ICES CM 2010/G41

Spatio-temporal Patterns of Near-surface Acoustic Backscatter in the Eastern Bering Sea Based on Multi-frequency Analysis

Mathieu Woillez[a,b], Patrick H. Ressler[a], Christopher D. Wilson[a] and John K. Horne[a,b]

[a] Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, USA
[b] School of Aquatic and Fishery Sciences, University of Washington, Seattle, USA

Introduction

Fisheries acoustic surveys often collect multi-frequency data to help classify the species composition of the observed layers and aggregations. Because fishes are usually the primary target species, relatively little sampling effort is typically spent to confirm the identity of the other organisms. Methods to characterize spatio-temporal patterns of these other scatterers based on the acoustic data could provide important ecological insights to facilitate an ecosystem approach to fisheries.

Objective

Develop a robust and objective method to partition acoustic data into clusters based on multi-frequency patterns in the data, and then characterize the spatio-temporal patterns of the resulting clusters.

Application

The method described here will be used to characterize a persistent, near-surface acoustic backscattering layer that exists throughout most of the eastern Bering Sea (EBS). The species composition of this layer, and its role in the EBS ecosystem is unknown.

Methods

Multi-frequency data

Acoustic data were collected at 18, 38, 120, and 200 kHz with Simrad EK60 echosounders in the EBS during summers 2004, 2006-2008. Noise from range attenuation (120 and 200 kHz) and ringing (18 kHz) was estimated. Only data with a signal-to-noise ratio ≥ 10 dB at any frequencies, and with S_v > -80 dB at one or more frequencies were used in subsequent analyses (e.g., Fig. 1).

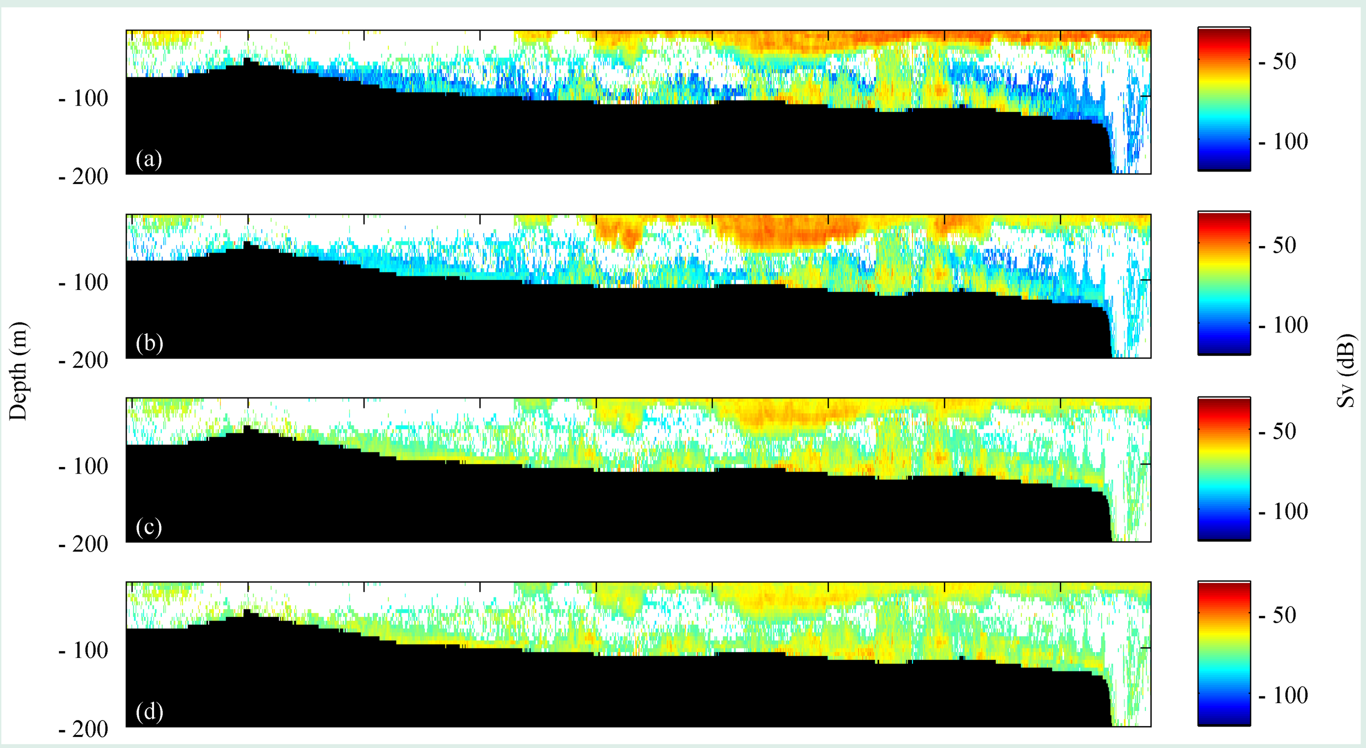

Figure 1. Volume backscattering strength (S_v in dB) at 18, 38, 120 and 200 kHz (a, b, c and d) that meet signal-to-noise and threshold criteria, transect 20, summer 2007.

Choosing the data support

Support is the point, the surface, or the volume over which a variable is measured or defined. As the amount of acoustic data was very large, aggregation was helpful for the continuation of the analysis. However, the spatial structure of the data might suggest the appropriate scale at which the phenomenon should be described. Variograms were used to determine the correct support. Directional indicator variograms on presence-absence of backscatter were computed and normalized by their sample variance to compare their spatial structure between years.

Unsupervised clustering using Gaussian mixture model

Let Y be the pair-wise frequency difference observation vector to be modeled by a mixture of K multivariate Gaussian distributions. Its vector of parameters (i.e. the mixing proportions π, the means μ and the covariance matrix Σ) can be learned using a recursive Expectation-Maximization algorithm (Anderson et al., 2007).

$$P(y_i \mid \Theta) = \sum_{k=1}^{K} \pi_k P(y_i \mid \mu_k, \Sigma_k)$$

An Initial estimation of the parameters is first made by using a K-means algorithm, then the Expectation step computes the probability of data point y_i belonging to cluster k given the current parameter vector.

$$P(k \mid y_i) = \frac{P(k)P(y_i \mid k)}{P(y_i)}$$

The Maximization step consists of updating the parameters.

$$\pi_k = \frac{1}{N}\sum_{i=1}^{N} P(k \mid y_i) \qquad \mu_k = \frac{\sum_{i=1}^{N} P(k \mid y_i)\, y_i}{\sum_{i=1}^{N} P(k \mid y_i)} \qquad \Sigma_k = \frac{\sum_{i=1}^{N} P(k \mid y_i)(y_i-\mu_k)(y_i-\mu_k)^T}{\sum_{i=1}^{N} P(k \mid y_i)}$$

Cluster labels were assigned to samples by considering clusters for which the probability of membership was maximized. The Bayesian Information Criterion (BIC) was used to determine the optimum number of clusters.

Ground truth of the primary target species

Aggregation in midwater was sampled with trawls to identify the species composition of much of the acoustic data. Trawl locations, where catch compositions were dominated by a single species group (typically >90 %), were used in another EBS study to develop species-specific pair-wise frequency differences classification (De Robertis et al., 2010). Here the species-specific pair-wise frequency differences were used to identify the clusters representing the primary target species, which are walleye pollock and euphausiids (Fig. 2).

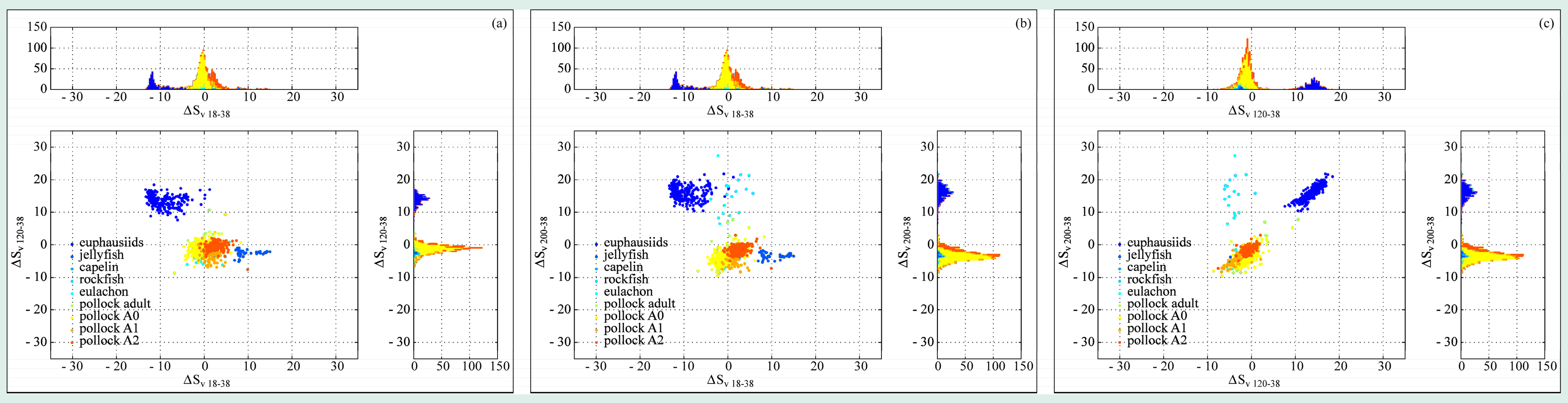

Figure 2. Pair-wise frequency differences from identified aggregations: scatter plot and histograms from $S_{v\ 18\text{-}38}$ and $S_{v\ 120\text{-}38}$ (a), $S_{v\ 18\text{-}38}$ and $S_{v\ 200\text{-}38}$ (b), $S_{v\ 120\text{-}38}$ and $S_{v\ 200\text{-}38}$ (c).

Cluster spatial patterns and inter-annual variability

Spatial patterns of each cluster were described using some spatial indices (Woillez et al., 2007). The indices characterized the location (centre of gravity), dispersion (inertia and isotropy), occupation (positive area) and aggregation (spreading and equivalent areas). They were used to describe both horizontal or vertical patterns.

To describe the inter-annual variability in cluster spatial patterns, a Multiple Factor Analysis (MFA) was used to provide a simultaneous representation of the acoustic clusters spatial patterns during the 4 surveyed years, in an approach similar to Woillez et al. (2007) and Petitgas and Poulard (2009).

Results

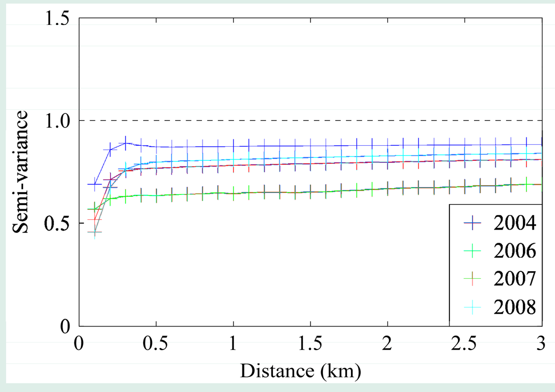

Figure 3. Normalized indicator variograms along transects computed for each surveyed year. Behavior near the origin was used to determine the correct support.

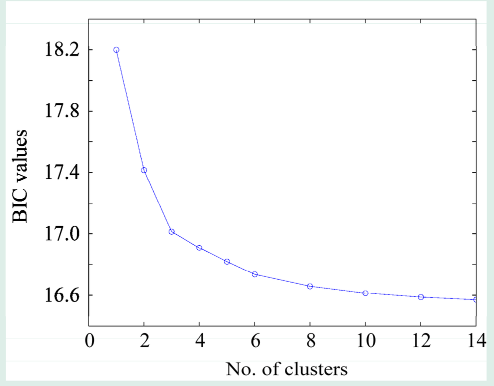

Figure 4. Bayesian Information Criterion versus the number of clusters.

- Data have been aggregated into 0.4 km bins horizontally (Fig. 3).
- The number of clusters used is a trade-off between a low value of the Bayesian Information Criterion and an interpretable number of clusters. Six clusters were chosen for the Gaussian mixture model (Fig. 4).
- Unsupervised clustering has been performed while pooling all years together to ensure consistency over the time series (Fig. 5). Some clusters represent major groups of scatterers identified in the ground truth data: for example, cluster 5-6 (red, orange) for euphausiids, cluster 3 (green) for swimbladdered fish (mainly walleye pollock), and cluster 1 (dark blue) for jellyfish.

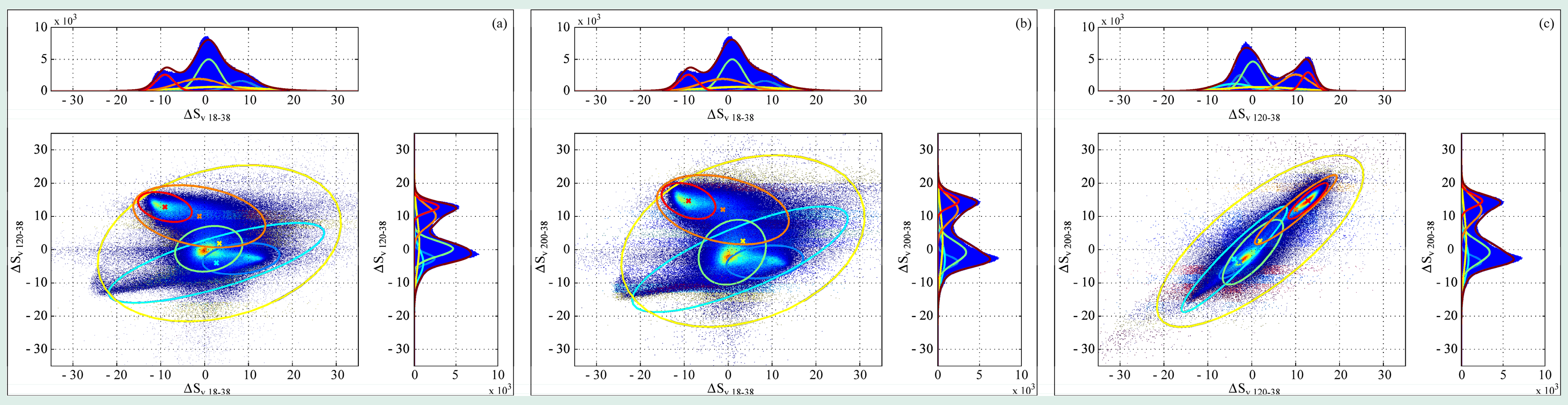

Figure 5. Pair-wise frequency differences in backscatter data from all 4 surveys: 1D and 2D histograms from $S_{v\ 18\text{-}38}$ and $S_{v\ 120\text{-}38}$ (a), $S_{v\ 18\text{-}38}$ and $S_{v\ 200\text{-}38}$ (b), $S_{v\ 120\text{-}38}$ and $S_{v\ 200\text{-}38}$ (c). The 6 fitted multivariate Gaussian distributions corresponding to each cluster have been represented on all graphs (colors red to blue).

- The clusters found are illustrated on an echogram. Spatial patterns appear which were not explicitly formulated in the clustering algorithms (Fig. 6).

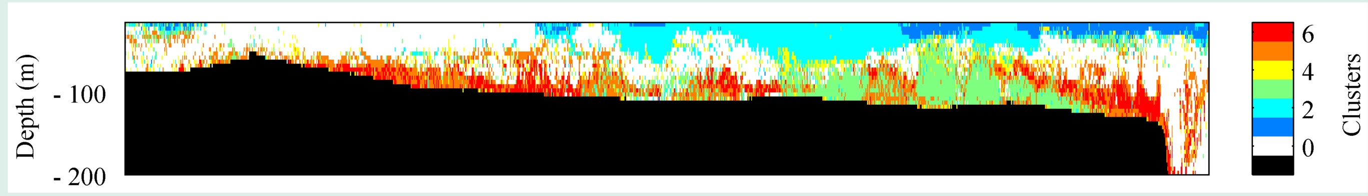

Figure 6. Corresponding clusters obtained by unsupervised clustering based on Gaussian Mixture Model for the transect 20, summer 2007. Data that did not meet signal-to-noise and threshold criteria are coded as 0, and the sea floor is coded as -1.

- The spatial patterns of each cluster can be described within and between surveys. For instance, the center of gravity (CG) of cluster 2 for summer 2007 is located at 172°W, 57.6°N and 28 m depth, and the mean CG of this cluster across all years is located at 171°W, 57.8°N and 30 m depth (Fig. 7a). Its occupation area decreased over the time series. Cluster 2 is one of the clusters comprising the recurring near-surface scattering layer found during summer in the EBS, along with the clusters 1, 4, and to a lesser extent 3 (Fig. 7b). Cluster 3 (mainly walleye pollock) was the least variable cluster across years in contrast to clusters 5 and 6 (euphausiids). It also had the highest value of occupation of space through the water column and northern most mean location.

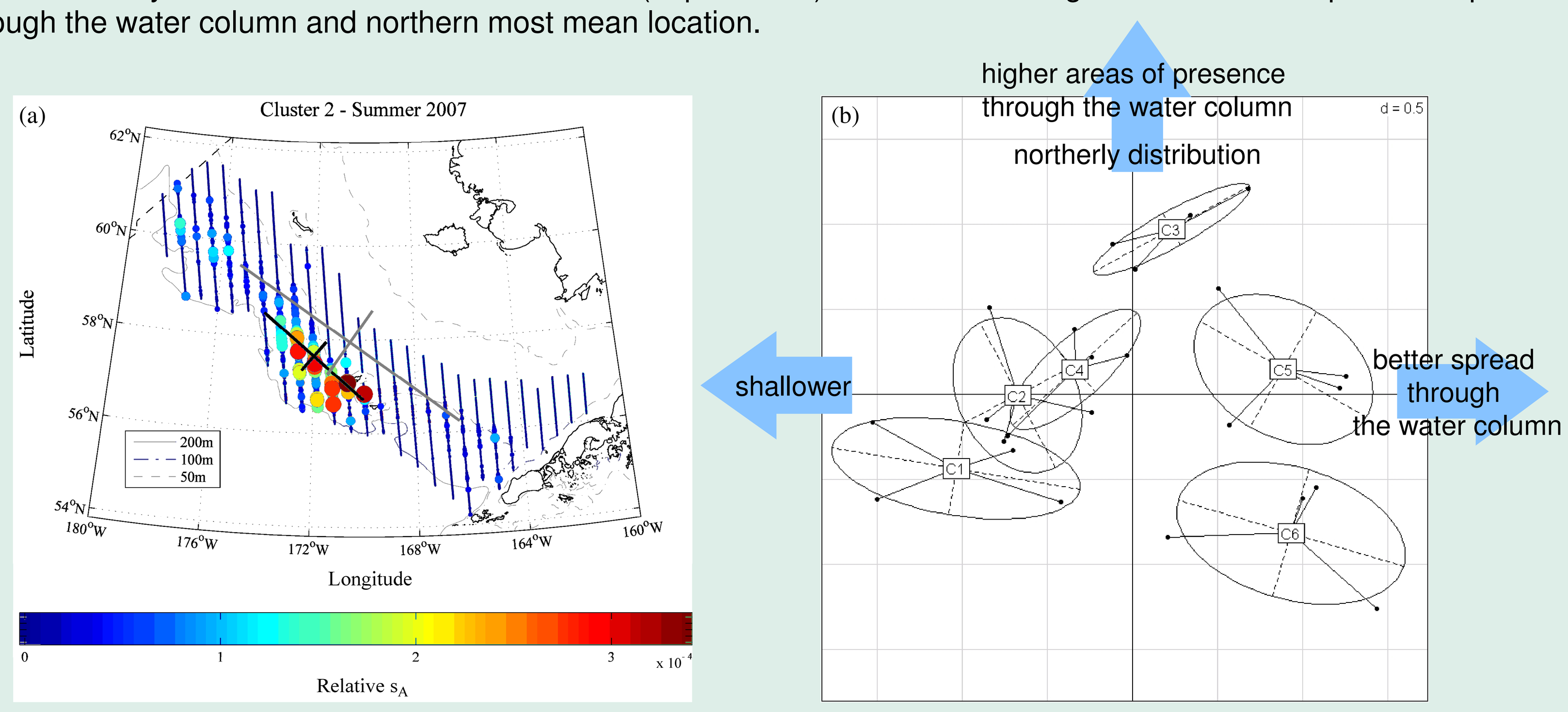

Figure 7. (a) Center of gravity and axes of inertia of the sample locations (grey) and of the relative s_A of cluster 2 (black) in summer 2007. (b) MFA representation of the cluster's spatial patterns of the surveyed ecosystem. The points represent the spatial distribution of each cluster in each year. The cluster labels represent the cluster-specific center of gravity, materializing the average cluster spatial pattern. Spatial indices that were correlated more than 3 times in the series of years with the MFA first two principal axes were used for the axes interpretation.

Conclusions

Frequency differencing, mixture modeling, and clustering for acoustic classification

The unsupervised clustering procedure identified known major scattering types (walleye pollock, euphausiids) and also several groups of unknown composition. Three clusters, of which one might be jellyfish, composed the near-surface layer.

Sampling recommendation

Unknown clusters of the near-surface layer should be sampled. Pair-wise frequency differences and spatial patterns will help to choose the most appropriate techniques and locations for sampling.

Monitoring the spatial distribution of a surveyed ecosystem

A multivariate spatial indicator that characterizes the spatial patterns over all clusters could be tracked over time and used for monitoring purposes.

References

Anderson, C.I.H., Horne, J.K., and Boyle, J. 2007. Classifying multi-frequency fisheries acoustic data using a robust probabilistic classification technique. Journal of the Acoustical Society of America, 121: 230-237

De Robertis, A., McKelvey, D.R., and Ressler, P.H. 2010. Development and application of an empirical multi-frequency method for backscatter classification. Canadian Journal of Fishery and Aquatic Sciences, 67: 1459-1474.

Petitgas, P., and Poulard, J.-C. 2009. A multivariate indicator to monitor changes in spatial patterns of age-structured fish populations. Aquatic Living Resources, 22(2): 165-171.

Woillez, M., Poulard, J.-C., Rivoirard, J., Petitgas, P. and Bez, N. 2007. Indices for capturing spatial patterns and their evolution in time with an application on European hake (*Merluccius merluccius*) in the Bay of Biscay. ICES Journal of Marine Science, 64: 537-550.

ICES
Annual Science Conference
20-24 September 2010
Nantes, France

Alaska Fisheries Science Center
NOAA FISHERIES

UNIVERSITY OF WASHINGTON

BERING SEA INTEGRATED ECOSYSTEM RESEARCH PROGRAM