ICES Journal of Marine Science, 2023, **0**, 1–17
DOI: 10.1093/icesjms/fsad102
Original Article

A Bayesian inverse approach to identify and quantify organisms from fisheries acoustic data

Samuel S. Urmy [1,*], Alex De Robertis [1], and Christopher Bassett [2]

[1]Resource Assessment and Conservation Engineering Division, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Seattle, WA 98115, USA
[2]University of Washington Applied Physics Laboratory, 1013 NE 40th St, Seattle, WA 98105, USA
[*] *Corresponding author: tel: +1 206 526 4682*; e-mail: sam.urmy@noaa.gov.

Identifying sound-scattering organisms is a perennial challenge in fisheries acoustics. Most practitioners classify backscatter based on direct sampling, frequency-difference thresholds, and expert judgement, then echo-integrate at a single frequency. However, this approach struggles with species mixtures, and discards multi-frequency information when integrating. Inversion methods do not have these limitations, but are not widely used because species identifications are often ambiguous and the algorithms are complicated to implement. We address these shortcomings using a probabilistic, Bayesian inversion method. Like other inversion methods, it handles species mixtures, uses all available frequencies, and extends naturally to broadband signals. Unlike previous approaches, it leverages Bayesian priors to rigorously incorporate information from direct sampling and biological knowledge, constraining the inversion and reducing ambiguity in species identification. Because it is probabilistic, a well-specified model should not produce solutions that are both wrong and confident. The model is based on physical scattering processes, so its output is fully interpretable, unlike some machine learning methods. Finally, the approach can be implemented using existing Bayesian libraries and is easily parallelized for large datasets. We present examples using simulations and field data from the Gulf of Alaska, and discuss possible applications and extensions of the method.

Keywords: artificial intelligence, big data, broadband, echo integration, EK80, fish, mesopelagic, species identification, wideband.

Introduction

Active acoustics have been used for many decades to measure the distribution and abundance of aquatic animals. Acoustic techniques have compelling advantages over other methods of sampling marine life: they can sample large volumes of water quickly, non-extractively, and with high spatial resolution. However, because the information contained in an animal's echo is limited, acoustics also have a perennial challenge: identification of the sound-scattering organisms. The typical solution is to sample them using nets, using the catches to estimate species composition and allocate backscatter accordingly, using scattering models or empirically established relationships. If backscatter is dominated by a few scatterer types that are well-retained by trawls, this approach works well and is the basis for many long-running acoustic-trawl fisheries surveys (Simmonds and MacLennan, 2005).

Several current trends present this approach with new challenges, as well as new opportunities. One is an increased interest in "ecosystem-based" surveys, targeting multiple species rather than a single dominant one. Another is the use of platforms besides research vessels, including autonomous underwater vehicles (AUVs, Benoit-Bird *et al.*, 2018), uncrewed surface vehicles (USVs, De Robertis *et al.*, 2021), vessels of opportunity (Honkalehto *et al.*, 2011), and stationary moorings and observatories (Urmy *et al.*, 2012; De Robertis *et al.*, 2017). These platforms offer expanded spatial coverage and temporal persistence compared to crewed vessels, at the cost of reduced or absent direct sampling. At the same time, technological developments such as remote camera systems (Robison, 1999; Williams *et al.*, 2010; Reisenbichler *et al.*, 2016), shadowgraph cameras (Ohman *et al.*, 2019), and environmental DNA (eDNA, Berger *et al.*, 2020; Shelton *et al.*, 2022) can all provide information on species composition, but have different biases and have seen only limited integration with acoustics. Finally, the advent of commercially available broadband (BB) echosounders should offer more information from the echoes themselves, with improved ability to identify scatterers remotely (Bassett *et al.*, 2018; Benoit-Bird and Waluk, 2020; Cotter *et al.*, 2021b). However, the increased information in BB data comes with corresponding increases in data volume and processing complexity, for which standard analysis procedures are still developing. Despite ongoing research on these new approaches (and the fact that nets can have their own detectability and selectivity biases, e.g. Williams *et al.*, 2015; De Robertis *et al.*, 2023) the traditional acoustic-trawl methodology remains the basis of almost every operational fisheries acoustic survey.

Even though direct sampling remains indispensable, there is also a long history of attempts at remote species identification using acoustics (Horne, 2000). Of these techniques, the most widely used is classification via the relative frequency response, or "frequency differencing" (Korneliussen *et al.*, 2018). This involves comparing the ratio of backscatter across frequencies. If the scatterers' frequency-specific target strengths (TS) are known, they can be used to define thresholds for frequency-dependent scattering that can be used for classification. This approach is simple to implement and works well for scatterers with distinct frequency responses (e.g. fish and zooplankton), but also has downsides. For one, choosing a threshold imposes an arbitrary cutoff, which may not be suitable in uncertain cases (though probabilistic extensions are possible, e.g. De Robertis *et al.*, 2010). In addition,

Received: 27 February 2023; **Revised:** 19 May 2023; **Accepted:** 22 May 2023

frequency differencing has to assume that scattering at each location is dominated by a single type of animal, and therefore is ill-suited to mixed aggregations of scatterers. More recently, a variety of machine learning and artificial intelligence (ML/AI) approaches have been applied to the problem (Roberts *et al.*, 2011; Brautaset *et al.*, 2020; Sarr *et al.*, 2021). The promise of these techniques is that with big-enough data, advanced algorithms will be able to detect subtle patterns that humans and simple models cannot. However, supervised ML requires labour-intensive manual labelling, and the trained model typically operates as a difficult-to-understand black box, which may rely on contextual information from the training dataset, rather than the actual features of the scatterers. For instance, neural networks for image recognition learn to expect "sheep" to appear as white patches on a green field. If the sheep are coloured orange, they are instead labelled as "flowers," and white rocks may be labelled "sheep." If the neural network is shown an image of a sheep in a person's arms, it calls it a dog; a sheep in water may become a polar bear, etc. (Shane, 2018). Unsupervised ML approaches do not require manual training and can discover new classes of scattering (Woillez *et al.*, 2012), but cannot easily tell the difference between a common mixture and an "unidentified scatterer." Ultimately, both frequency differencing and ML/AI are *classification* approaches, using multifrequency information to isolate areas of interest on an echogram, which are then integrated at a single frequency.

Instead of classifying and integrating sequentially, one can do both simultaneously by solving the "inverse problem"–i.e. given some observed backscatter, which animals, at what densities, actually generated it? Inverse approaches in fisheries and plankton acoustics have a long history (Holliday *et al.*, 1989; Stanton *et al.*, 2012; Chu *et al.*, 2016) and provide several compelling advantages. They can handle mixed scatterers, which are problematic for frequency-response classifiers and, unlike ML, they are based on transparent physics-based scattering processes. Inverse methods extend naturally from narrowband (NB) to BB data, and in some cases can be used to infer properties of individual scatterers, such as orientation and size (Conti and Demer, 2006; Stanton *et al.*, 2012; Loranger *et al.*, 2022). However, the inverse approach is seldom used in practical applications due to several challenges. It is more complex to implement than frequency differencing, and not available as a tool in most standard acoustic-processing software. More importantly, the inverse problem is often underdetermined, with more unknowns than equations, yielding ambiguous solutions. Even when solvable in principle, inversions of field data always involve some uncertainty. Besides one linearized approximate method (Chu *et al.*, 2016), the issue of uncertainty has been mostly ignored in the biological inversion literature. Without a practical approach to constrain these ill-posed problems and quantify their inherent uncertainties, the potential benefits of the inverse approach for fisheries acoustics will remain unrealized.

We address these challenges by implementing the inverse scattering problem as a Bayesian statistical model. Bayesian statistics have been used in geoacoustic inversion and sound propagation applications for decades (Gerstoft and Mecklenbräuker, 1998; Dosso, 2002), and are a natural fit for biological scattering as well. For this application, two features are especially worth highlighting. First, Bayesian inference allows (indeed, requires) that a *prior distribution* be specified for all parameters. The prior provides a natural place to impose constraints on an underdetermined inverse problem, leveraging biological knowledge, direct samples, or both. The "subjectivity" of the prior is sometimes perceived as a disadvantage of Bayesian statistics. However, we see it as a clear advantage: interpretation of acoustic backscatter always requires assumptions and auxiliary information, and expressing them as priors makes them explicit. In fact, as we illustrate below, it allows for incorporation of diverse kinds of auxiliary information beyond traditional research trawl catches. Second, instead of point estimates, a Bayesian model returns *posterior probability distributions* for all parameters being inferred. For an inverse biological scattering problem, these posteriors provide estimates of animal densities, incorporating all sources of uncertainty built into the model. Crucially, making the inverse scattering problem probabilistic makes it robust enough to run automatically on large volumes of data. The prior acts like a shock absorber for the solution, and the posterior provides a rigorous precision estimate. If the model is adequately specified, it should be honest about the quality of its solution: in other words, solutions should not be wrong and highly confident at the same time.

Bayesian inference has a reputation for complexity and computational expense, but modern probabilistic programming packages (Patil *et al.*, 2010; Gelman *et al.*, 2015; Ge *et al.*, 2018) make it simpler to implement, as well as flexible and extensible. It is possible, for instance, to include a physical acoustic scattering model inside a Bayesian statistical model, and perform inference on its parameters. By leveraging numerical techniques like Hamiltonian Monte Carlo (HMC, Betancourt and Girolami, 2013; Hoffman and Gelman, 2014) and variational inference (VI, Blei *et al.*, 2017), Bayesian models can also be run efficiently. In fact, because backscatter from a volume of water depends only on the scatterers in that volume, inversions can be run in parallel on an arbitrarily large number of processors (e.g. Regier *et al.*, 2018, 2019).

We begin this paper with a brief theoretical overview of the inverse problem for biological backscatter, followed by a description of how to solve it using freely available software libraries for Bayesian computation, in an implementation we call "automatic probabilistic echo solving" (APES). We then show how the method works using two simulated scenarios, one simple and one more complex, illustrating the effects of uncertainty, the number of acoustic frequencies, prior information, and ground truth on the model's solutions. Finally, we apply the Bayesian inverse approach to NB and BB data collected as part of a 2021 acoustic-trawl survey of Alaska pollock (*Gadus chalcogrammus*) in the Gulf of Alaska to demonstrate how probabilistic models allow the inverse approach to be applied at scale in the field.

Methods

Model structure

APES is simply a probabilistic implementation of a model of linear, non-coherent backscatter from aggregated targets. These conditions hold in most acoustic surveys of the water column, and are the central assumption behind all echo-integration surveys (Foote, 1983). If an ensonified volume of water V contains N individual scatterers at numerical density $n = N/V$, the mean volume backscattering coefficient (MVBC) s_v is equal to n times the average backscattering

cross-section $\langle\sigma_{bs}\rangle$, where the brackets denote a statistical average over all scatterers. (Backscattering cross-sections are often expressed logarithmically as $TS = 10\log_{10}(\sigma_{bs})$, MacLennan *et al.*, 2002.) Because this paper only deals with backscatter, we will drop the subscript "bs" for clarity. This relationship, $s_v = \langle\sigma\rangle\, n$, allows animal density to be estimated from integrated echo energy.

In general, $\langle\sigma\rangle$ for an individual scatterer depends on its identity and the acoustic frequency, so for a volume containing S types of scatterers ensonified at frequency f, this equation can be expanded as

$$s_{v,f} = \langle\sigma_{1,f}\rangle\, n_1 + \langle\sigma_{2,f}\rangle\, n_2 + \ \ldots\ + \langle\sigma_{j,f}\rangle\, n_j + \ \ldots\ + \langle\sigma_{S,f}\rangle\, n_S, \tag{1}$$

where $s_{v,f}$ is the MVBC at frequency f, $\langle\sigma_{j,f}\rangle$ is the mean backscattering cross-section for scatterer class j at frequency f, and n_j is the numerical density of the j^{th} scatterer class. Scatterer classes could be different species, different size classes of the same species, or abiotic sources of scatter such as bubbles, sediment, or turbulent microstructure (Monahan and Lu, 1990; Warren *et al.*, 2003). For a survey operating at F frequencies, (1) can be written more compactly in matrix form as

$$\boldsymbol{s_v} = \begin{bmatrix} \langle\sigma_{1,1}\rangle & \cdots & \langle\sigma_{1,S}\rangle \\ \vdots & \ddots & \vdots \\ \langle\sigma_{F,1,}\rangle & \cdots & \langle\sigma_{F,S}\rangle \end{bmatrix} \boldsymbol{n} \ = \ \Sigma\, \boldsymbol{n}, \tag{2}$$

where $\boldsymbol{s}_v$ is a length-F vector of frequency-specific MVBCs, $\boldsymbol{n}$ is a length-S vector of scatterer densities, and Σ is an $F \times S$ matrix whose i^{th}, j^{th} element is the backscattering cross-section of scatterer class j at frequency i. In other words, each column of Σ corresponds to the backscattering spectrum of a single class of scatterer. If we measure $\boldsymbol{s}_v$ and know (or assume) Σ, solving (2) for $\boldsymbol{n} \geq 0$ defines the basic inverse problem. Equation 2 is equally applicable to single-, multi-frequency, or BB data, though it will not have a unique solution if the number of scatterer classes is greater than the number of frequencies, or if multiple scatterers have collinear backscattering spectra–i.e. in linear algebra terms, if the scattering matrix Σ is rank-deficient.

In practice, scattering is influenced by numerous random processes and solutions to the inverse problem inevitably contain uncertainties, even if the equation is fully determined in theory. To account for these uncertainties, we embed it into a Bayesian probabilistic model. At its simplest, this model takes the form

$$\begin{aligned} \log_{10}(\mathbf{n}) &= [\text{Vector of prior distributions}] \\ \omega &\sim \text{Nonnegative prior on observation error variance} \\ \boldsymbol{\mu} &= 10\log_{10}(\Sigma\ \mathbf{n}) \\ \mathbf{S}_v &\sim \text{Normal}(\boldsymbol{\mu}, \omega) \end{aligned}$$

where $\sim$ means "is distributed as," indicating the quantity on the left-hand side is a random variable. Note that we formulate the model in terms of scatterer log-densities and decibel volume backscattering strengths S_v (MacLennan *et al.*, 2002). This avoids issues with numerical precision (Goldberg, 1991), which are especially problematic for gradient-based optimizers and samplers, causing slow or inaccurate model convergence. Additionally, because individual s_v measurements from aggregated scatterers are approximately Rayleigh distributed (Stanton *et al.*, 2018) their *SD* increases proportionally to scatterer density. Expressing the model logarithmically makes these multiplicative errors additive. By the central limit theorem, the assumption of normal observation errors is acceptable if enough individual samples are averaged into each cell.

A key decision when setting up the model is how to define the scattering matrix Σ. In the case where each scatterer has a well-defined TS spectrum, these simply specify the columns of the matrix. If any of these TS spectra are uncertain, we could make Σ a random variable within the model, which will propagate this uncertainty to the posterior estimates for $\boldsymbol{n}$. One or more of these spectra could also be modelled as functions $g_i(\boldsymbol{\theta}_i)$ of physical parameters (such as body shape, swimbladder size, tilt angle, material properties, etc.), denoted here by a generic parameter vector $\boldsymbol{\theta}$. This scattering model can then be incorporated as another component to the statistical model above, e.g.

$$\begin{aligned} \boldsymbol{\theta}_i &\sim \text{Prior distribution} \\ \langle\boldsymbol{\sigma}_i\rangle &= g_i(\boldsymbol{\theta}_i),\ \ i = 1, 2, \ldots, S \\ \Sigma &= [\langle\boldsymbol{\sigma}_1\rangle\langle\boldsymbol{\sigma}_2\rangle \ \ldots\ \langle\boldsymbol{\sigma}_S\rangle], \end{aligned}$$

where each column vector $\langle\boldsymbol{\sigma}\rangle_i$ is the average backscattering cross-section of scatterer class i across all frequencies. This type of model allows us to do inference on the properties of the scatterers themselves in addition to their densities in the water.

A second key decision is how to set the priors. This is especially important when Equation 2 is underdetermined. In such situations, which are common, an informed prior may be the only way to constrain an otherwise unsolvable problem. Prior distributions (including their means, variances, shapes, and extreme values) may be specified based on past measurements, literature values, or educated guesses based on contextual information. They can also be set based on direct samples or other *in-situ* data. Which of these methods is most appropriate will depend on the particular situation, the data available, and the analyst's judgement.

Model implementation

All models and analyses were implemented in the Julia programming language (Bezanson *et al.*, 2012, 2017) using the Turing.jl library for probabilistic programming (Ge *et al.*, 2018). Turing provides a concise syntax for defining Bayesian models and tools for fitting them to data, including Markov-chain Monte Carlo (MCMC) samplers and maximum a-posteriori (MAP) optimizers. Because it is written in Julia, Turing models can include any Julia code and run efficient inference on it by exploiting Julia's native capabilities for automatic differentiation. This is an advantage in acoustics, since it allows models to include arbitrarily complex scattering models (e.g. modal solutions for spheres and bubbles, distorted-wave Born approximations (DWBAs) including numerical integration, etc., Stanton and Chu, 2000; Jech *et al.*, 2015). Additionally, because the models are specified cell-by-cell, they can be fit in parallel on as many computer cores as desired, whether locally or in the cloud. We fit our models via MCMC using the highly-efficient no-U-turn sampler (NUTS, Hoffman and Gelman, 2014) with the standard target acceptance rate of 0.8 and 1 000 samples for adaptation and convergence. For each model, we sampled four replicate chains and checked them for convergence using the $\hat{R}$ statistic,

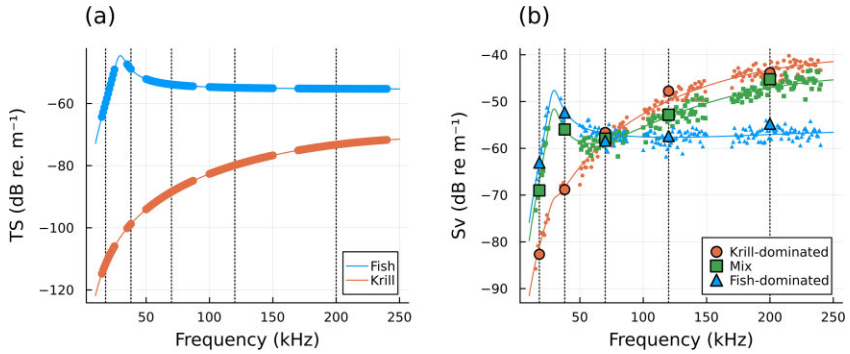

Figure 1. Simulation of mixed backscatter from larval fish and krill. (A) Target-strength spectra assumed for fish and krill (fine lines). Points show the frequencies at which BB measurements were made (at 1 kHz intervals), and the vertical dotted lines show the five NB frequencies. (B) Simulated volume-backscatter spectra, including stochastic noise, for three scenarios: one where backscatter is dominated by krill, one where both krill and fish make significant contributions, and one where backscatter is dominated by fish. Lines indicate the theoretical noise-free spectra, with points showing the noisy observations simulated for the BB (small symbols) and NB (large symbols) scenarios.

which compares within-chain variance to variance between chains (Vehtari *et al.*, 2021). A package implementing APES is available at https://github.com/ElOceanografo/ProbabilisticEchoInversion.jl, and scripts for all analyses presented here are available at https://github.com/ElOceanografo/APESExamples.

Simulated case studies

To test the probabilistic inverse method in a controlled setting where the true values of all quantities are known, we applied it to backscatter spectra simulated from several assumed mixtures of biological scatterers. This allowed us to see the effects of different mixture compositions, and of adding different types of auxiliary information.

Fish and krill

The first test problems were simulated mixtures of larval fish and euphausiids. Discriminating fish from zooplankton is perhaps the most common use of frequency differencing (Korneliussen *et al.*, 2018), but it is not always reliable for mixed-species aggregations (De Robertis *et al.*, 2010). In particular, backscatter from weakly scattering targets such as zooplankton may be mis-classified in the presence of stronger acoustic targets. In contrast, a full inverse method should (theoretically) be able to recover the densities of both animals. In many regions, mixtures of larval fish and zooplankton are relatively common and may be important ecologically, but are typically under-sampled by large-mesh research trawls targeting adult fish.

We assumed that the TS of a larval fish would be dominated by scattering from its swimbladder, approximated as a damped bubble with equivalent spherical radius (ESR) $a = 0.7$ mm and damping parameter $\delta = 0.3$, located at 200 m depth, giving it a resonant frequency of 22 kHz (Love, 1978). Krill TS was modelled based on a fluid-like deformed cylinder using the DWBA (McGehee *et al.*, 1998; Stanton and Chu, 2000). We assumed the same krill shape as McGehee *et al.* (1998), but reduced the body length to 25 mm to better match the size of krill encountered during our surveys, scaling down all dimensions proportionally. We also changed its sound-speed and density contrasts to 1.032 and 1.019, the average values for krill in Alaska waters (Lucca *et al.*, 2021). The DWBA implementation was implemented in the Julia package SDWBA.jl (Urmy, 2016). To simulate the stochastic variability present in real backscatter, for each observation, we generated three random draws at each frequency from a Rayleigh distribution with expected value of $\Sigma\, \boldsymbol{n}$, then took their average. This approximates the processes generating volume backscatter from a volume containing many animals (Stanton *et al.*, 2018), without the complexity of explicitly simulating the variability in individual TS due to variability in size, shape, material properties, tilt, etc. This simplification is justified for most echo integration applications, where the backscatter values being integrated are averages over large volumes of water containing many scatterers.

We simulated volume backscattering spectra for three different mixtures of fish and krill. In the first scenario, we assumed densities of 0.001 fish m^{-3} and 1 000 krill m^{-3}, making krill the dominant source of backscatter. The second scenario assumed 0.2 fish and 400 krill m^{-3}, for relatively equal backscatter contributions. The third scenario assumed 0.5 fish m^{-3} and 20 krill m^{-3}, making fish the dominant scatterer. Backscatter from each of these mixtures was simulated at 18, 38, 70, and 120 kHz, the NB frequencies traditionally used in fisheries acoustics. It was also observed at 1 kHz intervals across four frequency ranges (15–25, 35–38, 50–87, 100–150, and 170−240 kHz), representing the analysis bands achievable with a modern wideband echosounder (Bassett et al., 2018). The simulated TS and S_v spectra are shown in Figure 1.

To invert these three scenarios, we set up a model with two candidate scatterers. The first was krill with a TS spectrum fixed at the modelled mean value (Figure 1A). The second scatterer was a larval fish represented by a bubble with $\delta = 0.3$, but with ESR left as an unknown free parameter (i.e. corresponding to the generic function g in Equation 2). We gave the bubble radius a normal prior centred at 0.5 mm, +/− 0.2 mm, and truncated below 0.1 and above 1.5 mm. Fish and krill densities were given log-normal priors with median values of 0.1 and 200 animals m^{-3}, respectively–i.e. broad but physically realistic values. The coefficient of variation for the residual error was given an exponential prior with an expected value 0.1. The model was thus attempting to infer four unknown parameters: two animal densities, one animal size, and an error dispersion. We fit the model to the NB and BB data from each simulated mixture. Since there were data at four or more frequencies in all cases, this problem was well-posed and solvable in the absence of noise. For each inversion we drew 2 500

posterior samples from each of four parallel Markov chains, using NUTS as described above.

Mesopelagic mixture

The second set of test scenarios simulated a more challenging problem: a mesopelagic deep-scattering layer (DSL). Such layers are found worldwide and contain a potentially enormous biomass of fish (Irigoien *et al.*, 2014) although there are considerable uncertainties in those estimates (Proud *et al.*, 2018; Cotter *et al.*, 2021a). DSLs are typically composed of multiple scatterers in addition to fishes, including crustaceans, cephalopods, and gelatinous animals. Some, like most crustaceans and gelatinous animals, are relatively weak sound-scatterers, but others, like squids (Kawabata, 2005; Soule *et al.*, 2010) and physonect (i.e. gas-bearing) siphonophores (Barham, 1963; Proud *et al.*, 2018) can scatter sound strongly in a similar manner to swimbladdered fishes.

Untangling this mixed community with multiple similar TS spectra is made more challenging by the fact that no single direct-sampling gear captures all of them effectively. Midwater trawls retain fishes, squid, and larger crustaceans with size- and species-dependent efficiency (Kwong *et al.*, 2018), but are poor samplers of fragile gelatinous animals. Video cameras can count gelatinous animals and crustaceans (Robison, 1999), but not fishes or squid, which may be repelled (or attracted) by the vehicles on which the cameras are mounted (Stoner *et al.*, 2008). Finally, new approaches like environmental DNA (eDNA) can detect the presence of almost any taxon within a water mass, though it is still unclear to what extent the number of reads can be interpreted as a measure of abundance (but see Shelton *et al.*, 2022).

We simulated a mixture of animals commonly found in DSLs in the California Current, where the U.S. National Oceanic and Atmospheric Administration (NOAA) and the Canadian Department of Fisheries and Oceans conduct a biannual acoustic-trawl survey of Pacific hake (*Merluccius productus)*. Hake are commonly found in mesopelagic scattering layers during the daytime, along with other species such as squid, myctophids, and sergestid shrimp (Barham, 1957, 1963; Omori and Gluck, 1979). The TS spectrum of hake was interpolated between NB values for a generic large swimbladdered fish (De Robertis *et al.*, 2010). We modelled myctophids and siphonophores as damped bubbles with radii $a = 0.65$ and 0.58 mm, and $\delta = 0.3$ and 0.2, respectively. The squid were assumed to be 20 cm long, and their TS spectrum was approximated from that in Jones *et al.*, (2009). Sergestid shrimp were modelled using the McGehee (1998) deformed-cylinder krill model, this time scaled up to 4.5 cm body length. We simulated volume backscatter from a mixture of 0.0025 hake, 0.1 myctophids, 1.7 shrimp, and 0.07 siphonophores m^{-3}. Note that squid were not included in this mixture, i.e. their density was zero. Again, we simulated backscatter from the mixture at four NB frequencies and their corresponding wideband frequency ranges as the mean of random Rayleigh variables with expected value $\Sigma \boldsymbol{n}$, this time using 10 draws. The TS and S_v spectra for this simulation are shown in Figure 2.

For this problem, we assumed that the TS spectra of the five scatterers are fixed and known, unlike in the first problem, where one of them was a function of a free parameter. However, this problem had five scatterers and only four NB frequencies, so it was ill-posed, at least as a deterministic inversion in the NB case. We demonstrate several ways the solution could be constrained with the addition of different types of auxiliary information. For each of these scenarios, we ran the inverse problem using both NB and BB frequencies, drawing 1 000 MCMC samples from each of 4 chains, again using the standard NUTS sampler.

In the first scenario, the only constraints were prior distributions on the densities of all five scattering classes. We specified vague, but physically plausible, log-normal priors for all scatterers–i.e. normal distributions in the log domain, with means set at the common (base-10) logarithm of the true density and *SD* of 2 (i.e. two orders of magnitude in the linear domain). For squid, whose true density was 0, we set the prior mean at $\log_{10}(0.25)$, the same as hake. The imposition of even vague prior distributions is enough to make the problem numerically solvable, though it does not guarantee the solution will be accurate or informative.

In the second scenario, we assumed information was available from a video survey conducted by a remotely operated vehicle or autonomous underwater vehicle. Video surveys can reliably quantify slow-moving siphonophores and sergestid shrimp, but not faster swimmers like hake, myctophids, and squid. We assumed that the video survey estimated densities of 0.07 siphonophores m^{-3} and 1.7 sergestids m^{-3}, with 15% coefficient of variation (c.v.). We used these estimates to set the siphonophore and sergestid priors in the log-domain as normal variables centred at $\log_{10}(0.07)$ and $\log_{10}(1.7)$, with *SD*s of 0.06. All other priors were as in scenario 1.

The third scenario used constraints from a different hypothetical sampling method: environmental DNA analysis of water samples. Environmental DNA analyses detect the presence or absence of all species for which the appropriate primers are available, though it is not clear if they yield reliable estimates of absolute abundance. Here, we assumed the eDNA analysis revealed evidence for the presence of all scatterers except squid. This information then allowed us to set the prior for squid to approximately zero (-9.0 +/− 0.01 in the log domain). Again, all other priors were the same as in scenario 1.

The final scenario assumed that both the video survey and eDNA sample were available, constraining the siphonophore, sergestid, and squid priors. The hake and myctophid priors were the same as in scenario 1.

Application to field data

To demonstrate the utility of APES in real-world applications, we applied the method to data collected during the Alaska Fisheries Science Center's acoustic-trawl survey of the Gulf of Alaska in boreal summer 2021. This survey was conducted aboard the NOAA Ship *Oscar Dyson*, a 64 m noise-reduced fisheries research vessel, operating a 5-frequency (18, 38, 70, 120, and 200 kHz) Simrad EK80 scientific echosounder system (Kongsberg Maritime AS, Horten, Norway). This system was operated both in NB and wideband mode with all frequencies operating simultaneously; operating parameters are given in Table 1. Wideband pulses were linear frequency modulated up-sweeps with the first and last 10% of the signal tapered by a Hann window ("fast ramping" in EK80 settings). Non-linear cross-talk between BB channels is a potential issue for multi-channel BB echosounders transmitting simultaneously (Khodabandeloo *et al.*, 2021). While its impact on volume-scattering (as opposed to TS) measurements is generally small (<10% even in extreme cases), we minimized the

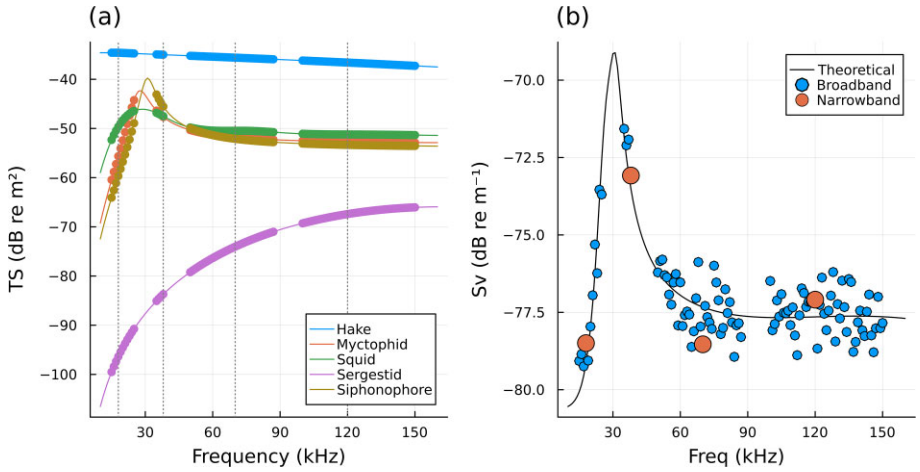

Figure 2. Simulation of backscatter from a mixture of mesopelagic organisms. (a) TS spectra of the five candidate scatterers. Thin lines show the average TS of each animal as a function of frequency. Points show the frequencies at which BB measurements were made, and the vertical dotted lines show the four NB frequencies. (b) Simulated volume backscattering spectrum from the mixed assemblage. The black line shows the theoretical noise-free backscatter spectrum, while small blue and large orange symbols show the noisy observations made at BB and NB frequencies, respectively.

Table 1. Settings and parameters for the EK80 acoustic system. Simultaneous pinging was used in both NB and BB configurations.

Frequency band (kHz)	Beam width (deg)	NB pulse length (ms)	NB transmit power (W)	BB pulse length (ms)	BB transmit power (W)
18	7	1.024	1000	—	—
38	7	0.512	2000	1.024	200
70	7	0.512	750	1.024	75
120	7	0.512	250	1.024	50
200	7	0.512	105	1.024	150

potential for this type of interference by restricting transmit power to low levels when in BB mode (Khodabandeloo *et al.*, 2021). All echosounder frequencies were calibrated using the standard-sphere method (Demer *et al.*, 2015) in both NB and wideband mode before and after the survey. During data collection, the ship travelled at a nominal speed of 12 kt (6.2 m s^{-1}) with the transducers deployed on its drop keel at a depth of 9.15 m.

Fitting models to logarithmic data requires care in how low backscatter values are handled. Typical practice is to set an arbitrary threshold (e.g. −70 dB re m^{-1}) and ignore values below it. This causes problems for inverse methods, since low backscatter values at one or more frequencies contain valuable information: any class that scatters sound in excess of measurement thresholds at those frequencies cannot be present. However, it is also important to exclude data where biological backscatter is near or below the level of background noise, since this will produce erroneous solutions. The general rule is not to threshold data, but to replace pixels with a low signal to noise ratio (SNR) with missing values. These are easy to account for in the model: the missing frequency is simply skipped when computing the log-likelihood, contributing no information to the inference. We quality controlled, preprocessed, and spatially averaged the raw NB acoustic data using Echoview version 11 (Echoview Software Pty Ltd, Hobart). Background noise was estimated and removed (De Robertis and Higginbottom, 2007), with pixels falling below 10 dB SNR designated as missing data. Noise spikes were manually removed, and the automatic bottom detection was checked and corrected where necessary.

Broadband acoustic data were processed to produce calibrated volume backscattering data following the approach described in Bassett *et al.*, 2017. Using the power budget equations in Demer *et al.* (2017), volume backscattering spectra were calculated for each of the five-frequencies using 1 m windows zero-padded to N = 512 points. Prior to zero-padding a Tukey window with a 10% cosine fraction was applied to the average range-compensated data from the individual echosounder sectors. Further processing steps included the Fourier transform and compensation for the power budget terms. The band edges were trimmed from the channels such that spectral bands 35–43 kHz, 46–86 kHz, 95–157 kHz, and 165–255 kHz were retained for the 38, 70, 120 and 200 kHz channels. After trimming, these spectra were combined into a single vector covering the full bandwidth at each window range and for each ping. Individual volume scattering spectra were calculated throughout the water column with 50% overlap (every 50 cm). Before fitting the inverse models, spectra were averaged into coarser spatial cells 5 m deep and 2 minutes in duration.

Midwater trawls were deployed regularly along the survey transects to confirm the identity of the scatterers and to provide length-frequency data for TS calculations. We used an LFS 1421 midwater trawl with 76.8 m headrope and footrope; its vertical opening was monitored using a Simrad FS70 netsonde and averaged 16.4 m while fishing. The trawl's meshes tapered from 650 cm at the headrope to 3.8 cm just forward of the cod end. The cod end itself was fitted with a knotless 3.2 mm nylon liner to increase retention of small organisms. Average speed during trawling was 1.8 m s^{-1} (3.4 knots).

Application at large scale: Aleutian shelf break

Our first dataset was taken from a south-north transect near Unimak Island in the eastern Aleutian Chain. This section began at 05:27 local time over the Aleutian Trench and ended on the continental shelf at 06:20, approximately 18 km north of its starting point (Figure 3). We chose it as a demonstration problem because it contained several partially overlapping scattering layers with different frequency responses. These

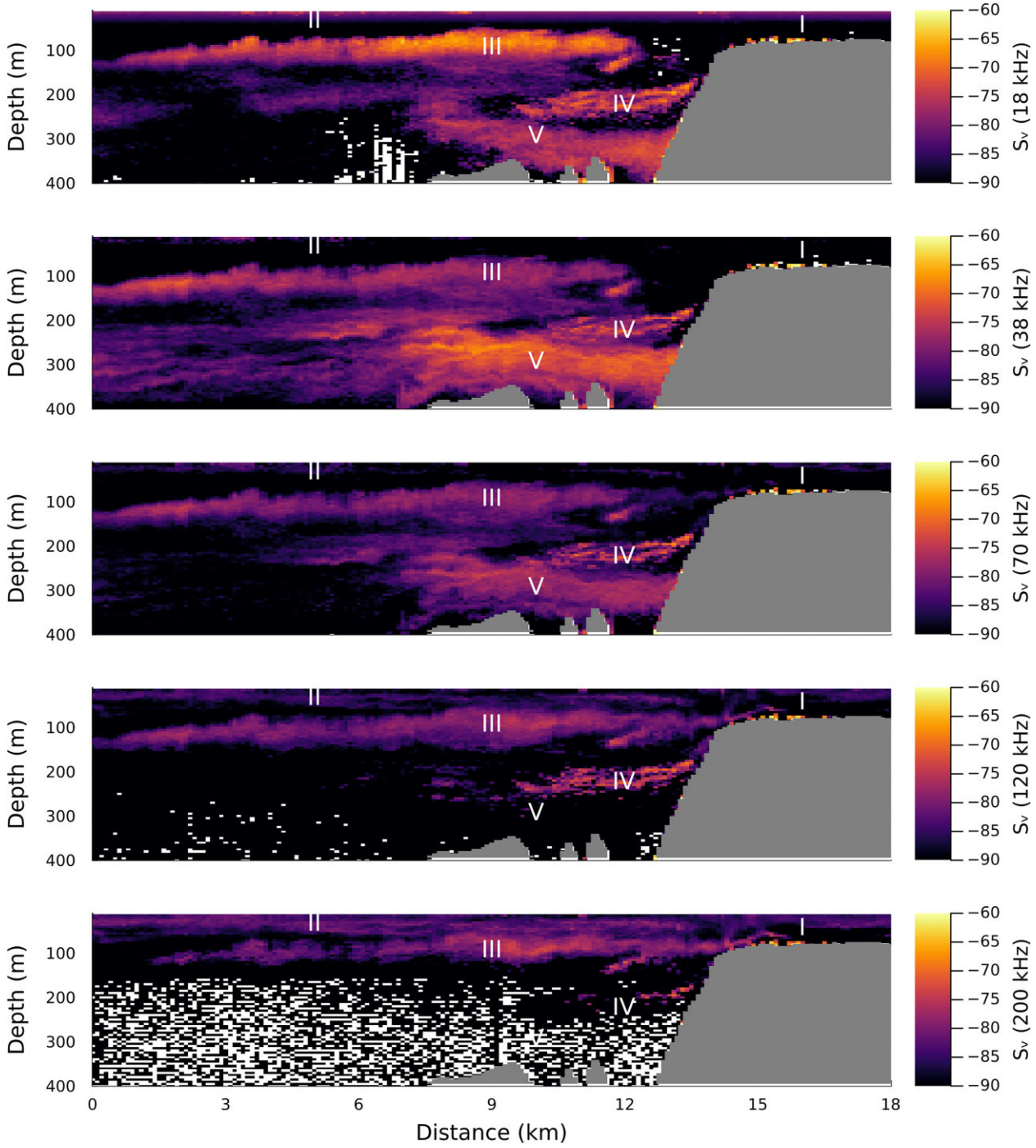

Figure 3. Observed volume backscatter (in dB re m^{-1}) at five NB frequencies (from top to bottom, 18, 38, 70, 120, and 200 kHz) near the shelf break south of Unimak Island in the Gulf of Alaska. White pixels indicate areas of low signal-to-noise ratio; these are treated as missing data in the inverse models. Several aggregations and scattering layers are labelled with numerals: (I) strong, near-bottom backscatter at all frequencies on the shelf, (II) a surface layer, highest at 200 kHz, (III) a layer near 100 m depth, strongest at 18 kHz, (IV) a slope-associated aggregation visible at all frequencies, and (V) a deep scattering layer strongest at 38 kHz.

data were collected in NB mode; after de-noising and quality control as described above, volume backscatter was averaged into cells 100 m along-track by 5 m depth.

A trawl targeting a high-backscatter aggregation just inshore of the shelf break confirmed that it was a mixed aggregation composed of ~85% pollock and 15% Pacific ocean perch (*Sebastes alutus*) (Levine *et al.*, 2022). Since these species have similar backscattering spectra (De Robertis *et al.*, 2010), we opted to include a single "large swimbladdered fish" spectrum as the model's first candidate scatterer type. A Methot trawl (Methot, 1986) targeting a zooplankton aggregation on the shelf on the previous transect found it was composed of euphausiids with mean length 24 mm, so the second candidate scatterer included in the model was a krill spectrum based on the same DWBA model as in the first simulation, with its body length scaled to 24 mm. Since these data also covered

mesopelagic depths off the shelf, we included a "myctophid-like" scattering spectrum based on the published frequency response of confirmed myctophid aggregations in this area (De Robertis *et al.*, 2010).

Finally, preliminary inspection of the data indicated that portions of the water column were occupied by spatially diffuse scatterers with a frequency response stronger at 18 kHz than the higher frequencies–a type of scattering that is often observed in the Gulf of Alaska and Bering sea (De Robertis *et al.*, 2010; Woillez *et al.*, 2012). These scatterers remain poorly characterized and are rarely analysed explicitly. Such a frequency response could be generated by a resonant scatterer with its resonant frequency near 18 kHz, so we included a bubble-like scatterer, using the same model as above, with its radius (and therefore its resonant frequency) as a free parameter. The depth of the bubble was set to the depth of the acoustic cell.

We set the priors for the densities of these four scatterers to log-normal distributions with mean abundance in the log domain of −3, −2, −2, and −2 individuals m^{-3}, respectively, and *SD*s of 3 (i.e. 3 orders of magnitude in the linear domain). The bubble radius was given a uniform prior between 0.01 and 2 mm.

Broadband inversion: Barnabas Trough

Though the echosounders were operated in NB mode for most of the 2021 survey, we collected a small-scale dataset in wideband mode in Barnabas Trough, Alaska (Figure 4), on the south shore of Kodiak Island. Barnabas Trough begins in Ugak Bay, and its main channel (∼25 km long, 8 km wide at its mouth, and 100 m deep) extends southwards to the shelf break. We ran four transects in a zig-zag pattern, starting in upper Ugak Bay at 22:30 local time on June 19, 2021, and finished in Barnabas trough at 2:00 on June 20. Earlier that day, as we surveyed northwards, we conducted two research trawls in upper Barnabas trough at 57.279°N, 152.523°W and 57.442°N, 152.643°W, targeting dense schools in the middle of the water column at ∼50 m depth. Both caught mixtures of pollock and Pacific herring (*Clupea pallasii*): the first 68% pollock and 32% herring, and the second 89% pollock and 11% herring.

Based on these results, we defined the inverse model with two fixed scattering spectra, for a generic swimbladdered fish and a 24 mm euphausiid, as well as a variable-radius spherical bubble. Myctophids were not included in this model because they are not found in abundance on the shelf at these shallow depths. The priors for the densities of these scatterers were set as in the shelf break case above.

Fitting models to real BB data required several changes from the NB procedures. For one, because the width of each transducer's beam narrows as frequency increases, volume backscatter from single off-axis targets appears to decline with increasing frequency (Medwin and Clay, 1998). This can introduce artefacts to the volume backscatter spectrum, namely a distinctive "sawtooth" pattern when all frequency channels are plotted together. In turn, the model may attempt to fit these spurious peaks using the resonant bubble-like scatterer. To help avoid this, we narrowed the prior for bubble ESR to a uniform distribution between 0.05 and 0.5 mm, which is closer to the probable range of swimbladder sizes for larval fish. To help improve the chances of convergence, we also estimated the maximum a posteriori values of the parameters and used the optimised values as the starting point for the Markov Chain. Optimization used the standard BFGS algorithm (Nocedal and Wright, 2006), implemented in Julia's Optim.jl package (Mogensen and Riseth, 2018).

Results

Simulated case study: fish and krill

All models were able to extract useful information from the data, with their posteriors differing significantly from the priors and encompassing the true parameter values within their 95% credible intervals (i.e. between the 2.5% and 97.5% quantiles of the posterior, Figure 5). All models met the convergence criterion ($0.99 < \hat{R} < 1.01$ for all parameters). The NB model struggled somewhat in the krill-dominated scenario, identifying a posterior mode for ESR that was smaller than the true value (Figure 5). Because this quantity determined the TS for fish used to partition the backscatter, the posteriors for the numerical densities were inaccurate as well. In contrast, the BB models for all three scenarios converged to narrow posteriors centered on the true ESR of 0.7 mm, allowing them to produce more accurate and precise estimates for fish TS and therefore fish density. The NB posteriors in all three scenarios were broader, reflecting the model's difficulty in determining the true resonant peak and densities from five individual frequencies .

The relative accuracy of the posteriors for fish and krill differed depending on the dominant source of backscatter. When fish contributed most of the backscatter, their density estimate was more accurate; the converse was true when krill were dominant. The model tended to overestimate the numerical density of the non-dominant scatterer, particularly for the NB models. In the fish-dominated scenario, the mean of the posterior for krill in the NB model was 1.6 times the true value (16 vs. 10 krill m^{-3}). In the krill-dominated scenario, the mean of the posterior for fish was 42 times the true value (0.04 vs. 0.001 fish m^{-3}). These overestimates were largely driven by the long tails on the respective posterior distributions (Figure 5), probably caused by the model attempting to account for random noise in the backscatter spectrum by increasing the density of the non-dominant scatterer. When BB information was available, these posteriors were much better constrained. The long tails were eliminated, and the posterior means were more representative of the posterior distributions–as well as being closer to the true simulated animal densities.

Simulated case study: mesopelagic mixture

In the simulated mixture of mesopelagic animals, the inverse model was again able to recover good estimates of the modelled densities, though with varying precision depending on the auxiliary information available (Figure 6). All Markov chains reliably converged to unimodal posteriors. As seen in the first case study, the posterior for every BB model was narrower than that for the corresponding NB model.

In the first scenario, with only acoustic data available, the models were still able to effectively update their priors to improve their estimates of the densities of nearly all the scatterers. The only exception was squid; these posteriors suggested densities well above the true value of zero (Figure 6). In the second scenario, with the addition of a hypothetical camera survey providing precise priors for sergestids and siphonophores, the posteriors for these two scatterers were

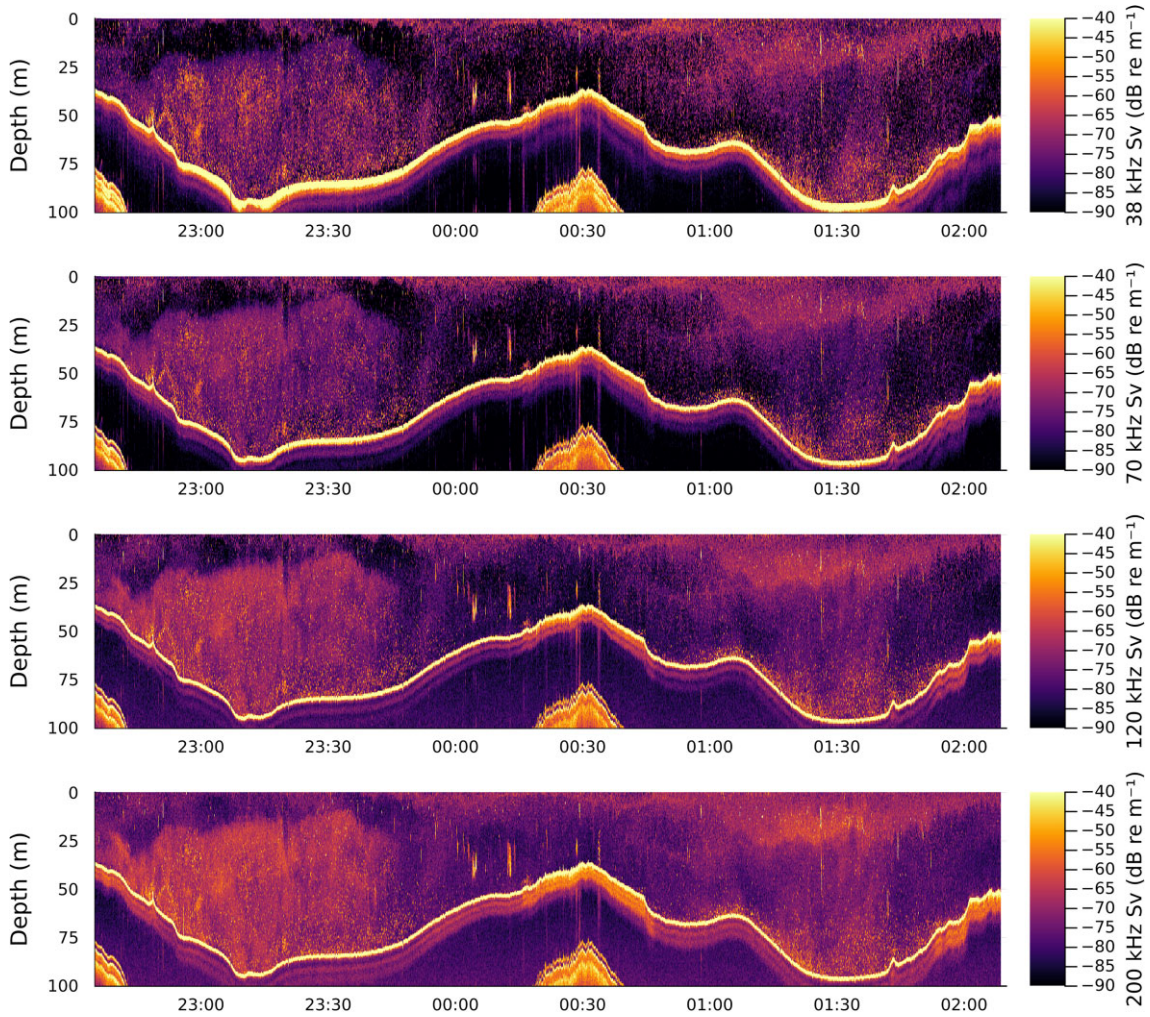

Figure 4. Echograms from upper Barnabas Trough, Gulf of Alaska, showing mean volume backscattering strength (S_v, in dB re m^{-1}) from BB, pulse compressed signals using the 38, 70, 120, and 200 kHz channels.

restricted to a narrow range around their true values. However, this information also improved the estimates for hake and myctophids, especially for the NB models. By fixing the proportion of backscatter from siphonophores, a degree of freedom was eliminated for the similarly-scattering hake and myctophids, constraining the upper limit of their posteriors. In the third scenario, eDNA sampling eliminated squid and confirmed the presence of all other scatterers. Here, a similar but less dramatic effect was seen on the hake posteriors: both the NB and BB posteriors became slightly more focused around the true value than in the initial acoustics-only case. The elimination of squid from the prior of course also eliminated them from the posterior. In the final scenario, where both video and eDNA information were available, the model made reasonably accurate estimates of all scatterer densities. The posteriors means for hake, myctophids, sergestids, and siphonophores differed from the true values by 21%, 59%, 0.7%, and 0.3% in the NB models, and by 4%, 1%, 6%, and 6% in the BB models.

Field study: Aleutian shelf break

Backscatter at the Aleutian shelf break showed several structures, with different frequencies dominating at different locations in the water column (labelled I—V in Figure 3). A small region of high backscatter at all frequencies (Layer I) was present near the seabed just inshore of the shelf break. Layer II spanned the length of the transect between the surface and 75 m depth, and was strongest at 200 kHz. Another layer (III), visible at all frequencies but most intense at 18 kHz, was centred near 100 m depth offshore of the shelf break. Around the 10 km mark, this layer appeared to merge with Layer II above it. A smaller, more irregular region of scattering (Layer IV) was located around 200 m depth just off the continental slope, with a relatively flat frequency response. Finally, a region of diffuse backscatter (Layer V) intersected the continental slope at about 300 m depth extending offshore for approximately 7 km and rising to 200 m depth. This layer was most intense at 38 kHz with much weaker scattering at 120 and 200 kHz.

Probabilistically inverting the 11 078 cells took 61 minutes when run in parallel on 16 processes on a high-performance consumer laptop, and produced realistic posterior densities for all four scatterers over most of the survey line. The main exceptions were at deeper depths for krill, where the lack of data at 120 and 200 kHz (where small, fluid-like scatterers such as krill echo most strongly) allowed the model to return posteriors for krill that were low but highly uncertian, and for bubbles in areas of low 18 kHz backscatter, where the model

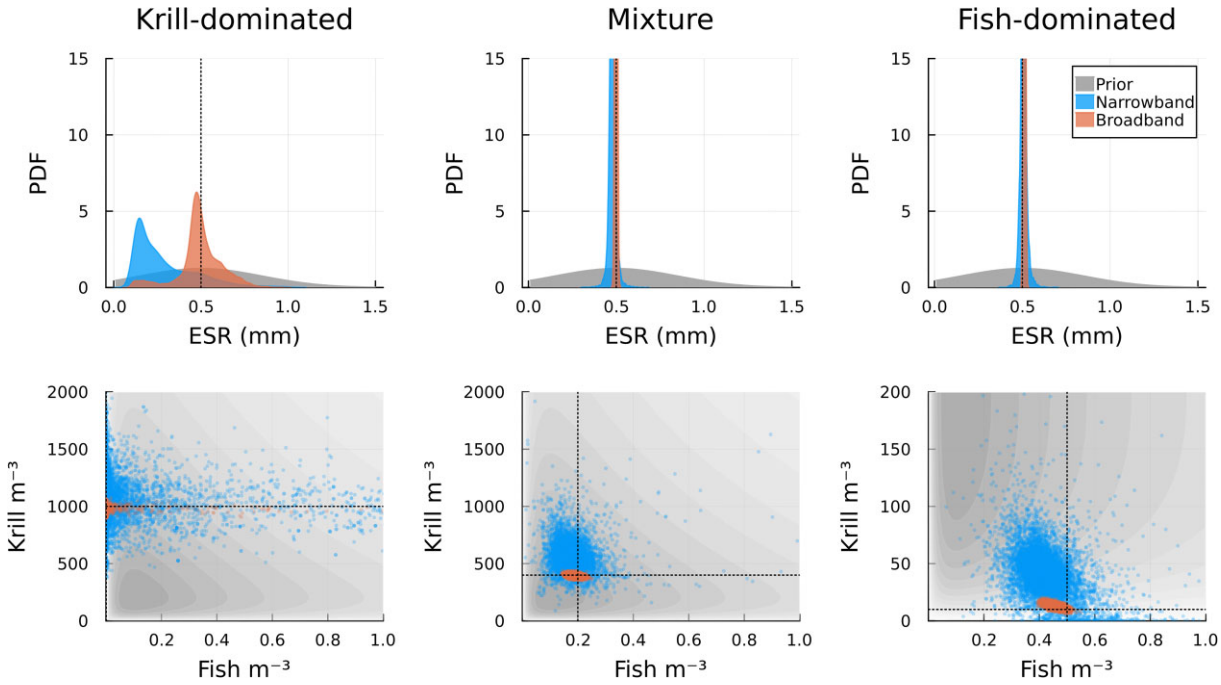

Figure 5. The inverse model was able to recover the true values for fish swimbladder size (equivalent spherical radius, ESR), as well as fish and krill densities, in three simulated scenarios. These scenarios simulated mixtures of larval fish and krill with backscatter dominated by krill (left column), dominated by fish (right column), and with similar contributions from both (centre column). The top row shows probability distributions for the ESR, with the prior in grey and kernel density estimates of the MCMC posteriors for NB (blue) and BB (orange) models. The vertical dotted line indicates the true radius. The bottom row shows bivariate scatterplots of the MCMC posteriors for fish and krill numerical densities from the NB (blue) and BB(orange) models, overlaid on their prior distribution (grey contours). Note that the prior is identical for all three scenarios, but the axes limits are different for each plot. Dotted lines indicate the true densities.

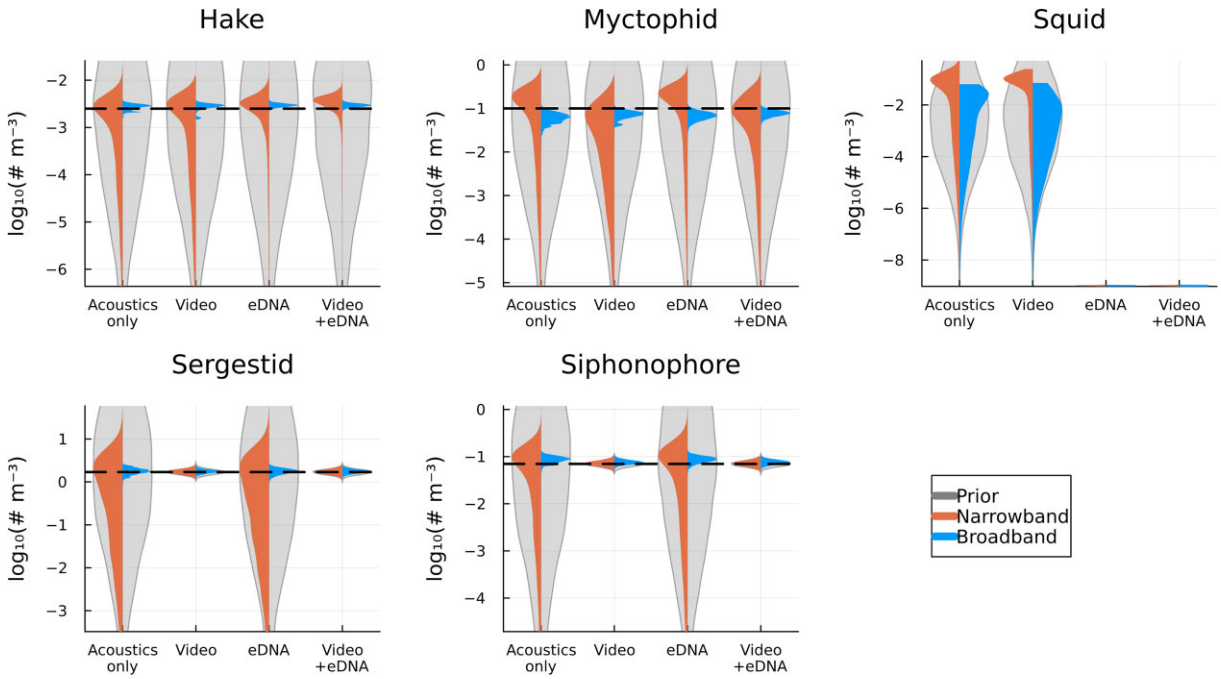

Figure 6. The accuracy and precision of Bayesian inverse solutions improved with the addition of auxiliary data and with the expansion of acoustic data from NB to BB. Each subplot shows the prior (grey) and posterior distributions (orange for NB, blue for BB) of scatterer density in each of four simulated scenarios. In the first scenario, only acoustic data were available. In the second, a video survey revealed the density of sergestid shrimp and siphonophores (reflected in their narrower priors), but none of the faster-swimming animals. In the third, environmental DNA (eDNA) confirmed the presence (but not abundance) of all scatterers except for squid. In the final scenario, both video counts and eDNA were available. Even though neither *in-situ* sampling method could quantify hake or myctophids, the constraints they imposed allowed the model to improve its estimates of their density relative to the true values (horizontal dashed lines).

struggled to fit a resonance peak (Figure 7). As in the first case study of simulated fish and krill, some of the posteriors were multi-modal, as the model did not converge to a single value for the bubble-like scatterer's ESR, so the posteriors for bubble density were also often multi-modal, requiring caution when summarising them (Figure S14).

The model attributed Layer I, the small region of high backscatter at the shelf break, mostly to large swimbladdered fishes (Figure 7a), consistent with the trawl catches, although it did not rule out the presence of myctophids or other bubble-like scatterers (Figure 7b, d). Off the shelf, the near-surface scattering Layer II was allocated primarily to krill, consistent with the high backscatter at 120 and 200 kHz (Figure 7c). Layer III below it displayed a somewhat more complex structure. Its base, at about 100 m depth, was inferred to be of a mixture of large fishes and myctophid-like scatterers at densities on the order of 1×10^{-4} m^{-3} and 1×10^{-2} m^{-3}, respectively (Figure 7a, b). Slightly shallower and closer to the shelf, between 6 and 12 km along track, the layer's composition shifted to bubble-like scatterers with a radius of ∼0.6 mm (Figure 7d and Figure 8). These scatterers also overlapped with the deepest portion of the krill in Layer I near the 10 km mark. Though krill were not the dominant contributors to backscatter here, they were by far the most common in numerical terms, with densities on the order of 0.3 m^{-3}, a hundred times more dense than the acoustically dominant bubble-like scatterers (Figure 7).

Layer IV, the smaller area of backscatter just off the mid-slope, was attributed to a mixture of large swimbladdered fish and myctophid-like scatterers (Figure 7a and b). The deepest layer (V) was dominated by myctophids, with peak densities near 0.02 m^{-3} (Figure 7b). However, it also appeared to contain lower densities of large fishes and small bubble-like scatterers, again with ESRs between 0.8 and 1 mm (Figure 8).

Alongside the posteriors' central tendencies, it is instructive to look at their relative precisions, presented here as coefficients of variation (Figure 7 and Figure 8). These give a rough indication of the model's confidence in its solutions. In most locations where the model inferred a high density of one of the scatterers, it also inferred a low value for the posterior's C.V., suggesting high confidence in the values.

Field study: broadband in Barnabas Trough

Backscatter in Barnabas Trough included large, diffuse clouds and layers, as well as small, discrete pelagic schools and dispersed echoes from individual targets. The schools were mostly observed in the upper 50 m, while the more diffuse types of scattering extended over most of the water column (Figure 4). Inverting the 1 315 binned and averaged BB spectra took ∼1 hour running on 16 parallel processes on a high-performance laptop. Consistent with the BB simulations in the first case study, most, but not all (69%), Markov chains converged to unimodal posteriors (Figure S15 C.

The model attributed the dense schools in the upper water column mainly to large swimbladdered fish, although some were inferred to contain krill or small bubble-like scattering as well (Figure 9). The diffuse scattering filling the main axis of the Trough was attributed to a mixture of krill and small bubble-like scatterers. On the first transect across the Trough, these two types of scatterers were relatively evenly mixed, with densities of both between 5 and 10 m^{-3}, though krill were somewhat more abundant near the top of the diffuse layer. On the second transect across the trough, krill were more concentrated in the upper water column, likely following upward vertical migration after sunset at 23:15, with higher densities ranging from 10 to 20 m^{-3}. On this second transect, bubble-like scatterers were spread across the water column, though their densities were highest around 20 m depth, overlapping with the highest concentration of krill, and in the deepest portion of the trough around 90 m. In the regions of highest density for each scatterer class, their posterior C.V.s were mostly <30%, indicating fairly confident model fits (Figure 9). In areas of lower inferred density for each scatterer the model was less confident, consistent with findings in the fish-krill simulation above.

The inferred size of small bubble-like scatterers tended to increase with depth in the water column, with mean ESRs of about 0.1 mm in the surface layer growing to ∼0.35 mm in the deepest parts of the trough. The models were usually able to converge on a single, well-constrained posterior distribution for bubble ESR (Figures 10B and S15), but where they could not, this uncertainty propagated through to their inferred numerical density: variability in a bubble's size translates into even greater variability in its resonant frequency and backscattering cross-section, so it is not surprising that a low-precision fit for this parameter would translate into a similarly low-precision estimate for its numerical density.

Discussion

These results demonstrate some of the advantages of the probabilistic inverse approach when compared to simpler methods of classification and integration. It can deconvolve mixtures of different scatterers and, with appropriate data, it may also be able to infer some physical properties of the scatterers themselves, such as their size. Perhaps more significantly, the probabilistic structure of the model allows it to incorporate diverse types of auxiliary information, providing a variety of practical ways to constrain the often-underdetermined inverse problem. Finally, the output of the model is itself probabilistic, providing an intrinsic uncertainty estimate, and a variety of diagnostics to assess its goodness-of-fit and reliability.

However, the method is not a panacea, and can fail in a variety of ways. Most fundamentally, the inverse problem could be ill-posed, due to too many candidate scatterers for the number of frequencies, collinear TS spectra, overly vague priors, or insufficient independent constraints. Alternatively, too few candidate scatterers, or the wrong set of candidates, might be included in the model. This could cause biassed or incorrect results as the model attempts to account for the observed scattering with the TS spectra it has available, even if they are not appropriate. Even if the model is well-formulated, the posteriors may be multimodal or degenerate, causing numerical challenges and poor convergence for the optimizer or MCMC algorithm (Betancourt and Girolami, 2013). And as always, noisy or otherwise poor-quality data are the enemies of acoustic inference. If the signal-to-noise ratio is too low, little can be done to mitigate it, and ideally the method would indicate this with large uncertainties on estimates, as it did here for krill and bubbles in deep water off the Aleutian shelf (Figure 7c and d). Ultimately, APES is simply a variant of Bayesian regression, and is subject to all the challenges and constraints inherent to any statistical model. However, because it is based

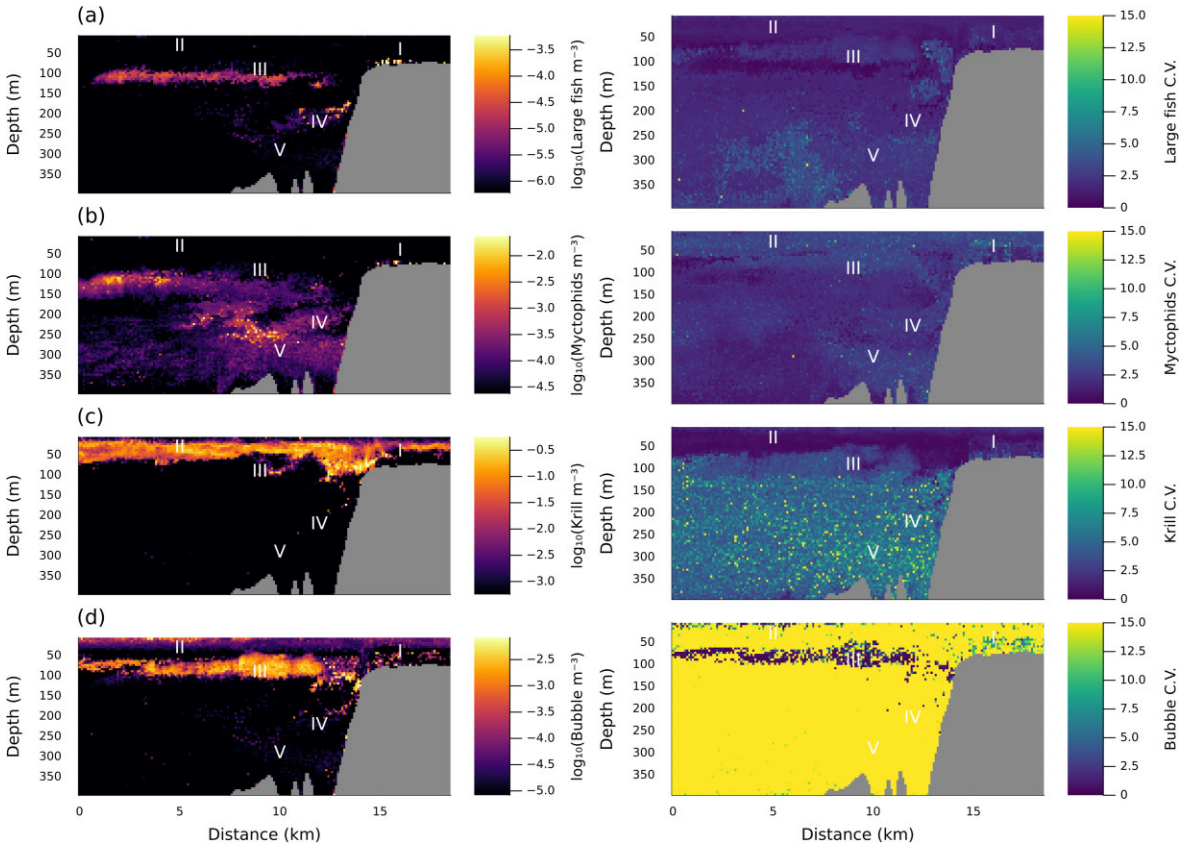

Figure 7. Estimated numerical densities and associated uncertainties for (a) large swimbladdered fish, (b) myctophids, (c) krill, and (d) small bubble-like scatterers along the Gulf of Alaska shelf transect. Left panels show the mean value of the posterior for numerical density on a logarithmic scale (note that scales differ between these plots). The right panels show the associated posterior coefficient of variation (C.V., i.e. the ratio of the *SD* to the mean).

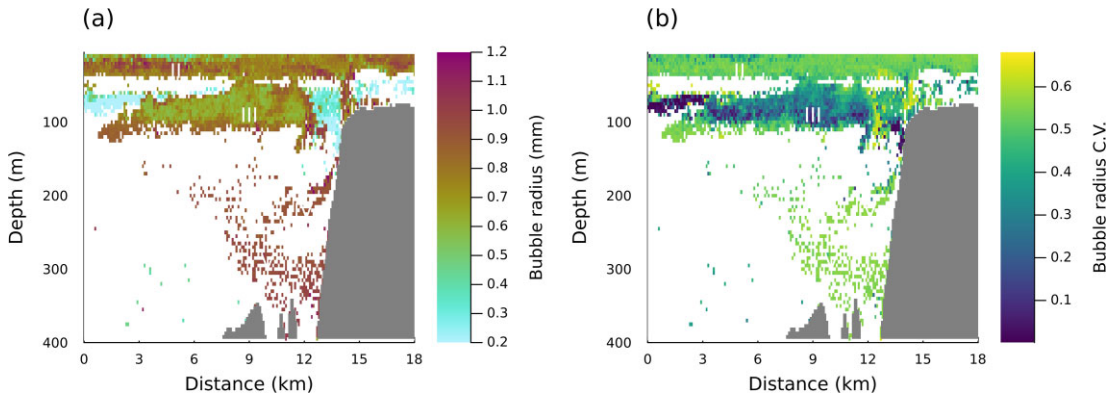

Figure 8. Inferred equivalent spherical radius for small bubble-like scatterers along the Gulf of Alaska shelf break transect. Left panel shows the mean of the posterior for bubble size. Right panel shows the posterior's coefficient of variation (C.V, the ratio of its *SD* to its mean). For clarity, areas where the estimated density of bubbles was lower than 1510^{-5} m^{-3} are not displayed (white pixels).

on the physical principles of acoustic scattering and linearity (Foote, 1983) that underlie all of fisheries acoustics, an insufficiently precise or otherwise suspect answer simply indicates that better data or model formulations are needed.

More than anything else, two decisions will determine the model's success or failure in a particular application. The first is how to define the backscattering cross-section matrix Σ. This could mean choosing an appropriate set of candidate scatterers to include in the model, or selecting appropriate physics-based scattering models to fit. Whether using a fixed set of predefined spectra or calculating them using acoustic theory, APES will be subject to the same tradeoffs as any statistical model. Including more candidate scatterers will reduce the chance of a biassed result due to omitting a species that was truly present, at the cost of making the posteriors for all species more diffuse, and their credible intervals wider. Similarly, building in parameterized scattering functions gives the model more flexibility to achieve an accurate fit, at the cost of some precision, and perhaps computation time.

Several strategies can be used to help set up an appropriate scattering matrix. The first, and most obvious, is to limit the candidates to species that could plausibly be present at

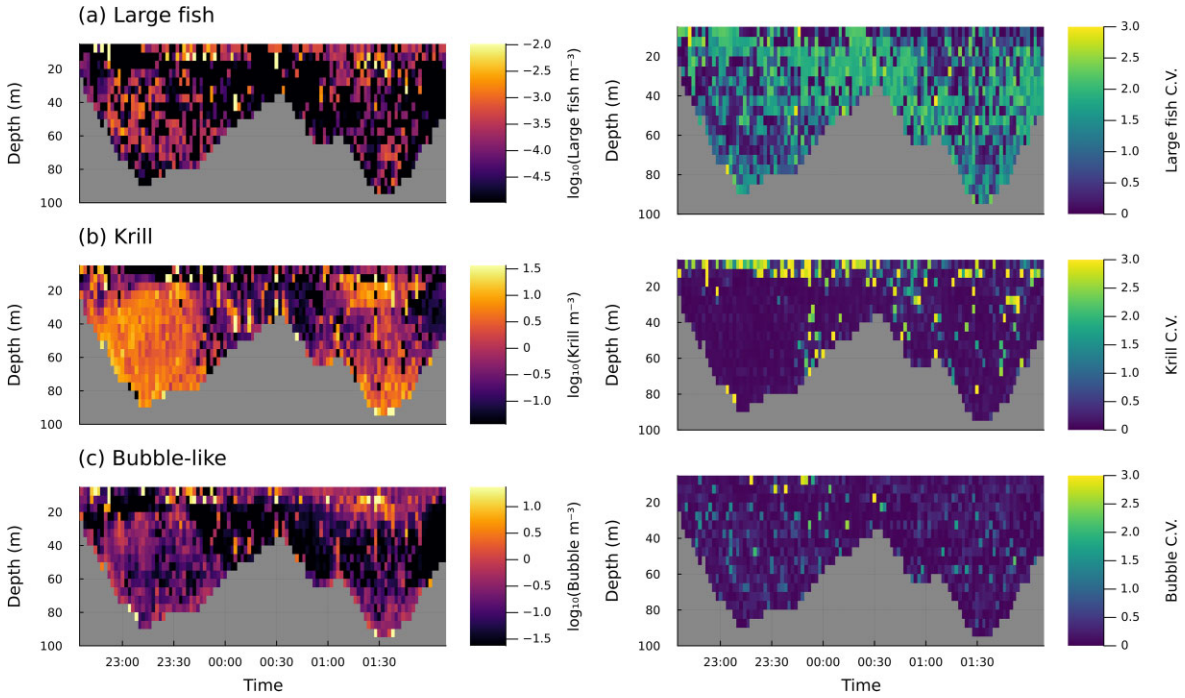

Figure 9. Estimated numerical densities and associated uncertainties for (a) large swimbladdered fish, (b) krill, and (c) small bubble-like scatterers along two transects across upper Barnabas Trough, Gulf of Alaska. Left panels show the logarithm of the posterior mean for numerical density, while the right panels show the associated posterior coefficient of variation (C.V., i.e. the ratio of the *SD* to the mean). Note that colour scales differ for each plot.

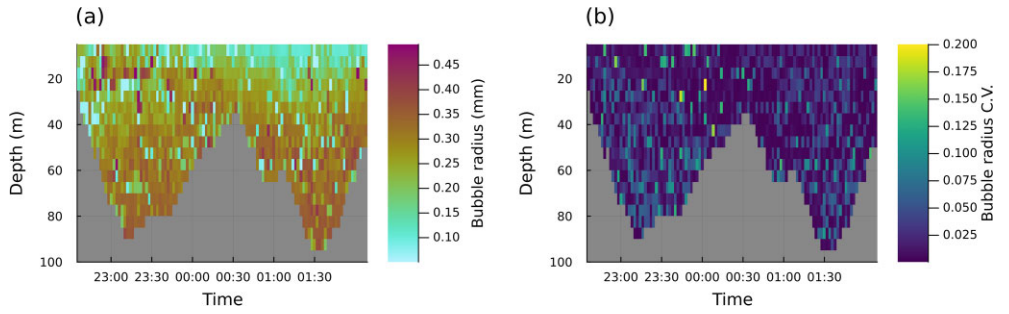

Figure 10. Estimated equivalent spherical radius (ESR) for the bubble-like scatterers along two transects in upper Barnabas Trough, Gulf of Alaska. (a) Posterior mean value of ESR. (b) Coefficient of variation (C.V., the ratio of the *SD* to the mean) for the posterior distribution of ESR.

significant densities, and to avoid including species with similar scattering spectra (unless you have an independent way to constrain them). Given a scattering matrix, its condition number can be calculated as a measure of collinearity among the individual scattering spectra (Chu *et al.*, 2016); if the condition number is too high, this could suggest a need to remove or group some of the candidates. Parametric scattering functions may reduce the total number of unknowns, e.g. instead of fitting five fixed spectra for different size classes of the same species, a single scattering function parameterized by length may be more parsimonious. When little prior information is available, unsupervised factorization or clustering algorithms can be used to identify possible constituent spectra (Woillez *et al.*, 2012). Though there is no guarantee that they will be biologically meaningful, if their shapes are recognizable as those of realistic scatterers (e.g. resonant, fluid-like, etc.) it may be helpful (Cotter *et al.*, 2021b).

The second major choice is how to set the priors for the scatterers' densities. This choice is especially critical when the scattering matrix is ill-conditioned. In the absence of other information, one can simply identify the maximum biologically realistic density n_{max} for each candidate scatterer, and choose a prior with negligible probability mass above that value. This is conceptually equivalent to the procedure of setting a maximum threshold $s_{v,max}$ for volume backscatter from a given class of scatterers (e.g. Ballón *et al.*, 2011; Uumati, 2013), where $s_{v,max} = n_{max}\,\sigma$. Knowledge of different species habitats, depth distributions, and geographical ranges could also be used to set priors, based on how abundant a species is likely to be at the time and place of the observed data. While these choices may appear somewhat vague, it is worth emphasising that most practitioners already make similar decisions implicitly, for instance by assuming that backscatter at a particular frequency is "dominated' by species A, or that backscatter values above a particular threshold could not have been generated by species B. Ideally, the priors can be set based on information from direct sampling, either as absolute densities, estimates of species composition, or presence/absences. An ad-

vantage of the Bayesian approach compared to current practice is that the model will reveal if the acoustic data are inconsistent with the priors. Even when direct samples are available, they will be at a different spatial resolution than the acoustic data, so some type of interpolation or extrapolation will be required. This is an important topic for future research.

Ultimately, if this method is to be used, it will need to be compared with, and incorporated into, existing workflows for acoustic surveys. On one level, this should be relatively straightforward–it is, after all, just a more complete implementation of the same theory and models already in use. On another level, however, it requires a shift in perspective. Most current practice relies on classifying backscatter as consistent with a species of interest via examination of echograms or applying frequency response thresholds, and then integrating backscatter at a single frequency to estimate that species' abundance (Korneliussen *et al.*, 2018). The idea of "classification," and the binary decision implicit in the framework, is thus deeply ingrained, so it is worth emphasising that APES is *not* a classification method–it is a probabilistic inversion. It does not assign backscatter to one category or another, but rather yields a multivariate probability distribution of animal densities. If desired, this can be summarised to give a "most probable constituent" or "dominant scatterer," but this is not required. Doing so gives up some of the method's advantages–namely, its ability to deconvolve mixed aggregations of different animals. We suspect that such mixtures may be more common than widely appreciated, given that a classification-based perspective will tend to overlook them.

The preceding discussion of the challenges and key decisions involved in the application of the probabilistic inverse method points the way to several immediate avenues for future research. One broad area is improving the reliability and efficiency of model-fitting procedures. This could include programmed checks for Markov chain convergence, for posteriors that fit the data poorly, or that are similar to the prior, indicating that little information was available from the data. Methods for automated model comparison and relevance determination (Marwala, 2018) could choose the most-likely subset of scatterers from a longer list of candidates, reducing the need for subjective analyst supervision. Where appropriate, quadratic approximations and VI (Blei *et al.*, 2017), could speed up inference. Even the most efficient algorithm, however, will require at least two orders of magnitude more computation than simple frequency differencing, which may pose challenges when fitting large survey-scale datasets. In these cases, the "embarrassingly parallel" structure of the cell-wise inverse problem can be exploited to run it on distributed computing systems, allowing the method to scale efficiently with data volume. In an astronomical application, Regier *et al.* (2018, 2019) used conceptually similar Bayesian models to infer the identity and parameters of 188 million stars and galaxies in 178 TB of telescope images. Running in parallel on a cluster with 665 000 processors, this task completed in less than 15 minutes. While few fisheries scientists have access to such supercomputers, their datasets are also orders of magnitude smaller, and commercial cloud-computing services mean scalable parallel data processing is both feasible and economical.

Another important question for future work is how to incorporate spatial information, including autocorrelation, into the models. At its most basic, this could mean choosing an appropriate size for the cells into which volume backscatter is averaged. This involves a tradeoff between spectral signal-to-noise ratio and speed of model fitting, which improve as the cells get coarser, and spatial resolution, which gets worse. More sophisticated partitions of the echogram are also possible. For instance, distinct schools, aggregations, or layers could be detected and segmented as a pre-processing step, with backscatter spectra from similar pixels averaged together before being passed to APES (Fernandes, 2009; Korneliussen *et al.*, 2016; Loranger *et al.*, 2022). On a regular grid, spatial dependence could be accounted for by smoothing the results from nearby cells with a moving average. Alternatively, an explicit spatial model could be used. One simple way to implement this would be to use the posterior for each depth layer in one profile of the water column as the prior for the corresponding layer in the next profile–essentially treating the echogram as a time series in a multivariate state-space model. Many elaborations on these themes are possible, and we believe they have good potential to improve the reliability of the basic inverse model.

Broadband echosounders are relatively new and have not been incorporated into most routine abundance surveys. Our examples indicate that the increased information in BB signals has the potential to greatly improve abundance estimates, and that probabilistic inverse models offer a flexible framework to incorporate this information in large-scale surveys. However, more work is needed to optimise spectral processing for probabilistic inversions. Volume scattering spectra are inherently noisy (Bassett et al., 2018) due to the complexities of animal anatomy, as well as their variable size, orientation, and spatial arrangement within aggregations. Thus, even single-species aggregations require some smoothing, and the optimal level of smoothing for different applications remains to be determined. Further challenges are introduced when scattering in a given volume is dominated by a small number of targets. Single echoes from an individual animal can have spectra that are considerably more complex than the relatively smooth average TS curves used in the model (Reeder *et al.*, 2004). Because a transducer's effective beamwidth changes during an FM sweep, backscatter from a single off-axis target will also appear to decrease with increasing frequency within each channel, creating a distinctive sawtooth-shaped spectrum. The echoes from a small number of targets within the same volume will interfere constructively and destructively, introducing spectral peaks and nulls whose spacing is related to the *in situ* spacing of the targets (Demer *et al.*, 2017). Non-linear crosstalk may also become more apparent in these situations (Khodabandeloo *et al.*, 2021).

While there are many aspects open for further research and development, we believe that the probabilistic inverse approach holds great potential for applied fisheries acoustics. The approach offers a consistent framework to incorporate all available information (e.g. BB observations, diverse types of *in-situ* data, and prior knowledge) into acoustic abundance estimates. The models are based on standard physical acoustic theory and basic principles of probability, making them more easily interpretable, and potentially more generalizable, than some ML/AI approaches. At the same time, the fact that APES is probabilistic makes it more robust than past inverse approaches used in fisheries acoustics. In effect, the error bars provided by the posterior distribution act like shock absorbers, making it safe to run automatically on large volumes of data. If there is insufficient information to solve the inverse problem precisely, the model will not report a precise poste-

rior. Conversely, when high-quality data are available, a narrow posterior distribution can provide more confidence–more than is often available in fisheries acoustics–that the inferred scatterers truly are what they appear to be. As echosounders are deployed more widely on varied platforms, the spectrum of frequencies available expands, and more varied forms of ground-truth become available, probabilistic approaches such as the one presented here will provide a way for fisheries acoustics to make improved measurements of the changing ocean environment.

Acknowledgements

We would like to thank the crew of the NOAA Ship *Oscar Dyson* for making our survey possible under the challenging conditions of a global pandemic, compounded by staffing shortages and the ensuing fatigue of long stretches at sea. The manuscript was greatly improved by early feedback from Cole Monahan and Beth Phillips, as well as discussions with Gemma Carroll. We also thank the three anonymous reviewers for their time and constructive comments. Richard D. Eastaville and Josephine Wolverqueen provided emotional support through all stages of the project, and Kirra Urmy provided the best possible motivation to finish the manuscript. Use of trade names does not imply endorsement by the National Marine Fisheries Service, NOAA. Likewise, the conclusions of this paper are those of the authors, not necessarily their employers.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Author contributions

S.S.U. conceived the study in discussion with A.D.R,. wrote the software and simulations, collected and analysed the field data, and drafted the manuscript. C.B. processed the raw broadband data. All authors discussed the results and revised the manuscript.

Conflict of interest statement

The authors declare no conflicts of interest.

Data availability

All data and scripts required to reproduce the analyses in this manuscript are available at https://github.com/ElOceanografo/APESExamples. Additionally, the inverse modeling algorithms are implemented in a software package, available at https://github.com/ElOceanografo/ProbabilisticEchoInversion.jl.

References

Ballón, M., Bertrand, A., Lebourges-Dhaussy, A., Gutiérrez, M., Ayón, P., Grados, D., and Gerlotto, F. 2011. Is there enough zooplankton to feed forage fish populations off Peru? An acoustic (positive) answer. Progress in Oceanography, 91: 360–381.

Barham, E. G. 1957. The Ecology of Sonic Scattering Layers in the Monterey Bay Area. Stanford University. 182 pp.

Barham, E. G. 1963. Siphonophores and the deep scattering layer. Science, 140: 826–828.

Bassett, C., De Robertis, A., and Wilson, C. D. 2018. Broadband echosounder measurements of the frequency response of fishes and euphausiids in the Gulf of Alaska. ICES Journal of Marine Science, 75: 1131–1142.

Benoit-Bird, K. J., Patrick Welch, T., Waluk, C. M., Barth, J. A., Wangen, I., McGill, P., Okuda, C. *et al.* 2018. Equipping an underwater glider with a new echosounder to explore ocean ecosystems. Limnology and Oceanography: Methods, 16: 734–749.

Benoit-Bird, K. J., and Waluk, C. M. 2020. Exploring the promise of broadband fisheries echosounders for species discrimination with quantitative assessment of data processing effects. The Journal of the Acoustical Society of America, 147: 411–427.

Berger, C. S., Bougas, B., Turgeon, S., Ferchiou, S., Ménard, N., and Bernatchez, L. 2020. Groundtruthing of pelagic forage fish detected by hydroacoustics in a whale feeding area using environmental DNA. Environmental DNA, 2: 477–492.

Betancourt, M. J., and Girolami, M. 2013. Hamiltonian Monte Carlo for Hierarchical Models. arXiv. http://arxiv.org/abs/1312.0906 (last accessed 28 June 2022).

Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. 2017. Julia: a fresh approach to numerical computing. SIAM Review, 59: 65–98.

Bezanson, J., Karpinski, S., Shah, V. B., and Edelman, A. 2012. Julia: a fast, dynamic language for technical computing. arXiv. http://arxiv.org/abs/1209.5145.

Blei, D. M., Kucukelbir, A., and McAuliffe, J. D. 2017. Variational Inference: a Review for Statisticians. Journal of the American Statistical Association, 112: 859–877.

Brautaset, O., Waldeland, A. U., Johnsen, E., Malde, K., Eikvil, L., Salberg, A.-B., and Handegard, N. O. 2020. Acoustic classification in multifrequency echosounder data using deep convolutional neural networks. ICES Journal of Marine Science, 77: 1391–1400.

Chu, D., Lawson, G. L., and Wiebe, P. H. 2016. Estimation of biological parameters of marine organisms using linear and nonlinear acoustic scattering model-based inversion methods. The Journal of the Acoustical Society of America, 139: 2885–2895.

Conti, S. G., and Demer, D. A. 2006. Improved parameterization of the SDWBA for estimating krill target strength. ICES Journal of Marine Science, 63: 928–935.

Cotter, E., Bassett, C., and Lavery, A. 2021a. Comparison of mesopelagic organism abundance estimates using *in situ* target strength measurements and echo-counting techniques. JASA Express Letters, 1: 040801.

Cotter, E., Bassett, C., and Lavery, A. 2021b. Classification of broadband target spectra in the mesopelagic using physics-informed machine learning. The Journal of the Acoustical Society of America, 149: 3889–3901.

De Robertis, A., and Higginbottom, I. 2007. A post-processing technique to estimate the signal-to-noise ratio and remove echosounder background noise. ICES Journal of Marine Science, 64: 1282–1291.

De Robertis, A., Levine, M., Lauffenburger, N., Honkalehto, T., Ianelli, J., Monnahan, C. C., Towler, R. *et al.* 2021. Uncrewed surface vehicle (USV) survey of walleye pollock, *Gadus chalcogrammus*, in response to the cancellation of ship-based surveys. ICES Journal of Marine Science, 78: 2797–2808.

De Robertis, A., Levine, R., Williams, K., and Wilson, C. 2023. Modifying a pelagic trawl to better retain small Arctic fishes. Deep Sea Research Part II: Topical Studies in Oceanography, 207: 105225.

De Robertis, A., Levine, R., and Wilson, C. D. 2018. Can a bottom-moored echo sounder array provide a survey-comparable index of abundance? Canadian Journal of Fisheries and Aquatic Sciences, 75: 629–640.

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

Demer, D. A., Andersen, L. N., Bassett, C., Chu, D., and Cutter, G. R. 2017. Evaluation of a Wideband Echosounder for Fisheries and Ma-

rine Ecosystem Science. ICES Cooperative Research Report No. 336. 70 pp. Available at https://doi.org/10.17895/ices.pub.2318 (last accessed 20 June 2023).

Demer, D. A., Berger, L., Bernasconi, M., Bethke, E., Boswell, K. M., Chu, D., Domokos, R. *et al.* 2015. Calibration of acoustic instruments. ICES Cooperative Research Report No. 326. 133 pp. International Council for the Exploration of the Sea, Copenhagen.

Dosso, S. E. 2002. Quantifying uncertainty in geoacoustic inversion. I. A fast Gibbs sampler approach. The Journal of the Acoustical Society of America, 111: 129–142.

Fernandes, P. G. 2009. Classification trees for species identification of fish-school echotraces. ICES Journal of Marine Science, 66: 1073–1080.

Foote, K. G. 1983. Linearity of fisheries acoustics, with addition theorems. The Journal of the Acoustical Society of America, 73: 1932.

Ge, H., Xu, K., and Ghahramani, Z. 2018. Turing: a Language for Flexible Probabilistic Inference. *In* Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics, pp. 1682–1690. PMLR. https://proceedings.mlr.press/v84/ge18b.html (last accessed 14 March 2022).

Gelman, A., Lee, D., and Guo, J. 2015. Stan: a Probabilistic Programming Language for Bayesian Inference and Optimization. Journal of Educational and Behavioral Statistics, 40: 530–543.

Gerstoft, P., and Mecklenbräuker, C. F. 1998. Ocean acoustic inversion with estimation of a posteriori probability distributions. The Journal of the Acoustical Society of America, 104: 808–819.

Goldberg, D. 1991. What every computer scientist should know about floating-point arithmetic. ACM Computing Surveys, 23: 5–48.

Hoffman, M. D., and Gelman, A. 2014. The No-U-Turn Sampler: adaptively Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research, 15: 1593–1623.

Holliday, D. V., Pieper, R. E., and Kleppel, G. S. 1989. Determination of zooplankton size and distribution with multifrequency acoustic technology. ICES Journal of Marine Science, 46: 52–61.

Honkalehto, T., Ressler, P. H., Towler, R. H., and Wilson, C. D. 2011. Using acoustic data from fishing vessels to estimate walleye pollock (*Theragra chalcogramma*) abundance in the eastern Bering Sea. Canadian Journal of Fisheries and Aquatic Sciences, 68: 1231–1242.

Horne, J. K. 2000. Acoustic approaches to remote species identification: a review. Fisheries Oceanography, 9: 356–371.

Irigoien, X., Klevjer, T. A., Røstad, A., Martinez, U., Boyra, G., Acuña, J. L., Bode, A. *et al.* 2014. Large mesopelagic fishes biomass and trophic efficiency in the open ocean. Nature communications, 5: 3271.

Jech, J. M., Horne, J. K., Chu, D., Demer, D. A., Francis, D. T. I., Gorska, N., Jones, B. *et al.* 2015. Comparisons among ten models of acoustic backscattering used in aquatic ecosystem research. The Journal of the Acoustical Society of America, 138: 3742–3764.

Jones, B. A., Lavery, A. C., and Stanton, T. K. 2009. Use of the distorted wave born approximation to predict scattering by inhomogeneous objects: application to squid. The Journal of the Acoustical Society of America, 125: 73–88.

Kawabata, A. 2005. Target strength measurements of suspended live ommastrephid squid, *Todarodes pacificus*, and its application in density estimations. Fisheries Science, 71: 63–72.

Khodabandeloo, B., Ona, E., Macaulay, G. J., and Korneliussen, R. 2021. Nonlinear crosstalk in broadband multi-channel echosounders. The Journal of the Acoustical Society of America, 149: 87–101.

Korneliussen, R. J., Berger, L., Campanlla, F., Chu, D., Demer, D., De Robertis, A., Domokos, R. *et al.* 2018. Acoustic target classification. ICES Cooperative Research Report No. 344. ICES. http://doi.org/10.17895/ices.pub.4567 (last accessed 28 June 2022).

Korneliussen, R. J., Heggelund, Y., Macaulay, G. J., Patel, D., Johnsen, E., and Eliassen, I. K. 2016. Acoustic identification of marine species using a feature library. Methods in Oceanography, 17: 187–205.

Kwong, L. E., Pakhomov, E. A., Suntsov, A. V., Seki, M. P., Brodeur, R. D., Pakhomova, L. G., and Domokos, R. 2018. An intercomparison of the taxonomic and size composition of tropical macrozooplankton and micronekton collected using three sampling gears. Deep Sea Research Part I: Oceanographic Research Papers, 135: 34–45.

Levine, M., Jones, D. T., and McGowan, D. W. 2022. Results of the Acoustic-Trawl Survey of Walleye Pollock (*Gadus chalcogrammus*) in the Gulf of Alaska, June-July 2021 (DY2021-04). Resource Assessment and Conservation Engineering Division, Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA.

Loranger, S., Jech, M. J., and Lavery, A. C. 2022. Broadband acoustic quantification of mixed biological aggregations at the New England shelf break. The Journal of the Acoustical Society of America, 152: 2319–2335.

Love, R. H. 1978. Resonant acoustic scattering by swimbladder-bearing fish. The Journal of the Acoustical Society of America, 64: 571–580.

Lucca, B. M., Ressler, P. H., Rodger Harvey, H., and Warren, J. D. 2021. Individual variability in sub-Arctic krill material properties, lipid composition, and other scattering model inputs affect acoustic estimates of their population. ICES Journal of Marine Science, 78: 1470–1484.

MacLennan, D. N., Fernandes, P. G., Maclennan, J. D., Fernandes, D. N., Maclennan, D. M., and Fernandes, P. G. 2002. A consistent approach to definitions and symbols in fisheries acoustics. ICES Journal of Marine Science, 59: 365–369.

Marwala, T. 2018. Automatic relevance determination. *In* Handbook of Machine Learning, pp. 59–75. World Scientific, Singapore.

McGehee, D. E., O'Driscoll, R. L., and Traykovski, L. V. M. 1998. Effects of orientation on acoustic scattering from Antarctic krill at 120 kHz. Deep-Sea Research Part II: Deep Sea Research Part II: Topical Studies in Oceanography, 45: 1273–1294.

Medwin, H., and Clay, C. S. 1998. Fundamentals of Acoustical Oceanography. Academic Press, Boston. 739pp. http://www.sciencedirect.com/science/article/pii/B9780124875708500167.

Methot, R. D. 1986. Frame trawl for sampling juvenile pelagic fish. CalCOFI Reports, 27: 267–278.

Mogensen, P. K., and Riseth, A. N. 2018. Optim: A mathematical optimization package for Julia. Journal of Open Source Software, 3:615–617.

Monahan, E. C., and Lu, M. 1990. Acoustically Relevant Bubble Assemblages and Their Dependence on Meteorological Parameters. IEEE Journal of Oceanic Engineering, 15: 340–349.

Nocedal, J., and Wright, S. J. 2006. Numerical Optimization. Springer Science, New York. 664pp.

Ohman, M. D., Davis, R. E., Sherman, J. T., Grindley, K. R., Whitmore, B. M., Nickels, C. F., and Ellen, J. S. 2019. Zooglider: an autonomous vehicle for optical and acoustic sensing of zooplankton. Limnology and Oceanography: Methods, 17: 69–86.

Omori, M., and Gluck, D. 1979. Life history and vertical migration of the pelagic shrimp *Sergestes similis* off the southern California coast. Fishery Bulletin, 77: 183–198.

Patil, A., Huard, D., and Fonnesbeck, C. J. 2010. PyMC: bayesian Stochastic Modelling in Python. Journal of statistical software, 35: 1–81.

Proud, R., Handegard, N. O., Kloser, R. J., Cox, M. J., and Brierley, A. S. 2019. From siphonophores to deep scattering layers: uncertainty ranges for the estimation of global mesopelagic fish biomass. ICES Journal of Marine Science, 76: 718–733.

Reeder, D. B., Jech, J. M., and Stanton, T. K. 2004. Broadband acoustic backscatter and high-resolution morphology of fish: measurement and modeling. The Journal of the Acoustical Society of America, 116: 747–761.

Regier, J., Miller, A. C., Schlegel, D., Adams, R. P., McAuliffe, J. D., andPrabhat. 2019. Approximate inference for constructing astronomical catalogs from images. The Annals of Applied Statistics, 13: 1884–1926.

Regier, J., Pamnany, K., Fischer, K., Noack, A., Lam, M., Revels, J., Howard, S. *et al.* 2018. Cataloging the Visible Universe through Bayesian Inference at Petascale. arXiv: 1801.10277 [astro-ph]. http://arxiv.org/abs/1801.10277.

Reisenbichler, K. R., Chaffey, M. R., Cazenave, F., McEwen, R. S., Henthorn, R. G., Sherlock, R. E., and Robison, B. H. 2016. Automating MBARI's midwater time-series video surveys: the transition from ROV to AUV. *In* OCEANS 2016 MTS/IEEE Monterey, pp. 1–9. IEEE Publishing, New York.

Roberts, P. L. D., Jaffe, J. S., and Trivedi, M. M. 2011. Multiview, Broadband Acoustic Classification of Marine Fish: a Machine Learning Framework and Comparative Analysis. IEEE Journal of Oceanic Engineering, 36: 90–104.

Robison, B. H. 1999. The Coevolution of Undersea Vehicles and Deep-Sea Research. Marine Technology Society Journal, 33: 65–73.

Sarr, J.-M. A., Brochier, T., Brehmer, P., Perrot, Y., Bah, A., Sarré, A., Jeyid, M. A. *et al.* 2021. Complex data labeling with deep learning methods: lessons from fisheries acoustics. ISA Transactions, 109: 113–125.

Shane, J. 2018. Do neural nets dream of electric sheep? [online]. Available at https://www.aiweirdness.com/do-neural-nets-dream-of-electric-18-03-02/ (last accessed 4 July 2023).

Shelton, A. O., Ramón-Laca, A., Wells, A., Clemons, J., Chu, D., Feist, B. E., Kelly, R. P. *et al.* 2022. Environmental DNA provides quantitative estimates of Pacific hake abundance and distribution in the open ocean. Proceedings of the Royal Society B: Biological Sciences, 289: 20212613. Royal Society.

Simmonds, J., and MacLennan, D. 2005. Fisheries acoustics: Theory and practice: Second edition. Blackwell Science, Oxford. 1–252pp.

Soule, M. A., Hampton, I., and Lipiński, M. R. 2010. Estimating the target strength of live, free-swimming chokka squid *Loligo reynaudii* at 38 and 120 kHz. ICES Journal of Marine Science, 67: 1381–1391.

Stanton, T. K., and Chu, D. 2000. Review and recommendations for the modelling of acoustic scattering by fluid-like elongated zooplankton: euphausiids and copepods. ICES Journal of Marine Science, 57: 793–807.

Stanton, T. K., Lee, W.-J., and Baik, K. 2018. Echo statistics associated with discrete scatterers: a tutorial on physics-based methods. The Journal of the Acoustical Society of America, 144: 3124–3171.

Stanton, T. K., Sellers, C. J., and Jech, J. M. 2012. Resonance classification of mixed assemblages of fish with swimbladders using a modified commercial broadband acoustic echosounder at 1–6 kHz. Canadian Journal of Fisheries and Aquatic Sciences, 69: 854–868.

Stoner, A. W. S. W., Ryer, C. H. R. H., Parker, S. J. P. J., Auster, P. J. A. J., and Wakefield, W. W. W. W. 2008. Evaluating the role of fish behavior in surveys conducted with underwater vehicles. Canadian Journal of Fisheries and Aquatic Sciences, 65: 1230–1243.

Urmy, S. S. 2016. SDWBA.jl: a Julia package for modeling acoustic backscatter from zooplankton. https://doi.org/10.5281/zenodo.594901 (last accessed 20 June 2023).

Urmy, S. S., Horne, J. K., and Barbee, D. H. 2012. Measuring the vertical distributional variability of pelagic fauna in Monterey Bay. ICES Journal of Marine Science, 69: 184–196.

Uumati, M. 2013. Acoustic investigations on bearded goby and jellyfish in the northern Benguela ecosystem. University of St Andrews, St Andrews, Scotland. 176pp. https://www.semanticscholar.org/paper/Acoustic-investigations-on-bearded-goby-and-in-the-Uumati/e4c4a4ca8f6ad553671ab29359148fb69758a3f4 (last accessed 28 June 2022).

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P. C. 2021. Rank-normalization, folding, and localization: an improved $\widehat{R}$for assessing convergence of MCMC. Bayesian Analysis, 16. http://arxiv.org/abs/1903.08008 (last accessed 21 July 2022).

Warren, J. D., Stanton, T. K., Wiebe, P. H., and Seim, H. E. 2003. Inference of biological and physical parameters in an internal wave using multiple-frequency, acoustic-scattering data. ICES Journal of Marine Science, 60: 1033–1046.

Williams, K., Horne, J. K., and Punt, A. E. 2015. Examining influences of environmental, trawl gear, and fish population factors on midwater trawl performance using acoustic methods. Fisheries Research, 164: 94–101.

Williams, K., Rooper, C. N., and Towler, R. 2010. Use of stereo camera systems for assessment of rockfish abundance in untrawlable areas and for recording pollock behavior during midwater trawls. Fishery Bulletin, 108: 352–362.

Woillez, M., Ressler, P. H., Wilson, C. D., and Horne, J. K. 2012. Multifrequency species classification of acoustic-trawl survey data using semi-supervised learning with class discovery. The Journal of the Acoustical Society of America, 131: EL184–EL190.

Handling editor: Richard O'Driscoll

Received: 27 February 2023; **Revised:** 19 May 2023; **Accepted:** 22 May 2023