Combining bottom trawls and acoustics in a diverse semipelagic environment: What is the contribution of walleye pollock (Gadus chalcogrammus) to near-bottom acoustic backscatter in the eastern Bering Sea?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Fisheries and Aquatic Sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>cjfas-2015-0481.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>06-May-2016</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Lauffenburger, Nathan; National Marine Fisheries Service- NOAA, Alaska Fisheries Science Center De Robertis, Alex; National Marine Fisheries Service - NOAA, Alaska Fisheries Science Center Kotwicki, Stan; National Marine Fisheries Service - NOAA,</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Acoustic Trawl Survey, Demersal Fishes, Species Diversity, Walleye Pollock, BIOMASS &lt; General</td>
</tr>
</tbody>
</table>
Combining bottom trawls and acoustics in a diverse semipelagic environment: What is the contribution of walleye pollock (Gadus chalcogrammus) to near-bottom acoustic backscatter in the eastern Bering Sea?

Authors: Nathan Lauffenburger¹#, Alex De Robertis¹, Stan Kotwicki¹

¹-Alaska Fisheries Science Center
National Marine Fisheries Service, NOAA
7600 Sand Point Way NE
Seattle, WA, 98115

# corresponding author: nathan.lauffenburger@noaa.gov, 206-526-4177
Abstract

The abundance of walleye pollock (*Gadus chalcogrammus*) in the eastern Bering Sea (EBS) is estimated in part through fisheries-independent acoustic trawl (AT) surveys, which currently use acoustic backscatter data down to 3 m above the bottom. A large portion of adult pollock are demersal and these estimates will become more accurate if the survey is extended closer to bottom. The purpose of this project was to assess the feasibility of extending the AT survey closer to the bottom by estimating the contributions of each demersal fish species to observed acoustic backscatter in the highly diverse near-bottom region. This was accomplished by fitting a regression model to simultaneously collected acoustic backscatter and bottom trawl (BT) catch data. Pollock were the dominant source of acoustic backscatter among demersal species accounting for $85.9 \pm 4.8\%$ of acoustic backscatter (mean ± standard deviation). A method was developed to extend the AT survey to within 0.5 m of the bottom and applied to the 1994-2014 surveys, pollock biomass increased by an average of $35 \pm 12\%$.

Keywords: acoustic trawl survey, acoustic backscatter, demersal fishes, species diversity, walleye pollock, biomass
Introduction

Acoustic-trawl (AT) surveys are commonly used to derive abundance estimates of pelagic species by combining acoustic backscatter with species and length information from trawls (Honkalehto et al. 2013, Jakobsen et al. 1997). However, it is more challenging to quantify the biomass of semipelagic species near the seafloor with the same methods due to different biases inherent in the acoustic and the bottom trawl methods (McQuinn et al. 2005). AT and bottom trawl (BT) surveys have been combined to produce a single estimate of a species biomass (Jakobsen et al. 1997, Godø and Wespestad 1993) but the accuracy tends to be in question, as the two methods measure different fractions of the population, which may overlap in space and time causing uncertainties (Lawson and Rose 1999). These include effective fishing height (Aglen 1996, Hjellvik et al. 2003), the effective sampling volume of trawls (Handegard and Tjøstheim 2009), and uncertainties due to bottom trawl geometry (Walsh 1996) and the acoustic dead zone where fish cannot be acoustically detected (Kotwicki et al. 2013, Ona and Mitson 1996).

There have been a number of studies aiming to quantify the factors that are associated with these uncertainties by direct comparison of acoustic and bottom trawl data simultaneously collected (Aglen 1996, Kotwicki et al. 2013, McQuinn et al. 2005). These projects are typically designed to examine hauls containing a targeted single species or a few species, such as Pacific cod (Gadus macrocephalus) or walleye pollock (Gadus chalcogrammus; hereafter referred to as pollock).

While this approach is beneficial in assessing fish behavior and the biases associated with acoustics and trawl methods for a single species, these studies do not generalize to describe the characteristics and challenges of a multi-species trawl survey. This study focuses on an approach to combine acoustic and bottom trawl data to assess the contributions of all species present (separated into groups) while accounting for the uncertainties introduced by the biases
described above. The results of this exercise will provide a more accurate AT abundance estimate by expanding the estimates of pollock closer to the seafloor.

Pollock are dominant semipelagic fish which play important commercial and ecological roles in the North Pacific. Pollock account for about 5% of global fish harvest, with annual harvests ranging from 4-7 million metric tons (Bailey et al. 1999). Since the late 1970s, the pollock stock abundance in the eastern Bering Sea (EBS) has been monitored by the National Marine Fisheries Service using AT and BT fisheries-independent survey methods (Lauth and Nichol 2013; Honkalehto et al. 2013). In the case of the AT survey, abundance estimates are determined from acoustic backscatter between ~15 m below the surface to 3 m above the seafloor, with ground-truth catch data from trawls. While the AT survey has a near-bottom blind zone where fish are either not detected or undersampled (Ona and Mitson 1996), this method can detect fish closer to the bottom than this 3 m height.

The acoustic data are used down to 3 m above the seafloor primarily for historical reasons and due to the uncertainty in the fraction of the backscatter in the lower 3 m that is caused by pollock. The AT survey was used in the past in conjunction with the BT survey, which has a nominal headrope height of 3 m (Von Szalay et al. 2007), to create a combined index of abundance (Bakkala et al. 1986; Karp and Walters 1994). This combined index has since been abandoned in favor of independent BT and AT indices of abundance due to the factors mentioned above (Ianelli et al. 2009). Presently, the BT survey is used as an abundance index of near-bottom pollock and the AT survey is utilized as a separate index of abundance for the midwater component of the population (Ianelli et al. 2009). In the current EBS pollock AT surveys, backscatter is measured (but not used) to within 0.5 m of the echosounder detected seafloor and midwater trawl catches are used to interpret the midwater sound scatters. Given that
a sizeable portion of the pollock population is located within the first 3 m off bottom (Kotwicki et al. 2005), use of the acoustic data down to 0.5 m above bottom will encompass a larger proportion of the population and consequently should produce more accurate abundance estimates. However, species diversity as estimated by the bottom trawls (Lauth and Nichol 2013) is much larger within the first 3 m off bottom compared to the species diversity above 3 m as detected by the midwater trawls (Honkalehto et al. 2013). For example, from 2006 to 2011, pollock accounted for 95 ± 4% of the catch by weight in the midwater trawls conducted during the AT survey, and 23 ± 3% of the BT survey trawls by weight (includes fish and non-fish species for both surveys). Therefore, it has been unclear what proportion of the near-bottom backscatter can be attributed to pollock. An improved understanding of the proportion of the backscatter within 3 m of the bottom that is attributable to pollock is required to extend the survey results to 0.5 m off bottom.

The goal of this project was to assess the feasibility of extending the AT survey down to a half meter above the echosounder detected seafloor by estimating the contribution of pollock and other species to the acoustic backscatter observed between 0.5 and 3 m. A statistical model was developed to establish the correlation between the catch of a given fish species and acoustic backscatter during each haul event from concurrent area-swept BT catch estimates and measurements of acoustic backscatter.

Methods

Bottom trawl catch and acoustic data were simultaneously collected as part of the Alaska Fisheries Science Center’s annual EBS BT survey (see Lauth and Nichol 2013 for details)
between the years 2006 and 2011. The chartered commercial fishing vessels F/V *Aldebaran*, F/V *Arcturus*, F/V *Northwest Explorer*, and F/V *Alaska Knight* were used for the BT surveys.

Trawl stations were conducted on a 20 x 20 nautical mile (nmi) grid with increased spatial resolution near St. Matthews and Pribilof Islands. At each station, an 83O112 eastern otter trawl (Lauth and Nichol 2013) was towed for 30 minutes at 3 knots.

The BT survey stations extend farther north-east onto the shelf than the AT survey track lines, so only bottom trawls within the AT survey area were chosen for analysis (Fig. 1). Each survey year includes data from two charter vessels, except for 2007, when acoustic data were only collected on one vessel. In 2007, efforts were made to ensure that the single vessel’s spatial coverage was well distributed over the EBS shelf so these data are still acceptable for this project. The data are widely distributed over the survey area, even though the number of processed stations varies from year to year. The analysis was restricted to US territorial waters where both AT and BT data are available.

Catch per unit effort data from BT stations were calculated using the area-swept method (Alverson and Pereyra 1969) whereby area swept was estimated by multiplying distance fished while the trawl was on bottom (as measured by a bottom contact sensor; Somerton and Weinberg 2001) by the average distance between wing tips measured using Netmind spread sensors (see Weinberg and Kotwicki 2008 for details). Random subsamples of up to 300 specimens per fish species were measured for fork length (total length was measured when fork length was not suitable for the species) to the nearest 1.0 cm for each tow (Lauth and Nichol 2013). The resulting length frequency data for the subsample was then extrapolated to the entire catch for each measured species. The BT catch was divided into 16 groups: pollock, Arctic cod (*Arctogadus glacialis*), eelpouts (family *Zoarcidae*), large flatfishes (adult length > 50 cm; order...
Pleuronectiformes), Pacific cod, rockfishes (Sebastes spp.), salmon (family Salmonidae), saffron cod (Eleginus gracilis), small flatfishes (adult length < 50 cm), skates, miscellaneous, capelin (Mallotus villosus), eulachon (Thaleichthys pacificus), Pacific herring (Clupea pallasii), sand lance (Ammodytes hexapterus), and other fishes. Five of the species groups (the last five in the previous list) only had measurements of weight collected from the catch and no length measurements were available. The mean lengths per haul for these species were estimated with a bias-corrected allometric weight-length relation (De Robertis and Williams 2008). The catch (C) per unit effort at length (j is index for length class) by species (sp is index for species) for each haul is computed as

\[ C_{sp,j} = \frac{N_j}{S} \]

where \( N_j \) is the number of fish at length j in the catch and S is the area-swept in nmi\(^2\).

The BT survey vessels were equipped with 38 kHz split-beam Simrad ES60 echosounders and consistent protocols for calibration, instrument settings, and data logging were followed (Honkalehto et al. 2011). Acoustic data collected during each trawl haul were processed following the semi-automated techniques described in Kotwicki et al. 2009. Nautical area scattering coefficients (\( s_A \); see MacLennan et al. 2002 for definitions of acoustic units) were computed and vertically integrated using the readEKRaw library from the EchoLab MATLAB (MATLAB 2013a) toolkit. These data were integrated with a -70 dB re 1 m\(^{-1}\) minimum threshold and the output was stored in 0.5 m vertical bins < 5 m above the ES60 sounder-detected bottom and in 1 m bins > 5 m above bottom. The application computed an alternate amplitude-based bottom detection which was used in cases when there was no sounder-detected
bottom available or when the sounder-detected bottom deviated more than the 0.5 m from previous bottom detections.

A total of 1,434 echograms (i.e., color representations of backscattering strength) were produced by the application for visual inspection, shown in Fig. 2. Trawl sites were excluded (totaling 643) from use in the analysis when bottom integrations, noise spikes (e.g. noise caused by hydraulics), and backscatter unlikely to be from fish were observed in the lower 3 m of the echograms. Bottom integration is when a line that is generated by the data acquisition software (to exclude backscatter from the seafloor in the acoustic analysis, i.e. bottom detection) drops below the bottom echo. Including the strong backscatter from the seafloor introduces artefacts that will mask the weaker biological signals. Backscatter unlikely to be from fish refers to backscatter in the echogram that is clearly not from adult fishes based on evaluation of the backscatter strength and schooling behavior: it is typically weak and diffuse. Non-fish backscatter is evident in Fig. 2 from the surface to 20 m depth and other fish species (most likely pollock) are seen below 50 m as individuals or small schools. The high percent of data exclusion (45%) was primarily due to bottom integrations as the automated processing software did not allow for manual correction of the sounder-detected bottom. A total of 791 hauls conducted over 6 years were used to fit the model. A periodic and systematic error that results in a maximum 1 dB (23%) bias in acoustic backscatter measurements made on individual transmissions with the ES60 was removed by fitting the error to the otherwise constant transmit pulse and correcting the data (Ryan and Kloser 2004). Only those data collected during periods when the trawl was in contact with the seafloor were used in the analysis.
Model Formulation

The premise of the model is that the observed acoustic backscatter, $B$, is proportional to the linear sum of the catch per unit effort of each species, $C_{sp}$. The basic form of the model is a multiple variable linear regression, i.e.,

$$ (2) \quad B = \sum_{sp} A_{sp} C_{sp} + I $$

where $A_{sp}$ are coefficients that are fit to each species and $I$ is the intercept from the model fit. As detailed below, the model becomes somewhat more complex when fish length is accounted for and the log-normal error assumption is used within the model. The quantities and symbols used in model development are summarized in Table 1.

In order to inspect the relationship between the demersal catch in each trawl and the acoustic data observed under the vessel during the haul, the nautical area scattering coefficient, $s_A$, was summed from 0.5 to 3 m above the seafloor. We assume that backscatter is proportional to fish abundance (Foote, 1983): thus depth-summed $s_A$ is proportional to the product of the density of fish in the area sampled by the beam and the acoustic backscattering cross section, $\sigma_{bs}$, across all length classes for each trawl. In this model, the density of species $sp$ in the beam is represented by the BT catch ($C_{sp}$) multiplied by $a_{sp}$, which is a species-specific parameter accounting for the average effect of an individual captured in the bottom trawl on acoustic backscatter.

Therefore, a model for each trawl event ($i$), each species ($sp$) and each length class ($j$) is constructed as

$$ (3) \quad (\sum_{h=0.5}^{3} s_A)_i = (\sum_{sp} [a_{sp} \cdot \sum_{j} C_{sp,j} \cdot \sigma_{bs,sp,j}])_i + I $$
The constant, $I$, is an intercept term that represents observed acoustic backscatter from species that were not quantified such as larval fishes and zooplankton, which were not retained in the trawl. Hereafter, for simplification, $B = \sum_{h=0.5}^{3} s_A$. Thus, eq. 3 becomes

$$B_i = (\sum_{sp} [a_{sp} \cdot \sum_{j} C_{sp,j} \cdot \sigma_{bs,sp,j}]) + I$$

The trawl-dependent right hand side of the equation includes a summation over all species groups, with $C_{sp,j}$ and $\sigma_{bs,sp,j}$ first summed along all length classes within each species group. The $\sigma_{bs}$ is the acoustic backscattering cross section (i.e., a linear measure of acoustic backscatter from an individual), which is length and species dependent. It is defined as

$$\sigma_{bs,sp,j} = 10^{TS_{sp,j}/10}$$

where $TS$ is the target strength relationship and takes the form

$$TS_{sp,j} = 20 \cdot \log_{10}(L) - b_{sp}.$$ 

This target strength equation, which assumes that backscatter is proportional to surface area can be substituted into eq. 5 and simplified as

$$\sigma_{bs,sp,j} = 10^{(2 \cdot \log_{10}(L) - b_{sp})/10}$$

$$\sigma_{bs,sp,j} = 10^{2 \cdot \log_{10}(L) \cdot 10^{-b_{sp}/10}}$$

$$\sigma_{bs,sp,j} = L^2 \cdot 10^{-b_{sp}/10}.$$ 

Using eq. 9 for $\sigma_{bs,sp,j}$, eq. 3 can be adjusted to separate out terms dependent on length and the terms only dependent on species:
(10) \[ B_i = \left( \sum_{sp} a_{sp} \cdot 10^{-b_{sp}/10} \cdot \sum_j C_{sp,j} \cdot L^2 \right)_i + I \]

To simplify, we define representative variables to use in the model.

(11) \[ T_{sp,L} = C_{sp,j} \cdot L^2 \]

(12) \[ A_{sp} = a_{sp} \cdot 10^{-b_{sp}/10} \]

where \( T_{sp,j} \) is the trawl catch scaled by length, which can be understood as the abundance expressed as fish surface area. The \( A_{sp} \) coefficient represents the size-normalized contribution of a single fish (of a given species) in the BT catch to the observed acoustic backscatter between 0.5 and 3 m above the bottom. This coefficient contains within it contributions from the target strength intercept, so \( b_{sp} \) does not need to be explicitly defined for any species category (see the discussion for further explanation of the \( A_{sp} \) coefficient). The final form of the expression used for fitting is

(13) \[ B_i = \left[ \left( \sum_{sp} A_{sp} \sum_j T_{sp,j} \right)_i + I \right] \cdot e^{\epsilon_i} \]

To determine values for parameters \( A_{sp} \), model fitting was performed using maximum likelihood, assuming log-normal error (\( e^{\epsilon_i} \)) with a negative log-likelihood (NLL) function

(14) \[ NLL = 0.5 N_T \log(2\pi\sigma^2) + \frac{\sum_{i=1}^{N_T} \left| \log(\tilde{B}_i) - \log(\tilde{B}_i) \right|^2}{2\sigma^2} \]

where \( N_T \) is the number of tows, \( \sigma^2 \) is the error variance, and \( \tilde{B}_i \) is the model prediction. A log-normal error was assumed following Kotwicki et al. 2013, who showed that a log-normal distribution is appropriate when predicting EBS BT catches from acoustic backscatter. All \( A_{sp} \) parameters were bounded to be greater than 0 as it is not possible for the fish to produce negative
backscattering values. Stations with negligible backscatter (< 2 m² nmi⁻²) were not used to fit because the model fit was disproportionally sensitive to changes at this magnitude. This is attributable to the log-normal error model and the use of an integration threshold, which introduces non-linearity when backscatter is low. Fitting was performed using the “mle2” function in the R package “bbmle” (Bolker 2012).

We performed forward variable selection starting with the model including only one of the species groups as a predictor. Each species group was individually tested in this single predictor model and an estimate of Akaike’s information criterion corrected for finite sample size was noted (AICc; Burnham and Angerson 2010). The species group that resulted in the minimum AICc was permanently incorporated into the model and the next species group was tested. Additions of species groups to the model continued in this manner until no reduction in the AICc was achieved. Profile likelihood confidence intervals were estimated using the R function “profile” (Bolker 2012).

Model diagnostics were performed using residual analyses that included scatter plots of the observed values and standardized residuals versus predicted values, and normal Q-Q plot. We also examined plots of standardized residuals versus all predictors. Variance inflation factors (VIF) were also calculated for all predictors in the final model to quantify the effects of possible multicollinearity in linear predictors (Kutner et al. 2004).

The predicted backscatter attributed to each species group by year was estimated by averaging trawl catch data (\( T_{sp} \)) over all hauls for each year, represented as \( \overline{T}_{sp,year} \). The predicted contribution of each species group by year to the observed backscatter from 0.5 to 3 m above the acoustic bottom detection is computed similarly to eq. 13 using the formula
Finally, the proportion of backscatter attributable to each species group by year ($Prop_{sp,year}$) was found by using

$$Prop_{sp,year} = \frac{s_{A,sp,year}}{\sum_{sp} s_{A,sp,year}}.$$

A sensitivity analysis was performed by year-based leave-one-out cross-validation (LOOCV). The data were fitted to the best model omitting one year at a time to examine how sensitive the estimates from eq. 16 were to the missing data. The proportion of each species to the backscatter in the omitted year was computed by refitting the model with the data from the target year excluded and then applying the resulting $A_{sp}$ parameter to the trawl catch data, $T_{sp,year}$ (eqs. 15 & 16) from the withheld year.

**Application to AT survey**

A method was developed to combine acoustic measurements from the AT survey with BT catches and the fitted parameters described above to extend the EBS pollock AT survey to 0.5 meters above the acoustic bottom detection. The best-fit model parameters, $A_{sp}$, for all significant fish species groups were used to estimate the proportion of backscattering attributable to pollock between 0.5 and 3 m off bottom in the summer EBS AT survey. The measured acoustic backscatter, $s_A$, from all 0.5 nmi elementary distance sampling units (EDSUs) in the AT survey within this depth layer was converted to pollock abundance by combining it with the size distribution of these pollock in nearby BT catch data. This near-bottom estimate was then added to the conventional (i.e., > 3 m above bottom) survey estimate.
For each 0.5 nmi EDSU, all BT survey CPUE data within a radius of 25 nmi were identified. At
the locations where there were no stations within this range, the range was incremented by 2 nmi
until there was at least one BT station for use (this occurred 1.8% of the time). A weighted
average (based on distance from the relevant EDSU) of the proportions from all selected BT
stations was found, using

\[
\overline{T_{sp}} = \sum_{i=1}^{N} \left( T_{sp,i} \cdot \frac{W_i}{\sum W_i} \right)
\]

where \( i \) represents each BT station, \( N \) is the number of BT stations within \( R_o \) (25 nmi) of the
current EDSU, and

\[
W_i = 1 - \frac{R_i}{R_o}
\]

where \( R_i \) is the distance (in nmi) from each BT station from the current EDSU. The weighted
mean proportion of backscatter attributable to pollock, \( \overline{Prop_{pk}} \), is found by placing \( \overline{T_{sp}} \) into eqs.
15 and 16. Near-bottom backscatter was summed between 0.5 and 3 m at the current EDSU and
multiplied by the weighted mean proportion, \( \overline{Prop_{pk}} \) to give an \( s_A \) that represents pollock
backscatter between 0.5 and 3 m off the bottom.

The size distribution of pollock in the 0.5 to 3 m layer for use in estimates of abundance at length
was computed in a similar fashion. A weighted length frequency \( \overline{L_{pk}} \) was determined for each
EDSU from the nearby bottom trawl catches using:

\[
\overline{L_{pk}} = \sum_{i=1}^{N} \left( L_{pk,i} \cdot \frac{W_i}{\sum W_i} \right)
\]

where \( L_{pk,i} \) is the length frequency for pollock at each BT station. Pollock biomass was then
computed using the methodology outlined in Honkalehto et al. (2008). These biomass estimates
were summed for each year and compared to biomass observed above 3 m off bottom for surveys between 1994 and 2014.

Results

The species composition of bottom trawl catches \( C_{sp,j} \) and catch normalized to surface area \( T_{sp,year} \) are presented in Fig. 3. The primary species groups in the catch for all years are small flatfishes and pollock. In 2010, Arctic cod were also abundant in the catch. The proportion of pollock for each year is larger than small flatfishes in terms of \( T_{sp,year} \) relative to \( C_{sp,j} \) (as pollock are on average larger). Pollock averaged 35% of the total \( C_{sp,j} \), and 53% of the total \( T_{sp} \).

The forward variable selection resulted in a final model indicating that the pollock, Arctic cod, large flatfish, rockfish and miscellaneous fish groups contributed significantly to the backscatter within 0.5 to 3 m off the bottom (Table 2). The other species groups were not significant and were excluded from the final model.

The model indicates that pollock accounted for approximately 90% of backscatter between 0.5 and 3 m off bottom in all years, except 2010, when large concentrations of Arctic cod were detected in both the BT and AT surveys (Fig. 4). In this year, pollock contributed 79% and Arctic cod contributed 12%.

The estimates of \( A_{sp} \) indicated that the expected backscatter per individual fish retained in the bottom trawl was highest for rockfishes and lowest for large flatfishes. The confidence intervals were small for the most abundant species (i.e., pollock and large flatfishes), indicating low
uncertainty in the contribution to observed backscatter. Greater uncertainty existed for the other significant species.

Diagnostic plots (Fig. 5) indicate that the assumption of log normal error was appropriate. Visual examination of standardized residuals against all predicted values and predictors (not shown) indicated that no trends were present in the residuals. Estimates of VIFs were below 1.02 for all predictors, indicating that multicollinearity of predictor variables was not a concern (Kutner et al. 2004).

The difference between the contribution of each species for each year left out in the cross-validation sensitivity analysis and the contribution of each species using the full model (Fig. 4) is shown in Table 3 (full model for each year – LOOCV model for each year). For most years and species, the change in percent contribution between the full model and the sensitivity analysis did not exceed 4%. The consistency between the annual estimates and the similarity to the full model estimates indicates that the coefficients fit in the model are robust to use for predictions outside of this dataset. An exception to the low variability in this sensitivity analysis is the year 2010, when Arctic cod was particularly high in the BT catch (see $T_{sp}$ in Fig. 3). Since there was very little Arctic cod caught during all other years, there are not enough data to reliably fit Arctic cod when 2010 is left out. Arctic cod was not significant using data from all other years, and it is dropped from the model and assigned a zero percent contribution to backscatter if 2010 is left out. The implications of this will be presented in the Discussion.

**Extending AT survey results to 0.5 m off bottom**
The model resulted in a substantial amount of pollock biomass from the 0.5 to 3 m off bottom acoustic backscatter (Fig. 6). The estimate of the near-bottom AT pollock biomass incorporating backscatter from 0.5 to 3 m above bottom averaged 0.79 million tons per year from 1994 to 2014 but varied from 0.42 million tons in 2006 to 1.52 million tons in 2014 (Fig. 6, top panel). The addition of these data increased pollock biomass by an average of 35%, but this varies interannually with a standard deviation of 12%. For example, the increase in pollock biomass from the AT survey ranged from 20% in 2010 to 60% in 2008 (Fig. 6, bottom panel).

Discussion

Species composition of the near-bottom fish community in the EBS habitat is diverse, with pollock only accounting for 35% of bottom trawl catch by number (from 2006 – 2011). However, our model shows that pollock dominates the acoustic backscatter within 0.5 and 3 m above bottom with an average contribution of 85% over the 6 years examined. The main conclusion of this study was that the AT survey index of abundance for pollock in the EBS can be improved by incorporating data below 3 m into the abundance estimates as a substantial portion of the pollock available to the AT survey is located in the layer between 0.5 to 3 m. Moreover, our results indicate that the proportion of the pollock biomass from 0.5 to 3 m relative to the biomass in the rest of the water column varies interannually (Fig. 6). This variability is of concern because it introduces additional error into AT index of abundance, which is not reflected in the present estimates of AT observation error (Walline, 2007). The large variability in the proportion of pollock distributed within the lower 3 m from year to year indicates that the observation error used in the pollock stock assessment to weight the contribution of each year (Ianelli et al. 2012) should include data from this layer to improve accuracy. The
implementation of our method introduces the uncertainty associated with estimation of the $A_{sp}$ parameter. However, this uncertainty is likely smaller than uncertainty associated with the variability of the proportion of pollock in the 0.5 to 3 m layer.

While the primary goal of this project was to determine the contribution of pollock to demersal backscatter and its implication for the AT survey index of abundance, it is of importance to understand the parameter coefficient $A_{sp}$ that is fit to the data. This coefficient represents the average contribution of a single fish (of a given species group) in the BT catch to the observed acoustic backscatter between 0.5 and 3 m above the bottom. The coefficient $A_{sp}$ serves as a unit conversion between the bottom-trawl catch and the acoustics (i.e., acoustic backscatter per unit surface area of the fish in the catch). The scaling to surface area corrects for the first-order effects of size: for many North Pacific species, acoustic backscatter increases roughly in proportion to fish surface area (e.g., Traynor 1996, Gauthier and Horne 2004). The coefficient $A_{sp}$ is proportional to the ratio of the selectivity of the acoustics to the selectivity of the BT catch data multiplied by the acoustic reflectivity (backscattering cross section) of an individual (scaled to 1 cm size) and is unique to each species group. Selectivity in the context of this study is defined as the ability of the survey method to accurately observe a fish species’ biomass.

One factor that affects the selectivity of the AT survey is properly quantifying the near-bottom acoustic dead zone (ADZ) where backscatter from fish cannot be distinguished from that of the seafloor (Kotwicki et al. 2013, Ona and Mitson 1996). In our method, a correction factor for the ADZ was not necessary because any fish species that was typically in the ADZ but caught in the bottom trawl would be represented by a smaller $A_{sp}$. In this case, the selectivity of the acoustics would be very small relative to the selectivity of the bottom trawl and this is captured in $A_{sp}$.  

https://mc06.manuscriptcentral.com/cjfas-pubs
Likewise, the selectivity of the bottom trawl can be influenced by a number of factors, including effective fishing height (Aglen 1996, Hjellvik et al. 2003) or effective sampling volume (Handegard and Tjøstheim 2009). Assuming that each species group has a mean diving and herding behavior (the model does not account for variation among hauls in fish behavior), the estimation of the proportion of acoustic backscatter within 0.5 to 3 m will be unbiased. For example, if pollock diving behavior results in an effective fishing height of 16 m off bottom (Kotwicki et al. 2013), the coefficient fit parameter $A_{pollock}$ will be reduced due to a higher bottom trawl catch than is represented by the acoustic backscatter between 0.5 and 3 m. However, when the proportion of acoustic backscatter attributable to pollock is computed, the multiplicative factor $C_{pollock}$ will reverse this influence (of a reduced coefficient) exactly and create an unbiased estimate of the proportion of acoustic backscatter attributable to pollock. Due to this effect, this model accounts for the average behaviors for all fish groups. An analogous argument can be made for species differences in target strength or acoustic backscatter per individual. A major advantage to this method is that the poorly understood uncertainties of combining acoustics with the bottom trawl are jointly embedded into the fitted $A_{sp}$ parameter. These various influences to the experimental method do not need to be assessed independently for all species present, which would be very challenging.

For the 6 years of data available to this study, pollock dominates the backscatter and the model predicted this result relatively well when the data from any year were excluded (i.e., different input data, see Table 2). This suggests that the model will perform well out of sample. However, the sensitivity analysis revealed that Arctic cod was not a significant contributor when the 2010 data were excluded. Arctic cod were only abundant in 2010, and if that year is excluded there is not a sufficient effect of Arctic cod on backscatter to result in a significant
coefficient for that species in the model. As with any predictive model, this illustrates that the training set used to develop the model needs to be representative of the conditions under which the model will be applied. If the method of using BT data to estimate biomass of pollock between 0.5 and 3 m off the bottom (as described above) is implemented in future AT surveys, the species catch composition should be monitored for significant changes. A situation could arise in which there is an increase in non-pollock species, such as Pacific herring (pelagic species that are strong sound scatterers) that is not captured in the full model fit presented in this paper. In this case, years with a substantial catch of these species would need to be included in the full model fit to accurately assess the significance of their contribution to acoustic backscatter. This caveat also applies to avoiding the use of this trained model for different geographical areas. However, the basic model framework can be used in different situations provided that suitable data are used to fit the parameters.

While there are many challenges associated with combining AT and BT methods in producing a combined index of abundance for semipelagic species (Jakobsen et al. 1997, Godø and Wespestad 1993), the proposed process can provide an estimate of the proportion of acoustic backscatter that is attributable to a given species in a relatively diverse demersal environment. The method has the advantage of a simple, empirically-derived compound parameter that incorporates all species present and accounts for the different biases inherent in the acoustic and bottom trawl survey methods. For example, Pacific cod is a relatively abundant species in the EBS BT survey (Fig. 3) and we suspected it to be one of the major contributors to acoustic backscatter, but it was not significant in this study and therefore is likely mostly distributed in the ADZ. This method is appropriate for use in any semipelagic habitat with available simultaneous acoustic and bottom trawl data and may help to refine other surveys that rely on
two independent data sets or those that attempt to combine AT and BT data from separate surveys.

Acknowledgements

We want to thank all the people who participated in the eastern Bering Sea bottom trawl survey, as this project was contingent on the availability of these data. We also thank David Somerton, Jim Ianelli, Chris Wilson, Jeff Napp, and Taina Honkalehto for reviews and discussions that improved the quality of this manuscript.

The findings and conclusions in the paper are those of the author(s) and do not necessarily represent the views of the National Marine Fisheries Service, NOAA.

Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA.
References


Table 1- Quantities used in the model

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>Model fit coefficients ($\text{#fish}^{-1}$)</td>
</tr>
<tr>
<td>$C$</td>
<td>Catch per unit effort data ($\text{#fish nmi}^{-2}$)</td>
</tr>
<tr>
<td>$I$</td>
<td>Intercept from model fit ($\text{m}^2 \text{nmi}^{-2}$)</td>
</tr>
<tr>
<td>$N_j$</td>
<td>Number of fish at length $j$ ($\text{#fish}$)</td>
</tr>
<tr>
<td>$S$</td>
<td>Area swept for bottom trawl ($\text{nmi}^2$)</td>
</tr>
<tr>
<td>$s_A$</td>
<td>Nautical area scattering coefficient ($\text{m}^2 \text{nmi}^{-2}$)</td>
</tr>
<tr>
<td>$B$</td>
<td>$s_A$ between 0.5 and 3 meters off bottom ($\text{m}^2 \text{nmi}^{-2}$)</td>
</tr>
<tr>
<td>$\sigma_{bs}$</td>
<td>Acoustic backscattering cross section ($\text{m}^2$)</td>
</tr>
<tr>
<td>$a$</td>
<td>Conversion factor from fish areal density to acoustic backscatter ($\text{#fish}^{-1}$)</td>
</tr>
<tr>
<td>$T$</td>
<td>Trawl catch scaled by length, $C$ multiplied by $L^2$ ($\text{#fish m}^2 \text{nmi}^{-2}$)</td>
</tr>
<tr>
<td>$b$</td>
<td>Y-intercept of the target strength relationship (dB re 1 $\text{m}^2$)</td>
</tr>
<tr>
<td>$i$</td>
<td>Index for haul</td>
</tr>
<tr>
<td>$j$</td>
<td>Index for length class</td>
</tr>
<tr>
<td>$sp$</td>
<td>Index for species</td>
</tr>
<tr>
<td>$N_t$</td>
<td>Number of tows</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>Error variance</td>
</tr>
<tr>
<td>$W$</td>
<td>Weight of each closest bottom trawl to the EDSU</td>
</tr>
<tr>
<td>$R_i$</td>
<td>Distance from BT stations to the EDSU (nmi)</td>
</tr>
<tr>
<td>$R_o$</td>
<td>Maximum range to use BT stations (25 nmi)</td>
</tr>
</tbody>
</table>
Table 2- Significant species group contributors to acoustic backscatter between 0.5 and 3 m above bottom, along with the intercept and the standard deviation. The fitted parameter values (second column) with the 95% confidence intervals (third and fourth columns) are given for each predicted value.

<table>
<thead>
<tr>
<th>Fitted quantities</th>
<th>Predicted values</th>
<th>Confidence intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2.50%</td>
</tr>
<tr>
<td>Pollock ($A_{pk}$)</td>
<td>2.52</td>
<td>2.21</td>
</tr>
<tr>
<td>Arctic cod ($A_{ac}$)</td>
<td>16.39</td>
<td>2.84</td>
</tr>
<tr>
<td>Large flatfish ($A_{lf}$)</td>
<td>0.85</td>
<td>0.16</td>
</tr>
<tr>
<td>Rockfish ($A_{rf}$)</td>
<td>93.59</td>
<td>8.63</td>
</tr>
<tr>
<td>Miscellaneous ($A_{mc}$)</td>
<td>11.63</td>
<td>3.92</td>
</tr>
<tr>
<td>Intercept ($I$)</td>
<td>3.43</td>
<td>1.95</td>
</tr>
<tr>
<td>Standard Deviation ($\sigma$)</td>
<td>1.02</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Table 3- Results from the LOOCV sensitivity analysis, where each value is the difference between the percent of total contribution to backscatter predicted for that year using data from all other years and the percent of total contribution to the backscatter for each year using data for every year (percentage for full model for each year – percentage of LOOCV model for each year).

<table>
<thead>
<tr>
<th></th>
<th>Pollock</th>
<th>Arctic cod</th>
<th>Large flatfishes</th>
<th>Rockfishes</th>
<th>Miscellaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>0.63</td>
<td>-3.34</td>
<td>-0.07</td>
<td>-0.68</td>
<td>3.44</td>
</tr>
<tr>
<td>2007</td>
<td>-0.70</td>
<td>-3.02</td>
<td>0.40</td>
<td>3.90</td>
<td>-0.58</td>
</tr>
<tr>
<td>2008</td>
<td>0.69</td>
<td>-1.84</td>
<td>-1.41</td>
<td>0.02</td>
<td>2.54</td>
</tr>
<tr>
<td>2009</td>
<td>0.68</td>
<td>-2.82</td>
<td>1.54</td>
<td>-0.86</td>
<td>1.46</td>
</tr>
<tr>
<td>2010</td>
<td>-16.52</td>
<td>14.87</td>
<td>0.42</td>
<td>-0.74</td>
<td>2.04</td>
</tr>
<tr>
<td>2011</td>
<td>1.90</td>
<td>-2.61</td>
<td>1.52</td>
<td>0.29</td>
<td>-1.10</td>
</tr>
</tbody>
</table>
Figure Captions:

Fig. 1- Bottom trawl stations for all years used in the analysis. Black dots show BT stations and polygons are bounds for the area covered by the AT survey. Only hauls within the AT survey area (those shown) were used for analysis.

Fig. 2- Example of a typical echogram examined by visual inspection to determine quality. The scale is volume backscattering strength with minimum display threshold of -70 dB re 1 m⁻¹. The black line to the left is the vertical distribution of backscatter throughout the water column. Two lines above the bottom mark 3 meters and 0.5 meter above the bottom detected bottom.

Fig. 3- Left: Total catch per unit effort ($C_{sp}$) for each species group (stacked) and for each year that data were used for the analysis. Right: The sum of quantity $T_{sp}$ (i.e., $C_{sp}$ expressed as fish surface area) for each species group (stacked) and for each year. Each unique shading pattern represents a different species group.

Fig. 4- Percent of the total backscatter contributed by each species that was significant in the likelihood model computed using eq. 16. Each unique shading pattern represents a different species group. Pollock is the primary contributor to the backscatter (compare with Fig. 3).

Fig. 5- Diagnostic plots for the model: (a) scatter plot of standardized residual scatter plot versus log(predicted); (b) scatter plot of the logarithm of the observed values versus the logarithm of the predicted values, with the line showing a $y = x$ relationship; (c) a Q-Q plot showing the sample quantiles (obtained from the residuals) versus the theoretical quantiles of normal distribution.

Fig. 6- Top panel: pollock biomass by year in million tons for the region between the surface and 3 m above bottom (solid line), the region between 0.5 and 3 m above bottom (dotted line), and the sum of these two regions (dashed line). Bottom panel: Percent increase from the region between the surface and 3 m above bottom to the total water column. This represents the change in pollock biomass not included in the acoustic survey in the past when acoustic backscatter between 0.5 and 3 m was not used.
Fig. 1- Bottom trawl stations for all years used in the analysis. Black dots show BT stations and polygons are bounds for the area covered by the AT survey. Only hauls within the AT survey area (those shown) were used for analysis.

254x190mm (300 x 300 DPI)
Fig. 2- Example of a typical echogram examined by visual inspection to determine quality. The scale is volume backscattering strength with minimum display threshold of -70 dB re 1 m$^{-1}$. The black line to the left is the vertical distribution of backscatter throughout the water column. Two lines above the bottom mark 3 meters and 0.5 meter above the bottom detected bottom.

254x190mm (300 x 300 DPI)
Fig. 3- Left: Total catch per unit effort ($C_{sp}$) for each species group (stacked) and for each year that data were used for the analysis. Right: The sum of quantity $T_{sp}$ (i.e., $C_{sp}$ expressed as fish surface area) for each species group (stacked) and for each year. Each unique shading pattern represents a different species group.
Fig. 4- Percent of the total backscatter contributed by each species that was significant in the likelihood model computed using eq. 16. Each unique shading pattern represents a different species group. Pollock is the primary contributor to the backscatter (compare with Fig. 3).
Fig. 5- Diagnostic plots for the model: (a) scatter plot of standardized residual scatter plot versus log(predicted); (b) scatter plot of the logarithm of the observed values versus the logarithm of the predicted values, with the line showing a y = x relationship; (c) a Q-Q plot showing the sample quantiles (obtained from the residuals) versus the theoretical quantiles of normal distribution.
Fig. 6- Top panel: pollock biomass by year in million tons for the region between the surface and 3 m above bottom (solid line), the region between 0.5 and 3 m above bottom (dotted line), and the sum of these two regions (dashed line). Bottom panel: Percent increase from the region between the surface and 3 m above bottom to the total water column. This represents the change in pollock biomass not included in the acoustic survey in the past when acoustic backscatter between 0.5 and 3 m was not used.