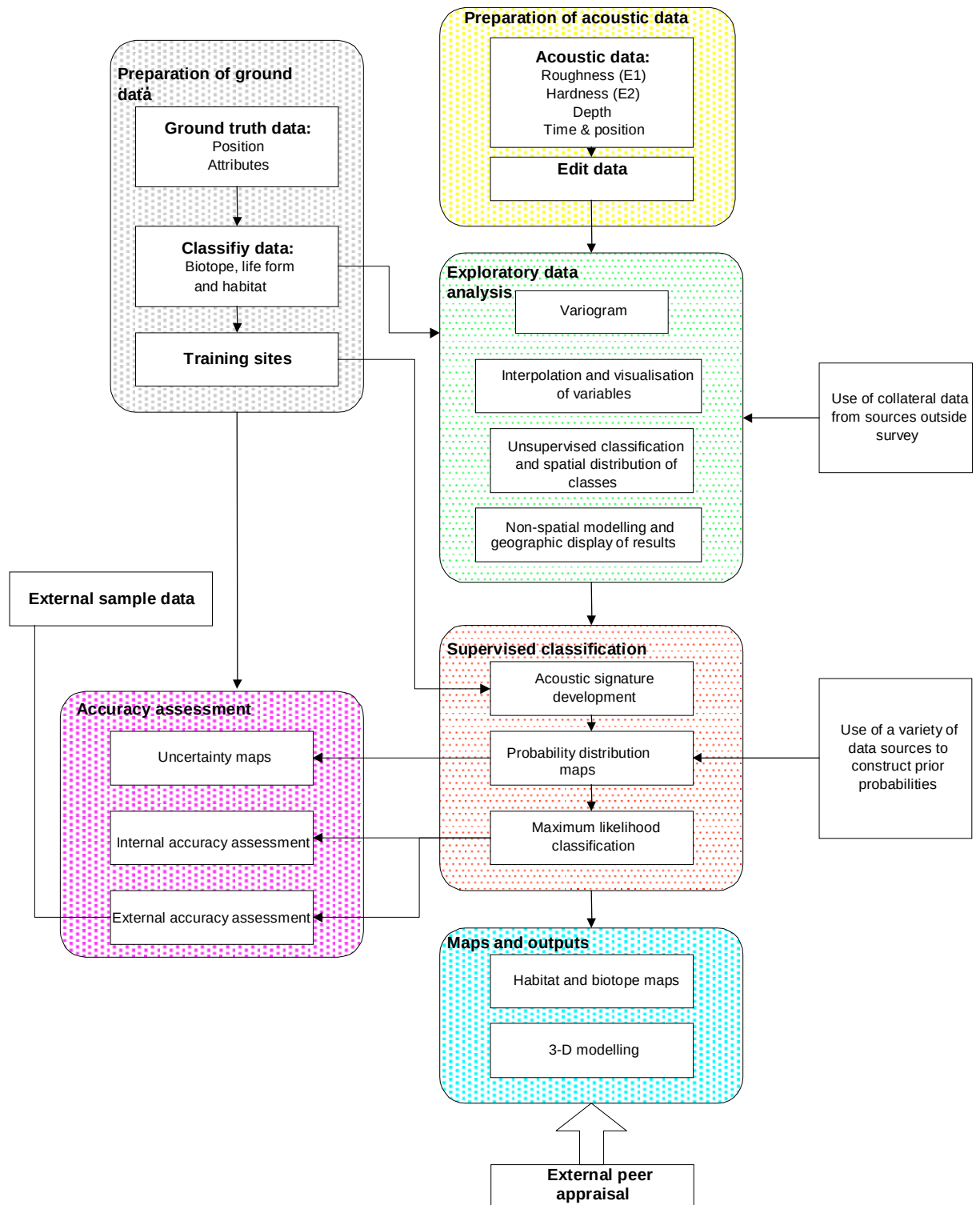
Electronic versions of some of the results from data analysis may also be required, although not every query can be anticipated. It is unlikely that queries can easily be answered from a re-analysis of the track and ground truth data by someone who is not familiar with the survey data and the analysis used. Therefore, it is important to anticipate the queries that are likely to be made of the data and to build in cost of the analysis required to furnish answers.

19.4.7 External peer appraisal

The best measure of success of the methodology is whether or not the maps prove to be useful. Do users have confidence in the maps and, if they are used in a predictive capacity, do users find that the maps stand up to scrutiny?

1. **An assessment of how well the map reflects expert local opinion:** However well the analyst regards the match between the map and the survey data, a map is unlikely to carry conviction if it conflicts with the view of local biological expert opinion. It is important, therefore, that maps are circulated for comment and any concerns expressed are explored by re-analysis of the data.
2. **A measure of the predictive ability of the map as tested by users:** Ultimately, there is only one way to test the usefulness of a map and that is by testing its predictive capability. Ideally, a user should plan further sampling based on the map to prove its accuracy.
3. **Ease of use:** Maps should be easy to read and use. This may entail a compromise between the level of detail available and simplicity. This is especially important for digital image processing where it is tempting to represent the interpretation of every pixel no matter how isolated it may be. Thus, filtering is usually required to simplify the image for ease of viewing. The user should know what level of filtering has been applied.

Figure 19.1. Key stages in analysis of the data and production of maps



20 Conclusions

The primary aim of the BROADSCALE Mapping Project was to develop a cost-effective methodology based on acoustic remote sensing that is capable of mapping habitats and the marine life of large areas of the seabed. The methodology has both a technical and an analytical component.

20.1 The broad scale survey and mapping methodology

20.1.1 Technology and field survey

The system was based largely on acoustic ground discrimination systems (AGDS) and sidescan sonar. AGDS are coarse resolution survey tools whilst sidescan sonar provides high resolution images. On the other hand, AGDS have powers of discrimination between sediment types that are superior to discrimination by sidescan sonar. The two technologies were seen as complementary and this, together with the constraints imposed by costs, determined the choice of equipment adopted for the development of the BMP methodology. However, other systems have been demonstrated to BMP or reviewed and are discussed in this report. It is hoped that the methodology adopted by the BMP is sufficiently generic to be able to absorb new or expensive technologies as they become available to the wider scientific community.

The technical limitations of AGDS have been stressed in this Technical Report because there is always the danger that potential users of the technology (or the maps derived from AGDS surveys) will have unrealistic expectations. Despite these limitations, AGDS have proved to be robust and useful tools for habitat and biotope discrimination. AGDS systems are also commonly installed on many vessels and satisfies the need for widespread use if the technology is to be applied to the survey of very large sea areas.

The deployment of video has been the technique of choice for rapid ground truthing and a system that can operate to 200m depth and deployed from small vessels was successfully used. The combined AGDS and video technology is relatively inexpensive and transportable. Other sampling methods (grabs, trawls and dredges) have also been used besides video, although it is often difficult to find a common descriptive framework for disparate sampling techniques. Sidescan sonar can be thought of as a sampling tool where the images have been used to construct training sites for the classification of AGDS data, albeit producing data from a far greater area of the seabed than other sampling techniques.

There are a number of different strategies for using the technologies that can be applied to a range of aims and objectives (Section 8). Thus, the methodology is readily adaptable to different application without many changes being required in technology and equipment. In particular, the methodology can be used at many different scales. This versatility, combined with the transfer of experience from one type of survey to another, is one of the main strengths of the methodology.

20.1.2 Analysis and map preparation

20.1.2.1 Classification

The BMP methodology relies on maximum likelihood classifiers available through image processing packages as the main route for the interpretation of the AGDS data. Other approaches can be taken, such as the use cluster analysis and parallelipiped classification (Murphy *et al.* 1995; Greenstreet *et al.* 1997) or regression models to establish links between AGDS data and habitats or biota (Pinn and Robertson 1998). However, the power of likelihood classifiers, the versatility of the probability distribution maps and the way they can be easily linked to other cartographic modules in image processing packages are advantages that make this the approach of choice for the BMP methodology.

It is easy to view resource maps and take them at face value, making the assumption that they are a faithful representation of reality. However, resource maps based on remote sensing are predictive, a

fundamental point which has been stressed in the Technical Report. The BMP has made use of various ways of assessing accuracy, confidence and levels of uncertainty that may be invested in biotope resource maps. These include variogram models of spatial variability, internal and external accuracy assessment through the use of error matrices and maps showing uncertainty.

The accuracy and confidence of the maps will vary depending upon the intensity of survey and the predictive power of a resource map should be tested against an external data source. However, even where survey intensity is high, external accuracy assessment always shows a moderately high level of uncertainty and variability in the classification. Confusion between two or more habitat or biotope categories arises because (a) in reality they occur as patches in a heterogeneous environment (or as intermediates between two or more categories), and (b) the remote sensing system and classification process cannot be precise about the discrimination between classes. Various ways to mitigate the effects of poor categorisation of the ground truth data on the interpretation of the acoustic data have been suggested in the Technical Report (see Section 15.5.3).

Part of the problem with poor matches between different interpretations is that the maximum likelihood classifier assigns pixels to the most probable class and this hides the real level uncertainty where two or more habitat or biotope classes have similar probabilities. An alternative to portraying the resources as 'hard' mutually exclusive categories is to use the probability distribution maps that underlie maximum likelihood classification. These have the advantage that shades of certainty can be represented. The disadvantage is that the information is much harder to represent since a single theme (maximum likelihood) is substituted by separate themes for each habitat or biotope category.

20.1.2.2 Cartographic modelling: Incorporating external knowledge

The basic process of image processing links the remotely sensed images to ground truth points through the classification process in an automated, systematic way. However, it is likely that there will be many other sources of information about the survey area, including personal experience and knowledge of local experts. There will be inevitable conflicts between this information and the habitat or biotope maps. This can be very frustrating for the analyst and may induce others to dismiss the maps as being inaccurate. The BMP has explored many ways in which some of this knowledge can be used in the production of maps.

Three main avenues have been used. Firstly, broad scale trends in the distribution of habitats or biotopes derived from the ground truth data or other sources of sample data can be incorporated into prior probability maps which are then included in the classification process. Secondly, simple models limiting the extent of biotopes (e.g., with respect to depth) can be used to restrict their distribution predicted from the AGDS data. These can be incorporated either as prior probability images or through overlaying the restrictions (as a geographic boundary) onto the resource maps and reclassifying the intersections. Thirdly, collateral data from other sources, such as bathymetric data, can be used to give better estimates of interpolated acoustic values.

20.2 Future developments of broad scale survey and mapping

20.2.1 Mapping wider sea areas

There is no doubt that broad scale habitat and biotope maps are a useful resource in their own right for supporting management, as can be judged by their use to date in areas of high conservation interest. For example, maps produced by *SeaMap* either during the BioMar program or the BMP have assisted English Nature in establishing or justifying the pSAC boundaries off Flamborough and the Isle of Wight as well as assisting with responses in the consultation process of various licences for the use of seabed resources. The mapping in the Firth of Lorn has also contributed to the definition of reef boundaries in this pSAC. The requirement for information (in terms of location) is difficult to anticipate, especially if issues arise away from areas of recognised conservation or economic interest. Information on the distribution of habitats and biota is lacking for much of the UK coastline making it extremely difficult to make an informed response based on local knowledge. Clearly, a comprehensive knowledge of the UK waters, even at a basic level, would be valuable for day-to-day management of

the seabed as well as for the planning of longer term strategies for coastal zone management. If these broad scale surveys are extended over very large sections of the U.K. coastline, this new knowledge of the distribution of biotopes will help in the management of the wider seas.

The BMP has demonstrated that it is possible to amalgamate data sets from many surveys that have used different AGDS. This is crucial to the further development of the main aims of the project, namely to survey large areas of the continental shelf of U.K. waters. If this were to be done, then it is inevitable that many different surveyors would be involved.

There are many problems to be overcome when combining the results from different surveys, even if they used similar techniques. Simply joining the interpretations of habitats or biotopes from one survey with those of another is possible, but this is likely to result in obvious mis-matches at the boundaries between surveys. An alternative process devised by BMP amalgamated tracks from different surveys that have been given probability values for each habitat or biotope class. These combined images are then interpolated to create a unified probability distribution map for each class which can then be subjected to maximum likelihood classification. This procedure was used successfully in Pembroke and the Firth of Lorn. The use of probabilities in this way has the potential for amalgamation of surveys using different technologies as long as predicted habitat or biotope distribution can be expressed as probabilities. This commonality between different surveys will be important if mapping is to keep pace with advances in technology without rendering previous mapping surveys obsolete.

20.2.2 Maps in data management systems

Biotope resource maps are just one layer of information of an electronic integrated data management system. They are intended to form part of an integrated data management system, based on GIS, that can be used to respond to requests for information. A single layer showing the most likely biotope distribution is easiest to view, but probability images for each biotope are even more versatile and can readily be incorporated into risk assessment models. In other words, whilst there will certainly be a need to present a clear and simple view of the natural resources of the seabed, more complex information can be held and manipulated within a GIS as part of a sophisticated management information/modelling system. Not only will such systems be able to retrieve relevant information in response to queries, but create predictions for 'what-if' queries based on modelling. Habitat and biotope distribution maps, especially those showing probabilities and uncertainties, provide the basic map-based information for management information/modelling support systems.

20.2.3 Maps in the study of dynamic processes

Biotope maps are essentially a snap-shot in time. How will distribution change over time? How can our knowledge of the dynamics of habitats and biotopes be incorporated into dynamic distribution models? Again, probability maps could form the basis of dynamic models by providing predictions on changing patterns of biotope distribution at different scales that can be tested through a highly targeted, stratified sampling program.

20.2.4 Maps as part of a wider integrated survey strategy

The lessons learnt from the use of mapping in conjunction with more traditional grab sample surveys could have enormous implications for the future design of monitoring surveys. Knowledge of the distribution of biota could structure future surveys leading to a cost-effective strategy that is based on refining our knowledge of the processes that determine biological community structure. The development of a more efficient use of data is based on an iterative process of (a) making predictions based on current knowledge (the generation of testable hypotheses) and (b) the design of surveys and/or statistical analyses to test the predictions. In a site-specific capacity this process will refine the knowledge of the distribution of biota and at a more general level increases the knowledge base of biological community structure that can be applied to other sites, times and situations (see Figure 20.1).

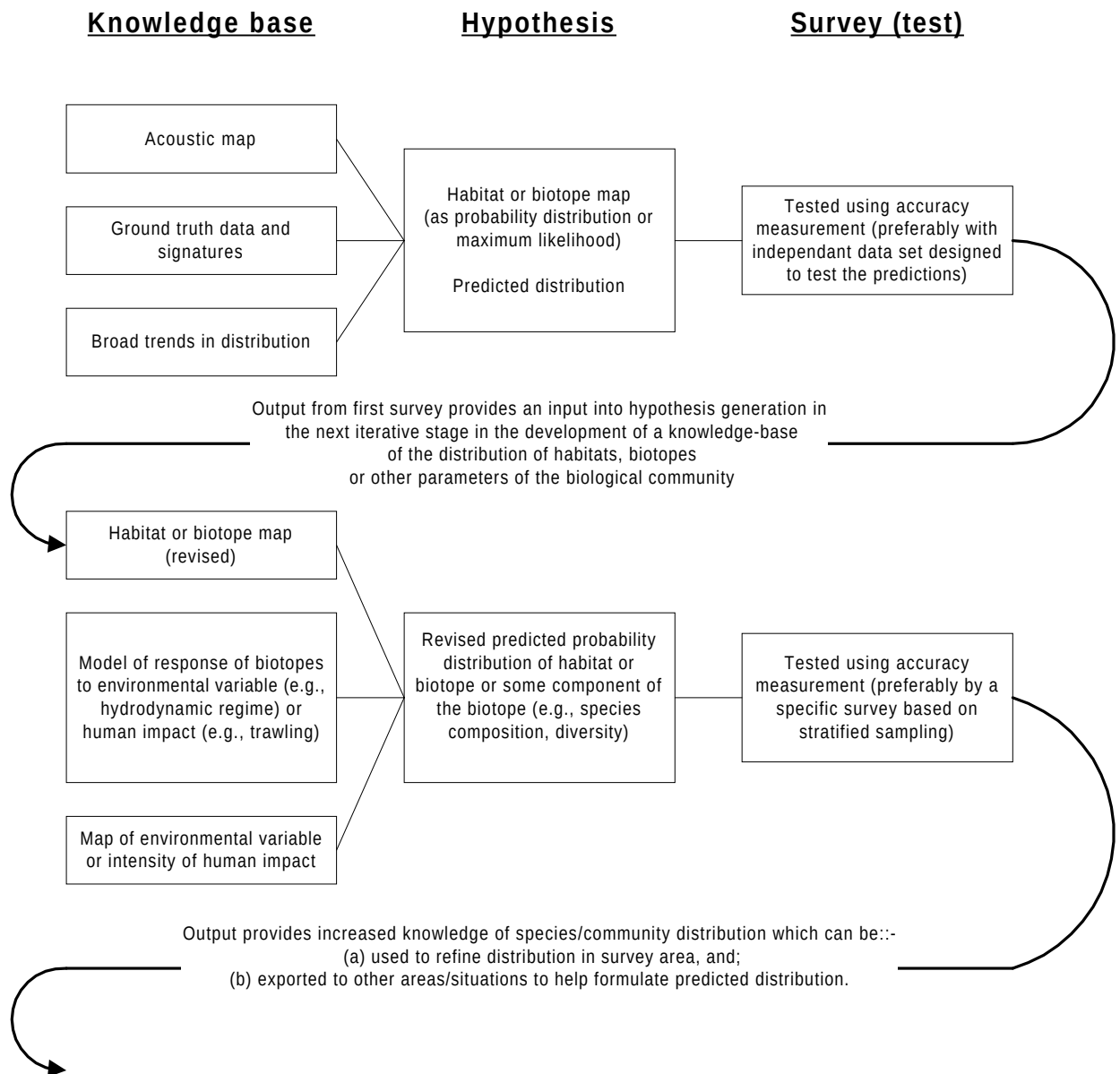


Figure 20.1. A possible scheme of iterative prediction/survey to build up a knowledge base of the distribution of biota.

The implementation of a scheme as outlined in Figure 20.1 will require the gathering together of appropriate survey methods, descriptive units, analytical techniques (including GIS operations, spatial statistics, geostatistics and other statistical approaches, such as the Bayesian techniques used in image processing) and presentation tools (three dimensional display and animation) to develop an integrated strategy for making the best use of the knowledge that is available.

Biotope maps will undoubtedly fulfil a central role in the support of scientifically-based decision making in the marine environment. The next step is to provide this spatial information for larger sea areas and in a format that can be used in the new generation of data management and risk assessment systems that are being developed for integrated marine management.

21 Glossary

Many of the specialist terms used in the text are printed in **bold** and are explained below. Many of the terms are used in the particular context of broad scale biotope mapping using acoustic remote sensing and have been adapted from definitions given in texts on remote sensing and geographical information systems. The reader is referred to such texts as Eastman (1997) and Wilkie & Finn (1996) for image processing, and DeMers (1997) for geographic information systems.

Acetate overlay	A method of overlaying information for a given area held as separate thematic maps by reproducing each map (using a common scale) on a clear sheet of acetate and then superimposing one on another.
Acoustic ground discrimination system (AGDS)	A system based on an echo sounder that measures parameters of the return signal (echo) for the purpose of characterising the acoustic reflectance properties of the sea floor.
Acoustic ground types	Categories of sea floor based on acoustic properties alone. See also unsupervised classification.
Acoustic track	Data composed of a string of single datum points that is acquired as the vessel steams over the survey area. Each datum contains acoustic information as well as time and location.
Aggregation	Mixture of two or more biotopes, each at or above the minimum biotope size, which cannot be separately distinguished acoustically using the track spacing adopted for a survey.
Association	A group of biotopes that are often found together and are
Atlas	A collection of thematic maps which can be viewed separately, rather than superimposed.
Bayesean statistics	An extension of classical probability theory in which existing knowledge can be combined with new evidence to arrive at a modified estimate as to whether a particular hypothesis is true (see also prior and posterior probabilities).
Bottom up	The process of piecing together detailed information to produce a broad picture.
Buffering	The process of creating areas of a given distance from a point or line.
Cartographic modelling	The process of transforming and combining information in GIS/image processing to create a map. The process can be quite complex and is often represented graphically by a flow diagram indicating the decisions that have been taken regarding analysis and alternative routes that may be taken, involving further decisions to be made as the process is followed.
Centroid	The geometric centre of a polygon that can be used to represent the attributes of that polygon as a single georeferenced point
Circalittoral	Zone at depths below that dominated by algae (see 'infralittoral').
Continuous coverage	A thematic map whose values can be read from all possible locations (also called a continuous surface)

Contouring	Process of connecting points of the same value on a map coverage of a continuous variable (e.g., bathymetry).
Digital image	An image composed of a regular grid of rows and columns of discrete, quantized digits or units.
Discrimination	The ability to distinguish between two or more biotopes or habitats on the basis of their acoustic signatures
Ensonify	Akin to 'illuminating' an area of the seabed, but with sound.
Errors of commission	Habitats or biotopes falsely represented at a particular location on a map when they are, in reality, not present.
Errors of omission	Habitats or biotopes omitted from a location on a map when they are, in reality, present.
Exclusivity	The property of a biotope having an acoustic signature that is substantially unique (not shared with other biotopes)
External accuracy measure	A measure of the accuracy of a map by comparing the predicted habitat or biotope with an external georeferenced sample data set.
Extrapolation	The process of predicting missing values or biotope classes some distance away from existing ground truth points on the basis of measured acoustic values and acoustic signatures.
Fidelity	The property of a biotope that is always found in association with a defined acoustic signature.
Footprint	The area and shape of the seabed which is ensonified by an echo sounder.
Georeferenced	Point, line or area represented in a database by geographical coordinates.
Geostatistics	Application of statistics to spatial patterns
Global	Associations between biotopes, habitats and acoustic properties that hold true over very large geographic areas and can be used to predict biotope or habitat occurrence at a local level.
Gridding	The calculation of values for nodes of a grid based on irregularly spaced data covered by that grid. This is needed for converting track data to raster images and deriving contours.
Ground truthing or ground validation	The process of collecting information about habitat and biotope features to aid the interpretation of remotely sensed data. Existing data, not necessarily collected for ground truthing, can be used for this purpose.
Hard (crisp) boundary	A clear boundary between two entities without any ambiguity at the transition.
Image processing	All the various processes that can be applied to remotely sensed data
Images	A recorded representation of an area produce by remote sensing techniques.
Infralittoral	Zone permanently submerged (i.e., below extreme low water spring

	tides) but at depths in which algal growth dominates rock surfaces. Lies above the ciralittoral zone (see definition).
Integrated survey strategy	A strategy in which broadly-based, general information (from published information and/or broad scale survey) provides the context for setting the aims and guiding the choice of appropriate techniques for more detailed survey. Often referred to colloquially as 'top-down' survey.
Internal accuracy	A measure of the consistency with which an interpretation of the acoustic data matches the ground truth data set used for signature development.
Interpolation	The process of predicting unknown values using known values at multiple locations around the unknown value.
Intersects	An overlay operation carried out on two vector maps where polygons are added or subtracted from each other (using Boolean algebra) to derive a third thematic map.
Isopleths	Map composed of lines connecting points of equal value (contours).
Iterative	A process of repeating part of the survey strategy with the benefit of the output from the previous cycle.
Kriging	An interpolation routine that depends upon the pattern of spatial dependence between points to calculate the optimal distance-weighting for the estimation of values.
Life form	Life forms are the conspicuous, morphological adaptations of organisms to their environment, particularly adaptations to trophic requirements, maintenance of their position in the physical milieu in which they live and avoidance of predation.
Likelihood	The probability of a habitat or biotope being present at a particular location ranging from 0 (certain absence) to 1 (certain presence).
Local	Associations between biotopes, habitats and acoustic properties that hold true only for the immediate locality of the locations from which the data were taken and cannot confidently be used to predict biotopes or habitats some distance away.
Local mean	The mean value of a variable as calculated from all the data within a given sector of the survey area without taking any spatial pattern into account. This is, in effect, what occurs in interpolation when the position of an estimated value lies beyond the range (see definition) of spatial dependence in the data.
Maximum likelihood	A statistical decision rule used in supervised classification to assign pixels to the class of highest probability according to acoustic signatures.
Mixed biotope	Area approximately equivalent to the dimensions of a biotope that is composed of elements from two or more biotope types is termed a mixture.
Moving window	A filtering technique involving the mathematical/statistical comparison of a pixel with those surrounding it. The filter moves through the whole digital image calculating a new filter-value for each pixel in turn.

Multispectral data	Data from a number of spectral bands collected at the same time. Acoustic values for hardness, roughness, variability, slope and depth can be regarded as spectral for the purposes of image processing acoustic data.
Nearest neighbour interpolation	An estimation of the value of a variable or category at a location made by taking the value of the nearest measured value.
Nesting	The process of placing fine scale, detailed data into context with broad scale, general data.
Overlay	The operation of comparing the values and their spatial distribution amongst 2 or more coverages by superimposition.
Parelleliped	Acoustic signature defined by maximum and minimum values of the various acoustic parameters. In the simplest cases, these equate to the <i>RoxAnn</i> boxes.
Patchy	Distribution of biotopes where two or more biotopes, each at or above the minimum biotope size, cannot be separately distinguished acoustically using the track spacing adopted for a survey.
Pixel	‘Picture element’. A single datum point in a digital image that is georeferenced and has data on spectral characteristics for the rectangular area of ground it represents. Acoustic values can be regarded as spectral for the purposes of image processing acoustic data.
Polygons (vector)	Irregularly shaped region with boundaries of line segments joined by georeferenced nodes that enclose an area with the same attribute or value throughout.
Posterior probability	An estimate as to whether a hypothesis is true based on a combination of a prior probability estimate and new evidence.
Predictive	Strategy for survey based on testing the predicted occurrence of a biotope or habitat. See also ‘integrated survey strategy’.
Prior probability	Knowledge as to whether a hypothesis (e.g., that a biotope is found at location ‘X’) is true before the hypothesis is tested against some new evidence. This is sometimes referred to as a ‘belief’ but is more likely to be a probability value between 0-1 based on some existing data.
Probability maps	Maps showing the likely distribution of a habitat or biotope in terms of probability ranging from zero (certainly absent) to one (certainly present).
Range	The distance over which there is spatial correlation between data.
Raster image	Continuous coverage composed of pixels.
Reclassification	Assigning new values to defined ranges of values in a data set.
Regions	See polygons
Remote sensing	The collection of information about an area using a sensor that is not in direct or intimate contact with the objects in the area.
Resolution	Spatial detail possible with any particular sensor.

Set of conditions	Conditions from a wide range of habitat and biological parameters that can be used to predict the occurrence of a biotope.
Signature	Parameters from a range of acoustic variables that can be used to predict the occurrence of a biotope.
Sill	The region of a variogram where variance reaches a maximum and then levels off. The range where the sill is reached indicates the maximum distance over which some spatial correlation between points that depends on the distance separating them can be assumed to be operating.
Sliver	When two polygons that are supposed to share a common boundary are overlain there is often a small region of discrepancy where they either overlap or do not meet. This results in small polygons being created that need to be removed.
Smart raster	Continuous coverage of pixels with multiple attributes contained in a linked table that can be queried and the results of that query displayed.
Stratified sampling	Sampling strategy where locations for ground truthing are selected on the basis of known acoustic ground type.
Supervised classification	The process of generating a thematic map from a digital image where pixels are assigned to classes defined by an analysis of the acoustic spectra contained within selected training sites representing the different classes.
Surface	A continuous coverage of a variable that has been produced by mathematical modelling, such as interpolation.
Thematic map	A map designed to show one variable or feature.
Top down	The process of regarding detail in a broader context, used to help structure detailed survey.
Training site	A sample of the acoustic ground type known to represent a given biotope or habitat class.
Unsupervised classification	The process of generating a thematic map from a digital image by first using statistical clustering techniques to group pixels into classes according to their acoustic similarity. The process can then continue by using the surveyor's knowledge of the area to label these areas by biotope or habitat type.
Variogram	A graph showing how the variance in the values between points varies with the distance between them (the 'lag').

22 References

- Anon., 1991. *The Admiralty simplified harmonic method of tidal prediction*. Taunton, UK., Hydrographer of the Navy.
- Bailey, T. C., and Gatrell, A. C., 1995. *Interactive Spatial Data Analysis*. Harlow, U.K., Longman.
- Blondel, P., and Murton, B. J., 1997. *Handbook of Seafloor Sonar Imagery*. Chichester, U.K., Wiley-Praxis.
- Buchanan, J. B., 1984. Sediment analysis. In *Methods for study of marine benthos*, (eds N. A. Holme and A. D. McIntyre). Oxford, Blackwell Scientific Publications.
- Bunker, F. S. D., and Foster-Smith, R. L., 1996. *Field Guide for Phase 1 Seashore Mapping*. Peterborough, UK., English Nature.
- Burns, D., Queen, C. B., and Chivers, R. C., 1985. An ultrasonic signal processor for use in underwater acoustics. *Ultrasonics*, **23**, 189-191.
- Burroughs, P. A., and McDonnell, R. A., 1998. *Principles of Geographical Information Systems*. Oxford, Oxford University Press.
- Cameron, T. D. J., Crosby, A., Balson, P. S., Jeffery, D. H., Lott, G. K., Bulat, J., and Harrison, D. J., 1992. *Geology of the southern North Sea*. British Geological Survey. London, HMSO.
- Chivers, R. C., Emerson, N., and Burns, D. R., 1990. New acoustic processing for underway surveying. *The Hydrographic Journal*, **56**, 9-17.
- Collins, M. B., and Voulgaris, G., 1993. Empirical field and laboratory evaluation of a real-time acoustic sea bed surveying system. *Proceedings of the Institute of Acoustics*, **26**(343-352).
- Collins, W., Gregory, R., and Anderson, J., 1996. A digital approach to seabed classification. *Sea Technology*, **37**(8), 83-87.
- Congleton, R. G., 1991. A review of assessing the accuracy of remotely sensed data. *Remote Sensing of the Environment*, **37**, 35-46.
- Connor, D. W., Dalkin, M. J., Hill, T. O., Holt, R. H. F., and Sanderson, W. G., 1997. *Marine nature conservation review: marine biotope classification for Britain and Ireland. Volume 2. Sublittoral biotopes. Version 97.06*. JNCC Report, No. 230. Peterborough, UK., Joint Nature Conservation Committee.
- Cracknell, A. P., and Hayes, L. W. B., 1993. *Introduction to Remote Sensing*. London, Taylor and Francis.
- Cressie, N. A. C., 1990. The origins of Kriging. *Mathematical Geology*, **22**, 239-252.
- Curran, T., 1995. Remote Sensing Technology for Subtidal Classification. *Kitasoo First Nations Fisheries Program Resource Inventory Committee of British Columbia*, Victoria, Canada.
- de Moustier, C., 1988. State of the art swath bathymetry survey systems. *International Hydrographic Review of Monaco*, **65**, 25-54.
- DeMers, M. N., 1997. *Fundamentals of Geographic Information Systems*. New York, John Wiley & Sons.

- Devillers, P., Devillers-Terschuren, J., and Ledant, J.-P., 1991. *CORINE Biotopes Manual: Habitats of the European Community*. Luxemburg. Commission of the European Communities, Directorate General of Environment, Nuclear Safety and Civil Protection.
- Doody, P., 1999. Integrated coastal management and information - The European perspective. *Integrating Information and Management on the Coastal Zone*. University of London, R. C. Earll and Coastnet.
- Downie, D. W., Donnan, D. W., and Davison, A. J., 1999. A review of Scottish Natural Heritage's work in subtidal marine biotope mapping using remote sensing. *International Journal of Remote Sensing*, **20**(3), 585-592.
- Eastman, R. J., 1997. *Idrisi for Windows User's Guide*. Worcester, MA, USA, Clark University.
- Foster-Smith, R. L., 1995. *SEASEARCH starter pack*. Edinburgh, Scottish Natural Heritage.
- Foster-Smith, R. L., Davies, J., and Sotheran, I., 1997. *BioMar Final Report*. Newcastle upon Tyne, Newcastle University.
- Foster-Smith, R. L., and Sotheran, I. S., 1999. *Broad scale remote survey and mapping of sublittoral habitats and biota of the Wash and the Lincolnshire and north Norfolk coasts*. Report of the Broadscale Mapping Project. Peterborough, English Nature.
- Foster-Smith, R. L., Walton, R., Strong, E., Davies, J., and Sotheran, I., 1999. *Trialing of acoustic ground discrimination systems (AGDS) and video sledge monitoring techniques in Loch Maddy*. (Contractors: SeaMap, Newcastle University). Edinburgh, Scottish Natural Heritage.
- Gardener, J. V., Butman, P. B., Mayer, L. A., and Clarke, J. H., 1998. Mapping U.S. continental shelves. *Sea Technology*, **39**(6), 10-17.
- Gilbert, C., 1999. How to pick a GPS receiver. *Mapping Awareness*, **13**(1), 22-24.
- Green, E. P., Mumby, P. J., Edwards, A. J., and Clark, C. D., 1996. A review of remote sensing for the assessment and management of tropical coastal resources. *Coastal Management*, **24**, 1-40.
- Greenstreet, P. R., Tuck, I. D., Grewar, G. N., Armstrong, E., Reid, D. G., and Wright, P. J., 1997. An assesment of the acoustic survey technique, RoxAnn, as a means of sampling seabed habitat. *ICES Journal of Marine Science*, **54**, 939-959.
- Hannan, M., 1999. Environmental Risk Assessment (ERA) in the coastal and offshore environment - Computer aided decision support. *Integrating Information and Management on the Coastal Zone*. University of London, R. C. Earll and Coastnet.
- Hogarth, A. N., 1999. SimCoast: Coupling expert knowledge with decision making in the coastal zone. *Integrating Information and Management on the Coastal Zone*. University of London, R. C. Earll and Coastnet.
- Hull, J., and Nunny, R., 1999. Mapping intertidal sediment distributions using the RoxAnn system, Dornoch Firth, NE Scotland .In *Sedimentary Processes in the Intertidal Zone*, (eds K. S. Black, D. M. Paterson, and A. Cramp), Geological Society, London, pp 273-282.
- Kaiser, M. J., Armstrong, P. J., Dare, P. J., and Flatt, R. P., 1998. Benthic communities associated with a heavily fished scallop ground in the English Channel. *Journal of the Marine Biological Association of the United Kingdom*, **78**, 1045-1059.

- Kaiser, M. J., and Spencer, B. E., 1996. The effects of beam trawl disturbance on infaunal communities in different habitats after trawling disturbance. *ICES Journal of Marine Science*, **65**, 346-358.
- Kavli, T., Carlin, M., and Madsen, R., 1993. Seabed classification using artificial neural networks and other nonparametric methods. *Proceedings of the Institute of Acoustics*, **15**, 144-148.
- Kavli, T., Weyer, E., and Carlin, M., 1994. Real time seabed classification using multi frequency echo sounders. *Proceedings of Oceanology International 94*, **4**, 1-9.
- Keckler, D., 1994. *Surfer for Windows*. Golden, Colorado, USA. Golden Software, Inc.
- Kraak, M. J., and Ormeling, F. J., 1996. *Cartography: Visualization of Spatial Data*. Harlow, U.K., Longman.
- Lawrence, G. R. P., 1971. *Cartographic Methods*. London , Methuen & Co Ltd.
- Magorrian, B. H., Service, M., and Clarke, W., 1995. An acoustic bottom classification survey of Strangford Loch, Northern Ireland. *Journal of the Marine Biological Association of the United Kingdom*, **75**, 987-992.
- Manly, B. F. J., 1997. *Randomization, Bootstrap and Monte Carlo Methods in Biology*. London, Chapman and Hall.
- Mitson, R. B., 1983. *Fisheries Sonar*. Farnham, U.K., Fishing News Books Ltd.
- Muirhead, K., and Cracknell, A. P., 1986. Review article: Airborne lidar bathymetry. *International Journal of remote Sensing*, **7**, 597.
- Murphy, L., Leary, T., and Williamson, A., 1995. Standardising seabed classification techniques. *Sea Technology*, **36**(7), 15-19.
- Nature Conservancy Council, 1990. *Handbook for Phase 1 habitat survey - a technique for environmental audit*. Peterborough, U.K., Nature Conservancy Council.
- Pinn, E.H., Robertson, M.R., Shand, C.W. and Armstrong, F., 1998. Broadscale benthic community analysis in the Greater Minch Area (Scottish west coast) using remote and non-destructive techniques. *International Journal of Remote Sensing*, **19**, 3039-3054.
- Riddy, P., and Masson, D. G., 1996. The Sea Floor - Exploring a Hidden World .In *Oceanography: An Illustrated Guide*, (eds C. P. Summerhayes and S. A. Thorpe). London, Manson Publishing Ltd, pp 300-313.
- Rukavina, N. A., and Caddell, S., 1997. Applications of an acoustic sea-bed classification system to freshwater environmental research and remediation in Canada. *Proceedings of the Fourth International Conference on Remote Sensing for Marine and Coastal Environments*, **1**, 317-329.
- Safer, A., 1997. Canada's multibeam platform: Advantages and applications. *Sea Technology*, **38**(3), 17-25.
- Schafer, W., 1972. *Ecology and Palaeontology of Marine Environments*. (Translator Oertel, I.) Edinburgh, Oliver and Boyd.
- Scherzinger, B., Woolven, S., and Field, M., 1997. Seafloor Mapping Solution. *Sea Technology*, **38**(3), 63-69.

- Schlagintweit, G. E. O., 1993. *Real-time acoustic bottom classification for hydrography: a field evaluation of RoxAnn*. Canadian Hydrographic Service and Department of Fisheries and Oceans.
- Sinclair, M., 1998. Australians get on board with new LADS. *Sea Technology*, **39**(6), 19-25.
- Sotheran, I. S., Foster-Smith, R. L., and Davies, J., 1997. Mapping of marine benthic habitats using image processing using raster based geographic information systems. *Estuarine and Coastal Shelf Science*, **44**(Supplement A), 25-31.
- Urick, R. J., 1983. *Principles of Underwater Sound*. New York, McGraw-Hill.
- Wilkie, D. S., and Finn, J. T., 1996. *Remote sensing imagery for natural resources monitoring: A guide for first time users*. New York, USA, Columbia University Press.

Appendix 1. Life forms and colours used in the biotope maps with equivalent MNCR biotope complexes and MNCR colours

Table 23.1. Summary of life forms and mapping colours. Note that littoral life forms have been included for completeness.

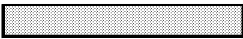
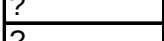
Life form & colour code	Red/green/blue
KELP Khaki	 155/ 155/ 50
FUCOIDS Crimson	 255/ 0/ 0
ALGAL TURF Dark purple	 128/ 0/ 128
LICHENS & ALGAE Grey/green	 80/ 120/ 80
MUSSEL & BARNACLE Light blue	 210/255/255
FAUNAL & ALGAL CRUSTS Light red	 254/220/190
SHORT FAUNAL TURF, CRUSTS & CUSHIONS Dark blue	 0/ 0/ 128
FAUNAL TURF Mid blue	 0/ 200/ 250
BIOGENIC SAND REEFS Orange	 255/ 210/ 0
BIOGENIC CALCAREOUS REEFS Black	 0/ 0/ 0
BRITTLE STAR BEDS Mid green	 80/200/80
MUSSEL BEDS Dark grey/blue	 35/ 120/ 140
OYSTER BEDS Orange/grey	 200/ 190/ 160
MAERL BEDS Pink/purple	 250/ 110/210
SEAGRASS BEDS Bright green	 0/ 255/ 0
SALTMARSH Pink	 255/ 200/ 250
SHINGLE, SHELL, COARSE SAND & MIXED SEDIMENTS Grey/brown	 135/122/105
SAND Yellow	 255/ 255/ 0
MUDDY (SILTY) SAND Green brown	 100/ 50/ 0
MUD Dark brown	 85/ 0/ 0

Table 23.2. Life forms and colour codes for littoral and sublittoral biotopes with associated MNCR biotope complexes and habitat classes (Connor *et al.*, 1997. *Marine Biotope Classification for Britain and Ireland*. Version 97.06). Biotope complex codes marked * include biotope code where required. See footnote for habitat complex codes. The suggested red/green/blue composition of the life form colours are given by the values in brackets beneath the colour boxes. See Connor *et al.* for explanation of MNCR mapping

Life form & colour code	MNCR Biotope complex	MNCR Biotope complex code	MNCR Habitat complex	MNCR Mapping colours
KELP Khaki  155/155/50	Kelp with cushion fauna, foliose red seaweeds or coralline crusts	KFaR	EIR	
	Kelp & red seaweeds	KR	MIR	
	Grazed kelp with algal crusts	GzK	MIR	
	Sand or disturbed kelp & seaweed	SedK	MIR	
	Silted kelp	K	SIR	
FUCOIDS Crimson  255/0/0	Robust fucoids or red seaweeds	FR	ELR	
	Dense fucoids	F	SLR	
	Fucoids, barnacles or ephemeral seaweeds	FX	SLR	
	Barnacles & fucoids	BF	MLR	
	Rock pools	Rkp	LR(other)	
	Overhangs & caves	Ov	LR(other)	
	Submerged fucoids, green & red seaweeds (lagoonal)	Lag	SIR	
ALGAL TURF Dark purple  128/0/128	Red seaweeds	R	MLR	
	Ephemeral green & red seaweeds	Eph	MLR	
	(also including biotope 'Ephemeral green & red seaweeds on variable salinity or disturbed mixed substrata')	FX EphX*	SLR	
	Algal turfs; biotopes from kelp biotope complexes:-			
	Foliose red seaweeds	KFaR FoR*	EIR	
	Red seaweeds on sand-covered rock	SedK PolAhn*	MIR	
	Rock pools (in part)	Rkp	LR(other)	
LICHENS & ALGAE Grey/green  80/120/80	Lichens or algal crusts	L	LR	
	Overhangs & caves (in part)	Ov	LR(other)	

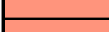
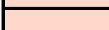
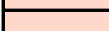
Life form & colour code	MNCR Biotope complex	MNCR Biotope complex code	MNCR Habitat complex	MNCR Mapping colours
MUSSEL & BARNACLE	Mytilus & barnacles	MB	ELR	
Light blue  210/255/255	Mytilus & fucoids	MF	MLR	
FAUNAL & ALGAL CRUSTS				
Light red  255/95/80	Robust faunal cushions & crusts (in part: planar areas))	SG CC.BalPom*	EIR	
	Grazed fauna	GzFa	MCR	
SHORT FAUNAL TURF, CRUSTS & CUSHIONS				
Dark blue  0/0/128	Robust faunal cushions & crusts (in part: vertical walls)	SG	EIR	
	Faunal crusts & short turfs	EFa	ECR	
	Fauna & seaweed (shallow vertical rock)	FaSwV	IR(other)	
FAUNAL TURF				
Mid blue  0/200/250	Alcyonium dominated communities	Alc	ECR	
	Barnacle & cushion sponge communities	BS	ECR	
	Mixed faunal turfs	XFa	MCR	
	Bryozoan/hydroid turfs	ByH	MCR	
	Ascidian communities	As	MCR	
	Brachiopod & solitary ascidian communities	BrAs	SCR	
	Faunal turfs (deep vertical rock)	FaV	CR(other)	
	Caves & overhangs (deep)	Cv	CR (other)	
BIOGENIC SAND REEFS				
Orange  255/210/0	Sabellaria alveolata reefs		MIR	
	Cirralittoral Sabellaria reefs	CSab	MCR	
	Biotope Sabellaria spinulosa & Polydora on mixed sediments	n/a SspiMX*	CMX	

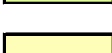
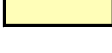
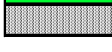
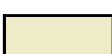
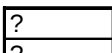
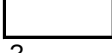
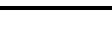
Life form & colour code	MNCR Biotope complex	MNCR Biotope complex code	MNCR Habitat complex	MNCR Mapping colours
BIOGENIC CALCAREOUS REEFS Black 0/0/0	Serpulid reefs	?	?	
BRITTLE STAR BEDS Mid green 80/200/80	Brittle star beds	Bri	MCR	
MUSSEL BEDS Dark grey/blue 35/120/140	Mytilus (mussel) beds on mixed substrata Mussel beds (sublittoral open coast) Sheltered Modilous(horse mussel) beds	MX M Mod	SLR MCR SCR	
OYSTER BEDS Orange/grey 200/190/160	Oyster beds	Oy	IMX	
MAERL BEDS Pink/purple 250/110/210	Maerl beds (open coasts/clean sediments) Maerl beds (muddy mixed sediments)	Mrl MrlMx	IGS IMX	
SEAGRASS BEDS Bright green 0/255/0	Zostera noltii Seagrass beds	Zos Sgr	LMS IMS	
SALTMARSH Pink 255/200/250	Saltmarsh	Sm	LMU	
SHINGLE, COARSE SAND & MIXED SEDIMENTS Grey/brown 135/120/105	Shingle & gravel shore Estuarine coarse sediment Shallow gravel faunal communities Shallow mixed sediment communities Estuarine sublittoral gravels & sands Estuarine sublittoral mixed sediments Habitat complex; Circalittoral mixed sediments Habitat complex: Circalittoral gravel & sand	Sh Est Fag FaMX EstGS EstMX n/a n/a	LGS LGS IGS IMX IGS IMX CMX CGS	

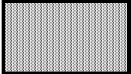
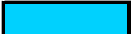
Life form & colour code	MNCR Biotope complex	MNCR Biotope complex code	MNCR Habitat complex	MNCR Mapping colours
SAND Yellow  255/255/0	Sand shores Shallow sand faunal communities	S FaS	LGS IGS	 
MUDDY SAND Brown  100/50/0	Muddy sand shores Shallow muddy sand faunal communities Habitat complex; Circalittoral muddy sands	MS FaMS n/a	LMS IMS CMS	  
MUD Dark brown  85/0/0	Sandy mud shores Soft mud shores Shallow marine mud communities Estuarine sublittoral muds Habitat complex; Circalittoral muds	SMu Mu MarMu EstMu n/a	LMU LMU IMU IMU CMU	    

MNCR Habitat codes: LR Littoral rock; LR(other) Rock pools, caves and overhangs; ELR Exposed littoral rock; MLR moderately exposed littoral rock; SLR Sheltered littoral rock; LGS Littoral gravels and sands; LMS Littoral muddy sands; LMU Littoral muds; IR(other) Vertical rock; EIR Exposed infralittoral rock; MIR Moderately exposed infralittoral rock; SIR Sheltered infralittoral rock; ECR Exposed circalittoral rock; MCR Moderately exposed circalittoral rock; SCR Sheltered circalittoral rock; IGS Infralittoral gravels and sands; CGS Circalittoral gravels and sands; IMS Infralittoral muddy sands; CMS Circalittoral muddy sands; IMU Infralittoral muds; CMU Circalittoral muds; IMX Infralittoral mixed sediments; CMX Circa littoral mixed sediments.

Table 23.3. The correlation between the MNCR mapping colours and associated MNCR habitat complexes and biotope complexes, and the life form colour coding system. See Table 1 for full explanation of codes.

MNCR Habitat complex	MNCR Mapping colour	MNCR Biotope complex	MNCR Biotope complex code	Life form	Life form colour code
ELR		Robust fucoids or red seaweeds	FR	Fucoids	
ELR		Mytilus & barnacles	MB	Mussels & barnacles	
EIR		Kelp with cushion fauna, foliose red seaweeds or coralline crusts	KFaR	Kelp	
EIR		Foliose red seaweeds	KFaR FoR*	Algal turf	
EIR		Robust faunal cushions & crusts (in part: planar areas)	SG	Faunal & algal crusts	
EIR		Robust faunal cushions & crusts (in part: vertical walls)	SG	Short fauna lturf, crusts & cushions	
ECR		Faunal crusts & short turfs	EFa	Short fauna lturf, crusts & cushions	
ECR		Alcyonium dominated communities	Alc	Faunal turf	
ECR		Barnacle & cushion sponge	BS	Faunal turf	
MLR		Barnacles & fucoids	BF	Fucoids	
MLR		Red seaweeds	R	Algal turf	
MLR		Ephemeral green & red seaweeds	Eph	Algal turf	
MLR		Mytilus & fucoids	MF	Mussels & barnacles	
MIR		Kelp & red seaweeds	KR	Kelp	
MIR		Grazed kelp with algal crusts	GzK	Kelp	
MIR		Sand or disturbed kelp & seaweed	SedK	Kelp	
MIR		Red seaweeds on sand-covered rock	SedK	Algal turf	
MIR		Sabellaria alveolata reefs	PolAhn*	Biogenic sand reefs	
MCR		Grazed fauna	GzFa	Faunal & algal crusts	
MCR		Mixed faunal turfs	XFa	Faunal turf	
MCR		Bryozoan/hydroid turfs	ByH	Faunal turf	
MCR		Ascidian communities	As	Faunal turf	
MCR		Cirralittoral Sabellaria reefs	CSab	Biogenic sand reefs	
MCR		Brittle star beds	Bri	Brittle star beds	
MCR		Mussel beds (sublittoral open coast)	M	Mussel beds	
SLR		Dense fucoids	F	Fucoids	
SLR		Fucoids, barnacles or ephemeral seaweeds	FX	Fucoids	
SLR		(also including biotope 'Ephemeral green & red seaweeds on variable salinity or disturbed mixed substrata')	FX EphX*	Algal turf	
SLR		Mytilus (mussel) beds on mixed substrata	MX	Mussel beds	
SIR		Silted kelp	K	Kelp	
SIR		Submerged fucoids, green & red seaweeds (lagoonal)	Lag	Fucoids	

MNCR Habitat complex	MNCR Mapping colour	MNCR Biotope complex	MNCR Biotope complex code	Life form	Life form colour code
SCR		Brachiopod & solitary ascidian communities	BrAs	Faunal turf	
SCR		Sheltered Modiolous(horse mussel) beds	Mod	Mussel beds	
LGS		Shingle & gravel shore	Sh	Shingle, coarse sand &	
LGS		Estuarine coarse sediment	Est	Shingle, coarse sand &	
LGS		Sand shores	S	Sand	
IGS		Maerl beds (open coasts/clean sediments)	Mrl	Maerl beds	
IGS		Shallow gravel faunal communities	Fag	Shingle, coarse sand & mixed sediments	
IGS		Estuarine sublittoral gravels & sands	EstGS	Shingle, coarse sand & mixed sediments	
IGS		Shallow sand faunal communities	FaS	Sand	
CGS		Habitat complex: Circalittoral gravel & sand	n/a	Shingle, coarse sand & mixed sediments	
LMS		Zostera noltii	Zos	Sea grass beds	
LMS		Muddy sand shores	MS	Muddy sand	
IMS		Seagrass beds	Sgr	Sea grass beds	
IMS		Shallow muddy sand faunal	FaMS	Muddy sand	
CMS		Habitat complex; Circalittoral muddy sands	n/a	Mud	
LMU		Saltmarsh	Sm	Saltmarsh	
LMU		Sandy mud shores	SMu	Mud	
LMU		Soft mud shores	Mu	Mud	
IMU		Shallow marine mud communities	MarMu	Mud	
IMU		Estuarine sublittoral muds	EstMu	Mud	
CMU		Habitat complex; Circalittoral muds	n/a	Mud	
IMX		Oyster beds	Oy	Oyster beds	
IMX		Maerl beds (muddy mixed sediments)	MrlMx	Maerl beds	
IMX		Shallow mixed sediment communities	FaMX	Shingle, coarse sand & mixed sediments	
IMX		Estuarine sublittoral mixed sediments	EstMX	Shingle, coarse sand & mixed sediments	
CMX		Biotope Sabellaria spinulosa & Polydora on mixed sediments	n/a SspiMX*	Biogenic sand reefs	
CMX		Habitat complex; Circalittoral mixed sediments	n/a	Shingle, coarse sand & mixed sediments	
LR		Lichens or algal crusts	L	Lichens & algae	
LR(other)		Rock pools	Rkp	Fucoids	
LR(other)		Overhangs & caves	Ov	Fucoids	
LR(other)		Rock pools (in part)	Rkp	Algal turf	

MNCR Habitat complex	MNCR Mapping colour	MNCR Biotope complex	MNCR Biotope complex code	Life form	Life form colour code
LR(other)	?	Overhangs & caves (in part)	Ov	Lichens & algae	
IR(other)	?	Fauna & seaweed (shallow vertical rock)	FaSwV	Short fauna lturf, crusts & cushions	
CR	?	Caves & overhangs (deep)	Cv	Faunal turf	
CR(other)	?	Faunal turfs (deep vertical rock)	FaV	Faunal turf	
?		Serpulid reefs	?	Calcareous biogenic reefs	

Appendix 2: Modified Folks triangle used in sediment analysis

