ICES CM 2004/T:13

Considerations in large-scale acoustic seabed characterization for mapping benthic habitats

J.M. Preston, A.C. Christney, W.T. Collins, R.A. McConnaughey, S.E. Syrjala

J.M. Preston, A.C. Christney, W.T. Collins: Quester Tangent Corporation (QTC), 201-9865 West Saanich Road, Sidney, BC, V8L 5Y8, Canada [tel: 1 250 656 6677, fax: 1 250 655 4696, e-mail: jpreston@questertangent.com]; R.A. McConnaughey, S.E. Syrjala: NMFS Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA, 98115, USA, [tel: 1 206 526 4150, fax: 1 206 526 6723, e-mail: bob.mcconnaughey@noaa.gov]

In 1999, NMFS Alaska and QTC collected about 18,000 line miles of seabed acoustic data at 38 and 120 kHz from the eastern Bering Sea. With four million echoes at each frequency, this data set permitted thorough explorations of some practical considerations that influence every acoustic seabed classification. Our unsupervised classification involved an objective determination of the optimal number of classes for each of the pre-classification methods we explored, allowing useful comparisons among methods. Stacking, one of the pre-classification steps, is the process of averaging sequential echoes to allow sediment information to express itself in spite of ping-to-ping variability. With stacks of fifty pings, feature spaces had more detail and better defined clusters, thus more classes in unsupervised classification, compared to stacks of five pings. Classification by echo shape requires changing the sampling rate to compensate for depth changes. While effective, this changes the apparent roughness and the amount of detail submitted to the feature-generating algorithms. Depth and stack size affect spatial resolution; the scale of the survey and the sharpness of sediment boundaries guide the surveyor's choice of spatial resolution. Even such a huge data set is a sampling of the sea bottom, and further sub-sampling simplified feature spaces further, reducing the optimal number of classes. The two frequencies differed in beam width and sediment penetration, and thus gave complimentary information. The influences of each of these pre-classification processes and other considerations are presented, with representative maps of acoustic diversity and with statistical comparisons, accompanied by preliminary correlations with fish census data.

Keywords: acoustic seabed classification, seabed characterization, benthic habitat, hydro-acoustic remote sensing

Fish and Invertebrate Abundance in the Eastern Bering Sea

Research trawl surveys are conducted annually in the eastern Bering Sea (EBS) by the Resource Assessment and Conservation Engineering Division of the Alaska Fisheries Science Center, National Marine Fisheries Service (NMFS), to provide data for stock assessments and management of the fishery resources in the region [1]. The surveys are conducted each June-August in a 463,400 km^2 region of the EBS shelf (Fig. 1) at depths between about 20 and 200 m. An 83-112 Eastern otter trawl is deployed from chartered vessels within each of 355 standard stations defined by a 20 by 20 nautical mile sampling grid. Each sample consists of a 30-min tow taken at a speed of 3 kn. The catch is individually processed to determine total weight and numbers by species and a variety of measurements and biological samples are collected from individual specimens.

In general, the EBS continental shelf is a relatively shallow and level basin, with an average depth of 60 m and an average slope less than 0.3 m/km. Mean grain size generally decreases with distance from shore and/or increasing water depth [2]. Sand predominates overall, with very little gravel and a conspicuous lack of mud over the inner and middle shelf. The inner shelf has mixed gravel and sand near shore changing to sand on the mid-shelf, and mixed sand and mud farther offshore. The middle shelf is largely mixed sand with mud extending in a broad band from southeast to northwest and generally following the bathymetry. This pattern is disrupted by a rather continuous swath of sand along a line extending from the Pribilof Islands to the western tip of Nunivak Island, with rocky material along its northwestern edge. Surficial sediments of the outer shelf are, again, largely mixed sand and mud, with mixed gravel and sand along the outer margin southeast of the Pribilofs. Mud is common near the northwest corner of the outer shelf and also west and northwest of the Pribilofs. Isolated patches of sand are found along the outer margin of the EBS shelf.

As a group, flatfish species occur throughout the EBS, although there are significant differences in distribution and abundance of individual species. Using published sediment descriptions and trawl survey data, McConnaughey and Smith [3] demonstrated spatially explicit relationships between

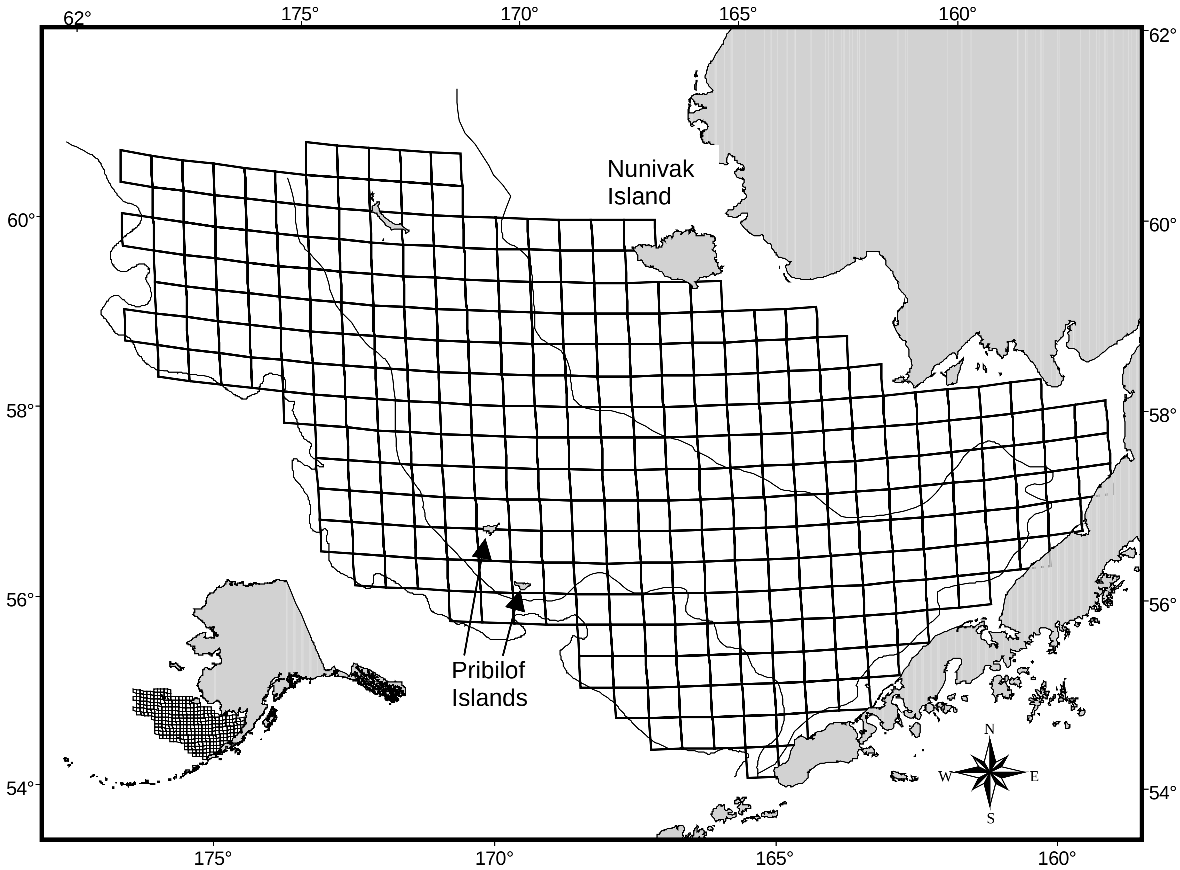

Fig. 1. Schematic diagram of the annual NMFS bottom trawl survey of the eastern Bering Sea continental shelf. Stations, at the centre of each square, are generally located 20 nautical miles apart. The 50, 100 and 200 m isobaths define the inner, middle and outer shelf areas.

pleuronectid flatfish abundance and surficial sediments there. Sediment textures in areas of highest- and lowest-abundance were compared for six species. Sand predominated in areas of high yellowfin sole (YFS) and rock sole (RS) abundance, while mixed sand and mud was most common in areas of lowest abundance. In contrast, mixed sand and mud predominated in areas preferred by flathead sole (FHS), Alaska plaice (AP) and arrowtooth flounder (ATF), with more diverse substrates in low-density areas. Areas of high and low Greenland turbot (GT) abundance had similar sediment textures (primarily mixed sand and mud). Food habits of these species were examined in these areas of high and low abundance. Species with highly restricted diets (AP) or piscivores with weak sediment associations (GT, ATF) had relatively inflexible food habits, whereas YFS, RS and FHS food habits varied considerably with sediment type. Uneven spatial distributions, strong sediment preferences and substrate-mediated food habits suggest that benthic-feeding species prefer certain sediment types because of adaptive differences in the availability and quality of benthic prey.

Although prey availability is likely to be influencing flatfish distributions in the EBS, rather than surficial sediments per se, it is clearly more cost-effective to map surficial sediments as a proxy. This is because benthic fauna have inherently high spatial and temporal variability, not to mention considerable sample-processing costs. However, because of the need for dedicated ship time and laboratory processing of samples, even traditional methods of collecting sediment data can be rather costly and inefficient. Moreover, broad scale mapping of sediment properties based on a limited number of samples may obscure spatial variability at biologically relevant scales. Compilations of historical data represent one option, if sufficient data are available and compatible methodologies have been used [2]. Synoptic imaging with sidescan and multibeam sonars may also be appropriate, although groundtruthing with traditional methods is required and interpretation of the sonograms includes a subjective component. The QTC VIEW™ system [4-6] is another class of acoustic tool developed expressly for seabed classification. It operates in the background using a vessel’s single-beam echosounder and statistical classification methods. Because numerous studies, including work in Alaska [7,8], have demonstrated that the QTC VIEW™ system is able to detect and map seabed types with distinct acoustic properties, the present study was undertaken to evaluate the system's utility for large-scale fish habitat characterizations as mandated by U.S. statute.

Acoustic Segmentation

In QTC VIEW™ and the associated software suite QTC IMPACT™, seabed characterization is based on statistical segmentation of a survey area into regions from which echoes are acoustically similar. Statistical methods are used because inversion of models of high-frequency backscatter is rarely unique, and thus has limited applicability. For surveys over a wide range of sediments, including mixtures, a phenomenological approach such as this is the preferred practical choice and can be used with uncalibrated echo sounders.

Systematically exploiting details in backscatter is the basis of acoustic segmentation. Echo characteristics depend on the type of sediment for several reasons, particularly the way that the backscatter cross-section varies with angle of incidence in different ways with different sediments. Suitable echo details include amplitude at nadir and at other grazing angles, rise time, echo duration, and spectral character. Details like these, and others with explanations that are more obscure, are called features. The key steps in making good features are quality control of the raw echoes, pre-processing to address ping-to-ping variability and depth dependence, and design of the feature algorithms. This paper discusses pre-processing and related considerations. References 4-6 contain information on the other steps as implemented in Quester Tangent (QTC) products.

Spatial Resolution and Stacking

Features are not made from individual echoes because ping-to-ping variability has been found to mask sediment dependence. QTC acoustic bottom classification is based on echo details, and these details can easily be obscured by noise in any single ping. In any set of pings that have the same underlying details, stacking improves classification resolution by better exposing echo details to the feature-generating algorithms. The stacking process is to shift echo time series forward or backward to align the samples at which the echoes were deemed to start, and then to sum them. Dividing by the number summed is optional; dividing would not affect the results because all echoes are treated identically. A window of each stacked time series, starting shortly before the bottom pick and long enough to contain the full echo, is presented to the feature algorithms. Stacks of five echoes are the usual practice; a goal of this work was to explore the effects of larger stacks.

Fig. 2. Idealised footprint of a stack of seven echoes. The ship advances the distance a between pings and each ping has a footprint of radius r.

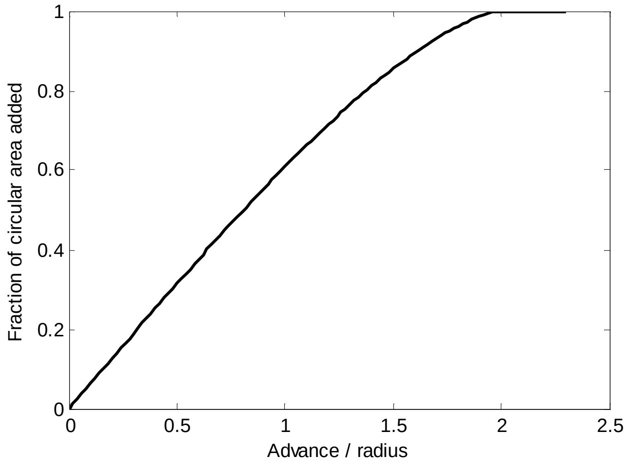

Fig. 3. The fraction, ξ, of the full circular area that a ping, other than the first one, adds to the area insonified by a stack of pings.

The area, *A*, imaged by each stack of n echoes is the sum of the non-overlapping parts of *n* circles, where the radius of each circle is $r = depth \times \tan(beamwidth/2)$, taking the beam pattern to be a cone. The distance between circle centres is *v*/*f*, where *v* is the ship speed and *f* the ping frequency. As sketched in Figure 2, this distance is here called the advance, $a = v/f$. To calculate the area insonified by a stack of pings, start by considering that the first ping in the stack contributes the full circular area, πr^2. Subsequent pings add the same full area if $a \geq 2r$, that is, if their footprints do not overlap, but less if they do. Figure 3 shows the fraction of the full circular area that each subsequent echo adds. The insonified area for the stack is the sum of the areas of the first circle and the fraction ξ of the other *n*-1 circles.

$$A = \pi r^2 + \xi(n-1)\pi r^2 = \pi r^2(1+\xi(n-1)) \tag{1}$$

Let us consider two models of the boundaries between surficial sediments: sharp and gradual compared to the distance the ship travels while acquiring the number of echoes that are stacked together. With sharp boundaries, stacking could introduce artificial classes as the survey ship crosses a boundary, as echoes of two distinct types are added together to create a composite echo that does not correspond to a real sediment type. With gradual boundaries, echoes from legitimate intermediate sediment might be recognised as a class if the number being stacked is large enough to suppress noise and allow subtle echo details to be exposed to the feature algorithms.

This suggests that the number of echoes to be stacked should be as large as possible, limited only by the spatial resolution that must be achieved to meet the survey requirements. The number of classes that can be resolved should increase with the number of echoes stacked together. However the boundaries between classes must be gradual compared to the distance the ship travels while acquiring a stack of echoes, or artificial classes could be introduced. Another upper limit to the stack size is that there must be enough stacks for the subsequent statistical processing.

<u>Depth compensation</u>

Depth compensation is the process of changing the sampling rate of each echo time series so that the echo duration, in number of samples, is what it would have been had it come from the reference depth. For example, an echo from water 200-m deep is about twice as long as one from 100 m, so every second data point from the deep echo is discarded, which halves the sampling rate. Without compensation, features that depend on rise time, decay time, and other time-related echo details would be very different even if both these echoes were from the same sediment. Reference 9 is a recent paper on depth compensation by changing the sampling rate.

Usually, one chooses a reference depth near the mean (arithmetic or, preferably, geometric) depth of the survey area. If the echo is from deeper than the reference depth its time series is decimated to reduce the number of samples; if shallower it is interpolated. In principle, the choice of reference depth should have no effect on classifications. In practice, there are two issues: having enough data points and distorting echo details. The echo time series might end too soon for echoes from near the deep end of the sounder's phase window. The user cannot predict how many data points will be needed before a deep echo is shrunk by interpolation. This was not an issue in the Bering Sea data set that is discussed below. Distortion can occur if the actual sample rate was barely adequate for the echo bandwidth and more than two or three points are interpolated between pairs of recorded data points. Details can also be damaged by shrinking the time series too much, as this may involve discarding data points that carry useful information, depending again on the natural echo bandwidth. Thus fidelity of the echo time series argues, on the one hand, for a deep reference depth so that the echo details are represented by many points in the time series and, on the other hand, for a medium reference depth to avoid interpolating by a factor of more than about three.

Bering Sea Survey

Acoustic data were collected during the 1999 annual echo-integration-trawl survey [10]. Acoustic classification was an incidental part of this acoustic survey to quantitatively assess midwater fish abundance. Acoustic data for classification were collected without interfering with the primary goals. For the mapping research, the overall aim was to seek optimal conditions for acoustic classification of benthic habitats in the EBS, that is, to select the combination of resolution, stacking, and depth compensation that best characterize acoustic diversity there. Optimal classifications, in this context, strike a balance between high information content (many classes) and high confidence (few classes).

After connecting the QTC VIEW™ full waveform acquisition system (ISAH-S) to the Simrad EK-500 scientific echosounder on the NOAA ship *Miller Freeman*, over 9,000 miles of track line data were collected in the EBS between June and August 1999. The sounder operated simultaneously at two

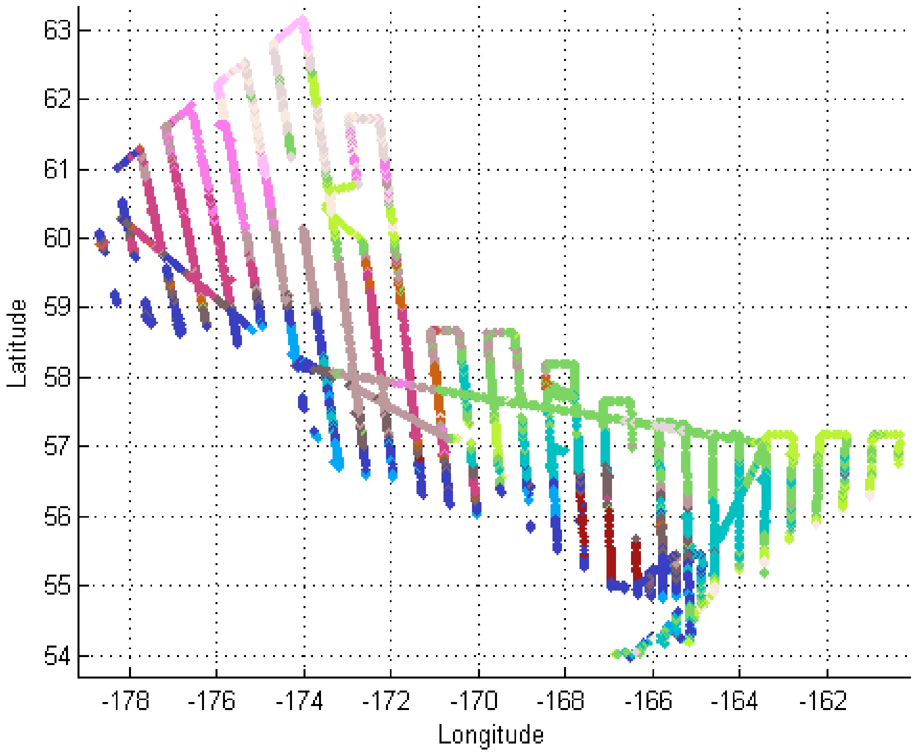

Fig. 4. Map of the 14 classes in the optimal model for the sub-sampled data set for 38 kHz, stacks of 50 echoes, and reference depth of 90 m. Similar colours denote classes that are acoustically similar, thus near each other in feature space.

frequencies, 38 kHz and 120 kHz. The raw data consist of about four million echo envelopes at each frequency. Depths of interest were between 40 and 155 m; echoes from beyond this depth range were filtered, which eliminated concerns about inadequate length of the echo time series when compensating for depth. Three stack sizes were used in post-processing with QTC IMPACT™: 5, 15 and 50 echoes, each with reference depths of 90 and 150 m. There were thus twelve combinations of stack size, reference depth, and sonar frequency.

The well-established methods used by QTC for seabed classification are based on the details of echo formation at the sea bottom. Surface roughness, acoustic impedance, and volume homogeneity all are characteristic of the sediment type and all influence details of the echo of a vertical-incidence echo sounder. QTC has developed a suite of algorithms that extract from the echoes features rich in seabed character. Applying principal components analysis combines the many features into three combinations that contain a very large fraction of the variance. Clustering analysis in three-dimensional space (called Q space) groups the echoes into distinct clusters based on their acoustic diversity. The acoustic diversity directly represents substrate diversity. Figure 4 is one of the better maps of classes, better, that is, in the context of the considerations of this paper.

Ranking Pre-Classification Processes

A goal of this work was to discover which pre-classification process, namely stacking and depth compensation, was best suited to long-term repetitive acoustic surveys in the EBS. Two steps were required. First generate the optimal set of clusters for each process, and second compare the clusters statistically.

Objectivity was the over-riding consideration in clustering. QTC clustering had, until this project, been a user-guided process. An objective method has recently been developed [11]. It determines both the optimal number of clusters and the best assignment of data points, one from each stack, to the clusters. The method is to minimize the Bayesian Information Content (BIC) [12] of candidate splittings, thus moving to an optimal segmentation through simulated annealing.

Objective clustering yields a statistically optimal result for each of the twelve processes, but is not a method for ranking them. This required specialized statistical techniques for comparing the quality of clusters [13, Appendix E].

For both frequencies, this statistical analysis showed that the optimal reference depth was the shallower one, 90 m, and that the larger the stack size the better defined were the clusters. These rankings, obtained from sub-samples of between 4788 and 4961 stacks from each pre-classification process, are in the right-hand column of Table 1. The same table shows that the optimal number of clusters increased with the number of echoes in a stack.

Frequency (kHz)	Echoes in Stack	Reference Depth (m)	Optimal Number of Clusters			Ranking within frequency
			Lowest BIC	Not statistically different	Conclusion	
38	5	90	11	5, 6	5	6
	15		10		10	4
	50		14		14	1
	5	150	9	6, 7, 8	6	5
	15		9		9	3
	50		12		12	2
120	5	90	6		6	6
	15		10	4, 7, 8	4	5
	50		13	12	12	1
	5	150	5		5	2
	15		8		8	4
	50		10		10	2

Table 1. Optimal number of clusters in the Bering Sea data set as determined by objective clustering for the six variations of processing parameters at each sonar frequency.

Spatial Resolution and Stack Size

For a linear survey with few track crossings, as is the case here, the length of the set of footprints (Fig. 2) is more pertinent than their area. For the largest stacks, 50, and in the deepest water, 150 m, the length is about 240 m, which is 0.002 degrees of latitude. The track plots cover 9 degrees of latitude. For habitat studies, this along-track spatial resolution of about 2 parts in ten thousand is adequate. Thus the stack size can be chosen to optimise the classification, as the trade-off in resolution is acceptable.

Suppressing ping-to-ping variability by stacking more consecutive echoes together allows the sediment information in the echo shape and spectral nature to express itself in spite of noise-like variability. In feature space, large-stack clusters are better separated from their neighbours because the “smearing” from ping-to-ping variability has been reduced. To define optimal classes, one should stack as many echoes as possible, consistent with the required spatial resolution.

Reference Depth

The results show that limiting distortion in interpolation is more important than expanding each echo time series to display the maximum detail. For both frequencies, the channel with the best statistical scores had been processed with a reference depth of 90 m, that is, near the mean depth rather than near the deepest depth. With an intermediate reference depth, the decimation and interpolation factors by which the echo time series is shrunken or expanded are small. For example, consider an echo from a shallow area, 40 m deep. Interpolating to 90 m is an expansion of 90/40 = 2.25, and to 150 m is a ratio of 150/40 = 3.75. Interpolation factors above three exceed a rule of thumb for distortion, in that many points are being inserted between the actual data points. Our results imply that the additional echo detail available with the deeper reference depth is less important to the classifications than the distortion that occurs with interpolating with a factor this large.

Optimal Sonar Frequency

Whether 38 kHz or 120 kHz is the better choice for sediment classification in the Bering Sea is not an issue that can be settled by any statistical criterion. The two frequencies produce different information about the bottom, because of their different penetration distances, wavelengths, and beamwidths. Careful assessment and comparison of the maps is a method for deciding which sonar frequency is the better choice for routine surveys in the Bering Sea. The classifications based on 38 kHz echoes are more contiguous, have more consistent track crossings, and conform better to our broad understanding of the sediment types in the basin. Classifications based on 120 kHz echoes tended to be more variable along the survey tracks (based on maps that are not shown here). This is consistent with the expected sensitivity to near-surface differences and small-scale bottom details, which might be as relevant, or more relevant, to fish habitat than the information from a frequency that penetrates deeper. Thus it seems that 38 kHz is a better choice for classifying the Bering Sea into sediment types and 120 kHz is a better choice based on near-surface differences.

Conclusions

Objective statistical methods have been used to determine which of several processing methods gave optimal acoustic classification results. This involved developing several statistical methods for acoustic classification: a variant of simulated annealing to divide the records into classes in an optimal way and methods for comparing the quality of the resulting clusters. The data set used for this work was the largest data set ever used for acoustic classification.

The results given here are based on statistical comparisons and indicators, not on correlations with sediment or fish-census data. It was found, statistically, that feature spaces derived from stacks of 50 pings had more detail and better defined clusters, thus more classes in unsupervised classification, than those from stacks of five pings. The trade-off in spatial resolution was acceptable in this case. Classification by echo shape requires changing the sampling rate of the echo time series to compensate for the effects of depth on echo duration. While effective, changing the sampling rate also changes the apparent sizes of objects on the seabed and the amount of detail submitted to the feature-generating algorithms. It was found that a reference depth near the mean of the survey depths was preferable to a depth near the deepest in the survey.

There would be substantial benefits to fishery management if these acoustic classifications were shown to be correlated with distribution and abundance data from stock assessment surveys. Once demonstrated, it would then be prudent to determine the physical attributes of the acoustic classes in order to understand the mechanism of association, and also to investigate the feasibility of integrating data from multiple acoustical platforms so as to improve the efficiency of the mandated mapping effort. Statistical analyses of this type are currently underway using the acoustic data discussed here.

References

1. E. Acuna and S. Kotwicki, “2003 bottom-trawl survey of the eastern Bering Sea continental shelf. U.S. Dept. Commer., AFSC Processed Report 2004-02, Seattle, 2004.
2. K.R. Smith and R. A. McConnaughey, “Surficial sediments of the eastern Bering Sea continental shelf: EBSSED database documentation”, U.S. Dept. Commer., NOAA Tech. Memo. NMFS-AFSC-104, 1999.
3. R.A. McConnaughey and K.R. Smith, “Associations between flatfish abundance and surficial sediments in the eastern Bering Sea”, Can. J. Fish. Aquat. Sci. 57, 2410-2419, 2000.
4. J. M. Preston, A. Rosenberger, and W. T. Collins, “Bottom classification in very shallow water”, Proceedings of the fifth European Conference on Underwater Acoustics, Lyon, France, 293-299, 2000.
5. R. Freitas, S. Silva, V. Quintino, A.M. Rodrigues, K. Rhynas, and W.T. Collins, “Acoustic seabed classification of marine habitats: studies in the western coastal-shelf area of Portugal”, ICES J Marine Science, 60, 599-608, 2003.
6. J.T. Anderson, R.S. Gregory, and W.T. Collins, “Acoustic classification of marine habitats in coastal Newfoundland”, ICES J Marine Science, 59, 156-167, 2002.
7. P.G. von Szalay, “The feasibility of using single beam seabed classification systems to identify and quantify slope rockfish habitat in the Gulf of Alaska”, M.S. Thesis, School of Fisheries, University of Washington, Seattle, 1998.
8. P.G. von Szalay, and R.A. McConnaughey, “The effect of slope and vessel speed on the performance of a single beam acoustic seabed classification system”, Fish. Res. (Amst.) 56, 99-112, 2002.
9. Z. Lubniewski and E. Pouliquen, “Sensitivity of echo parameters to seafloor properties and depth variability”, Proceedings of the seventh European Conference on Underwater Acoustics, Delft, Holland, 763-768, 2004.
10. T. Honkalehto, N. Williamson , S. de Blois, and W. Patton “Echo-integration-trawl survey results for Walleye Pollock *Theragra chalcogramma* on the Bering Sea shelf and slope during summer 1999 and 2000”, available at http://www.afsc.noaa.gov/Publications/AFSC-TM/NOAA-TM-AFSC-126.pdf, 2002.
11. J. M. Preston, A. C. Christney, L. S. Beran, and W. T. Collins, “Statistical seabed segmentation – from images and echoes to objective clustering”, Proceedings of the seventh European Conference on Underwater Acoustics, Delft, Holland, 813-818, 2004.
12. C. Fraley and A.E. Raftery, “How many clusters? Which clustering method? Answers via model-based cluster analysis”, Computer Journal, 41, 578-588, 1998.
13. Quester Tangent Corporation, “QTC VIEW™ acoustic seabed classification, Bering Sea, Miller Freeman 1999”, project report to National Marine Fisheries Service, Seattle, 2004.