



Environmental Effects Monitoring Program

2022 Annual Report

January 13, 2023

Fundy Ocean Research Center for Energy
PO Box 2573, Halifax, Nova Scotia, B3J 3N5
902-406-1166
fundyforce.ca

What's New?

FORCE ATTENDS INTERNATIONAL CONFERENCE ON OCEAN ENERGY IN SAN SEBASTIÁN, SPAIN

From October 18 – 20 a small contingent from FORCE attended the ICOE meeting in San Sebastián, Spain. The FORCE team gave presentations to an audience of international delegates and attended sessions on a variety of topics surrounding marine renewable energy. The event also allowed for several in person meetings with tidal energy developers regarding their current and future project plans. [read more](#)

SEABIRD RADAR FEASIBILITY STUDY

FORCE has received a technical report from Strum Consulting which highlights challenges with the radar data as well as options for moving forward with the feasibility study. This work will continue into next year to fully explore the potential utility of radar to monitor for seabirds at the FORCE site. [read more](#)

RAP 2022 FISH TAGGING COMPLETE

The 2022 fish tagging under the Risk Assessment Program (RAP) in collaboration with our partners at the Mi'kmaw Conservation Group (MCG), local fishers, and DFO Science is complete. [read more](#)

VITALITY PLATFORM RECOVERED

After communication with the Vitality platform was interrupted over the summer it was recovered on October 17th to diagnose the cause of the communication failure. Post recovery testing determined that the failure was caused by corrosion of one of the platform grounding electrical cables. All other components continued to function as expected. During its deployment the platform successfully acquired data on current speeds, harbour porpoise localization, and provided a live video feed of underwater life to guests at the FORCE visitor center. [read more](#)

Executive Summary

Tidal stream energy devices are an emerging renewable energy technology that use the ebb and flow of the tides to generate electricity. These devices are in various stages of research, development, operation and testing in countries around the world.

FORCE was established in 2009 after undergoing a joint federal-provincial environmental assessment with the mandate to enable the testing and demonstration of tidal stream devices. Since that time, more than 100 related research studies have been completed or are underway with funding from FORCE, Net Zero Atlantic (formally the Offshore Energy Research Association (OERA)), and others. These studies have considered physical, biological, socioeconomic, and other research areas.

The current suite of monitoring programs implemented by FORCE build off those initiated during 2016-2020 that were conducted in anticipation of tidal stream energy device deployments at FORCE's tidal demonstration site. These efforts are divided into two components: FORCE monitoring activities (>100 metres from a device), and developer or 'device-specific' monitoring led by project developers (\leq 100 metres from a device) at the FORCE site. All plans are reviewed by FORCE's independent Environmental Monitoring Advisory Committee (EMAC) and federal and provincial regulators prior to implementation.

FORCE monitoring presently consists of monitoring for fish, marine mammals, seabirds, lobster, and marine sound. During monitoring from 2016 through 2020, FORCE completed:

- ~564 hours of hydroacoustic fish surveys;
- more than 5,083'C-POD' marine mammal monitoring days;
- bi-weekly shoreline observations;
- 49 observational seabird surveys;
- four drifting marine sound surveys and additional sound monitoring; and
- 11 days of lobster surveys

FORCE submitted its 2021-2023 proposed EEMP to regulators in early 2021 and is awaiting feedback. The 2021-2023 EEMP is designed to prepare for effects testing with the deployment of operational tidal stream energy devices and adheres to the principles of adaptive management by evaluating existing datasets to ensure appropriate monitoring approaches are being implemented. Moreover, the plan adopts internationally accepted standards for monitoring where possible, including feasibility assessments for new monitoring approaches that are planned to be implemented. The 2021-2023 EEMP has been implemented as designed and reviewed by FORCE's environmental monitoring advisory committee (EMAC). Device deployments are pending and there has not been an opportunity for effects testing under the 2021-2023 proposed EEMP.

Since the beginning of the 2021-2023 EEMP, FORCE has completed:

- 8 days of lobster surveys;
- a preliminary radar feasibility study to monitor for seabirds; and
- bi-weekly shoreline observations

FORCE is working with academic and Indigenous partner organizations to advance the Risk Assessment Program (RAP) for tidal stream energy. This program seeks to develop credible and

statistically robust encounter rate models for migratory and resident fish species in Minas Passage with tidal stream energy devices. This will be accomplished by combining physical oceanographic data related to flow and turbulence in the Minas Passage with hydroacoustic tagging information for various fish species in the region curated by the Ocean Tracking Network at Dalhousie University. Since the start of the project, FORCE has established a high-resolution radar network in Minas Passage and has started to quantify hydrodynamic features in the region and build the tidal flow atlas required for the program. FORCE has also started modelling the spatiotemporal distributions for the nine species for which sufficient acoustic tracking data is available and is developing species distribution maps for each species. In partnership with FORCE, the Mi'kmaw Conservation Group (MCG) and local fishers have completed the fish tagging component of the program that is required for species distribution and encounter rate model validation. To share the results of the modelling work, FORCE is currently exploring the development of a user-friendly graphical user interface as a tool for regulators, rights holders, stakeholders, industry, and academia to explore the models. A more advanced science-based decision support tool that can be used run the models is also being explored. Ultimately, this work will contribute towards understanding the risk of tidal stream energy development for fishes in the Bay of Fundy and will assist in the development of future environmental effects monitoring programs.

This report provides a summary of monitoring activities and data analyses completed at the FORCE site up to the end of 2022. In addition, it also highlights findings from international research efforts, previous data collection periods at the FORCE site, and additional research work that is being conducted by FORCE and its partners. This includes supporting fish tagging efforts with Acadia University and the Ocean Tracking Network, radar research projects, and subsea instrumentation platform deployments through the Fundy Advanced Sensor Technology (FAST) Program. Finally, the report presents details regarding future research and monitoring efforts at the FORCE test site. This includes work in support of the 2022 EEMP and the RAP program.

All reports, including quarterly monitoring summaries, are available online at www.fundyforce.ca/document-collection.

Contents

What's New?	1
Executive Summary	2
Appendices	5
Introduction	6
About FORCE	6
Background	7
Tidal Stream Energy Device Deployments	8
International Experience & Cooperation	9
FORCE Monitoring Activities	11
Monitoring Objectives	12
Lobster	13
Fish	15
Marine Mammals	17
Passive Acoustic Monitoring	17
Observation Program	18
Marine Sound (Acoustics)	19
Seabirds	20
Developer Monitoring Activities	21
Other FORCE Research Activities	21
Risk Assessment Program	21
Fundy Advanced Sensor Technology (FAST) Activities	25
Platform Projects	25
Vitality Project	26
Fish Tracking	27
Discussion	29
References	30

Appendices

- Appendix I Acronyms
- Appendix II Viehman et al. (2022). The ups and downs of using active acoustic technologies to study fish at tidal energy sites. *Frontiers in Marine Science* 9: 851400. Doi: 10.3389/fmars.2022.851400
- Appendix III Lowe et al. (2022). Echofilter: A deep learning segmentation model improves the automation, standardization, and timeliness for post-processing echosounder data in tidal energy streams. *Frontiers in Marine Science* 9: 867857. Doi: 10.3389/fmars.2022.867857
- Appendix IV Bangley et al. (2022). Modelling the probability of overlap between marine fish distributions and marine renewable energy infrastructure using acoustic telemetry data. *Frontiers in Marine Science* 9: 851757. Doi: 10.3389/fmars.2022.851757
- Appendix V Hasselman et al. (2022). Editorial: Novel technologies for assessing the environmental and ecological impacts of marine renewable energy systems. *Frontiers in Marine Science* 9: 990327. Doi: 10.3389/fmars.2022.990327

Introduction

This report outlines monitoring activities occurring at the Fundy Ocean Research Centre for Energy test site in the Minas Passage, Bay of Fundy during 2022. Specifically, this report highlights results of environmental monitoring activities conducted by FORCE and other research and development activities conducted at the FORCE site. This report also provides a summary of international research activities around tidal stream energy devices.

About FORCE

FORCE was created in 2009 to lead research, demonstration, and testing for high flow, industrial-scale tidal stream energy devices. FORCE is a not-for-profit entity that has received funding support from the Government of Canada, the Province of Nova Scotia, Encana Corporation, and participating developers.

FORCE has two central roles in relation to the demonstration of tidal stream energy converters in the Minas Passage:

1. Host: providing the technical infrastructure to allow demonstration devices to connect to the transmission grid; and
2. Steward: research and monitoring to better understand the interaction between devices and the environment.

The FORCE project currently consists of five undersea berths for subsea tidal energy device generators, four subsea power cables to connect the devices to land-based infrastructure, an onshore substation and power lines connected to the Nova Scotia Power transmission system, and a Visitor Centre that is free and open to the public from May to November annually. These onshore facilities are located approximately 10 km west of Parrsboro, Nova Scotia.

The marine portion of the project is located in a 1.6 km x 1.0 km tidal demonstration area in the Minas Passage. It is also identified as a Marine Renewable-electricity Area under the Province's Marine Renewable-energy Act. This area consists of five subsea berths that are leased to tidal energy companies¹ selected by the Nova Scotia Department of Natural Resources and Renewables. Current berth holders at FORCE are:

- Berth A: Minas Tidal Limited Partnership
- Berth B: Rio Fundo Operations Canada Limited²
- Berth C: Sustainable Marine Energy (Canada)³
- Berth D: Big Moon Power Canada
- Berth E: Halagonia Tidal Energy Limited⁴

Research, monitoring, and associated reporting is central to FORCE's steward role, to assess whether tidal stream energy devices can operate in the Minas Passage without causing significant adverse effects on the environment, electricity rates, and other users of the Bay.

¹ Further information about each company may be found at: fundyforce.ca/partners

² On April 30, 2019 the Department of Energy and Mines approved the transfer of the Project Agreement and FIT approvals from Atlantis Operations (Canada) Ltd. to Rio Fundo Operations Canada Ltd.

³ On May 15, 2019 the Department of Energy and Mines issued an approval for Black Rock Tidal Power to change its name to Sustainable Marine Energy (Canada) Ltd. with the transfer of assets from SCHOTTEL to Sustainable Marine Energy.

⁴ Berth E does not have a subsea electrical cable provided to it.

As part of this mandate FORCE has a role to play in supporting informed, evidence-based decisions by regulators, industry, rightsholders, the scientific community, and the public. As deployments of different technologies are expected to be phased in over the next several years, FORCE and regulators will have the opportunity to learn and adapt environmental monitoring approaches as lessons are learned.

Background

The FORCE demonstration project received its environmental assessment (EA) approval on September 15, 2009 from the Nova Scotia Minister of Environment. The conditions of its EA approval⁵ provide for comprehensive, ongoing, and adaptive environmental management. The EA approval has been amended since it was issued to accommodate changes in technologies and inclusion of more berths to facilitate provincial demonstration goals.

In accordance with this EA approval, FORCE has been conducting an Environmental Effects Monitoring Program (EEMP) to better understand the natural environment of the Minas Passage and the potential effects of tidal stream energy devices as related to fish, seabirds, marine mammals, lobster, marine sound, benthic habitat, and other environmental variables. All reports on site monitoring are available online at: www.fundyforce.ca/document-collection.

Since 2009, more than 100 related research studies have been completed or are underway with funding from FORCE, Net Zero Atlantic (formally the Offshore Energy Research Association (OERA)) and others. These studies have considered socioeconomics, biological, and other research areas.⁶

Monitoring at the FORCE site is currently focused on lobster, fish, marine mammals, seabirds, and marine sound and is divided into developer (≤ 100 m from a device) and FORCE led (> 100 m from a device) monitoring. As approved by regulators, individual berth holders complete monitoring in direct vicinity of their device(s), in recognition of the unique design and operational requirements of different technologies. FORCE completes site level monitoring activities as well as supporting integration of data analysis between these monitoring zones, where applicable.

All developer and FORCE monitoring programs are reviewed by FORCE's Environmental Monitoring Advisory Committee (EMAC), which includes representatives from scientific, First Nations, and local fishing communities.⁷ These programs are also reviewed by federal and provincial regulators prior to device installation. In addition, FORCE and berth holders also submit an Environmental Management Plan (EMP) to regulators for review prior to device installation. EMP's include environmental management roles and responsibilities and commitments, environmental protection plans, maintenance and inspection requirements, training and education requirements, reporting protocols, and more.

⁵ FORCE's Environmental Assessment Registration Document and conditions of approval are found online at: www.fundyforce.ca/document-collection.

⁶ Net Zero Atlantic Research Portal (<https://netzeroatlantic.ca/research>) includes studies pertaining to infrastructure, marine life, seabed characteristics, socio-economics and traditional use, technology, and site characterization.

⁷ Information about EMAC may be found online at: www.fundyforce.ca/about-us

Tidal Stream Energy Device Deployments

Since FORCE's establishment in 2009, tidal stream energy devices have been installed at the FORCE site three times: once in 2009/2010, November 2016 – June 2017, and July 2018 – present. Given the limited timescales in which a device has been present and operating at the FORCE site, environmental studies to-date have largely focused on the collection of baseline data and developing an understanding of the capabilities of monitoring devices in high flow tidal environments.

On July 22, 2018, CSTV installed a two-megawatt OpenHydro turbine at Berth D of the FORCE site and successfully connected the subsea cable to the turbine. CSTV confirmed establishment of communication with the turbine systems on July 24. On July 26, 2018, Naval Energies unexpectedly filed a petition with the High Court of Ireland for the liquidation of OpenHydro Group Limited and OpenHydro Technologies Limited.⁸ For safety purposes, the turbine was isolated from the power grid that same day. On September 4, 2018, work began to re-energize the turbine, but soon afterwards it was confirmed that the turbine's rotor was not turning. It is believed that an internal component failure in the generator caused sufficient damage to the rotor to prevent its operation. Environmental sensors located on the turbine and subsea base continued to function at that time except for one hydrophone.

As a result of the status of the turbine, the monitoring requirements and reporting timelines set out in CSTV's environmental effects monitoring program were subsequently modified under CSTV's Authorization from Fisheries and Oceans Canada. The modification requires that CSTV provide written confirmation to regulators monthly that the turbine is not spinning by monitoring its status during the peak tidal flow of each month. This began October 1, 2018 and was expected to continue until the removal of the turbine; however, as a result of the insolvency of OpenHydro Technology Ltd., all developer reporting activities by CSTV ceased as of March 1, 2019. FORCE subsequently provided monthly reports to regulators confirming the continued non-operational status of the CSTV turbine from March 2019 – May 2020 and received authorization from the Nova Scotia Department of Environment on June 2, 2020, to conclude these monthly reports.

In September 2020, Big Moon Canada Corporation (Big Moon) was announced as the successful applicant to fill berth D at the FORCE test site following a procurement procedure administered by Power Advisory LLC. As part of the agreement, Big Moon has provided a \$4.5 million security deposit to remove the non-operational CSTV turbine currently deployed at berth D, and has until December 31, 2024 to raise the turbine. The project start date for BigMoon is largely dependent on the economic recovery from the COVID-19 pandemic and the potential impact to Big Moon's supply chain. As such, the project start date is not known at this time.

Additional devices are expected to be deployed at the FORCE site in the coming years. In 2018, Sustainable Marine Energy (formerly Black Rock Tidal Power) installed a PLAT-I system in Grand Passage, Nova Scotia under a Demonstration Permit.⁹ This permit allows for a demonstration of the 280 kW system to help SME and its partners learn about how the device operates in the marine environment of the Bay of Fundy. On May 11, 2022, SME announced it has successfully delivered the first floating tidal stream energy to Nova Scotia's power grid.¹⁰ Also in 2018, Natural Resources Canada announced a \$29.8 million contribution to Halagonia Tidal Energy's project at

⁸ See original news report: <https://www.irishexaminer.com/breakingnews/business/renewable-energy-firms-with-more-than-100-employees-to-be-wound-up-857995.html>.

⁹ To learn more about this project, see: <https://novascotia.ca/news/release/?id=20180919002>.

¹⁰ To learn more about this project, see: <https://www.sustainablemarine.com/press-releases/sustainable-marine-delivers-first-floating-tidal-power-to-nova-scotia-grid>.

the FORCE site through its Emerging Renewable Power Program.¹¹ The project consists of submerged turbines for a total of nine megawatts – enough capacity to provide electricity to an estimated 2,500 homes.

Each berth holder project will be required to develop a device-specific monitoring program, which will be reviewed by FORCE's EMAC and federal and provincial regulators including Fisheries and Oceans Canada, the Nova Scotia Department of Environment, and the Nova Scotia Department of Energy and Mines prior to device installation.

Overall, the risks associated with single device or small array projects are anticipated to be low given the relative size/scale of devices (Copping 2018). For example, at the FORCE site a single two-megawatt OpenHydro turbine occupies ~ 1/1,000th of the cross-sectional area in the Minas Passage (Figure 1). A full evaluation of the risks of tidal stream energy devices, however, will not be possible until more are tested over a longer-term period with monitoring that documents local impacts, considers far-field and cumulative effects, and adds to the growing global knowledge base.



Figure 1: The scale of a single turbine (based on the dimensions of the OpenHydro turbine deployed by CSTV, indicated by the red dot and above the blue arrow) in relation to the cross-sectional area of the Minas Passage. The Passage reaches a width of ~ 5.4 km and a depth of 130 m.

International Experience & Cooperation

The research and monitoring being conducted at the FORCE test site is part of an international effort to evaluate the risks tidal energy poses to marine life (Copping 2018; Copping and Hemery 2020). Presently, countries such as China, France, Italy, the Netherlands, South Korea, the United Kingdom, and the United States (Marine Renewables Canada 2018) are exploring tidal energy, supporting environmental monitoring and innovative R&D projects. Tidal energy and other marine renewable energy (MRE) technologies such as tidal range, tidal current, wave, and ocean thermal energy offer significant opportunities to replace carbon fuel sources in a meaningful and permanent manner. Some estimates place MRE's potential as exceeding current human energy needs (Lewis et al. 2011; Gattuso et al. 2018). Recent research includes assessments of operational sounds on marine fauna (Schramm et al. 2017; Lossent et al. 2018; Robertson et al. 2018; Pine et al. 2019), the utility of PAM sensors for monitoring marine mammal interactions with turbines (Malinka et al. 2018) and collision risk (Joy et al. 2018b), demonstrated avoidance behavior by harbour porpoise around tidal turbines (Gillespie et al. 2021), a synthesis of known effects of marine renewable energy devices on fish (Copping et al. 2021), and the influence of tidal turbines on fish behavior (Fraser et al. 2018).

Through connections to groups supporting tidal energy demonstration and R&D, FORCE is working to inform the global body of knowledge pertaining to environmental effects associated

¹¹ To learn more about this announcement, see: <https://www.canada.ca/en/natural-resources-canada/news/2018/09/minister-sohi-announces-major-investment-in-renewable-tidal-energy-that-will-power-2500-homes-in-nova-scotia.html>.

with tidal power projects. This includes participation in the Fundy Energy Research Network¹², the Bay of Fundy Ecosystem Partnership¹³, TC114¹⁴, the Atlantic Canadian-based Ocean Supercluster¹⁵, and OES-Environmental¹⁶.

FORCE will continue to work closely with OES-Environmental and its members to document and improve the state of knowledge about the interactions of MRE devices interactions with the marine environment. To that end, Dr. Hasselman served as a guest editor alongside Dr. Huidong Li (Pacific Northwest National Laboratory), Dr. Emma Cotter (Woods Hole Oceanographic Institute) and Dr. James Joslin (University of Washington) for a special issue of *Frontiers in Marine Science* entitled '[Novel Technologies for Assessing the Environmental and Ecological Impacts of Marine Renewable Energy Systems](#)'. The editorial team advertised the special issue on the Tethys website and received nine abstracts from researchers developing cutting-edge technologies for monitoring around marine renewable energy devices. Full manuscript submissions were due by January 9, 2022 and the special issue is now published¹⁷. A selection of the published papers are found in Appendices II – V.

Additionally, OES-Environmental is pursuing the development of new research topics for the 2024 State of the Science Report related to i) knowledge of environmental effects as the tidal energy industry scales up from single devices to arrays, ii) understanding the cumulative impacts of marine renewable energy with other anthropogenic effects, and iii) an ecosystem approach for understanding environmental effects, including interactions between trophic levels, between ecosystems and between ecosystem services. Dr. Hasselman is involved in the development of all three of these topics but is leading the effort to understand the environmental effects of 'scaling up'.

On April 7th, FORCE hosted a half-day workshop in Halifax focused on advancing understanding of the precautionary principle as it applies to the regulation of the marine renewable energy sector. The purpose of this workshop was to bring together stakeholders, rights holders and regulators to better understand the precautionary principle and the benefits and challenges it brings to regulating and advancing tidal development. The workshop consisted of three speakers who shared their knowledge on the precautionary principle which had 39 participants (29 in person and 10 virtual). Presentations were followed by breakout group discussions. A report on key takeaways is currently in development and will be available later this year.

¹² FERN is a research network designed to " coordinate and foster research collaborations, capacity building and information exchange" (Source: fern.acadiau.ca/about.html). FORCE participates in the Natural Sciences, Engineering, and Socio-Economic Subcommittees of FERN.

¹³ BoFEP is a 'virtual institute' interested in the well-being of the Bay of Fundy. To learn more, see www.bofep.org.

¹⁴ TC114 is the Canadian Subcommittee created by the International Electrotechnical Commission (IEC) to prepare international standards for marine energy conversion systems. Learn more: tc114.oreg.ca.

¹⁵ The OSC was established with a mandate to "better leverage science and technology in Canada's ocean sectors and to build a digitally-powered, knowledge-based ocean economy." Learn more: www.oceansupercluster.ca.

¹⁶ OES Environmental was established by the International Energy Agency (IEA) Ocean Energy Systems (OES) in January 2010 to examine environmental effects of marine renewable energy development. Member nations include: Australia, China, Canada, Denmark, France, India, Ireland, Japan, Norway, Portugal, South Africa, Spain, Sweden, United Kingdom, and United States. Further information is available at <https://tethys.pnnl.gov>.

¹⁷ Novel Technologies for Assessing the Environmental and Ecological Impacts of Marine Renewable Energy Systems - <https://www.frontiersin.org/research-topics/19503/novel-technologies-for-assessing-the-environmental-and-ecological-impacts-of-marine-renewable-energy#articles>

Dr. Hasselman chaired a session at the Pan American Marine Energy Conference in Ensenada, Mexico on June 17th entitled 'The role of MRE test centers in facilitating MRE development'. The afternoon session included 5 presentations which provided an overview of the role of marine renewable energy (MRE) test centers in device testing and their capacity to demonstrate the utility of monitoring technologies and approaches for understanding environmental effects of MRE devices. The presentations and round table discussions built on the first workshop held at PAMEC 2020 in Costa Rica and fostered dialogue around the value of test centers like FORCE and their role in building social license and addressing questions relevant to the establishment of MRE technologies regionally.

On August 22nd Dr. Hasselman attended the American Fisheries Society 152nd annual meeting held in Spokane, Washington. During a symposium on offshore wind in the United States, Dr. Hasselman presented on the Risk Assessment Program (RAP) species distribution modelling framework. Species distribution models are a commonly used tool for understanding the effects of offshore wind projects as well, and it is encouraging that this approach is applicable across marine renewable energy sectors.

From October 18 – 20 a small contingent from FORCE attended the International Conference on Ocean Energy (ICOE) in San Sebastián, Spain which consisted of Lindsay Bennett, Dan Hasselman, and Joel Culina. Dr. Hasselman participated in a session entitled "Determining the environmental impact of a marine energy installation" and gave a presentation titled "Scaling up our understanding of environmental effects of MRE devices from single devices to large-scale commercial arrays". Dan co-authored two additional presentations entitled "Cumulative effects assessment for marine renewable energy" presented by Dr. Lenaïg Hemery (OES-Environmental) and "The ecosystem approach applied to marine renewable energies: application and knowledge gaps" presented by Dr. Georges Safi at France Energie Marines. Dan also participated in a workshop co-hosted by OES-Environmental and Aquaterra entitled "Environmental monitoring around deployed marine energy devices" as well as several meetings with tidal energy developers regarding current and future projects. Lindsay participated in a session entitled "Slashing consenting time to 12 months (or less)" and gave a presentation titled "Building collaborative approaches to risk assessment and project approval". Finally, Joel co-authored a poster presented at the conference entitled "Can we trust turbulence models?" which won first prize at the poster session.

FORCE Monitoring Activities

FORCE has been leading site-level monitoring for several years, focusing on a variety of environmental variables. FORCE's previous environmental effects monitoring program (2016-2020) was developed in consultation with SLR Consulting (Canada)¹⁸ and was strengthened by review and contributions by national and international experts and scientists, DFO, NSE, and FORCE's EMAC. The most recent version of the EEMP (2021-2023) was developed in consultation with Atlantis Watershed Consultants Ltd. with input from national and international experts, including FORCE's EMAC, and has been submitted to regulators for approval. The 2021-2023 EEMP has been modified from the 2016-2020 EEMP based on results of previous monitoring activities, experience and lessons learned. This is consistent with the adaptive management approach inherent to the FORCE EEMP – the process of monitoring, evaluating

¹⁸ This document is available online at: www.fundyforce.ca/document-collection.

and learning, and adapting (AECOM 2009) that has been used at the FORCE site since its establishment in 2009.¹⁹

FORCE's EEMP currently focuses on the impacts of operational tidal stream energy devices on lobster, fish, marine mammals, and seabirds as well as the impact of device-produced sound. Overall, these research and monitoring efforts, detailed below, were designed to test the predictions made in the FORCE EA. As mentioned in the Executive Summary, since the beginning of the 2016-2020 EEMP, FORCE has completed approximately:

- 564 hours of hydroacoustic fish surveys;
- more than 5,083'C-POD' (marine mammal monitoring) days;
- bi-weekly shoreline observations;
- 49 observational seabird surveys;
- four drifting marine sound surveys and additional bottom-mounted instrument sound data collection; and
- 11 days of lobster surveys.

Since the beginning of the 2021-2023 EEMP, FORCE has completed:

- 8 days of lobster surveys;
- a preliminary radar feasibility study to monitor for seabirds; and
- bi-weekly shoreline observations

The following pages provide a summary of the site-level monitoring activities conducted at the FORCE site up to the end of 2022 including data collection, data analyses performed, initial results, and lessons learned, that builds on activities and analyses from previous years. Where applicable, this report also presents analyses that have integrated data collected through developer and FORCE monitoring programs to provide a more complete understanding of device-marine life interactions.

Monitoring Objectives

The overarching purpose of environmental monitoring is to test the accuracy of the environmental effect predictions made in the original EA. These predictions were generated through an evaluation of existing physical, biological, and socioeconomic conditions of the study area, and an assessment of the risks the tidal energy demonstration project poses to components of the ecosystem.

A comprehensive understanding of device-marine life interactions will not be possible until device-specific and site-level monitoring efforts are integrated, and additional data is collected in relation to operating tidal stream energy devices. Further, multi-year data collection will be required to consider seasonal variability at the FORCE test site and appropriate statistical analyses of this data will help to obtain a more complete understanding of device-marine life interactions.

Table 1 outlines the objectives of the site-level monitoring activities conducted at the FORCE demonstration site. FORCE led site-level monitoring summaries will be updated as devices are scheduled for deployment at FORCE. At this time, and considering the scale of device

¹⁹ The adaptive management approach is necessary due to the unknowns and difficulties inherent with gathering data in tidal environments such as the Minas Passage and allows for adjustments and constant improvements to be made as knowledge about the system and environmental interactions become known. This approach has been accepted by scientists and regulators.

deployments in the near-term at FORCE, it is unlikely that significant effects in the far-field will be measurable (SLR Consulting 2015). Far-field studies such as sediment dynamics will be deferred until such time they are required. However, recent discussions with scientists serving on FORCE's EMAC suggests that the natural variability inherent to the upper Bay of Fundy ecosystem far exceeds what could be measured by far-field monitoring efforts. Moreover, the scale of tidal power development would need to surpass what is possible at the FORCE tidal demonstration site to extract sufficient energy from the system to have any measurable effects. In short, far-field monitoring would be futile unless tidal power development transitions from demonstration scale to commercial arrays. As more devices are scheduled for deployment at the FORCE site and as monitoring techniques are improved, monitoring protocols will be revised in keeping with the adaptive management approach. These studies will be developed in consultation with FORCE's EMAC, regulators, and key stakeholders.

Table 1: The objectives of each of the environmental effects monitoring activities, which consider various Valued Ecosystem Components (VECs), led by FORCE.

FORCE Environmental Effects Monitoring VEC	Objectives
Lobster	<ul style="list-style-type: none"> to determine if the presence of a tidal stream energy device affects commercial lobster catches
Fish	<ul style="list-style-type: none"> to test for indirect effects of tidal stream energy devices on water column fish density and fish vertical distribution to estimate probability of fish encountering a device based on fish density proportions in the water column relative to device depth in the water column
Marine Mammals	<ul style="list-style-type: none"> to determine if there is permanent avoidance of the study area during device operations to determine if there is a change in the distribution of a portion of the population across the study area
Marine Sound (Acoustics)	<ul style="list-style-type: none"> to conduct ambient sound measurements to characterize the soundscape prior to and following deployment of the tidal stream energy device
Seabirds	<ul style="list-style-type: none"> to understand the occurrence and movement of bird species in the vicinity of tidal stream energy devices to confirm FORCE's Environmental Assessment predictions relating to the avoidance and/or attraction of birds to tidal stream energy devices

Lobster

FORCE conducted a baseline lobster catchability survey in fall 2017 (NEXUS Coastal Resource Management Ltd. 2017). This catch-and-release survey design was conducted over 11 days and consisted of commercial traps deployed at varying distances around the future location of the CSTV turbine deployment planned for 2018. Captured lobsters were measured (carapace length), had their sex and reproductive stage determined (male, female, and berried female), and shell condition evaluated. This baseline survey captured 351 lobsters and reported a high catchability rate (> 2.7 kg/trap).²⁰ Preliminary qualitative analyses indicated that catch rates declined during the survey and were associated with increasing tidal velocities; a statistically significant negative

²⁰ This is classified as 'high' according to DFO's Catch Per Unit Effort (CPUE) index (Serdynska and Coffen-Smout, 2017).

relationship was detected between catch rates and maximum tidal range. No significant difference in catch rates was detected across separate locations from the proposed turbine deployment site. Cumulatively, these results suggested that the impact of turbines may be higher on lobster catchability than anticipated in the EA (AECOM 2009), but a repeat of the study in the presence of an operational device is required to verify this prediction.

Indeed, a repeat of this catchability survey was planned for fall 2018 in the presence of an operational turbine to test the EA prediction (with pre-installation and operating turbine collection periods) that tidal stream energy devices will have minimal have impacts on lobster populations within the FORCE test site (AECOM 2009). However, given the non-operational status of the CSTV turbine, the objectives of the 2018 survey effort could not be achieved, and the survey has been postponed until an operational device is present at the site.

In 2019, FORCE commissioned TriNav Fisheries Consultants Ltd. to redesign FORCE's lobster monitoring program based on feedback from regulators to include a more statistically robust study design for monitoring lobster at the FORCE test site. TriNav Fisheries Consultants evaluated the efficacy of using a variety of methods including divers and hydroacoustic tags to track lobster movements. However, given the strong tidal flows and brief window available during periods of slack tide, divers are not a viable option due to safety concerns. Ultimately, TriNav Fisheries Consultants identified the combination of a modified catchability survey design and a mark-recapture study using conventional tags as the best approach for monitoring lobster at the FORCE site. This study design was implemented in fall 2021 in partnership with the Fishermen and Scientists Research Society (FSRS; Figure 1) and with the assistance of a local lobster fisher. There were two phases to the study – each centered around the two neap tide phases in September to ensure trap recovery. During each phase, nine experimental lobster traps were deployed in and around the FORCE tidal demonstration site. Traps were hauled after 24 hours and lobsters were measured, assessed, and tagged prior to being released back to the water. The first phase of the study occurred during August 29-September 2, and the second phase took place during September 27-October 1. The study captured 582 lobster and tagged and released 477 of them – some of which were recaptured during the commercial lobster season in LFA 35, and their tag numbers and capture coordinates reported to FORCE. Preliminary results suggest a high catchability rate during the fall survey which is comparable to available commercial data from DFO. The final report from this monitoring program is currently undergoing edits and will be available soon. Shaun Allain, FORCE Environmental Program Manager, presented the initial results of this survey at the FSRS annual conference and AGM on March 24th.



Figure 1: Lobster scientist from the Fishermen and Scientist Research Society showing a tagged lobster prior to release.

Fish

FORCE has been conducting mobile fish surveys since May 2016 to test the EA prediction that tidal stream energy devices are unlikely to cause substantial impacts to fishes at the test site (AECOM 2009). To that end, the surveys are designed to:

- test for indirect effects of tidal stream energy devices on water column fish density and fish vertical distribution; and
- estimate the probability of fish encountering a device based on any ‘co-occurrence’ relative to device depth in the water column.

Moreover, these surveys follow a ‘BACI’ (Before/After, Control/Impact) design to permit a comparison of data collected before a device is installed with data collected while a device is operational at the FORCE site, and in relation to a reference site along the south side of the Minas Passage. These 24-hour mobile surveys encompass two tidal cycles and day/night periods using a scientific echosounder, the Simrad EK80, mounted on a vessel, the Nova Endeavor (Huntley’s Sub-Aqua Construction, Wolfville, NS). This instrument is an active acoustic monitoring device and uses sonar technology to detect fish by recording reflections of a fish’s swim bladder.

Analyses of hydroacoustic fish surveys completed during baseline studies in 2011 and 2012 (Melvin and Cochrane 2014) and surveys during May 2016 – August 2017 (Daroux and Zydlewski 2017) evaluated changes in fish densities in association with diel stage (day/night), tidal stage (ebb/flood), and device presence or absence (an OpenHydro turbine was present November 2016 – June 2017). Results support the EA prediction that tidal stream devices have minimal impact on marine fishes. However, additional surveys in relation to an operating device are required to fully test this prediction.

In 2019, the University of Maine conducted a thorough analysis for 15 fish surveys conducted by FORCE from 2011-2017. The hydroacoustic data set included six ‘historical’ surveys conducted between August 2011 and May 2012, and nine ‘contemporary’ surveys conducted between May 2016 and August 2017. The analyses included comparisons of fish presence/absence and relative fish density with respect to a series of temporal (historical vs. contemporary, or by survey), spatial (CLA vs. reference study area, or by transect) and environmental (tide phase, diel state, or with/against predicted tidal flow) explanatory variables. The report identified a statistically significant difference in fish presence/absence and relative fish density between the historical and contemporary data sets that may be attributable to differences in the survey design/execution between the time periods, or could reflect changes in fish usage of the site. As such, remaining analyses were restricted to the contemporary data sets. The results revealed that: i) data collection during the ebb tide and at night are important for understanding fish presence in the CLA, ii) various explanatory variables and their additive effects should be explored further, and iii) increasing the frequency of surveys during migratory periods (consecutive days in spring/fall) may be required to understand patterns and variability of fish presence and density in Minas Passage. Importantly, the report suggested a statistically significant difference in fish presence/absence and relative density between the CL and reference site, suggesting that the reference site may not be sufficiently representative to serve as a control for the CLA, and for testing the effects of an operational device on fish density and distribution in Minas Passage. Additional work is underway using data from eight additional contemporary fish surveys (2017-2018) to determine whether this finding is biologically meaningful, or whether it is simply a statistical artefact of how the data was aggregated in the original analysis.

Because complex hydrodynamic features of the Minas Passage introduce turbulence and bubbles into the water column that interfere with the use of hydroacoustics, FORCE’s mobile fish surveys have been optimized for collecting data during the best neap tidal cycle per month when turbulence is greatly reduced. However, this approach limits the number of surveys that can be conducted, and regulators have suggested that the scope of the program be expanded so that survey results are more representative of how fish use the Minas Passage. To that end, FORCE conducted multiple fish surveys during each of three neap tidal cycles in fall 2020 (i.e., September 25, 27, 29; October 7, 9, 13; and October 24, 26, 29) to determine whether variation in fish density and distribution for any given survey within a neap cycle was representative of the other surveys conducted during that same time frame. Previous work comparing stationary and mobile hydroacoustic surveys in Minas Passage found that the temporal representative range of a 24-hr mobile was approximately three days (Viehman et al. 2019).

A recent study ([Viehman et al. \(2022\)](#), Appendix II) examined entrained air contamination in echosounder data collected at the FORCE test site. It found that fish abundance estimates in the lower 70% of the water column and current speeds less than 3 m/s were well represented in that there was little contamination of the data set from entrained air. However, undersampling of the upper water column and faster speeds strongly affected fish abundance estimates especially during strong spring tides. This means that data collected during neap tides are more likely to yield a more accurate picture of fish abundance and distribution than those collected during spring tides. The study also highlighted how estimates of fish abundance may be affected differently depending on where fish are in the water column. For example, (hypothetical) fish located at mid-depths were omitted from the data more often as current speeds increased. These findings indicate a complex and dynamic ecosystem where the interactions of water movement and fish distribution affect our ability to infer how fish populations may interact with tidal power devices in the Minas Passage. The use of acoustic telemetry being studied under the RAP program could be used concurrently with echosounders to fill gaps in datasets and optimize what can be learned about fish abundance and distribution at tidal energy sites.

Another issue with the entrained air found in high flow environments is the need to remove the contaminated data prior to analysis which is often a tedious and time-consuming process. Existing algorithms used to identify the depth-of-penetration of entrained air are insufficient for a boundary that is discontinuous, depth-dynamic, porous, and varies with tidal flow speed. Using a case study from data obtained at the FORCE test site a recent study ([Lowe et al. \(2022\)](#), Appendix III) described the development and application of a deep machine learning model called Echofilter. Echofilter was found to be highly responsive to dynamic range of turbulence conditions in the data and produced an entrained-air boundary line with an average error of less than half that of the existing algorithms. The model had a high level of agreement with human data trimming. This resulted in 50% reduction in the time required for manual edits to the data set when using currently available algorithms to trim the data.

FORCE is currently working towards a development of a comprehensive fish synthesis that will bring together existing knowledge of fish distribution, abundance, and use of the Minas Passage using existing literature from stock assessments, prior hydroacoustic surveys, acoustic telemetry-based surveys, as well as other relevant sources of information. This synthesis will focus on species of conservations concern, cultural relevance, and commercial and recreational value. The results of this synthesis project will be available in 2023 and will help to determine the extent to which questions regarding fish and tidal energy project permitting have been answered and identifying remaining knowledge gaps. Dr. Graham Daborn at Acadia University is leading this work and a final report is expected early in 2023.

Marine Mammals

Since 2016, FORCE has been conducting two main activities to test the EA prediction that project activities are not likely to cause significant adverse residual effects on marine mammals within the FORCE test site (AECOM 2009):

- passive acoustic monitoring (PAM) using ‘click recorders’ known as C-PODs; and
- an observation program that includes shoreline, stationary, and vessel-based observations.

Passive Acoustic Monitoring

The first component of FORCE’s marine mammal monitoring program involves the use of PAM mammal detectors known as C-PODs, which record the vocalizations of toothed whales, porpoises, and dolphins.²¹ The program focuses mainly on harbour porpoise – the key marine mammal species in the Minas Passage that is known to have a small population that inhabits the inner Bay of Fundy (Gaskin 1992). The goal of this program is to understand if there is a change in marine mammal presence in proximity to a deployed tidal stream energy device and builds upon baseline C-POD data collection within the Minas Passage since 2011.

From 2011 to early 2018, more than 4,845 ‘C-POD days’²² of data were collected in the Minas Passage. Over the study period, it was found that harbour porpoise use and movement varies over long (i.e., seasonal peaks and lunar cycles) and short (i.e., nocturnal preference and tide

²¹ The C-PODs, purchased from Chelonia Limited, are designed to passively detect marine mammal ‘clicks’ from toothed whales, dolphins, and porpoises.

²² A ‘C-POD day’ refers to the number of total days each C-POD was deployed times the number of C-PODs deployed.

stage) timescales. This analysis, completed by Sea Mammal Research Unit (Canada) (Vancouver, BC), showed some evidence to suggest marine mammal exclusion within the vicinity of CSTV turbine when it was operational (November 2016 – June 2017) (Joy et al. 2018a). This analysis revealed that the C-PODs in closest proximity to the turbine (230 m and 210 m distance) had reduced frequency of detections, but no evidence of site avoidance with a device present and operating. These findings also revealed a decrease in detections during turbine installation activities, consistent with previous findings (Joy et al. 2017), but requiring additional data during an operational device to permit a full assessment of the EA predictions.

This monitoring program demonstrates the prevalence of harbour porpoise at FORCE, with the species being detected on 98.8% of the 1,888 calendar days since monitoring with C-PODs commenced in 2011. Harbour porpoise detections at FORCE varies seasonally, with peak activity occurring during May – August, and lowest detections during December – March. Harbour porpoise detections also vary spatially, with C-PODs deployed at locations W2 and S2 recording the greatest detection rates, and D1 values typically low. Mean lost time across C-PODs, due to ambient flow noise saturating the detection buffer on the C-POD, averaged 22.6%. Interestingly, an analysis against past datasets that controlled for time of year, indicated that the effects of the non-operational CSTV turbine structure had no detectable effect on the rate of harbour porpoise detection.

SMRU provided their 4th year final report of harbour porpoise monitoring using C-PODs at the FORCE test site (Palmer et al. 2021). The report describes the results of C-POD deployments #11-12 (i.e., 1,043 days of monitoring from August 2019 – September 2020), and places the results in the broader context of the overall marine mammal monitoring program at FORCE. The final report includes summary data that revealed that harbour porpoise was detected on a least one C-POD every day, with a median value of 11 and 17 minutes of porpoise detections per day during deployments 11 and 12, respectively. The mean percent lost time due to ambient flow and sediment noise was 19.5% and 23.8%, respectively, comparable to previous deployments. Overall, the final report supports previous findings of monitoring activities that harbour porpoise are prevalent at the FORCE test site.

The final report also reiterates that sufficient baseline data has been collected to meet the goals of the EEMP. As such, FORCE has recommended in its 2021-2023 EEMP proposal that the collection of additional baseline harbour porpoise data using C-PODs be suspended until an operational device is deployed at the FORCE site. Upon receiving confirmation that a device will be deployed at the tidal demonstration area, FORCE will deploy C-PODs prior to the construction phase to begin collecting data and assessing any changes to harbour porpoise detections in the presence of an operational device. FORCE is currently working with SMRU to continue with this monitoring program when operational devices are present.

*Harbor porpoise (*Phocoena phocoena*) monitoring at the FORCE Test Site, Canada featured on Tethys (by FORCE and SMRU): <https://tethys.pnnl.gov/stories/harbor-porpoise-phocoena-phocoena-monitoring-force-test-site-canada>*

Observation Program

FORCE's marine mammal observation program in 2022 includes observations made during bi-weekly shoreline surveys, stationary observations at the FORCE Visitor Centre, and marine-

based observations during marine operations. All observations and sightings are recorded, along with weather data, tide state, and other environmental data. Any marine mammal observations will be shared with SMRU Consulting to support validation efforts of PAM activities when C-PODs are deployed.

FORCE uses an Unmanned Aerial Vehicle (UAV) for collecting observational data along the shoreline and over the FORCE site using transects by programming GPS waypoints in the UAV to standardize flight paths. FORCE staff received training to operate FORCE's UAV and have acquired UAV pilot certification by successfully passing the 2019 Canadian Drone Pilot Basic Operations Examination, administered by Transport Canada. These staff are now licensed to safely operate the UAV at the FORCE site. FORCE also hosts a public reporting tool that allows members of the public to report observations of marine life: mmo.fundyforce.ca

Marine Sound (Acoustics)

Marine sound – often referred to as ‘acoustics’ or ‘noise’ – monitoring efforts are designed to characterize the soundscape of the FORCE test site. Data collected from these monitoring efforts will be used to test the EA predictions that operational sounds produced from functioning tidal stream energy devices are unlikely to cause mortality, physical injury or hearing impairment to marine animals (AECOM 2009).

Results from previous acoustic analyses completed at the FORCE site indicate that the CSTV turbine was audible to marine life at varying distances from the turbine, but only exceeded the threshold for behavioural disturbance at very short ranges and during particular tide conditions (Martin et al. 2018). This is consistent with findings at the Paimpol-Bréhat site in France where an OpenHydro turbine was also deployed – data suggests that physiological trauma associated with a device is improbable, but that behavioural disturbance may occur within 400 m of a device for marine mammals and at closer distances for some fish species (Lossent et al. 2018).

In previous years, regulators have encouraged FORCE to pursue integration of results from multiple PAM instruments deployed in and around the FORCE test site. To that end, FORCE, and its partner JASCO Applied Sciences (Canada) Ltd. pursued a comparative integrated analysis of sound data collected by various hydrophones (i.e., underwater sound recorders) deployed autonomously and mounted on the CSTV turbine. That work revealed that flow noise increased with the height of the hydrophone off the seabed but had little effect on hydrophones deployed closer to the sea floor. The comparative integrated analysis provided valuable information about future marine sound monitoring technologies and protocols while building on previous acoustics analyses at the FORCE site.

In its 2021-2023 EEMP proposal, FORCE has recommended conducting a test survey in the presence of an operational device using an internationally recognized standard methodology for monitoring sound (International Electrotechnical Commission 2019). This would permit the feasibility of the approach to be tested in the Minas Passage to ensure the method can be implemented as described. This work is pending an operational device being deployed at the FORCE tidal demonstration area. FORCE will work with JASCO to collect and analyze marine sound data once a device is deployed.

Seabirds

FORCE's seabird monitoring program is designed to test the EA prediction that project activities are not likely to cause adverse residual effects on marine birds within the FORCE test area (AECOM 2009). However, there has been limited opportunity to determine potential effects of an operational device on seabirds at the FORCE test site and to test the EA predictions.

Since 2011, FORCE and Envirosphere Consultants Ltd. (Windsor, NS) have collected observational data from the deck of the FORCE Visitor Centre, documenting seabird species presence, distribution, behaviour, and seasonality throughout the FORCE site (Envirosphere Consultants Ltd. 2017). Envirosphere Consultants Ltd. recently published the results of their monitoring from 2010-2012 and demonstrated that the species and seasonal cycles of seabirds in Minas Passage reflect patterns that are typical of the inner Bay of Fundy and the northeast Atlantic coast of North America. The report also highlights the importance of the Minas Passage as a migratory pathway for black scoter (*Melanitta americana*) and Red-throated loon (*Gavia stellata*).

In 2019, FORCE commissioned Envirosphere Consultants Ltd. and Dr. Phil Taylor (Acadia University) to synthesize the results of its observational seabird surveys (2011-2018) at the FORCE test site, and to evaluate advanced statistical techniques for analysing seabird count data in relation to environmental predictor variables. The seabird count data were examined using Generalized Additive Models (GAMs) to characterize seabird abundance and to better understand the potential impacts of tidal stream energy devices on seabirds at the FORCE test site. The results of the analyses revealed that overall model fit is suitable to characterize count data for some species, and that there are clear patterns of effects of time of year, wind speed and direction, tide height and time of day on the number of seabirds observed. However, the analyses also revealed that not all species reported at FORCE have been observed frequently enough to be modelled effectively using the GAM approach. This is due in part to the variability in count data that is particularly relevant for modelling abundance of migratory species that are only present at the FORCE site for brief periods during annual migrations. This is consistent with observational data collected over the course of these surveys that have demonstrated that the FORCE site has a lower abundance of seabirds in relation to other areas of the Bay of Fundy, and even other regions of Atlantic Canada. Given these results, the report recommends that future monitoring and analyses focus on locally resident species (i.e., great black-backed gull, herring gull, black guillemot, and common eider) so that the EA predictions can be tested most effectively. This work contributes to the development of appropriate analytical methods for assessing the impacts of tidal power development in the Minas Passage on relevant seabird populations and supports the continued responsible development of tidal energy at FORCE.

In 2022 FORCE began working with Strum Consulting to test radar-based seabird monitoring capabilities and to adapt existing data processing algorithms and statistical analysis tools for quantifying seabird use of the FORCE site. Strum has provided a technical report which highlights challenges and options to move forward with this approach. Challenges with the quality of the radar data limited the assessment and the full study could not be completed. This feasibility study will continue in 2023 with FORCE providing a new data set to Strum to work through some of the challenges in locating avian targets.

Developer Monitoring Activities

While FORCE completes site-level monitoring activities at the FORCE site, device specific monitoring is led by individual berth holders. Like the FORCE monitoring programs, the developer monitoring plans and reports undergo review by FORCE's EMAC and regulators.

In September 2018, it was confirmed that that CSTV turbine rotor was not spinning. Since that time, CSTV had been providing written confirmation to regulators monthly that the turbine is not operational by monitoring its status during the peak tidal flow of each month. However, because of the insolvency of OpenHydro Technology Ltd., all reporting activities by CSTV ceased as of March 1, 2019. Data collection from the turbine-mounted ADCPs to confirm the turbine is no longer spinning was managed and reported by FORCE to regulators monthly from March 2019 – May 2020 but was discontinued following an amendment to this requirement.

As additional developer, device-specific environmental effects monitoring programs are required and implemented for deployed tidal stream devices, berth holder updates will be included as appendices to future reports.

Other FORCE Research Activities

Risk Assessment Program

The Risk Assessment Program (RAP) for tidal stream energy is a collaborative effort between FORCE, academic partners, First Nations, and industry to advance our understanding of the environmental risks of tidal stream development in Minas Passage. The greatest potential risk of tidal stream energy device operations continues to be perceived by regulators and stakeholders as collisions between marine animals and turbine blades (Copping and Hemery 2020). However, these types of interactions are difficult to observe directly due to the environmental conditions under which they would occur (i.e., fast flowing, turbid waters) and using the suite of environmental monitoring instrumentation currently available (i.e., standard oceanographic and remote sensing instruments intended for use in more benign marine conditions) (Hasselman et al. 2020), but can be modeled using appropriate baseline data. The objective of the RAP program is to develop statistically robust encounter rate models for migratory and resident fishes with tidal stream energy devices in the Bay of Fundy using a combination of physical oceanographic data related to flow and turbulence in the Minas Passage and hydroacoustic tagging data for various fish species curated by the Ocean Tracking Network (OTN) at Dalhousie University.

Recent research has revealed how hydrodynamics (flow and turbulence-related features) in tidal stream environments can influence the distribution of marine animals, including fish (Lieber et al. 2018, 2019; McInturf et al. 2019). The Minas Passage is characterized by a series of turbulent hydrodynamics features (i.e., vortices, eddies, whirlpools, wakes, and shear currents) that could impact the spatiotemporal distribution of various fishes. The RAP will use a series of ADCP data collection efforts combined with a high-resolution radar network to create the first spatiotemporal flow atlas of the Minas Passage to understand these hydrodynamic features. Two Nortek Signature 500 autonomous ADCP's (Figure 2) were deployed in the tidal demonstration area on January 27th. One of the ADCP's was successfully recovered on May 4th however, the second unit could not be recovered due to an unforeseen issue with the acoustic release recovery mechanism. Further attempts to recover the unit were made this summer but unfortunately the unit could not be recovered. In lieu of the data from the missing ADCP, data from a previous study was used as a replacement. Concurrently, hydroacoustic data for various migratory and resident fish species in the Bay of Fundy that is curated by OTN will be compiled and analysed to

understand their spatiotemporal distributions. The hydrodynamic and hydroacoustic data will then be combined with information about device specific parameters (e.g., turbine blade length, swept area, turbine height off the seabed) to develop encounter rate models for various fish species. These models will then be refined and validated through a series of hydroacoustic tagging efforts, ultimately leading to the development of a user-friendly Graphical User Interface (GUI) similar to what is available for the offshore wind energy industry in the United Kingdom (McGregor et al. 2018). Ultimately, the RAP will contribute towards improving our understanding of the risks of tidal stream energy development for fishes of commercial, cultural, and conservation importance in the Bay of Fundy, and will assist in the development of future environmental effects monitoring programs.



Figure 2: Two Nortek Signature 500 autonomous ADCP's fitted in aluminum frames during deployment at the FORCE tidal demonstration area.

Since the program commenced in April 2020, OTN has acquired acoustic tag data from 22 contributors, covering nine species of fish in the Bay of Fundy (i.e., alewife (*Alosa pseudoharengus*), American shad (*A. sapidissima*), American eel (*Anguilla rostrata*), Atlantic salmon (*Salmo salar*), Atlantic sturgeon (*Acipenser oxyrinchus oxyrinchus*), Atlantic tomcod (*Microgadus tomcod*), spiny dogfish (*Squalus acanthias*), striped bass (*Morone saxatilis*), and white shark (*Carcharodon carcharias*)). FORCE has also established a high-resolution radar network in Minas Passage and has begun quantifying hydrodynamic features (turbulence, flow etc.) of Minas passage (Figure 3). The integration of physical habitat variables with acoustic tag data commenced in 2021, including the development of species distribution models for each species and species distribution maps. This work continues in 2022 with additional tagging currently underway to further validate model predictions. In collaboration with the Mi'kmaw Conservation Group (MCG) (Figure 4) fish tagging has been completed on 25 Atlantic salmon smolts, 24 American shad, 50 alewives, 4 Atlantic Sturgeon, and 15 spiny dogfish. The Atlantic Salmon smolts were tagged by DFO Science as their contribution to the project. Shad and alewife tags were purchased by FORCE berth holder Sustainable Marine as part of their contribution to RAP.

The acoustic receiver array (Figure 5) for detecting tagged fish was deployed in 2021 between early June and late August and again from September to early December. Due to the dynamic nature of the Minas Passage the equipment required extensive repairs which has delayed redeployment of the array. This time was used to reassess the positioning of acoustic receivers on the mooring/SUBs packages and develop a more streamlined design to alleviate extensive drag and reduce damage (Figure 6). The array was redeployed in May 2022 with the new design configuration and was recovered in early September. The data sets from the receivers have been downloaded and are being used to improve the models that the program has developed. The new design configuration of the acoustic receivers on the SUBs packages appears to be working well as minimal damage was noted following this deployment of the receiver line.

The first peer reviewed published research to come out of the RAP project is now publicly available ([Bangley et al. \(2022\)](#), Appendix IV) and is an important milestone for the project by demonstrating the potential utility of species distribution modelling of acoustic tag detections in predicting interactions with renewable energy devices. The report also showed the importance of physical oceanographic variables in influencing species distribution in a highly dynamic marine environment. The results of the study suggest that during fall (October-December) striped bass in Minas Passage are more likely to be associated with warmer water temperatures, very simple or very complex seafloor types and turbulent water conditions that are associated with ebbing tides. When mapped out, higher probabilities of striped bass presence occurred mostly in nearshore environments and were most widespread during the late ebb tide stage.



Figure 3: One of two high-resolution radars constructed near the FORCE site to be used for the Risk Assessment Program.



Figure 4: Acoustic tagging of spiny dogfish from the Minas Basin by RAP partner organization Mi'kmaq Conservation Group in 2022.

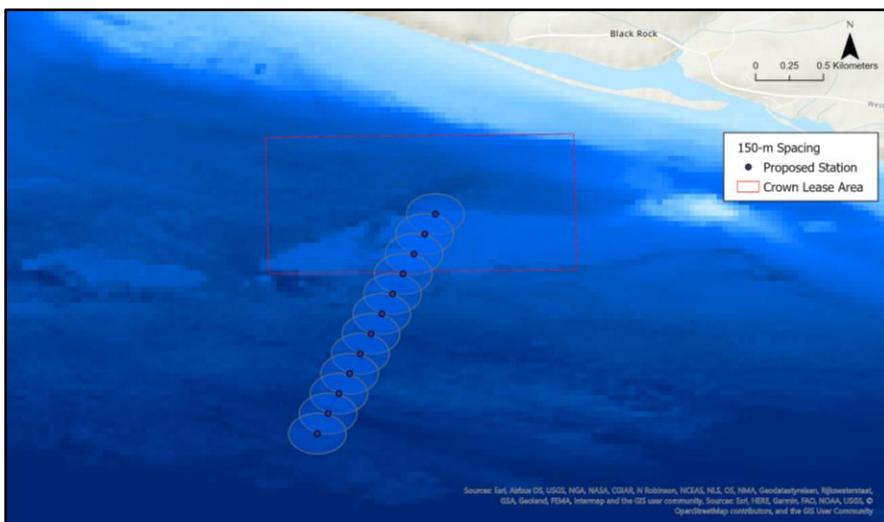


Figure 5: Acoustic receiver array deployment configuration in Minas Passage.



Figure 6: New positioning of acoustic receivers and release mechanism on a SUB package.

Fundy Advanced Sensor Technology (FAST) Activities

FORCE's Fundy Advanced Sensor Technology Program is designed to advance capabilities to monitor and characterize the FORCE site. Specifically, the FAST Program was designed to achieve the following objectives:

- 1) To advance capabilities of site characterization;
- 2) To develop and refine environmental monitoring standards and technologies; and
- 3) To enhance marine operating methodologies.

FAST combines both onshore and offshore monitoring assets. Onshore assets include a meteorological station, video cameras, an X-band radar system, and tide gauge. Offshore assets include modular subsea platforms for both autonomous and cabled data collection and a suite of instrumentation for a variety of research purposes. Real-time data collected through FAST assets will be broadcasted through the Canadian Integrated Ocean Observing System (CIOOS) later this year. Static ADCP data is currently available on the CIOOS website.²³

Platform Projects

The first and largest of the FAST platforms houses an instrument called the Vectron. Developed in partnership with Nortek Scientific (Halifax, NS), Memorial University (St. John's, NL), and Dalhousie University (Halifax, NS), the Vectron is the world's first stand-alone instrument to remotely measure, in high resolution, turbulence in the mid-water column. Measurements and analysis from the Vectron will help tidal energy companies to better design devices, plan marine operations, and characterize the tidal energy resource.

A smaller platform called FAST-3 was equipped with an upward looking echosounder and deployed during 2017-2018 to monitor fish densities at the FORCE site. FORCE and its partners, including Echoview Software completed data processing and analysis in 2019. This data was integrated with the mobile hydroacoustic surveys that FORCE conducts as part of its EEMP to evaluate the temporal and spatial representativeness of each method and to determine the

²³ This is available online at: https://catalogue.cioosatlantic.ca/dataset/ca-cioos_db15458d-df2c-4efb-b5a0-791e7561a0cb

degree to which results were corroborative (Figure 7). Although the spatial representative range of the stationary results could not be determined from the mobile data, it did reveal strong tidal and diel periods in fish density estimates at the site, with greater variation over shorter time frames than over the course of a year. These findings reinforce the importance of 24-hr data collection periods in ongoing monitoring efforts. The report reveals that collecting 24 hours of data allows the tidal and diel variability to be quantified and isolated from the longer-term trends in fish density and distribution that need to be monitored for testing the EA predictions. This project was funded by Natural Resources Canada (NRCan), the NSDEM, and Net Zero Atlantic (formally OERA).

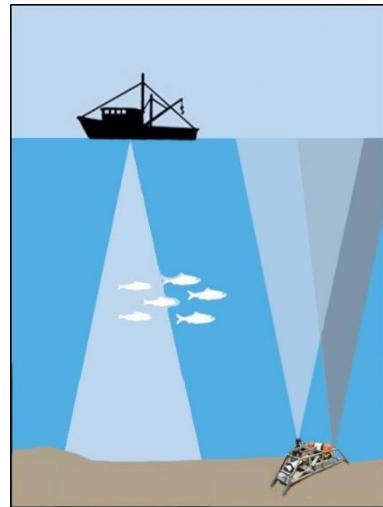


Figure 7: A representation of the data collection methods of the FORCE site-level fish EEMP and the FAST-3 platform.

Vitality Project

FORCE is actively participating in a new research and development program called the VITALITY Innovation Ecosystem Activity Project that is focused on integrating tidal stream data from the FORCE test site into CIOOS. CIOOS is a national online digital platform for sharing, discovering, and accessing ocean data in Canada, and data that is integrated into CIOOS is visible regionally and nationally. FORCE's component of the VITALITY project has three primary objectives:

1. Integration of FORCE's resource characterization and relevant environmental monitoring data (real time and static) into CIOOS to support better data accessibility and preservation,
2. Incorporation of industry and other stakeholder's data into CIOOS (i.e., industry use case), and
3. Installation and commissioning of a cabled subsea node at the FORCE site with applied R&D sensors whose real-time data will be integrated into CIOOS.

To that end, FORCE and its project partner Dalhousie University have recently developed a cabled subsea platform that includes an ADCP for measuring tidal current flow, waves and water temperature, a video camera for providing live stream video, and an array of hydrophones for testing the real-time detection of harbour porpoise. The platform underwent a deployment in the intertidal zone near the FORCE test site for initial testing this spring which was deemed a success. (Figure 8). Once the intertidal testing was completed, the platform was recovered from the intertidal zone and re-deployed in closer proximity to the FORCE site to test capabilities in the dynamic tidal conditions of the Minas Passage. This deployment took place on May 25th and the platform successfully streamed live data back to the FORCE visitor centre. On August 10th communication with the platform was interrupted and all efforts to re-establish a communication link failed.

To diagnose the cause of the communication failure, the platform was recovered on October 17th. Post recovery testing determined that the failure was caused by corrosion of one of the platform grounding electrical cables. The likelihood of corrosion increases in a high flow environment like Minas Passage, and we are currently discussing design adjustments to address this issue going forward. All other components continued to function as expected and had the electrical cabling not corroded, the system could potentially have remained deployed long term. A long-term platform monitoring system would be an asset to FORCE and our berth holders in the future.

During its deployment, the Vitality platform successfully acquired 75 days of real-time acoustic doppler current profiler (ADCP) data which gathers ocean current, wave, and temperature data that is vital for our environmental monitoring program and technology developers. The platform also collected data from a hydrophone array that was configured in conjunction with Dalhousie University to gather harbour porpoise localization data. This data helps contribute to our benchmark environmental monitoring program so that we can better understand the marine mammal population at the FORCE site. Lastly, the platform also had a standard definition underwater video camera and LED light with a live feed back to our visitor centre where guests and staff could view the offshore underwater environment in real time. The deployment of the camera was a successful test to see if we could get a feed of underwater life and future deployments will save the video files to contribute to our catalogue of local fish in the area.



Figure 8: The cabled subsea platform developed for the VITALITY project just prior to deployment at the FORCE test site.

Video of the VITALITY platform being deployed at the FORCE Test Site:
<https://vimeo.com/718028837>

Fish Tracking

To enhance fish monitoring and to expand its data collection capacity, FORCE partnered with the Ocean Tracking Network (OTN)²⁴ and attached one VEMCO²⁵ fish tag receiver (a VR2W receiver) to each C-POD mooring/SUBS (Streamlined Underwater Buoyancy System) package (see above). These receivers are used to supplement OTN's ongoing data collection program within

²⁴ Ocean Tracking Network's website: www.oceantrackingnetwork.org.

²⁵ VEMCO is "the world leader in the design and manufacture of acoustic telemetry equipment used by researchers worldwide to study behaviour and migration patterns of a wide variety of aquatic animals." Learn more: www.vemco.com.

the Minas Passage and are referred to as ‘Buoys of Opportunity.’ Upon retrieval of the C-PODs and receivers, instruments are shared with OTN where data is offloaded prior to redeployment. This effort will support increased knowledge of fish movement within the Minas Passage, which has applicability beyond tidal energy demonstration, as well as complement FORCE’s hydroacoustic data collection efforts that do not allow for species identification. No C-POD mooring/SUBS have been deployed since 2020, however ongoing data collection for fish monitoring is occurring through the RAP acoustic receiver line.

OTN data managers are in the process of acquiring information, including species identification, and sharing this with FORCE. Initial results show that the OTN receivers deployed by FORCE have detected tags from the following projects:

- Maritimes Region Atlantic salmon marine survival and migration (Hardie, D.C., 2017);
- Quebec MDDEFP Atlantic Sturgeon Tagging (Verreault, G., Dussureault, J., 2013);
- Gulf of Maine Sturgeon (Zydlowski, G., Wippelhauser, G. Sulikowski, J., Kieffer, M., Kinnison, M., 2006);
- OTN Canada Atlantic Sturgeon Tracking (Dadswell, M., Litvak, M., Stokesbury, M., Bradford, R., Karsten, R., Redden, A., Sheng, J., Smith, P.C., 2010);
- Darren Porter Bay of Fundy Weir Fishing (Porter, D., Whoriskey, F., 2017);
- Movement patterns of American lobsters in the Minas Basin, Minas Passage, and Bay of Fundy Canada (2017);
- Shubenacadie River Monitoring Project: Tomcod (Marshall, J., Fleming, C., Hunt, A., and Beland, J., 2017);
- MA Marine Fisheries Shark Research Program (Skomal, G.B., Chisholm, J., 2009);
- UNB Atlantic Sturgeon and Striped Bass tracking (Curry, A., Linnansaari, T., Gautreau, M., 2010);
- Inner Bay of Fundy Striped Bass (Bradford, R., LeBlanc, P., 2012);
- Minas Basin Salmon Kelt (McLean, M., Hardie, D., Reader, J., Stokesbury, M.J.W., 2019);
- New York Juvenile White Shark Study (Tobey Curtis); and
- Massachusetts White Shark Research Program (Greg Skomal)

Further information about these Buoys of Opportunity, and the projects listed above, can be found on OTN’s website: <https://members.oceantrack.org/project?ccode=BOOFORCE>

Starting in 2018, FORCE has worked in collaboration with Dr. Mike Stokesbury at Acadia University to install additional VEMCO receivers of a new design on FORCE’s C-POD moorings/SUBS packages. These new receivers are expected to be even more effective in picking up acoustic detections in high flow environments, where tag signals can be obscured by noise. This partnership will contribute additional information regarding movement patterns of Atlantic salmon, sturgeon, striped bass, and alewife in Minas Passage and Basin. This work is sponsored by the OERA, NRCAN, NSDEM, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Canadian Foundation for Innovation (CFI).²⁶

²⁶ Information about this project, and others funded through this program, is available online at:
<https://netzeroatlantic.ca/sites/default/files/2020-04/2020-04-09%20NRCAN%20Public%20Report%20Final%20-%20Resize.pdf>

Discussion

The 2021-2023 EEMP has provided a strategic opportunity for FORCE and its partners to learn from previous experiences, incorporate regulatory advice, and to re-evaluate approaches to research and monitoring in the high flows of the Minas Passage. The EEMP is designed to prepare for effects testing with the deployment of operational devices and adheres to the principles of adaptive management by evaluating existing datasets to ensure appropriate monitoring approaches are being implemented. Moreover, the plan adopts internationally accepted standards for monitoring where possible, including feasibility assessments for new monitoring approaches that are planned to be implemented.

FORCE invested in the development of its internal scientific capacity by hiring a PhD level hydroacoustician (Dr. Louise McGarry). This assisted FORCE with tackling data that required processing, analyses, and integration with other data sets. Dr. McGarry also assisted with the development of study designs to help advance our understanding of how fish utilize the Minas Passage. FORCE assisted in knowledge transfer to ASL Environmental (Victoria, BC) about hydroacoustics research in high flow environments so that ASL can be used as a resource to help support the sector in the future.

While the 2020 COVID19 outbreak initially impacted our ability to gather data at our site and conduct marine operations – all of which require multiple people working in close proximity – our operations and monitoring data collection activities have resumed. As such, FORCE and its partners have resumed conducting monitoring, engaging in meaningful assessments of monitoring technology capabilities, and providing data analyses and interpretation that advance our ability to effectively monitor the effects of tidal stream energy devices in high flow environments, and specifically at the FORCE test site. Reports from FORCE's partners and updates are routinely subjected to review by FORCE's EMAC and regulators, along with continued results from FORCE's ongoing monitoring efforts.

FORCE continues to implement lessons learned from the experiences of local and international partners, build local capacity, and enhance skills development, test new sensor capabilities, and integrate results from various instruments. Cumulatively, these efforts provide an opportunity for adaptive management and the advancement and refinement of scientific approaches, tools, and techniques required for effectively monitoring the device and site-level areas of tidal stream energy devices in dynamic, high-flow marine environments.

Ongoing monitoring efforts will continue to build on the present body of knowledge of marine life-device interactions. While it is still early to draw conclusions, initial findings internationally and at the FORCE test site have documented some disturbance of marine mammals primarily during marine operations associated with device installation/removal activities, but otherwise have not observed significant effects.

FORCE will continue to conduct environmental research and monitoring to increase our understanding of the natural conditions within the Minas Passage and, when the next device(s) are deployed and operating, test the EA prediction that tidal energy is unlikely to cause significant harm to marine life. In the longer-term, monitoring will need to be conducted over the full seasonal cycle and in association with multiple different device technologies to understand if tidal energy can be a safe and responsibly produced energy source. FORCE will continue to report on progress and release results and lessons learned in keeping with its mandate to inform decisions regarding future tidal energy projects.

References

- AECOM. 2009. Environmental Assessment Registration Document - Fundy Tidal Energy Demonstration Project Volume I: Environmental Assessment. Available from Available at www.fundyforce.ca.
- Bangley, C.W., Hasselman, D.J., Flemming, J.M., Whoriskey, F.G., Culina, J., Enders, L., and Bradford, R.G. 2022. Modeling the probability of overlap between marine fish distributions and marine renewable energy infrastructure using acoustic telemetry data. *Front. Mar. Sci.* **9**: 851757. doi:10.3389/fmars.2022.851757.
- Copping, A.E. 2018. The State of knowledge for environmental effects - driving consenting/permitting for the marine renewable energy industry. Available from <https://tethys.pnnl.gov/sites/default/files/publications/The%20State%20of%20Knowledge%20Driving%20Consenting%20Permiting%20for%20the%20MRE.pdf>.
- Copping, A.E., and Hemery, L.G. 2020. OES-Environmental 2020 State of the Science Report: Environmental Effects of Marine Renewable Energy Development Around the World. Report for Ocean Energy Systems (OES). *In* State Sci. Rep. doi:10.2172/1632878.
- Copping, A.E., Hemery, L.G., Viehman, H., Seitz, A.C., Staines, G.J., and Hasselman, D.J. 2021. Are fish in danger? A review of environmental effects of marine renewable energy on fishes. *Biol. Conserv.* **262**: 109297. Elsevier Ltd. doi:10.1016/j.biocon.2021.109297.
- Daroux, A., and Zydlewski, G. 2017. Marine fish monitoring program tidal energy demonstration site – Minas Passage. : 34. Orono, ME.
- Envirosphere Consultants Ltd. 2017. Marine seabirds monitoring program – tidal energy demonstration site – Minas Passage, 2016-2017.
- Fraser, S., Williamson, B.J., Nikora, V., and Scott, B.E. 2018. Fish distributions in a tidal channel indicate the behavioural impact of a marine renewable energy installation. *Energy Reports* **4**: 65–69. Elsevier Ltd. doi:10.1016/j.egyr.2018.01.008.
- Gaskin, D.E. 1992. Status of the harbour porpoise, *Phocoena phocoena*, in Canada. *Can. F. Nat.* **106**(1): 36–54.
- Gattuso, J.P., Magnan, A.K., Bopp, L., Cheung, W.W.L., Duarte, C.M., Hinkel, J., Mcleod, E., Micheli, F., Oschlies, A., Williamson, P., Billé, R., Chalastani, V.I., Gates, R.D., Irisson, J.O., Middelburg, J.J., Pörtner, H.O., and Rau, G.H. 2018. Ocean solutions to address climate change and its effects on marine ecosystems. *Front. Mar. Sci.* **5**(OCT). doi:10.3389/fmars.2018.00337.
- Gillespie, D., Hastie, G., Palmer, L., Macaulay, J., and Sparling, C. 2021. Harbour porpoises exhibit localized evasion of a tidal turbine. *Aquat. Conserv. Freshw. Ecosyst.*: 1–10. doi:10.1002/aqc.3660.
- Hasselman, D.J., Barclay, D.R., Cavagnaro, R., Chandler, C., Cotter, E., Gillespie, D.M., Hastie, G.D., Horne, J.K., Joslin, J., Long, C., McGarry, L.P., Mueller, R.P., Sparling, C.E., Williamson, B.J., and Staines, G.J. 2020. Environmental monitoring technologies and techniques for detecting interactions of marine animals with turbines. *In* Report for Ocean Energy Systems (OES).
- Hasselman, D.J., Li, H., Cotter, E., and Joslin, J. 2022. Editorial: Novel technologies for assessing the environmental and ecological impacts of marine renewable energy systems. *Front. Mar. Sci.* **9**: 990327. doi:10.3389/fmars.2022.990327.
- International Electrotechnical Commission. 2019. Marine Energy - Wave, tidal and other water current converters - Part 40: Acoustic characterization of marine energy converter.
- Joy, R., Robertson, F., and Tollit, D. 2017. FORCE Marine Mammal Environmental Effects Monitoring Program - 1st Year (2017) Monitoring Report.
- Joy, R., Wood, J., and Tollit, D. 2018a. FORCE echolocating marine mammal environmental effects monitoring program - 2nd year (2018) monitoring report.

- Joy, R., Wood, J.D., Sparling, C.E., Tollit, D.J., Copping, A.E., and McConnell, B.J. 2018b. Empirical measures of harbor seal behavior and avoidance of an operational tidal turbine. *Mar. Pollut. Bull.* **136**: 92–106. Elsevier. doi:10.1016/j.marpolbul.2018.08.052.
- Lewis, A., Estefen, S., Huckerby, J., Musial, W., Pontes, T., and Torres-Martinez, J. 2011. Ocean Energy. In *Renewable Energy Sources and Climate Change Mitigation: Special Report of the Intergovernmental Panel on Climate Change*. Edited by O. Edenhofer, R. Pichs-Madruga, Y. Sokona, K. Seyboth, P. Matschoss, and S. Kadner. Cambridge University Press, Cambridge, Massachusetts. pp. 497–534.
- Lieber, L., Nimmo-Smith, W.A.M., Waggitt, J.J., and Kregting, L. 2018. Fine-scale hydrodynamic metrics underlying predator occupancy patterns in tidal stream environments. *Ecol. Indic.* **94**: 397–408. Elsevier. doi:10.1016/j.ecolind.2018.06.071.
- Lieber, L., Nimmo-Smith, W.A.M., Waggitt, J.J., and Kregting, L. 2019. Localised anthropogenic wake generates a predictable foraging hotspot for top predators. *Commun. Biol.* **2**(1): 1–8. Springer US. doi:10.1038/s42003-019-0364-z.
- Lossent, J., Lejart, M., Folegot, T., Clorennec, D., Di Iorio, L., and Gervaise, C. 2018. Underwater operational noise level emitted by a tidal current turbine and its potential impact on marine fauna. *Mar. Pollut. Bull.* **131**(May 2017): 323–334. Elsevier. doi:10.1016/j.marpolbul.2018.03.024.
- Lowe, S.C., McGarry, L.P., Douglas, J., Newport, J., Oore, S., Whidden, C., and Hasselman, D.J. 2022. Echofilter: A deep learning segmentation model improves the automation, standardization, and timeliness for post-processing echosounder data in tidal energy streams. *Front. Mar. Sci.* **9**: 867857. doi:10.3389/fmars.2022.867857.
- Malinka, C.E., Gillespie, D.M., Macaulay, J.D.J., Joy, R., and Sparling, C.E. 2018. First in situ passive acoustic monitoring for marine mammals during operation of a tidal turbine in Ramsey Sound, Wales. *Mar. Ecol. Prog. Ser.* **590**: 247–266. doi:10.3354/meps12467.
- Marine Renewables Canada. 2018. State of the Sector Report: Marine Renewable Energy in Canada.
- Martin, B., Whitt, C., and Horwich, L. 2018. Acoustic data analysis of the OpenHydro opencentre turbine at FORCE: final report.
- McGregor, R.M., King, S., Donovan, C.R., Caneco, B., and Webb, A. 2018. A Stochastic Collision Risk Model for Seabirds in Flight. Available from <https://www2.gov.scot/Resource/0053/00536606.pdf>.
- McInturf, A.G., Steel, A.E., Buckhorn, M., Sandstrom, P., Slager, C.J., Fangue, N.A., Klimley, A.P., and Caillaud, D. 2019. Use of a hydrodynamic model to examine behavioral response of broadnose sevengill sharks (*Notorynchus cepedianus*) to estuarine tidal flow. *Environ. Biol. Fishes* **102**(9): 1149–1159. Environmental Biology of Fishes. doi:10.1007/s10641-019-00894-3.
- Melvin, G.D., and Cochrane, N.A. 2014. Investigation of the vertical distribution, movement and abundance of fish in the vicinity of proposed tidal power energy conversion devices. Final Report for the Offshore Energy Research Association. Research Project 300-170-09-12.
- NEXUS Coastal Resource Management Ltd. 2017. Lobster Catchability Study Report.
- Palmer, K.J., Wood, J., and Tollit, D.J. 2021. FORCE Marine Mammal EEMP - Yer 4 Final Report.
- Pine, M.K., Schmitt, P., Culloch, R.M., Lieber, L., and Kregting, L.T. 2019. Providing ecological context to anthropogenic subsea noise: Assessing listening space reductions of marine mammals from tidal energy devices. *Renew. Sustain. Energy Rev.* **103**(July 2018): 49–57. Elsevier Ltd. doi:10.1016/j.rser.2018.12.024.
- Robertson, F., Wood, J., Joslin, J., Joy, R., and Polagye, B. 2018. Marine Mammal Behavioral Response to Tidal Turbine Sound. (206). doi:10.2172/1458457.
- Schramm, M.P., Bevelhimer, M., and Scherelis, C. 2017. Effects of hydrokinetic turbine sound on the behavior of four species of fish within an experimental mesocosm. *Fish. Res.* **190**:

- 1–14. Elsevier B.V. doi:10.1016/j.fishres.2017.01.012.
- SLR Consulting. 2015. Proposed Environmental Effects Monitoring Programs 2015-2020 for Fundy Ocean Research Center for Energy (FORCE).
- Viehman, H., Hasselman, D., Boucher, T., Douglas, J., and Bennett, L. 2019. Integrating hydroacoustic approaches to predict fish interactions with in-stream tidal turbines. Available from https://netzeroatlantic.ca/sites/default/files/2020-04/FISH_FINDER_Final_Report_31-03-2020_0.pdf.
- Viehman, H.A., Hasselman, D.J., Douglas, J., and Boucher, T. 2022. The ups and downs of using active acoustic technologies to study fish at tidal energy sites. *Front. Mar. Sci.* **9**: 851400. doi:10.3389/fmars.2022.851400.

Appendix I

Acronyms

AAM	Active Acoustic Monitoring
ADCP	Acoustic Doppler Current Profiler
AMAR	Autonomous Multichannel Acoustic Recorder
BACI	Before/After, Control/Impact
BC	British Columbia
BoFEP	Bay of Fundy Ecosystem Partnership
CFI	Canadian Foundation for Innovation
CIOOS	Canadian Integrated Ocean Observing System
CLA	Crown Lease Area
cm	Centimetre(s)
CPUE	Catch Per Unit Effort
CSTV	Cape Sharp Tidal Venture
DFO	Department of Fisheries and Oceans (Canada)
DEM	Department of Energy and Mines (Nova Scotia)
EA	Environmental Assessment
EEMP	Environmental Effects Monitoring Program
EMAC	Environmental Monitoring Advisory Committee
EMP	Environmental Management Plan
FAD	Fish Aggregation Device
FAST	Fundy Advanced Sensor Technology
FAST-EMS	Fundy Advanced Sensor Technology – Environmental Monitoring System
FERN	Fundy Energy Research Network
FORCE	Fundy Ocean Research Center for Energy
GPS	Global Positioning System
hr	Hour(s)
IEA	International Energy Agency
kg	Kilogram(s)
km	Kilometre(s)
kW	Kilowatt(s)
m	Metre(s)
MET	Meteorological
MRE	Marine Renewable Energy
MREA	Marine Renewable-electricity Area
NL	Newfoundland and Labrador
NRCan	Natural Resources Canada
NS	Nova Scotia
NSDEM	Nova Scotia Department of Energy and Mines
NSE	Nova Scotia Department of Environment
NSERC	Natural Sciences and Engineering Research Council
NSPI	Nova Scotia Power Inc.
OERA	Offshore Energy Research Association of Nova Scotia
OES	Ocean Energy Systems
ONC	Ocean Networks Canada
ORJIP	Offshore Renewables Joint Industry Programme
OSC	Ocean Supercluster
OTN	Ocean Tracking Network
PAM	Passive Acoustic Monitoring
Q1/2/3	Quarter (1, 2, 3), based on a quarterly reporting schedule

R&D	Research and Development
TC114	Technical Committee 114
SUBS	Streamlined Underwater Buoyancy System
SME	Sustainable Marine Energy (Canada)
UAV	Unmanned Aerial Vehicle
UK	United Kingdom
VEC(s)	Valuable Ecosystem Component(s)

Appendix II



The Ups and Downs of Using Active Acoustic Technologies to Study Fish at Tidal Energy Sites

Haley A. Viehman^{1*}, Daniel J. Hasselman², Jessica Douglas³ and Tyler Boucher³

¹ Echoview Software Pty Ltd, Hobart, TAS, Australia, ² Fundy Ocean Research Center for Energy, Halifax, NS, Canada,

³ Independent Researcher, Halifax, NS, Canada

OPEN ACCESS

Edited by:

Wei-Bo Chen,
National Science and Technology
Center for Disaster Reduction
(NCDR), Taiwan

Reviewed by:

Michael Dadswell,
Acadia University, Canada
Philippe Blondel,
University of Bath, United Kingdom

*Correspondence:

Haley A. Viehman
haley.viehman@echoview.com

Specialty section:

This article was submitted to
Ocean Solutions,
a section of the journal
Frontiers in Marine Science

Received: 09 January 2022

Accepted: 07 March 2022

Published: 31 March 2022

Citation:

Viehman HA, Hasselman DJ,
Douglas J and Boucher T (2022) The
Ups and Downs of Using Active
Acoustic Technologies to Study
Fish at Tidal Energy Sites.
Front. Mar. Sci. 9:851400.
doi: 10.3389/fmars.2022.851400

Active acoustic instruments (echosounders) are well-suited for collecting high-resolution information on fish abundance and distribution in the areas targeted for tidal energy development, which is necessary for understanding the potential risks tidal energy devices pose to fish. However, a large proportion of echosounder data must often be omitted due to high levels of backscatter from air entrained into the water column. To effectively use these instruments at tidal energy sites, we need a better understanding of this data loss and how it may affect estimates of fish abundance and vertical distribution. We examined entrained air contamination in echosounder data from the Fundy Ocean Research Center for Energy (FORCE) tidal energy test site in Minas Passage, Nova Scotia, where current speeds can exceed $5 \text{ m}\cdot\text{s}^{-1}$. Entrained air depth was highly variable and increased with current speed, and contamination was lowest during neap tides. The lower 70% of the water column and current speeds $<3 \text{ m}\cdot\text{s}^{-1}$ were generally well-represented in the dataset. However, under-sampling of the upper water column and faster speeds strongly affected simulated fish abundance estimates, with error highly dependent on the underlying vertical distribution of fish. Complementary sensing technologies, such as acoustic telemetry and optical instruments, could be used concurrently with echosounders to fill gaps in active acoustic datasets and to maximize what can be learned about fish abundance and distribution at tidal energy sites.

Keywords: active acoustics, hydroacoustics, fish, entrained air, data quality, marine renewable energy, tidal energy, MHK

1 INTRODUCTION

The tidal energy sector is a nascent industry, and the potential environmental effects of marine hydrokinetic (MHK) devices on fish continues to be an area of concern for regulators and stakeholders of the marine environment (Copping et al., 2021). Predicting fish interactions with MHK devices, and therefore potential device effects, requires information on fish presence, abundance, and distribution at a resolution and scale that is rarely required elsewhere. Spatial resolution must be on the order of meters for data to be related to an individual MHK device, and collected throughout the water column and/or across tidal channels that can be kilometers wide. Similarly, fine temporal resolution (seconds to minutes) may be required to capture shifts in fish distribution that affect MHK device encounter

rates but years of observations may be needed to characterize seasonal patterns and longer-term population shifts. Active acoustic instruments are excellent tools for collecting this high-resolution information across large spaces and periods of time. This technology includes single beam, split beam, and multibeam echosounders utilizing single or multiple frequencies in narrow- or broad-band modes (Demer et al., 2015). Active acoustics is a vital component of fisheries stock assessments worldwide, given these instruments' unequaled capacity to rapidly and non-invasively sample large volumes of water (Horne, 2000; Simmonds and MacLennan, 2005). Echosounders have been employed in studies of fish at tidal energy sites around the world, as well (e.g. Viehman et al., 2015; Fraser et al., 2017; Viehman et al., 2018; Gonzalez et al., 2019; Williamson et al., 2019; Scherelis et al., 2020; Whitton et al., 2020).

Tidal channels are characterized by fast currents and complex hydrodynamics that pose unique challenges to active acoustics technology, which can hamper the translation of raw data to information that can be used by scientists, developers, and regulators of the tidal energy industry. The primary challenge is the high prevalence of air bubbles entrained into the water column, which scatter the sound transmitted by echosounders. Air entrainment is a common occurrence in the open ocean, with the primary source of entrainment being breaking waves (Woolf, 2001; Baschek et al., 2006). Air plumes in the open ocean commonly extend to depths of 10–15 m, but the extreme hydrodynamic conditions in areas with strong tidal currents can draw bubbles to depths well over 100 m (Baschek et al., 2006). Though bubbles entrained in the water column tend to be very small (e.g. < 1 mm diameter; Woolf, 2001; Baschek et al., 2006), they are strong scatterers of sound. The sound scattered by clouds of bubbles observed at tidal energy sites is similar to, or stronger than, that scattered by fish (for example, in the 120 kHz data assessed here, volume backscatter of the entrained air layer averaged -46 dB re 1 m²m⁻³), and the two scatterer types cannot be separated in active acoustics data if they inhabit the same volume of water. Measurements containing backscatter from entrained air must therefore be removed from acoustic datasets prior to analyzing backscatter from fish.

Studies at tidal energy sites have utilized different methods to remove backscatter from entrained air. The majority of methods exploit the distinct temporal and/or morphological characteristics of the bubble plumes to differentiate them from fish backscatter, including occurrence and duration in time and surface connectivity (Fraser et al., 2017; Scherelis et al., 2020). Features with the designated characteristics are then removed from the dataset, either manually or with some mix of automated and manual steps. Removal has included omitting just the contaminated data points (Fraser et al., 2017; Whitton et al., 2020), or a fixed depth range plus the entire water column when air extends further (Viehman et al., 2018). Other studies have kept only the lowermost portion of the water column as the depths of primary interest, ignoring the upper layers (Viehman et al., 2015; Gonzalez et al., 2019). Regardless of the method, the result is omitting a large amount of water that could contain fish but is unable to be effectively sampled by active acoustics instruments.

Omitting the entrained air layer is likely to affect acoustically derived estimates of fish abundance and vertical distribution, and

therefore our ability to estimate encounters with MHK devices. Moreover, it is possible that different fish species' or life stages' contributions to acoustic measurements will be unequally affected by removing different portions of the water column, given depth preferences that are often species- or life-stage-specific. For example, in the northwest Atlantic, Atlantic salmon post-smolts and adults (*Salmo salar*) tend to be found within the upper 10 m of the water column (Dutil and Couturier, 1988; Sheehan et al., 2012). Other species utilize the entire water column more generally (e.g. Atlantic herring, *Clupea harengus*, Hussey et al., 2012; Viehman et al., 2018; Atlantic mackerel, *Scomber scombrus*, Castonguay and Gilbert 1995), while others are typically associated with the bottom (e.g. Atlantic cod, *Gadus morhua*, Hobson et al., 2007). American eel (*Anguila rostrata*) have exhibited distinct vertical migrations to take advantage of favorable tidal currents, a behavior known as selective tidal stream transport (STST; Parker and McCleave, 1997). At present, it is unclear whether depth preferences observed in lower-energy environments will persist within highly energetic tidal channels, and there is some evidence that they may differ (Stokesbury et al., 2016; Lilly et al., 2021).

Though data contamination by entrained air is an issue at all tidal energy sites, we have yet to examine the resulting data loss in detail (e.g. its magnitude and spatiotemporal distribution), or how this loss could affect our acoustically derived estimates of fish abundance and vertical distribution. This information would be particularly helpful in the planning stages of a study or environmental monitoring plan, when steps can be taken to address any expected limitations of the active acoustic dataset. These steps may include, for example, the simultaneous use of complementary technologies and sampling techniques.

In this paper, we examined the entrained air layer in active acoustic data collected at the FORCE tidal energy test site. We developed a method for identifying and removing the data points contaminated by entrained air, quantified entrained air depth and resulting data loss, and demonstrated the effects of this data loss on estimates of fish abundance and vertical distribution obtained from simulated vertical distributions of fish. The active acoustic data assessed in this paper are from a fixed-location split beam, narrowband, scientific-grade echosounder, which is the type most used for assessing the abundance and vertical distribution of fishes over long periods of time or space. Our goal was to provide researchers, developers, and regulators of the tidal energy industry with the information they need to utilize active acoustics technology to its fullest potential, and to mitigate the limitations imposed on it by this exceptionally challenging environment.

2 MATERIALS AND METHODS

2.1 Data Collection

Data were collected at the Fundy Ocean Research Center for Energy (FORCE) tidal energy test site, in the Bay of Fundy, Nova Scotia, Canada (Figure 1). Instruments were installed on the Fundy Advanced Sensor Technology subsea platform, FAST-3 (Figure 2). This stationary platform was deployed on the seafloor

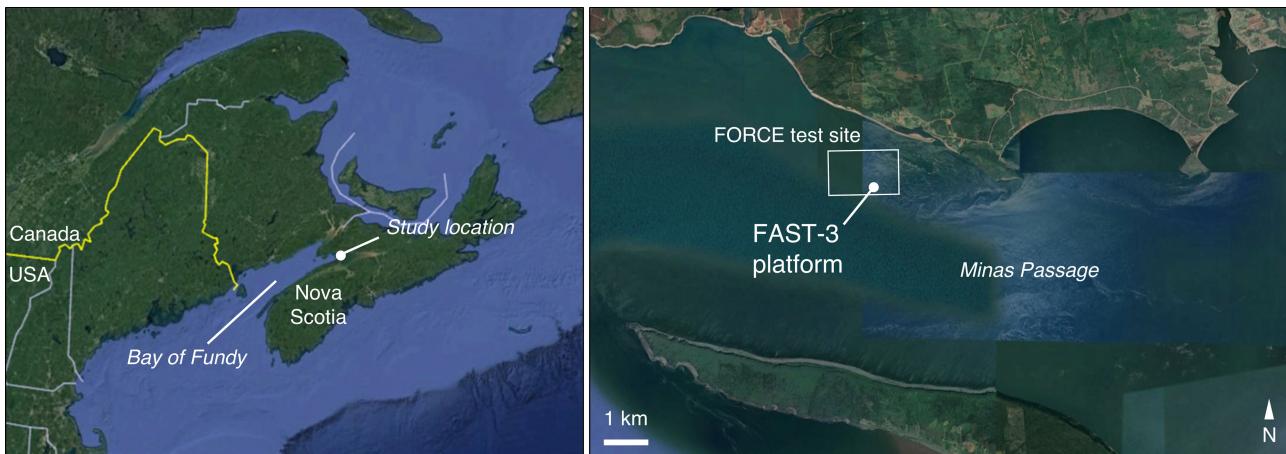


FIGURE 1 | Study location in the Minas Passage of the Bay of Fundy, Canada. The location of Minas Passage is indicated by the filled circle in the left-hand panel, and the study site is shown on the right.

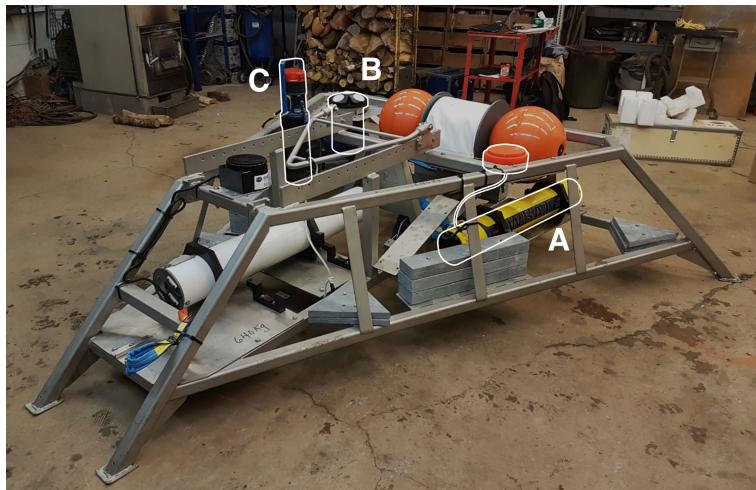


FIGURE 2 | FAST-3 platform deployed at the FORCE Tidal energy test site from 30 Mar to 23 May 2018. Equipment included (A) Simrad WBAT EK80 echosounder, (B) Nortek Signature 500 ADCP, (C) Aanderaa SeaGuard RCM.

at $45^{\circ}21'47.34''$ N, $64^{\circ}25'38.88''$ W, and was in place for 53 days from 30 March to 23 May 2018. At this location, water column depth averaged 33 m at low tide and 43 m at high tide.

Active acoustic data were collected by a Simrad EK80 WBAT echosounder with a 120 kHz split beam transducer (7° half-power beam angle), located 0.7 m above the seafloor and facing upward. Data were collected in 5-min recording periods every half hour, with a ping rate of 1 Hz, pulse duration of 0.128 ms, transmit power of 125 W, and maximum recording range of 60 m. Collection settings were chosen based on pilot data collected near this site in February 2017.

Measurements of current velocity throughout the water column were collected by a Nortek Signature 500 acoustic

doppler current profiler (ADCP). The ADCP's face was located at 0.7 m above the seafloor. Data were collected in 5-min bursts every 15 min, alternating with echosounder measurements to avoid acoustic interference between the two instruments. The sample rate during each burst was 2 Hz, the blanking distance was 1 m, and the cell size was 1 m.

Water temperature and salinity at the platform were measured by an Aanderaa SeaGuard RCM every half hour.

2.2 Data Processing

2.2.1 Active Acoustic Data

Active acoustic data processing was carried out using Echoview[®] software (12.1, Myriax, Hobart, Australia). We developed a data

processing routine in Echoview that detected the surface and entrained air layer, minimizing the need for manual correction as much as possible. The template developed for this process is provided in supplementary materials with a detailed explanation of all steps.

Briefly, the surface was detected with a line, and the boundary of the surface dead zone was delineated below this (0.16 m below on average; Ona and Mitson, 1996). A line was also defined at 2x the acoustic nearfield distance from the transducer face (Simmonds and MacLennan, 2005), and acted as the lower analysis limit in all following steps. Entrained air was defined morphometrically as clusters of backscatter which extended downward from the surface, similar to Fraser et al. (2017). Detection of these clusters required a series of separate processing steps, including smoothing the raw volume backscatter (S_V) data, applying a minimum data threshold, and using Echoview's schools detection algorithm to detect contiguous clusters of backscatter that surpassed this threshold. Clusters which were connected to the surface were isolated and expanded in depth and time, and a line was drawn below the resulting backscatter to establish the lower extent of the entrained air layer. The maximum depth of this layer was limited by the acoustic nearfield, 2.4 m above the seafloor.

All processing steps and settings were chosen by iteratively reviewing the performance of the processing routine on a subset of data files that represented a wide range of entrained air contamination, until the level of necessary manual corrections to the surface and entrained air lines was deemed acceptably low. All data files were then batch-processed in Echoview using the finalized routine. The resulting Echoview files were reviewed manually to make any necessary corrections to the surface and entrained air lines.

Once all necessary corrections were made, the surface and entrained air line depths were exported, and we calculated the average water column depth and entrained air depth for each 5-min data recording period. For each recording period, we also calculated the number of samples (individual datapoints) omitted due to the entrained air layer. We converted this number to a percent of analyzable samples, which was more comparable over time as water level changed. We defined analyzable samples as all samples between the nearfield and surface dead zone because samples outside of these boundaries would always be excluded from acoustic analysis.

Echosounder data were calibrated using calibration sphere measurements obtained at a calm location off-site, before and after the deployment. As environmental conditions changed significantly over the course of the deployment (temperature and salinity shifts caused the speed of sound to increase from $1452 \text{ m}\cdot\text{s}^{-1}$ to $1477 \text{ m}\cdot\text{s}^{-1}$), acoustic data were split into sections to which different calibration parameters were applied. Details of data calibration are supplied in supplementary materials.

2.2.2 ADCP Data

ADCP measurements were first corrected for platform tilt and compass declination using Ocean Contour (version 2.1.5, Ocean Illumination Ltd., Canada). We obtained average horizontal speed and direction for each 1-m cell of every ADCP burst. The first

measurement cell was centered 2 m from the transducer face. Measurements from the uppermost 10% of the water column could not be used due to interference from side lobes, so we removed these upper cells prior to calculating water column average speed and direction. We then interpolated these speed and direction values in time to obtain water column averages at the midpoint of each echosounder recording period. All future references to current speed or direction measurements refer to these interpolated water column averages.

Slack tide was defined as current speed $< 0.5 \text{ m}\cdot\text{s}^{-1}$, which captured the period of time when current direction was shifting between ebb and flood. In this dataset, slack tide defined in this way (by current speed and direction) occurred approximately 15–30 min after the time of lowest or highest water. Spring and neap tides were identified in the current velocity time series as maxima and minima in peak flow speed.

2.2.3 SeaGuard RCM Data

Conductivity and temperature readings from the SeaGuard RCM were used in the calculation of sound speed, for calibrating echosounder data (see supplementary material).

2.3 Data Analysis

There was no way to predict how many fish were omitted from the acoustic dataset by removing the entrained air layer. We therefore demonstrated how entrained air contamination affects estimates of fish abundance and distribution by constructing five hypothetical fish distribution scenarios that we then subjected to different levels of contamination and data removal. Analysis was carried out in R software version 4.1.2 (R Core Team, 2021).

The five vertical distribution scenarios each spanned one tidal cycle, which was split into 24 equally spaced time segments (tide bins; approximately 30 min each). All recording periods from the acoustic dataset were partitioned into these tide bins, and for each tide bin we calculated mean water column depth and the 5th, 50th (median), and 95th percentiles of entrained air depth. The mean water column depth from each tide bin defined the hypothetical water column in each fish distribution scenario. The water column was then split into 1 m depth bins to be populated with some number of fish. For simplicity, total fish abundance was held constant over time (1000 fish per tide bin, 24000 fish total). The fish distribution scenarios we generated were:

1. Fish utilizing the entire water column: for each tide bin, 1000 fish were distributed randomly into all water column bins, from the seafloor to the surface.
2. Surface-oriented fish: for each tide bin, 1000 fish were distributed into the upper 10 bins of the water column. To simulate a gradual increase in fish abundance towards the surface (as observed previously; e.g. Viehman et al., 2018), fish were assigned to depth bins following a beta distribution which peaked in the 2–3 m depth bins.
3. Bottom-oriented fish: for each tide bin, 1000 fish were assigned to the lowermost 10 m of the water column, using the same method as for Scenario 2 but with fish abundance increasing towards the sea floor and peaking in the lowermost bin.

4. Selective tidal stream transport (STST): fish were bottom-oriented during the flood tide (as in Scenario 3) and surface-oriented during ebb tide (as in Scenario 2), transitioning through the mid-water-column during slack tides. This scenario represented STST for a species migrating outward toward the open ocean, utilizing the current during ebb tide.
5. Mixed fish assemblage: Scenarios 1-4 were combined to represent a mix of species exhibiting different depth preferences and vertical movements. 50% of fish were randomly distributed, 20% were surface-oriented, 20% were bottom-oriented, and 10% exhibited STST. The proportions of fish exhibiting each vertical distribution were chosen arbitrarily for illustration purposes, as these proportions are not yet known for fishes utilizing Minas Passage.

To simulate the effects of entrained air contamination on acoustically-derived estimates of fish abundance, we removed counts from any depth bins within the entrained air layer. The 5th, 50th, and 95th percentile air layer depths represented “best”, “middle”, and “worst” contamination conditions, respectively. We also omitted fish below the nearfield range, as that portion of active acoustic data would not be useable either. We calculated “observed” fish abundances as the water column sums for each tide bin in these reduced datasets (making the assumption that all fish would be equally detectable by the echosounder). We then compared observed abundances to the known water column

sums (“actual” fish abundance), which was 1000 fish per tide bin.

For scenario 5, we also compared actual and observed fish vertical distribution for each stage of the tide: low (tide bins 1 and 24), high (tide bin 13), flood (tide bins 2 to 12), and ebb (tide bins 14 to 23). The vertical distribution for each tidal stage was constructed by breaking the water column into depth bins which spanned 5% of the total water column height (to account for changing water level), then summing the numbers of fish contained within each percentage bin.

3 RESULTS

The entrained air detection method worked well, with only a small number of files requiring manual adjustments to the automatically detected surface and entrained air lines (approximately 6% and 3%, respectively). Most entrained air was easily identifiable as backscatter extending downward from the surface, whereas most backscatter likely to be from fish did not overlap with the surface (**Figure 3**).

Despite the entrained air layer detection algorithm generally working well (**Figure 4A**), there were still instances where it was difficult to differentiate backscatter from bubbles or fish based on appearance alone. Some backscatter could have been either aggregated fish or partial, detached bubble plumes (**Figure 4C**). This ambiguous backscatter needed to be

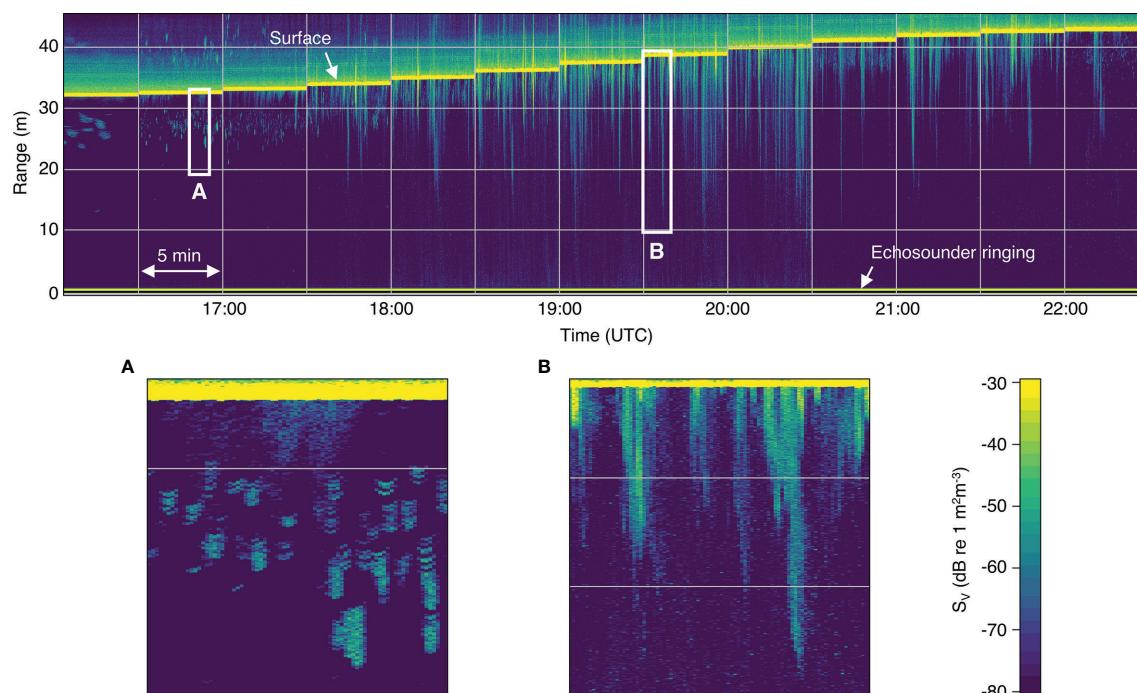


FIGURE 3 | Volume backscatter (S_v) echogram from a typical flood tide (on 22 May 2018). Low tide is on the left and high tide is on the right. Vertical gridlines separate the 5-min echosounder recording periods, which began every half hour (times shown in UTC). Horizontal gridlines indicate 10-m range bins (measured upwards from the transducer face). All following echograms use the same grid and color scale shown here. **(A)** Backscatter from small aggregations of fish visible near the surface near low tide. **(B)** Bubble plumes extending far into the water column during the peak flow.

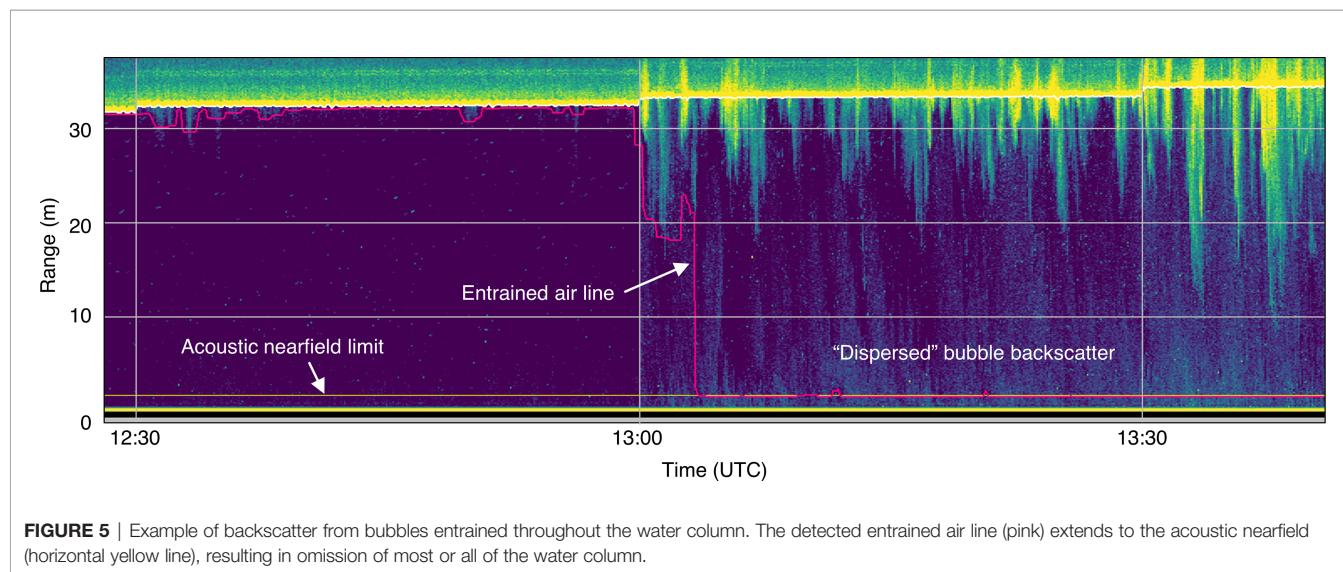
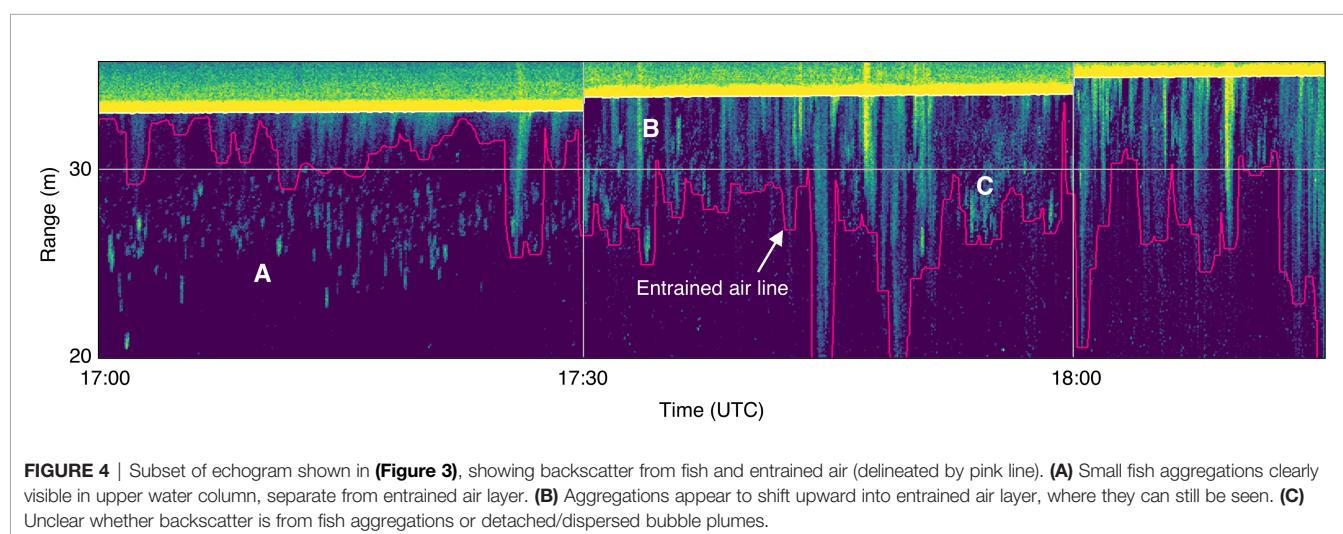
classified manually based on the appearance of the surrounding water column and neighboring recording periods. There were also many periods where fish were evident within bubble plumes but inseparable from plume backscatter, and therefore omitted (**Figure 4B**).

Backscatter from entrained air was not always confined to dense plumes of bubbles. At peak current speeds, when the plumes were most obvious, it was clear that the remaining water column was also subject to additional backscatter that often surpassed the same minimum threshold applied to the plumes (**Figure 5**). This more dispersed backscatter was likely also related to bubbles, given its strong association with deep bubble plumes, and it was therefore considered to be part of the entrained air layer. This situation is the cause of all recording periods that were missing 100% of their analyzable samples.

The final dataset consisted of 2583 5-min recording periods. Across all recording periods, 29% of all analyzable samples were

removed due to contamination from entrained air. Entrained air depth varied greatly over time, from the surface to the nearfield-exclusion line (**Figure 6B**). Consequently, the percentage of analyzable samples that would be omitted from any given recording period also varied from near 0% up to 100% (**Figure 6C**). Overall, 4% of recording periods were missing all of their analyzable samples, 16% were missing at least half of their samples, and 41% were missing at least a quarter. Almost all recording periods missing 100% of their samples occurred during peak flow near spring tides, when current speeds were highest (**Figures 6A, C**, orange bars). During neap tides, data loss in a given recording period did not often exceed 50% (**Figure 6C**, purple bars).

Due to entrained air extending downward from the surface, the lower water column was sampled more consistently than the upper water column. Across all recording periods, the uppermost 5% of the water column was only sampled 15% of the time, whereas the



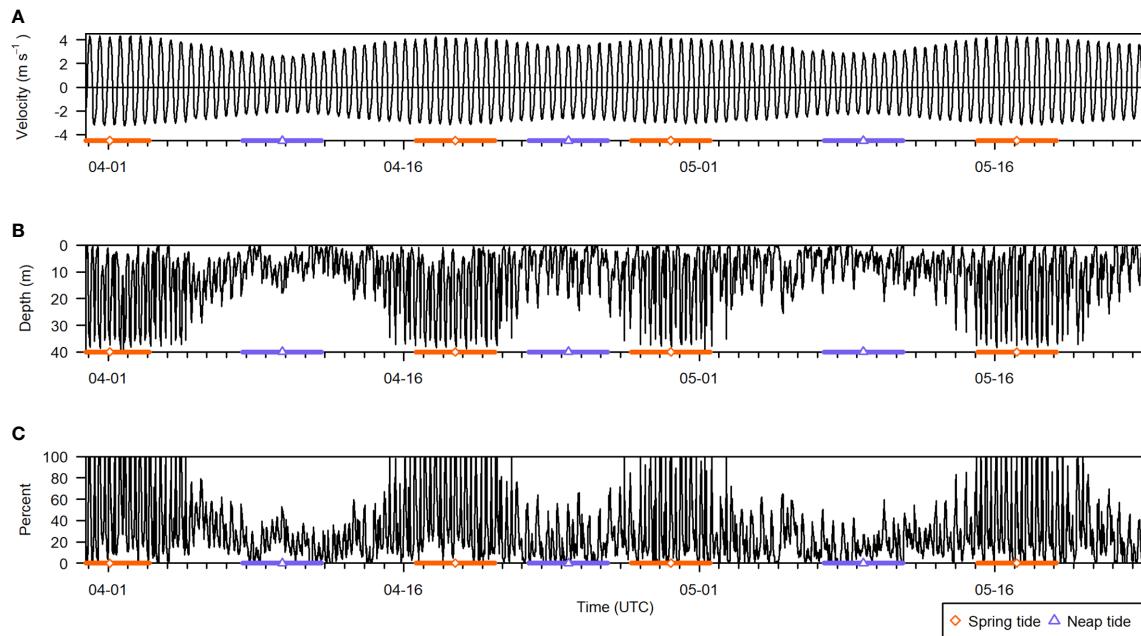


FIGURE 6 | Summary of all active acoustic recording periods from the deployment, spanning 30 March to 23 May 2018. **(A)** Current velocity (negative is ebb direction, positive is flood), **(B)** entrained air depth, and **(C)** percent of analyzable samples that were contaminated by entrained air in each recording period. The times of spring and neap tides are indicated by the orange diamond and purple triangle symbols, respectively, and the colored bars span 2 days on either side. Note that the entrained air layer depth stops at the acoustic nearfield, located 2.4 m above the seafloor.

5% above the nearfield was sampled 93% of the time (**Figure 7A**). The uppermost water column was almost exclusively sampled at current speeds less than $2\text{ m}\cdot\text{s}^{-1}$, and current speeds over $3\text{ m}\cdot\text{s}^{-1}$ were only well sampled in the lower half of the water column (i.e. in proportions similar to total recording periods, **Figure 7A**, upper panel). The fastest current speeds, greater than $4\text{ m}\cdot\text{s}^{-1}$, were very rarely sampled without contamination from entrained air, and only in the lower 45% of the water column.

There was a noticeable difference between depths and current speeds sampled during spring and neap tides (**Figures 7B, C**). Most current speeds greater than $3\text{ m}\cdot\text{s}^{-1}$ occurred during spring tides (**Figure 7B**, upper panel), but were not well sampled anywhere in the water column (**Figure 7B**, lower panel). During spring tide, contamination by entrained air at these faster speeds resulted in omitting at least 20% of recording periods throughout the water column, and more closer to the surface.

Conversely, during neap tides, the current speeds sampled in the lowermost 75% of the water column largely reflected the current speeds measured across all neap tide periods. Moreover, bins in the lower 70% of the water column were contaminated less than 10% of the time. Though surface depth bins were still under-sampled relative to lower bins, data collected during neap tides spanned the most representative range of current speeds for the largest portion of the water column.

The unequal representation of current speeds across depths was due to the correlation of entrained air depth with current speed (**Figure 8A**). Higher current speeds resulted in greater air contamination and data loss. The highest current speeds recorded

during either flood or ebb tide (occurring near spring tides) were often correlated with 100% contaminated samples (**Figure 8B**), though peak speeds were lower during ebb than flood (**Figure 9C**). The recording periods missing all or nearly all samples were mainly due to the “dispersed” bubble backscatter shown in **Figure 5**.

The correlation of entrained air depth with current speed meant the uncontaminated portion of the water column grew and shrank in an approximately 6-hour cycle, aligned with the tidal currents. This was very clear when data were summarized by tide bin (**Figure 9**).

The 5 hypothetical fish distribution scenarios are shown in **Figures 10A–E**, along with samples removed according to the 5th, 50th, and 95th percentile entrained air depth for each tide bin (hatchlines), and acoustic nearfield (crosshatched area along the bottom). Different levels of entrained air contamination had clear effects on fish abundances obtained from each of the 5 distribution scenarios (**Figure 11**). The magnitude of the impact on “observed” fish abundance over the course of the tidal cycle varied according to the underlying vertical distribution of fish. Generally, error in abundance estimates was greatest whenever fish were most concentrated in the upper water column (**Figures 11B, D**). For scenarios with fish in the upper- and mid-water-column, omission of data in the entrained air layer generated a distinct tidal pattern in observed fish abundance, as fewer fish were detected at higher current speeds (**Figures 11A, B, D, E**). This was true for all three entrained air levels applied to the simulated scenarios. Observed abundance of fish inhabiting the lowermost water column was primarily affected by the exclusion of data in the acoustic nearfield

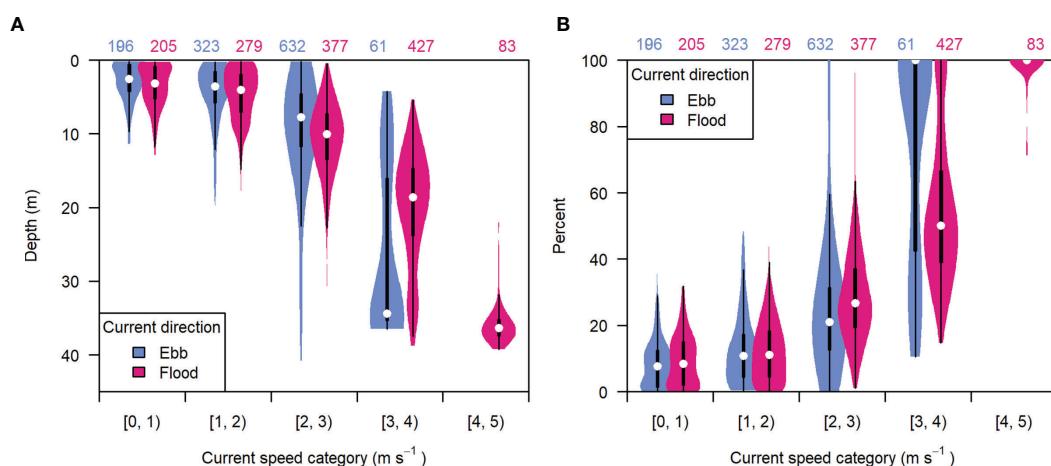
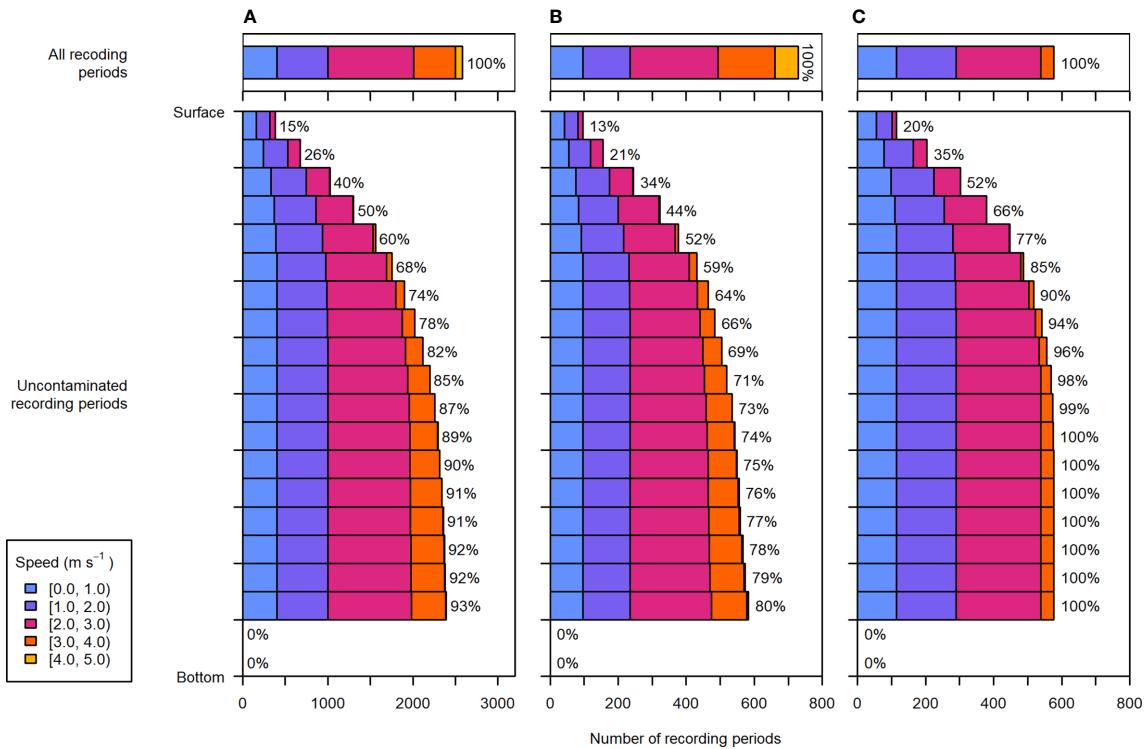


FIGURE 8 | The distribution of **(A)** entrained air depth and **(B)** percent of analyzable samples missing from each recording period, grouped by current speed category and tidal current direction (ebb or flood). Light blue indicates ebb tide, dark pink indicates flood tide. White points are the median value, boxes span the interquartile range (IQR), whiskers extend to 1.5*IQR, and violins span the minimum and maximum values in each group. Numbers at the top indicate the number of recording periods in each group.

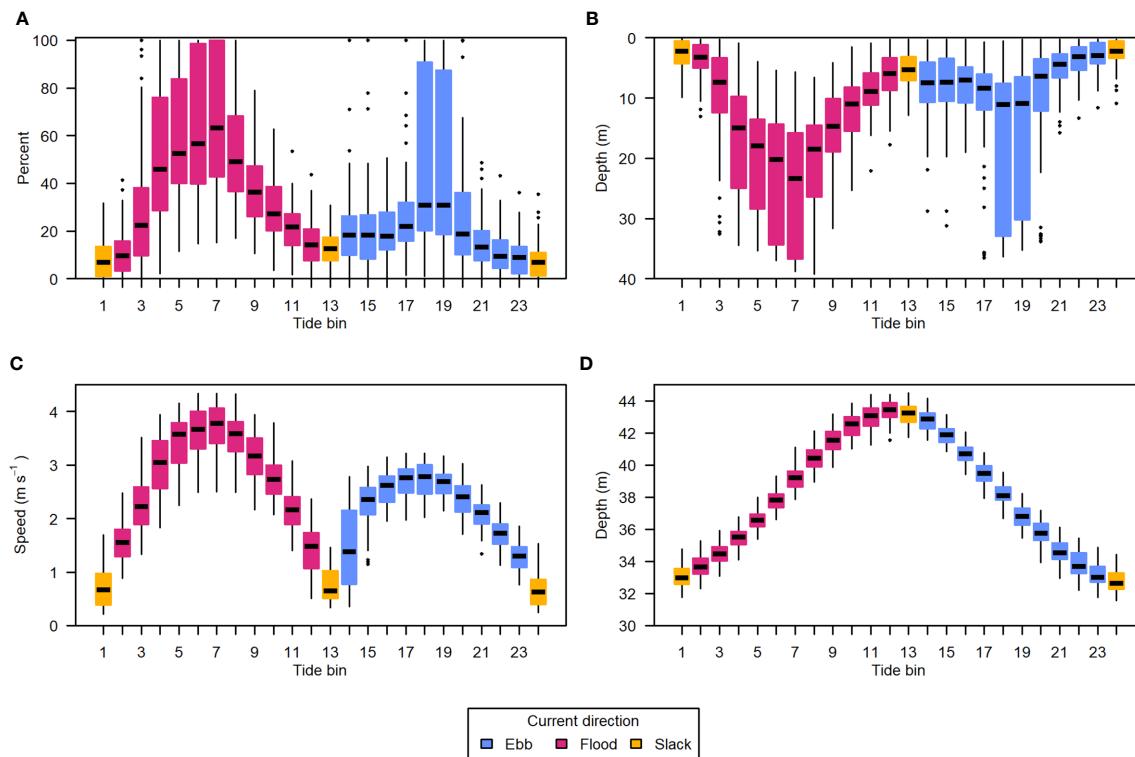


FIGURE 9 | Data from all recording periods summarized by position in the tidal cycle. **(A)** Percent of analyzable samples missing due to entrained air contamination, **(B)** depth of entrained air layer relative to the surface, **(C)** water column mean current speed, and **(D)** water column total depth. The tidal cycle was divided into 24 equal bins, each spanning approximately half an hour. Light blue indicates ebb tide, dark pink indicates flood tide, and yellow indicates the tide bins containing slack tides. Horizontal black lines are the median of each group, boxes span the interquartile range (IQR), whiskers extend to 1.5*IQR, and points are outlying values.

(a constant negative bias; **Figure 11C**, solid red line); however, lower-water-column observed abundances were also affected by the more extreme level of entrained air contamination (**Figure 11C**, dashed red line).

The observed vertical distribution of fish was also heavily affected by the differing levels of entrained air contamination, as demonstrated with scenario 5 (**Figure 12**). Estimates of fish abundance in the uppermost portion of the water column were most affected, particularly during the running tides (ebb and flood) when entrained air extended the farthest. Even the best case situation, using the 5th percentile of entrained air depths, resulted in excluding the majority of fish in the upper 10% (3.2–4.5 m depth) of the water column in all tidal stages, and the upper 20% (6.3–8.9 m depth) during flood tide. Due to the height of the acoustic nearfield above the sea floor, fish in the lowermost layers of the water column were also noticeably under-sampled.

4 DISCUSSION

Active acoustics technologies provide more detail and breadth of information on fish throughout the water column than any other sampling method currently available. However, entrained air poses a significant problem for active acoustics data collected at

tidal energy sites, and this must be considered when developing a study or environmental monitoring plan. The magnitude of entrained air contamination varies by site, and will be heavily dependent on local conditions (e.g. hydrodynamics, bathymetry, and weather; Baschek et al., 2006; Jech et al., 2021). The FORCE tidal energy test site has some of the fastest tidal currents on the planet ($> 5 \text{ m s}^{-1}$; Karsten et al., 2013), and its complex bathymetry and resulting dynamic current regime makes it one of the more challenging locations to use active acoustics instruments. Though the FORCE site is heavily affected by entrained air, the considerations discussed below will likely apply to echosounder users at other tidal energy test sites, as well.

Backscatter from entrained air contaminated 30% of all samples in our active acoustic dataset, and most of these were in the upper water column. However, contamination by entrained air varied greatly over time. The entrained air layer regularly spanned the entire water column during spring tides, though it rarely surpassed the middle water column during neap tides. So, while there were multiple days in a row with high levels of entrained air contamination, there were also periods of time with “best case” contamination levels, which would yield lower error rates in acoustically derived estimates of fish abundance. Peak current speeds were lower during neap tides than spring tides, but were well represented in the data throughout much of

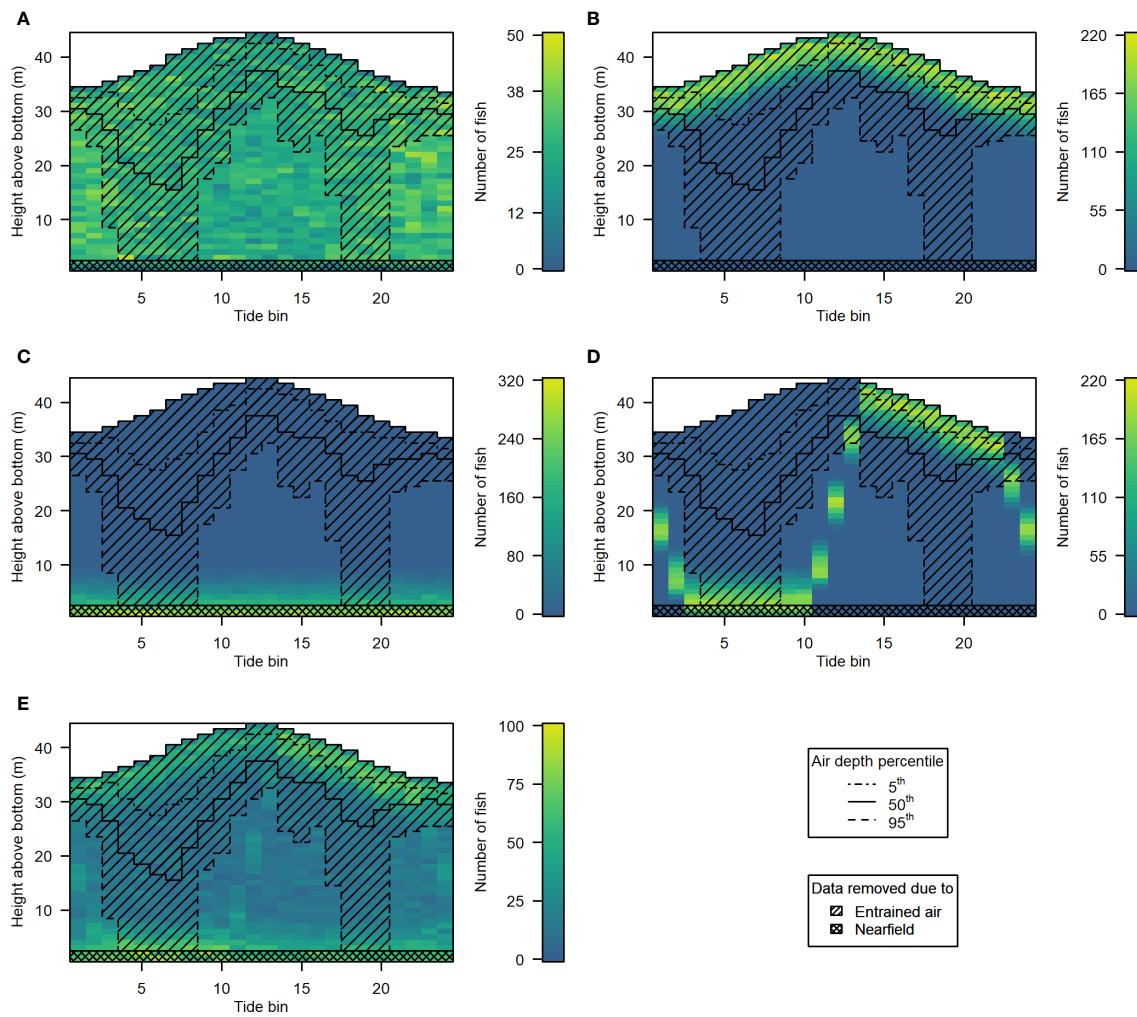


FIGURE 10 | Hypothetical fish distributions generated for scenarios 1 to 5 [(A-E), respectively]. Color indicates the number of fish in each depth bin. Datapoints excluded by the 5th (dot-dash line), 50th (solid line), and 95th (dashed line) percentile depth of the entrained air layer are indicated by the hatched area. The crosshatched rectangle covering the lowermost 2 m of the water column indicates the data that would be omitted due to the height of the acoustic nearfield above the sea floor.

the water column. Active acoustics data collected near neap tides are therefore likely to consistently yield more complete information on fish abundance and vertical distribution than data collected closer to spring tides.

That being said, we found that the distribution of entrained air backscatter over the shorter time scales (e.g. during a tidal cycle) could magnify the error introduced to estimates of fish abundance and vertical distribution. In our simulations, the tidally fluctuating extent of the entrained air layer generated false tidal patterns in observed fish abundance, depending on the underlying vertical distribution of fish. The largest errors occurred when fish were mainly present in the uppermost layers of the water column, as this generated the strongest tidal pattern in estimated abundance (Figures 11B, D). Fish in the mid-water-column were increasingly omitted as current speed increased (Figures 11A, E). Abundance estimates of fish in the

lowermost layers were mainly affected by the omission of data due to the height of the instrument above the sea floor and the extent of the acoustic nearfield, which introduced a constant negative bias (Figure 11C). Given the many species- and life-stage-specific depth preferences of fish, the prevalence of entrained air will therefore influence the extent to which different species are likely to be sampled by active acoustic instruments (for now ignoring other species-specific factors that affect detectability, such as their acoustic scattering properties; Horne, 2000).

The spatiotemporal fish distributions that we simulated were generalized examples of some commonly exhibited depth preferences among fish, and these may apply to many of the species likely to be in Minas Passage. For example, Scenario 1 may represent pelagic fish species that use most of the water column over the course of a day, including Atlantic herring,

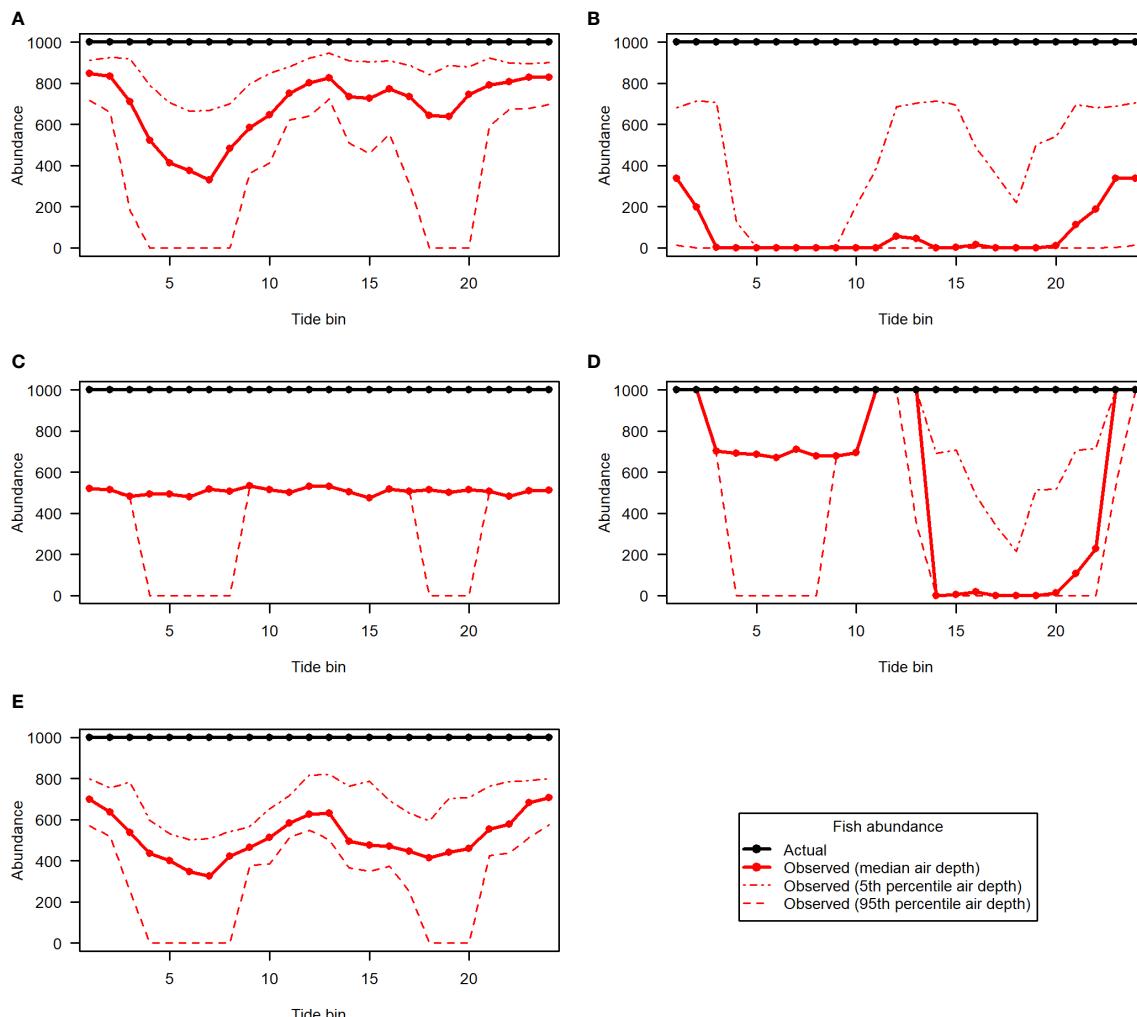


FIGURE 11 | Fish abundance over the course of the tidal cycle for simulated fish distribution scenarios 1 to 5 [(**A–E**), respectively]. Actual abundance (black) is shown in contrast to abundance that would be observed under different levels of entrained air contamination (red lines).

Atlantic mackerel, and striped bass (Castonguay and Gilbert 1995, Redden et al., 2014; Keyser et al., 2016; Viehman et al., 2018). Atlantic salmon, typically found in the uppermost 10 m in the northwest Atlantic, may be well-represented by Scenario 2 (Dutil and Couture, 1988; Sheehan et al., 2012). The Minas Basin is inhabited by a large number of demersal species, such as Atlantic cod, Atlantic sturgeon (*Acipenser oxyrinchus*), winter flounder (*Pseudopleuronectes americanus*), white hake (*Urophycis tenuis*), and dogfish (*Squalus acanthias*), among many others (Parker et al., 2007). Such species are likely to be on the seafloor or in the lowermost meters of the water column (e.g. Hobson et al., 2007), and therefore represented best by Scenario 3. Silver- and yellow-phase American eels have exhibited STST (Scenario 4) when migrating or moving around their home range, though with more frequent vertical movements during a tide and not always traversing the whole water column (Parker and McCleave, 1997). Other species have also exhibited STST, such as Atlantic cod

(though with smaller vertical movements above the seafloor; Arnold et al., 1994; Hobson et al., 2009), and possibly Atlantic mackerel (Castonguay and Gilbert 1995). The cyclic changes represented by Scenario 4 could also be extended to diel differences in vertical distribution, which would bring fish into and out of the under-sampled layers of the water column on a 24-hour cycle (rather than 12-hour). Many species and life stages of fish exhibit some level of diel vertical migration; e.g. Atlantic herring (Huse et al., 2012; Viehman et al., 2018) and alosids (American shad, *Alosa sapidissima*; Alewife, *A. pseudoharengus*; and river/Blueback herring, *A. aestivalis*; Stone and Jessop, 1992). Scenario 5 may represent a mixed species assemblage, which is more realistic for this location; however, the proportions of fish exhibiting each type of distribution were chosen somewhat arbitrarily, as there is little information to base these on.

It is unknown whether species-specific depth preferences will persist in high-speed tidal channels. Apart from STST, most

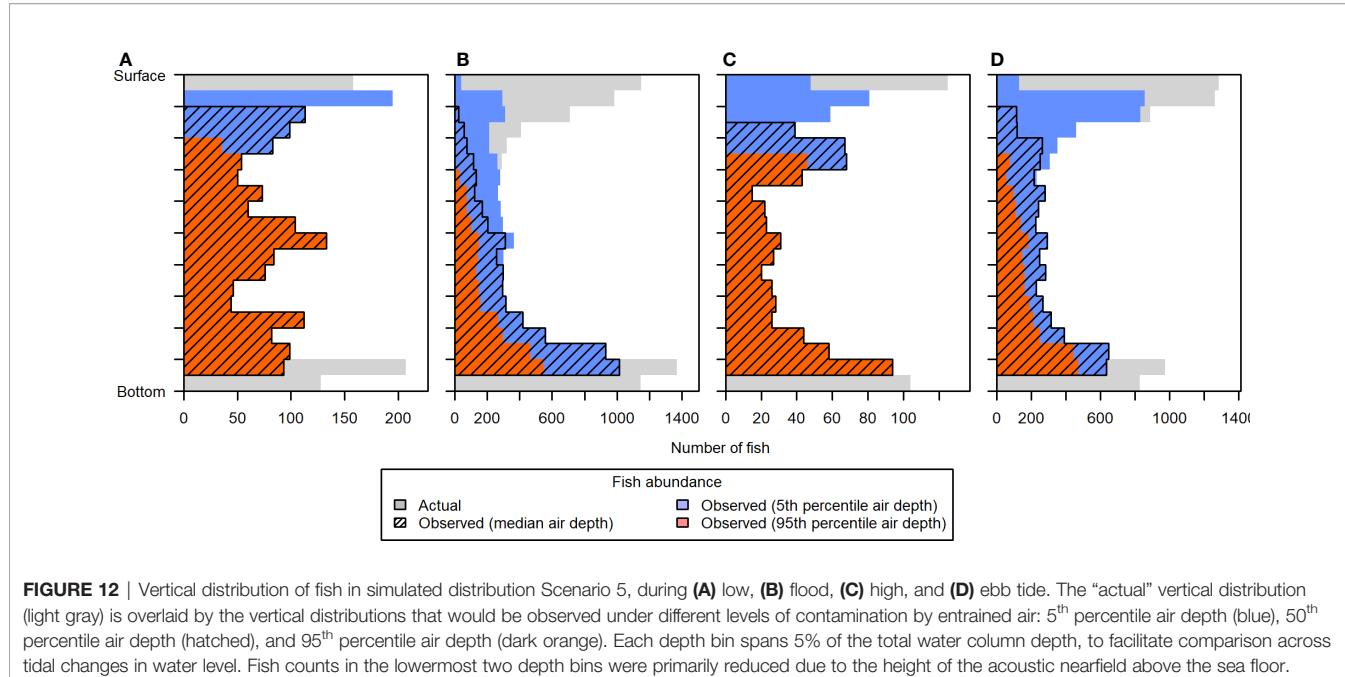


FIGURE 12 | Vertical distribution of fish in simulated distribution Scenario 5, during **(A)** low, **(B)** flood, **(C)** high, and **(D)** ebb tide. The “actual” vertical distribution (light gray) is overlaid by the vertical distributions that would be observed under different levels of contamination by entrained air: 5th percentile air depth (blue), 50th percentile air depth (hatched), and 95th percentile air depth (dark orange). Each depth bin spans 5% of the total water column depth, to facilitate comparison across tidal changes in water level. Fish counts in the lowermost two depth bins were primarily reduced due to the height of the acoustic nearfield above the sea floor.

knowledge of different species’ depth distributions and vertical movements comes from measurements obtained in less energetic environments. Some information exists for tidal channels. Atlantic sturgeon, for example, are normally a demersal species, but acoustically-tagged sub-adults were found to transit Minas Passage pelagically (Stokesbury et al., 2016), slightly deeper during ebb tide than flood tide (Lilly et al., 2021). This could increase their detectability by active acoustics instruments (deployed as presented here), as individuals would be more likely to be in the middle-water-column rather than in the omitted layers near the sea floor. Eight acoustically tagged silver-stage American eels have been detected in the FORCE test site, and though they were mainly detected during ebb tide, they did not appear to exhibit the vertical motions associated with STST which this species has displayed elsewhere, instead utilizing most of the water column (Redden et al., 2014). Striped bass have been detected at the FORCE test site from summer through winter, carrying out diel vertical migrations from 20–40 m depth during the day to the upper 30 m at night, except at temperatures below 1°C (Redden et al., 2014; Keyser et al., 2016). If Atlantic sturgeon, American eel, and striped bass all move pelagically at the FORCE site, then their availability to sampling by active acoustics may be best represented here by Scenario 1 (e.g., greater error in estimated abundance at peak flow). A better understanding of how different species utilize the water column in high-flow areas is necessary to assess their likelihood of sampling by active acoustic instruments.

Tidal and diel shifts in fish depth appear to be common across tidal energy sites, and these shifts could additionally influence the effects of entrained air on acoustically derived estimates of fish abundance and distribution. In Minas Passage, active acoustic measurements of fish (expected to be mainly overwintering Atlantic herring) found them to be more evenly spread out in

the water column at night than during the day (Viehman et al., 2018), which was also observed throughout the year for a mixed fish assemblage in Cobscook Bay, USA (Viehman et al., 2015). In a tidal channel in Tasmania, Australia, fish were more closely associated with the surface at higher current speeds (Scherelis et al., 2020). In the Holyhead Deep, UK, European sprat (*Sprattus sprattus*) carried out diel vertical migrations linked to the depth of light penetration (Whitton et al., 2020), and in Admiralty Inlet, USA, the vertical location of fish and zooplankton changed on a 24-hour cycle (Gonzalez et al., 2019). Periodic vertical movements such as these could bring fish into and out of the entrained air layer at regular intervals. The possible interaction of this periodic movement with tidal patterns in entrained air depth could mask or generate patterns in observed fish abundance over time (as seen for Scenarios 4 and 5; **Figures 11D, E**). These considerations also apply to fish shifting their depth usage in response to deployed MHK devices; for example, avoiding a device by moving higher or lower in the water column, and therefore potentially into or out of the entrained air layer. There has been some evidence that marine animals (including fish and marine mammals) may change their swimming behavior in response to device presence (Williamson et al., 2021).

While the upper water column and higher current speeds ($> 3 \text{ m}\cdot\text{s}^{-1}$) were under-sampled in this dataset, the lower 70% of the water column was generally well-sampled for current speeds up to $3 \text{ m}\cdot\text{s}^{-1}$ (**Figure 7A**). This is a large amount of data that can yield information on fish use of particular depth bins and how their depth may be influenced by a range of current speeds, all of which can inform our understanding of their likelihood of encountering an operating MHK device. However, information gained from a subset of the full range of depths and current speeds experienced at a site should not be assumed representative of the remaining, under-sampled depths and speeds. This is due

to the above links between species, current speed, and depth usage, but also to other potential effects of current speed on fish behavior. For example, at a tidal energy site in the Pentland Firth, UK, fish school abundance and physical size was found to change as current speed surpassed $1 \text{ m}\cdot\text{s}^{-1}$, potentially indicating an effect of physical forcing from tidal currents on schooling behavior (Fraser et al., 2018; Williamson et al., 2019). In these environments dominated by extreme physical forcing by tidal currents, it remains important to determine the extent to which information gathered at greater depths and lower speeds can be extrapolated (if at all). This could be examined at tidal energy sites that may have lower levels of entrained air contamination, or in future data collected with additional, complimentary sensors.

Additional sensing technologies will be essential for filling the gaps in active acoustics datasets that are left by entrained air, and for providing the necessary context for interpreting results. Acoustic telemetry has already provided valuable insight into when different species are likely to be present and where they are likely to be in the water column, and therefore how likely they are to be sampled with active acoustics in a deployment such as ours. Acoustically tagged individuals can be tracked over large distances, providing much-needed spatial context for the narrow volume sampled by an echosounder. Acoustic telemetry can help answer essential questions for building probability of encounter models, such as the proportion of a given fish population likely to come into the vicinity of a tidal turbine, and whether fish are actively swimming or drifting passively with the current. This adds to the information active acoustics provides for such models, which is fine-scale information on fish presence in the depths spanned by a given device, and how this changes over short and long time scales (for many more fish than can be tagged).

As with active acoustics, the efficiency of some acoustic telemetry systems can be reduced by current speed (Redden et al., 2014; Keyser et al., 2016; Tsitrin, 2019), resulting in fewer observations of fish location and depth during the time periods of greatest interest. This drop in detection probability could be related to the number of pulses that must be received from a given tag to allow a detection (Redden et al., 2014), the chance of a fish moving quickly past a receiver between acoustic tag transmissions (Keyser et al., 2016), as well as severe tilting of tethered receiver moorings in faster currents (Sanderson et al., 2017). These issues could be mitigated with appropriate choice of acoustic tags, mooring design, and receiver deployment (Sanderson et al., 2017; Sanderson et al., 2021). Recent experiments have shown drifting receivers could improve long-term tracking of individuals transiting Minas Passage, which wouldn't necessarily be possible with fixed receiver arrays (Sanderson et al., 2021). A combination of active acoustics and acoustic telemetry, using both stationary and drifting receivers, could yield a much more complete picture of fish use of a tidal energy site and their chance of encountering MHK devices.

Fish activity within the entrained air layer itself may be quantifiable using optical techniques. While bubble plumes are largely "opaque" to active acoustic instruments, cameras may be less affected unless bubble density is very high. Video has been used for studying fish interactions with tidal energy turbines

(Hammar et al., 2013; Broadhurst et al., 2014; Matzner et al., 2017), and in many other underwater applications requiring fish detection (e.g. Davidsen et al., 2005; Ellis and Bell, 2008). Optical systems cannot be used at night without additional lighting, which can affect fish behavior (Marchesan et al., 2005), and turbid or debris-laden water reduces fish detectability substantially (Ellis and Bell, 2008; Matzner et al., 2017). However, during daylight and with a few meters of visibility, there is an opportunity for video to be utilized for fish detection within the entrained air layer (Pattison et al., 2020). If optical data could be collected concurrently with an active acoustic system, ensuring sampled volumes overlap (or nearly do), results could help us understand how fish presence in the entrained air layer compares to abundance lower in the water column, and to what extent acoustically derived information from greater depths might be extrapolated upward.

Additional sensing technologies can help address another gap in active acoustics data analysis, which is the species and sizes of detected fish. This information would be helpful to those assessing the risk posed by tidal energy turbines, particularly when threatened or endangered species may be present. Information on fish species and length is also required to convert acoustic backscatter values to quantities of fish (Horne, 2000), unless fish are spread out enough to be detected and counted individually (e.g. Shen et al., 2016). Active acoustics data cannot usually provide identification of the detected scatterers to the species level without additional supporting information, which is typically obtained with trawls (Horne, 2000). The highly energetic and dynamic conditions at tidal energy sites often make them very difficult to sample safely or efficiently with trawls (Vieser et al., 2018), particularly at the spatial and temporal resolution required for classifying backscatter from a mixed assemblage within a rapidly changing environment. To date, most active acoustic studies at tidal energy sites have lacked physical sampling and stopped short of converting fish backscatter to estimates of abundance or biomass (Viehman et al., 2015; Fraser et al., 2018; Viehman et al., 2018; Staines et al., 2019; Williamson et al., 2019; Scherelis et al., 2020), with only one able to carry out concurrent trawling of a distinct layer of schools (Whitton et al., 2020).

Stereo optical camera or video systems may be useful alternatives to physically sampling fish at tidal energy sites. In recent years, species and length estimates from stereo camera systems have been found suitable for converting active acoustics backscatter to biological quantities, including in "untrawlable environments" (Rasmussen et al., 2021). Stereo optical systems are additionally non-lethal to sampled fish, less cumbersome than midwater trawls, and offer greater spatial resolution than trawls can provide (Boldt et al., 2018). Integrated optical-acoustic systems have been explored for MRE site monitoring, though so far only alongside high-frequency multibeam echosounders (Cotter and Polagye, 2020). Some challenges will need to be overcome for optical sensors to inform analysis of active acoustic data collected throughout the water column. As previously mentioned, optical systems require adequate lighting and water clarity for fish detection and identification. They also sample a much smaller volume than active acoustic instruments, which

can complicate comparison to the larger volume sampled acoustically and can result in low sample sizes (Boldt et al., 2018). In addition to optical systems, acoustic tag detections could provide insight on the species in the area of an echosounder; however, only the species that were tagged would be detected, and any effects of high flow on detection probability would need to be addressed.

Using multiple acoustic frequencies could also broaden the information that can be gained from an active acoustic dataset. Data from multiple frequencies could aid in identifying different groups of scatterers (e.g. air bubbles, fish with and without swim bladders, zooplankton, etc.) based upon their frequency response (Horne, 2000; Korneliusson, 2018). The frequency response alone may not always be sufficient to identify fish to the species level without supporting information on which species are likely to be present. However, it is possible that the frequency response could be used to improve identification and removal of backscatter from entrained air bubbles. The entrained air detection method we used here relied mainly on morphological characteristics of the backscatter, which, for entrained air, mainly took the form of plumes extending downward from the surface. This is similar to methods used at other tidal energy locations (Viehman et al., 2015; Fraser et al., 2018; Whitton et al., 2020). Manual scrutiny of the data showed that backscatter from entrained air did not always take this form (e.g., when the entire water column appeared to be contaminated by additional backscatter; **Figure 7**), and there were many near-surface backscatter features that were not easily classified as fish or bubbles based on morphological criteria alone (**Figure 5**). Adding a frequency response filter to the morphological one applied here could improve backscatter classification, and further ensure that remaining backscatter is likely to be from fish.

Multiple acoustic frequencies could also aid in characterizing the entrained bubbles themselves, which would be useful for assessing whether they are likely to affect the performance of surface-mounted echosounders transmitting sound through the air layer to quantify fish below (Dalen and Løvik, 1981; Vagle and Farmer, 1991; Jech et al., 2021). To our knowledge, frequency response has not yet been used for identifying or characterizing entrained bubbles at tidal energy sites. However, this approach would be worth exploring in new or existing multifrequency datasets, as it can inform data collection moving forward.

5 CONCLUSION

Active acoustic technologies are well-suited for collecting information on fish abundance and distribution throughout the water column, with the resolution and breadth required for predicting the likelihood of fish occurring at the same depths as MHK devices. This information can add to our understanding of potential encounter rates, and therefore risk devices pose to fish. However, the prevalence of entrained air at tidal energy sites often masks large portions of the upper water column from echosounders, particularly at high current speeds. In the dataset examined, the lower 70% of the water column was well-represented for current speeds under $3 \text{ m}\cdot\text{s}^{-1}$, but the upper water column and faster current speeds were under-sampled in

comparison. These under-sampled depths and periods of time constitute gaps in the active acoustic dataset that limit our ability to accurately measure fish abundance and vertical distribution, and therefore their potential overlap with MHK devices. Additional technologies, such as acoustic telemetry and optical systems, could be used concurrently with active acoustics to help fill these gaps and maximize the information that can be extracted from active acoustics data. While other tidal energy sites may experience less data contamination from entrained air, patterns in data loss are likely to be similar. The possible influence of these patterns on acoustically derived measurements of fish abundance and vertical distribution must be considered when planning a study or environmental monitoring plan at a tidal energy site, and when interpreting results from active acoustic datasets.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

HV, TB, and JD carried out preparation and planning of equipment deployment, including testing equipment settings and calibrating instruments. TB and JD deployed and retrieved the instruments. HV and JD processed the data. HV and DH contributed to the design of this study. HV completed the analysis, wrote the manuscript, and produced the tables and figures. All authors contributed to manuscript revision. All authors contributed to the article and approved the submitted version.

FUNDING

Data collection was funded by the Offshore Energy Research Association (project number 300-208).

ACKNOWLEDGMENTS

We would like to thank Huntley's Diving and Marine Services (Mike Huntley and the crew of the Nova Endeavor) for making the safe deployment and retrieval of the FAST-3 platform possible. We also thank the members of the FORCE team, Dr. Louise P. McGarry, Dr. Joel Culina, and Lilli Enders, for assistance with understanding and interpreting the data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.851400/full#supplementary-material>

REFERENCES

- Arnold, G. P., Greer Walker, M., Emerson, L. S., and Holford, B. H. (1994). Movements of Cod (*Gadus Morhua L.*) in Relation to the Tidal Streams in the Southern North Sea. *ICES J. Mar. Sci.* 51, 207–232. doi: 10.1006/jmsc.1994.1021
- Baschek, B., Farmer, D. M., and Garrett, C. (2006). Tidal Fronts and Their Role in Air-Sea Gas Exchange. *J. Mar. Res.* 64, 483–515. doi: 10.1357/002224006778715766
- Boldt, J. L., Williams, K., Rooper, C. N., Towler, R. H., and Gauthier, S. (2018). Development of Stereo Camera Methodologies to Improve Pelagic Fish Biomass Estimates and Inform Ecosystem Management in Marine Waters. *Fish. Res.* 198, 66–77. doi: 10.1016/j.fishres.2017.10.013
- Broadhurst, M., Barr, S., and Orme, D. (2014). *In-Situ* Ecological Interactions With a Deployed Tidal Energy Device, an Observational Pilot Study. *Ocean Coast. Manage.* 99, 31–38. doi: 10.1016/j.ocecoaman.2014.06.008
- Castonguay, M., and Gilbert, D. (1995). Effects of Tidal Streams on Migrating Atlantic Mackerel, *Scomber Scombrus L.* *ICES J. Mar. Sci.* 52, 941–954. doi: 10.1006/jmsc.1995.0090
- Copping, A. E., Hemery, L. G., Viehman, H., Seitz, A. C., Staines, G. J., and Hasselman, D. J. (2021). Are Fish in Danger? A Review of Environmental Effects of Marine Renewable Energy on Fishes. *Biol. Cons.* 262, 1–13. doi: 10.1016/j.biocon.2021.109297
- Cotter, E., and Polagye, B. (2020). Automatic Classification of Biological Targets in a Tidal Channel Using Multibeam Sonar. *J. Atmos. Ocean. Technol.* 37, 1437–1455. doi: 10.1175/JTECH-D-19-0222.1
- Dalen, J., and Løvik, A. (1981). The Influence of Wind-Induced Bubbles on Echo Integration Surveys. *J. Acoust. Soc. Am.* 69, 1653–1659. doi: 10.1121/1.385943
- Davidson, J., Svenning, M.-A., Orell, P., Yoccoz, N., Dempson, J. B., Niemelä, E., et al. (2005). Spatial and Temporal Migration of Wild Atlantic Salmon Smolts Determined From a Video Camera Array in the Sub-Arctic River Tana. *Fish. Res.* 74, 210–222. doi: 10.1016/j.fishres.2005.02.005
- Demer, D. A., Berger, L., Bernasconi, M., Bethke, E., Boswell, K., Chu, D., et al. (2015). Calibration of Acoustic Instruments. *ICES Cooperative Res. Rep.* No. 326. doi: 10.17895/ices.pub.5494
- Dutil, J. D., and Coutu, J. M. (1988). Early Marine Life of Atlantic Salmon, *Salmo Salar*, Postsmolts in the Northern Gulf of St. Lawrence. *Fish. Bull.* 86, 197–212.
- Ellis, W. L., and Bell, S. S. (2008). Tidal Influence on a Fringing Mangrove Intertidal Fish Community as Observed by *in Situ* Video Recording: Implications for Studies of Tidally Migrating Nekton. *Mar. Ecol. Prog. Ser.* 370, 207–219. doi: 10.3354/meps07567
- Fraser, S., Nikora, V., Williamson, B. J., and Scott, B. E. (2017). Automatic Active Acoustic Target Detection in Turbulent Aquatic Environments. *Limnol. Oceanogr-Meth.* 15, 184–199. doi: 10.1002/lom3.10155
- Fraser, S., Williamson, B., Nikora, V., and Scott, B. (2018). Fish Distributions in a Tidal Channel Indicate the Behavioural Impact of a Marine Renewable Energy Installation. *Energy Rep.* 4, 65–69. doi: 10.1016/j.egyr.2018.01.008
- Gonzalez, S., Horne, J. K., and Ward, E. J. (2019). Temporal Variability in Pelagic Biomass Distributions at Wave and Tidal Sites and Implications for Standardization of Biological Monitoring. *Int. J. Mar. Energy* 2, 15–28. doi: 10.36688/imej.2.15-28
- Hammar, L., Andersson, S., Eggertsen, L., Haglund, J., Gullström, M., Ehnborg, J., et al. (2013). Hydrokinetic Turbine Effects on Fish Swimming Behaviour. *PloS One* 8, e84141. doi: 10.1371/journal.pone.0084141
- Hobson, V. J., Righton, D., Metcalfe, J. D., and Hays, G. C. (2007). Vertical Movements of North Sea Cod. *Mar. Ecol. Prog. Ser.* 347, 101–110. doi: 10.3354/meps07047
- Hobson, V. J., Righton, D., Metcalfe, J. D., and Hays, G. C. (2009). Link Between Vertical and Horizontal Movement Patterns of Cod in the North Sea. *Aquat. Biol.* 5, 133–142. doi: 10.3354/ab00144
- Horne, J. K. (2000). Acoustic Approaches to Remote Species Identification: A Review. *Fish Oceanogr* 9, 356–371. doi: 10.1046/j.1365-2419.2000.00143.x
- Huse, G., Utne, K. R., and Fernö, A. (2012). Vertical Distribution of Herring and Blue Whiting in the Norwegian Sea. *Mar. Biol. Res.* 8, 488–501. doi: 10.1080/17451000.2011.639779. 5–6.
- Jech, J. M., Schaber, M., Cox, M., Escobar-Flores, P., Gastauer, S., Haris, K., et al. (2021). Collecting Quality Echosounder Data in Inclement Weather. *ICES Cooperative Res. Rep.*, 352. doi: 10.17895/ices.pub.7539
- Karsten, R., Swan, A., and Culina, J. (2013). Assessment of Arrays of in-Stream Tidal Turbines in the Bay of Fundy. *Phil. Trans. R. Soc A.* 371, 1985. doi: 10.1098/rsta.2012.0189
- Keyser, F. M., Broome, J. E., Bradford, R. G., Sanderson, B., and Redden, A. M. (2016). Winter Presence and Temperature-Related Diel Vertical Migration of Striped Bass (*Morone Saxatilis*) in an Extreme High-Flow Passage in the Inner Bay of Fundy. *Can. J. Fish. Aquat. Sci.* 73, 1777–1786. doi: 10.1139/cjfas-2016-0002. 12.
- Korneliussen, R. J. (2018). Acoustic Target Classification. *ICES Cooperative Res. Rep.* No. 344. doi: 10.17895/ices.pub.4567
- Lilly, J., Dadswell, M. J., McLean, M. F., Avery, T. S., Comolli, P. D., and Stokesbury, M. J. W. (2021). Atlantic Sturgeon Presence in a Designated Marine Hydrokinetic Test Site Prior to Turbine Deployment: A Baseline Study. *J. Appl. Ichthyol.* 37, 826–834. doi: 10.1111/jai.14274. 6.
- Marchesan, M., Spoto, M., Verginella, L., and Ferrero, E. A. (2005). Behavioural Effects of Artificial Light on Fish Species of Commercial Interest. *Fish. Res.* 73, 171–185. doi: 10.1016/j.fishres.2004.12.009. 1–2.
- Matzner, S., Trostle, C., Staines, G., Hull, R., Avila, A., and Harker-Klimes, G. (2017). *Triton: Igiugig Fish Video Analysis (PNNL-26576)*. Report to the U.S. Department of Energy (Richland, Washington: Pacific Northwest National Laboratory).
- Ona, E., and Mitson, R. B. (1996). Acoustic Sampling and Signal Processing Near the Seabed: The Deadzone Revisited. *ICES J. Mar. Sci.* 53, 677–690. doi: 10.1006/jmsc.1996.0087
- Parker, S. J., and McCleave, J. D. (1997). Selective Tidal Stream Transport by American Eels During Homing Movements and Estuarine Migration. *J. Mar. Biol. Ass. U.K.* 77, 871–889. doi: 10.1017/S0025315400036237
- Parker, M., Westhead, M., and Service, A. (2007). *Ecosystem Overview Report for the Minas Basin, Nova Scotia* (Fisheries and Oceans Canada, Dartmouth, NS: Oceans and Habitat Report 2007-05).
- Pattison, L., Serrick, A., and Brown, C. (2020). *Testing 360 Degree Imaging Technologies for Improved Animal Detection Around Tidal Energy Installations*. Final Report to the Offshore Energy Research Association of Nova Scotia (OERA) (Dartmouth, NS: Applied Oceans Research Group, Nova Scotia Community College).
- Rasmussen, L. K., Fields, S. A., Blume, M. T. O., Lawrence, K. A., and Rankin, P. S. (2021). Combined Video-Hydroacoustic Survey of Nearshore Semi-Pelagic Rockfish in Untrawlable Habitats. *ICES J. Mar. Sci.* 0, 1–17. doi: 10.1093/icesjms/fsab245
- R Core Team. (2021). *R: A Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing). Available at: <https://www.R-project.org/>.
- Redden, A. M., Stokesbury, M. J. W., Broome, J. E., Keyser, F. M., Gibson, A. J. F., Halfyard, E. A., et al. (2014). Acoustic Tracking of Fish Movements in the Minas Passage and FORCE Demonstration Area: Pre-Turbine Baseline Studies (2011–2013). *Final Report to the Offshore Energy Research Association of Nova Scotia and Fundy Ocean Research Centre for Energy* (Acadia University, Wolfville, NS: Acadia Centre for Estuarine Research Technical Report No. 118).
- Sanderson, B., Buhariwalla, C., Adams, M., Broome, J., Stokesbury, M., and Redden, A. (2017). Quantifying Detection Range of Acoustic Tags for Probability of Fish Encountering MHK Devices, In: *EWTEC 2017: Proceedings of the 12th European Wave and Tidal Energy Conference*; 2017 Aug 27 – Sep 1 (Cork, Ireland).
- Sanderson, B., Stokesbury, M. J. W., and Redden, A. M. (2021). Using Trajectories Through a Tidal Energy Development Site in the Bay of Fundy to Study Interaction of Renewable Energy With Local Fish. *J. Ocean Technol.* 16, 51–70. 1.
- Scherelis, C., Penesis, I., Hemer, M. A., Cossu, R., Write, J. T., and Guihen, D. (2020). Investigating Biophysical Linkages at Tidal Energy Candidate Sites: A Case Study for Combining Environmental Assessment and Resource Characterization. *Renew. Energy* 159, 399–413. doi: 10.1016/j.renene.2020.05.109
- Sheehan, T. F., Reddin, D. G., Chaput, G., and Renkawitz, M. D. (2012). SALSEA North America: A Pelagic Ecosystem Survey Targeting Atlantic Salmon in the Northwest Atlantic. *ICES J. Mar. Sci.* 69, 1580–1588. doi: 10.1093/icesjms/fss052. 9.
- Shen, H., Zydlowski, G. B., Viehman, H. A., and Staines, G. (2016). Estimating the Probability of Fish Encountering a Marine Hydrokinetic Device. *Renew. Energy* 97, 746–756. doi: 10.1016/j.renene.2016.06.026
- Simmonds, J., and MacLennan, D. N. (2005). *Fisheries Acoustics: Theory and Practice* (Oxford: Blackwell Science) 437 p.

- Staines, G., Zydlewski, G., and Viehman, H. (2019). Changes in Relative Fish Density Around a Deployed Tidal Turbine During on-Water Activities. *Sustainability*. 11, 2. doi: 10.3390/su11226262
- Stokesbury, M. J. W., Logan-Chesney, L. M., McLean, M. F., Buhariwalla, C. F., Redden, A. M., Beardsall, J. W., et al. (2016). Atlantic Sturgeon Spatial and Temporal Distribution in Minas Passage, Nova Scotia, Canada, a Region of Future Tidal Energy Extraction. *PloS One* 117, e0158387. doi: 10.1371/journal.pone.0158387
- Stone, H. H., and Jessop, B. M. (1992). Seasonal Distribution of River Herring *Alosa Pseudoharengus* and *A. Aestivalis* Off the Atlantic Coast of Nova Scotia. *Fish. Bull.* 90, 376–389.
- Tsitrin, E. (2019). Sofishticated Tracking: Improved Protocol for Acoustic Tagging of Sensitive Clupeid Fishes, and Applications in Investigating Postspawning Migration of Alewife (*Alosa pseudoharengus*) in Minas Basin, Bay of Fundy. [Master's thesis]. Wolfville (NS): Acadia University.
- Vagle, S., and Farmer, D. M. (1991). The Measurement of Bubble-Size Distributions by Acoustical Backscatter. *J. Atmos. Ocean. Technol.* 9, 630–644. doi: 10.1175/1520-0426(1992)009<0630:TMOBSD>2.0.CO;2
- Viehman, H., Boucher, T., and Redden, A. (2018). Winter and Summer Differences in Probability of Fish Encounter (Spatial Overlap) With MHK Devices. *Int. J. Mar. Energy* 1, 9–18. doi: 10.36688/imej.1.9-18. 1.
- Viehman, H., Zydlewski, G., McCleave, J., and Staines, G. (2015). Using Hydroacoustics to Understand Fish Presence and Vertical Distribution in a Tidally Dynamic Region Targeted for Energy Extraction. *Estuaries Coast* 38, 215–226. doi: 10.1007/s12237-014-9776-7. 1.
- Vieser, J. D., Zydlewski, G. B., and McCleave, J. D. (2018). Finfish Diversity and Distribution in a Boreal, Macrotidal Bay. *Northeastern Nat.* 25, 545–570. doi: 10.1656/045.025.0403. 4.
- Whitton, T. A., Jackson, S. E., Hiddink, J. G., Scoulding, B., Bowers, D., Powell, B., et al. (2020). Vertical Migrations of Fish Schools Determine Overlap With a Mobile Tidal Stream Marine Renewable Energy Device. *J. Appl. Ecol.* 57, 729–741. doi: 10.1111/1365-2664.13582
- Williamson, B. J., Blondel, P., Williamson, L. D., and Scott, B. E. (2021). Application of a Multibeam Echosounder to Document Changes in Animal Movement and Behaviour Around a Tidal Turbine Structure. *ICES J. Mar. Sci.* 78, 1253–1266. doi: 10.1093/icesjms/fsab017
- Williamson, B., Fraser, S., Williamson, L., Nikora, V., and Scott, B. (2019). Predictable Changes in Fish School Characteristics Due to a Tidal Turbine Support Structure. *Renew. Energy* 141, 1092–1102. doi: 10.1016/j.renene.2019.04.065
- Woolf, D. K. (2001). “Bubbles,” in *Encyclopedia of Ocean Sciences* (Cambridge, MA: Academic Press), 352–357. doi: 10.1006/rwos.2001.0062

Conflict of Interest: Author HV was employed by Echoview Software Pty Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Viehman, Hasselman, Douglas and Boucher. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Appendix III



Echofilter: A Deep Learning Segmentation Model Improves the Automation, Standardization, and Timeliness for Post-Processing Echosounder Data in Tidal Energy Streams

OPEN ACCESS

Scott C. Lowe^{1,2*}, Louise P. McGarry^{3*}, Jessica Douglas³, Jason Newport^{4,5}, Sageev Oore^{1,2}, Christopher Whidden^{1,4} and Daniel J. Hasselman³

Edited by:

Wei-Bo Chen,
National Science and Technology
Center for Disaster
Reduction (NCDR), Taiwan

Reviewed by:

Matthew Lewis,
Bangor University, United Kingdom
Wen-Ray Su,
National Science and Technology
Center for Disaster
Reduction (NCDR), Taiwan

***Correspondence:**

Scott C. Lowe
scott.lowe@dal.ca
Louise P. McGarry
louise.mcgarry@fundyforce.ca

Specialty section:

This article was submitted to
Ocean Solutions,
a section of the journal
Frontiers in Marine Science

Received: 02 February 2022

Accepted: 17 June 2022

Published: 09 August 2022

Citation:

Lowe SC, McGarry LP, Douglas J, Newport J, Oore S, Whidden C and Hasselman DJ (2022) Echofilter: A Deep Learning Segmentation Model Improves the Automation, Standardization, and Timeliness for Post-Processing Echosounder Data in Tidal Energy Streams. *Front. Mar. Sci.* 9:867857.
doi: 10.3389/fmars.2022.867857

¹Faculty of Computer Science, Dalhousie University, Halifax, NS, Canada, ²Vector Institute, Toronto, ON, Canada,

³Fundy Ocean Research Centre for Energy, Dartmouth, NS, Canada, ⁴DeepSense, Dalhousie University, Halifax, NS, Canada,

⁵Marine Environmental Research Infrastructure for Data Integration and Application Network, Halifax, NS, Canada

Understanding the abundance and distribution of fish in tidal energy streams is important for assessing the risks presented by the introduction of tidal energy devices into the habitat. However, tidal current flows suitable for tidal energy development are often highly turbulent and entrain air into the water, complicating the interpretation of echosounder data. The portion of the water column contaminated by returns from entrained air must be excluded from data used for biological analyses. Application of a single algorithm to identify the depth-of-penetration of entrained air is insufficient for a boundary that is discontinuous, depth-dynamic, porous, and varies with tidal flow speed.

Using a case study at a tidal energy demonstration site in the Bay of Fundy, we describe the development and application of deep machine learning models with a U-Net based architecture that produce a pronounced and substantial improvement in the automated detection of the extent to which entrained air has penetrated the water column.

Our model, Echofilter, was found to be highly responsive to the dynamic range of turbulence conditions and sensitive to the fine-scale nuances in the boundary position, producing an entrained-air boundary line with an average error of 0.33 m on mobile downfacing and 0.5–1.0 m on stationary upfacing data, less than half that of existing algorithmic solutions. The model's overall annotations had a high level of agreement with the human segmentation, with an intersection-over-union score of 99% for mobile downfacing recordings and 92–95% for stationary upfacing recordings. This resulted in a 50% reduction in the time required for manual edits when compared to the time required to manually edit the line placement produced by the currently available algorithms. Because of the improved initial automated placement, the implementation of the models permits an increase in the standardization and repeatability of line placement.

Keywords: machine learning, deep learning, hydroacoustics, entrained air, marine renewable energy, tidal energy, environmental monitoring, marine technology

1 INTRODUCTION

The need for clean, non-carbon emitting, alternatives for power production is well established (IPCC, 2021). With advancements in technology, energy extraction from kinetic marine sources (ocean current, tidal energy streams, and wave) have recently emerged as potential contributions to the suite of renewable energies for the generation of electricity (Cada et al., 2007; Roberts et al., 2016; Coppings et al., 2020; IRENA, 2020). In the case of energy extraction from tidal energy streams, tidal turbines are introduced into nearshore, coastal ecosystems that can be important habitats of major biological importance to fish for migration, nursery, and feeding activities (Blaber et al., 2000; Melvin and Cochrane, 2012; DFO, 2018; Tsirrin et al., 2022). The development of this nascent industry is therefore introducing new technologies and new but uncertain risks into the marine environment (DFO, 2008). In regions where fish stocks are managed or listed for special protections, regulators require monitoring for potential effects for fish as a condition for licensing tidal and other marine renewable energy (MRE) projects. Depending on local flow and bathymetric characteristics in the nearshore environments, the current flows suitable for tidal energy development can be turbulent (Cornett et al., 2015; Melvin and Cochrane, 2015; Williamson et al., 2017; Perez et al., 2021; Wolf et al., 2022), entraining persistent and deeply penetrating air into the water column. An efficient backscatterer of sound, the presence of air in the water column complicates the post-processing activities for data collected with acoustic instruments.

Hydroacoustic methods, applied to data collected with scientifically calibrated echosounders, are used to quantify the distribution and abundance of fish in the marine environment (Johannesson and Mitson, 1983; Fernandes et al., 2002; Benoit-Bird and Lawson, 2016). Echosounders emit a pulse of sound (a “ping”) into the water and record the magnitude of the returned backscatter (the “echo”) (Simmonds and MacLennan, 2005). The advantage of echosounders is the ability to sample the full water column in high spatiotemporal resolution. However, to achieve the goals of biological analyses for fish presence and distribution, backscatter recorded from physical interfaces must be excluded, including from the seafloor or sea surface (sea-air interface) and those portions of the water contaminated by backscatter from entrained air. The international standard solution to this is to use the software Echoview (Echoview Software Pty Ltd., Hobart, Australia), which enables advanced visualization and post-processing of hydroacoustic data. Echoview includes a library of highly configurable, parameterized algorithms by which to achieve the work of post-processing, including defining the boundaries of the region suitable for biological analyses.

The classical algorithms of Echoview generally produce appropriate placement for the lines designating the seafloor and sea surface given their continuous, strongly reflective, and non-porous natures. In contrast, the boundary of the entrained-air penetration is indistinct, porous, and discontinuous, and formed of local features that can only be distinguished from

biological features through their broader context. The profile of entrained air is further complicated for recordings at sites where the penetration of entrained air is influenced by tidal flow speeds which can range from slack tide to 5 ms^{-1} (10 knots), e.g. Bay of Fundy (Karsten et al., 2011). These characteristics limit the potential for classical algorithms to successfully identify the extent of entrained air within the water column. This lack of automation has important consequences for hydroacoustic data post-processing and analyses:

- substantial and time-consuming manual edits are required to refine the ping-by-ping demarcation of the ambit of entrained air,
- the quantity of edits generates analyst fatigue putting at risk the regions where the full force of analyst attention is needed for discerning usable data, and
- standardization and/or repeatability is impossible to achieve between analysts and within the work of a single analyst.

Machine learning is a methodology that enables the construction of models through the use of example input/output data. In particular, deep learning allows us to build the complex models which are necessary to solve challenging tasks which would otherwise require a human to laboriously perform (LeCun et al., 2015; Schmidhuber, 2015; Goodfellow et al., 2016). Deep learning models have revolutionized computer vision over the last 10 years (Krizhevsky et al., 2012; He et al., 2016; Bengio et al., 2021), have been successfully applied to image segmentation tasks (Ronneberger et al., 2015; Redmon et al., 2016; Minaee et al., 2022), and have attained human-level or superhuman performance at narrow tasks (Karpathy, 2014; Russakovsky et al., 2015; He et al., 2016; Santoro et al., 2016; Silver et al., 2017). We hypothesised that a deep neural network would be able to solve the task of placing the entrained-air line correctly. Hence we deployed machine learning methods, with a convolutional neural network architecture inspired by U-Net (Ronneberger et al., 2015) and EfficientNet (Tan and Le, 2019), to determine whether such models can generate an entrained-air line with better placement than the existing classical algorithms (as implemented in Echoview) and thus reduce the amount of human labour needed to complete this task.

In the deep learning framework, an artificial neural network is instantiated with a particular architectural design (with randomly initialized parameters), and its parameters are iteratively updated through gradient descent in order to maximize performance at the objective task. Through this training process, the network learns to approximate a function that maps a set of input stimuli to the correct outputs. In the context of this work, the input to the model was a 2-D image-like representation of the hydroacoustic recording for which the axes are depth and time, and the intensity at each pixel is the volume backscattering strength (S_v , dB re: 1 m^{-1}); we refer to this input as an *echogram*. The model's main output is a prediction of the depth of the entrained-air boundary line for each point in time (each ping). In addition to this, our model also predicts the depths of the seafloor and sea surface boundary lines, and (for each datapoint) whether the

echosounder was active (emitting pings) or passive (in listening-only mode).

Our final implementation, *Echofilter*, is openly available under the AGPLv3 license. Python source code and a stand-alone Windows executable are available at <https://github.com/DeepSenseCA/echofilter>, with command line interface (CLI) and application programming interface (API) documentation available at <https://DeepSenseCA.github.io/echofilter/>.

2 MATERIALS AND EQUIPMENT

2.1 Data Sources

Hydroacoustic data was collected from two tidal energy demonstration sites within the Bay of Fundy in Nova Scotia, Canada: Minas Passage in which flow speeds can exceed 5 m s^{-1} (Karsten et al., 2011) and Grand Passage in which flow speeds can achieve 2.5 m s^{-1} (Guerra et al., 2021).

“Stationary” data was collected using a calibrated Simrad EK80 WBAT 7° split-beam echosounder operating in continuous wave (CW) mode at 120 kHz in Minas Passage and in Grand Passage. The echosounder, with its transducer in an upward facing orientation was attached to a platform deployed to the seafloor (see **Figure 1**). The seawater depth at the platform location varied with tide height from 29 m to 44 m at Minas Passage, and 14 m to 20 m at Grand Passage. The echosounder was deployed in Minas Passage for three 2-month periods in 2018. Data was recorded for 5 minutes every half hour. Passive data collection with the echosounder in listening-only mode to document system self-noise and record levels of ambient sound present at 120 kHz was collected during two of the three deployments. There were two deployments of the echosounder in Grand Passage during late 2019 and early 2020. In both cases, the echosounder was deployed for less than 14 days. Data collection cycle in Grand Passage consisted of one-hour continuous data collection in alternating hours. Short durations of passive data were collected each hour.

“Mobile” data was collected from the Minas Passage site using a calibrated Simrad EK80 WBT 7° split-beam echosounder operating in CW mode at 120 kHz. The transducer was deployed

in a downward facing orientation attached via polemount to the vessel. The mobile survey pattern consisted of a set of six parallel transects, each of length 1.8 km and separated by 200 m, encompassing the Minas Passage multi-berth tidal energy demonstration site in the northern portion of the Passage, plus three reference transects located across the Passage near the southern shore. For ten of the seventeen mobile surveys, one additional transect was added to sample a region of interest in the demonstration site. The mobile surveys consisted of discrete 24-hour data collection periods during which the grid of transects was traversed four times, weather permitting. A completed grid consisted of one with-the-current and one against-the-current traverse of each transect. Seventeen such surveys were conducted between May 2016 and October 2018. Seawater depths ranged from 13 m to 67 m. Mobile data collection included periods of passive data collection with the transiting of each transect. No mobile data was collected in Grand Passage.

The echosounder data files were imported into Echoview (version 10.0) and post-processed in the typical way: (i) assigned calibration parameters, (ii) examined the data and removed noise, (iii) removed the passive data from further processing, (iv) set a line at constant range from the transducer face (1.7 m in this case) by which to exclude the transducer nearfield and, (v) applied Echoview algorithms to estimate, for each ping, the position of the seafloor (for downfacing echosounder) or sea surface (for upfacing echosounder) and the depth-of-penetration of the entrained air. In order to exclude the acoustic deadzone inherent in echosounder data (Simmonds and MacLennan, 2005), a one-meter offset was applied to the bounding line (seafloor or sea surface) and to the entrained-air line.

2.2 Data Partitioning

The full suite of Echoview files were divided into sets of files for training, validating, and testing the machine-learning models. The mobile downfacing dataset collected at Minas Passage consisted of 17 surveys, repeated at (a subset of) the same 10 transects on 17 different days spanning the course of three years. We selected two transects and placed all recordings from these in the test set. The remaining data was partitioned into training,

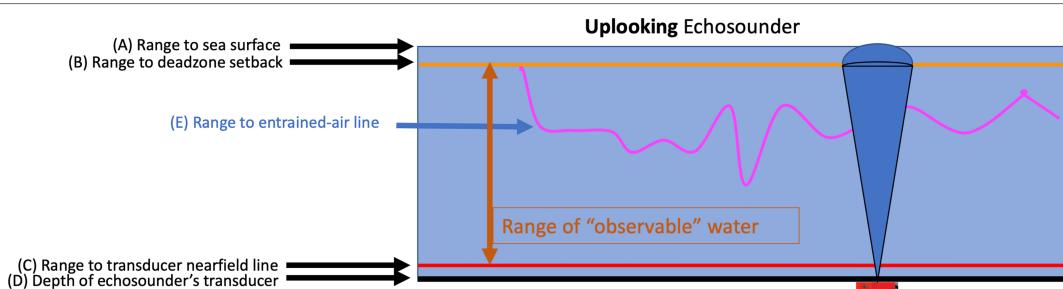


FIGURE 1 | Illustration of the boundaries defining the range of the observable water column for an upfacing echosounder. **(A)** Range to sea surface. **(B)** Deadzone setback: the range in which biomass within the acoustic beam is indistinguishable from the strong surface echo (Simmonds and MacLennan, 2005). **(C)** Nearfield line: a line set at a constant range to exclude the transducer “nearfield” in which the sound pulse is not yet organized (Simmonds and MacLennan, 2005). **(D)** Position of the upfacing transducer. **(E)** Entrained-air line: depth range of observable water is restricted to that not contaminated by entrained air.

validation, and (unused) “test2” partitions with an 80/10/10% split, stratified against the season in which the data was collected (winter vs non-winter) to ensure an equal split of the sparser winter recordings. The stationary data was grouped into blocks of 6 hours of consecutive recordings, and these blocks were partitioned at random (without stratification). We placed 80% of the MP:sta \uparrow and GP:sta \uparrow data files in the training partition. For the MP:sta \uparrow dataset, 10% of the data was used for model validation and 10% for final testing. Due to its smaller size, we did not use any GP:sta \uparrow data for the model validation process and kept the remaining 20% of the data for testing. The number of recordings and pings for each partition of each dataset is indicated in **Table 1**.

Files used for manual evaluation were selected from the MP:sta \uparrow and GP:sta \uparrow validation and test partitions, and chosen to ensure 24-hour coverage, with both neap tide and spring tide included. Stratifying the samples in this way ensured we would see examples of model performance under best-case and worst-case entrained-air scenarios. We also selected files from the MP:sta \uparrow and GP:sta \uparrow training partitions to inspect, in order to determine whether errors were due to applying the models on new data, indicating an issue with the model’s ability to generalize (overfitting), or whether there was also a problem on the training data, indicating an issue with the model design (underfitting).

3 METHODS

3.1 Echoview Algorithm

As a baseline to benchmark our models against, we used Echoview algorithms to generate seafloor (downfacing recordings), sea surface (upfacing recordings), and entrained-air boundary lines.

Two separate Echoview algorithmic approaches were used for estimating the ambit of entrained air within the stationary and mobile datasets, respectively. For data collected during stationary surveys, periods of recorded passive-data were excluded and a 2-D Gaussian blur was applied to the remaining echogram using the Echoview “XxY Convolution” operator. We used a 13-by-13 kernel, and a standard deviation of $\sigma = 2.0$ in both depth (over return sample indices) and time (over ping indices) dimensions. We then used the “Threshold Offset” line picking operator, with a minimum threshold boundary of -80 dB, to search below the surface line and define the ambit of entrained air for each ping. The position identified within each ping was then used as the automated demarcation between entrained air and water column in the original S_v echogram.

For data collected during mobile surveys, the “Best Bottom Candidate” line picking operator was used to estimate the ambit of entrained air at each ping. Because the Best Bottom Candidate operator identifies the first instance of strong signals deeper than weak signals in the water column, we first inverted the intensity of the S_v echogram by multiplying the values by -1 and adding -150 to each result. We then used the Best Bottom Candidate operator, parameterized with -70 dB for the minimum S_v for a good pick and for the discrimination level, to identify the interface between entrained air (“weak”) values, and water column (“strong”) values in the inverse echogram. The position identified for each ping was then used as the automated demarcation between entrained air and water column in the original S_v echogram. This standardized protocol was used for the last 8 mobile surveys (surveys 10 through 17). For the first 9 mobile surveys, the protocol was inconsistent (sometimes including smoothing operations on the line, and offsets of varying sizes) and yielded variable outputs; these surveys were excluded from our benchmarking analysis described in Section 4.1.

The Best Bottom Candidate line picking operator was used to estimate the position of the seafloor (downfacing recordings), and the sea surface (upfacing recordings). For seafloor detection, the default settings were used and included a bottom offset of 0.5 m, except for the first two of the seventeen mobile surveys for which some of the parameters were adjusted. For sea surface detection, the default settings were used except for the backstep discrimination level which was halved to -25 dB.

The entrained-air and seafloor lines produced by the Echoview algorithms were used as seed lines which expert human annotators, with reference to the S_v echogram including a minimum S_v threshold set to -66 dB, then manually adjusted to create corrected, finalized annotations. These human-refined annotations were used as the targets for training the machine learning model.

3.2 Data Preprocessing

Annotated data was stored in Echoview EV files, which contain both the S_v data and human-generated annotations for the boundary lines. The EV files were opened in Python using win32com to interface with Echoview’s programming interface (API), and exported into several files. The surface, seafloor, and entrained-air lines were exported into Echoview line (EVL) file format. The S_v data was exported into CSV format twice as follows. The first S_v CSV file (“raw S_v CSV”) was exported with all EV exclusion settings disabled, and contained the entire S_v

TABLE 1 | Summary of datasets used in this study: their recording locations, mobility, and recording orientation.

Dataset	Location	Mobility	Orientation	No. Recordings			No. Pings		
				Train	Val	Test	Train	Val	Test
MP:mob \downarrow	Minas Passage	Mobile	Downfacing (\downarrow)	727	91	245	1.21M	148K	394K
MP:sta \uparrow	Minas Passage	Stationary	Upfacing (\uparrow)	7,249	919	875	2.45M	305K	300K
GP:sta \uparrow	Grand Passage	Stationary	Upfacing (\uparrow)	118	0	28	0.36M	0	96K

We indicate the sizes of the dataset partitions, in terms of the number of contiguous recordings (duration dependent on dataset), and the total duration of the recordings measured in number of pings (k, thousand; M, million).

data in the EV file. The second S_v CSV file (“clean S_v CSV”) used the exclusion settings as implemented in the EV file so that all data which should be excluded from ecosystem analyses was masked out, appearing as the NaN indicator value -9.9×10^{37} in the output CSV. This means all datapoints above the entrained-air line, below the seafloor line (for downfacing recordings), passive data, bad data time periods where the analyst deemed a sequence of pings to contain data throughout the water column too contaminated by returns from entrained air or suspended sediment to use at all, and other miscellaneous localized “bad data” caused by anomalous events such as a rope drifting into view which the analyst had labelled for exclusion, were removed from the output (set to the NaN indicator value), leaving only the datapoints deemed as “good data” by the analyst.

Since this export process requires using Echoview to read in and export the data from the EV file, and Echoview is only available for Windows, this first step of the data processing pipeline must be performed on a Windows system with a licensed copy of Echoview installed. The remaining steps in the data processing and model training pipeline only require Python and can be run on any operating system.

The CSV files and EVL files were loaded into Python with a custom data loader. The depth resolution (and number of datapoints) per ping sometimes differed during a recording session, resulting in data with an uneven sampling resolution; we addressed this by finding the modal depth resolution across pings and linearly interpolating the data for each ping onto the same array of depth sample points. We created a “target mask” based on the location of NaN-values in the clean S_v CSV. This target mask corresponds to the overall target for the network’s output. The depth lines loaded from the EVL files were linearly interpolated onto the same set of timestamps as the S_v data.

We observed some discrepancies between the depth lines and the mask, which was caused by (1) off-by-one differences when the line threshold is applied in Echoview compared with our own interpolation of the line; (2) analysts using boxes or freehand regions to annotate exclusion regions which are adjacent to the boundary lines. We handled this by identifying the upper and lower contiguous extent of the masked out area to generate new lines from the mask. For the entrained-air line, we primarily used the deepest extent of the two options as provided via the line annotation and the mask annotation. For the seafloor line, we primarily used the original line annotation as the network’s output target, but we also produced a second line (with more aggressive removal) which extended higher up the water column to include any additional masked out area. The spare “aggressive” version of the seafloor line was included as an auxiliary target during training.

The surface line annotations were mostly unchanged by the annotators from the output produced by Echoview’s algorithms. These were observed to be mostly accurate, but contained occasional large jumps in value. These outliers were detected and removed by using a median filter as follows. We applied a median filter with a kernel length of 201 and observed the residual between the raw signal and the median filter. Values more than 5 standard deviations (robustly estimated from the interquartile range, $\sigma = \text{iqr}/1.35$) were set to the median value. We then applied a

median filter with a kernel length of 31 and removed anomalous values more than 4 standard deviations (robustly estimated from the interquartile range, $\sigma = \text{iqr}/2.56$) from the median. The second step was repeated until no anomalies were removed. Additionally, if the surface line was ever deeper than the entrained-air line, we set it to be the same depth as the entrained-air line. We found this anomaly removal process produced surface lines of sufficient quality. For downfacing samples, the surface line was set at 0 (coincident with the transducer face).

Passive data annotations were taken as hard-coded on/off cycles where known *a priori*. Otherwise, passive data collection periods were identified using a bespoke algorithm. The first S_v responses, corresponding to depths closest to the echosounder, have large intensities when the echosounder is active and much lower values when the echosounder is passive. We identified passive data periods by observing the first 38 depth sample points (after our interpolation step onto a common sampling grid). We took the difference in S_v between consecutive pings, and then the median across the first 38 depth samples for each ping. Median differences which exceeded ± 25 dB were identified as boundary points between passive and active recording periods.

Bad data periods were identified as collections of consecutive pings for which all the data was masked out. Periods of passive data recording were excluded from the bad data periods. Bad data periods in which the entrained-air line was at or below the seafloor line throughout the entire period were also excluded.

Bad data patches were identified by the “pixels” in the echogram which were masked out for any reason not already covered by being above the entrained-air line, below the seafloor line, during a period of passive data collection, or during a period of time identified as a bad data period.

Our data was comprised of both upfacing and downfacing echosoundings. In the recording data structure, and exported CSV files, the y -dimension is stored as increasing distance from the echosounder. To standardize our inputs to the network, we flipped the orientation of the upfacing data such that increasing indices in the y -dimension corresponded to increasing depth within the water column.

The number of timepoints per file was much larger than we could reasonably supply to the network as a single input “image”. Moreover, it is important that a single training batch contains a diversity of training data. To prevent the system from having to read in the contents of an entire recording file when needing to select only a small subset of the data to present for each training step, we broke the training data into chunks (shards) each with a length of 128 samples.

The pipeline for converting the CSV and EVL data into the preprocessed training shards can be executed with the command `echofilter-generate-shards`.

3.3 Training Inputs

When analysing echosounder data, it is common practice to offset the seafloor and entrained-air boundary lines by a fixed distance, 1 m for the echosounders used here. The purpose of the fixed-distance offsets are to exclude those portions of the data near boundaries, such as the sea surface or seafloor or the entrained-air

boundary, that may be biased due to the echosounder deadzone (see **Figure 1**) which is a function of the shape of the spherically spreading beam intersecting with a surface (Simmonds and MacLennan, 2005). In addition, it generates a buffer between the boundary of the entrained air and the data reserved for biological analyses, so as to exclude returns from entrained air adjacent to, but not connected to, the pronounced entrained-air boundary. This ensures processing errs on the side of excluding slightly more data, instead of accidentally including bad data. Some datasets had an offset of 1 m included in the line definitions, whereas others did not. We standardized this by subtracting offsets from the lines which had them included. Consequently, the model's target output is to predict the exact boundary locations, and offsets can be added to its outputs as appropriate via optional settings in the Echofilter API.

Each training input image was normalized independently, based on the distribution of S_v values within the training input. Normalization was performed by subtracting the median over all S_v values, and dividing by a robust estimate of the standard deviation derived from the interdecile range ($\sigma = \text{iqr}/2.56$). A small number of NaN values were present in the raw S_v data, and these were set to a value of -3 after the normalization step.

The maximum apparent range of the echogram can in some cases be several times further than the actual depth of the water column. This is because the depth dimension corresponds to the time-of-flight of the signals, the maximum of which is determined by a maximum range parameter chosen by the operator of the echosounder, which may be held the same across many recordings and thus may be much larger than the local depth of the water column. In order to get the most precise output for the entrained-air lines from the trained model, we would like to zoom in on only the salient region of the image: the water column, extending from seafloor to sea surface. This allows the model to predict the boundary point with sufficiently high granularity. However, since the depth of the seafloor is not necessarily known *a priori*, the model needs to be able to determine the depth of the seafloor, or range to sea surface, from the full echogram as well. For testing, we thus use a two-step approach. First, the full echogram is presented to the network and the seafloor and/or surface lines are predicted. These outputs are used to zoom in on the water column. Second, this zoomed-in echogram is presented to the network, and precise seafloor, surface, and entrained-air lines are generated.

Inputs to the network are samples from the distribution of plausible echograms. During training, inputs to the network were drawn from the training partition and augmented with several operations. (i) Temporal stretching, stretch/squashed by a factor sampled log-uniformly from [0.5, 2]. (ii) Random depth cropping. With $p = 0.1$, the depth was left at the full, original extent. With $p = 0.1$, the echogram was zoomed in on the range from the shallowest surface depth to the deepest seafloor depth (the "optimal" zoom). With $p = 0.4$, the echogram was zoomed in to a random range of depths close to the optimal zoom, stretched or squashed by up to 25%, but never so much as to remove more than 25% of the entrained-air line or (for downfacing recordings) more than 50% of the seafloor line. With $p = 0.4$, the echogram was zoomed in to a random range of depths between the full original extent and the "optimal" extent. Depth upper and lower

limits were selected uniformly across the appropriate range. (iii) Random reflection in the time (ping) dimension, performed with $p = 0.5$. (iv) "Color" jitter. We applied a "brightness" augmentation by offsetting normalized S_v values by a random additive offset chosen uniformly from [-0.5, +0.5], and a "contrast" augmentation by multiplying normalized S_v values by a random multiplicative factor chosen uniformly from [0.7, 1.3]. The same random offset and factor were used for each pixel in an echogram input. The order of the brightness and contrast augmentations was randomly selected for each input. (v) Elastic grid deformation, performed with $p = 0.5$. Elastic deformation was performed separately in the depth and time dimensions, to create an elastic grid deformation. We chose to deform the dimensions separately, instead of jointly as per a standard elastic deformation where space is stretched/squashed in a 2-D manner, because our targets are mostly at the ping level (depth of lines at each ping, whether the ping is passively or actively sampled, etc.) and apply to the entire column of data. A standard 2-D elastic deformation would break the relationship between our input and target; performing a joint elastic deformation on the echogram input would make it challenging to relate the input to the targets. We used $\sigma = 8$ in the time dimension, $\sigma = 16$ in the depth dimension, and $\alpha = 0.1$ in both dimensions. The echogram was interpolated in 2-D, with the interpolation order randomly selected from linear, quadratic, and cubic (equal weighting).

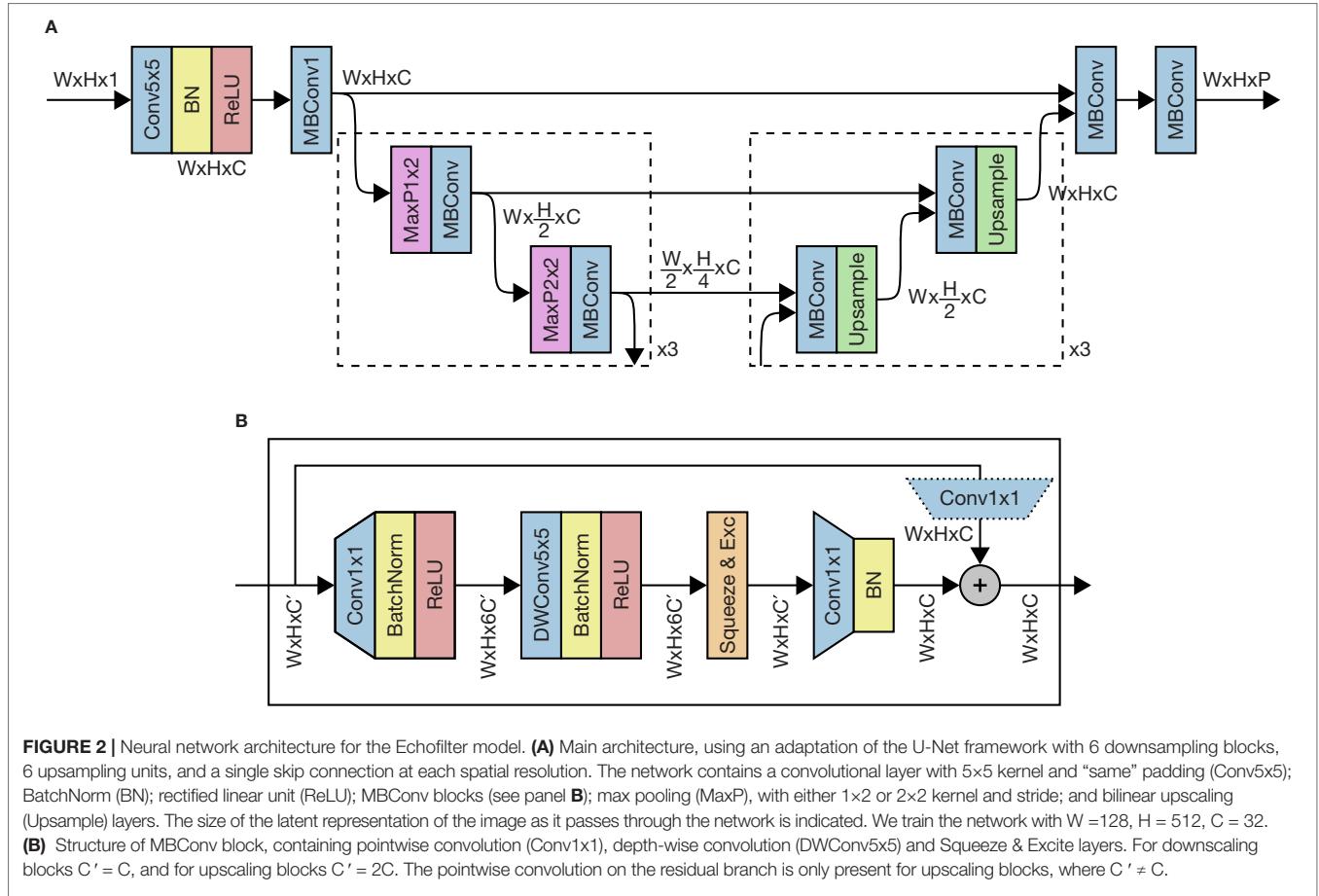
Finally, the echogram was rescaled to size (128, 512) pixels (time-by-depth) for presentation to the network with nearest-neighbour interpolation.

3.4 Model Architecture

The model architecture used is a U-Net (Ronneberger et al., 2015) with EfficientNet MBCConv blocks (Howard et al., 2017; Tan and Le, 2019), illustrated in **Figure 2**. This architecture is a convolutional neural network (CNN) with residual skip connections across blocks, 6 encoder layers where the size is spatially compressed, 6 decoder layers where the size is expanded back to the original input dimensions, and skip connections from the encoder to decoder blocks. The network has a backbone width of 32 channels throughout, and each MBCConv block is inverse residual with an expansion factor of 6 (except the very first block, which has an expansion factor of 1). We used depthwise-separable convolutions with a kernel size of 5, and ReLU activations. We used Squeeze & Excite attention layers (Hu et al., 2019) on each block with a reduction factor of 2. In total, our final models each had 1.63M trainable parameters.

Since the input is rectangular, with higher resolution in the depth dimension, we downscaled the time dimension at a slower rate than the depth dimension. Downscaling was performed with max-pooling using a kernel size and stride of either 1×2 or 2×2 (alternating blocks). The depth dimension was downsampled after every block, whilst the time dimension was downsampled every other block.

The decoder branch was a mirror of the encoder: upscaling in the depth dimension after every block, and in the time dimension every other block. Upscaling was performed using bilinear interpolation with torch.nn.Upsample.



The model has 10 output planes. These correspond to the probability (represented in logit form) that a pixel is at the boundary point for: (1) the entrained air (expanded), (2) the entrained air (original), (3) the seafloor line (expanded), (4) the seafloor line (original), (5) the surface line; and the probability (logit) that a pixel is within (6) a passive data period, (7) a bad data period (vertical region), (8) a miscellaneous bad data patch (to accompany expanded lines), (9) a miscellaneous bad data patch (to accompany original lines), (10) a miscellaneous bad data patch (to accompany original seafloor/expanded entrained air).

In practice, the expanded/original lines are almost identical and their pseudo-replication during training was superfluous, but their inclusion did indirectly increase the contribution of the entrained-air and seafloor lines towards the overall loss term. When performing inference with the model, we discard outputs 2, 4, 9, and 10.

For the Bifacing model, these 10 output planes are replicated three times. One is the standard output, the second are logits which are updated only on downfacing inputs, and the third are logits which are updated only on upfacing inputs. In this way, the model learns to represent conditional probabilities $P(\text{boundary} \mid \text{upfacing})$, etc. After training the model, we can ask it to predict the boundaries and masks agnostic of the orientation

of the recording, or conditioned on the orientation (upfacing or downfacing).

3.5 Model Training

The model was optimized with gradient descent to minimize a loss function. The loss function acts as a proxy for the task of interest; a high loss corresponds to worse performance on the task, and a low loss to better performance. We constructed our loss function as the sum of several terms, each corresponding to one of the output planes produced by the model. The loss terms for the seafloor, sea surface, and entrained-air lines were each the cross-entropy between the column of logits across all depths for a single ping against a one-hot representation of the depth of the line. The loss terms for the passive collection and bad data periods were binary cross-entropy between the model’s output for that ping (a single scalar, after collapsing the depth-dimension with log-avg-exp; Lowe et al., 2021) and the target value. The loss term for the localized bad data regions was binary cross-entropy. Outlying surface line values detected with our algorithm during preprocessing were masked out from the training objective. We took the mean over pings for all loss terms. We took the mean over the batch dimension; for outputs conditioned on the orientation of the echosounder, we masked out irrelevant samples

before taking the batch-wise mean. When training the bifacing model with conditioning signals, all stimulus presentations were double-counted and the entire loss was divided by two to correct for this.

The model was optimized using the RangerVA optimizer (Wright, 2019), which combines RAdam, Lookahead, and gradient centralization (Zhang et al., 2019; Liu et al., 2020; Yong et al., 2020; Tong et al., 2022), with a weight decay of 1×10^{-5} . We used a batch size of 12 samples, and stratified the batches to contain the same ratio of downfacing and upfacing samples as available in the aggregated training set. The learning rate (LR) followed a cyclic learning rate schedule (Smith, 2015; Smith and Topin, 2017; Smith, 2018). In each cycle, the learning rate was warmed up for the first 10% of training, held constant for 40% of training, then warmed down for the last 50%. During the LR warmup period, the momentum was decreased from a maximum of $\beta_1 = 0.98$ to a base of $\beta_1 = 0.92$, and then increased back to 0.98 during the LR warmdown period. Both the LR and momentum were increased and decreased with cosine annealing. The second moment parameter was held constant at $\beta_2 = 0.999$ throughout training. In the first cycle, the model was trained for 100 epochs with a maximum learning rate of $LR = 0.012$. In subsequent cycles, the training duration was progressively doubled and maximum learning rate halved. We trained two models: the Bifacing model was trained for three cycles (700 epochs), whilst the Upfacing-only model was trained for two cycles (400 epochs). The model parameters were saved at the end of each cycle for subsequent analysis. We chose to stop the cyclic training process when the model's validation performance had reached a plateau.

The Upfacing model was trained on MP:sta↑ and GP:sta↑ datasets, which contain only upfacing S_v recordings. The Bifacing model was trained on the MP:mob↓ dataset in addition to the MP:sta↑ and GP:sta↑ datasets. To address the smaller size of the GP:sta↑ dataset, we upsampled it by presenting echograms drawn from it twice per epoch instead of once (for both models).

The model architecture and training hyperparameters were each selected over a series of manual searches against the validation partition with short training durations of 5 or 20 epochs.

The network was trained using PyTorch 1.2.0 and CUDA 10.2. The model training and testing were done on the DeepSense high performance computing cluster with each training cycle or test using a 20 Core IBM Power8NVL 4.0 GHz compute node with 512 GB of RAM and a pair of NVIDIA Tesla P100 GPUs with 16 GB of GPU memory.

The Echofilter model can be trained using the command `echofilter-train`, with training parameters set at the command prompt.

3.6 Model Output Post-Processing

The neural network model is configured to generate predictions for each output type at the pixel level. That is to say, for each pixel in the input echogram, the network predicts a set of output variables at that particular pixel. For the passive data and bad

data periods, we convert this 2-D output into a 1-D time series by taking the log-avg-exp over the depth dimension (Lowe et al., 2021).

We converted the model's output into lines as follows. For each boundary line, our model predicts the probability that each pixel is the location of said boundary. We integrated this probability across depth to create a cumulative probability density estimate, and identified the depth at which the cumulative probability exceeded 50%. In so doing, we generate a boundary depth prediction for every ping.

For the purposes of the machine learning model, all salient information needed to produce its outputs is contained in data at, or immediately surrounding, the water column. However, some echosound recordings have much greater range than this, extending out beyond the water column with a large number of samples. In order to put the echogram into the network, we scale the depth dimension down to 512 pixels. For echograms much larger than the water column, this step incurs a loss of information, since the water column may occupy only a small fraction of the 512 pixel resolution.

In order to alleviate this issue, the Echofilter protocol may run the echogram through the network twice, once zoomed out and once zoomed in on the water column. In the first instance, the echogram is "zoomed out" to the maximum extent and scaled down to 512 pixels. The depth of the seafloor or sea surface line is noted (the choice of line depending on echosounder orientation), and used to estimate the extent of the water column. Using a robust estimate of the standard deviation of depths in this line, we set our limit to be 4 standard deviations out from the mean of the line, or the furthest extent of the line, whichever is least distal. For upfacing recordings, we zoom in on the range from the deepest recording up to this depth minus an additional 2 m. For downfacing recordings, we zoom in on the range from the shallowest recording depth down to this depth plus an additional 2 m. After cropping the echogram down to this range of depths, we scale it down to 512 pixels and present it to the network again. The output from the second, "zoomed-in" presentation is used to determine the final entrained-air, surface/seafloor lines and other outputs.

This "zoom+repeat" technique provides gains (see Section 4.1.2), but we expect it to be needlessly expensive when only a small fraction of the echogram is outside the water column. For this reason, we only perform the second presentation if more than 35% of the echogram would be cropped out. This setting can be controlled with the `--autozoom-threshold` argument to Echofilter.

In our analysis, we observed that Echofilter's predictions of the locations of "bad data" were not sufficiently accurate. Furthermore, the mask can include a large number of small disconnected areas, which results in a inconveniently large number of regions to import into Echoview. In order to counter this, we can merge together regions with small gaps in between them, and impose a minimum size threshold on regions to be included in the output. We merged together consecutive passive regions annotations provided by the model with a gap smaller than 10 pings, and similarly for bad-data period labels. Any

remaining regions shorter than 10 pings in length were omitted from the final output. For bad data patches, any patch with an area smaller than 25 ping-metres was omitted from the final output. In extremis, we can omit all bad-data annotations from Echofilter’s region outputs.

An alternative solution to noisy outputs is to spatially smooth the output probabilities. We can apply a Gaussian smoothing kernel across each output plane before converting the logits into probabilities, and subsequently into lines and regions. However, we did not find this process yielded better results.

Lines and regions produced by Echofilter are exported into Echoview line (EVL) and region (EVR) files so they can be imported into Echoview. Additionally, the Echofilter command line supports saving lines and regions directly into the EV file which it is processing (Windows OS and a licensed copy of Echoview required), removing the subsequent step of manually importing the files.

Inference using a pretrained model can be performed on EV (Windows-only) or CSV files with the command `echofilter`. Pre-processing and post-processing options can be configured be set at the command prompt.

3.7 Performance Quantification

To compare the quality of the outputs from the Echoview algorithm and our Echofilter models, we used a selection of metrics to quantify their performances. The results shown in Section 4.1 were determined by using these performance measurements on the test partition of each dataset. The test partition was neither seen by the model during training, nor used to optimize the model architecture and training process.

3.7.1 Intersection-Over-Union

The model’s output was evaluated using the intersection-over-union score (IoU), also known as the Jaccard index metric (Jaccard, 1912), and Jaccard similarity coefficient score. This metric is commonly used to evaluate the performance of image segmentation models within the field of computer vision. The IoU of two masks is calculated by assessing their overlap; it is the ratio of the size of the intersection of the two masks against their union:

$$\text{IoU}(\text{annotated}, \text{predicted}) = \frac{\text{Area}(\text{annotated} \cap \text{predicted})}{\text{Area}(\text{annotated} \cup \text{predicted})}. \quad (1)$$

For this study, one mask identifies the data marked as “good” by a human annotator, and the other mask is the data marked as “good” by the model. A higher IoU is better, indicating the two masks are better aligned. We chose to use this performance metric (instead of accuracy, etc.) because it is robust against padding the echogram with irrelevant range outside of the water column (below the seafloor for downfacing recordings, or above the sea surface for upfacing recordings).

For the MP:sta↑ and GP:sta↑ datasets, the IoU measurements we report are the total area of the mask intersections across the whole test set, divided by the total area of the union of the two masks (i.e. the division operation performed after the summation). For the MP:mob↓ dataset, the IoU reported

is the average IoU over all the EV files in the test set (i.e. the division operation performed before the mean). In both cases, we determine the standard error (SEM) by considering the distribution of IoU scores over EV files. For any recording where the target mask is all marked as False (no good data), the intersection of the predicted area with the target area is always 0, and any prediction from the model results in an anomalously minimal score. Consequently, we excluded examples where the target was an empty mask when measuring the SEM.

3.7.2 Mean Absolute Error

We performed further evaluation of the model’s outputs using the mean absolute error (MAE) performance metric. The MAE is defined as

$$\text{MAE} = \frac{1}{n} \sum_{\forall i} |y_i - \hat{y}_i|, \quad (2)$$

where y_i is the target value for the i -th ping, \hat{y}_i is the predicted value generated by the model, and n is the number of pings to average over. We applied the MAE to measure the quality of the output lines. In this context, the MAE corresponds to the average distance (across pings) of the model’s line from the target line. A smaller MAE is better, indicating the model’s line is (on average) closer to the target line.

When measuring the MAE of the lines, we excluded pings which were marked as being within a passive or bad data region in the target annotations. To find the overall MAE, for MP:mob↓ we determined the MAE within each file and averaged over files; for MP:sta↑ and GP:sta↑ we averaged over all pings across all files in the test set, weighting each ping equally.

Additionally, we report the standard error of the MAE. This is determined by computing the MAE for each test file, and measuring the standard error across these independent measurements.

3.7.3 Root-Mean-Square Error

We measured the root-mean-square error (RMSE) in a similar manner to the MAE. The RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{\forall i} (y_i - \hat{y}_i)^2}, \quad (3)$$

where y_i is the target value for the i -th ping, \hat{y}_i is the predicted value generated by the model, and n is the number of pings to average over.

When measuring the overall RMSE on the test data, we first found the average MSE over the whole test set, then took the square root. Sample/file weighting, and standard error determination, was performed the same way as for the MAE.

3.7.4 Cumulative Error Distribution

We were particularly interested in how much human labour would be saved by the improvement in the annotations. There clearly must exist some error threshold below which errors in the annotation have no significant impact on downstream analysis and hence do not need to be fixed by the analyst. We speculate

that this error tolerance threshold may be at around 0.5 m to 1.0 m, since the lines are offset by 1.0 m before performing biological analyses to ensure all “bad data” is excluded. Hence we can crudely estimate what fraction of the model’s output needs to be adjusted by the analyst by considering what fraction of pings are within 0.5m or 1.0m of the target line.

Since we can not be sure what the appropriate error tolerance is — and the tolerable error threshold may vary depending on the application — we can evaluate the performance of the model over a range of potential tolerance values by considering the cumulative distribution of the absolute error. Such a plot shows the fraction of outputs which are within a certain absolute error threshold, and is similar to a receiver operating characteristic (ROC) curve. If we seek to optimize this curve without assigning any particular error tolerance threshold, we can consider the total area above the curve (the expected rejection rate over all error tolerance thresholds), which we seek to minimise. This area is precisely equal to the MAE metric.

3.7.5 Test Data Weighting

For the MP:mob \downarrow dataset, test recordings were taken from two held out transects — no recordings from these transects were presented during training. This allows us to evaluate the performance of the model at novel recording locations which the model has not seen before. Unfortunately, the protocol for annotating the seafloor was not consistent for the first 9 of the 17 MP:mob \downarrow surveys; hence we evaluated the seafloor line and overall IoU only on test data from the final 8 surveys. The entrained air, passive data collection, bad data periods, and bad data patches were evaluated on all 17 of the MP:mob \downarrow surveys.

For the MP:sta \uparrow and GP:sta \uparrow datasets, our target surface lines were generated with the Echoview surface line detector, followed by automated anomaly detection, as described in Section 3.2. However, in some cases the Echoview algorithm fatally failed to detect the water–surface boundary, placing the surface line impossibly close to the echosounder, or impossibly far away. When evaluating the surface line on the test set, we dropped recordings where the “target” surface line depth was outside the known range of low to high tide water depths for that recording site (Minas Passage: 28.5 m to 44 m; Grand Passage: 13.5m to 20 m). This allowed us to evaluate the Echofilter model’s surface line predictions against sane target values, but this selectively removed almost all the occasions where the Echoview algorithm’s predictions were wrong, severely compromising our ability to evaluate the performance of the Echoview algorithm at generating the surface line.

4 RESULTS

We measured the performance of our models using coarse-grained quantitative metrics (Section 4.1), and compared to the output of algorithms built into Echoview as a baseline. To further contextualise the level of performance attained by our models, we measured the level of agreement between expert annotators separately annotating the data (Section 4.2). We also evaluated the performance by detailed investigation with qualitative outputs (Section 4.3), and finally we evaluated the practical

output of the model by measuring the amount of time taken to audit and correct the model output (Section 4.4).

4.1 Quantitative Evaluation

We evaluated the overall performance of our final model by comparing the final “good data” mask produced by the model with that of the target labels. The target mask indicates which values within the echogram should be included in biological analyses. This mask excludes all values above the entrained-air line, below the seafloor line, during passive data collection regions, or marked as “bad data”. Our model produces outputs corresponding to each of these elements, and combining these outputs allows us to generate a final output mask. We measured the alignment between the two masks using the Intersection-over-Union (IoU), described in Section 3.7.1.

For our purposes, the most important output from the Echofilter model was the entrained-air line, which provides segmentation between the air entrained into the water column, and the rest of the water column. To provide a human-interpretable measurement of the error in the placement of this line, we measured the mean absolute error (MAE) and root-mean-square error (RMSE) between the depth of model’s entrained-air line and the target. See Section 3.7.2 and Section 3.7.3 for more details.

Other outputs from the model were evaluated similarly, using the IoU, MAE, and/or RMSE. In all cases, we show the performance of the model on the test partition, which was held out during all stages of model development and training.

4.1.1 Performance Break-Down Across Outputs

We investigated the performance of the final Upfacing (@400 epoch) and Bifacing (@700 epoch) models across all outputs produced by the network, and compared the quality of these outputs against the Echoview algorithm. The results were evaluated against the target annotations produced by a human expert, except for the surface line where the target was taken from the line produced by the Echoview algorithm but with anomalous values rejected (see Section 3.7.5).

We considered the IoU for each output, the results for which are shown in Table 2. The overall IoU compares the overall mask produced by removing pixels above the entrained-air line, below the seafloor line (if downfacing), during periods marked as passive data collection, and bad data annotations; we compare the mask obtained with the model against that from human annotation. For the entrained-air line, the IoU measurement considers the area beneath the entrained-air line — for upfacing recordings this extends to the echosounder, and for downfacing recordings it extends to the seafloor line provided by the expert’s annotation. Similarly the IoU measurement for the surface line extends from the surface to the echosounder, and is only measured for upfacing recordings. For the seafloor line, we compare the area from the seafloor line to the echosounder. For passive data region annotations, we compare the set of pings identified as passive by the model with the target annotations, performing a 1-D IoU calculation. The vertical bad data periods are measured in the same way as the passive data region annotations, using

TABLE 2 | Final model performance (agreement with manual annotation) for each output.

Output	MP:mob↓		MP:sta↑			GP:sta↑		
	Echoview	Echofilter Bifacing	Echoview	Echofilter Upfacing	Echofilter Bifacing	Echoview	Echofilter Upfacing	Echofilter Bifacing
Intersection-over-Union (%) ; larger is better								
Overall	96.80 ± 0.34	99.15 ± 0.08	90.41 ± 0.76	95.08 ± 0.34	94.91 ± 0.35	87.66 ± 1.05	92.10 ± 1.00	92.97 ± 1.00
Entrained-air	97.37 ± 0.31	99.11 ± 0.09	91.63 ± 0.72	96.05 ± 0.29	95.96 ± 0.28	89.06 ± 1.03	94.49 ± 0.50	94.95 ± 0.29
Surface	—	—	99.83 ± 0.05	99.82 ± 0.02	99.83 ± 0.02	—	98.59 ± 1.22	99.86 ± 0.01
Seafloor	99.33 ± 0.08	99.79 ± 0.03	—	—	—	—	—	—
Air–Seafloor	96.81 ± 0.34	99.16 ± 0.08	—	—	—	—	—	—
Passive	—	99.78 ± 0.06	—	100.0 ± 0.00	100.0 ± 0.00	—	99.97 ± 0.01	100.00 ± 0.00
Bad data period	—	—	—	40.58 ± 7.64	38.92 ± 7.42	—	24.68 ± 7.65	25.78 ± 8.07
Patch (anomaly)	—	0.00 ± 0.00	—	0.30 ± 0.12	0.30 ± 0.11	—	0.20 ± 0.07	0.20 ± 0.07
Mean Absolute Error (m; smaller is better)								
Entrained-air	1.178 ± 0.295	0.325 ± 0.031	2.187 ± 0.147	0.981 ± 0.044	1.005 ± 0.045	1.252 ± 0.198	0.577 ± 0.074	0.532 ± 0.031
Surface	—	—	0.062 ± 0.018	0.063 ± 0.007	0.062 ± 0.007	—	0.235 ± 0.232	0.024 ± 0.002
Seafloor	0.279 ± 0.032	0.089 ± 0.012	—	—	—	—	—	—
Root-Mean-Square Error (m; smaller is better)								
Entrained-air	6.436 ± 0.390	1.281 ± 0.064	4.275 ± 0.196	2.181 ± 0.085	2.228 ± 0.088	2.995 ± 0.508	1.244 ± 0.137	1.104 ± 0.060
Surface	—	—	1.323 ± 0.140	0.149 ± 0.020	0.134 ± 0.019	—	3.019 ± 1.073	0.035 ± 0.003
Seafloor	2.017 ± 0.187	0.292 ± 0.024	—	—	—	—	—	—
Proportion of pings where line placed within 0.5 m of target (%) ; larger is better								
Entrained-air	71.24 ± 1.60	88.30 ± 0.82	52.67 ± 2.09	61.28 ± 1.26	61.11 ± 1.27	61.18 ± 2.58	69.85 ± 2.54	70.17 ± 1.47
Surface	—	—	99.39 ± 0.15	99.67 ± 0.27	99.68 ± 0.26	—	99.44 ± 0.57	99.98 ± 0.01
Seafloor	89.27 ± 1.02	97.26 ± 0.70	—	—	—	—	—	—
Proportion of pings where line placed within 1.0 m of target (%) ; larger is better								
Entrained-air	80.20 ± 1.33	93.09 ± 0.59	58.34 ± 1.99	74.76 ± 0.97	74.33 ± 1.01	69.10 ± 2.70	83.70 ± 1.84	84.93 ± 0.92
Surface	—	—	99.44 ± 0.15	99.87 ± 0.15	99.88 ± 0.15	—	99.46 ± 0.57	100.00 ± 0.00
Seafloor	95.30 ± 0.49	98.75 ± 0.53	—	—	—	—	—	—
Proportion of pings where line placed within 2.0 m of target (%) ; larger is better								
Entrained-air	87.58 ± 1.01	96.46 ± 0.38	68.08 ± 1.80	86.50 ± 0.66	86.12 ± 0.69	80.61 ± 2.26	93.33 ± 1.17	94.20 ± 0.45
Surface	—	—	99.52 ± 0.15	99.93 ± 0.11	99.94 ± 0.10	—	99.46 ± 0.57	100.00 ± 0.00
Seafloor	98.48 ± 0.18	99.78 ± 0.16	—	—	—	—	—	—

The performance of the final Upfacing (@400 epoch) and Bifacing (@700 epoch) models, with thresholded zoom+repeat, merging/ignoring small output regions; compared against the performance of the Echoview algorithm as a baseline. Bold: best model. Italic: no significant difference from best (two-sided Wilcoxon signed-rank test, $p>0.05$).

a 1-D IoU. The IoU for the bad data patches is a comparison of the area marked as bad data by the model with a target mask indicating the locations of bad data patches.

We found the entrained-air and seafloor boundaries produced by both models had statistically significantly higher agreement with the human annotation than the lines produced by the Echoview algorithm (two-sided Wilcoxon signed-rank test, $p<0.05$). There was no significant difference between the outputs from the two models, except from the entrained-air line on MP:sta↑ where the magnitude of the difference in performance was small. Correspondingly, the quality of the overall output of the models were not significantly different from each other, but were significantly better than the Echoview algorithm.

The passive region annotations are highly accurate, reaching 100% accuracy on MP:sta↑ and GP:sta↑. On the MP:mob↓ dataset, the Bifacing model attains an IoU of 99.78%.

The bad data period annotations were challenging for the model to replicate, attaining an IoU of only 40% on MP:sta↑ and 25% on GP:sta↑. The anomalous bad data patches were impossible for the network to learn with any meaningful reliability, with an IoU of $\leq 0.3\%$. The poor performance of both of these annotations yields an increase in performance when small outputs are ignored (as seen in Section 4.1.2). On GP:sta↑, the bad data period annotations are sufficiently poor to yield an increase in performance when they are dropped entirely (see Section 4.1.2).

We measured the mean absolute error (MAE) of the entrained-air, surface, and seafloor lines (described in Section 3.7.2). As shown in Table 2, we found that the Bifacing model placed the entrained-air line on MP:mob↓ with only 0.325 m average error — a significant reduction (two-sided Wilcoxon signed-rank test, $p<0.05$) over the Echoview algorithm, which

had over three times as much error on average. For the MP:sta↑ and GP:sta↑ datasets, the two Echofilter models had comparable performance, a significant reduction in error against the Echoview algorithm baseline which had more than twice the error of Echofilter on both the upfacing, stationary datasets. Our findings when evaluating the entrained-air lines using the RMSE metric, and the proportion of pings within fixed distances of the target lines, were the same as with MAE.

The larger absolute error on the GP:sta↑ dataset (0.53 m to 0.58 m) and MP:sta↑ dataset (1.0 m) is indicative of the increased difficulty intrinsic to these datasets collected at sites where

the persistence, depth-of-penetration, variability of depth-of-penetration, and proportion of water column contaminated by entrained air exceeded that typically found at the transect locations sampled by the MP:mob↓ surveys. In particular, we note that the average and standard deviation of the depth-of-penetration for the entrained air was (12.4 ± 4.8) m for MP:sta↑, (6.7 ± 2.1) m for GP:sta↑, but only (2.4 ± 2.2) m for MP:mob↓. This corresponds to $(34 \pm 13)\%$ of the water column for MP:sta↑, $(41 \pm 12)\%$ for GP:sta↑, but only $(6.2 \pm 5.3)\%$ for MP:mob↓.

As shown in **Figures 3A–C**, the cumulative error distribution curve produced by Echoview is dominated by Echofilter for all

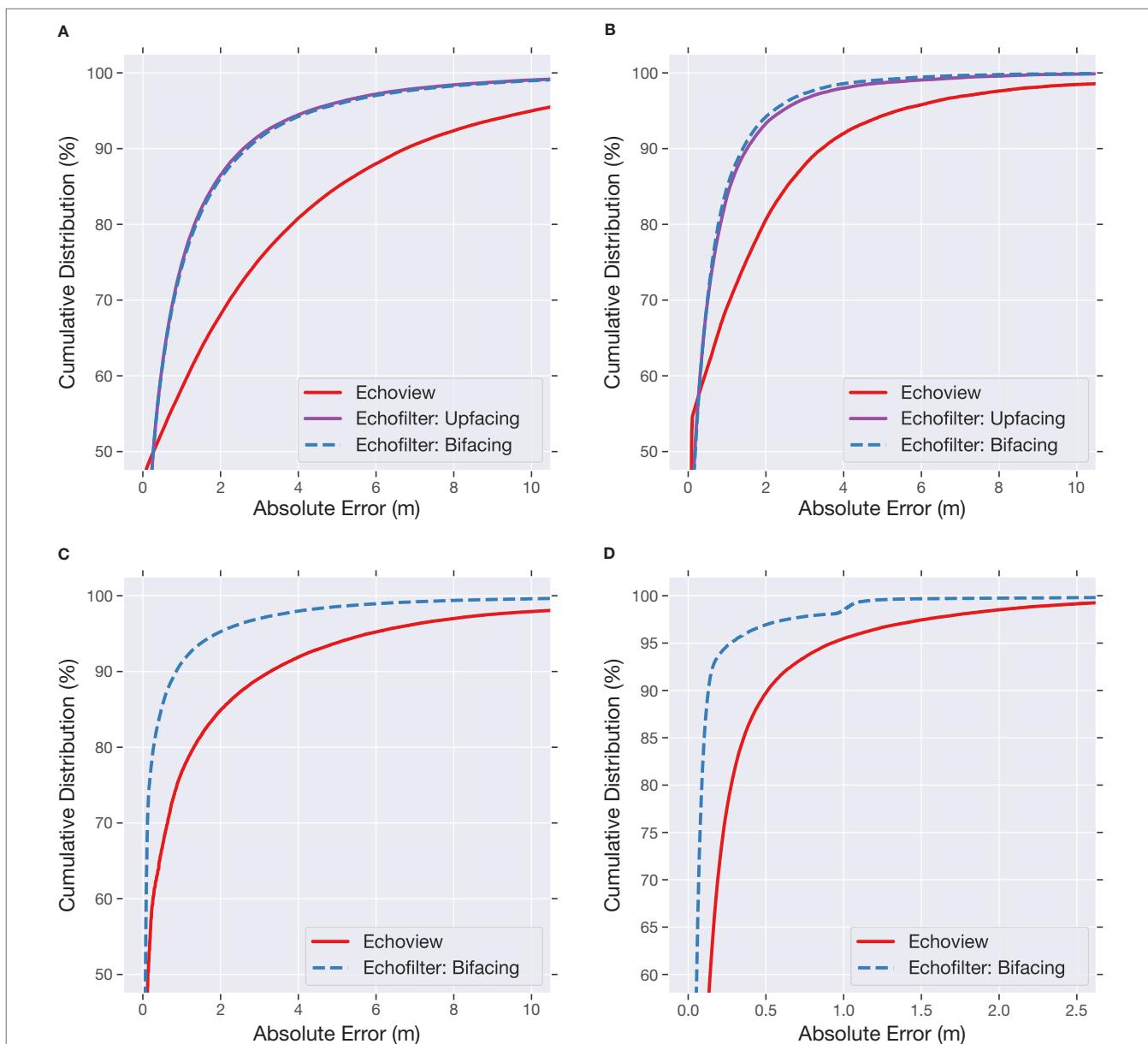


FIGURE 3 | Cumulative distribution for the absolute error of entrained-air (**A–C**) and seafloor (**D**) lines generated by the models (Echoview: red; Upfacing@400ep: magenta; Bifacing@700ep: blue). This indicates (on the y-axis) the fraction of pings where the output line was within a given threshold distance (x-axis) of the target line. **(A)** Error in entrained-air line on MP:sta↑. **(B)** Error in entrained-air line on GP:sta↑. **(C)** Error in entrained-air line on MP:mob↓. **(D)** Error in seafloor line on MP:mob↓.

values in the range of interest. Echoview appears to outperform Echofilter when using very narrow error thresholds (error < 0.2 m; the easiest 50% of the data), however this is an artifact of the manual data annotation process, in which Echoview was used to generate initial annotations which were then corrected as needed by a human expert.

As shown in **Table 2**, we found that the Bifacing model placed the seafloor line with very low error (only 0.09 m on average) on the MP:mob \downarrow test set. Again, this was significantly lower than the Echoview algorithm, for which the MAE was three times higher (0.28 m). Similar results were seen with the RMSE and proportion of pings within a tolerance threshold. As shown in **Figure 3**, the Echofilter cumulative error distribution for the seafloor line dominates Echoview. We note there is a step jump in Echofilter model performance at 1.0 m error; this is caused by inconsistencies in the training data annotation in the first 9 surveys which impacted the model fit (evaluated only on the last 8 surveys).

For the surface line annotation, we find that the Echofilter models have a MAE comparable to Echoview on MP:sta \uparrow and outperform Echoview when considering the RMSE and fraction of pings within 0.5 m to 2.0 m error threshold. This is because the Echofilter models do not produce the anomalous surface line depths seen with the Echoview line generation, which we removed from our training and target lines. On GP:sta \uparrow , the Bifacing model produced better surface lines than the Upfacing model. Manual inspection of the results demonstrates that the Upfacing model is sometimes confused by reflections in the additional range of these recordings (following the erroneous training targets generated by Echoview), whilst the Bifacing model was not confused by these reflections.

4.1.2 Impact of Post-Processing Steps

We evaluated the performance of the final Upfacing and Bifacing models before and after each post-processing step impact described in Section 3.6. Our results are shown in **Table 3**.

Compared with applying the model only once on the full echogram, using the two step “zoom+repeat” stimulus

presentation provided a statistically significant increase in the entrained-air line placement as evaluated by the MAE (two-sided Wilcoxon signed-rank test, $p < 0.05$) on all datasets, and for both the Upfacing and Bifacing model. The overall IoU also significantly increased, except for the Bifacing model on MP:sta \uparrow where “zoom+repeat” caused a very small, but statistically significant, decrease.

We also considered using a threshold of 0.35 to determine when to do the zoom+repeat step, following our assumption that a second application of the model on a zoomed-in echogram is not necessary when less than 35% of the echogram data is outside the surface-seafloor extent. We found that using this threshold had no impact on the performance of the models on the MP:mob \downarrow and GP:sta \uparrow datasets, where the range of the data extended far outside the surface-seafloor extent and hence zoom+repeat was *de facto* always applied. On the MP:sta \uparrow dataset, where the recording range was not much further than the distance from seafloor to sea surface, there was a significant decrease in performance when a threshold was used to determine when to apply a second round of the model. This suggests that the zoom+repeat protocol should always be used in order to yield the best annotation with the model. Nonetheless, the rest of our results present in this paper use the (faster to perform) thresholded zoom+repeat, with a threshold of 0.35.

The remaining optional post-processing steps were considered with thresholded zoom+repeat in place. We found no significant differences in the overall IoU when small regions were merged together or dropped from the output (changing the way regions are handled has no effect on the entrained-air line placement). Omitting bad data regions and patches entirely had a positive impact on the overall performance on the GP:sta \uparrow data, but a negative impact on MP:sta \uparrow data. This was because the bad data period predictions (as seen in **Table 2**) were notably worse on GP:sta \uparrow than MP:sta \uparrow . There was no impact on MP:mob \downarrow data because the model did not predict any bad data regions on this test dataset.

TABLE 3 | Impact of post-processing steps on the model performance metrics.

Model	Overall IoU (%)			Entrained-air MAE (m)		
	MP:mob \downarrow	MP:sta \uparrow	GP:sta \uparrow	MP:mob \downarrow	MP:sta \uparrow	GP:sta \uparrow
Echoview algorithm	96.80 \pm 0.34	90.41 \pm 0.76	87.66 \pm 1.05	1.178 \pm 0.295	2.187 \pm 0.147	1.252 \pm 0.198
Upfacing w/o zoom	–	95.06 \pm 0.34	88.06 \pm 3.88	–	0.987 \pm 0.045	0.629 \pm 0.076
w/zoom+repeat	–	95.11 \pm 0.35	92.09 \pm 1.01	–	0.950 \pm 0.041	0.574 \pm 0.071
w/thresholded z+r	–	94.27 \pm 0.46	92.07 \pm 1.01	–	0.981 \pm 0.044	0.577 \pm 0.074
+ ignore small regions	–	95.08 \pm 0.34	92.10 \pm 1.00	–	0.981 \pm 0.044	0.577 \pm 0.074
+ ignore all “bad data”	–	94.77 \pm 0.44	93.01 \pm 0.76	–	0.981 \pm 0.044	0.577 \pm 0.074
+ logit smoothing	–	94.27 \pm 0.46	92.48 \pm 0.86	–	1.099 \pm 0.046	0.623 \pm 0.095
Bifacing w/o zoom	98.59 \pm 0.09	94.90 \pm 0.35	88.35 \pm 3.93	0.402 \pm 0.030	1.004 \pm 0.045	0.589 \pm 0.047
w/zoom+repeat	99.16 \pm 0.08	94.86 \pm 0.40	92.95 \pm 1.01	0.325 \pm 0.031	0.979 \pm 0.044	0.532 \pm 0.031
w/thresholded z+r	99.16 \pm 0.08	94.90 \pm 0.35	92.95 \pm 1.01	0.325 \pm 0.031	1.005 \pm 0.045	0.532 \pm 0.031
+ ignore small regions	99.15 \pm 0.08	94.91 \pm 0.35	92.97 \pm 1.00	0.325 \pm 0.031	1.005 \pm 0.045	0.532 \pm 0.031
+ ignore all “bad data”	99.15 \pm 0.08	94.74 \pm 0.42	93.45 \pm 0.64	0.325 \pm 0.031	1.005 \pm 0.045	0.532 \pm 0.031
+ logit smoothing	98.90 \pm 0.08	94.35 \pm 0.43	93.12 \pm 0.67	0.385 \pm 0.030	1.103 \pm 0.051	0.570 \pm 0.047

Bold: best pre-processing option. *Italic:* no significant difference from best (two-sided Wilcoxon signed-rank test, $p > 0.05$).

We considered the effect of logit smoothing on the model's final output by applying this postprocessing step, in addition to thresholded zoom+repeat and ignoring all bad data annotations, with a Gaussian kernel size of 1. We found that logit smoothing had a significant negative impact on the accuracy of the entrained-air line placement, and on the overall mask output, for all datasets.

4.1.3 Impact of Model Training Duration

We investigated the impact of training time on the final model outputs. We compared the output of each of the models at the end of each stage of the cyclic training process. For this analysis, we used thresholded zoom+repeat, and merged/ignored small regions in the model output.

As shown in **Table 4**, we found that further training cycles improved the performance on MP:sta↑ and MP:mob↓, though with diminishing returns. Additional training *reduced* the performance on GP:sta↑, but the reduction was not statistically significant.

4.2 Inter-Annotator Agreement Benchmarking

The extent to which air is entrained in the water column is not observed directly, and can only be estimated based on the echosounder recordings. With training and experience, human annotators can learn which datapoints correspond to entrained air and which to fish populations within the water column. However, without a ground truth measurement, the annotations are subjective and will differ between annotators.

With this in mind, it is difficult to know how well we could expect an ideal model to perform at the task. It is infeasible to

expect perfect agreement between the model and the human annotations, since human annotators do not always agree amongst each other and are not necessarily consistent in their choice of line placement. We endeavoured to quantify how well our model performs by measuring the agreement between two human annotators, which acts as a baseline to estimate the Bayes error rate.

We selected 10 EV files from the Grand Passage stationary-upfacing dataset (GP:sta↑), ensuring that the selected files were composed of sufficiently complex data so that any differences in line placement between each annotator would be highlighted. Annotations were generated by Echoview using the preexisting workflow. The Echoview annotations were edited independently by JD and LPM in order to create two sets of finalized annotations for all 10 files. We then created annotations using Echofilter (models Upfacing@400ep and Bifacing@700ep, using thresholded zoom+repeat, and dropping small regions). While both annotators are experts in this field, JD was the most experienced at handling this data — her annotations constituted the majority of the annotations used to train the models. Consequently, we treated JD's annotations as the ground truth labels, and measured the performance of the other annotation methods in comparison to her labels.

As shown in **Table 5**, we found that the level of agreement in placement of the entrained-air line between Echofilter and JD exceeded that of LPM, with higher IoU and a smaller average distance between the line depths, though the difference was not statistically significant ($p>0.05$). This suggests our model outputs have an accuracy comparable to human-level performance at this task.

TABLE 4 | Performance of models after each training cycle (different total training durations).

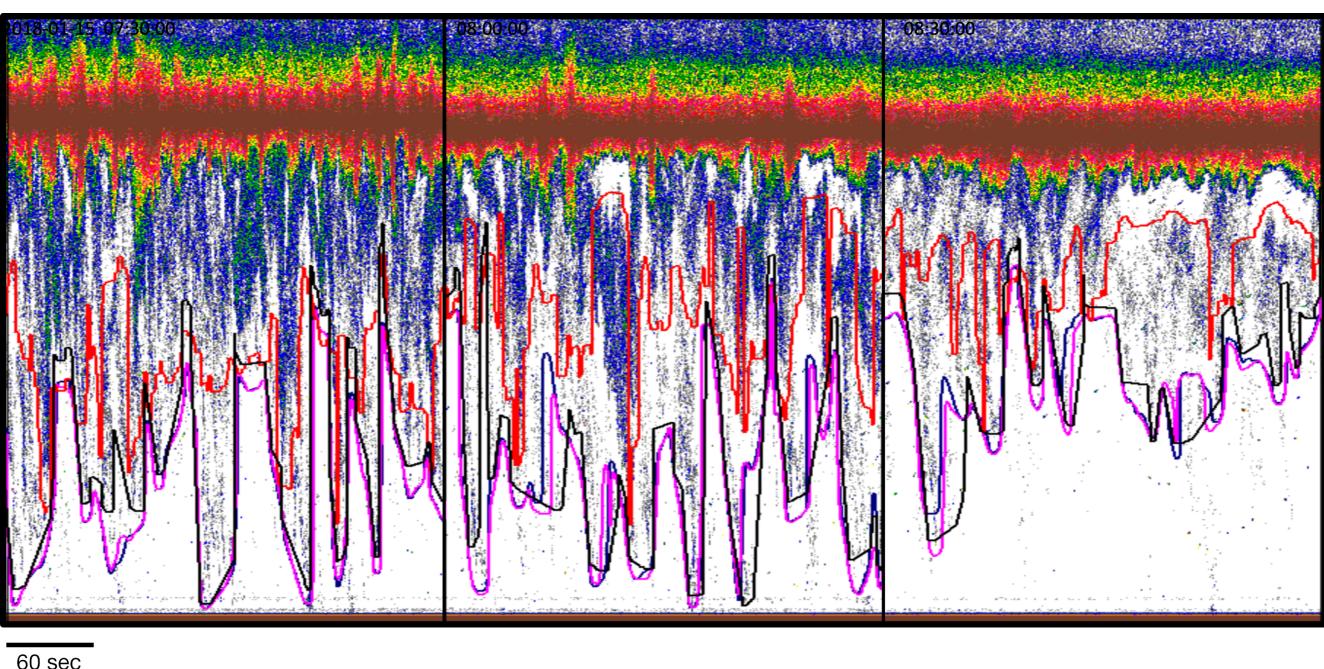
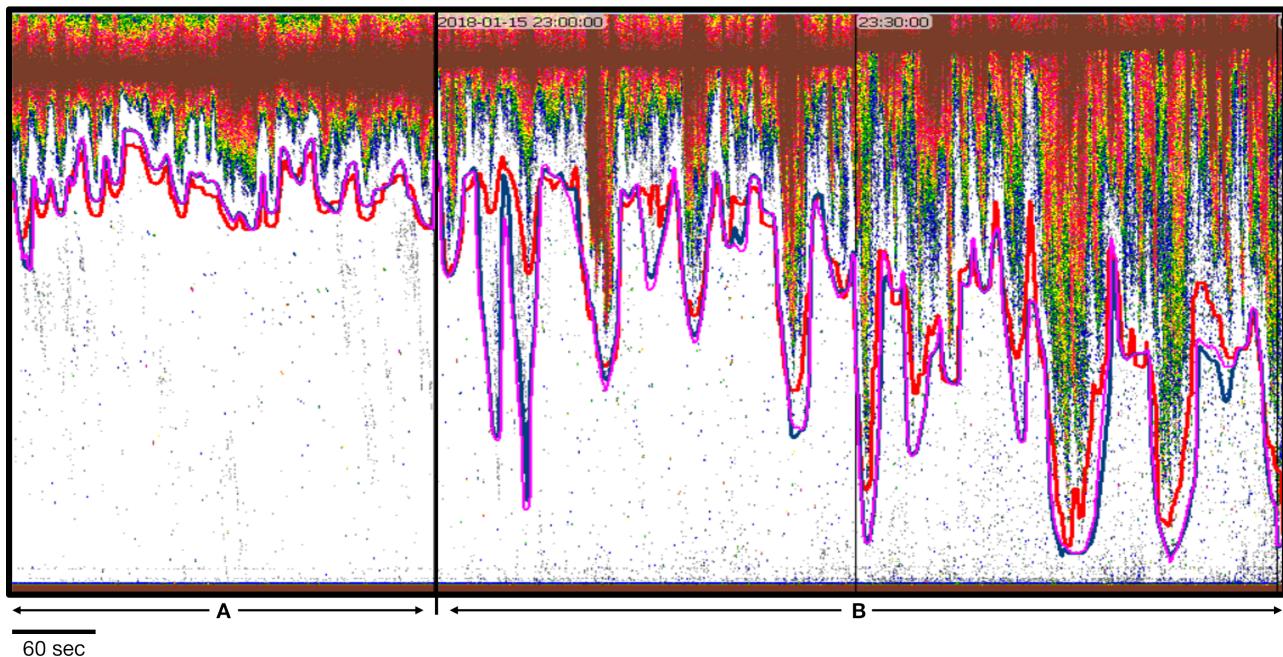
Model	Overall IoU (%)			Trained-air MAE (m)		
	MP:mob↓	MP:sta↑	GP:sta↑	MP:mob↓	MP:sta↑	GP:sta↑
Upfacing 100ep	—	95.05 ± 0.33	93.32 ± 0.87	—	1.001 ± 0.046	0.520 ± 0.033
400ep	—	95.08 ± 0.34	92.10 ± 1.00	—	0.981 ± 0.044	0.577 ± 0.074
Bifacing 100ep	98.93 ± 0.10	94.93 ± 0.32	93.52 ± 0.69	0.369 ± 0.034	1.036 ± 0.047	0.513 ± 0.032
400ep	99.02 ± 0.09	94.97 ± 0.33	93.18 ± 0.92	0.329 ± 0.028	1.022 ± 0.047	0.520 ± 0.032
700ep	99.15 ± 0.08	94.91 ± 0.35	92.97 ± 1.00	0.325 ± 0.031	1.005 ± 0.045	0.532 ± 0.031

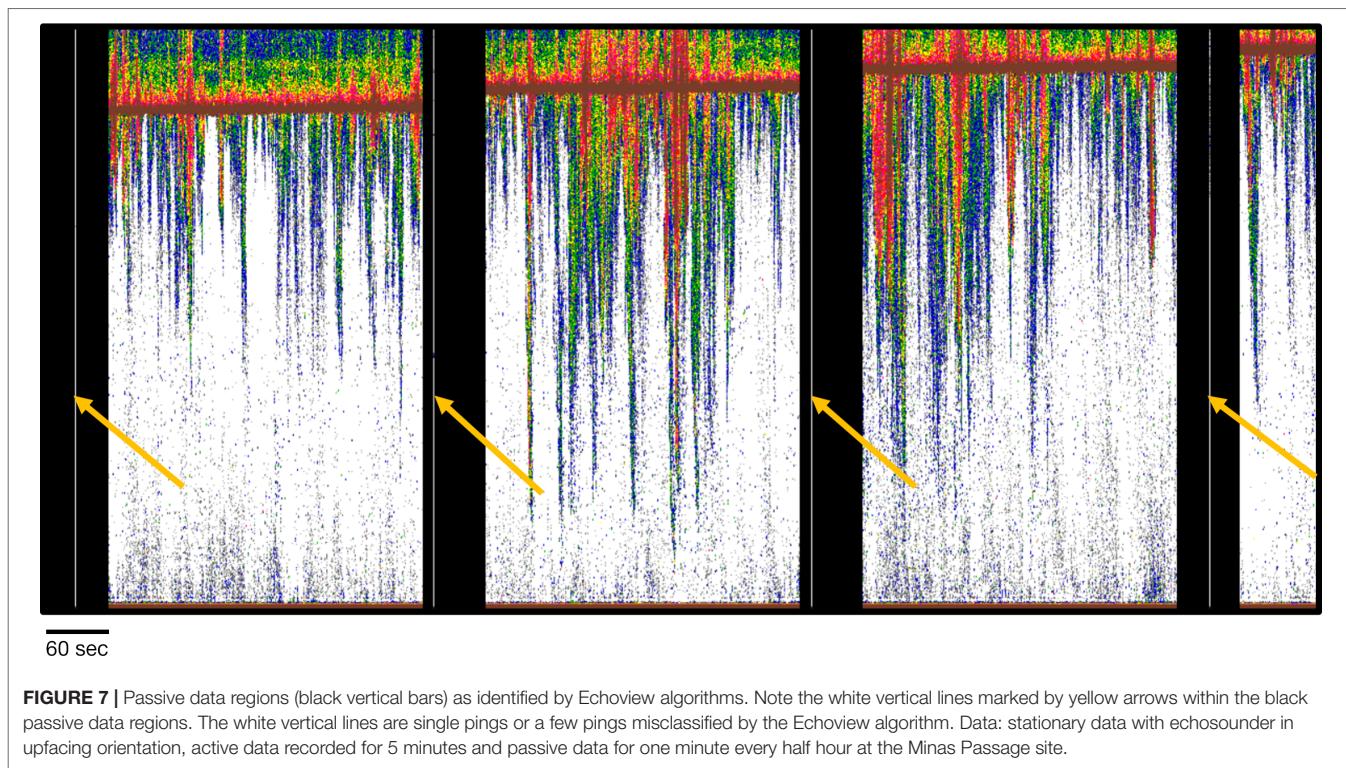
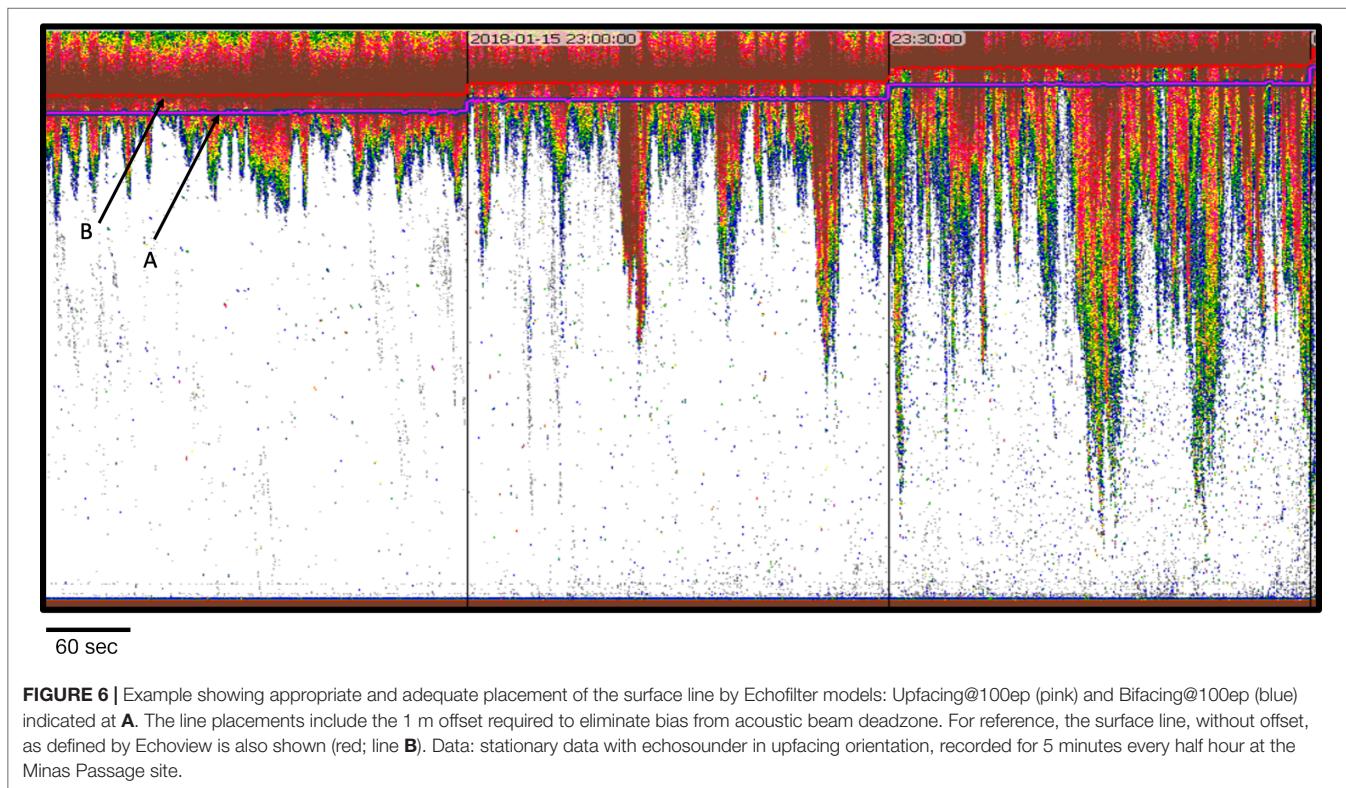
Bold: best training duration. *Italic:* no significant difference from best (two-sided Wilcoxon signed-rank test, $p>0.05$).

TABLE 5 | Comparison of agreement between several annotation sources.

Annotator	IoU (%; larger is better)				Δ Trained-air (m)	
	Overall	Entrained-air	Bad data period	Patch	MAE	RMSE
Human expert (LPM)	90.7 ± 1.2	92.4 ± 0.9	97.8 ± 25.8	0.29 ± 0.10	0.86 ± 0.10	1.63 ± 0.17
Echoview	88.7 ± 1.4	90.9 ± 1.1	—	—	1.05 ± 0.12	2.01 ± 0.19
Echofilter: Upfacing	90.5 ± 2.9	93.2 ± 0.9	70.7 ± 24.0	0.32 ± 0.09	0.76 ± 0.05	1.25 ± 0.11
Echofilter: Bifacing	91.3 ± 1.3	93.0 ± 1.0	92.8 ± 25.5	0.38 ± 0.12	0.78 ± 0.04	1.27 ± 0.09

We compared several annotation methods against expert labels created by JD. The intersection-over-union (IoU) across all recordings is shown, in addition to the mean absolute error (MAE) and root-mean-square error (RMSE) for the placement of the entrained-air separation line ($n = 10$, \pm inter-recording standard error). Note that JD used Echoview to generate seed annotations for refinement into finalized annotations. *Bold:* model with best agreement with target (JD) annotations. *Italic:* no significant difference from best (two-sided Wilcoxon signed-rank test, $p>0.05$).





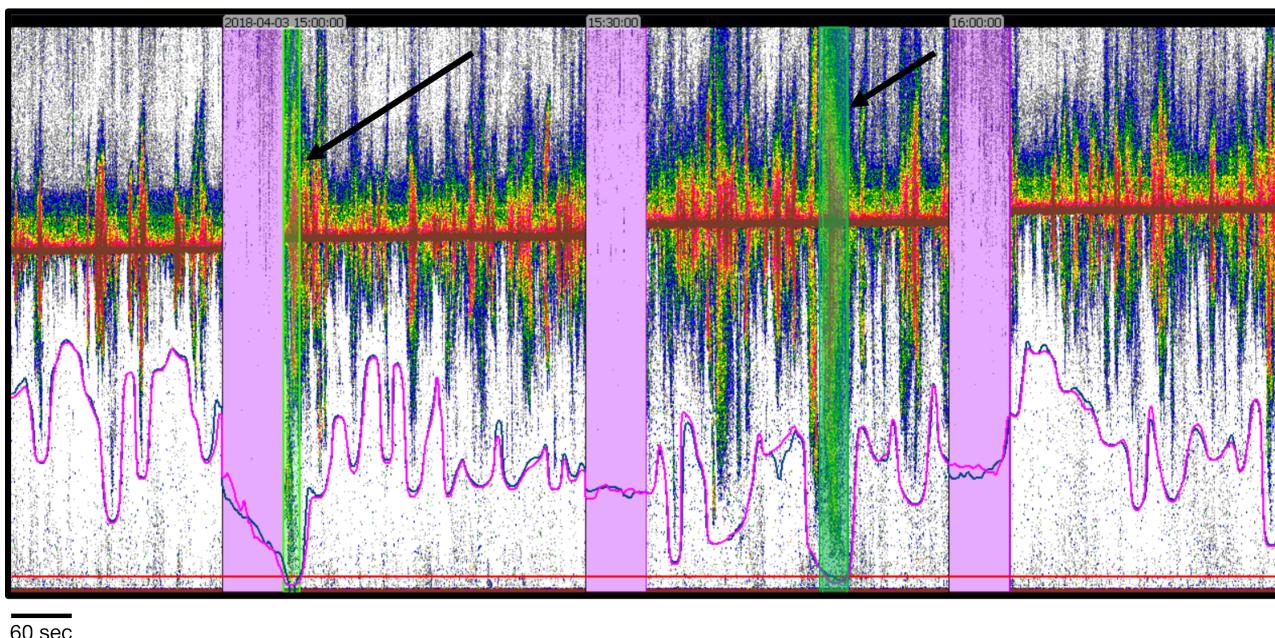


FIGURE 8 | Example of passive data regions (pink bars) and bad data periods (green bars) defined by Echofilter. Echofilter successfully identified the contiguous pings of passive data collection. Bad data periods designated by Echofilter were pings where the entitled-air line (pink or blue lines in this example) intersects with the bottom analytical line (in this case the line designating the outer boundary of the transducer nearfield). Although these cases meet the criteria, in each case the hydroacoustic analyst would designate the entire recording period as a bad data region due to the strength, penetration, and persistence of entitled-air signals. Data: stationary data with echosounder in upfacing orientation, active data recorded for 5 minutes and passive data for one minute every half hour at the Minas Passage site.

4.3 Manual Evaluation of Model Outputs

Manual investigation of the Echofilter results were carried out by JD and LPM on a Windows 10 operating system, using Echoview 10, or Echoview 11 newly released at the time of testing. The performance of Echofilter was evaluated on 24 Echoview files, selected from the test partition as described in Section 2.

During model development, a series of iterative testing and upgrades to Echofilter was undertaken. Echofilter was run on the entire set of test files, applying both models (Bifacing and Upfacing, with thresholded zoom+repeat, and logit-smoothing) to the data for comparative purposes. The results were examined for adequacy and appropriateness of the placement of lines (sea surface and entitled-air), the identification of the passive data collection periods and identification of bad data regions. Issues with the outputs were investigated in detail, and used to make changes to the model architecture design, training paradigm, or to the format of input and target data provided to the model during training. This process was iterated until any additional improvements were marginal and inconsequential.

By the end of testing and upgrades to the models, both models (Bifacing and Upfacing) produced appropriate automated initial placement of the boundary lines. Most importantly, the model placement of entitled-air boundary lines were visibly superior to the line placements as produced by the Echoview algorithms, as shown in Figure 4. The model results proved to be much more responsive than the Echoview algorithms to the entitled-air ambient characteristics across the varying tidal flow

rates (Figure 4). In some cases, the automated prediction of the entitled-air line placement as produced by Echofilter were far superior to that produced by Echoview; see Figure 5. Note that Echofilter entitled-air lines as defined by each model (Bifacing and Upfacing) were essentially equivalent, although not identical.

As shown in Figure 6, the Echofilter models produced appropriate and adequate automated placement of the surface line, including a user-defined offset; in this case 1 m. Likewise, the Echofilter models produced appropriate and adequate identification of the passive data regions that will be excluded from biological analyses. We found the Echofilter passive data region identification was superior to the Echoview algorithms implemented to automate the identification of passive data regions. The Echoview algorithms would, not uncommonly, exclude a ping or few pings from within the passive data region, thereby inappropriately designating those pings for inclusion in biological analyses, as shown in Figure 7. No such occurrences were noted in the Echofilter results (e.g. Figure 8).

In addition to the passive data regions, there are two additional types of bad data regions that are not uncommon to echosounder data. The first type, is a contiguous time period marked to be removed from analysis. As shown in Figure 8, these bad data regions are identified by Echofilter when the position of the entitled-air line resolves to a position intersecting or extending below the bottom line, whether that line is the seafloor or the line designating the transducer nearfield exclusion line. In other words, when the position of the entitled-air line indicates that

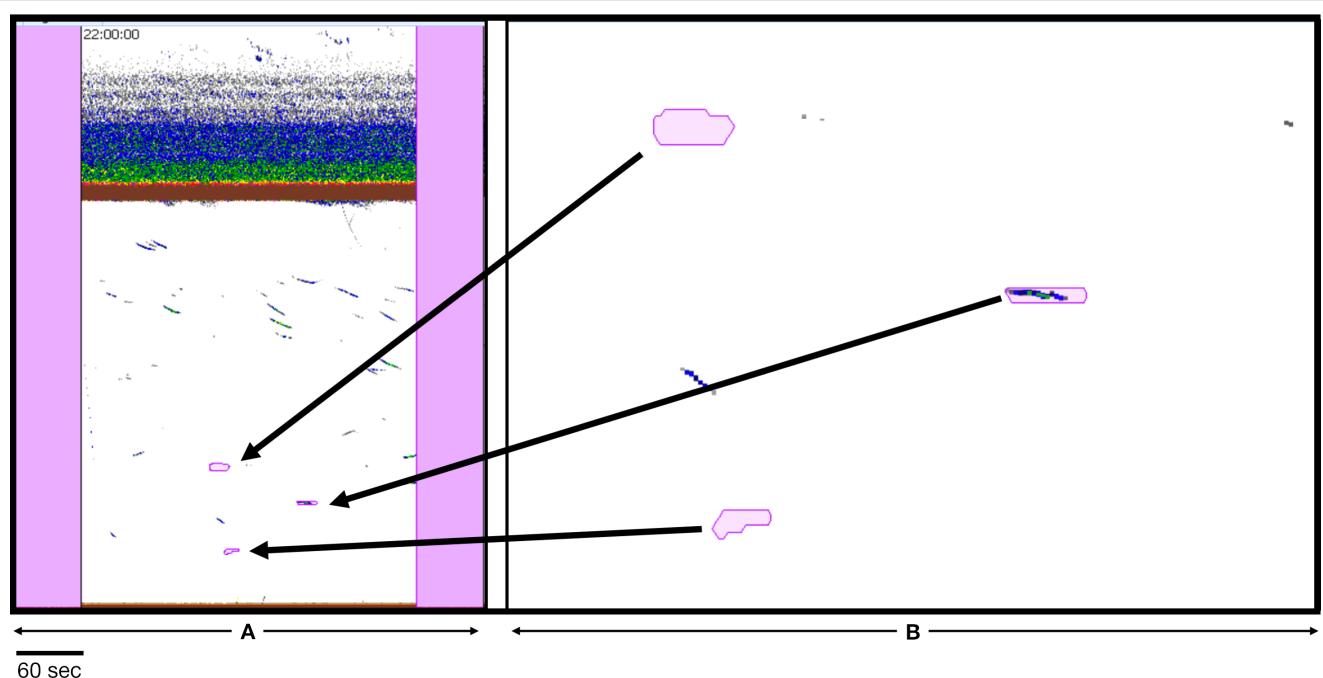


FIGURE 9 | Example of false positive “patch” bad data regions identified by Echofilter. **(A)** A 5-minute section of echogram with passive data regions (pink rectangles) on either side. **(B)** Enlargement to show the contents within each patch. Empty patches are false positive. The patch containing color samples within it would be classified as fish by the hydroacoustic analyst. It was likely identified as a bad data region by Echofilter because of its nearly horizontal position. The data on which the models were trained contain occurrences of unidentified interference which appear as horizontal lines. Those were classified as bad data regions by the analyst prior to training. Both models (Bifacing@100ep and Upfacing@100ep) designate true and false positives, but differently. Bifacing@100ep results appear to include fewer false positives. Data: stationary data with echosounder in upfacing orientation, active data recorded for 5 minutes and passive data for one minute at Minas Passage Site.

the entrained air has penetrated the entire depth of the water column. Such occurrences are not uncommon in the Minas Passage and Grand Passage datasets, sometimes occurring for just a few pings and other times the penetration occurs throughout an entire 5-minute data collection period. The single criteria of intersecting or penetrating below the bottom line is insufficient for defining all pings that should be excluded in their entirety. Figure 8 provides an example of just such a case: less than 50% of the water column remains after the entrained-air exclusion. In that case, if the goal of the analyses is to understand metrics

within the full water column, that data collection period would need to be excluded in its entirety.

The second type of bad data region, a “patch” of bad data, can be characterized as forming randomly shaped discrete patches. Within the original test segment of 24 files, only three had occurrences of the patch-type bad data region. Two additional EV files containing patch-type bad data regions were identified from the validation and training segments for manual inspection of the patch-type results only. Both Echofilter models performed poorly, generating false positives as illustrated in Figure 9.

4.4 Time-Savings Analysis

We sought to evaluate the amount of time-savings that the Echofilter model would offer, relative to the existing workflow using Echoview algorithms. Five of the Echoview files from the MP:sta↑ test partition were selected for a time test. The files were selected to represent each tide and phase combination: flooding spring tide, ebbing spring tide, flooding neap tide, and ebbing neap tide, plus one file with especially noisy data for which neither Echoview or Echofilter would likely render a well-placed entrained-air line. Annotations were initialized twice: once using the preexisting workflow utilizing Echoview algorithms, and once using Echofilter with the Upfacing@100ep model, with logit-smoothing enabled. The initial entrained-air line in each of the ten files was audited and edited by the hydroacoustic analyst

TABLE 6 | Results from the time-to-edit experiment.

Tide	Phase	Edit order		Edit time (MM:SS)		
		Echoview	Echofilter	Echoview	Echofilter	Reduction
Spring	Flood	2	6	8:06	4:04	50%
Spring	Ebb	3	4	8:00	4:04	49%
Neap	Flood	8	7	8:18	4:30	46%
Neap	Ebb	1	5	7:08	3:57	45%
Bad file		9	10	4:28	1:51	59%
Overall Mean				7:12	3.42	49%

A hydroacoustic analyst used the entrained-air lines produced by either Echoview or Echofilter to seed their annotations. We compare the amount of time needed to convert the seed lines into “correct” annotation lines. **Bold:** best model (shortest duration).

(JD), while recording the amount of time taken to do so. We randomized the order in which tasks (file and seed annotation source) were completed, except the especially noisy “bad file” which was evaluated later.

Our results, shown in **Table 6**, demonstrate that using the annotations generated by Echofilter results in less time taken for the human annotator to complete their task. For typical data files, the time taken to finalize annotations was consistently 45%–50% shorter when using annotations produced by Echofilter as the seed instead of annotations produced by Echoview. For an especially noisy file, the reduction in time was even larger, at 59%. The reduction in time was statistically significant ($p < 0.001$; paired Student’s *t*-test).

5 DISCUSSION

5.1 Impact of Echofilter Model

We have described the implementation of a deep learning model, Echofilter, which can be used to generate annotations to segment entrained air appearing in hydroacoustic recordings at tidal energy sites. Our goal was to produce an automated, model-based approach to the placement of a line appropriately defining the boundary between that portion of the water column contaminated by acoustic returns from entrained air, and that portion of the water column appropriate for biological analyses. This was motivated by the need for reliable, timely analyses and subsequent reporting to assist regulators, developers, and stakeholders in understanding the risks to fish imposed by the deployment of tidal energy devices into marine ecosystems.

We found the deep learning models we implemented produced significantly and appreciably better placement of the entrained-air line than the Echoview algorithms. For mobile, downward-facing recordings, the average error was 0.33 m, less than a third of Echoview’s 1.2 m average error. For stationary, upward-facing recordings, the average error was 0.5 m to 1.0 m depending on dataset, consistently less than half the error seen with Echoview algorithms (1.2 m to 2.2 m). Furthermore, the surface, seafloor, and passive region placement were also superior to those produced using Echoview. The model’s overall annotations had a high level of agreement with the human segmentation, with an intersection-over-union score of 99% for mobile downfacing recordings and 92% to 95% for stationary upfacing recordings. As such, Echofilter provides a complete automated line placement and passive data identification methodology.

The most challenging segmentation line to place correctly is the entrained-air line, which currently can require time-consuming manual placement due to the lack of a well-placed automated solution. We found that the increase in accuracy of the automated placement of the entrained-air line provided by Echofilter corresponded to a 50% reduction in the time required for a hydroacoustician to audit and correct the line placement. Our quantitative analysis has shown that the Echofilter models produce lines which are closer (in distance) to the line placement defined by the human expert. Additionally, we note that the

ML models are more sensitive to the fine-scale nuances in the boundary position of the entrained air; when the model places the line incorrectly, the errors tend follow the correct shape of the entrained air but are offset by some amount, and hence require only a simple, coarse edit to shift the line in some region to the correct offset. In contrast, when the Echoview algorithm is incorrect, the shape is incorrect and corrections to the line involve time-consuming fine-scale edits instead. Since coarse-scale edits are less cognitively taxing and far fewer edits are required, far less analyst fatigue is invoked during manual corrections of the model-placed entrained-air line, thereby allowing the analyst to bring the full-force of their intellect, training, and analytical skills to modifying placement of the line segments as necessary. Additionally, the reduction in the number of fine-scale edits provides the opportunity for an increase in the standardization and repeatability of line placement, within an analyst’s work and among analysts.

Machine learning applied to the hydroacoustic data by which we quantify fish distribution and abundance has garnered improvements to the work flow and increased the efficiency of the work by 50%, improvements that haven’t been achieved any other way. The machine learning contribution to assessing the ecological impacts of introducing marine renewable energy devices into the marine habitat is the improved analytical consistency and substantial improvements in the timeliness of analyses and subsequent reporting.

5.2 Limitations Associated With Echofilter

We developed Echofilter with the goal of increasing the efficiency and standardization of the post-processing of hydroacoustic data collected in dynamic marine environments such as tidal channels. The model was thoroughly evaluated on data recorded from upward-facing stationary echosounders at two tidal energy demonstration sites in the Bay of Fundy. The models have not been evaluated on data collected in other regions, with other instrumentation, or in other deployment configurations. Consequently, the performance of Echofilter on data collected under conditions that differ substantially from those used for model development may be heavily impacted and require some level of re-training to ensure accurate results, which is a non-trivial procedure.

In addition to the entrained-air boundary line, Echofilter predicts the depths of the surface (for upfacing recordings) and the seafloor (downfacing). Our performance metrics indicate that these lines are all placed accurately, however we have not thoroughly inspected the model’s output on downfacing recordings and can not confirm the integrity of the seafloor line.

In addition to the lines, our model attempts to predict regions which should be excluded from biological analyses. However, it was not possible for the model to learn these annotations with sufficient accuracy to be usable for downstream tasks. Consequently, it is not possible to automate away a need for manual inspection of the data. A hydroacoustician must always inspect the recordings themselves in order to annotate regions to exclude from analysis, and adjust lines as necessary.

5.3 Accessing Echofilter

To ensure the broader community can utilize our model described in this paper, we have released the final implementation, Echofilter, under the AGPLv3 license. Python source code and a stand-alone Windows executable are available at <https://github.com/DeepSenseCA/echofilter>. Additionally, the command line interface (CLI) and application programming interface (API) documentation is available at <https://DeepSenseCA.github.io/echofilter/>.

We hope this tool will prove useful to tidal energy researchers, and the wider hydroacoustic community.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. Requests to access these datasets should be directed to FORCE, info@fundyforce.ca.

AUTHOR CONTRIBUTIONS

Manuscript written by SL and LM. Data conversion was performed by JN and SL. Model architecture design and training was performed by SL. Interface and API development by SL. Model evaluation was performed by LM and JD. Identification of features of specific value to the hydroacoustics community was performed by LM. Project oversight by DH, CW, and SO.

REFERENCES

- Bengio, Y., Lecun, Y. and Hinton, G. (2021). Deep Learning for AI. *Commun. ACM* 64, 58–65. doi: 10.1145/3448250
- Benoit-Bird, K. J. and Lawson, G. L. (2016). Ecological Insights From Pelagic Habitats Acquired Using Active Acoustics. *Annu. Rev. Mar. Sci.* 8, 463–490. doi: 10.1146/annurev-marine-122414-034001
- Blaber, S. J. M., Cyrus, D. P., Albaret, J.-J., Ching, C. V., Day, J. W., Elliott, M., et al. (2000). Effects of Fishing on the Structure and Functioning of Estuarine and Nearshore Ecosystems. *ICES J. Mar. Sci.* 57, 590–602. doi: 10.1006/jmsc.2000.0723
- Cada, G., Ahlgren, J., Bahleda, M., Bigford, T., Stavrakas, S. D., Hall, D., et al. (2007). Potential Impacts of Hydrokinetic and Wave Energy Conversion Technologies on Aquatic Environments. *Fisheries* 32, 174–181. doi: 10.1577/1548-8446(2007)32[174:PIOHAW]2.0.CO;2
- Copping, A. E., Hemery, L. G., Overhus, D. M., Garavelli, L., Freeman, M. C., Whiting, J. M., et al. (2020). Potential Environmental Effects of Marine Renewable Energy Development—the State of the Science. *J. Mar. Sci. Eng.* 8. doi: 10.3390/jmse8110879
- Cornett, A., Toupin, M. and Nistor, I. (2015). “Appraisal of the IEC Technical Specification for Tidal Energy Resource Assessment at Minas Passage, Bay of Fundy, Canada,” in *Proc. 2015 European Wave and Tidal Energy Conference (EWTEC)* (Nantes, France).
- DFO (2008). *Potential Impacts of, and Mitigation Strategies for, Small-Scale Tidal Generation Projects on Coastal Marine Ecosystems in the Bay of Fundy* (Tech. rep., Fisheries and Oceans Canada (DFO) Canadian Science Advisory Secretariat). Science Response 2008/013.
- DFO (2018). *Delineating Important Ecological Features of the Evangeline-Cape Blomidon-Minas Basin Ecologically and Biologically Significant Area (EBSA)* (Tech. rep., Fisheries and Oceans Canada (DFO) Canadian Science Advisory Secretariat). Science Response 2018/005.
- All authors contributed to the article and approved the submitted version.

FUNDING

SL was funded by Mitacs award IT16140. LPM was supported by the Offshore Energy Research Association of Nova Scotia (OERA) under the Pathway Program; funding provided by Natural Resources Canada, grant number ERPP-RA-05.

ACKNOWLEDGMENTS

This work was supported in part by Mitacs through the Mitacs Accelerate program, and by Offshore Energy Research Association (OERA; <https://oera.ca/>). Computations were performed on the DeepSense (<https://deepsense.ca/>) high-performance computing platform. DeepSense is funded by ACOA, the Province of Nova Scotia, the Centre for Ocean Ventures and Entrepreneurship (COVE), IBM Canada Ltd., and the Ocean Frontier Institute (OFI). We thank Echoview Software Pty Ltd. and Haley Viehman, Ph.D., for providing support and technical advice for this project. We thank Paul C. Hines of Hines Ocean S&T for his helpful discussions during preparation of the manuscript. Additionally, we thank Tyler Boucher, Aurélie Daroux, Mili Sanchez, and Haley Viehman, for (in addition to JD and LM) creating the expert annotations for the data used in this study. This work has previously appeared as an arXiv preprint (Lowe et al., 2022), available at doi: 10.48550/arxiv.2202.09648.

- Fernandes, P., Gerlotto, F., Holliday, D., Nakken, O. and Simmonds, E. (2002). Acoustic Applications in Fisheries Science: The ICES Contribution. *ICES Mar. Sci. Symp.* 215, 483–492. doi: 10.17895/ices.pub.8889
- Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning* (MIT Press). Available at: <http://deeplearningbook.org>
- Guerra, M., Hay, A. E., Karsten, R., Trowse, G. and Cheel, R. A. (2021). Turbulent Flow Mapping in a High-Flow Tidal Channel Using Mobile Acoustic Doppler Current Profilers. *Renewable Energy* 177, 759–772. doi: 10.1016/j.renene.2021.05.133
- He, K., Zhang, X., Ren, S. and Sun, J. (2016). “Deep Residual Learning for Image Recognition,” In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (New York, NY, USA: IEEE), pp. 770–778. doi: 10.1109/CVPR.2016.90
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., et al. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv:1704.04861. doi: 10.48550/arxiv.1704.04861.
- Hu, J., Shen, L., Albanie, S., Sun, G. and Wu, E. (2019). Squeeze-and-Excitation Networks. arXiv preprint arXiv:1709.01507. doi: 10.48550/arxiv.1709.01507.
- IPCC (2021). “Summary for Policymakers,” in *Climate Change 2021: The Physical Science Basis. Contributions of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Eds. Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R. and Zhou, B. (Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press), 3–32.
- IRENA (2020). *Innovation Outlook: Ocean Energy Technologies* Abu Dhabi: Tech. rep., International Renewable Energy Agency (IRENA).
- Jaccard, P. (1912). The Distribution of the Flora in the Alpine Zone. *New Phytol.* 11, 37–50. doi: 10.1111/j.1469-8137.1912.tb05611.x

- Johannesson, K. A. and Mitson, R. B. (1983). *Fisheries Acoustics: A Practical Manual for Aquatic Biomass Estimation*, vol. 240 of *FAO Fisheries Technical Paper* (Rome, Italy: Food and Agriculture Organization of the United Nations).
- Karpathy, A. (2014). What I Learned From Competing Against a ConvNet on ImageNet. In: *Andrej Karpathy Blog*. Available at: <http://karpathy.github.io/2014/09/02/what-i-learned-from-competing-against-a-convnet-on-imagenet/> (Accessed 2022-05-01).
- Karsten, R., Greenberg, D., Tarbotton, M., Culina, J., Swan, A., O'Flaherty-Sproul, M., et al. (2011). *Assessment of the Potential of Tidal Power From Minas Passage and Minas Basin* (Acadia University). Tech. Rep. 300-170-09-11.
- Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012). "ImageNet Classification With Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems*, vol. 25. Eds. Pereira, F., Burges, C., Bottou, L. and Weinberger, K. (Redhook, NY, USA: Curran Associates, Inc.).
- LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep Learning. *Nature* 521, 436–444. doi: 10.1038/nature14539
- Liu, L., Jiang, H., He, P., Chen, W., Liu, X., Gao, J., et al. (2020). "On the Variance of the Adaptive Learning Rate and Beyond," In *Proceedings of the Eighth International Conference on Learning Representations (ICLR 2020)*, doi: 10.48550/arXiv.1908.03265
- Lowe, S. C., McGarry, L. P., Douglas, J., Newport, J., Oore, S., Whidden, C., et al. (2022). Echofilter: A Deep Learning Segmentation Model Improves the Automation, Standardization, and Timeliness for Post-Processing Echosounder Data in Tidal Energy Streams. arXiv preprint arXiv:2202.09648. doi: 10.48550/arxiv.2202.09648.
- Lowe, S. C., Trappenberg, T. and Oore, S. (2021). LogAvgExp Provides a Principled and Performant Global Pooling Operator. arXiv preprint arXiv:2111.01742. doi: 10.48550/arxiv.2111.01742.
- Melvin, G. D. and Cochrane, N. A. (2012). *A Preliminary Investigation of Fish Distributions Near an In-Stream Tidal Turbine in Minas Passage, Bay of Fundy* (Tech. rep., Canadian Technical Report of Fisheries and Aquatic Sciences), Tech. rep. 3006.
- Melvin, G. D. and Cochrane, N. A. (2015). Multibeam Acoustic Detection of Fish and Water Column Targets at High-Flow Sites. *Estuaries Coasts* 38, 227–240. doi: 10.1007/s12237-014-9828-z
- Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., and Terzopoulos, D. (2022). Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 44, 3523–3542. doi: 10.1109/TPAMI.2021.3059968
- Perez, L., Cossu, R., Grinham, A. and Penesis, I. (2021). Seasonality of Turbulence Characteristics and Wave-Current Interaction in Two Prospective Tidal Energy Sites. *Renewable Energy* 178, 1322–1336. doi: 10.1016/j.renene.2021.06.116
- Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016). "You Only Look Once: Unified, Real-Time Object Detection," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (New York, NY, USA: IEEE), 779–788. doi: 10.1109/CVPR.2016.91
- Roberts, A., Thomas, B., Sewell, P., Khan, Z., Balmain, S. and Gillman, J. (2016). Current Tidal Power Technologies and Their Suitability for Applications in Coastal and Marine Areas. *J. Ocean Eng. Mar. Energy* 2, 227–245. doi: 10.1007/s40722-016-0044-8
- Ronneberger, O., Fischer, P. and Brox, T. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Eds. Navab, N., Hornegger, J., Wells, W. M. and Frangi, A. F. (Cham: Springer International Publishing), 234–241. doi: 10.1007/978-3-319-24574-4_28
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vision* 115, 211–252. doi: 10.1007/s11263-015-0816-y
- Santoro, A., Bartunov, S., Botvinick, M., Wierstra, D. and Lillicrap, T. (2016). One-Shot Learning With Memory-Augmented Neural Networks. arXiv preprint arXiv:1605.06065. doi: 10.48550/arxiv.1605.06065.
- Schmidhuber, J. (2015). Deep Learning. *Scholarpedia* 10, 32832. doi: 10.4249/scholarpedia.32832. Revision 184887.
- Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., et al. (2017). Mastering the Game of Go Without Human Knowledge. *Nature* 550, 354–359. doi: 10.1038/nature24270
- Simmonds, J. and MacLennan, D. (2005). *Fisheries Acoustics Theory and Practice*. 2nd edn (Oxford, UK: Blackwell Publishing).
- Smith, L. N. (2015). No More Pesky Learning Rate Guessing Games. arXiv preprint arXiv:1506.01186. doi: 10.48550/arxiv.1506.01186.
- Smith, L. N. (2018). A Disciplined Approach to Neural Network Hyper-Parameters: Part 1 – Learning Rate, Batch Size, Momentum, and Weight Decay. arXiv preprint arXiv:1803.09820. doi: 10.48550/arxiv.1803.09820.
- Smith, L. N. and Topin, N. (2017). Super-Convergence: Very Fast Training of Residual Networks Using Large Learning Rates. arXiv preprint arXiv:1708.07120. doi: 10.48550/arxiv.1708.07120.
- Tan, M. and Le, Q. (2019). "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, vol. 97, 6105–6114. doi: 10.48550/arxiv.1905.11946
- Tong, Q., Liang, G. and Bi, J. (2022). Calibrating the Adaptive Learning Rate to Improve Convergence of ADAM. *Neurocomputing* 481, 333–356. doi: 10.1016/j.neucom.2022.01.014
- Tsitrin, E., Sanderson, B. G., McLean, M. F., Gibson, A. J. F., Hardie, D. C. and Stokesbury, M. J. W. (2022). Migration and Apparent Survival of Post-Spawning Alewife (*Alosa pseudoharengus*) in Minas Basin, Bay of Fundy. *Anim. Biotelemetry* 10:11. doi: 10.1186/s40317-022-00277-z
- Williamson, B. J., Fraser, S., Blondel, P., Bell, P. S., Waggitt, J. J. and Scott, B. E. (2017). Multisensor Acoustic Tracking of Fish and Seabird Behavior Around Tidal Turbine Structures in Scotland. *IEEE J. Oceanic Eng.* 42, 948–965. doi: 10.1109/JOE.2016.2637179
- Wolf, J., Dominicis, M. D., Lewis, M., Neill, S. P., O'Hara Murray, R., Scott, B., et al. (2022). "9.04 - Environmental Issues for Offshore Renewable Energy," in *Comprehensive Renewable Energy*, 2nd edn. Ed. Letcher, T. M. (Oxford: Elsevier), 25–59. doi: 10.1016/B978-0-12-819727-1.00036-4
- Wright, L. (2019). "Ranger - a Synergistic Optimizer," in *GitHub Repository*. Available at: <https://github.com/lessw2020/Ranger-Deep-Learning-Optimizer>. Revision @8d636a5.
- Yong, H., Huang, J., Hua, X. and Zhang, L. (2020). "Gradient Centralization: A New Optimization Technique for Deep Neural Networks," in *Computer Vision – ECCV 2020*. Eds. Vedaldi, A., Bischof, H., Brox, T. and Frahm, J.-M. (Cham: Springer International Publishing), 635–652. doi: 10.48550/arxiv.2004.01461
- Zhang, M., Lucas, J., Ba, J. and Hinton, G. E. (2019). "Lookahead Optimizer: K Steps Forward, 1 Step Back," in *Advances in Neural Information Processing Systems*, vol. vol. 32 . Eds. Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E. and Garnett, R. (Redhook, NY, USA: Curran Associates, Inc.). doi: 10.48550/arxiv.1907.08610

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Lowe, McGarry, Douglas, Newport, Oore, Whidden and Hasselman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Appendix IV



Modeling the Probability of Overlap Between Marine Fish Distributions and Marine Renewable Energy Infrastructure Using Acoustic Telemetry Data

Charles W. Bangley^{1*}, Daniel J. Hasselman², Joanna Mills Flemming¹, Fredrick G. Whoriskey³, Joel Culina², Lilli Enders⁴ and Rod G. Bradford⁵

OPEN ACCESS

Edited by:

Maria Grazia Pennino,
Spanish Institute of Oceanography
(IEO), Spain

Reviewed by:

Raul Vilela,
BioConsult SH, Germany
Lenaig G Hemery,
Pacific Northwest National Laboratory
(DOE), United States

***Correspondence:**

Charles W. Bangley
Charles.Bangley@dal.ca

Specialty section:

This article was submitted to
Marine Conservation and
Sustainability,
a section of the journal
Frontiers in Marine Science

Received: 10 January 2022

Accepted: 17 May 2022

Published: 26 July 2022

Citation:

Bangley CW, Hasselman DJ,
Flemming JM, Whoriskey FG,
Culina J, Enders L and Bradford RG
(2022) Modeling the Probability of
Overlap Between Marine Fish
Distributions and Marine Renewable
Energy Infrastructure Using
Acoustic Telemetry Data.
Front. Mar. Sci. 9:851757.
doi: 10.3389/fmars.2022.851757

¹ Department of Mathematics and Statistics, Dalhousie University, Halifax, NS, Canada, ² Fundy Ocean Research Centre for Energy, Dartmouth, NS, Canada, ³ Ocean Tracking Network, Dalhousie University, Halifax, NS, Canada, ⁴ Acadia University, Wolfville, NS, Canada, ⁵ Bedford Institute of Oceanography, Fisheries and Oceans Canada, Dartmouth, NS, Canada

Understanding the spatiotemporal distributions of migratory marine species at marine renewable energy sites is a crucial step towards assessing the potential impacts of tidal stream turbines and related infrastructure upon these species. However, the dynamic marine conditions that make tidal channels attractive for marine renewable power development also make it difficult to identify and follow species of marine fishes with existing technologies such as hydroacoustics and optical cameras. Acoustic telemetry can resolve some of these problems. Acoustic tags provide unique individual ID codes at an ultrasonic frequency, which are then detected and recorded by acoustic receivers deployed in the area of interest. By matching detection locations of fish species with environmental conditions at proposed sites for tidal energy infrastructure, species distribution models can be developed to predict the probability of species occurrence at sites of current and planned tidal power development. This information can be used to develop statistically robust encounter rate models to aid in quantifying the risk of tidal power development to migratory fish species. We used this approach to develop a predictive model of striped bass (*Morone saxatilis*) distribution within Minas Passage in the upper Bay of Fundy, Nova Scotia. Model results suggested increased probability of striped bass presence in Minas Passage during late ebb tide conditions and at relatively high water temperatures. We demonstrate the potential utility of species distribution modeling of acoustic tag detections in predicting interactions with renewable energy infrastructure, and show the importance of physical oceanographic variables influencing species distributions in a highly dynamic marine environment.

Keywords: species distribution analysis, tidal stream energy impact, acoustic telemetry, minas passage, boosted regression tree (BRT) models, striped bass (*Morone saxatilis*), encounter risk

INTRODUCTION

Tidal stream power has an opportunity to significantly contribute to the transition to renewable energy for coastal communities in the vicinity of extreme tidal environments. Development is already occurring in northern Scotland and is being explored in parts of the Bay of Fundy in US and Canadian waters (Sparling et al., 2020; Copping et al., 2021). The Minas Passage in the upper Bay of Fundy has particularly high potential for significant power generation, with some estimates suggesting turbines in the area could extract over 2.5 GW of power each tidal cycle (Karsten et al., 2008).

While in-stream tidal power generation lacks the large-scale environmental consequences of fossil fuels and does not affect access to essential habitat as much as dam-based riverine hydropower, tidal power generation may have the potential for harmful interactions with marine species including collisions between animals and rotating turbine blades (Sparling et al., 2020; Copping et al., 2021). In Canadian waters, documenting and mitigating potential harm to fish species is a requirement for permitting the deployment of tidal turbines and other power generation structures under the *Canadian Environmental Assessment Act*, *Fisheries Act*, and *Species at Risk Act*. Globally, some research has been done to assess the potential for collision risk with tidal turbines, though typically this has focused on marine mammals (Gillespie et al., 2021; Onoufriou et al., 2021). The risk of collision is typically framed as an encounter risk, defined as the risk of an individual animal or group of animals entering the turbine's area of effect and presumably either colliding with the turbine or executing some kind of avoidance behavior (Wilson et al., 2007). Encounter rate modeling has traditionally been derived from predator-prey encounter models originally developed by Gerritsen and Strickler (1977) with turbine swept-area representing the area of potential encounter with the "predator." This approach provided a framework that has been adapted to collision risk modeling (Wilson et al., 2007). However, because different species transiting through areas where they may encounter tidal stream turbines are likely to differ in their movement patterns, density, and avoidance behaviors, there is a need for species-specific information on potential encounter rate with tidal power infrastructure.

Species-specific information on spatial and temporal fish distributions in high-energy environments are lacking due to the impracticality of most fishing activities and fishery-independent survey methods in these areas. Hydroacoustic methods have been used to measure baseline fish densities at potential tidal power sites in the Bay of Fundy (Viehman et al., 2015) but it is difficult to translate these data to species-specific presence/absence or abundance data without accompanying trawl survey data to cross-reference acoustic targets. Typical fishery-independent survey methods such as stratified trawling and gillnetting are impossible and likely hazardous in Minas Passage. This leaves intertidal fishing weirs along the margins of the passage as the only practical fishing method that could be used to sample finfish in the immediate area (Dadswell et al., 2020). However, there has been extensive tagging of multiple species with acoustic transmitters in the area through the

activities of various institutions, including Fisheries and Oceans Canada (DFO), The Mi'kmaw Conservation Group (MCG), Acadia University, and Dalhousie University, as well as by international researchers tagging highly migratory species outside of the vicinity of Minas Basin. Deployment metadata, including species and tagging location and tag detection data for many of these projects, are stored on the database of the Ocean Tracking Network (OTN), which organizes acoustic telemetry data with the goal of facilitating collaborative data sharing agreements among individual researchers and projects (Iverson et al., 2018). Acquiring cooperative agreements with researchers conducting acoustic telemetry research on species of fishery and conservation interest in Minas Passage to pool and use their data on species presence and movements provides an unprecedented opportunity to advance our understanding of potential interactions between tidal stream power and fishes.

Acoustic telemetry requires both the transmitters carried by the animals and deployment of receivers to detect the ultrasonic signals. While individual receivers or small arrays of receivers are capable of local tracking, large-scale receiver networks like the OTN and its associated regional nodes allow for tracking at large regional or even continental spatial scales (Bangley et al., 2020b). With this rich scale of data available, it is possible to match tag detections with environmental data to generate predictive species distribution models, which can be used to predict the probability of species presence based on environmental conditions (Bangley et al., 2020a). If both acoustic tag detections and environmental data are recorded in the local area around proposed tidal stream turbine deployment, we suggest it is possible to use this approach to predict the probability of species of interest co-occurring in space and time with tidal stream turbines. Here we demonstrate this approach for a single species using multiple years of environmental and hydrodynamic data recorded in Minas Passage and matched with tag detections of striped bass (*Morone saxatilis*).

METHODS

Study Area

The Minas Passage connects Minas Basin to the greater Bay of Fundy and the rest of the Gulf of Maine. At only 5 km in width, Minas Passage represents the only path for both tidal currents and marine species. During both flood and ebb tide, tidal currents are forced through the passage at velocities that can exceed 5 m/s (Karsten et al., 2011). Due to this high-energy environment, most of the bottom habitat is made up of exposed bedrock and boulder habitat. A volcanic plateau rises to a near-uniform depth off the north shore of Minas Basin near Black Rock. Due to its relatively consistent bathymetry and position in some of the strongest tidal currents it makes up the majority of the area within the Crown Lease Area (CLA) overseen by the Fundy Ocean Research Centre for Energy (FORCE) (**Figure 1**). The CLA was established to assess the viability of tidal power development in Minas Passage and is an active site of research and monitoring related to the establishment of tidal stream power.

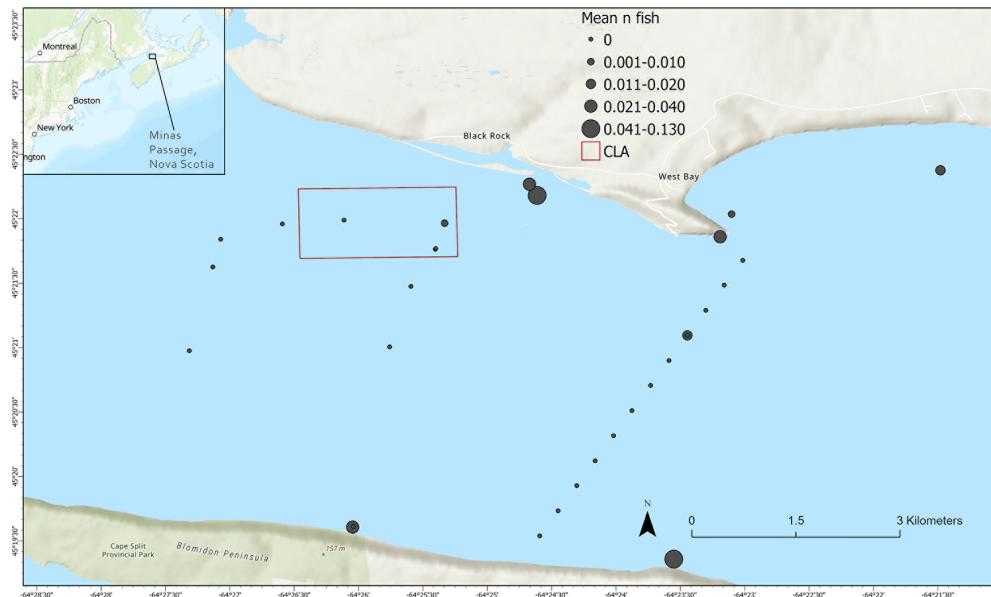


FIGURE 1 | Study area within Minas Basin in the upper Bay of Fundy, including receiver locations with mean number of striped bass detected during October–December 2017–2020, the borders of the FORCE Crown Lease Area (CLA), and the boundaries of the area where species distribution model results will be mapped.

Minas Passage is the only point of entry and exit for most migratory species that occupy Minas Basin (CSAS, 2018). This includes transient species such as Atlantic sturgeon (*Acipenser oxyrinchus*) and sharks as well as anadromous species that spawn in the freshwater tributaries of Minas Basin. These anadromous species include alewife (*Alosa pseudoharengus*), American eel (*Anquilla rostrata*), American shad (*Alosa sapidissima*), Atlantic tomcod (*Microgadus tomcod*), and striped bass. Aside from their ecological importance, these species are economically and culturally significant to local communities.

Focal Species

The Minas Basin population of striped bass spawns in the Shubenacadie River and Grand Lake, Nova Scotia and is considered the northernmost confirmed breeding population for the species (Bradford et al., 2015). A portion of the population remains within Minas Basin year-round and overwinters in Grand Lake, while some larger individuals forage within and transit through Minas Passage on larger-scale coastal migrations (Paramore and Rulifson, 2001; Rulifson et al., 2008; Bradford et al., 2015). Sub-adult and adult striped bass are found throughout Minas Passage from May to October but depart over the course of October as they migrate to overwintering habitats in Grand Lake or the greater Bay of Fundy (Kendall et al., 2018). Oceanic migration in this population has been confirmed through otolith microchemistry and correlates with differences in coloration and diet from individuals that overwinter in freshwater habitats, with likely ocean migrants making up at least 25% of striped bass found in the Shubenacadie River (Paramore and Rulifson, 2001). However, acoustically

tagged striped bass have been detected on receivers in Minas Passage during winter at what would be expected to be lethal temperatures (Keyser et al., 2016). The Bay of Fundy striped bass population is designated as Endangered by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC), and as the only currently confirmed spawning population in the Bay of Fundy, Minas Basin striped bass are of high conservation interest and potential ecological importance (Bradford et al., 2015).

Initial Range Testing

To measure the potential effect of current velocity on tag detection efficiency and range, we deployed five receiver stations equipped with VR2 69-kHz receivers and two sentinel tag stations with 69-kHz transmitters in a line perpendicular to the prevailing current direction (**Supplementary Figure S1**). Receiver stations were deployed with approximately 50 m spacing between stations and sentinel tag stations were deployed 75 m from either end of the receiver line. Receivers and sentinel tags were mounted to the posterior tail assembly of streamlined sub-surface floats designed to pivot and orient into the prevailing current. Each sub-surface float was anchored using a 500-lb single anchor chain link connected using a 2-m length of riser chain. The range test was conducted within the CLA from April 9th to May 11th, 2021 and was timed to match a full monthly tidal cycle. Signed current velocity at each station was derived from the FVCOM model (Chen et al., 2006) in 10-minute increments for the entirety of the deployment period.

The sentinel tags were Innovasea V13 or V16 69-kHz transmitters and transmitted at randomized 1140–1320-second intervals. Power level was 143 dB for the V13 transmitter and 158

dB for the V16, comparable to the power levels of tags deployed on striped bass used in this study. Detection efficiency was calculated as the percentage of expected transmissions that were detected on each receiver and matched with the current velocity at time of detection and the distance between the receiver and the sentinel tag. These data were used to interpolate a matrix of detection efficiency as a function of current velocity and distance.

Model Overview

We used boosted regression tree (BRT) analysis to model the probability of striped bass presence based on hydrodynamic conditions in a section of Minas Passage. BRT analysis is ideal for species distribution modeling because it is insensitive to most error distributions common in ecological data, multicollinearity of explanatory variables, and outliers (Elith et al., 2008; Dormann et al., 2013; Dedman et al., 2017). These features allow BRT analysis to be among the best-performing approaches to predictive species distribution modeling (Valavi et al., 2021). We performed our analysis in R version 4.1.2 (R Core Team, 2021) using the package *gbm.auto*, which automates many of the analytical and mapping steps (Dedman et al., 2017). In the following paragraphs we briefly summarize BRT modeling procedures, but for a more detailed description of each step and the underlying statistical theory see Elith et al. (2008) and see Dedman et al. (2017) for details on how each stage is performed in the *gbm.auto* package.

BRT analysis is a form of regression tree modeling, which divides data based on cut points in the range of each explanatory variable that result in reduced variance in the resulting sub-groups, referred to as branches. The number of branches at each split, or division at a cut point, is defined by tree complexity (tc) and splits typically occur at divisions between greater and lesser values of the response variable. Boosting reduces the variance of individual regression tree models by iteratively replicating trees through a stagewise machine learning process in which information from the previous tree is used to reduce the deviance in the next. The learning rate (lr) of a BRT model represents the contribution of each individual tree to reducing the deviance in the next. At each tree iteration, a portion of the data referred to as the bag fraction (bf) (Elith et al., 2008) is randomly selected and used to cross-validate the remainder of the data, which are used to train the model. This process is repeated until there is no longer significant deviance between successive tree iterations. Generally, a minimum of 1,000 tree iterations is considered sufficient to allow the model to reach the minimum possible deviance (Elith et al., 2008).

When applied to spatial species distribution modeling, BRT analysis requires two types of data: samples and grids (Dedman et al., 2015). Samples data include both the response variable and explanatory variables matched in space and time. These data are used to investigate the relationships between the response and explanatory variables and train the predictive model. Grids data are explanatory variable data recorded or interpolated in a regularly-spaced grid across the area of interest for the species distribution model. The predictive model generated from the samples data is applied to the grids data, generating predictive

map surfaces showing predicted values of the response variable across the area of interest (Dedman et al., 2015; Dedman et al., 2017).

Samples Data

In this model, our samples data included hourly presence/absence of acoustically-tagged striped bass at acoustic receivers as the response variable, and physical/environmental variables derived from FORCE surface water flow radar, current speed/water flow model data, and other sources matched to acoustic receiver locations each hour as the explanatory variables. The measurements used as explanatory variables in this model were sea surface height (m), sea surface height gradient, divergence of horizontal velocity, vorticity, signed water current velocity (m/s), bathymetry standard deviation (m), and water temperature (°C).

Hydrodynamic variables (u and v velocity components, sea surface height, sea surface height gradient, divergence, vorticity) were derived from water surface wave field data collected by two overlapping X-band radar installations covering Minas Passage: one at the FORCE test site and one on the end of Cape Sharp. Data were collected during periods when the wave field was coherent. Radar data were collected during December 2020-April 2021, covering four full tidal cycles. Current velocity components were extracted from the wave field using the Geometric Current Triangulation (GemCuT) algorithm (Bell and McCann, unpublished as cited by Locke, 2019) and computed on a regular 150 x 150-m grid covering the study area at a 20-min temporal resolution. Gaps in the radar data set were filled using unified tidal analysis and prediction/harmonic analysis (Codiga, 2011).

All radar-derived measurements were validated using the FORCE region hydrodynamic model and concurrent acoustic Doppler current profiler (ADCP) data. The FORCE hydrodynamic model uses the Finite Volume Community Ocean Model (FVCOM) as model solver (Chen et al., 2006) and has itself been validated against an extensive number of acoustic Doppler current profilers (ADCPs) deployed in the region (FORCE, unpublished data). Hydrodynamic conditions in Minas Passage are relatively consistent over the course of a tidal cycle and radar data covered most of the tidal variability seen historically (FORCE, unpublished data). Because of this, we were able to hind-cast radar-derived metrics over our study time period using relationships between tide stage, sea surface height, and the other hydrodynamic metrics to predict variable values from historically modeled and recorded sea surface height data.

Signed current velocity (referred to as current velocity from here) was calculated using the u and v velocity components derived from the FORCE radar data in equation 1:

$$vel = (u^2 + v^2)^{1/2} \times s$$

where *s* is 1 or -1 based on the current direction. Current direction was assigned by calculating the direction angle from the u and v velocity components using equation 2:

$$dir\ angle = 180 + \frac{180}{\pi[\tan^{-1}(u, v)]}$$

Based on the prevailing current directions in Minas Passage (Viehman et al., 2019), directional angles between 30 and 205° represented an incoming tide and assigned 1 for s in equation 1, while directional angles greater than 205° or less than 30° represented an outgoing tide and assigned -1 for s in equation 1

Bathymetry standard deviation was derived from high-resolution (2 m) multibeam sonar data collected through the Minas Passage and was assumed to be constant at each samples and grids location across the entire time frame of our study. Temperature was recorded by onboard sensors on four Innovasea HR2 acoustic receivers deployed in the vicinity of the CLA. Temperatures from all four receivers were compared and any measurements from a single receiver with a difference in standard deviation greater than 1°C were considered anomalous and removed. After removal of anomalous temperatures nearly all temperatures were within 1 SD of the others, which we interpreted as supporting the assumption that temperature would be constant throughout the modeled area due to the Minas Passage being an extremely well-mixed system. We applied the mean temperature recorded on the HR receivers uniformly across the modeled area. Except for bathymetry standard deviation, these variables were summarized by hour and matched with the date, hour, and location of each receiver.

All variables in the samples data were tested for multicollinearity using a pairwise Pearson's correlation test. This was conducted using the corr.test function in the R package psych (Revelle, 2022). We used thresholds of correlation strength from Dormann et al. (2013) to assess whether collinearity was occurring. If strong correlation was found between pairs of variables, model runs with one of the variables dropped were used to determine the effect of removing the variable on model performance.

Both striped bass tag detections and environmental data covered the time period of fall through early winter, which was a period during which seasonal changes in the use of Minas Passage by striped bass are expected, including the outward migration of the portion of the population that leaves Minas Basin (Rulifson et al., 2008). This time period covered the dates of October 1st through December 31st and within the years 2017–2020. During this time period, receivers were deployed at 86 stations within the study area (Figure 1). For each receiver, only the hours that individual receiver could be confirmed to have been in the water based on deployment and recovery dates and times from the OTN database were included in our analysis. Deployment periods during which the receiver was recovered from a location other than the most recent deployment location (i.e. due to detaching from the mooring and being recovered elsewhere) were removed from analysis due to our inability to confirm the receiver's location for the full deployment. Striped bass were tagged at their spawning grounds in the Shubenacadie River with Innovasea V16 model 69-kHz transmitters (Innovasea, Halifax, Nova Scotia) surgically implanted into the body cavity, as part of Fisheries and Oceans Canada (DFO) studies on movements within Minas Basin. Tagging occurred during 2016–2018, and primarily targeted individuals greater than 60 cm fork length. The tagged fish were detected on

Innovasea VR2W 69-kHz receivers deployed in Minas Passage during the study period (Figure 1). Striped bass tag detections were summarized as the total number of unique individuals detected during each hour at each receiver, and we assumed that hours during which no individuals were detected represented absence of the species.

Grids Data

Grids data were derived from the same data sources as the samples data and mapped across an area of Minas Passage in which radar data were consistently high quality (southwest corner 45.33990/-64.46124, northeast corner 45.37634/-64.37122) (Figure 1). Grid cell size was 150 x 150 m, which corresponds to the average spatial resolution of the FORCE radar data. Eight grids were generated, each representing the average conditions during slack high, early, mid, and late ebb, slack low, and early, mid, and late flood tide stages in October. The grids used for this demonstration represented mean conditions during each of these tide stages, based on the values of each explanatory variable in each grid cell between October 1st and 31st from 2009 through 2020.

Because Minas Passage is subject to eddies and other complex hydrology, tide stages were defined using a combination of signed current velocity and sea surface height at a single common datum located centrally in the passage (45.34392/-64.32384). Signed current velocity and sea surface height data were obtained at this location by extracting the values nearest to this point from the environmental grids using the filter function from the R package dplyr (Wickham et al., 2021). Based on histograms of current velocity frequency, we classified current velocities between -1.0 and 1.0 as the slack range and defined tide stages as follows: slack high tide was defined as velocity within the slack range with a sea surface height greater than 2 m, slack low tide was defined as velocity within the slack range with sea surface heights less than -2 m, ebb tide was defined as any negative velocity value outside the slack range, and flood tide was defined as any positive velocity value outside the slack range. Flood and ebb tide were further divided into early, mid, and late stages to capture the variation within these parts of the tidal cycle. Current velocity less than -2.5 m/s during ebb tide denoted mid ebb stage, current velocity greater than -2.5 m/s with sea surface height greater than 0 m denoted early ebb stage, and current velocity greater than -2.5 m/s with sea surface height less than 0 m denoted late ebb stage. During flood tide, current velocity greater than 2.5 m/s denoted mid flood, less than 2.5 m/s with sea surface height less than 0 m denoted early flood, and velocity less than 2.5 m/s with sea surface height greater than 0 m denoted late flood stage.

BRT Model Training and Evaluation

Once samples and grids data were prepared, we performed BRT model training and results mapping using the R package *gbm.auto* (Dedman et al., 2017), which automates most of the necessary steps. A binary presence/absence BRT model was used to predict the probability of striped bass presence in each of the four grids. The model was trained by testing combinations of

initial tc, lr, and bf values until the best-performing model was found. Models that generated at least 1,000 tree iterations were evaluated based on the cross-validation score (CV score), mean deviance, and % deviance explained, with the model showing the greatest CV score and % deviance explained and the lowest mean deviance selected as the best-performing model. To assess the potential for model overfitting, we observed the difference between the cross-validated area under operator curve (CV AUC) and training area under operator curve (training AUC), with lower differences indicating lower likelihood of overfitting (Hijmans and Elith, 2013).

Marginal effect plots were generated from the model results and used to illustrate the relationships between striped bass presence probability and each explanatory variable. The relative influence of each explanatory variable was measured as the percentage of tree splits attributed to that variable. Mapped results showed the probability of at least one striped bass being present in each grid cell. To determine how representative the hydrodynamic data used to train the model were of the conditions found in the grids data, maps of “unrepresentativeness” were generated, indicating how dissimilar each grid cell was to the conditions in the samples data.

RESULTS

A total of 69 unique individual striped bass were detected within Minas Passage during October–December 2017–2020, accounting for 254 hourly presence records. Response variable distributions varied (**Figure 2**): sea surface height and sea surface height gradient showed upper and lower peaks, while temperature skewed towards higher values and while bathymetry standard deviation towards the lower values. Current velocity metrics showed a relatively normal distribution and vorticity and divergence values were mostly grouped close to 0 (**Figure 2**).

Multicollinearity tests showed significant ($p < 0.001$) but weak ($r < 0.2$) correlations between current velocity and sea surface height, sea surface height gradient, vorticity, and divergence (**Table S1**). Vorticity and divergence both showed a significant ($p < 0.001$) and strong ($r > 0.9$) correlation, as did sea surface height and sea surface height gradient. Removal of either variable in each pair led to a decline in model performance based on CV and AUC values, so all variables were left in the final model. Pairwise correlations between all other variables were both insignificant ($p > 0.05$) and weak (**Table S1**).

Detection efficiency was greater than 60% at least as far as 350 m at low current velocities, but tapered off quickly as current speed increased on both flood and ebb tide (**Supplementary Figure 2**). Relatively high detection efficiency to at least 150 m occurred at current velocities up to -2.0 m/s during ebb tide and 2.5 m/s during flood tide, and was functionally zero at velocities beyond -3.0 m/s at ebb tide and greater than 3.5 m/s at flood tide. Detection efficiency appeared to reach zero at a lower current velocity during ebb tide but initially tapered off more sharply during flood tide (**Supplementary Figure 2**).

The best performing BRT model had a tc of 7, lr of 0.001, and bf of 0.60, and generated 1950 tree iterations (**Table 1**). Training

and CV AUC values showed a difference of less than 0.094, indicating that overfitting was unlikely. A CV score of over 0.77 was indicative of good performance, and the model explained 66.77% of the deviance (**Table 1**).

Temperature was the most influential variable, accounting for 24.1% of tree splits, followed by bathymetry standard deviation (18.9%), current velocity (15.4%), vorticity (14.6%), and sea surface height (13.3%) (**Figure 3**). Temperatures greater than approximately 12°C had a positive effect on presence probability, which peaked at 14°C. Bathymetry standard deviation had a negative effect between 3 and 10 m, but relatively lower and higher values had a positive effect. Current velocities associated with ebb tide (< 0 m/s) and vorticity less than -0.05 and greater than 0.05 had positive effects on presence probability. Sea surface height had a small but positive effect at values associated with near slack low (-4 to -6 m) and slack high (4 to 6 m) conditions (**Figure 3**).

Maximum presence probability rarely exceeded 0.4 except during late ebb tide, with areas of relatively higher presence probability generally distributed close to shore around Cape Sharp or south of the CLA in the main channel (**Figures 4, 5**). The generally low presence likelihood likely reflected the effect of temperature on striped bass presence during October. Areas of relatively greater or lower striped bass presence probability varied between tide stages. During the outgoing tide (**Figures 4A–D**), areas of elevated presence probability were distributed close to shore around Cape Sharp and into West Bay at the early ebb stage, increased to peak presence probability across a broader area by late ebb tide, and reduced at slack low tide. During the outgoing tide (**Figures 5A–D**), presence probability reduced during early flood stage and was relatively low throughout the study area through slack high tide. Striped bass presence probability within the CLA was relatively low during all tide stages, though it was elevated in small areas to the north and southwest of the CLA boundary (**Figures 4, 5**).

Unrepresentativeness was generally low to moderate (approximately 0.15–0.4) across most of the modeled area, indicating that samples data were fairly representative of environmental variation (**Supplementary Figures 3, 4**). Unrepresentativeness was generally lowest near shore and, during the mid flood through early ebb tide stages, extending into the channel off the end of Cape Sharp. Unrepresentativeness was greatest in the middle of the channel during all tide stages (**Supplementary Figures 3, 4**).

DISCUSSION

Overall our model performed well and produced results that are realistic in the context of local striped bass ecology. BRT marginal effect plots suggested an increased probability of striped bass presence at warmer temperatures, bathymetry at very low or high complexity, and velocity, vorticity, and sea surface height conditions associated with ebb tide. Mapped results reflected these relationships with the greatest maximum presence probability occurring during late ebb tide, with a general rise in presence probability during ebb tide and fall by

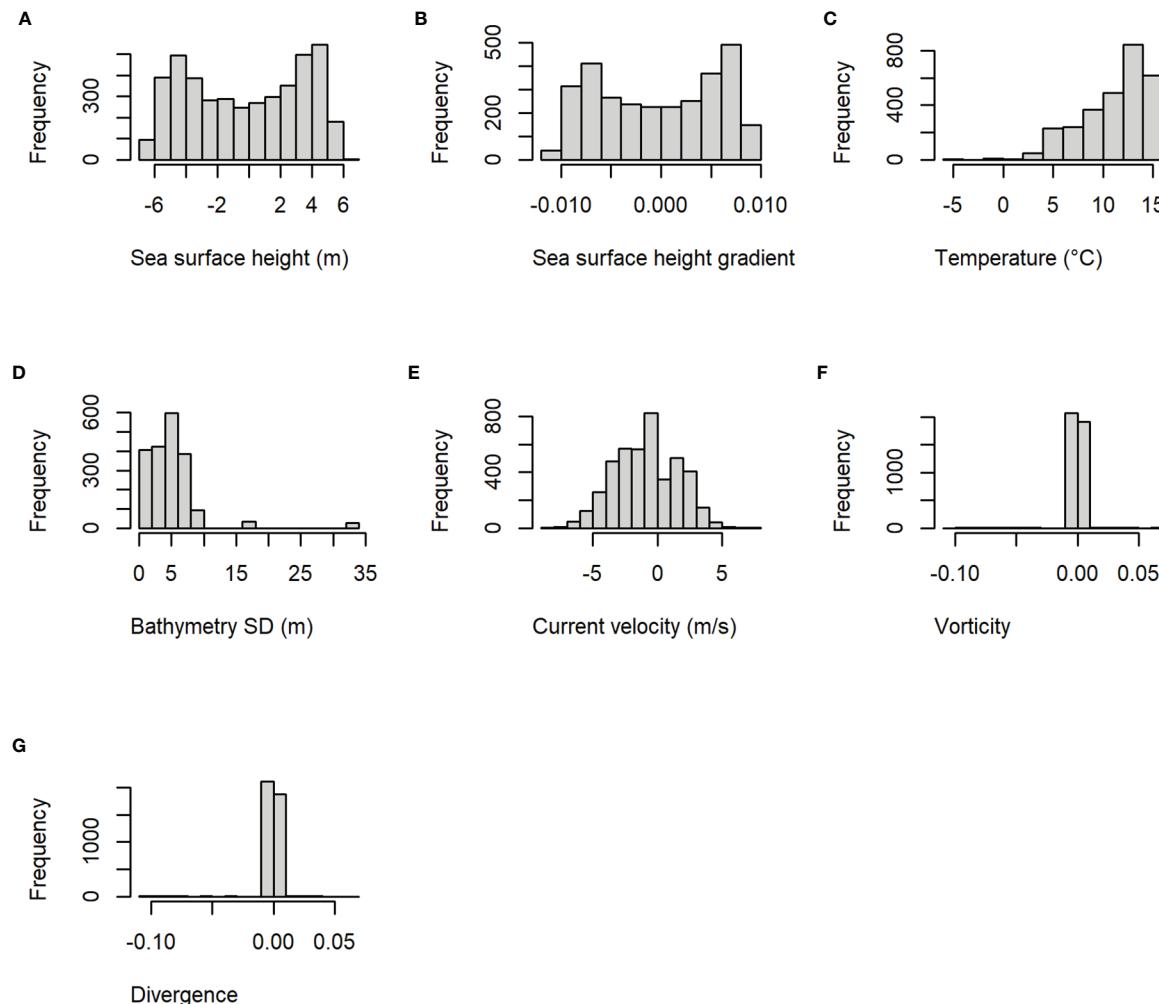


FIGURE 2 | Histograms showing the frequency of explanatory variable values from training data used to develop the final boosted regression tree model of striped bass presence probability during October–December 2017–2020. Variables are (A) sea surface height, (B) sea surface height gradient, (C) temperature, (D) bathymetry standard deviation (SD), (E) current velocity, (F) vorticity, and (G) divergence.

early flood tide. Mapped results also showed relatively low presence probability within the CLA during all tide stages with areas of greater presence probability occurring along shore and farther out in the open passage, though discrete areas of elevated presence probability were distributed just outside the boundaries.

Temperature was the most influential variable on the model, reflecting its clear role as an essential biological driver of striped bass movement from Minas Basin. Informed by the initial analysis of temperature records from four HR receivers, we assumed the relatively uniform temperature throughout Minas Passage was due to the high energy mixing of water. The relationships identified by the model between striped bass presence probability and temperature are likely to be broadly applicable in the study area. To our knowledge, no study thus far has documented stable and predictable temperature micro regimes in the study area that would contradict this assumption and our analysis seems to support it. Our model

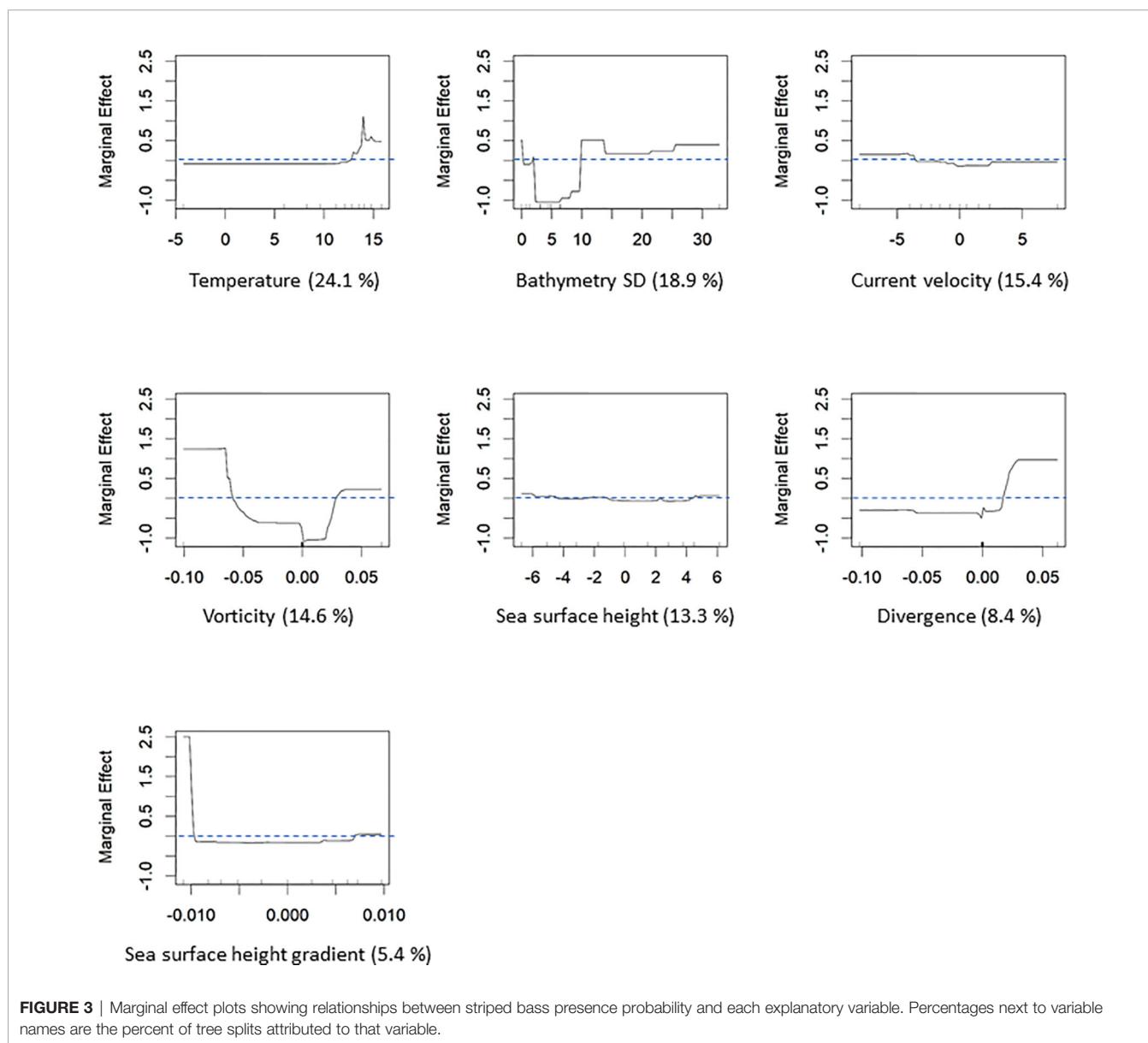
showed a sharp increase in striped bass presence probability at temperatures greater than 12°C, suggesting that the species should be more likely to occur within Minas Passage at relatively warm temperatures. This apparent threshold response occurred at a considerably warmer temperature than the lower incipient lethal range for striped bass from the local population (approximately 0–2°C) but may be as close to thermally optimal conditions for striped bass as possible during October (Coutant et al., 1984; Cook et al., 2006). It is worth noting that 35% of striped bass tagged by Keyser et al. (2016) were detected in the Minas Passage in the winter months of December–March. The species may show a different relationship with hydrodynamic conditions at lower temperatures later in the winter. Season-specific modeling should be performed to determine whether the relative importance and effects of environmental variables change between seasons.

TABLE 1 | Starting parameters and evaluation metrics for the best-performing boosted regression tree model.

Parameter/Metric	Value
Tree complexity	7
Learning rate	0.001
Bag fraction	0.60
N trees	1950
Training AUC	0.999
CV AUC	0.905
Training – CV AUC	0.094
CV score	0.771
% deviance explained	66.77

AUC, area under operator curve, which is shown for both the training and cross-validation (CV) data.

Bathymetry standard deviation was the second most influential variable, with increased striped bass presence probability at considerably lesser and greater bottom complexity than the most commonly occurring measurements. This may represent two different habitat use patterns within Minas Passage. The preference for less-complex bottom habitat may reflect midwater foraging or travel through the passage during seasonal migrations, while occurrence at more complex habitats may represent foraging for benthic prey or sheltering behavior. American lobster (*Homarus americanus*) is an important prey species for large adult striped bass (Nelson et al., 2003). Lobster occur in Minas Passage (CSAS, 2018) and may be more available in complex bottom habitat. A portion of the striped bass used to develop this model carry transmitters equipped with depth sensors, and further analysis of the fish's position in the water column relative to



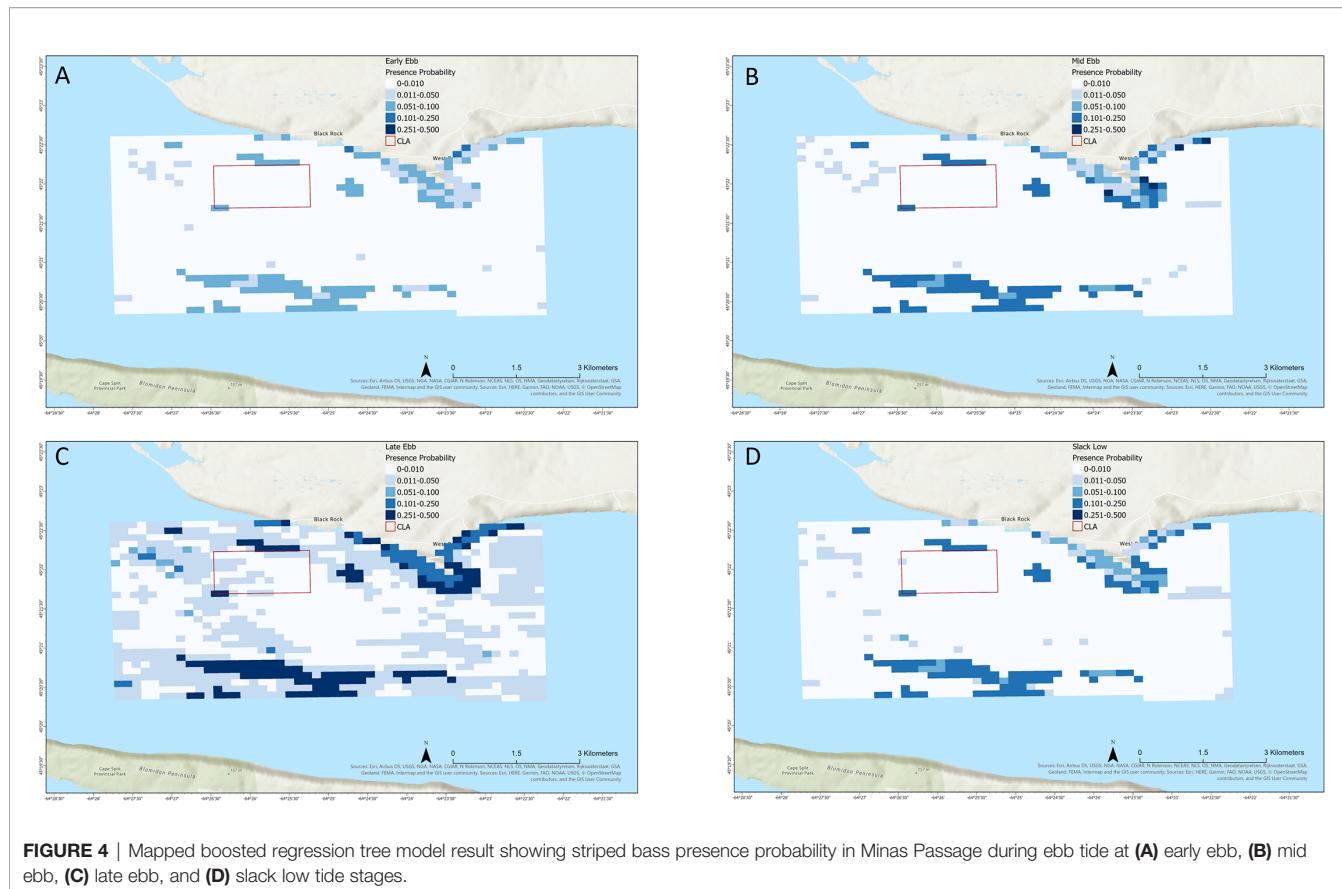


FIGURE 4 | Mapped boosted regression tree model result showing striped bass presence probability in Minas Passage during ebb tide at **(A)** early ebb, **(B)** mid ebb, **(C)** late ebb, and **(D)** slack low tide stages.

bathymetry standard deviation may provide deeper insight into striped bass habitat use within Minas Passage.

Variables directly related to tidal forces in Minas Passage, including current velocity, vorticity, and sea surface height were similarly influential, suggesting that tidal forces play an important role in local fish distributions. Predicted presence probability was most positively influenced by ebbing current velocities, relatively high vorticity, and the highest and lowest sea surface heights. Combined, these suggest that striped bass are more likely to be present in Minas Passage during the later stages of ebb tide. This may represent striped bass taking advantage of outgoing tidal currents during their fall outward migration while selecting for less extreme currents to maintain control over their swimming speed and direction. Such conditions are likely less energetically demanding on swimming performance and may represent periods during which striped bass are able to most efficiently emigrate from Minas Basin. During all tide stages, areas of relatively high presence probability were distributed close to shore, particularly near Cape Sharp and within West Bay. These areas may be refuges or staging areas for striped bass during higher-energy tide stages. These near-shore areas are also the sites of the most complex eddies in the study area, which may be the reason for the positive relationship between striped bass presence probability and relatively turbulent water. In open ocean environments, mesoscale eddies can create complex horizontal boundaries that aggregate prey and attract larger

predators (Godo et al., 2012). Eddies and wakes are created consistently in the vicinity of Cape Sharp during each tidal cycle, which may similarly attract striped bass and other large, mobile predators in Minas Passage.

Broome et al. (2015) noted few striped bass tag detections at current speeds greater than 3 m/s. This may be due to behavioral avoidance of high-flow conditions, but could also be an artifact of reduced tag detection efficiency at greater current velocities (Sanderson et al., 2017). Our own range testing confirmed a sharp reduction in tag detection efficiency at current speeds greater than 2.5 m/s during both ebb and flood tide. Future model iterations will explore this in greater detail and incorporate tag detection efficiency directly into predictions of presence probability.

Mapped predictions suggest that during October the probability of this species overlapping in space and time with tidal power devices deployed in the CLA should be relatively low. Areas of relatively elevated presence probability just outside the CLA boundaries are most likely due to bottom habitat complexity at the edges of the volcanic plateau. Widespread striped bass occurrence within the CLA is most likely during late ebb tide, but even during this tide stage the greatest presence probability does not exceed 0.05. It should be noted that these predictive maps represent a static snapshot of presence probability during average environmental and hydrodynamic conditions in each tide stage, and do not account for movement

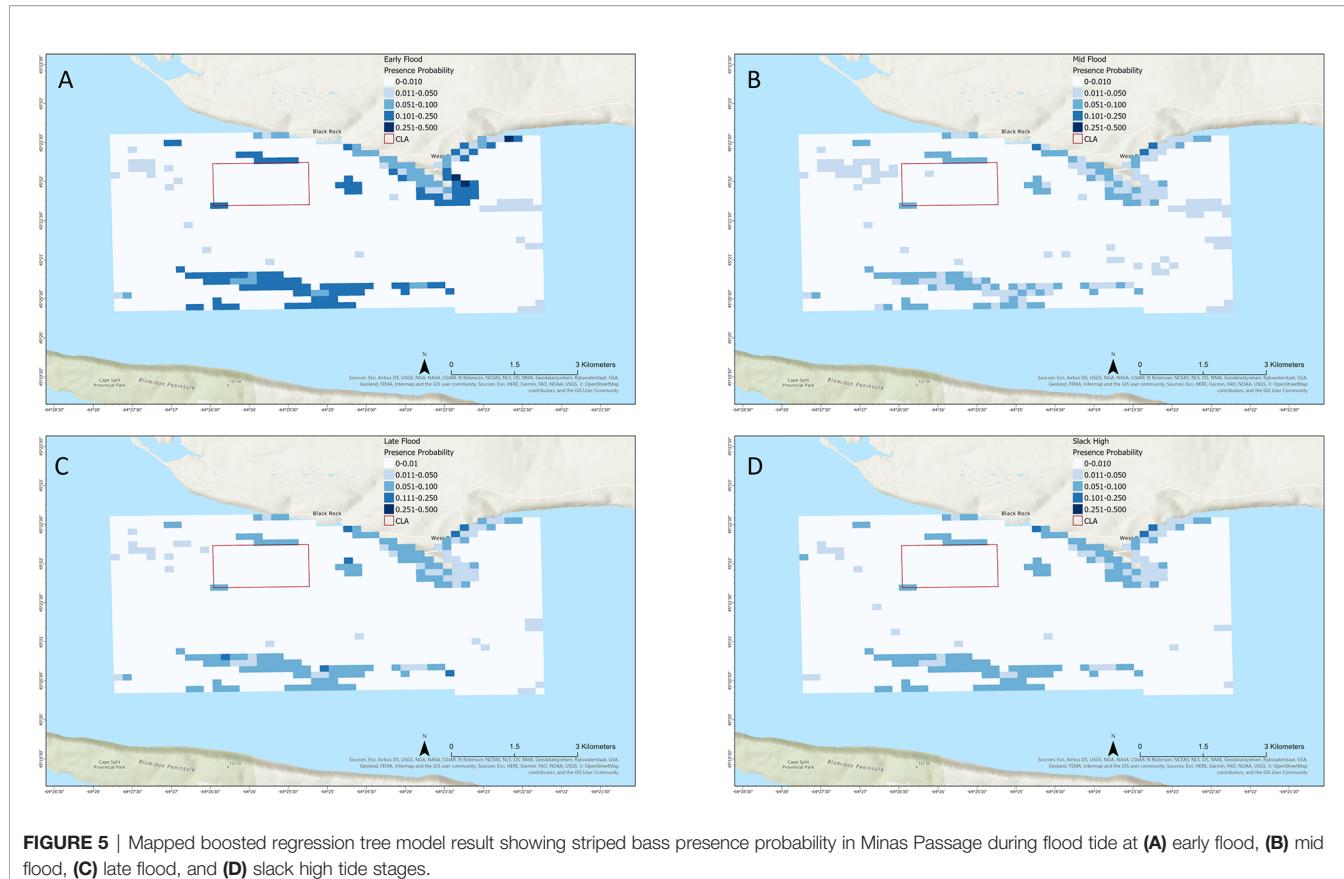


FIGURE 5 | Mapped boosted regression tree model result showing striped bass presence probability in Minas Passage during flood tide at **(A)** early flood, **(B)** mid flood, **(C)** late flood, and **(D)** slack high tide stages.

trajectories of individual fish. Therefore, these maps should be used to assess the probability of striped bass co-occurring in space and time with a tidal power device under a given set of conditions and represents the most basic assessment of potential encounter risk. In locations where fish and tidal stream devices are predicted to overlap, a more fine-scale analysis incorporating the vertical position and movement behaviors of the fish and the physical characteristics of the tidal stream device (height off the bottom, swept area, rotation speed, etc.) should be undertaken to assess the true risk of encounter or collision (Wilson et al., 2007). This modeling approach provides a method of assessing potential overlap at relatively large spatial scales and will help narrow down locations and times where more detailed encounter risk assessment is needed.

While all the variables used in this model have plausible ecological explanations for their relationships with species presence, we were limited by those variables that were available with sufficient spatial and temporal coverage for use in the model. Ideally, we would have included variables such as salinity, dissolved oxygen, and turbidity, but these data were not available for Minas Passage at the spatial and temporal scales needed for this modeling effort. The combination of variables we were able to use provided ecologically plausible results, but there is the possibility that variables we were unable to include may have important influences on the distribution of striped bass and

other species in Minas Passage. Should sufficient data for other variables become available, they will be included in future model iterations.

Some of our variables showed strong and significant correlation but dropping them from the model caused a decline in model performance. This suggests that while all hydrodynamic and environmental variables likely interact in some way within Minas Passage, each variable had at least some independent influence on striped bass presence probability. Such situations are common in ecological studies, where correlation does not necessarily mean collinearity and the true relationships between variables require interpretation based on how they interact in the study system (Dormann et al., 2013).

In this demonstration, we have made two key assumptions: transmissions from tagged fish are equally likely to be detected in all environmental conditions including tide stage, and environmental associations of tagged fish are representative of the general habits of the species and population. Our first assumption is likely violated by the dynamic nature of the Minas Passage environment. Current speed and turbulence can reduce the detection efficiency and range of acoustic transmissions by increasing ambient noise or by pushing the fish through the receiver's detection range before the tag's intermittent transmission can be detected (Sanderson et al., 2017). When building a model of presence probability, this is

problematic because it reduces the likelihood that all actual fish presences are documented and that a lack of detection represents a true absence. Our initial range testing shows that detection efficiency declines sharply at current velocities greater than 2–2.5 m/s associated with both peak ebb and peak flood tide stages, though it may be relatively reliable at current velocities occurring during other tide stages. Therefore, mapped presence probability results are likely more reliable during slack, early, and late tide stages than during peak flood and ebb tide. However, Minas Passage is a complex environment, and finer-scale currents and eddies likely mean that detection efficiency will vary spatially at any tide stage. We continue to conduct long-term range testing to assess detection efficiency of acoustic transmitters at varying distances from receivers under as close to the full spectrum of hydrodynamic conditions in Minas Passage as possible. In future model iterations, the results of this range testing will allow us to better measure the uncertainty of the predicted presence probabilities from our models, and to make any necessary adjustments to the modeling approach and receiver deployment strategy.

The second assumption, that the behavior of tagged fish is representative of the general population, is likely only true for the proportion of large adult striped bass that either forage in Minas Passage or make seasonal migrations out of Minas Basin to the ocean. For striped bass, currently available data suggest this may represent the only portion of the population that regularly occurs in Minas Passage (Paramore and Rulifson, 2001; Rulifson et al., 2008), but direct comparison with tag detections of smaller individuals will be necessary to confirm. Moving forward, this assumption will be tested by incorporating new tag detections that have been recorded since the completion of model training, including new transmitter deployments on species of interest, particularly demographics that are not well-represented in current or historical tag detection data. These new tag detection locations will be compared with predictive SDM maps to determine whether these detections that were not part of the original model training fall within or near areas of greater predicted presence probability. This will be used to validate model predictions, and new tag detections may also be incorporated into new model iterations to determine if this improves predictive performance.

As of this writing, our range testing and follow-up tagging studies testing model assumptions are underway but not complete. Therefore, current model results for assessing the risk of potential encounter with tidal stream turbines in Minas Passage are preliminary and should be used with caution until they have been sufficiently validated. This work has demonstrated a model providing testable predictions of animal distributions based on acoustic tag detections in a highly dynamic environment. The model's initial predictions seem realistic in the context of local striped bass ecology during the autumn study period. The expansion of the work to include data from other species and yearly seasonal cycles will provide a generalized tool useful to tidal power developers seeking to quantify risk and eliminate or reduce the potential for negative interactions between power infrastructure and marine species.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The animal study was reviewed and approved by Department of Fisheries and Oceans Canada.

AUTHOR CONTRIBUTIONS

CWB conceived of and conducted the statistical analyses and wrote the manuscript. DH organized the project, provided background material, and contributed to analysis and manuscript writing. FGW organized and provided access to fish tag and receiver deployment data, and contributed to manuscript writing. JMF contributed to statistical analyses and manuscript writing. JC organized and provided environmental data and contributed to manuscript writing. LE helped develop and describe methods for deriving environmental data and contributed to manuscript writing. RB provided fish tag detection data and metadata, and contributed to manuscript writing.

FUNDING

Natural Resources Canada, Emerging Renewable Power Program - the primary source of funding for this research. The grant number is ERPB-RA-07.

ACKNOWLEDGMENTS

We would like to acknowledge all the researchers at Fisheries and Oceans Canada who participated in tagging striped bass and contributed tag detection data for this modeling effort. Brian Sanderson has been instrumental in the development and analysis of the initial range test and continued range testing. Mike Stokesbury, Joe Beland, Darren Porter, and Charlie Fleming have provided valuable insight into fish movement patterns and telemetry logistics in Minas Basin and are partners in the development and execution of follow-up model validation studies. Jessica Douglas, Louise McGarry, Shaun Allain, and Dylan DeGrace assisted with field ops planning and initial analysis. We would also like to thank all field crew and data team members working with Darren Porter, The Mi'kmaq Conservation Group, Acadia University, and the Ocean Tracking Network for deployment and recovery of acoustic receivers and organization of tag detection data. Funding for this project was provided by Natural Resources Canada (award ERPP-RA-07).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.851757/full#supplementary-material>

Supplementary Figure 1 | Map of receiver and sentinel tag deployments used in the preliminary range test within the FORCE Crown Lease Area. Arrows designate the prevailing tidal current directions.

Supplementary Figure 2 | Matrix of the efficiency (% of average expected transmissions received) of detection of high-power (143 dB) transmissions from 69-kHz sentinel tags over current velocity (m/s) and range from the receiver (m), based

on results of a preliminary range test within the FORCE Crown Lease Area. Figure by B. Sanderson.

Supplementary Figure 3 | Unrepresentativeness maps showing the difference between explanatory variable values in the samples and grids data during October–December 2017–2020 ebb tide at A) early ebb, B) mid ebb, C) late ebb, and D) slack low tide stages.

Supplementary Figure 4 | Unrepresentativeness maps showing the difference between explanatory variable values in the samples and grids data during October–December 2017–2020 during flood tide at A) early flood, B) mid flood, C) late flood, and D) slack high tide stages.

REFERENCES

- Bangley, C. W., Curtis, T. H., Secor, D. H., Latour, R. J., and Ogburn, M. B. (2020a). Identifying Important Juvenile Dusky Shark Habitat in the Northwest Atlantic Ocean Using Acoustic Telemetry and Spatial Modeling. *Mar. Coast. Fish.* 12, 348–363. doi: 10.1002/mcf2.10120
- Bangley, C. W., Whoriskey, F. G., Young, J. M., and Ogburn, M. B. (2020b). Networked Animal Telemetry in the Northwest Atlantic and Caribbean Waters. *Mar. Coast. Fish.* 12, 339–347. doi: 10.1002/mcf2.10128
- Bradford, R. G., Halfyard, E. A., Hayman, T., and LeBlanc, P. (2015). *Overview of 2013 Bay of Fundy Striped Bass Biology and General Status*. Ottawa, ON: Fisheries and Oceans Canada.
- Broome, J., Redden, A., Keyser, F., Stokesbury, M., and Bradford, R. (2015). “Passive Acoustic Telemetry Detection of Striped Bass at the FORCE TISEC Test Site in Minas Passage, Nova Scotia, Canada,” in *Proceedings of the 3rd Marine Energy Technology Symposium*. 5. (Washington, DC)
- Chen, C., Beardsley, R. C., and Cowles, G. (2006). An Unstructured-Grid, Finite-Volume Coastal Ocean Model (FVCOM) System. *Oceanography* 19, 78–89. doi: 10.5670/oceanog.2006.92
- Codiga, D. L. (2011). “Unified Tidal Analysis and Prediction Using the UTide Matlab Functions,” in *Technical Report 2011-01* (Narragansett, RI, USA: Graduate School of Oceanography, University of Rhode Island), 59pp.
- Cook, A. M., Duston, J., and Bradford, R. G. (2006). Thermal Tolerance of a Northern Population of Striped Bass *Morone saxatilis*. *J. Fish. Biol.* 69, 1482–1490. doi: 10.1111/j.1095-8649.2006.01211.x
- Copping, A. E., Hemery, L. G., Viehman, H., Seitz, A. C., Staines, G. J., and Hasselman, D. J. (2021). Are Fish in Danger? A Review of Environmental Effects of Marine Renewable Energy on Fishes. *Biol. Conserv.* 262, 109297. doi: 10.1016/j.biocon.2021.109297
- Coutant, C. C., Zachman, K. L., Cox, D. K., and Pearman, B. L. (1984). Temperature Selection by Juvenile Striped Bass in Laboratory and Field. *Trans. Am. Fish. Soc.* 113, 666–671. doi: 10.1577/1548-8659(1984)113<666:TSBJSB>2.0.CO;2
- CSAS (2018). *Delineating Important Ecological Features of the Evangeline-Cape Blomidon-Minas Basin Ecologically and Biologically Significant Area (EBSA)*. Ottawa, ON: Fisheries and Oceans Canada.
- Dadswell, M. J., Spares, A. D., Potter, E., and Porter, D. (2020). Diversity, Abundance and Size Structure of Fishes and Invertebrates Captured by an Intertidal Fishing Weir at Bramber, Minas Basin, Nova Scotia. *Proc. Nova Scotian Inst. Sci.* 50, 283–318. doi: 10.15273/pnsis.v50i2.10003
- Dedman, S., Officer, R., Brophy, D., Clarke, M., and Reid, D. G. (2015). Modelling Abundance Hotspots for Data-Poor Irish Sea Rays. *Ecol. Model.* 312, 77–90. doi: 10.1016/j.ecolmodel.2015.05.010
- Dedman, S., Officer, R., Clarke, M., Reid, D. G., and Brophy, D. (2017). Gbm.auto: A Software Tool to Simplify Spatial Modelling and Marine Protected Area Planning. *PLOS One* 12, e0188955. doi: 10.1371/journal.pone.0188955
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., et al. (2013). Collinearity: A Review of Methods to Deal With It and a Simulation Study Evaluating Their Performance. *Ecography* 36, 27–46. doi: 10.1111/j.1600-0587.2012.07348.x
- Elith, J., Leathwick, J. R., and Hastie, T. (2008). A Working Guide to Boosted Regression Trees. *J. Anim. Ecol.* 77, 802–813. doi: 10.1111/j.1365-2656.2008.01390.x
- Gerritsen, J., and Strickler, J. R. (1977). Encounter Probabilities and Community Structure in Zooplankton: A Mathematical Model. *J. Fish. Res. Bd. Can.* 34, 73–82. doi: 10.1139/f77-008
- Gillespie, D., Palmer, L., Macaulay, J., Sparling, C., and Hastie, G. (2021). Harbour Porpoises Exhibit Localized Evasion of a Tidal Turbine. *Aquat. Conserv.* 31, 2459–2468. doi: 10.1002/aqc.3660
- Godo, O. R., Samuelsen, A., Macaulay, G. J., Patel, R., Hjollo, S. S., Horne, J., et al. (2012). Mesoscale Eddies are Oases for Higher Trophic Marine Life. *PLoS One* 7, e30161. doi: 10.1371/journal.pone.0030161
- Hijmans, R. J., and Elith, J. (2013). *Species Distribution Modeling With R*, Vol. 87. doi: 10.1.1.190.9177
- Iverson, S. J., Fisk, A. T., Hinch, S. G., Mills Flemming, J., Cooke, S. J., and Whoriskey, F. G. (2018). The Ocean Tracking Network: Advancing Frontiers in Aquatic Science and Management. *Ocean. Track. Net.: Adv. Aquat. Res. Manage.* 01, 1041–1051. doi: 10.1139/cjfas-2018-0481@cjfas-otn.issue01
- Karsten, R. H., Greenberg, D., and Tarbottom, M. (2011). “Assessment of the Potential of Tidal Power From Minas Passage and Minas Basin,” in *Final Report to OER/OETR, Project Number 300-170-09-11*. Halifax, NS: Offshore Energy Environmental Research Association (now Net Zero Atlantic).
- Karsten, R. H., McMillan, J. M., Lickley, M. J., and Haynes, R. D. (2008). Assessment of Tidal Current Energy in the Minas Passage, Bay of Fundy. *Proc. Instit. Mechan. Eng. Part A: J. Power Energy* 222, 493–507. doi: 10.1243/09576509JPE555
- Kendall, V. J., Stewart, P. L., and Levy, H. A. (2018). “Spatial and Temporal Delineation of Ecological Features of the Evangeline-Cape Blomidon-Minas Basin Ecologically and Biologically Significant Area (EBSA),” in *DFO Canadian Science Advisory Secretariat Research Document 2018/013*. Ottawa, ON: Fisheries and Oceans Canada.
- Keyser, F. M., Broome, J. E., Bradford, R. G., Sanderson, B., and Redden, A. M. (2016). Winter Presence and Temperature-Related Diel Vertical Migration of Striped Bass (*Morone saxatilis*) in an Extreme High-Flow Passage in the Inner Bay of Fundy. *Can. J. Fish. Aquat. Sci.* 73, 1777–1786. doi: 10.1139/cjfas-2016-0002
- Locke, J. G. (2009). *X-Band Radar as a Site Assessment Tool in Minas Passage* Vol. 119 (Wolfville, NS: Acadia University).
- Nelson, G. A., Chase, B. C., and Stockwell, J. (2003). Food Habits of Striped Bass (*Morone saxatilis*) in Coastal Waters of Massachusetts. *J. Northwest. Atlantic. Fish. Sci.* 32, 1–25. doi: 10.2960/J.v32.a1
- Onoufriou, J., Russell, D. J. F., Thompson, D., Moss, S. E., and Hastie, G. D. (2021). Quantifying the Effects of Tidal Turbine Array Operations on the Distribution of Marine Mammals: Implications for Collision Risk. *Renewable Energy* 180, 157–165. doi: 10.1016/j.renene.2021.08.052
- Paramore, L. M., and Rulifson, R. A. (2001). Dorsal Coloration as an Indicator of Different Life History Patterns for Striped Bass Within a Single Watershed of Atlantic Canada. *Trans. Am. Fish. Soc.* 130, 663–674. doi: 10.1577/1548-8659(2001)130<663:DCAAIO>2.0.CO;2
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. (Vienna, Austria: R Foundation for Statistical Computing). Available at: <https://www.R-project.org/>.

- Revelle, W. (2022). *Psych. Procedures for Personality and Psychological Research*. (Evanston, Illinois, USA: Northwestern University). Available at: <https://CRAN.R-project.org/package=psychVersion=2.2.3>.
- Rulifson, R. A., McKenna, S. A., and Dadswell, M. J. (2008). Intertidal Habitat Use, Population Characteristics, Movement, and Exploitation of Striped Bass in the Inner Bay of Fundy, Canada. *Trans. Am. Fish. Soc.* 137, 23–32. doi: 10.1577/T06-174.1
- Sanderson, B., Buhariwalla, C., Adams, M., Broome, J., Stokesbury, M., and Redden, A. (2017). *Quantifying Detection Range of Acoustic Tags for Probability of Fish Encountering MHK Devices* Vol. 27, 10. (Cork, Ireland: Proceedings of the 12th European Wave and Tidal Energy Conference).
- Sparling, C. E., Seitz, A. C., Masden, E., and Smith, K. (2020). “OES-Environmental 2020 State of the Science Report, Chapter 3: Collision Risk for Animals Around Turbines,” in *US Department of Energy Office of Science and Technical Information*. Washington, DC: US Department of Energy, 38 p. doi: 10.2172/1632881
- Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., and Elith, J. (2021). Predictive Performance of Presence-Only Species Distribution Models: A Benchmark Study With Reproducible Code. *Ecol. Monogr.* 92, e01486. doi: 10.1002/ecm.1486
- Viehman, H., Hasselman, D., Boucher, T., Douglas, J., and Bennett, L. (2019). “Integrating Hydroacoustic Approaches to Predict Fish Interactions With In-Stream Tidal Turbines,” in *Report to the Fundy Ocean Research Centre for Energy, Project No. 300-208*. Dartmouth, NS: Fundy Ocean Research Centre for Energy
- Viehman, H. A., Zydlewski, G. B., McCleave, J. D., and Staines, G. J. (2015). Using Hydroacoustics to Understand Fish Presence and Vertical Distribution in a Tidally Dynamic Region Targeted for Energy Extraction. *Estuar. Coast.* 38, 215–226. doi: 10.1007/s12237-014-9776-7
- Wickham, H., François, R., Henry, L., and Müller, K. (2021) *Dplyr: A Grammar of Data Manipulation. R Package Version 1.0.7*. Available at: <https://CRAN.R-project.org/package=dplyr>.
- Wilson, B., Batty, R. S., Daunt, F., and Carter, C. (2007). *Collision Risks Between Marine Renewable Energy Devices and Mammals, Fish, and Diving Birds*. (Oban, Scotland: Scottish Association for Marine Science).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Bangley, Hasselman, Flemming, Whoriskey, Culina, Enders and Bradford. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Appendix V



OPEN ACCESS

EDITED AND REVIEWED BY
Anna Milena Zivian,
Ocean Conservancy, United States

*CORRESPONDENCE
Daniel J. Hasselman
dan.hasselman@fundyforce.ca

SPECIALTY SECTION
This article was submitted to
Ocean Solutions,
a section of the journal
Frontiers in Marine Science

RECEIVED 09 July 2022
ACCEPTED 15 August 2022
PUBLISHED 25 August 2022

CITATION
Hasselman DJ, Li H, Cotter E and
Joslin J (2022) Editorial: novel
technologies for assessing the
environmental and ecological impacts
of marine renewable energy systems.
Front. Mar. Sci. 9:990327.
doi: 10.3389/fmars.2022.990327

COPYRIGHT
© 2022 Hasselman, Li, Cotter and
Joslin. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](#). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Editorial: Novel technologies for assessing the environmental and ecological impacts of marine renewable energy systems

Daniel J. Hasselman^{1*}, Huidong Li², Emma Cotter³
and James Joslin⁴

¹Fundy Ocean Research Centre for Energy (FORCE), Halifax, NS, Canada, ²Earth Systems Science Division, Pacific Northwest National Laboratory (DOE), Richland, WA, United States, ³Coastal Sciences Division, Pacific Northwest National Laboratory (DOE), Richland, WA, United States, ⁴MarineSitu Inc., Seattle, WA, United States

KEYWORDS

environmental effects, new technologies, climate change, marine renewable energy, novel approaches

Editorial on the Research Topic

[Novel technologies for assessing the environmental and ecological impacts of marine renewable energy systems](#)

The continued expansion and worldwide adoption of renewable energy systems, including marine renewable energy (MRE) technologies, is essential for addressing climate change ([IPCC, 2019](#); [IRENA, 2020](#); [IPCC, 2022](#)). Globally, the amount of potentially harvestable tidal stream and wave energy from nearshore areas is sufficient to meet current worldwide energy demand ([Mørk et al., 2010](#); [IRENA, 2020](#)). However, the share of MRE in global electricity generation falls far short of this potential due to the current small scale of deployments; typically, single devices or small-scale arrays. MRE expansion to larger, commercial scales may contribute to addressing the climate crisis, but is hampered by a variety of factors, including uncertainty about their environmental effects ([Neill et al., 2012](#); [Kempener & Neumann, 2014a](#); [Kempener & Neumann, 2014b](#); [Copping et al., 2016](#)).

Environmental monitoring around MRE devices has typically relied on standard oceanographic and remote-sensing instruments not intended for use in the complex hydrodynamic conditions typical of tidal channels and the nearshore regions where MRE development is planned ([Hasselman et al., 2020](#)). Exposure of environmental monitoring instruments to dynamic marine conditions has revealed challenges that have stimulated pioneering research and innovations in the technologies and approaches for understanding the effects of MRE devices on marine ecosystems. This Research Topic has compiled contributions from authors leading cutting-edge research advancing our understanding of the environmental and ecological effects of MRE development, thereby facilitating the expansion of the sector and accelerating progress in addressing the climate crisis.

Advances in machine learning are enhancing our understanding of the environmental effects of MRE devices. Multibeam imaging sonars and optical cameras are frequently used to monitor for interactions of marine animals with MRE devices, but post-processing is laborious, and the accurate identification of species remains challenging. Using convolutional neural networks and Kalman filters, [Kandimalla et al. \(2022\)](#) developed an automated real-time deep-learning framework for the accurate detection, tracking, species identification, and enumeration of fish recorded using a DIDSON imaging sonar and optical cameras. This achieved relatively high performance, though results were highly dependent on the quality of training data. Although the method was developed using data collected from a river and hydroelectric facility, it could be applied to monitoring MRE devices with site-specific retraining.

Machine learning is also enhancing the application of echosounders for monitoring fish in tidal channels. Turbulent hydrodynamics can entrain air in the water column that must be excluded before analyses, but the boundary of entrained air is porous, and its penetration depth can vary, complicating its identification and removal. Using echosounder data from tidal channels in Nova Scotia, [Lowe et al. \(2022\)](#) applied a deep learning approach to develop ‘Echofilter’ – a new model that accurately (>95%) identifies the boundary of entrained air, and reduces the post-processing time for raw echosounder data by 50%. Echofilter improves the standardization and repeatability of this process by removing the subjectivity inherent to manual post-processing.

It is also important to understand the implications of removing data contaminated by entrained air on estimates of fish abundance and distribution at MRE sites. Using echosounder data from Nova Scotia, [Viehman et al. \(2022\)](#) found little influence of entrained air on estimates of fish abundance and vertical distribution from the lower 70% of the water column and during current speeds < 3 m/s. However, the upper water column and faster current speeds were under-sampled, limiting accurate quantification of fish abundance and distribution at those times. These results highlight the value of complementary technologies that monitor animal movements for understanding potential environmental effects of MRE devices.

One of these technologies is acoustic telemetry. [Bangley et al. \(2022\)](#) demonstrate an approach to develop a predictive species distribution model for migratory fish species using fish implanted with acoustic tags that can be detected at various monitored locations. The authors matched physical oceanographic variables with tag detections of the species at receiver stations and used boosted regression tree analyses to generate a predictive species distribution model for striped bass in the Bay of Fundy. This framework can be applied to other fish telemetry datasets, and turbine specific parameters can be integrated to generate encounter rate models for quantifying the risk of MRE devices.

The integration of complementary monitoring technologies into subsea monitoring packages to facilitate continuous operation over extended periods and provide monitoring data

in real-time is a noteworthy advance in facilitating the expansion of the MRE sector. [Gillespie et al. \(2022\)](#) describe the development of a cabled subsea monitoring platform equipped with high-frequency multibeam sonars and a tetrahedral array of high-frequency hydrophones for monitoring the fine-scale movements of marine mammals around operational tidal stream turbines. The results proved the system to be highly reliable, and the platform will be deployed close to an operational turbine in 2022.

Work from tidal channels has improved our understanding of how hydrodynamics can influence species distributions (e.g., turbulent features may increase the availability of prey to predators). Knowledge of these associations is important for understanding potential environmental effects of MRE development. [Slingsby et al. \(2022\)](#) used drone imagery to quantify associations of diving seabirds (auks) with turbulence features. They found that auks primarily oriented themselves across the flow, and that density distribution was influenced by current velocity and tide phase, frequently coinciding with kolk-boils at the sea surface. This work highlights the value of drones for environmental monitoring and collection of seabird data that is difficult using conventional survey methods.

Cost-effective and practical monitoring approaches are needed to advance the MRE sector. [Fraser and Waggett \(2022\)](#) describe an approach for providing site-specific data on diving seabird behavior and prey assemblages using shore-based observation and baited fish traps. The information gathered using this approach provides metrics that inform environmental impact assessments and collision risk models for seabirds and site-specific data on prey assemblages in a cost-effective manner that will facilitate the responsible development of the MRE sector.

The studies compiled herein highlight recent advances for understanding the environmental and ecological effects of MRE development. Additional innovations will be needed to help facilitate the deployment of MRE devices at scales that can help address climate change, and this should include social science research on social, cultural and economic impacts.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

Author JJ was employed by MarineSitu Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Bangley, C. W., Hasselman, D. J., Flemming, J. M., Whoriskey, F. G., Culina, J., Enders, L., et al (2022). Modeling the Probability of Overlap Between Marine Fish Distributions and Marine Renewable Energy Infrastructure Using Acoustic Telemetry Data. *Front. Mar. Sci.* 9, 1–13. doi: 10.3389/fmars.2022.851757
- Copping, A. E., Sather, N. K., Hanna, L., Whiting, J., Zydlewski, G. B., Staines, G., et al. (2016) Annex IV 2016 state of the science report: Environmental effects of marine renewable energy development around the world. Available at: <https://doi.org/10.1097/JNN.0b013e3182829024>.
- Fraser, S., Waggett, J. J., and Hasselman, D. J. (2022). Practical Approaches for Providing Empirical Data on Seabird Behavior and Prey Assemblages in Tidal Channels, 1–17. doi: 10.3389/fmars.2022.851476
- Gillespie, D., Oswald, M., Hastie, G., and Sparling, C. (2022). Marine Mammal HiCUP: A High Current Underwater Platform for the Long-Term Monitoring of Fine-Scale Marine Mammal Behavior Around Tidal Turbines. *Front. Mar. Sci.* 9, 1–10. doi: 10.3389/fmars.2022.850446
- Hasselman, D. J., Barclay, D. R., Cavagnaro, R., Chandler, C., Cotter, E., Gillespie, D. M., et al. (2020). “Environmental monitoring technologies and techniques for detecting interactions of marine animals with turbines,” in A. E. Copping and L. G. Hemery (Eds.) *OES-Environmental 2020 State of the Science Report: Environmental Effects of Marine Renewable Energy Development Around the World. Report for Ocean Energy Systems (OES)*. pp. 176–212. doi: 10.2172/1633202
- IPCC (2019). “Special report on the ocean and cryosphere in a changing climate,” in *The ocean and cryosphere in a changing climate*. (Cambridge UK and New York, NY, USA: Cambridge University Press). Available at: <https://doi.org/10.1017/9781009157964>.
- IPCC (2022). “Climate change 2022: Mitigation of climate change,” in *Contribution of working group III to the sixth assessment report of the intergovernmental panel on climate change*. Eds. J. M. P. R. Shukla, J. Skea, R. Slade, A. Kourdajie, R. V. Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa and S. Luz (Cambridge UK and New York, NY, USA: Cambridge University Press). Available at: <https://doi.org/10.1017/9781009157926>.
- IRENA (2020). *Innovation outlook - ocean energy technologies*. (Abu Dhabi: International Renewable Energy Agency).
- Kandimalla, V., Richard, M., Smith, F., Quirion, J., Torgo, L., and Whidden, C. (2022). Automated Detection, Classification and Counting of Fish in Fish Passages With Deep Learning. *Front. Mar. Sci.* 8, 1–15. doi: 10.3389/fmars.2021.823173
- Kempener, R., and Neumann, F. (2014a). “Tidal energy technology brief,” in *IRENA ocean energy technology brief 3*. (Abu Dhabi: International Renewable Energy Agency). Available at: <https://doi.org/10.1002/9781119014492.ch4>.
- Kempener, R., and Neumann, F. (2014b). “Wave energy technology brief,” in *IRENA ocean energy technology brief 4*. (Abu Dhabi: International Renewable Energy Agency). Available at: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2014/Wave-Energy_V4_web.pdf.
- Lowe, S. C., McGarry, L. P., Douglas, J., Newport, J., Oore, S., Whidden, C., and Hasselman, D. J. (2022). Echofilter: A Deep Learning Segmentation Model Improves the Automation, Standardization, and Timeliness for Post-Processing Echosounder Data in Tidal Energy Streams. *Front. Mar. Sci.* 9, 1–21. doi: 10.3389/fmars.2022.867857
- Mørk, G., Barstow, S., Kabuth, A., and Pontes, M. T. (2010). “Assessing the global wave energy potential,” in *Proceedings of the ASME 2010 29th international conference on ocean, offshore, and Arctic engineering, June 6–11, 2010* (Shanghai, China: International Conference on Ocean, Offshore and Arctic Engineering).
- Neill, S. P., Jordan, J. R., and Couch, S. J. (2012). Impact of tidal energy converter (TEC) arrays on the dynamics of headland sand banks. *Renewable Energy* 37 (1), 387–397. doi: 10.1016/j.renene.2011.07.003
- Slingsby, J., Scott, B. E., Kregting, L., Mcilvenny, J., Wilson, J., Langlois, S., et al (2022). Using unmanned aerial vehicle (UAV) imagery to characterise pursuit-diving seabird association with tidal stream hydrodynamic habitat features. *Front. Mar. Sci.* 9, 1–17. doi: 10.3389/fmars.2022.820722
- Viehman, H. A., Hasselman, D. J., Douglas, J., and Boucher, T. (2022). The ups and downs of using active acoustic technologies to study fish at tidal energy sites. *Front. Mar. Sci.* 9, 851400. doi: 10.3389/fmars.2022.851400