Development and application of an empirical multifrequency method for backscatter classification

Alex De Robertis, Denise R. McKelvey, and Patrick H. Ressler

Abstract: We evaluated the feasibility of identifying major acoustic scatters in North Pacific ecosystems based on empirical measurements of relative frequency response. Acoustic measurements in areas where trawl catches were dominated by single taxa indicated that it might be possible to discern among key groups of scatterers such as fish with gas-filled swimbladders, euphausiids, myctophids, and jellyfish. To establish if walleye pollock (*Theragra chalcogramma*), a key species in the ecosystem, can be separated reliably from other groups under prevailing conditions, we developed a method based on the normal deviate (or *Z* score) to identify backscatter consistent with the pollock relative frequency response. We evaluated the performance of the method by comparing it with the traditional method of species identification (i.e., directed trawl catches and subjective interpretation of echograms) during five large-scale acoustic surveys of the eastern Bering Sea. Pollock abundance estimates employing the multifrequency method were highly correlated with those using the traditional method, which indicates that the multifrequency method performs well in this situation. In this environment, multifrequency methods will allow more inferences to be drawn when direct sampling of organisms is limited and will also complement existing abundance surveys by improving species classification and providing information about key nontarget species.

Résumé : Nous évaluons la faisabilité d'identifier les principales rétrodiffusions acoustiques dans les écosystèmes du Pacifique Nord, d'après des mesures empiriques des fréquences relatives des réponses. Les mesures acoustiques dans les régions dans lesquelles les prises au chalut sont dominées par un seul taxon indiquent qu'il devrait être possible de discerner les groupes principaux d'organismes responsables de la rétrodiffusion, tels que les poissons avec vessie natatoire remplie de gaz, les euphausiidés, les myctophidés et les méduses. Afin d'établir si la goberge de l'Alaska (*Theragra chalcogramma*), une espèce dominante de l'écosystème, peut être distinguée de façon fiable des autres groupes dans les conditions actuelles, nous avons mis au point une méthode basée sur l'écart normal (ou cote *Z*) pour identifier la rétrodiffusion correspondant aux fréquences relatives des réponses de la goberge. Nous avons évalué la performance de la méthode en la comparant avec la méthode traditionnelle d'identification des espèces (c'est-à-dire des récoltes dirigées au chalut et une interprétation subjective des échogrammes) durant cinq inventaires acoustiques à grande échelle dans l'est de la mer de Béring. Il existe une forte corrélation entre les estimations de l'abondance des goberges de l'Alaska basées sur la méthode des fréquences multiples et celles tirées de la méthode traditionnelle, ce qui indique que la méthode des fréquences multiples fonctionne bien dans les circonstances. Dans cet environnement, les méthodes des fréquences multiples permettront de faire plus de déductions lorsque l'échantillonnage direct des organismes est restreint et elles viendront compléter les inventaires actuels d'abondance en améliorant la classification des espèces et fournissant des renseignements sur les espèces importantes non ciblées.

[Traduit par la Rédaction]

Introduction

Acoustics is widely used for remote sensing of marine zooplankton and nekton (e.g., Holliday and Pieper 1995; Simmonds and MacLennan 2005). This approach offers many advantages for studies of the distribution and behavior of aquatic animals, including continuous sampling of large volumes of water and high-resolution measurements of vertical distribution and aggregation structure. The methodology for quantifying acoustic backscatter from echosounders is well developed (reviewed in Simmonds and MacLennan 2005). One of the major challenges in the use of acoustics in this context is the conversion of acoustic backscatter to biologically relevant information such as species composition and body size (Holliday and Pieper 1995; MacLennan and Holliday 1996; Horne 2000). The traditional approach to species identification (hereafter referred to as the traditional identification method) has been to combine acoustic

Received 29 October 2009. Accepted 15 May 2010. Published on the NRC Research Press Web site at cjfas.nrc.ca on 24 August 2010. J21493

Paper handled by Associate Editor J. Michael Jech.

A. De Robertis,[1] D.R. McKelvey, and P.H. Ressler. National Marine Fisheries Service, Alaska Fisheries Science Center, Resource Assessment and Conservation Engineering Division, 7600 Sand Point Way NE, Seattle, WA 98115, USA.

[1]Corresponding author (e-mail: Alex.DeRobertis@noaa.gov).

measurements with information from trawls or optical instruments and visual interpretation of echograms to interpret acoustic backscatter (e.g., McClatchie et al. 2000; Jech and Michaels 2006).

It has long been recognized that acoustic backscatter strength from zooplankton and fish is frequency-dependent and that this frequency dependence can be used to make inferences about the species composition and size distribution of the acoustic scatterers (reviewed in Holliday and Pieper 1995; Jech and Michaels 2006). As a result, broadband acoustic measurements (e.g., Simmonds et al. 1996; Zakharia et al. 1996) or, more commonly, narrowband measurements at multiple frequencies (e.g., Greenlaw 1979) have been used to make inferences about the identity of acoustic scatterers. Much of the early development of multifrequency acoustics was conducted in the context of zooplankton studies. In the frequency range commonly used (10s to 100s of kHz), zooplankton backscatter is highly frequency- and size-dependent, as these small organisms span the transition from Rayleigh to geometric scattering, which is not the case for larger fish with swimbladders (Holliday and Pieper 1995; Horne and Jech 1999; Horne 2000). Multifrequency approaches to backscatter classification have recently proliferated in the fisheries acoustics community due to (*i*) increased interest in information about species that are not subject to commercial fisheries (Pikitch et al. 2004; Koslow 2009) and (*ii*) increased availability of instrumentation, methods, and software for multifrequency acoustic measurement via multiple narrowband echosounders (Higginbottom et al. 2000; Korneliussen and Ona 2002; Korneliussen et al. 2008). This has led to the implementation of a diverse range of approaches using the difference in backscatter measurements at multiple frequencies (i.e., the relative frequency response) to improve acoustic species identification (Korneliussen and Ona 2003; Lavery et al. 2007; Benoit-Bird 2009). Substantial progress has been made, but the utility and effectiveness of any particular approach is situation-dependent: classification success depends on the species composition, relative abundance, relative frequency response, and spatial overlap of the species assemblage present in the environment.

Objective assessment of the effectiveness of multifrequency species identification methods under realistic conditions remains a challenge. The primary difficulty is that only a small subset of acoustic measurements can be verified with other sampling methods. Additionally, there can be substantial uncertainty even where verification of the species present is attempted because of sampling biases (McClatchie et al. 2000). In practice, many assessments of classification efficacy have relied on case studies of small data sets (e.g., comparison of trawl ground-truthed echograms). However, these data sets may not be representative of the conditions encountered during a survey in general, as there may be spatial or temporal changes in the extent to which organisms overlap or lower signal-to-noise ratios (SNR) in the survey measurements than in the case study. Notable exceptions to this are the studies of Jech and Michaels (2006), who compare the results of a multifrequency method with the results of multiple acoustic surveys, and Anderson et al. (2007), who provide a statistical measure of classification certainty.

Here, we evaluate the potential for use of multifrequency backscatter classification methods to improve interpretation of acoustic backscatter during long-standing acoustic-trawl surveys of walleye pollock (*Theragra chalcogramma*, hereafter referred to as pollock) in Alaska, USA. During these surveys, acoustic backscatter measurements are allocated to species based on the examination of echograms by experienced analysts and targeted trawling of aggregations with pelagic and bottom trawls (e.g., Karp and Walters 1994; Honkalehto et al. 2002). The approach taken here is empirical: we extract the relative frequency response from single-species aggregations sampled concurrently with acoustics and a trawl. We then develop a multifrequency method based on the normal deviate, or *Z* score, with which we compare unclassified acoustic measurements with the relative frequency response of trawl ground-truthed organisms. Finally, we assess the efficacy of the *Z*-score method by comparing pollock abundance estimates based on traditional species identification methods (i.e., trawling and echogram interpretation) with those based solely on the *Z*-score summary of relative frequency response during five large-scale surveys of the eastern Bering Sea shelf.

Materials and methods

Data collection

Concurrent acoustic measurements and trawl hauls were used to estimate the frequency response in areas where backscatter was dominated by a single taxonomic group. Measurements were made in the eastern Bering Sea and Gulf of Alaska during 13 echo integration-trawl surveys of pollock conducted from 2004 to 2007 aboard the NOAA ships *Miller Freeman* or *Oscar Dyson*. Acoustic backscatter was measured at 18, 38, 120, and 200 kHz with Simrad EK60 echosounders (for details, see De Robertis et al. 2008). A 70 kHz EK60 was installed on *Oscar Dyson* in 2007 and operated during one cruise. During some cruises in 2004 and 2005, a Simrad EK500 echosounder was used to measure 38 kHz backscatter. Echosounders were operated at power levels recommended to minimize harmonic distortion (Tichy et al. 2003; Korneliussen et al. 2008) and were calibrated at the start and end of each cruise. The spatial and physical characteristics of the measurements at different frequencies were as similar as possible following the recommendations of Korneliussen et al. (2008). All frequencies were transmitted simultaneously once per second with a 1.0 ms pulse length. The transducers were mounted as close together as possible on a centerboard extending 9 m below the waterline during all cruises except for one, where the 18 kHz transducer was mounted 11.3 m forward of the other transducers. In this case, a first-order correction was made by shifting the data in time to correspond as closely as possible to the spatial location sampled by the other frequencies (Korneliussen et al. 2008). Acoustic measurements from trawl locations were processed using Echoview (Higginbottom et al. 2000).

Aggregations detected by the echosounder were sampled with trawls to determine the species composition and fish size distribution during the surveys. The gear used depended on the suspected identity of the organisms: fish were sampled with a large midwater trawl or a bottom trawl, and smaller organisms (e.g., young-of-the-year fish, macrozooplankton) were sampled using smaller trawls (for a description of the trawl gear, see Honkalehto et al. 2002). Trawl depths were monitored using a netsonde to ensure that the trawl fished at the depth of the targeted aggregations. Catch compositions from 395 trawl hauls were screened to identify cases in which a single species or group (hereafter referred to as taxa) was likely to dominate the measured backscatter (Table 1).

Measurement of relative frequency response

The relative frequency response of aggregations dominated by a single taxon was measured in the immediate vicinity of the trawl path. Analysis was restricted to a region extending ± 0.5 nautical miles (nmi) from the fishing location and 20 m above to 20 m below the depths sampled by the trawl. A horizontal offset was applied to the analysis regions to account for the spatial separation of the trawl and the transducers (Zhou et al. 1994). Within these regions, the echograms were examined to ensure that there were no obvious signs of another taxon (e.g., fish schools did not overlap with diffuse backscatter likely to be from plankton). Backscatter passing these criteria was assigned to the taxon dominating the trawl catch for further analysis.

The relative frequency response was estimated from measurements of volume backscattering strength S_v (dB re 1 m^{-1}). Pairwise frequency differences (i.e., $\Delta S_{v,i-j} = S_{v,i} - S_{v,j}$, where i and j are indices denoting frequency in kHz) were computed to describe the relative frequency response. Subtracting logarithmic quantities in this manner is equivalent to computing the ratio in backscattering strength expressed in linear units. The relative frequency responses from aggregations where the trawl catch was dominated by pollock, capelin, euphausiids, jellyfish, myctophids, Pacific ocean perch, and eulachon (Table 1) were compiled into a database.

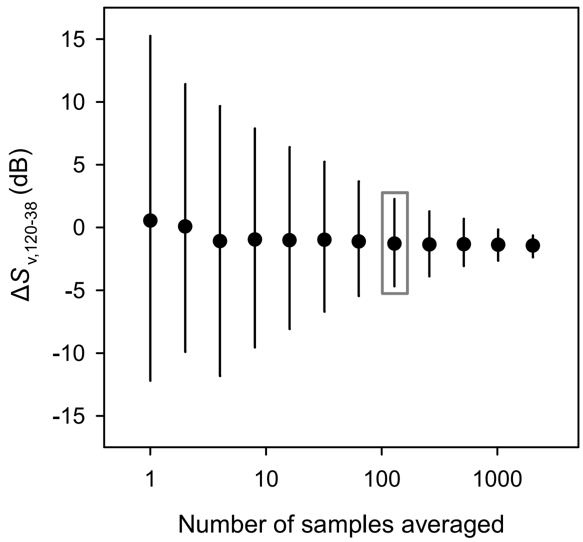
At each sampling location, S_v measurements from each taxon were averaged vertically and across multiple pings into larger analysis cells with the averaging conducted in the linear domain. The choice of analysis cell size represents a trade-off between decreasing variability in the observed relative frequency response and minimizing violations of the assumption that backscatter is dominated by a single organism. To guide selection of analysis cell size, the sensitivity of ΔS_v to spatial averaging was evaluated on five aggregations of pollock. $\Delta S_{v,120-38}$ at the original vertical resolution (0.19 m) was averaged at a series of spatial scales ranging from 1 to 2105 samples. The mean $\Delta S_{v,120-38}$ varied by only a small amount across the spatial scales, but the measurements became less variable as the spatial averaging increased (Fig. 1). The degree to which averaging combines backscatter from multiple species is more difficult to quantify, as this depends on the number of taxa present and their spatial overlap. To gain insight as to the degree to which increased averaging will result in combining backscatter from multiple taxa into an analysis cell, we reviewed a large number of echograms from previous surveys. We found that at moderate scales of averaging (~10s of metres), mixing of species was unlikely to be a concern. Based on these considerations, backscatter measurements collected within the boundaries of the trawl region were averaged vertically by 5 m and horizontally across 5 pings (hereafter referred to as

Table 1. Catch composition used to identify instances where a single taxon was likely to dominate backscatter at the trawl site.

			Catch composition (%)				Criteria for inclusion		
			Number		Weight				
Taxon	Taxonomic classification	Size (mm)	Minimum	Median	Minimum	Median	Minimum number	Minimum weight (kg)	No. of hauls
Capelin	*Mallotus villosus*	60–121 (SL)	97	98			5000		2
Eulachon	*Thaleichthys pacificus*	100–240 (FL)	85	95			4000		4
Euphausiids	Primarily *Thysanoessa* spp.	15–30 (TL)	100	100			1000		27
Jellyfish	*Chrysaora* spp. (96% by weight)	80–550 (BD)	94	100	98	100		20	5
Myctophids	*Stenobrachius leucopsarus* (81% by number)	40–111 (SL)	75	91			1000		3
Pacific ocean perch	*Sebastes alutus*	280–470 (FL)	81	99	84	99	200		5
Walleye pollock	*Theragra chalcogramma*								
Age 3+, winter		>300 (FL)	96	100	95	100	100		10
Age 3+, summer		>300 (FL)	96	99	96	99	100		25
Age 0		<100 (FL)	91	96			800		7
Age 1		160 (FL)	94	99			2000		11
Age 2		220 (FL)	93	97	87	94	6000		3

Note: Size measurements are approximate. The code in parentheses following size indicates the method of measurement, which differs by organism and includes jellyfish bell diameter (BD), fork length (FL), standard length (SL), defined as snout to hypural plate, and euphausiid total length (TL), defined as rostrum to telson.

Fig. 1. The relative frequency response at 38 kHz and 120 kHz as a function of averaging window. Results are for a subset of walleye pollock (*Theragra chalcogramma*) measurements. The circles show the mean $\Delta S_{v,120-38}$, and the error bars denote the 10th and 90th percentiles. The 5 ping × 5 m analysis cell resolution used in this study is indicated by the box.



a cell), which corresponds to a horizontal displacement of ~7 m at trawling speed (~3.3 kn) and ~24 m at survey speed (~11.5 kn).

The relative frequency response for each analysis cell was computed as the six unique pairwise frequency differences ($\Delta S_{v,i-j}$) possible from the combination of four frequencies. Only a small subset of measurements included 70 kHz, but we computed $\Delta S_{v,38-70}$ in those cases to identify whether large differences were present at this frequency. To minimize the effects of background noise, we estimated the signal-to-noise ratio (SNR) at each frequency using the method of De Robertis and Higginbottom (2007) and excluded all $\Delta S_{v,i-j}$ with SNR < 10 dB for either frequency in the pair. In addition, only cells with $S_v > -70$ dB for at least one of the frequencies were used in further analyses.

The aggregations varied in size, and this resulted in a variable number of analysis cells among sampling locations. In some instances, some frequencies did not meet the SNR criteria. To accommodate these characteristics while maximizing the amount of data included in the summary of observed frequency differences, we developed a stratified bootstrap scheme that weighted the contributions of each independent sample (i.e., trawl locations) equally and allowed for the inclusion of cases in which not all frequencies were above the SNR threshold. For each of 10 000 iterations, 50 samples from each trawl haul with >50 cells meeting the S_v threshold and SNR criteria at a given frequency pair were randomly selected with replacement. The mean and standard deviation (SD) of $\Delta S_{v,i-j}$ for each iteration were stored, and the mean and SD of $\Delta S_{v,i-j}$ for a given taxon were computed as the mean of each quantity over all iterations. Statistics for each frequency pair were compiled independently; a given cell can thus contribute to some or all of the pairwise frequency differences, depending on SNR.

Comparison of measurements when trawling and surveying

The relative frequency response was measured with the vessel trawling to maximize the probability of correct ground-truthing. However, acoustic backscatter strength (but not relative frequency response) measured while trawling has been observed to differ from that made while a vessel is surveying due to differential behavioral reactions to the vessel (Barange and Hampton 1994; De Robertis and Wilson 2006). We therefore compared the relative frequency response measured while trawling (~3.3 kn) with that measured at survey speed (~11.5 kn) immediately before or after the trawl to test for differences.

Analysis of a subset of the pollock (n = 29) and euphausiid (n = 14) trawl locations used in this analysis revealed that relative frequency responses made while trawling and surveying were similar. The difference in relative frequency response (i.e., $\Delta S_{v,\mathrm{survey}} - \Delta S_{v,\mathrm{trawl}}$) averaged over all pairwise combinations of 18, 38, 120, and 200 kHz was –0.16 dB for euphausiids and 0.15 dB for pollock when the vessel was at survey speed compared with trawling. Given that the relative frequency response when trawling and surveying was not significantly different (t test, $P > 0.05$ in both cases) and that frequency differences of this magnitude were small compared with the variability among sampling locations (see below), we concluded that relative frequency responses measured while the vessels are trawling are equivalent to those observed when surveying.

The *Z* score as a summary of relative frequency response

We developed a method to classify backscatter measurements based on relative frequency response. The approach is to summarize the relative frequency response and evaluate the degree to which an unclassified measurement corresponds to the relative frequency response of a known class of scatters in the database of ground-truthed frequency differences. The basis of the method is that the distribution of the cell-averaged ΔS_v approximates a normal distribution (Fig. 2). If ΔS_v is normally distributed, the deviations of an observed frequency difference $\Delta S_{v,k,l}$ from frequency pair k observed in sample l can be summarized by the normal deviate or Z score:

$$(1) \qquad Z_{k,l,m} = \frac{\Delta S_{v,k,l} - \mu_{k,m}}{\sigma_{k,m}}$$

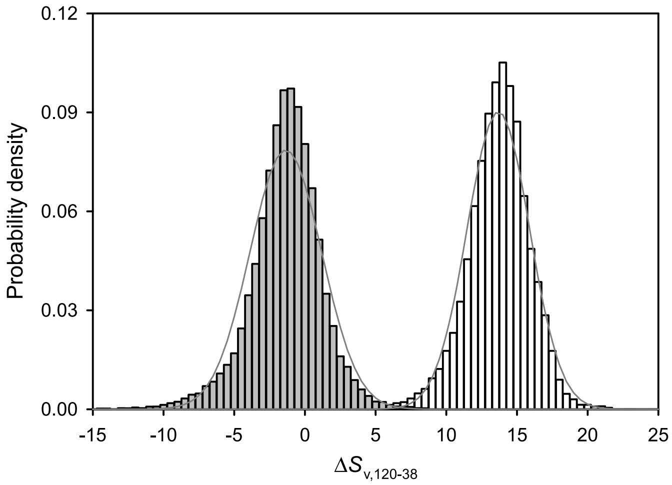
where $\mu_{k,m}$ and $\sigma_{k,m}$ are the sample mean and SD of ΔS_v estimated from the ground-truthed relative frequency response measurements for frequency pair k and taxon m.

The deviation of the analysis cell from the Z score expected for group m is summarized over all frequency pairs as follows:

$$(2) \qquad \overline{Z}_{l,m} = \frac{\sum_{k=1}^{n} |Z_{k,l,m}|}{n}$$

where $|Z_{k,l,m}|$ is the absolute value of eq. 1, and n is the number of frequency pairs used. $\overline{Z}_{l,m}$ has units of SDs from the expectation. For example, $\overline{Z}_{l,m}$ = 1 means that the average

Fig. 2. Histogram of $\Delta S_{v,120-38}$ for 5 ping × 5 m analysis cells identified as euphausiids (open bars) and walleye pollock (*Theragra chalcogramma*; shaded bars) based on trawl catches. The corresponding probability density for a normal distribution with the observed mean and variance is given by the gray line.



pairwise frequency difference observed for cell *l* differs by 1 SD from the expectation for taxon *m*.

S_v measurements were evaluated for consistency with taxon *m* on the basis of $\overline{Z}_m$. A threshold of 2 SDs ($Z_{thresh} \leq 2$) was applied to identify candidate cells that are consistent with the frequency difference observed for taxon *m*:

$$(3) \quad \text{if} \quad \overline{Z}_{l,m} \leq Z_{thresh}, \quad \text{then} \quad S_{v,l,m} = S_{v,l}, \quad \text{otherwise,} \quad S_{v,l,m} = -999$$

An S_v of –999 dB is used if the measurement falls below the threshold value because –999 approaches zero in the linear domain. In the case where a cell produced a $\overline{Z}_{l,m}$ consistent with multiple taxa (i.e., $\overline{Z}_{l,m} < Z_{thresh}$ for multiple *m*), the sample was assigned to the taxon with the lowest $\overline{Z}$. This criterion required that the backscatter in a given cell be assigned in its entirety to the single most likely class.

The average *Z* score of the relative frequency response provides a convenient measure of confidence in the cell classification, expressed in units of SDs relative to the measured relative frequency response for taxon *m*. For example, identification of a cell as consistent with pollock is more likely to be correct for a $\overline{Z}_{pollock}$ of 0.5 compared with one with a $\overline{Z}_{pollock}$ of 1.9, as 1.9 SDs is further in the tail of a normal distribution. Although we performed assignment to taxonomic classes (e.g., pollock vs. euphausiids) as a binary decision in this application, we used $\overline{Z}$ to interpret the results of the classification during postprocessing of acoustic surveys.

The availability of software to conveniently implement multifrequency analyses is a practical concern in many applications such as fish abundance surveys. We implemented the techniques described here using Echoview, a commercially available software application (Higginbottom et al. 2000). This allows processing to be conducted in real time during an acoustic survey. Instructions for implementing the *Z*-score method using this software are available from the authors.

Comparison of multifrequency and traditional species classification

We compared species classifications using the *Z*-score method with those made with traditional species identification methodology during five large-scale surveys of pollock in the eastern Bering Sea. During these surveys, experienced analysts partitioned the echo integral and allocated a portion of the backscatter to pollock by "scrutinizing" the echograms following traditional practice (Simmonds and MacLennan 2005; Jech and Michaels 2006). This involves using a combination of subjective examination of echo traces, prior knowledge, and ~100 targeted trawl hauls (Karp and Walters 1994; Honkalehto et al. 2002). The species classification in this survey is considered to be relatively straightforward (Honkalehto et al. 2002; De Robertis et al. 2008), and results are likely to be reliable compared with those of other situations. Pollock abundance estimates made with the traditional method provide a basis for comparison with the multifrequency estimates. The key difference between these methods is that the multifrequency method is based solely on the observed relative frequency response, whereas the traditional method relies primarily on the subjective judgment of the analyst and trawl sampling.

Estimates of pollock backscatter from surveys in 2004, 2006, 2007, 2008, and 2009 (no survey in 2005) using the traditional identification method used during the survey were compared with abundance estimates in which pollock identification was based solely on the *Z*-score method. Prior to *Z*-score classification, S_v at 18, 38, 120, and 200 kHz were averaged into 5 ping (horizontal) × 5 m (vertical) analysis cells to a maximum depth of 500 m. Backscatter from the transmit pulse and bottom echo were excluded prior to averaging. The SNR of S_v data was considered before *Z*-score classification to minimize biases introduced by background noise (cf. De Robertis and Higginbottom 2007). In practice, only 120 and 200 kHz were noise-limited in the depth range analyzed. If SNR at 200 kHz was <10 dB, the cell was classified using data from the remaining three frequencies. If SNR at 120 kHz was also <10 dB, the cell was excluded from further analysis. An exception to this procedure occurred in the analysis of the 2006 survey, when noise due to ringing of the transmit pulse from the 18 kHz transducer caused low SNR at short range. The ringing, which was very temporally consistent, was subtracted from the 2006 18 kHz S_v data, and the corrected data were used in subsequent computations.

For each analysis cell, we computed $\overline{Z}_m$ relative to four taxa (pollock, euphausiids, jellyfish, and myctophids). Only results for pollock will be presented here as they can be directly compared with the survey results. However, the other classes are relevant to this analysis in that a cell was allocated to pollock only when $\overline{Z}_{pollock}$ was lower than $\overline{Z}_m$ relative to the other taxa. Pollock backscatter identified using the traditional survey method and the *Z*-score method was integrated at 38 kHz using an S_v integration threshold of –70 dB re 1 m^{-1}. Nautical area backscattering coefficients (s_A, $m^2 \cdot nmi^{-2}$) were allocated to pollock using the survey methodology and the *Z*-score method and averaged into 0.5 nmi long × 20 m deep elementary distance sampling units (EDSUs) for comparison. In addition, the mean

$\overline{Z}_{pollock}$ for all cells allocated to pollock in each EDSU was computed.

The 2004 survey presented a special case: in this survey, age-0 pollock contributed a sizeable amount of backscatter at 38 kHz. The contribution from these fish was not quantified during the survey, which enumerated only age-1 and older pollock (Honkalehto et al. 2005). However, the relative frequency response of scattering from age-0 pollock is similar to that of older pollock (see below). The *Z*-score estimate of pollock abundance, which includes contributions from all age groups of pollock, was substantially higher than the survey estimate in 2004. This discrepancy is attributable, at least in part, to the exclusion of age-0 pollock from the survey estimate. To avoid this effect, we limited the comparison in 2004 to areas in which backscatter was identified as age-1 and older pollock using the traditional method.

Results

Observed relative frequency response

The relative frequency response of pollock was similar across age classes (Fig. 3), despite the large variation in body size with age (Table 1). Furthermore, the relative frequency response of adult fish in the summer (when the fish were not in spawning condition) overlapped broadly with that of adult fish in the winter when many of the fish were in spawning condition (Fig. 3). This similarity indicates that it will be difficult to separate pollock age groups or maturity states on the basis of relative frequency response at the frequencies used in this study. Because of the broad overlap in relative frequency response, all pollock ΔS_v samples were pooled together in subsequent analyses.

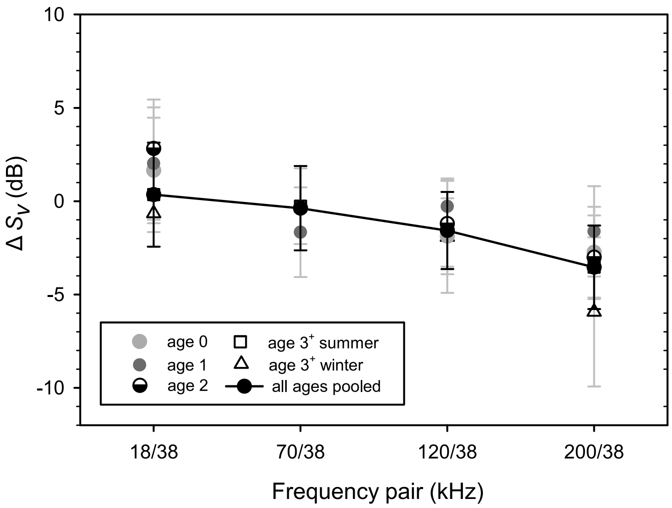
The compiled estimates of relative frequency response (Fig. 4; Table 2) indicated that in some cases, ΔS_v differs substantially among taxa. Pollock (n = 56), Pacific ocean perch (n = 5), and capelin (n = 2), all of which have swimbladders, have relatively flat relative frequency responses that overlap broadly, which will make separation of these species on the basis of ΔS_v difficult. Observations of eulachon (n = 4), which do not have a swimbladder, suggest a similar relative frequency response to pollock, Pacific ocean perch, and capelin, except for increased backscatter at 120 kHz relative to the other frequencies. Scattering layers dominated by myctophids (n = 3) exhibited elevated backscattering at 38 kHz relative to 18 kHz and 120 kHz. Locations (n = 5) where near-surface trawl catches were dominated by large *Chrysaora* spp. jellyfish (primarily *Chrysaora melanaster*) exhibited substantially stronger backscatter at 18 kHz than at the other frequencies. Adult euphausiids (n = 27) differed substantially from the other taxa sampled; their backscatter strength increased strongly with frequency.

Degree of overlap in *Z* scores in key taxa

To assess the potential to discriminate among taxa on the basis of *Z* score, we evaluated the degree to which *Z* scores ≤ 2 overlap among taxa. We computed the mean fraction of analysis cells for each taxon that met the criterion of $\overline{Z}_m \leq 2$, with the *Z* score based on the observed values for each taxon, m. The calculations for myctophids exclude 200 kHz as there was insufficient SNR to warrant inclusion of 200 kHz measurements. The results (Table 3) suggest that $\overline{Z}$, the mean normal deviate of all pairwise ΔS_v, may be useful to distinguish among backscatter from some of the taxa considered. As seen by the diagonal in the matrix, most of the cells in taxon m were consistent with the $\overline{Z}_m \leq 2$ criterion. This suggests that the criterion will recover most of the backscatter in a taxon, which is to be expected as the test and training data sets are the same. For example, 96% of the pollock analysis cells passed the $\overline{Z}_{pollock} \leq 2$ criterion. This behavior also indicates that in most cases, the distribution of ΔS_v is reasonably well approximated by a normal distribution.

The entries along a row in Table 3 indicate the degree to which $\overline{Z}$ overlaps among taxa and provides an indication of the potential for successful discrimination among taxa. For example, the first and last rows indicate that euphausiids and myctophids can likely be distinguished from the other taxa on the basis of $\overline{Z}$. In contrast, pollock, Pacific ocean perch, eulachon, and capelin have broad overlap in $\overline{Z}$, which indicates that the potential to discriminate among these taxa on the basis of $\overline{Z}$ is low.

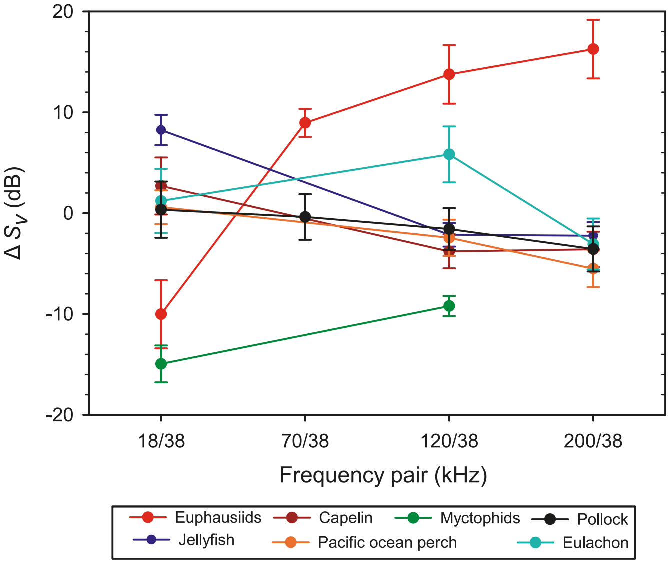
Fig. 3. Observed frequency response for walleye pollock (*Theragra chalcogramma*). Means and standard deviations of ΔS_v are shown for five groups of pollock and for all groups pooled. The solid error bars correspond to the pooled values and the shaded error bars correspond to the age-specific values.



Comparison of survey and multifrequency estimates

Initial visual assessment of synthetic echograms indicated that the multifrequency identification of pollock was consistent with the expectations based on experience in interpreting echograms during the survey (e.g., Fig. 5). In these surveys, few adult pollock are found at depths <40 m, but there is often a near-surface scattering layer of unknown composition (e.g., Fig. 5*a*). Application of the $\overline{Z}_{pollock}$ criteria to the multifrequency backscatter recorded during the survey (Figs. 5*b* and 5*c*) generally resulted in synthetic echograms that were consistent with the expectation of pollock distribution based on trawling and interpretation of the school morphology on echograms (i.e., compare Figs. 5*a* and 5*d*).

Fig. 4. Means and standard deviations of ΔS_v estimated for groups of fish and invertebrates commonly encountered during acoustic-trawl surveys in the Gulf of Alaska and eastern Bering Sea.



When averaged over the larger 0.5 nmi × 20 m EDSU scale, analysis cells consistent with pollock (i.e., $Z_{\text{thresh}} \leq 2$) at >40 m depth generally exhibited a $\overline{Z}_{\text{pollock}} < 1.5$, particularly when acoustic density was high (Fig. 6*a*). However, this was not always the case in the <40 m stratum: in some years, areas were observed where substantial backscatter was allocated to pollock, but $\overline{Z}_{\text{pollock}}$ was often >1.5 but below the $Z_{\text{thresh}} \leq 2$ criterion (Fig. 6*b*, the red (i.e., high $\overline{Z}_{\text{pollock}}$) and large diameter (high s_A) symbols observed in the offshore part of the middle of the survey grid). The difference between the shallow and deeper $\overline{Z}$ illustrates the utility of $\overline{Z}$ as a quality control measure. The shallow backscatter identified as pollock using the $\overline{Z}_{\text{pollock}} < 2$ criterion at the analysis cell scale is questionable, as the spatially averaged response is not as consistent with the measured relative frequency response of pollock as the deeper backscatter (i.e., $\overline{Z}_{\text{pollock}}$ is high for the shallower strata). These dubious identifications were largely eliminated with the application of the additional criterion of $\overline{Z}_{\text{pollock}} < 1.5$ at the EDSU scale.

Multifrequency pollock identification yielded pollock abundance and distribution similar to that obtained with the traditional identification method (Fig. 7). On a transect basis, pollock backscatter resulting from the methods agreed well in almost all cases (Fig. 8). The high coefficient of determination (r^2) indicates that, depending on the year, 93%–99% of the variance in survey pollock abundance was explained by the *Z*-score method. The traditional method typically identified slightly more pollock backscatter than the *Z*-score method. This is expected as the multifrequency criteria used here did not encompass the entire range of relative frequency response observed for pollock (e.g., Table 3). There was also an outlier: for one of the 2007 transects, the multifrequency estimate was much higher than the survey estimate. This was in an area where large aggregations of Pacific ocean perch were identified during the survey. This species has a similar relative frequency response to pollock

Table 2. Summary of ΔS_v statistics measured by taxon.

Taxon	No. of hauls	No. of cells	$\Delta S_{v,18\text{–}38}$	$\Delta S_{v,120\text{–}38}$	$\Delta S_{v,200\text{–}38}$	$\Delta S_{v,120\text{–}18}$	$\Delta S_{v,200\text{–}18}$	$\Delta S_{v,200\text{–}120}$
Euphausiids	28	11 340	−10.0±3.4	13.8±2.9	16.3±2.9	24.9±4.8	27.3±4.4	2.3±1.4
Jellyfish	5	1 755	8.2±1.5	−2.1±1.2	−2.2±1.3	−10.4±1.7	−11.0±1.8	−0.1±1.1
Walleye pollock	56	18 303	0.4±2.8	−1.6±2.1	−3.5±2.2	−1.9±2.9	−3.9±2.7	−2.0±1.8
Capelin	2	942	2.7±2.8	−3.8±1.7	−3.6±1.8	−6.5±3.2	−5.8±2.2	−0.4±1.5
Pacific ocean perch	5	2 538	0.6±1.7	−2.4±1.8	−5.5±1.8	−3.0±2.0	−5.2±1.9	−2.4±1.1
Eulachon	4	2 795	1.2±3.2	5.8±2.8	−3.1±2.5	4.6±3.3	−3.3±2.1	−2.1±1.8
Myctophids	3	1 048	−15.0±1.8	−9.2±1.0	n/a	5.9±1.6	n/a	n/a

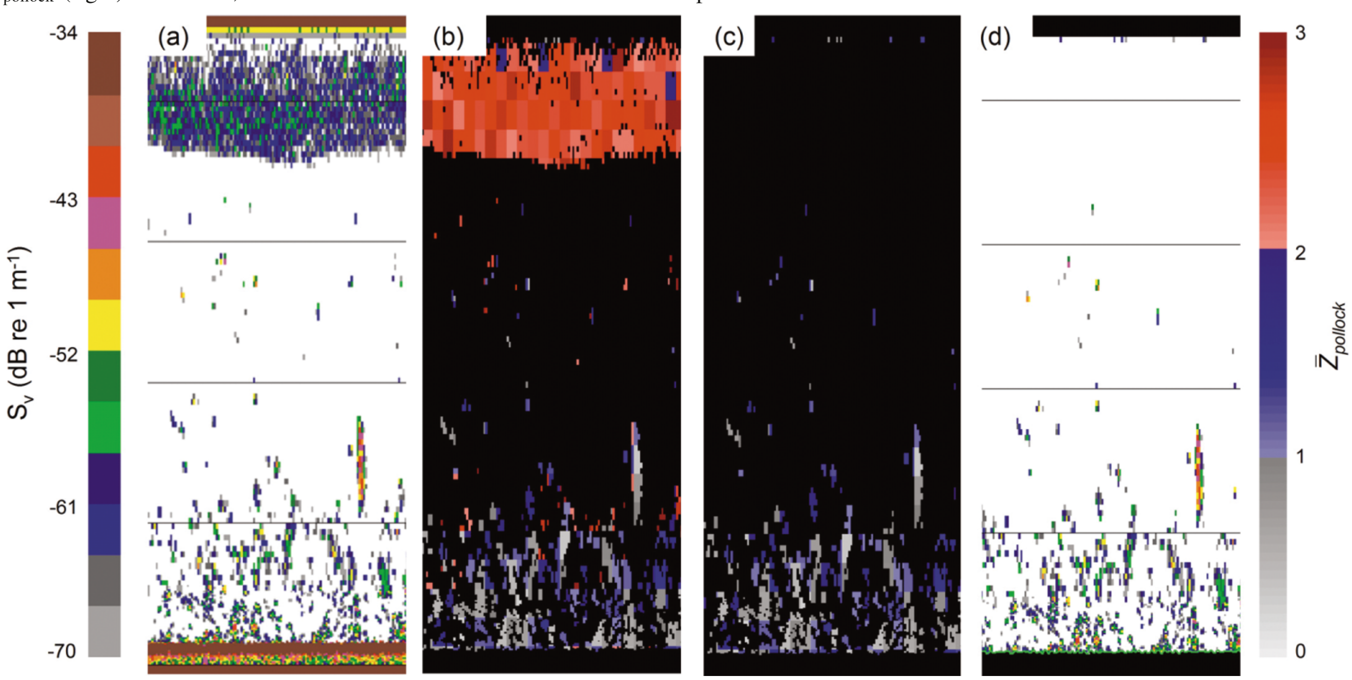
Note: At each frequency pair, the mean ± 1 standard deviation of the frequency difference for 5 ping × 5 m analysis cells is estimated from the bootstrap analysis. The number of hauls and the number of cells refer to the frequency pair with the most available data points. Some frequency pairs are based on fewer observations, as the number of cases that meet the analysis criteria is frequency and class dependent. n/a, no valid measurements are available.

Table 3. *Z*-score technique applied to the ground-truthed in situ multifrequency measurements.

	Data set						
Z score	Euphausiids	Jellyfish	Capelin	Pacific ocean perch	Eulachon	Walleye pollock	Myctophids
$Z_{Euphausiid}$	0.99	0.00	0.00	0.00	0.00	0.00	0.00
$Z_{Jellyfish}$	0.00	0.75	0.68	0.05	0.03	0.48	0.00
$Z_{Capelin}$	0.00	0.29	0.99	0.94	0.85	0.99	0.00
$Z_{Pacific\ ocean\ perch}$	0.00	0.01	0.78	0.78	0.76	0.78	0.00
$Z_{Eulachon}$	0.00	0.00	0.79	0.83	0.94	0.94	0.00
$Z_{Pollock}$	0.00	0.10	0.90	0.85	0.90	0.96	0.00
$Z_{Myctophid}$	0.00	0.00	0.00	0.00	0.00	0.00	0.98

Note: Each row shows the taxon, *m*, used as a reference for the *Z* score, and the columns indicate the trawl ground-truthed analysis cells for which the *Z* score was computed. The numbers indicate the fraction of analysis cells where $\overline{Z}_m \leq 2$. Rows indicate the fraction of analysis cells for a given taxon that is consistent with the *Z* score listed in each row. For example, the first row shows the fraction of cells for which $\overline{Z}_{euphausid} \leq 2$. Columns show the fraction of analysis cells for a given taxon that is consistent with $\overline{Z} \leq 2$ for the *Z* score listed in each row. Results represent the average for all haul locations.

Fig. 5. Example of implementation of the *Z*-score method in the eastern Bering Sea: (*a*) original 38 kHz echogram showing backscatter from demersal and pelagic walleye pollock (*Theragra chalcogramma*) and a near-surface layer of unknown composition; (*b*) synthetic echogram of $\overline{Z}_{pollock}$ summarizing the frequency responses at 18, 38, 120, and 200 kHz; results are shown for all samples above the integration threshold of –70 dB at 38 kHz; (*c*) as in (*b*), but implementing the $\overline{Z}_{pollock} < 2$ and $\overline{Z}_{pollock} = \min(\overline{Z}_m)$ criteria used to identify pollock; (*d*) 38 kHz echogram showing all samples with a frequency response consistent with that of pollock. Color scales for 38 kHz S_v (left) and $\overline{Z}_{pollock}$ (right) are shown, and black horizontal lines demarcate 25 m depth intervals.

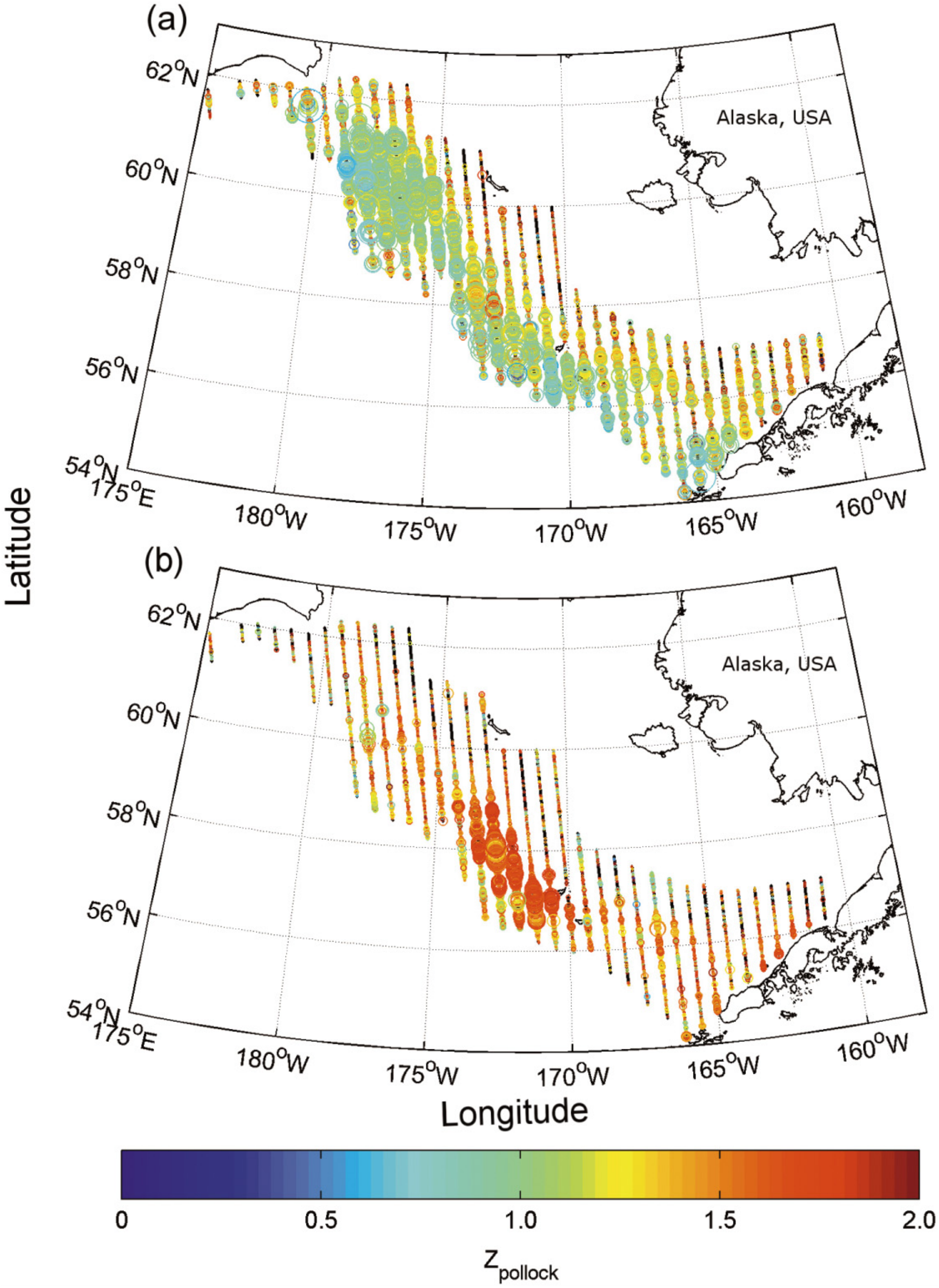


(Fig. 4; Table 3), and these aggregations were misidentified as pollock by the multifrequency method. This outlier illustrates one of the limitations of the method: when the relative frequency response of two taxa is similar, the misidentification rate will be high. The comparatively good agreement for the other transects indicates that this is not a widespread problem for this survey area.

On a survey-wide basis, there was good agreement between the survey and multifrequency classification results (Fig. 9). In 2004, the *Z*-score method identified less pollock s_A than the traditional method. However, in 2004, the *Z*-score method was applied only to cells classified as pollock using the traditional method due to the influence of age-0 pollock, which were not quantified by the traditional method. This constrained the *Z*-score method to be lower than or equal to the traditional estimate. In the other years, pollock abundance from the survey and multifrequency pollock were closer. The impact of applying the additional $\overline{Z}_{pollock} < 1.5$ criterion at the EDSU level differed among years in depth zones. In 2007 and 2008, there were substantial amounts of backscatter in the near-surface strata with $\overline{Z}_{pollock} > 1.5$ (e.g., Fig. 9), and the EDSU level thresholding made the survey and pollock abundances more comparable. In contrast, this step had little impact in 2004, 2006, and 2009.

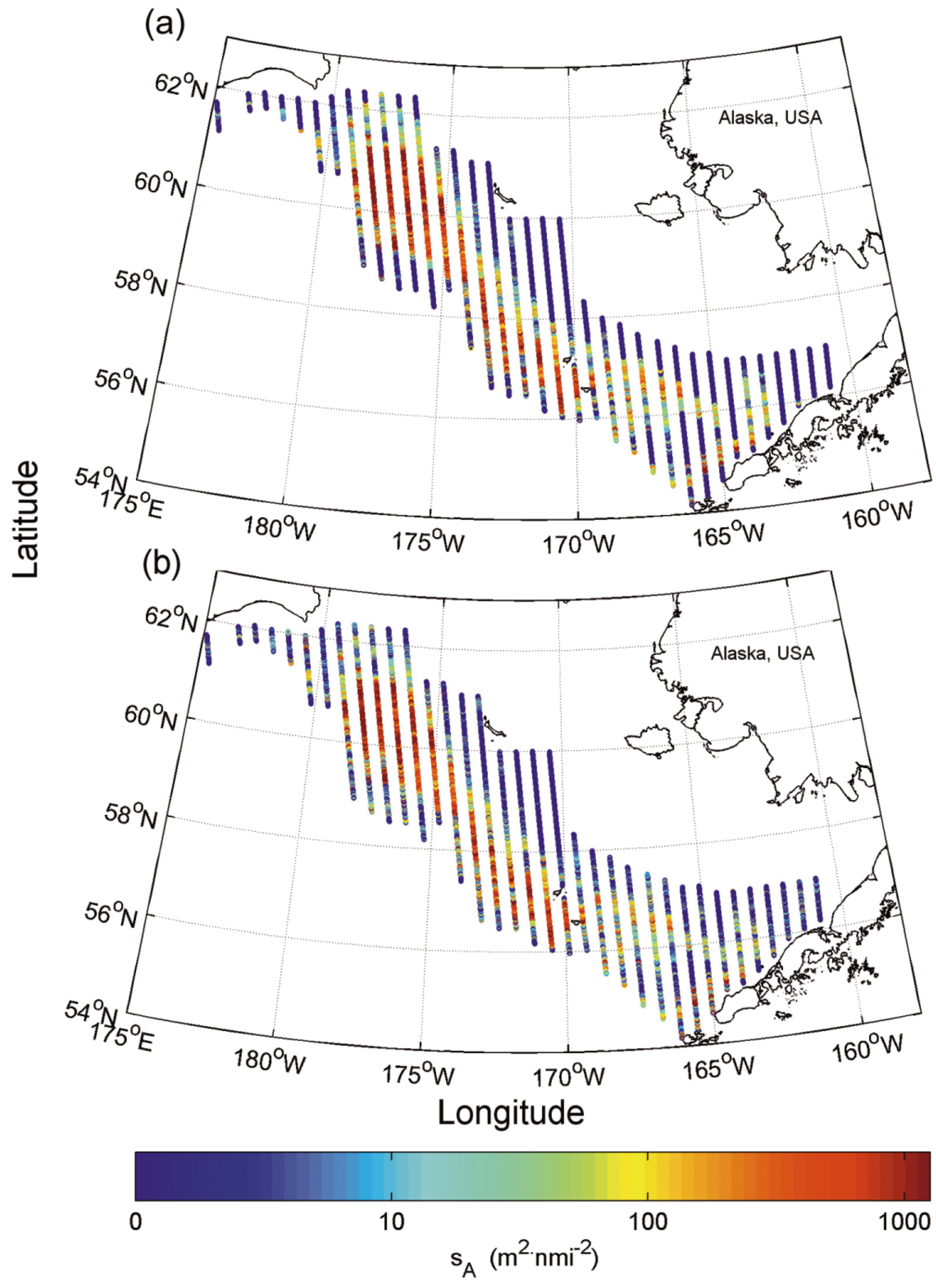
During the surveys, the multifrequency method associated a variable amount of the total 38 kHz echo integral

Fig. 6. Map of Z score for backscatter consistent with walleye pollock (*Theragra chalcogramma*) (i.e., $\overline{Z}_{\text{pollock}} < 2$). Results are shown for the strata (*a*) > 40 m and (*b*) < 40 m. $\overline{Z}_{\text{pollock}}$ was used to identify areas where the multifrequency identification is questionable. Symbol color represents $\overline{Z}_{\text{pollock}}$, and symbol size is proportional to acoustic backscatter. Questionable identifications (red symbols) in areas where high backscatter was observed (large symbols) are much more prevalent in the near-surface zone. Results are shown for 2007 as this survey has a large number of questionable identifications in near-surface waters.



with pollock. A range of 6% to 77% of 38 kHz backscatter along a transect was associated with pollock by the Z-score method. On a survey basis, 30%–58% of total 38 kHz backscatter was classified as pollock, depending on the year. Inspection of $\overline{Z}_{\text{pollock}}$ indicates that the confidence in the multifrequency identifications was fairly consistent among transects and years (Fig. 10). The overall range of $\overline{Z}_{\text{pollock}}$ was fairly narrow, suggesting that there are no strong spatial gradients in the degree to which the classified backscatter conforms to the measured relative frequency response of pollock. Similarly, $\overline{Z}_{\text{pollock}}$ in 2006–2009 overlapped broadly (Fig. 10), which indicates that there was a similar degree of confidence in the classification for these data sets. In contrast, $\overline{Z}_{\text{pollock}}$ was slightly lower (~0.1 SDs) in 2004 (Fig. 10). This is consistent with differences in application of the method: in 2004, the Z-score method was applied only to those cells identified as pollock during the survey. This presumably resulted in fewer misidentifications, which is reflected in the small, but consistent, difference in $\overline{Z}_{\text{pollock}}$.

Fig. 7. Estimates of walleye pollock (*Theragra chalcogramma*) acoustic backscatter based on allocating backscatter using (*a*) traditional survey methodology and (*b*) the *Z*-score multifrequency method. Results are for the 2007 echo integration-trawl survey of pollock in the eastern Bering Sea.



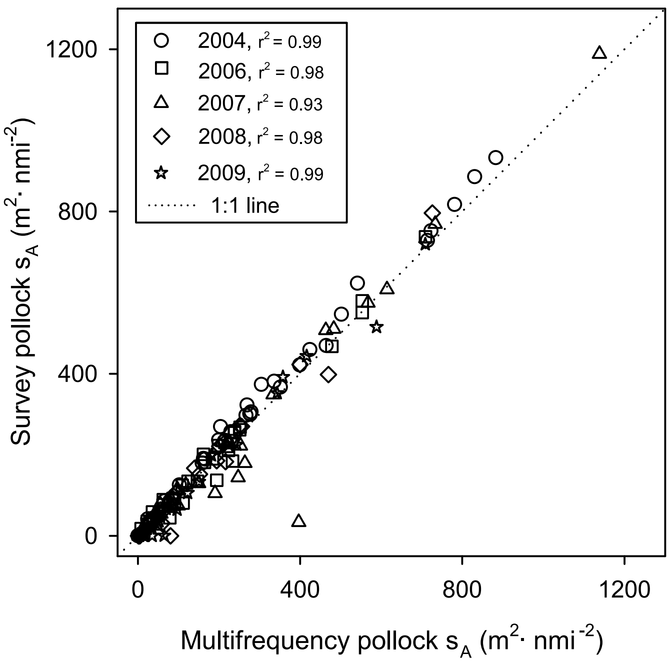
Discussion

Empirical measurement of relative frequency response

Our measurements indicate that use of multiple discrete frequencies in the range commonly used in fisheries acoustics (18–200 kHz) has the potential to improve the ability to discriminate among some of the primary acoustic scatterers encountered during surveys in the eastern Bering Sea and the Gulf of Alaska. Overall, our observations are consistent with previous studies that have demonstrated that one can potentially distinguish among organisms with contrasting acoustic properties, e.g., fish with or without swimbladders, resonant scatterers, and zooplankton, on the basis of relative frequency response (e.g., Korneliussen and Ona 2002; Lavery et al. 2007).

Fish with gas-filled swimbladders (e.g., pollock, capelin, Pacific ocean perch) exhibited a similar relative frequency response. This similarity in relative frequency response is likely because the contribution from the swimbladder is the dominant source of the backscattered signal from these fish (Foote 1980, 1985), which suggests that it will be difficult to discriminate among these species on the basis of multifre-

Fig. 8. Comparison of walleye pollock (*Theragra chalcogramma*) backscatter based on multifrequency identification of pollock with pollock backscatter identified in the acoustic survey for five surveys between 2004 and 2009. Each point represents a transect mean. The figure legend lists the r^2 resulting from linear regression of the multifrequency and survey pollock abundances in each year. The transects can be seen in Fig. 7.



quency measurements at these frequencies. Although only two observations were made, capelin appeared to have a relatively stronger backscatter at 18 kHz than at higher frequencies, which is consistent with previous observations of capelin (Fernandes et al. 2006).

The other taxa considered appeared to have relative frequency responses that differed from those of the fish with swimbladders. Eulachon, which lack swimbladders, exhibited a tendency toward a stronger backscattering strength at 120 kHz than at 18, 38, and 200 kHz, which is consistent with the theoretical predictions of Gauthier and Horne (2004). Euphausiids exhibited a strong increase in backscatter strength with frequency, which is distinct from the other taxa considered and consistent with previous studies (Madureira et al. 1993; Stanton et al. 1996; McKelvey and Wilson 2006). Areas where net catches were dominated by *C. melanaster* jellyfish exhibited disproportionately stronger backscatter at 18 kHz. This is consistent with previous target strength measurements of the jellyfish *Chrysaora hysoscella* at 18–200 kHz, in which a strong peak at 18 kHz relative to other frequencies has been observed (Brierley et al. 2004). Scattering layers in which net catches were dominated by myctophids exhibited substantially stronger backscatter at 38 kHz than at 18 kHz or 120 kHz, which may be caused by resonance of small gas-filled swimbladders at 38 kHz (Kloser et al. 2002). The primary species of myctophids in our study area (Sinclair and Stabeno 2002) have swimbladders, although some species (including *Stenobrachius leucopsarus*, the numerically dominant species) undergo an ontogenetic shift from a primarily gas-filled swimbladder to a lipid-filled swimbladder (Butler and Pearcy 1972). These results indicate that it should be possible to distinguish backscatter from swimbladdered fishes and euphausiids based on the large difference in observed relative frequency response as has been previously reported (e.g., Miyashita et al. 1997; McKelvey and Wilson 2006), and likely jellyfish and myctophids as well.

The relative frequency response for pollock was similar for different age and size groups: the relative frequency response of an age-0 pollock was similar to that of an adult weighing more than 1000 times as much. This lack of size dependence in the relative frequency response is not surprising, as the frequencies used do not span the transition from Rayleigh to geometric scattering for pollock. As a consequence, it will be difficult to invert pollock size distributions from measurements of relative frequency response with the frequencies commonly used in fisheries acoustics (Holliday and Pieper 1995; Horne and Jech 1999). In addition, the relative frequency response was similar across a wide range of pollock depths and aggregation types, and during the summer when there is no spawning compared with in the winter when pollock are in spawning condition. This indicates that although target strength is strongly dependent on factors such as fish size, maturity state, behavior, and depth distribution (e.g., Foote 1985; Ona 2003), these factors did not have a large impact on the relative frequency response of pollock. This is likely because the relative frequency response is based on the difference in target strength among frequencies (Gauthier and Horne 2004), which will often be lower than the absolute change in target strength at a single frequency. However, the effect of depth on relative frequency response is likely to be important for resonant scatters such as myctophids, which undergo swimbladder compression during vertical migrations, thus changing the resonant frequency of the swimbladder (Godø et al. 2009).

This study relied on in situ measurement of the relative frequency response of trawl ground-truthed backscatter (e.g., Brierley et al. 1998; McKelvey and Wilson 2006; Korneliussen et al. 2009). However, the dominant organism in the trawl catch may not be the most abundant organism in the environment, as the trawls capture organisms selectively (Nakashima 1990; McClatchie et al. 2000; Wiebe and Benfield 2003) and the trawl and acoustic samples are not completely spatially and temporally coincident. In addition, organisms differ substantially in their scattering properties, and the most abundant organisms are not necessarily the primary contributors to acoustic backscatter (Stanton et al. 1996). Thus, there is some uncertainty in our ground-truthed identifications of the dominant scatters. This is particularly the case for the myctophid and jellyfish identifications, which are more poorly constrained than those of the other taxa. Although unlikely, it is possible that much of the backscatter measured in these trawls was, in fact, backscattering from taxa that were not observed in the catch, for example, siphonophores with gas-filled pneumatophores (Stanton et al. 1996; Lavery et al. 2007). In addition, for some taxa, there were few replicate samples as they rarely dominated the trawl catch. Thus, we have the highest certainty for the measurements on pollock and euphausiids for which the

Fig. 9. Multifrequency walleye pollock (*Theragra chalcogramma*) abundance estimate relative to the traditional survey estimate as a function of identification method for samples > 40 m and < 40 m. For each survey, three bars are shown, representing pollock identified by the traditional technique (survey), pollock identified as analysis cells with $\overline{Z}_{pollock} \leq 2$ (Z_{cell}), and pollock identified as EDSUs with an average $\overline{Z}_{pollock} \leq 1.5$ (Z_{EDSU}). The dotted line indicates agreement with the survey estimate, as show by the first bar in each group. The 2004 comparisons differ from the other cases in that the comparisons are limited to cells assigned to pollock by the survey due to the presence of large numbers of age-0 pollock that were not included in the survey estimate in 2004.

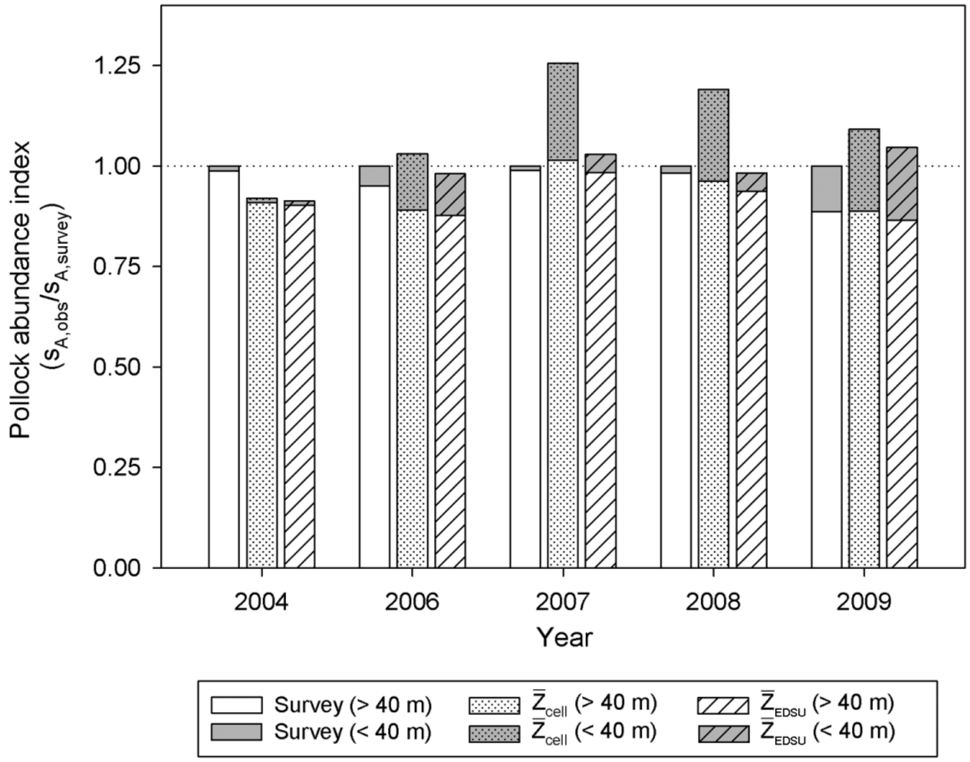
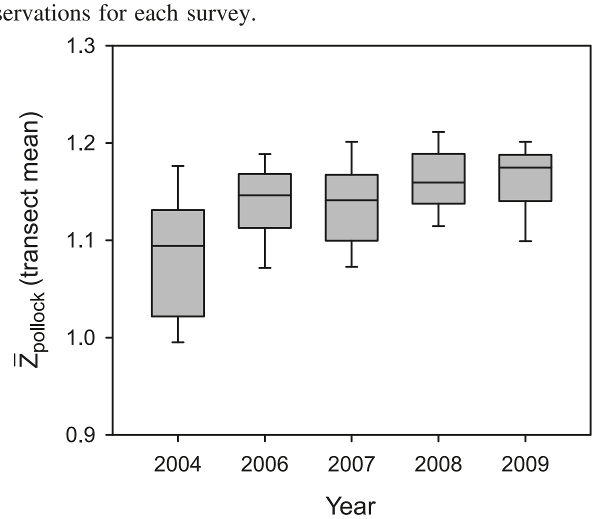


Fig. 10. Box plots of transect mean $\overline{Z}_{pollock}$ of analysis cells classified as walleye pollock (*Theragra chalcogramma*). The horizontal lines within the boxes represent the median values. The box plots demarcate the 10th, 25th, 50th, 75th, and 90th percentiles of the observations for each survey.



most measurements were available, and the identity of the acoustic scatters is less ambiguous. Because of these uncertainties, and the limited replicate samples, the results for other taxa should be treated as provisional until substantiated in future studies.

Z-score approach to multifrequency classification

We developed a method to classify backscatter by comparing the mean normal deviate of the relative frequency response in an acoustic measurement with the empirically measured relative frequency response for specific taxa. We chose the *Z* score to summarize the relative frequency response. This method assumes that ΔS_v is approximately normally distributed or, at least, that most of the distribution is encompassed by ±2 SDs of the mean. We purposely chose a method that could be implemented in real time during existing acoustic surveys using commercially available software for practical regions. Although this constrained us from applying more sophisticated classification techniques (e.g., Zakharia et al. 1996; Anderson et al. 2007; Korneliussen et al. 2009), which are potentially more effective, the *Z*-score approach is robust and requires few assumptions. In addition, the mean *Z* score accounts for more variable estimates at certain frequency pairs (i.e., less weight is given to frequency pairs with a higher observed SD).

A key benefit of the approach is that $\overline{Z}$ provides an intuitive index of classification certainty. The units of $\overline{Z}$ are such that they describe how well the sample conforms to the measured frequency difference for a given class: a $\overline{Z}$ of 1 indicates that the sample has relative frequency response with an average of 1 SD relative to what was observed in the reference samples for a given species. It is important to recognize that all methods for remote species identification are likely to fail under some circumstances. To avoid making improper interpretations, it is important to be able to identify these situations. In this context, $\overline{Z}$ is a useful metric; a substantial increase in the average $\overline{Z}$ suggests a substantial violation of the underlying assumptions and suspect results.

We used only the relative frequency response, which is largely dependent on the morphology of the organism, and avoided incorporating other measures that are likely to be more variable in an effort to make a more robust classification method. We did not use measures of the absolute backscatter strength (e.g., a threshold, or an absolute S_v) as the basis for classification. Such measures are a function of the number of organisms in the range-dependent volume sampled, as well as the target strength of the organism, and will be sensitive to changes in the abundance, orientation, behavior, size, or depth distribution of organisms. In addition, we did not include other descriptors such as geographic location, vertical distribution, environmental, or statistical descriptions of school shape (Reid et al. 2000) in our classification method, as these are spatially and temporally variable. For example, pollock are found in a wide depth range and a variety of spatial patterns, including solitary individuals, pelagic and demersal layers, and discrete schools of a broad range of sizes (Burgos and Horne 2008; Stienessen and Wilson 2008). Measurements of aggregation morphology from our ground-truthed backscatter are unlikely to encompass the diversity of aggregations for the taxa under study here. We chose not to formally include this information (e.g., Lawson et al. 2001; Fernandes 2009; Korneliussen et al. 2009) in an effort to develop a more general approach. However, this type of information can be used to complement and assess the validity of the multifrequency analysis.

Validation of the *Z*-score method

The *Z*-score method makes several assumptions: (*i*) the observed relative frequency response is generally applicable to the species over the study area (ΔS_v is spatially and temporally consistent and normally distributed); (*ii*) the relative frequency response of taxa does not overlap; and (*iii*) at the scale of an analysis cell, spatial overlap among taxa is low.

To evaluate whether these assumptions hold in practice for pollock, we compared the *Z*-score estimate of pollock backscatter with the traditional estimate, which is based on trawl sampling and expert judgment of echograms. The species classification methods are independent, as the traditional method combines prior knowledge about geographic distribution, depth distribution, and aggregation structure with targeted trawling on selected aggregations, whereas the *Z*-score method uses only the relative differences in backscatter at different frequencies. Overall, we observed good agreement between the methods, particularly after applying the $\overline{Z}_{\text{pollock}} < 1.5$ at the EDSU level criterion, which primarily excluded near-surface backscatter with a relative frequency response that overlapped with that of pollock but was, on average, only marginally consistent with pollock. The nested approach of thresholding $\overline{Z}_{\text{pollock}}$ at the scale of both the analysis cell and EDSU has the effect of classifying marginal analysis cells (i.e., $1.5 \leq \overline{Z}_{\text{pollock}} \leq 2$) as consistent with pollock only when the surrounding analysis cells in the EDSU are more consistent with the expectation for pollock.

The consistency between the estimates of pollock abundance based on the two methods indicates that for pollock in the eastern Bering Sea, the assumptions of the *Z*-score method are largely met. Similarly, the low $\overline{Z}_{\text{pollock}}$ for backscatter identified as pollock suggests that the assumptions are not consistently violated. This indicates that an abundance index of pollock approximating that from a traditional survey can be made with species identification based solely on the basis of relative frequency response in this environment. However, one must keep in mind that the eastern Bering Sea midwater community is dominated by a few abundant species (e.g., Honkalehto et al. 2002), as is typical in many high-latitude ecosystems. In this environment, the assumption that backscatter with a relative frequency response similar to that of pollock is attributable to pollock is not unreasonable. Species classification on the basis of relative frequency response will be more difficult in more diverse environments where these assumptions do not hold.

Areas for refinement

Although the method is effective as implemented, there are several refinements that may offer improvements. The *Z*-score framework can be used to incorporate new information on relative frequency response based on measurements as reported here, or theoretical predictions (e.g., Gauthier and Horne 2004; Lavery et al. 2007) as they become available. The *Z*-score approach weights the information at all frequency pairs equally. This is reasonable when attempting to distinguish a taxon from other, unknown sources of scattering as it uses all available information. However, this approach may be less effective when the relative frequency response of an organism is known to differ from other species present only at some frequency pairs (e.g., eulachon), as the similarity at the other frequency pairs will reduce the apparent difference in relative frequency response. In these cases, one might apply criteria using all frequencies as a preclassification step to isolate backscatter that is generally consistent with the taxa of interest and then add additional requirements at the specific frequency pairs at which these taxa differ to maximize the probability of discrimination.

Selection of an appropriate scale of spatial averaging is a central consideration in multifrequency backscatter classification. Averaging over several pings reduces random variability in measurements of relative frequency response and minimizes biases introduced by differences in the volume sampled at the different frequencies (Korneliussen and Ona 2002). At the same time, relatively fine-scale averaging is desirable to minimize the extent to which backscatter from multiple taxa will be observed in a single analysis cell (Demer et al. 2009). Performance at either very low or very high levels of averaging will be poor. The impact of averaging on variance is fairly easy to evaluate, but understanding the degree to which monospecific aggregations overlap in analysis cells of a given size remains a challenge as this will change over space and time. Multifrequency discrimination can be improved in cases where organisms form well-defined monospecific aggregations by preclassifying fish aggregations and combining all samples in an aggregation prior to multifrequency analysis to maximize the degree of averaging (Fernandes 2009; Korneliussen et al. 2009). In addition, combining preclassification at a fine spatial resolution to remove samples that are inconsistent with the taxon of interest, followed by classification based on the relative frequency response in spatially averaged analysis cells, is promising (Demer et al. 2009).

Utility of multifrequency classification

The *Z*-score method has the potential to improve the ability to interpret backscatter during acoustic surveys. Additional information gained from the relative frequency response can be used to improve interpretation of acoustic backscatter strength in several ways. The output of the method can be used quantitatively, as applied to pollock in the example given here. Alternatively, it can be used as a preprocessing step prior to subjective backscatter allocation. The method will eliminate backscatter that is not consistent with a target organism, which is useful even if it does not separate the target organism from all other taxa, as this changes the interpretation paradigm from "what is" to "what is not" backscatter from the organism of interest (c.f., Jech and Michaels 2006). The relative frequency response can be summarized in real time, and synthetic echograms can be used as an additional source of information to interpret backscatter visually or to better inform the decision of where to target a trawl. For example, if one observes a subtle shift in the relative frequency response, one might choose to allocate a trawl to ensure that the species composition has not changed. In addition, the method provides an index of certainty in backscatter interpretation, which can be used to identify areas or times where there is less certainty in classifications made based on the relative frequency response or subjective interpretation. For example, the survey average *Z* score can be used to quantify interannual or regional changes in uncertainty in species allocation during a survey time series in situations where species classification is based either on relative frequency response or on the traditional methodology.

Multifrequency information can also provide valuable information on nontarget species (Koslow 2009) from existing surveys of commercially important species at a modest increase in effort and cost. In the eastern Bering Sea survey area considered here, the relative frequency response indicates that euphausiids can be reliably separated from other dominant scattering groups. We are currently developing an index of euphausiid abundance for the eastern Being Sea, with species identification based on the *Z*-score method and a limited number of trawl samples for verification. The multifrequency approach will also likely be helpful when making inferences from acoustic backscatter when sampling is limited or not possible, for example, when making opportunistic observations from platforms such as fishing vessels, moorings, drifters, and autonomous vehicles.

Acknowledgements

Members of the Alaska Fisheries Science Center's (AFSC) Midwater Assessment and Conservation Engineering Program collected data at sea, and members of AFSC's Research Fishing Gear Program constructed and maintained the trawl gear. The efforts of the officers and crews of the NOAA ships *Miller Freeman* and *Oscar Dyson* made this work possible. Amanda Koltz conducted early work on the relative frequency response of rockfish, and Tess Brandon contributed to the postprocessing of classification results. Chris Wilson, Stephane Gauthier, Steven Barbeaux, Sandra Parker-Stetter, and Kelly Benoit-Bird provided insightful comments on earlier drafts of the manuscript. Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA

References

Anderson, C.I.H., Horne, J.K., and Boyle, J. 2007. Classifying multi-frequency fisheries acoustic data using a robust probabilistic classification technique. J. Acoust. Soc. Am. **121**(6): EL230–EL237. doi:10.1121/1.2731016. PMID:17552574.

Barange, M., and Hampton, I. 1994. Influence of trawling on *in situ* estimates of Cape horse mackerel (*Trachurus trachurus capensis*) target strength. ICES J. Mar. Sci. **51**(1): 121–126. doi:10.1006/jmsc.1994.1012.

Benoit-Bird, K.J. 2009. The effects of scattering-layer composition, animal size, and numerical density on the frequency response of volume backscatter. ICES J. Mar. Sci. **66**(3): 582–593. doi:10.1093/icesjms/fsp013.

Brierley, A.S., Ward, P., Watkins, J.L., and Goss, C. 1998. Acoustic discrimination of Southern Ocean zooplankton. Deep-Sea Res. **45**(7): 1155–1173. doi:10.1016/S0967-0645(98)00025-3.

Brierley, A.S., Axelsen, B.E., Boyer, D.C., Lynam, C.P., Didcock, A.S., Boyer, H.J., Sparks, C.A.J., Purcell, J.E., and Gibbons, M.J. 2004. Single-target echo detections of jellyfish. ICES J. Mar. Sci. **61**(3): 383–393. doi:10.1016/j.icesjms.2003.12.008.

Burgos, J.M., and Horne, J.K. 2008. Characterization and classification of acoustically detected fish spatial distributions. ICES J. Mar. Sci. **65**(7): 1235–1247. doi:10.1093/icesjms/fsn087.

Butler, J.L., and Pearcy, W.G. 1972. Swimbladder morphology and specific gravity of myctophids off Oregon. J. Fish. Res. Board Can. **29**: 1145–1150.

De Robertis, A., and Higginbottom, I. 2007. A post-processing technique to estimate the signal-to-noise ratio and remove echosounder background noise. ICES J. Mar. Sci. **64**: 1282–1291. doi:10.1093/icesjms/fsm112.

De Robertis, A., and Wilson, C.D. 2006. Walleye pollock respond to trawling vessels. ICES J. Mar. Sci. **63**(3): 514–522. doi:10.1016/j.icesjms.2005.08.014.

De Robertis, A., Hjellvik, V., Williamson, N.J., and Wilson, C.D. 2008. Silent ships do not always encounter more fish: comparison of acoustic backscatter recorded by a noise-reduced and a conventional research vessel. ICES J. Mar. Sci. **65**(4): 623–635. doi:10.1093/icesjms/fsn025.

Demer, D.A., Cutter, G.R., Renfree, J., and Butler, J.L. 2009. A statistical–spectral method for echo classification. ICES J. Mar. Sci. **66**(6): 1081–1090. doi:10.1093/icesjms/fsp054.

Fernandes, P.G. 2009. Classification trees for species identification of fish-school echotraces. ICES J. Mar. Sci. **66**(6): 1073–1080. doi:10.1093/icesjms/fsp060.

Fernandes, P.G., Korneliussen, R.J., Lebourges-Dhaussy, A., Masse, J., Iglesias, M., Diner, N., and Ona, E. 2006. The SIMFAMI project: species identification methods from acoustic multifrequency information. Final Report to the EC No. Q5RS-2001-02054. Available from Rolf Korneliussen Institute of Marine Research, P.O. Box 1870 Nordnes, N-5817 Bergen, Norway.

Foote, K.G. 1980. Importance of the swimbladder in acoustic scattering fish: a comparison of gadoid and mackerel target strengths. J. Acoust. Soc. Am. **67**(6): 2084–2089. doi:10.1121/1.384452.

Foote, K.G. 1985. Rather-high-frequency sound scattering by swimbladdered fish. J. Acoust. Soc. Am. **78**(2): 688–700. doi:10.1121/1.392438.

Gauthier, S., and Horne, J.K. 2004. Potential acoustic discrimination within boreal fish assemblages. ICES J. Mar. Sci. **61**(5): 836–845. doi:10.1016/j.icesjms.2004.03.033.

Godø, O.R., Patel, R., and Pedersen, G. 2009. Diel migration and swimbladder resonance of small fish: some implications for analyses of multifrequency echo data. ICES J. Mar. Sci. **66**(6): 1143–1148. doi:10.1093/icesjms/fsp098.

Greenlaw, C.F. 1979. Acoustical estimation of zooplankton populations. Limnol. Oceanogr. **24**(2): 226–242. doi:10.4319/lo.1979.24.2.0226.

Higginbottom, I.R., Pauly, T.J., and Heatley, D.C. 2000. Virtual echograms for visualization and post-processing of multiple-frequency echosounder data. *In* Proceedings of the Fifth European Conference on Underwater Acoustics, 10–13 July 2000, Lyon, France. Office for Official Publications of the European Communities, Luxembourg. pp. 1497–1502.

Holliday, D.V., and Pieper, R.E. 1995. Bioacoustical oceanography at high frequencies. ICES J. Mar. Sci. **52**(3–4): 279–296. doi:10.1016/1054-3139(95)80044-1.

Honkalehto, T., Patton, W.E., de Blois, S., and Williamson, N.J. 2002. Echo integration-trawl survey results for walleye pollock (*Theragra chalcogramma*) on the Bering Sea shelf and slope during summer 2000. NOAA Technical Memorandum NMFS-AFSC-126. Available from http://www.afsc.noaa.gov/Publications/AFSC-TM/NOAA-TM-AFSC-126.pdf.

Honkalehto, T., McKelvey, D., and Williamson, N.J. 2005. Results of the echo integration-trawl survey of walleye pollock (*Theragra chalcogramma*) on the U.S. and Russian Bering Sea shelf in June and July 2004. AFSC Processed Rep. 2005-02. Available from http://www.afsc.noaa.gov/Publications/ProcRpt/PR%202005-02.pdf.

Horne, J.K. 2000. Acoustic approaches to remote species identification: a review. Fish. Oceanogr. **9**(4): 356–371. doi:10.1046/j.1365-2419.2000.00143.x.

Horne, J.K., and Jech, J.M. 1999. Multi-frequency estimates of fish abundance: constraints of rather high frequencies. ICES J. Mar. Sci. **56**(2): 184–199. doi:10.1006/jmsc.1998.0432.

Jech, J.M., and Michaels, W.L. 2006. A multifrequency method to classify and evaluate fisheries acoustics data. Can. J. Fish. Aquat. Sci. **63**(10): 2225–2235. doi:10.1139/F06-126.

Karp, W.A., and Walters, G.E. 1994. Survey assessment of semi-pelagic gadoids: the example of walleye pollock, *Theragra chalcogramma*, in the Eastern Bering Sea. Mar. Fish. Rev. **56**: 8–22.

Kloser, R.J., Ryan, T., Sakov, P., Williams, A., and Koslow, J.A. 2002. Species identification in deep water using multiple acoustic frequencies. Can. J. Fish. Aquat. Sci. **59**(6): 1065–1077. doi:10.1139/f02-076.

Korneliussen, R.J., and Ona, E. 2002. An operational system for processing and visualizing multi-frequency acoustic data. ICES J. Mar. Sci. **59**(2): 293–313. doi:10.1006/jmsc.2001.1168.

Korneliussen, R.J., and Ona, E. 2003. Synthetic echograms generated from the relative frequency response. ICES J. Mar. Sci. **60**(3): 636–640. doi:10.1016/S1054-3139(03)00035-3.

Korneliussen, R.J., Diner, N., Ona, E., Berger, L., and Fernades, P.G. 2008. Proposals for the collection of multifrequency acoustic data. ICES J. Mar. Sci. **65**(6): 982–994. doi:10.1093/icesjms/fsn052.

Korneliussen, R.J., Heggelund, Y., Eliassen, I.K., and Johansen, G.O. 2009. Acoustic species identification of schooling fish. ICES J. Mar. Sci. **66**(6): 1111–1118. doi:10.1093/icesjms/fsp119.

Koslow, J.A. 2009. The role of acoustics in ecosystem-based fishery management. ICES J. Mar. Sci. **66**(6): 966–973. doi:10.1093/icesjms/fsp082.

Lavery, A.C., Wiebe, P.H., Stanton, T.K., Lawson, G.L., Benfield, M.C., and Copley, N. 2007. Determining dominant scatterers of sound in mixed zooplankton populations. J. Acoust. Soc. Am. **122**(6): 3304–3326. doi:10.1121/1.2793613. PMID:18247742.

Lawson, G.L., Barange, M., and Freon, P. 2001. Species identification of pelagic fish schools on the South African continental shelf using acoustic descriptors and ancillary information. ICES J. Mar. Sci. **58**(1): 275–287. doi:10.1006/jmsc.2000.1009.

MacLennan, D.N., and Holliday, D.V. 1996. Fisheries and plankton acoustics: past, present and future. ICES J. Mar. Sci. **53**(2): 513–516. doi:10.1006/jmsc.1996.0074.

Madureira, L.S.P., Everson, I., and Murphy, E.J. 1993. Interpretation of acoustic data at two frequencies to discriminate between Antarctic krill (*Euphuasia superba* Dana) and other scatters. J. Plankton Res. **15**(7): 787–802. doi:10.1093/plankt/15.7.787.

McClatchie, S., Thorne, R.E., Grimes, P., and Hanchet, S. 2000. Ground truth and target identification for fisheries acoustics. Fish. Res. **47**(2–3): 173–191. doi:10.1016/S0165-7836(00)00168-5.

McKelvey, D.R., and Wilson, C.D. 2006. Discriminant classification of fish and zooplankton backscattering at 38- and 120-kHz. Trans. Am. Fish. Soc. **135**(2): 488–499. doi:10.1577/T04-140.1.

Miyashita, K., Aoki, I., Seno, K., Taki, K., and Ogishima, T. 1997. Acoustic identification of isada krill, *Euphausia pacifica* Hansen, off the Sanriku coast, north-eastern Japan. Fish. Oceanogr. **6**(4): 266–271. doi:10.1046/j.1365-2419.1998.00042.x.

Nakashima, B.S. 1990. Escapement from a Diamond IX midwater trawl during acoustic surveys for capelin (*Mallotus villosus*) in the Northwest Atlantic. J. Cons. Int. Explor. Mer, **47**: 76–82 doi:10.1093/icesjms/47.1.76.

Ona, E. 2003. An expanded target-strength relationship for herring. ICES J. Mar. Sci. **60**(3): 493–499. doi:10.1016/S1054-3139(03)00031-6.

Pikitch, E.K., Santora, C., Babcock, E.A., Bakun, A., Bonfil, R., Conover, D.O., Dayton, P., Doukakis, P., Fluharty, D., Heneman, B., Houde, E.D., Link, J., Livingston, P.A., Mangel, M., McAllister, M.K., Pope, J., and Sainsbury, K.J. 2004. Ecosystem-based fishery management. Science (Washington, D.C.), **305**(5682): 346–347. doi:10.1126/science.1098222.

Reid, D., Scalabrin, C., Petigas, P., Masse, J., Aukland, R., Carrera, P., and Gregorakarakos, S. 2000. Standard protocols for the analysis of school based data from echo sounder surveys. Fish. Res. **47**(2–3): 125–136. doi:10.1016/S0165-7836(00)00164-8.

Simmonds, E.J., and MacLennan, D.N. 2005. Fisheries acoustics. 2nd ed. Blackwell Science Ltd., Oxford, UK.

Simmonds, E.J., Armstrong, F., and Copland, P.J. 1996. Species identification using wideband backscatter with neural network and discriminant analysis. ICES J. Mar. Sci. **53**(2): 189–195. doi:10.1006/jmsc.1996.0021.

Sinclair, E.H., and Stabeno, P. 2002. Mesopelagic nekton and associated physics of the southeastern Bering Sea. Deep Sea Res. Part II Top. Stud. Oceanogr. **49**(26): 6127–6145. doi:10.1016/S0967-0645(02)00337-5.

Stanton, T.K., Chu, D., and Wiebe, P.H. 1996. Acoustic scattering characteristics of several zooplankton groups. ICES J. Mar. Sci. **53**(2): 289–295. doi:10.1006/jmsc.1996.0037.

Stienessen, S., and Wilson, C.D. 2008. Juvenile walleye pollock aggregation structure in the Gulf of Alaska. *In* Resiliency of gadid stocks to fishing and climate change. *Edited by* G.H. Kruse, K. Drinkwater, J.N. Inaelli, D.L. Stram, V. Wespedtad, and D. Woodby. Alaska Sea Grant College Program, Fairbanks, Alaska. pp. 271–287.

Tichy, F.E., Solli, H., and Kaveness, H. 2003. Non-linear effects in a 200-kHz sound beam and the consequences for target strength measurement. ICES J. Mar. Sci. **60**(3): 571–574. doi:10.1016/S1054-3139(03)00033-X.

Wiebe, P.H., and Benfield, M.C. 2003. From the Hensen net toward four-dimensional biological oceanography. Prog. Oceanogr. **56**(1): 7–136. doi:10.1016/S0079-6611(02)00140-4.

Zakharia, M.E., Magand, F., Hetroit, F., and Diner, N. 1996. Wideband sounder for fish species identification at sea. ICES J. Mar. Sci. **53**(2): 203–208. doi:10.1006/jmsc.1996.0023.

Zhou, M., Nordhausen, W., and Huntley, M. 1994. ADCP measurements of the distribution and abundance of euphausiids near the Antarctic Peninsula in winter. Deep-Sea Res. **41**(9): 1425–1445. doi:10.1016/0967-0637(94)90106-6.