Acoustic Methods for Observing Biological Structure in the Ocean Interior

by

Emmett Culhane

B.A. English and Economics, Haverford College, 2013
M.S. Environmental Science, Yale University School of the Environment, 2020

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

September 2026

Signature of Author _____

Joint Program in Oceanography/Applied Ocean Science & Engineering
Massachusetts Institute of Technology &
Woods Hole Oceanographic Institution
June 7, 2026

Certified by _____

Dr. Camrin D. Braun
Thesis Supervisor
Woods Hole Oceanographic Institution

Accepted by _____

Dr. Greg Fournier
Chair, Joint Committee for Biological Oceanography
Massachusetts Institute of Technology/
Woods Hole Oceanographic Institution

Acoustic Methods for Observing Biological Structure in the Ocean Interior

by

Emmett Culhane

Submitted to the MIT-WHOI Joint Program in Oceanography/Applied Ocean Science and Engineering on June 7, 2026, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biological Oceanography

ABSTRACT

The mesopelagic zone (200–1,000 m depth) hosts taxonomically diverse communities of fishes, crustaceans, cephalopods, and gelatinous organisms that mediate globally significant fluxes of carbon, oxygen, and nutrients through diel vertical migration, respiration at depth, and trophic transfer. Current estimates of mesopelagic fish biomass span 1–20 Gt, and this uncertainty propagates directly into estimates of their biogeochemical contribution, which has been estimated in regional studies to approach or exceed gravitational particle flux. This uncertainty persists because the ocean interior is opaque to electromagnetic radiation, placing it beyond the reach of satellite remote sensing, and because conventional sampling methods (nets, trawls, optical systems) cannot scale to the volume of the ocean interior, which is the largest contiguous habitat on Earth. Active acoustic sensing—in which echosounders transmit sound pulses and record backscatter from insonified targets as a function of range—is the only existing technology capable of observing these communities continuously, at depth, and over the full water column. However, the quantitative interpretation of acoustic backscatter is subject to large uncertainties, particularly in the resonance scattering regime where gas-bearing organisms produce strongly enhanced, frequency-dependent returns whose amplitude depends on inclusion size, depth-dependent compression, and tissue properties. This thesis develops methods, analyses, and instrumentation that advance the capacity for large-scale acoustic observation of mesopelagic biological structure. **Chapter 2** presents a systematic uncertainty analysis of volume backscattering strength derived from shipboard Acoustic Doppler Current Profilers (ADCPs), using Monte Carlo propagation, Sobol sensitivity decomposition, and comparison with calibrated echosounder data from seven paired cruises to quantify parametric uncertainty and evaluate the robustness of structural property estimates. **Chapter 3** develops a convolutional autoencoder framework for analyzing archive-scale ADCP echogram imagery across the Pacific Ocean; unsupervised clustering identifies six regional acoustic modes corresponding to distinct oceanographic provinces, predictable from satellite-derived environmental variables—a step toward extending satellite surveillance of surface ocean properties to biological structure in the ocean interior. **Chapter 4** presents a basin-scale characterization of shallow scattering layers across the Indo-Pacific using calibrated echograms from the Australian IMOS Ships of Opportunity program, resolving the distribution, environmental controls, and acoustic subtypes of an incompletely characterized intermediate-depth feature. **Chapter 5** presents the design, construction, and validation of EchoBot, an open-source wideband split-beam echosounder built entirely from non-proprietary

components, with all hardware designs, PCB layouts, and signal processing software released under open licenses. Together, these contributions expand the usable acoustic archive, introduce new analytical methods for extracting ecological information from heterogeneous acoustic data, characterize biological structure at scales not previously resolved, and provide open instrumentation to support the observational capacity needed to reduce uncertainty in the ocean carbon budget and improve constraints on the role of mesopelagic communities in Earth system biogeochemistry.

Thesis Supervisor: Camrin D. Braun
Title: Assistant Scientist II, Biological Oceanography
Woods Hole Oceanographic Institution

For Albert

Acknowledgments

This work was supported by a NASA Future Investigators in NASA Earth and Space Science and Technology (FINESST) Fellowship in Earth Science Research and Carbon Cycling (Grant #21-EARTH21-0403), and the MIT-WHOI Joint Programs Office. Research related travel was supported by the WHOI Academic Programs Office and the MIT Student Assistance Fund.

Foremost, I would like to thank my advisor, Camrin Braun, for his guidance and support at every stage of this work, both academic and personal—but especially personal. He has been an excellent supervisor and mentor in every respect, and I am grateful to now count him as a friend.

I owe a special thanks to Peter Gaube, who gave me the opportunity to work as a summer fellow at the University of Washington Applied Physics Laboratory in 2019. Pete served as my co-advisor when I began the PhD at UW, and his support has been a constant since. He is also now a valued friend.

I thank my committee members—Glen Flierl and Mei Sato—my exam chair, Aran Mooney, and Simon Thorrold, who chaired my general exam, for dedicating their time and energy to this thesis. Their critiques and guidance improved this work throughout.

I am grateful to the many collaborators and co-authors who contributed to the research presented here: Ali DellaPenna, Martin Arostegui, Roland Proud, Andone Lavery, Bob Petitt, and Jennifer Johnson.

At MIT and WHOI, I thank Sam Silva and Ellie Hennelly for help with many last-minute on-call housing requests, the CRL team for providing laboratory space, and Jaida Elcock, Seth Cones, Laura McDonnell, and Jacob Partida for their friendship.

I thank my mom and dad for always being in my corner.

Most of all, I want to thank my lovely wife Callie for her support throughout, and for being such a wonderful mother to our son. Albert, I can't wait to show you the big wild world.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Biological Structure and the Mesopelagic Observation Gap

The ocean interior is the largest habitable space on the planet. Between 200 and 1,000 m depth, the mesopelagic zone hosts dense, taxonomically diverse communities of fishes, crustaceans, cephalopods, and gelatinous organisms that aggregate into persistent scattering layers and undertake diel vertical migrations spanning hundreds of meters (Kaartvedt et al. 2012b; St. John et al. 2016). These structures organize the vertical distribution of biomass, mediate the downward transport of carbon through respiration and excretion at depth, and support trophic linkages between surface primary production and deep-ocean food webs (Braun et al. 2022; Steinberg and Landry 2017). The ocean absorbs approximately 25% of anthropogenic carbon dioxide emissions (Friedlingstein et al. 2023) and stores roughly 50 times more carbon than the atmosphere on centennial to millennial timescales (Sabine et al. 2004). The biological component of this storage—and the mesopelagic communities that drive much of the active carbon flux—remains among the least constrained terms in the global carbon budget (Archibald et al. 2019). This lack of constraint stems directly from the difficulty of observing these communities at the scales required.

The observational challenge is fundamental. Abiotic ocean properties—temperature, salinity, dissolved oxygen, nutrient concentrations—can be measured routinely by in-situ sensors (CTDs, Argo floats, gliders) and, in the case of surface properties, by satellite remote sensing, yielding global coverage at high spatial and temporal resolution (Garcia et al. 2024; Roemmich et al. 2009). Biological quantities in the ocean interior have no equivalent observational infrastructure. The organisms of interest are motile, with Reynolds numbers ranging from $\sim 10^1$ (copepods) to $\sim 10^6$ (mesopelagic fishes); at these scales, organisms are patchily distributed across volumes far exceeding the capture area of nets and trawls, making representative sampling impractical (Kaartvedt et al. 2012b; Wiebe et al. 2023). Net and trawl sampling is further limited by its inability to resolve continuous vertical structure or capture diel variability, as each tow integrates over a depth interval and a single time point. More fundamentally, the ocean interior is opaque to electromagnetic radiation: visible light is attenuated to less than 1% of surface irradiance by 100 m depth in typical open-ocean conditions, and no electromagnetic signal penetrates usefully beyond ~200 m (Aksnes et al. 2017). Satellite remote sensing—the technology that has enabled global-scale monitoring of terrestrial ecosystems, atmospheric composition, and ocean surface properties—cannot observe the ocean interior at all. Optical imaging at mesopelagic depths requires artificial illumination, which alters organism behavior, and is limited in range to meters. The result is that the largest habitat on Earth, by volume, is effectively invisible to the primary tools of modern Earth observation.

This observational gap has direct consequences. Current estimates of mesopelagic fish biomass alone span 1–20 Gt (Irigoien et al. 2014; Proud et al. 2019, 2020), a range that encompasses

values from roughly equivalent to the biomass of all commercially harvested fish stocks to an order of magnitude greater. This uncertainty is not limited to total biomass. It extends to the depth distribution, vertical coherence, migration behavior, and environmental associations of scattering communities—the structural properties that determine how these organisms interact with their environment and mediate biogeochemical fluxes. These uncertainties propagate into estimates of active carbon flux (Archibald et al. 2019; Davison et al. 2013), respiratory oxygen demand (Bianchi et al. 2013), nutrient regeneration, and the trophic support available to higher predators (Braun et al. 2022; Willis et al. 2024). At the scale of the global carbon budget, they represent a gap in export mass balance that limits our ability to close the ocean carbon cycle, project the ocean's response to climate change, and assess the sustainability of proposed mesopelagic fisheries (St. John et al. 2016; Thibault et al. 2025). Major ecosystem services—carbon sequestration, nutrient cycling, support of commercial fisheries—depend on communities we cannot yet quantify at the precision required for Earth system modeling.

1.2 Acoustic Remote Sensing of the Ocean Interior

Active acoustics are uniquely suited to observing biological structure in the ocean interior. Acoustic waves propagate through seawater with low attenuation relative to electromagnetic radiation, traveling thousands of kilometers under favorable conditions (Simmonds and MacLennan 2008). Active acoustic instruments (echosounders) transmit sound pulses and record the intensity of backscattered returns as a function of range, producing depth-resolved, temporally continuous measurements of biological scattering throughout the water column (Benoit-Bird and Lawson 2016; Simmonds and MacLennan 2008). This capacity resolves the depth, continuity, and diel movement of scattering layers over spatial scales inaccessible to nets or optical methods.

The quantitative interpretation of acoustic backscatter is subject to large uncertainties arising from the physics of sound scattering by marine organisms. In the Rayleigh scattering regime, where the acoustic wavelength is much larger than the scatterer ($ka \ll 1$), backscatter increases steeply with frequency and is sensitive to organism density contrast but insensitive to morphology (Stanton et al. 1998). In the geometric regime ($ka \gg 1$), backscatter depends on the projected cross-sectional area and surface properties of the target (Simmonds and MacLennan 2008). Between these limits, in the resonance regime ($ka \sim 1$), organisms possessing gas inclusions (e.g., fish swim bladders, siphonophore pneumatophores) produce strongly enhanced, frequency-dependent backscatter whose amplitude and spectral shape depend on inclusion size, depth-dependent compression, and the elastic properties of surrounding tissue (Bassett et al. 2020; Cotter et al. 2021a; Stanton et al. 2010). For the mixed, multi-species assemblages that compose mesopelagic scattering layers, these scattering

regimes overlap across frequencies and taxa, and the mapping from received acoustic intensity to biological quantities (biomass, abundance, species composition) is nonlinear, frequency-dependent, and poorly constrained (Dornan et al. 2019; Proud et al. 2019). The practical consequence is the order-of-magnitude uncertainty in mesopelagic fish biomass noted above: from acoustic surveys alone, we cannot yet distinguish whether the ocean interior contains 2 or 20 Gt of fish (Proud et al. 2019).

1.3 Thesis Motivation and Overview

This thesis develops methods for extracting ecological information from acoustic observations of the ocean interior, with emphasis on biological structure that can be observed robustly across instruments, archives, and spatial scales. This work was funded, in part, by a NASA FINESST grant (#21-EARTH21-0403) whose broader aim is to develop tools that allow satellite measurements of the surface ocean to be translated into estimates of interior ocean biological content. The premise is that if the physical and chemical properties observable from space—sea surface temperature, sea surface height, chlorophyll concentration, light attenuation, mixed layer depth—shape the composition and distribution of mesopelagic communities through mechanistic relationships, then those relationships can be learned from paired surface-subsurface observations and used to predict biological state variables across the global ocean. Treating biological community structure as a state variable, analogous to temperature or salinity (Wunsch 2006), requires several capabilities that do not yet exist at sufficient maturity: quantification of the uncertainties inherent in different acoustic data sources and instruments; methods for extracting ecological information from large, heterogeneous acoustic archives; conceptual models of community state that relate acoustic observations to environmental forcing; expanded observational coverage through both data rescue and new instrumentation; and ground-truth measurements of the frequency-dependent scattering properties of organisms at depth, linked to surface-observable environmental conditions, so that acoustic measurements can be decomposed into estimates of biological stocks. The four chapters of this thesis each address one or more of these requirements.

Chapter 2 presents a systematic analysis of uncertainty in ADCP-derived volume backscattering strength (S_v). Using Monte Carlo propagation, Sobol sensitivity decomposition, and comparison with concurrent calibrated echosounder data from seven paired cruises, the chapter quantifies the impact of parametric assumptions in the S_v computation on correspondence with calibrated observations before and after regression calibration, and evaluates the robustness of structural property estimates such as weighted mean depth and diel migration timing.

Chapter 3 develops a convolutional autoencoder framework for analyzing archive-scale ADCP

echogram imagery across the Pacific Ocean, leveraging the structural robustness established in Chapter 2. The approach compresses each echogram into a learned feature space that captures structural patterns without requiring calibration. Unsupervised clustering identifies six regional modes that correspond to distinct oceanographic provinces and can be predicted from satellite-derived environmental data.

Chapter 4 presents a basin-scale characterization of shallow scattering layers (SSLs)—intermediate-depth aggregations between the surface community and the deep scattering layer—using calibrated echograms from the Australian IMOS Ships of Opportunity program. Environmental modeling, automated segmentation, and unsupervised clustering resolve the distribution, environmental controls, and acoustic subtypes of SSLs across the Indo-Pacific.

Chapter 5 presents the design, construction, and validation of EchoBot, an open-source, open-license wideband split-beam echosounder built from non-proprietary components. EchoBot is the first open-source, open-license wideband split-beam echosounder, with all hardware designs, PCB layouts, and signal processing software released under open licenses to support expanded acoustic observation and community-driven development.

Together, these chapters advance the capacity for large-scale acoustic observation of mesopelagic communities by expanding the usable archive, developing new analytical methods, characterizing incompletely described ecological features, and providing open-source instrumentation for future ground-truth and deployment.

Chapter 2

Parameter Error in ADCP-Derived Volume Backscattering Strength

Abstract

Shipboard ADCPs are the most widely deployed acoustic instruments in the ocean. Although designed to measure current velocity, they also record echo intensity from which volume backscattering strength (S_v) can be derived. However, the parameters required for this conversion are not needed for current computation, are rarely archived, and vary across instruments. We present a consolidated parameter registry for the Teledyne (TRDI) instrument classes represented in the Joint Archive for Shipboard ADCP (JASADCP) and use it to propagate S_v uncertainty across 145 cruises spanning 4 instrument classes on 3 frequencies. We then evaluate parameter correction and regression calibration using 7 paired OS-75 ADCP–echosounder datasets from 5 vessels, calibrated against echosounders at both 70 and 38 kHz. S_v uncertainty is dominated by the receiver sensitivity slope (k_c) and noise floor (E_r); all other parameters contribute <4%. After regression calibration, k_c is the only parameter producing depth-structured residual bias—invisible to RMSE but producing NASC bias up to +430% per depth layer. Depth-dependent regression eliminates this bias within-cruise but its depth coefficient is spatially unstable. We propose an empirical approach to estimating k_c from concurrent echosounder data and show that the resulting values are stable across the geographic subsets tested and do not distort the biological depth profile. Structural metrics from normalized profiles are invariant to parameterization state and calibration method, confirming the reliability of archival ADCP data for structural biological analyses.

2.1 Introduction

Shipboard ADCPs record echo intensity as a byproduct of velocity profiling. Converting this signal to volume backscattering strength (S_v) requires instrument-specific parameters that are rarely preserved in data archives (Deines 1999; Mullison 2017).

ADCP-derived S_v has been used to study mesopelagic scattering layers, diel vertical migration, and biomass distribution (Bianchi and Mislan 2016; Flagg and Smith 1989; Heywood et al. 1991), and has been compared to concurrent echosounder observations (Brierley et al. 1998; Lee et al. 2004; Receveur et al. 2020b). However, the sensitivity of these estimates to parameter uncertainty has not been systematically quantified. Understanding how parameter error propagates, persists after calibration, and produces depth-stratified bias is essential for determining whether ADCP S_v can be used quantitatively and for establishing best practices.

We address four operationally relevant questions:

1. How much does ADCP-derived S_v vary across the plausible range of each input parameter, and how is that variance distributed across parameters and depth?
2. How much does correct parameterization reduce the bias between uncalibrated ADCP S_v and concurrent echosounder observations, and which parameters drive that improvement?
3. If ADCP S_v is calibrated via regression against a concurrent echosounder, do the input parameters still matter, and if so, how do they affect the calibrated output?
4. To what extent are structural estimates from ADCP backscatter—scattering layer depth and diel migration timing—robust to parameter error, calibration method, and frequency mismatch?

We use 145 archival cruises for uncertainty propagation and 7 paired ADCP–echosounder datasets for calibration evaluation (Section 2.2.3), supported by a consolidated parameter registry (Section 2.2.2).

2.2 Methods

2.2.1 S_v equations

Volume backscattering strength (S_v, dB re 1 m^{-1}) is computed from ADCP echo intensity using the broadband equation (Mullison 2017):

$$S_v = C + 10\log_{10}\big((T_x + 273.16)\,R^2\big) - L\mathrm{DBM} - P\mathrm{DBW} + 2\alpha R + 10\log_{10}\big(10^{k_c(E-E_r)/10} - 1\big) \quad (2.1)$$

where C is the system constant (dB), T_x is transducer temperature (°C), R is along-beam slant range (m), LDBM is transmit pulse length (dB re m), PDBW is transmit power (dB re W), α is sound absorption (dB/m) (Francois and Garrison 1982), k_c is the receiver sensitivity slope (dB/count), E is echo intensity (counts), and E_r is the noise floor (counts). The signal-to-noise term is the corrected formulation from TRDI (Teledyne RD Instruments 1998), differing from the Deines (1999) approximation $k_c(E - E_r)$ by properly subtracting noise power before log-conversion. For narrowband instruments, k_c is temperature-dependent: $k_c = 127.3/(T_e + 271)$, valid only for $E < 200$ counts (Teledyne RD Instruments 1998).

2.2.2 Parameter registry

We assembled a parameter registry for all TRDI instrument classes in the JASADCP archive from manufacturer documentation (Mullison 2017; Teledyne RD Instruments 1998) and unpublished TRDI correspondence (J. Mullison, S. Idle, 2020), building on prior work by Leung (2020, Appendix A).

Table 2.1: Parameter registry — broadband instruments.

Parameter	OS-38	OS-75	OS-150	WH-300	Source
C (dB)	−172.19	−164.26	−156.01	−140.87	Mullison 2017
PDBW (dB re W)	24.0	24.0	21.0	17.5	Mullison 2017
k_c mean (dB/count)	0.37	0.39	0.42	0.42	Mullison/Idle emails
k_c sigma	0.004	0.019	0.012	—	Mullison emails
E_r mean (counts)	14	19	22	in situ	Mullison emails
E_r sigma (counts)	4.83	2.93	2.08	—	Mullison emails

Table 2.2: Parameter registry — narrowband instruments.

Parameter	75 kHz	150 kHz	Source
K2 (VM/DR)	2.5/2.2	4.3/3.6	FST-003
Ks	1.09×10^5	4.17×10^5	FST-003
K1c (VM/DR) (W)	5.4/6.1	3.3/3.9	FST-003
k_c	127.3/(T_e+271)	—	FST-003

2.2.3 Datasets

A set of 145 cruises from the JASADCP archive (NOAA NCEI), spanning 4 instrument classes (71 OS-75, 42 OS-38, 22 NB-150, 6 OS-150) was used for uncertainty propagation. Seven paired OS-75 ADCP–echosounder datasets from 5 vessels provide ground-truth comparisons at two frequency pairings: 70/75 kHz (close match) and 38/75 kHz (cross-frequency). Profiles are matched by nearest-neighbor time within 15 minutes, restricted to 20–600 m depth.

Table 2.3: Paired ADCP–echosounder datasets.

Dataset	Vessel	Region	EK freq	Matched profiles
70/75 kHz				
Investigator 2016	RV Investigator	S. Ocean	70 kHz	1,162
Investigator 2019	RV Investigator	S. Ocean	70 kHz	8,230
Shimada SH1906	NOAA Shimada	NE Pacific	70 kHz	948
Falkor FK191012	R/V Falkor	C. N. Pacific	70 kHz	71
38/75 kHz				
NAAMES	RV Atlantis	N. Atlantic	38 kHz	5,366
SO287	RV SONNE	Atlantic	38 kHz	26,065
SO283	RV SONNE	Benguela	38 kHz	260

2.2.4 Parameterization states

We compare ADCP S_v under three sequential parameterization states. Environmental inputs (WOA23 monthly T/S/pH, colocated to each cruise) are identical across states.

Table 2.4: Parameterization states.

State	E_r	k_c	Requires
Default	Registry mean (19)	Registry mean (0.39)	Nothing
Partial	Cruise-minimum	Registry mean	ADCP data only
Full	Cruise-minimum	EK-optimized	Paired EK cruise or factory k_c

The fully corrected k_c is determined empirically by sweeping k_c, fitting a fresh regression at each value, and identifying the zero-crossing of the post-regression residual depth slope. We perform this both constrained to the registry $\pm 2\sigma$ range and unconstrained; divergence indicates the best-fit k_c may lie outside published bounds. This empirical approach assumes that k_c is the dominant source of post-regression depth structure—an assumption we test via geographic subset analysis in Section 2.3.3.

2.2.5 Comparison metrics

We diagnose parameter error through physics-based and structural metrics:

S_v **bias**: mean depth-binned difference (ADCP − EK).

NASC bias: log-mean backscatter integration error per 100 m layer after regression.

WMD (weighted mean depth): computed over 200–600 m from per-profile normalized daytime profiles (solar noon ±3 h).

Migration timing: top-hat sigmoid fit to surface-layer (0–100 m) S_v over the diel cycle.

2.2.6 Uncertainty propagation

Monte Carlo sampling (500 realizations per profile) draws all uncertain parameters from their registry distributions (Table 2.1) to produce S_v ensembles at each depth across 145 JASADCP cruises. Variance is decomposed using Sobol sensitivity analysis (Sobol 1993) to isolate each parameter's marginal contribution.

2.2.7 Regression calibration

Regression calibration against concurrent echosounders is widely used for ADCP S_v (Lee et al. 2004; Receveur et al. 2020b), but it is unclear whether such calibration absorbs parameter errors or merely masks them. Within-cruise OLS ($S_{v,\mathrm{EK}}$ = slope × $S_{v,\mathrm{ADCP}}$ + intercept) is fit per dataset and parameterization state. We examine residual depth profiles rather than RMSE alone, and sweep k_c across [0.32, 0.55] with a fresh regression at each value. We additionally test whether a depth-dependent regression ($S_{v,\mathrm{EK}} = \beta_0 + \beta_1 S_{v,\mathrm{ADCP}} + \beta_2 \cdot \mathrm{depth}$) can absorb k_c bias statistically, comparing the spatial stability of its depth coefficient to that of the empirical k_c across distinct water masses.

2.3 Results

2.3.1 S_v uncertainty and variance attribution

S_v uncertainty is dominated by two parameters with a depth-dependent partitioning (Fig. 2.1): k_c accounts for 74% of OS-75 variance above 100 m, E_r for 67% below 400 m. Total 2σ uncertainty

ranges from 3–16 dB (OS-75) to 0–17 dB (NB-150). All remaining parameters are collectively minor (C 3.4%, PDBW 0.2%, T/S 0–3%).

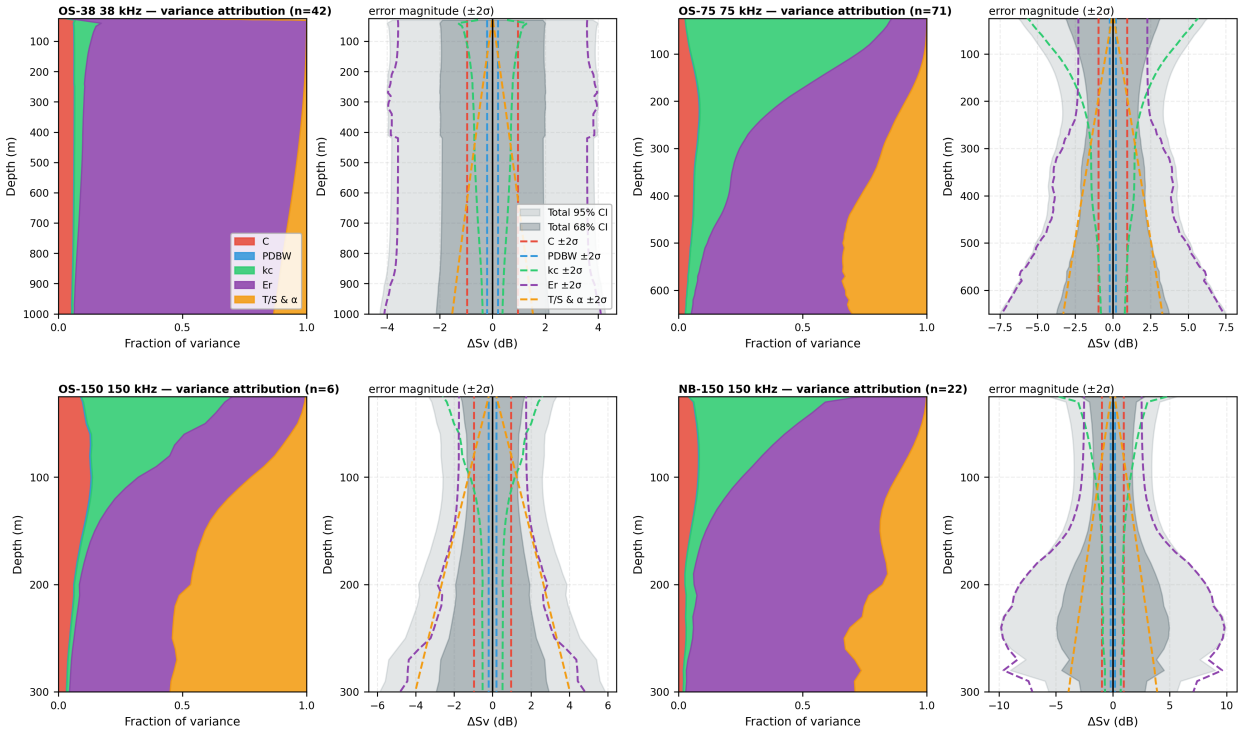

Figure 2.1: S_v variance attribution and 2σ uncertainty envelope for 4 ADCP instrument classes.

2.3.2 Uncalibrated ADCP S_v bias by parameterization state

Correcting E_r reduces uncalibrated bias by 4–5 dB (e.g., Investigator: −16.2 to −10.6 dB), with improvement scaling with the E_r discrepancy. Correcting k_c improves Shimada by 3.4 dB but can worsen mean offset on other datasets because it reshapes the vertical profile rather than shifting it uniformly (Table 2.5, Fig. 2.2).

Table 2.5: Uncalibrated S_v bias (dB) by parameterization state.

Dataset	Default	Partial	Full	E_r effect	k_c effect
Investigator 2016	−16.2	−10.6	−11.2	+5.6	−0.6
Investigator 2019	−15.8	−11.6	−12.2	+4.2	−0.6
Shimada	−8.0	−8.0	−4.6	0.0	+3.4
Falkor	−18.5	−13.3	−21.0	+5.2	−7.7
NAAMES	−10.4	−8.7	−9.3	+1.7	−0.6
SO287	−10.1	−5.8	—	+4.3	—
SO283	−12.8	−10.3	—	+2.5	—

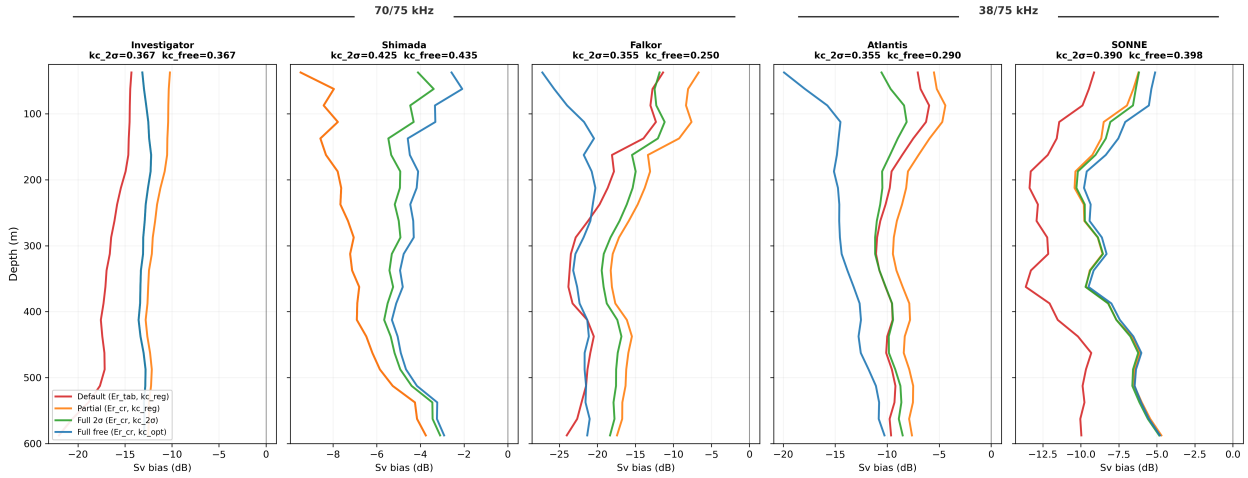

Figure 2.2: Uncalibrated S_v bias profiles under three parameterization states, stratified by frequency pairing.

2.3.3 Regression-calibrated ADCP S_v and surviving parameter error

k_c error produces depth-structured bias that persists through regression, while all other parameter errors are absorbed. Post-regression RMSE varies by <1 dB across the k_c sweep, but the residual depth slope varies by up to 4.0 dB/100 m—a diagnostic that RMSE averages away.

Correcting E_r reduces depth slope where E_r is wrong (e.g., Inv2016: +1.62 to +0.61 dB/100 m). Empirical k_c correction drives slope to near-zero on all datasets (Inv2019: −0.01, Shimada: −0.01, Falkor: +0.07 dB/100 m), confirming k_c as the dominant surviving error source. The progression from default to fully corrected residual profiles is shown in Figure 2.3.

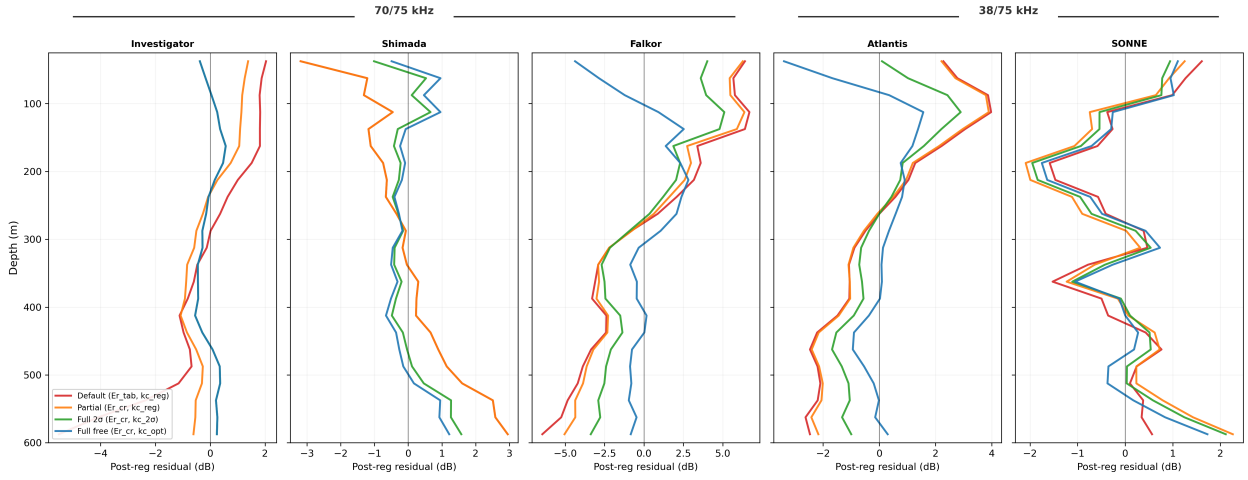

Figure 2.3: Post-regression residual depth profiles under four parameterization states for 5 vessels. Full k_c correction drives depth slope to near-zero.

Figure 2.4 shows the NASC consequence at four k_c values per vessel. At registry extremes, bias ranges from +430% (Falkor, 0–100 m) to −78% (Falkor, 500–600 m) with sign reversal between shallow and deep. The constrained and unconstrained k_c optima reveal that for 2 of 5 instruments (Investigator, NAAMES) the registry range is adequate, while for 3 (Shimada, Falkor, SONNE) the best-fit k_c lies outside published bounds—with in-situ values spanning a 10σ range of the manufacturer distribution.

Figure 2.4: NASC % bias at 4 k_c values per vessel: registry extremes and constrained/unconstrained optima.

Multiple regression and validation of empirical k_c

A depth-dependent regression (2D) eliminates depth-structured residuals within-cruise (slopes <0.05 dB/100 m vs 0.3–2.2 for 1D). However, geographic subset analysis reveals that the statistical depth coefficient and the physics-based k_c parameter behave fundamentally differently when transferred across water masses (Figs. 2.5–2.6).

The 2D depth coefficient β_2 diverges between regions—at 500 m the correction differs by 2.4 dB for NAAMES (subtropical vs subpolar, 25° latitude span)—producing transfer slopes of ±0.6 dB/100 m. In contrast, $k_{c,\text{opt}}$ converges to the same value regardless of region ($\Delta k_c = 0.3\sigma$ for NAAMES, 0.1σ for Investigator) and transfers with slopes <0.04 dB/100 m. This likely occurs because β_2 conflates the k_c artifact with real oceanographic depth structure, while k_c reflects only the receiver sensitivity. This contrast is shown directly in Figure 2.5.

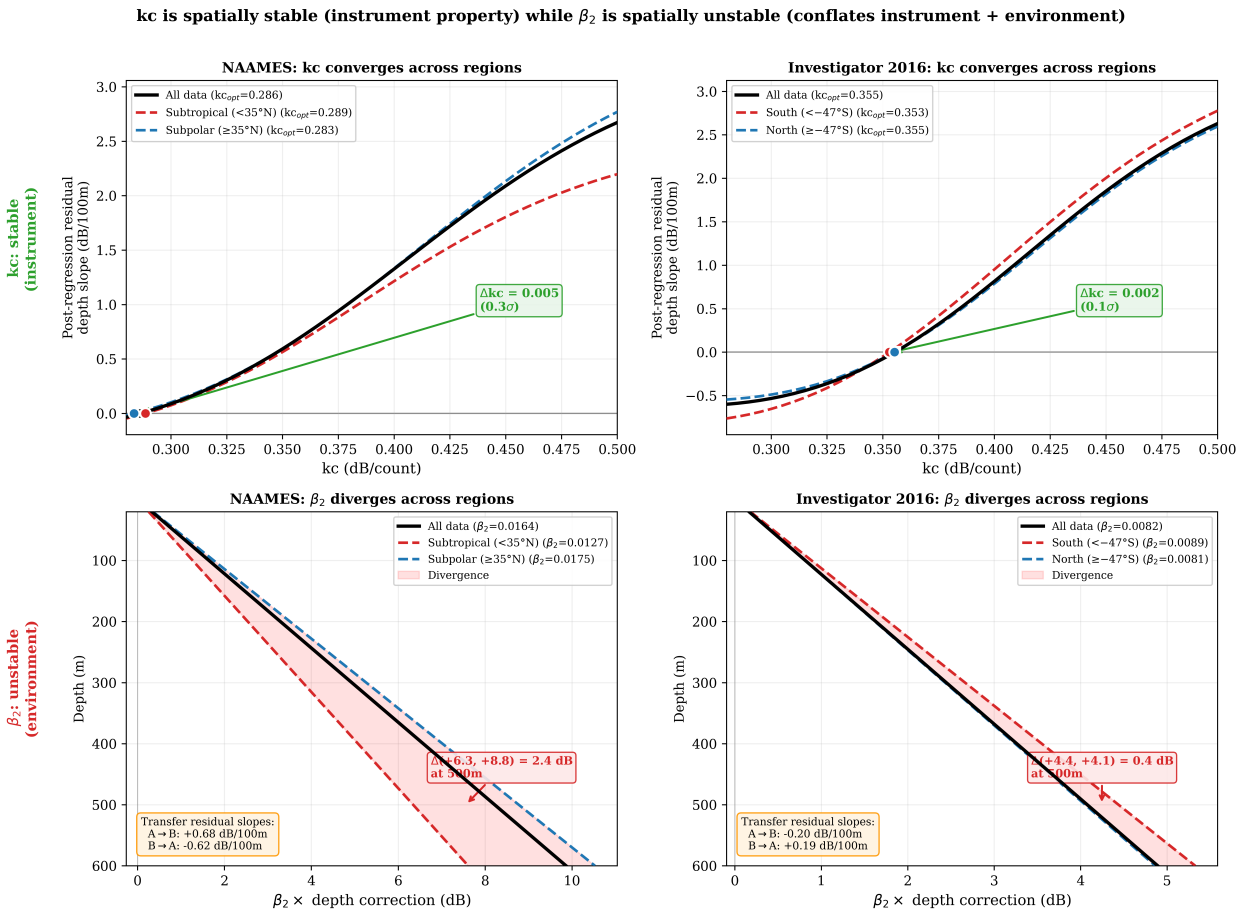

Figure 2.5: k_c converges across geographic subsets (top) while β_2 diverges (bottom).

When k_c is correct, the depth covariate adds nothing: 1D+$k_{c,\text{opt}}$ and 2D+$k_{c,\text{opt}}$ residual profiles overlap, producing identical performance on 5 of 7 datasets (Fig. 2.6).

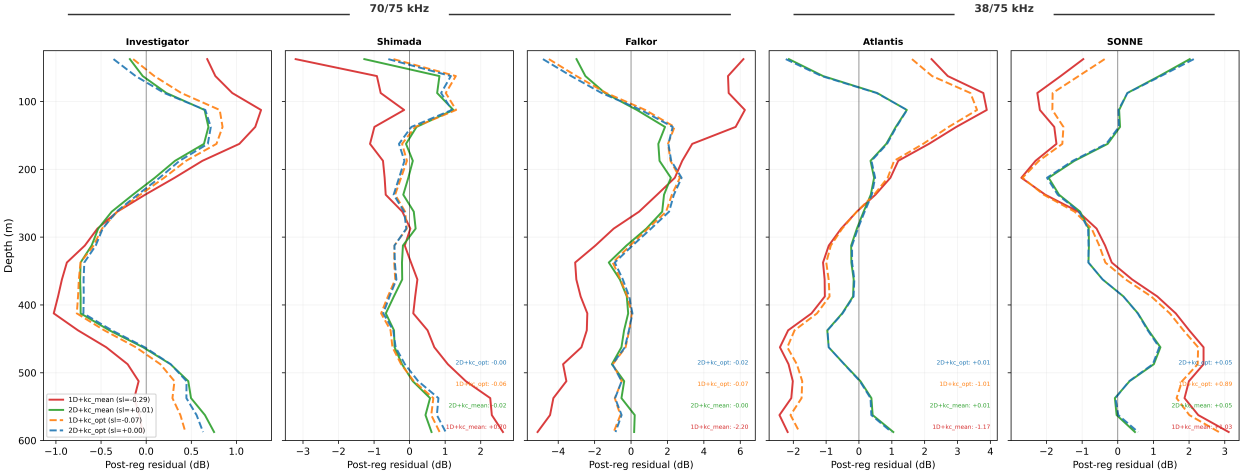

Figure 2.6: Post-regression residual profiles under 4 conditions for 5 vessels.

2.3.4 Cross-frequency robustness

Despite the frequency mismatch between 70/75 and 38/75 kHz pairings, the calibration pipeline produces comparable results across both: near-zero residual depth slopes (0.04 vs 0.06 dB/100 m), similar NASC bias magnitudes, and overlapping k_c ranges (0.25–0.44 vs 0.29–0.46). The regression absorbs frequency-dependent target strength differences through its slope and intercept, consistent with existing frequency-mismatched calibration workflows (Receveur et al. 2020b).

2.3.5 Structural metric robustness

Structural metrics from normalized profiles are invariant to parameterization state and calibration method across 70 paired days. WMD MAE is 6–9 m (70/75 kHz) and 18–22 m (38/75 kHz); migration timing MAE is ∼1 h regardless of state. Default = Regressed and Full = Reg+$k_{c,\mathrm{opt}}$ = 2D+$k_{c,\mathrm{opt}}$, confirming that normalization absorbs any regression. The larger 38/75 kHz WMD MAE likely reflects frequency-dependent scattering differences rather than parameter effects.

2.4 Discussion

2.4.1 S_v uncertainty and the parameter registry

The dominance of k_c and E_r (Section 2.3.1) means that these two parameters represent the primary targets for reducing ADCP S_v uncertainty. The registry assembled here offers a resource for operationalizing archival datasets where instrument parameters are not otherwise available. The 10σ spread of in-situ k_c values (Section 2.3.3) indicates that registry bounds substantially underestimate true variability, making per-instrument determination essential. k_c appears stable across deployments—same-instrument cruise pairs yield consistent values across years and ocean basins—and can be determined from a single paired echosounder cruise.

2.4.2 k_c error survives regression

That RMSE conceals depth-structured k_c bias (Section 2.3.3) means residual depth profiles should be examined routinely. The mechanistic reason is that k_c multiplies $(E - E_r)$, which varies nonlinearly with depth as signal decays. This distortion cannot be absorbed by linear regression. Given that 3 of 5 instruments have k_c outside the registry $\pm 2\sigma$ range, published bias estimates at registry extremes may understate the true impact.

2.4.3 Depth-dependent regression does not substitute for correct parameterization

Depth-dependent regression offers an appealing alternative to explicit k_c determination, but it does not generalize. β_2 conflates the k_c artifact with real oceanographic depth structure, causing it to diverge between water masses and distort real vertical structure when transferred. Once k_c is correct, β_2 adds nothing. The empirical k_c optimization recovers the same value from oceanographically independent subsets (Section 2.3.3), supporting its interpretation as a genuine instrument property.

2.4.4 Structural metrics are robust without calibration

The invariance of structural metrics to all calibration states (Section 2.3.5) retrospectively supports Bianchi and Mislan (2016) and indicates that archival ADCP data are reliable for structural analyses without calibration. Combined with cross-frequency robustness (Section 2.3.4), this opens the possibility of constructing integrated multi-frequency datasets for large-scale studies of scattering layer structure.

2.4.5 Limitations

The JASADCP archive provides only beam-averaged echo intensity; the concave SNR term introduces a systematic bias (Jensen's inequality) that is expected to be small relative to inter-instrument k_c variability but is irreducible from archival data. The E_r cruise-minimum estimator may be biased by transient noise, and the empirical k_c may absorb compensating errors. Broadband receiver ripple ($\sim\pm1.5$ dB) (Deines 1999) is unmodeled. The paired analyses use only OS-75; extension to other classes requires further validation.

ADCP backscatter can be used quantitatively when E_r and k_c are constrained. In the absence of concurrent echosounder data, factory calibration documentation becomes essential. Best practice should include archiving per-beam echo intensity, factory k_c and E_r, serial numbers, and firmware configuration.

2.5 Conclusions

1. ADCP S_v uncertainty is 3–16 dB (2σ, OS-75), dominated by k_c and E_r. All other parameters contribute <4%.

2. E_r correction reduces uncalibrated bias by 4–5 dB. k_c correction eliminates depth-structured distortion.

3. k_c error survives regression—invisible to RMSE, it produces NASC bias up to +430% per depth layer. In-situ k_c spans 10σ of the manufacturer registry.

4. Depth-dependent regression eliminates k_c bias within-cruise but is spatially unstable. Empirical k_c is stable across the geographic subsets tested and preserves vertical biological structure.

5. Structural metrics are invariant to parameterization and calibration method. The calibration pipeline is robust across echosounder frequencies (38–70 kHz).

6. Recommended workflow: E_r (cruise-minimum) $\rightarrow$ T/S (WOA23 monthly) $\rightarrow$ k_c (one paired echosounder cruise per instrument, then reusable).