Received: 4 October 2022 | Accepted: 1 March 2023

DOI: 10.1111/2041-210X.14095

Methods in Ecology and Evolution

RESEARCH ARTICLE

Identification of fish sounds in the wild using a set of portable audio-video arrays

Xavier Mouy[1,2] | Morgan Black[3] | Kieran Cox[3] | Jessica Qualley[3] | Stan Dosso[1] | Francis Juanes[3]

[1]School of Earth and Ocean Sciences, University of Victoria, 3800 Finnerty Road, Victoria, British Columbia V8P 5C2, Canada

[2]Integrated Statistics, Inc., Under Contract to National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, 166 Water Street, Woods Hole, Massachusetts 02543, USA

[3]Department of Biology, University of Victoria, 3800 Finnerty Road, Victoria, British Columbia V8P 5C2, Canada

Correspondence
Xavier Mouy
Email: xavier.mouy@outlook.com

Funding information
Canadian Healthy Oceans Network; Liber Ero Foundation; Mitacs; Natural Sciences and Engineering Research Council of Canada; CFI/BCKDF

Handling Editor: Camille Desjonquères

Abstract

1. Associating fish sounds to specific species and behaviours is important for making passive acoustics a viable tool for monitoring fish. While recording fish sounds in tanks can sometimes be performed, many fish do not produce sounds in captivity. Consequently, there is a need to identify fish sounds in situ and characterise these sounds under a wide variety of behaviours and habitats.
2. We designed three portable audio-video platforms capable of identifying species-specific fish sounds in the wild: a large array, a mini array and a mobile array. The large and mini arrays are static autonomous platforms than can be deployed on the seafloor and record audio and video for one to two weeks. They use multichannel acoustic recorders and low-cost video cameras mounted on PVC frames. The mobile array also uses a multichannel acoustic recorder, but mounted on a remotely operated vehicle with built-in video, which allows remote control and real-time positioning in response to observed fish presence. For all arrays, fish sounds were localised in three dimensions and matched to the fish positions in the video data. We deployed these three platforms at four locations off British Columbia, Canada.
3. The large array provided the best localisation accuracy and, with its larger footprint, was well suited to habitats with a flat seafloor. The mini and mobile arrays had lower localisation accuracy but were easier to deploy, and well suited to rough/uneven seafloors. Using these arrays, we identified, for the first time, sounds from quillback rockfish *Sebastes maliger*, copper rockfish *Sebastes caurinus* and lingcod *Ophiodon elongatus*. In addition to measuring temporal and spectral characteristics of sounds for each species, we estimated mean source levels for lingcod and quillback rockfish sounds (115.4 and 113.5 dB re 1 μPa, respectively) and maximum detection ranges at two sites (between 10.5 and 33 m).
4. All proposed array designs successfully identified fish sounds in the wild and were adapted to various budget, logistical and habitat constraints. We include here building instructions and processing scripts to help users replicate this

methodology, identify more fish sounds around the world and make passive acoustics a more viable way to monitor fish.

KEYWORDS
acoustic localisation, copper rockfish, fisheries, lingcod, passive acoustic monitoring, quillback rockfish, video cameras

1 | INTRODUCTION

Over 1000 species of fishes worldwide are currently known to be soniferous (Kaatz, 2002; Looby et al., 2021; Rountree et al., 2006). It is likely that many more species produce sounds, but their repertoires have not yet been identified (Looby et al., 2022). Fish can produce sound incidentally while feeding or swimming (e.g. Amorim et al., 2004; Moulton, 1960) or intentionally for communication (Bass & Ladich, 2008; Ladich & Myrberg, 2006). For example, fish sound spectral and temporal characteristics can convey information about male status and spawning readiness to females (Montie et al., 2016), or male body condition (Amorim et al., 2015). It has been speculated that some species of fish may also emit sound to orient themselves in the environment (i.e. by echolocation, Tavolga, 1977). As is the case for marine mammal vocalisations, fish sounds can typically be associated with a specific species and sometimes to specific behaviours (Ladich & Myrberg, 2006; Lobel, 1992). Recently, Parmentier et al. (2021) used fish sounds to identify a new cryptic species of humbug damselfish in French Polynesia. Several populations of the same species can also have different acoustic dialects (Parmentier et al., 2005). Consequently, researchers can measure the temporal and spectral characteristics of recorded fish sounds to identify which species of fish are present in a particular environment, to infer their behaviour and, in some cases, to potentially identify and track a specific population (Luczkovich et al., 2008).

Using passive acoustics to monitor fish can complement existing monitoring techniques such as net sampling (Portt et al., 2006), active acoustics (Godøl et al., 2014) or acoustic tagging (Pittman et al., 2014). Passive acoustics presents several advantages: It is non-intrusive, can monitor continuously for long periods of time and can cover large geographical areas. However, to use passive acoustics to monitor fish, their sounds must first be characterised and catalogued under controlled conditions. This can be achieved in various ways. The most common way to identify species and behaviour-specific sounds is to capture and isolate a single fish or several fish of the same species in a controlled environment (typically a fish tank) and record the sounds they produce (e.g. Riera et al., 2018, 2020; Širović et al., 2009). Such an experimental setup precludes sound contamination from other species and allows visual observation of the behaviour of the animal. While these studies provide important findings on fish sound production, they do not always result in sounds that fish produce in their natural environments. To partially address this issue, other studies record fish in natural environments but constrained in fishing net pens to ensure they remain in sufficient proximity of the hydrophones (e.g. Cott et al., 2014). This also presents some challenges as other fish species outside the pen can potentially be recorded.

Passively recording fish in their natural environment has many advantages, especially in terms of not disrupting the animals. However, it provides less control over external variables and also presents many technical challenges. Remotely operated vehicles (ROVs) equipped with video cameras and hydrophones have been used by Sprague and Luczkovich (2004) and Rountree and Juanes (2010). Locascio and Burton (2015) deployed fixed autonomous passive acoustic recorders and conducted diver-based visual surveys to document the presence of fish species. They also developed customised underwater audio and video systems to verify sources of fish sounds and to understand their behavioural contexts. Most of these monitoring techniques are limited by high power consumption and data storage space requirements, and are typically only deployed for short periods of time. Cabled ocean observatories equipped with hydrophones and video cameras provide valuable data for more extended time periods but by their nature are constrained to fixed locations and are expensive to deploy and maintain (Sirovic et al., 2012; Wall et al., 2014). Rountree et al. (2006) noted the need for the research community to develop longer term and affordable autonomous video and audio recorders that are more versatile than the current technology and facilitate cataloguing fish sounds in situ.

A key consideration when cataloguing fish sounds in the wild is the ability to localise the sounds recorded. In most cases, having only a single omnidirectional hydrophone and a video camera is not sufficient. Several fish can produce sounds at the same time and it is important to know which fish in the video recording produced the sound. Although numerous methods have been developed for the large-scale localisation of marine mammals based on their vocalisations (see reviews in Adam & Samaran, 2013; Zimmer, 2011), only a handful of studies have been published to date on the localisation of fish sounds. Wilson et al. (2019), D'Spain and Batchelor (2006), Mann and Jarvis (2004) and Spiesberger and Fristrup (1990) localised distant groups of fish. Putland et al. (2018) localised individual oyster toadfish in two dimensions (2D) using a 20 m long linear array fixed to a dock. Parsons et al. (2009, 2010) and Locascio and Mann (2011) conducted finer scale three-dimensional (3D) localisation and monitored individual fish in aggregations. Gervaise et al. (2019), Ferguson and Cleary (2001) and Too et al. (2019) also performed fine-scale acoustic localisation on sounds produced by invertebrates. Fine-scale localisation is extremely valuable as it can not only be used

with video recordings to identify the species and behaviour of the animals producing sounds, but can also be used to track movements of individual fish, estimate the number of vocalising individuals near the recorder, and measure source levels of the sounds. The latter represents critical information needed to estimate the distance over which fish sounds can propagate before being masked by ambient noise (Locascio & Mann, 2011; Radford et al., 2015).

Once fish sounds are catalogued, passive acoustics alone (without video recordings) can be used for monitoring the presence of fish in space and time. Many of the soniferous fish species are of commercial interest which makes passive acoustic monitoring a powerful and non-intrusive tool that could be used for conservation and management purposes (Davis et al., 2017; Gannon, 2008; Luczkovich et al., 2008; Rountree et al., 2006; Van Parijs et al., 2009). Sounds produced while fish are spawning have been used to document spatiotemporal distributions of mating fish (Bolgan et al., 2017; Lowerre-Barbieri et al., 2011; Luczkovich et al., 2008; Parsons et al., 2016, 2017; Sánchez-Gendriz & Padovese, 2017). Recently, Di Iorio et al. (2018) monitored the presence of fish in Mediterranean tapeweed *Posidonia oceanica* meadows in the Western Mediterranean Sea using passive acoustics over a 200 km area. Finally, Rountree and Juanes (2017) demonstrated how passive acoustics could be used to detect an invasive fish species in a large river system.

As described here, passive acoustics could be a very powerful tool to monitor fish populations and behaviour. However, its capabilities are currently limited significantly by the fact that many fish sounds have not yet been linked to specific species. This is, in part, because there is no readily available instrumentation capable of easily identifying sounds that fish produce in their natural habitat. Here, we propose three audio-video array designs (with their associated analysis software) that address this important research need by localising fish sounds in 3D and matching localised sounds to individual fish captured by video cameras. These systems are portable, adapted to a variety of coastal habitats, and straightforward to build and replicate. By making such hardware and software accessible, we provide the necessary tools that will help expand the worldwide fish sound catalogue and therefore make passive acoustics a more viable tool to monitor fish populations.

2 | MATERIALS AND METHODS

2.1 | Description of the audio-video arrays

Three audio-video arrays, referred to as the large, mini and mobile arrays, were developed to acoustically localise and visually identify fish producing sounds. Each array was designed with different constraints in mind. The large array was configured for the most-accurate 3D acoustic localisation, the mini array for easier deployments in constrained locations or on rough/uneven seafloors and the mobile array for dynamic real-time spatial sampling over shorter time periods (hours rather than days or weeks). We provide detailed building instructions and deployment procedures of these three audio-video arrays in the Supporting Information (Supporting_Information.pdf).

2.1.1 | Large array

The large array is a static platform deployed on the seafloor that records audio and video data for one to 2 weeks (Figure 1). It uses six M36-V35-100 omnidirectional hydrophones (GeoSpectrum Technologies Inc.) connected to an AMAR-G3R4.3 acoustic recorder (JASCO Applied Sciences) with a PVC housing rated to 250 m depth. Four of the hydrophones (1–4 in Figure 1) are connected to the first acquisition board of the recorder via a 4 m long 4-to-1 splitter cable. The two other hydrophones (5, 6 in Figure 1) are connected, via a 3 m long 2-to-1 splitter cable, to the second acquisition board of the recorder. The recorder is set to acquire acoustic data continuously as 30-min wave files, at a sampling frequency of 32 kHz, with a bit depth of 24 bits, and with a pre-digitalisation analog gain of 6 dB. An external battery pack (BP) with 48 D-cell batteries is used to power the recorder, which,

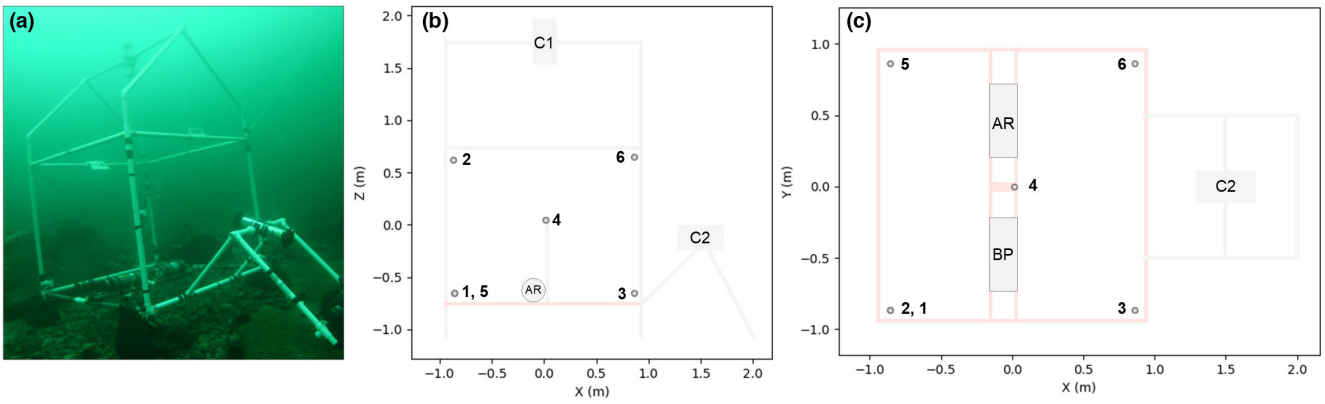

FIGURE 1 Large audio-video array. (a) Photograph of the large array deployed in the field. (b) Side view and (c) top view diagrams of the array with dimensions. The six hydrophones are represented by the numbered grey circles. The top and side video cameras are indicated by C1 and C2, respectively. Note that C1 is not represented in (c) for clarity. The acoustic recorder and its battery pack are indicated by AR and BP, respectively. Grey and red lines represent the PVC structure of the array (red indicating the square base of the array).

using this configuration, allows the system to acquire data for up to 35 days. An end-to-end calibration was performed for each hydrophone using a piston-phone type 42AA precision sound source (G.R.A.S. Sound & Vibration A/S) at 250 Hz. System gain measured on all hydrophones was – 167.3 ± 0.2 dB re FS/μPa, where FS is the full digitalisation scale (i.e., amplitude values between −1 and 1). Positions of the hydrophones within the large array (Figure 1) were defined to maximise the accuracy of the acoustic localisation (see optimisation procedure described in Section 2.3 and in the Supporting Information). Two low-cost autonomous FishCam cameras (Mouy et al., 2020) are used to record video inside the large array. One (C1) is located at the top of the array and is oriented downward towards the seafloor, and the other (C2) is located on the side of the array, pointing horizontally towards hydrophone 4 (Figure 1). Each camera is set to record video continuously during the day (from 5:00 to 21:00 local time) and to shutdown during the night. Video data are recorded on 300-s h264 files, with a frame rate of 10 frames per second, a resolution of 1600×1200 pixels, and an ISO of 400. Both cameras emit different sequences of beeps at 3 kHz every 4 h for time-synchronising the video and acoustic data. The autonomy of the FishCams is storage-limited, dependent on the underwater light conditions, and typically ranges from 8 to 14 days (see Mouy et al., 2020). All instruments are secured to a tent-frame shaped PVC frame of 2 m width, 2 m length and 3 m height (Figure 1). All structure elements (PVC tubes) are perforated to avoid having air pockets that could reflect sounds and impact the localisation accuracy.

2.1.2 | Mini array

Like the large array, the mini array is a static platform deployed on the seafloor. It can record audio and video data for approximately 1 week and has a much smaller footprint than the large array.

The mini array uses four HTI-96-MIN omnidirectional hydrophones (High Tech Inc.) connected to a SoundTrap ST4300HF acoustic recorder (Ocean Instruments). The recorder is set to acquire temperature every 10 s, and acoustic data continuously on 15-min wave files, at a sampling frequency of 48 kHz, and with a bit depth of 16 bits. Using this configuration, the recorder has an autonomy of approximately 7 days. An end-to-end calibration was performed for each hydrophone using a piston-phone type 42AA precision sound source (G.R.A.S. Sound & Vibration A/S) at 250 Hz. System gain measured on all hydrophones was – 168.2 ± 0.2 dB re FS/μPa. Coordinates of the hydrophones are indicated in Figure 2. The mini array uses a single FishCam video camera facing horizontally and placed below the four hydrophones. It is set to record video using the same configuration used for the large array. The recorder, hydrophones, and camera are secured to a PVC frame of 1.2 m width, 1.3 m length and 1.3 m height (Figure 2).

2.1.3 | Mobile array

Unlike the static large and mini arrays, the mobile array is remote controlled, can be re-positioned dynamically on the seafloor in response to observed fish presence and can collect audio and video data over periods of approximately 2 h.

The mobile array uses a SoundTrap ST4300HF acoustic recorder (Ocean Instruments) with four HTI-96-MIN omnidirectional hydrophones (High Tech Inc.) secured on top of a Trident (Sofar Ocean Technologies) underwater ROV (Figure 3). The recorder is set to acquire temperature every 10 s, and acoustic data continuously on 15-min wave files, at a sampling frequency of 48 kHz, and with a bit depth of 16 bits. Both the recorder and hydrophones are the same as used for the mini array, and thus the system gain is the same (i.e. – 168.2 ± 0.2 dB re FS/μPa). The Trident underwater ROV uses a 100 m tether and is controlled by a JXD S192K gaming tablet

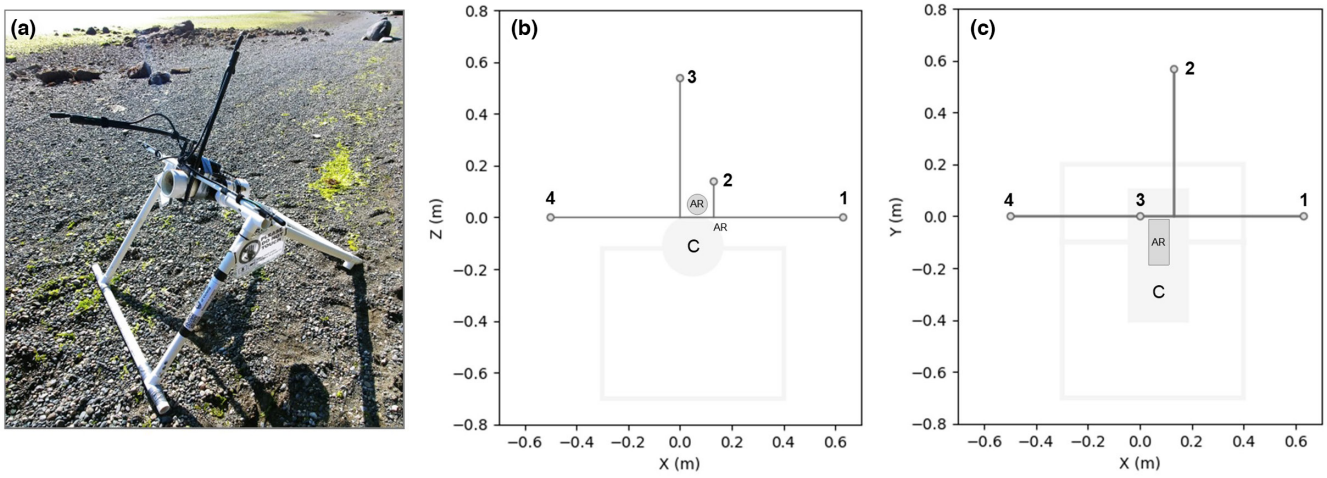

FIGURE 2 Mini audio-video array. (a) Photograph of the mini array on land before deployment. (b) rear view and (c) top view diagrams of the mini array with dimensions. The four hydrophones are represented by the numbered grey circles. The camera and the acoustic recorder are indicated by C and AR, respectively. Grey and black lines represent the frame of the array.

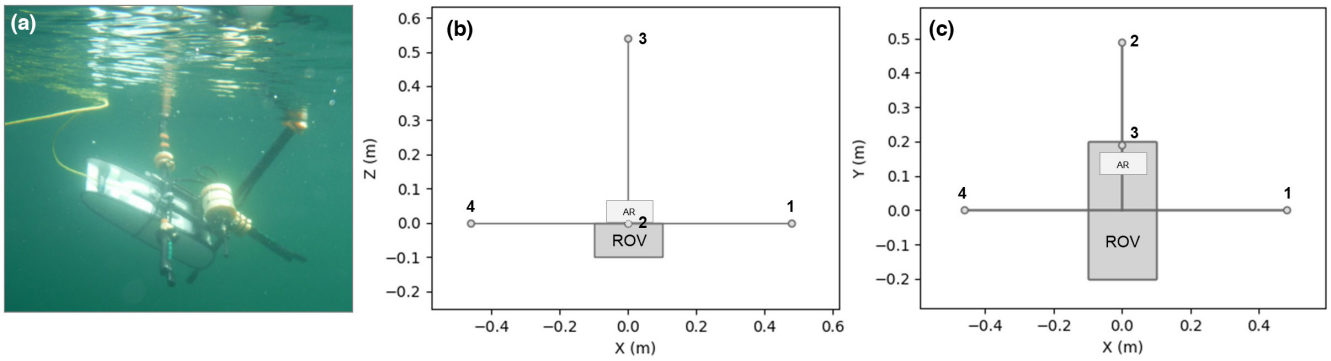

FIGURE 3 Mobile audio-video array. (a) Photograph of the mobile array deployed in the field. (b) rear view and (c) top view diagrams of the mobile array with dimensions. The four hydrophones are represented by the numbered grey circles. The underwater ROV and the acoustic recorder are indicated by ROV and AR, respectively. The video camera is located at the front of the ROV (coordinates: (0, 0.2, – 0.05) m) facing forward. Black lines represent the frame of the array.

(JinXing Digital Co. Ltd) connected via WiFi to the Trident surface hub. Video from the ROV's camera is transmitted in real-time to the JXD controller and is recorded with a resolution of 1920×1080 pixels and a frame rate of 30 frames per second. Time, water temperature, depth and horizontal orientation of the ROV are recorded and displayed on each video recording. The hydrophones and recorder are attached to a small frame made of polyethylene pipe and secured to the bottom of the ROV using four M3 screws. Hydrophones are secured to the frame and are placed at the end of each arm of the frame that extend to approximately 0.5 m on each side, in front and above the ROV (Figure 3). The array is made slightly negatively buoyant to remain stable when resting on the seafloor with the thrusters turned off.

2.2 | Analysis workflow

Data collected by the audio-video arrays are analysed in four steps as summarised in Figure 4. First, acoustic transients are automatically detected on the signal from one of the hydrophones. Second, the time-difference of arrivals (TDOAs) between hydrophones of the array are estimated via cross-correlation. Third, the 3D localisation of the detected events is estimated using linearised inversion for the large array, and fully non-linear inversion for the mini and mobile array (Gervaise et al., 2019; Mouy et al., 2018). Finally, 3D acoustic localisations are matched with the video recordings to identify the species and behaviour of the fish that emitted the sound. Only acoustic localisations that are within the field of view of the video cameras, and have uncertainties in each dimension less than 1 m, are kept for further analysis. Video recordings corresponding to the time of the selected localisations are manually inspected and experienced taxonomists visually identify individual fish matching the position of the acoustic localisation. Identification is performed to the species level and, when possible, to specific behaviours. Synchronisation beeps emitted by the video cameras are identified in the audio data and used to ensure the video and audio data are time-aligned. Acoustic localisation was tested in the field for all three arrays using controlled sound sources. A detailed description of each processing step and the results from the localisation of controlled sources can be found in the Supporting Information (Supporting_Information.pdf).

2.3 | Optimisation of hydrophone placement

For the large array, the placement of the hydrophones was defined so as to minimise the overall localisation uncertainty. This was achieved using the simulated annealing optimisation algorithm (Kirkpatrick et al., 1983) and followed the procedure developed in Dosso and Sotirin (1999). The optimisation consisted of finding the *x*, *y* and *z* coordinates of the six hydrophones (18 parameters) that minimises the average localisation uncertainty of 600 simulated sound sources placed on a 2 m radius sphere around the centre of the array. For the same array footprint (i.e., 2 m×2 m, Figure 5a,b), the spatial capacity of the large array to localise with an uncertainty below 50 cm is more than seven times larger than the hydrophone array used in Mouy et al. (2018) (i.e., 33.5 m^3 vs 4.2 m^3, Figure 5). This means that fish sounds can be localised accurately even when the fish are located up to a meter outside the array. The simulated annealing approach was not used for the mini and mobile arrays as the placement of the hydrophones was mostly dictated by the mechanical constraints of these platforms. Further details on the simulated annealing process can be found in the Supporting Information (Supporting_Information.pdf).

2.4 | Localisation capabilities

Because of their different hydrophone apertures, the three audio-video arrays do not have the same localisation capabilities. Errors in the TDOA measurements have greater impact on the localisations performed with the mini and mobile arrays than with the large array. Figure 6 depicts the estimated 50-cm localisation uncertainty isolines for the three arrays using the same TDOA errors (i.e. 0.12 ms) and shows that both the mini

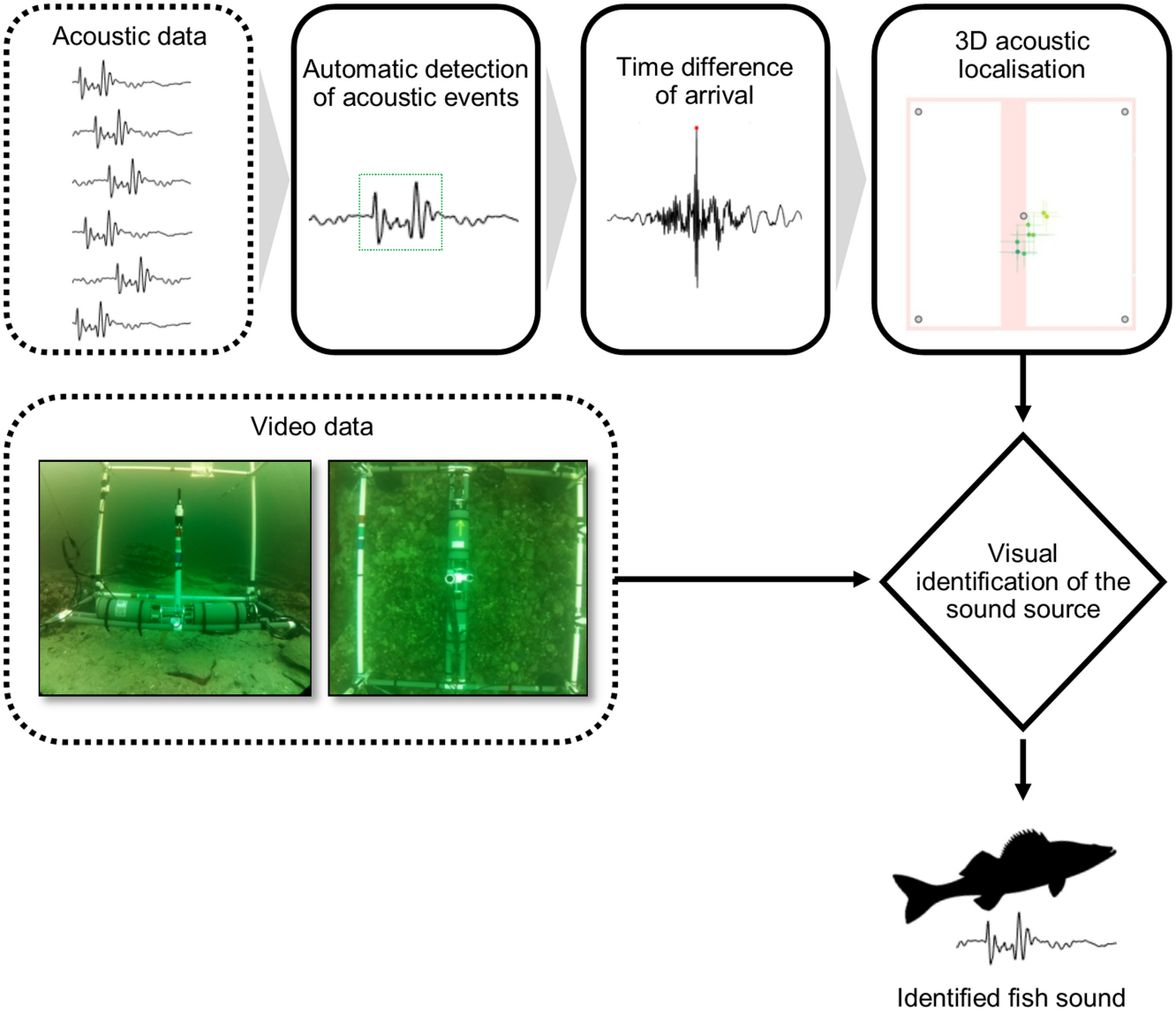

FIGURE 4 Overview of the processing workflow.

and mobile arrays are better suited for localising sound sources at close range (<1 m from the centre of the arrays), while the large array can localise sources accurately at a much greater range.

2.5 | Characterisation of identified fish sounds

Identified fish sounds are characterised by measuring their pulse frequency, pulse repetition rate and duration, where a pulse is defined as a positive/negative amplitude pair. Each fish sound is typically made up of one or more pulses. All measurements are performed on the waveform as in Casaretto et al., 2015 (Figure 7). The pulse duration, T_{pulse}, is measured as the time separating the two first consecutive amplitude peaks of a pulse, and is used to calculate the pulse frequency in hertz (i.e. $1 / T_{pulse}$). The pulse repetition interval, T_{rep} (also referred to as pulse interval in Casaretto et al., 2015), is measured as the duration between the first peak of two consecutive pulses and is used to calculate the pulse repetition rate in pulses per second (i.e. $1 / T_{rep}$). The duration, T_{dur}, is the duration of the fish sound in seconds. There is only one duration measurement per fish sound. However, fish sounds with multiple pulses (typically grunts) have several measurements of pulse frequency and pulse repetition rate (e.g. Figure 7).

2.6 | Estimation of source levels

The acoustic source levels are calculated for the localised fish sounds by applying estimated propagation loss values to the received levels using (Urick, 1983),

$$SL = RL + PL, \quad (1)$$

where SL is source level, RL is received level and PL is the propagation loss (all in dB re 1μPa). Received levels are calculated for each fish

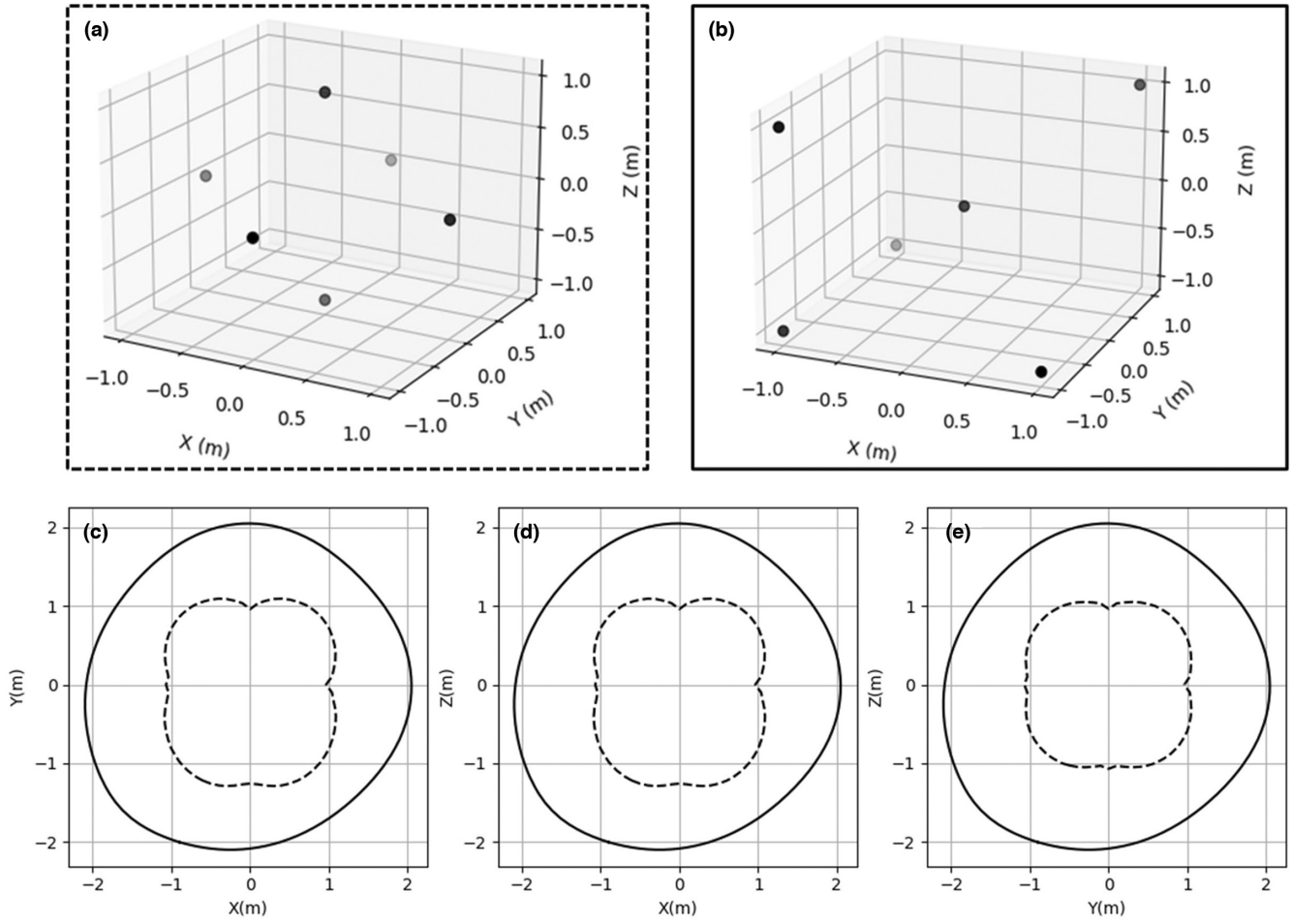

FIGURE 5 Comparison of localisation uncertainties between the large array from this study and the array used by Mouy et al. (2018). (a) Hydrophone geometry used by Mouy et al. (2018). (b) Hydrophone geometry of the large array as defined by the simulated annealing optimisation process. Estimated 50-cm localisation uncertainty isoline of the large array (solid line) and that of Mouy et al. (2018) (dashed line) in the (c) *XY*, (d) *XZ* and (e) *YZ* plane. Uncertainties for each array were computed on a 3D grid of 3 m^3 with points located every 10 cm, and using a standard deviation of data errors of 0.12 ms.

sound after converting the amplitude of the digitised signal, $x(t)$, into acoustic pressure values using

$$P(t) = 10^{-S_g/20} x(t), \tag{2}$$

where S_g is the system gain, in dB re FS/μPa, measured in the calibrations described in Section 2.1. The root-mean-square (RMS) received sound pressure level is defined (in dB re 1 μPa) as

$$RL = 20\ \log_{10}\left(\sqrt{\frac{1}{T_{\text{dur}}}\int_{T_{\text{dur}}} P^2(t)dt}\right). \tag{3}$$

Source levels are calculated by assuming spherical spreading of the acoustic wave. Additionally, given the short distance between the hydrophones and the fish and the low frequency of fish sounds, absorption losses are considered negligible. Therefore, the propagation loss in Equation 1 is defined by

$$PL = 20\ \log_{10}(R), \tag{4}$$

where R is the distance in meters between the source (i.e. the localised fish) and the receiver (hydrophone). Source levels are estimated using data from the hydrophone closest to the fish location and by band-pass filtering the acoustic recording in the frequency band of the fish sound (fourth-order Butterworth filter).

2.7 | Estimation of detection ranges

Estimating detection range is key in designing passive acoustic monitoring programs as it helps to (1) define the distance over which fish sounds can be detected, (2) determine how many recorders are required for the area of interest and (3) assess if passive acoustic monitoring is suitable for an area, given its ambient noise conditions. If we assume that fish sounds with a received level below the ambient noise are not detectable (i.e. detection threshold of 0 dB), and that sound waves spread spherically without absorption, then the maximum distance R_{max} at which fish sounds can be detected is estimated as

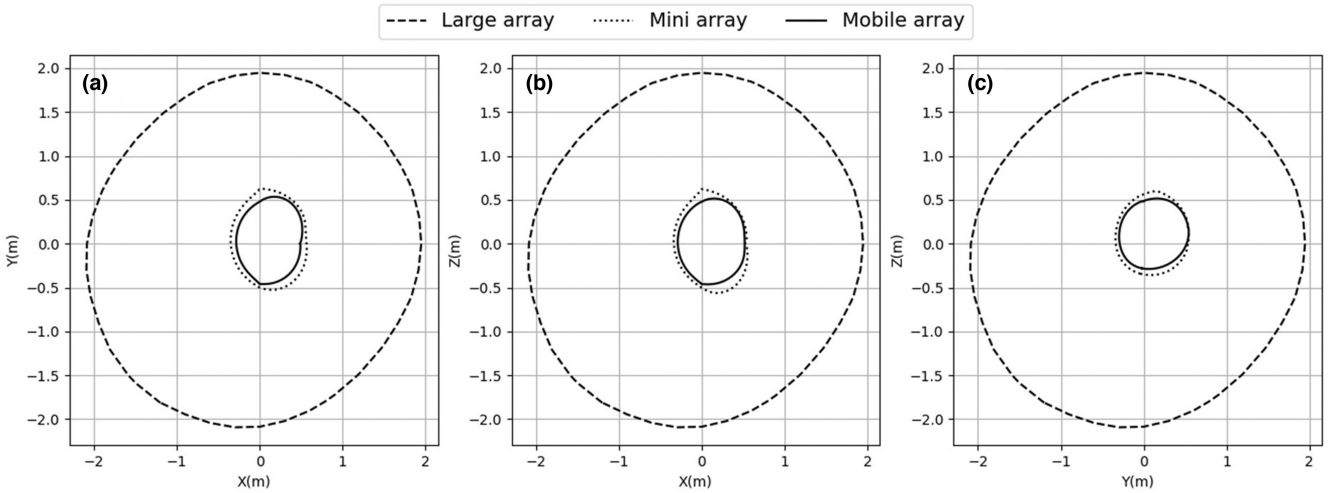

FIGURE 6 Comparison of localisation uncertainties between the three audio-video arrays. Estimated 50-cm localisation uncertainty isoline of the large (dashed line), mini (dotted line) and mobile (solid line) arrays in the (a) *XY*, (b) *XZ* and (c) *YZ* plane. Uncertainties for each array were computed on a 3D grid of 3 m^3 with points located every 1 cm, and using a standard deviation of data errors of 0.12 ms.

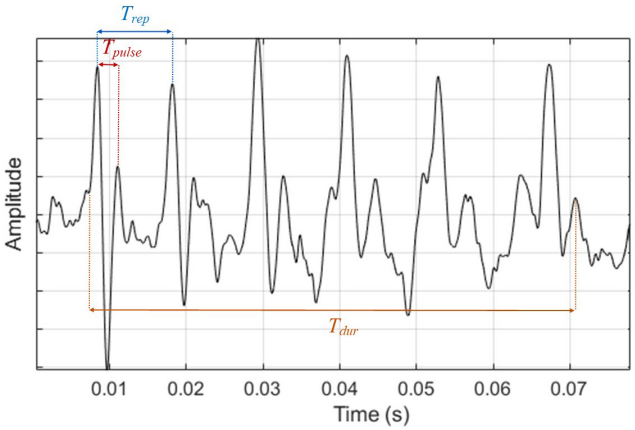

FIGURE 7 Waveform of a fish grunt composed of six pulses. The pulse duration, T_{pulse}, and pulse repetition interval, T_{rep}, are measured on the waveform representation of the fish sound to calculate the pulse frequency and pulse repetition rate, respectively. T_{dur} represents the duration of the sound.

$$R_{max} = 10^{\frac{SL-NL}{20}}, \quad (5)$$

where *SL* is the source level, and *NL* is the noise level at the monitoring location.

2.8 | Software implementation

All the detection, localisation and optimisation algorithms described in this paper were implemented in Python 3.8 using the library ecosound (Mouy, 2021), which relies on pandas (McKinney & Team, 2015), NumPy (Harris et al., 2020), scikits-learn (Pedregosa et al., 2011), Dask (Dask Development Team, 2016) and xarray (Hoyer & Hamman, 2017). Jupyter Notebooks allowing the reproduction of the results from this study are on the GitHub repository of this paper (see Data Availability Statement section).

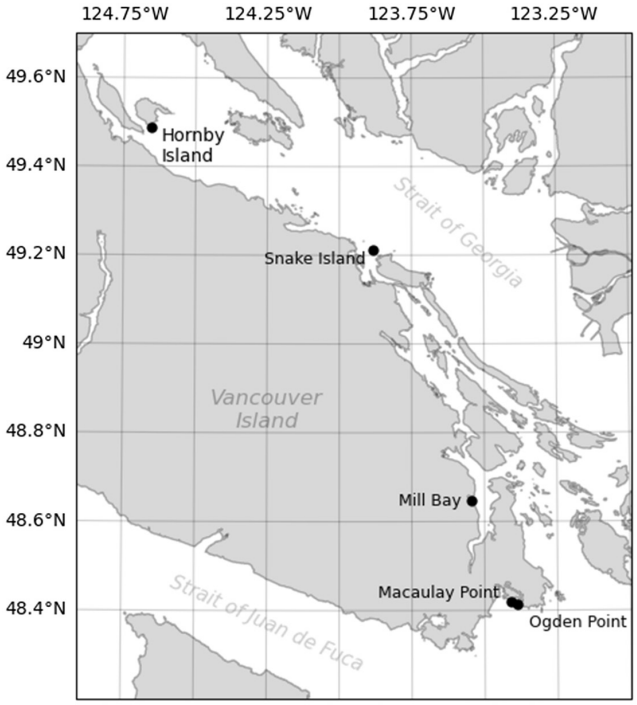

FIGURE 8 Map of the deployment locations. Black dots indicate the locations where the audio-video arrays were deployed.

2.9 | Data collection in the field

The large, mini and mobile arrays were deployed at five sites off the east coast of Vancouver Island, British Columbia, Canada (Figure 8, Table 1). These sites were selected to cover a variety of habitats and fish species. Ogden Point is a well known SCUBA diving and shore

TABLE 1 Deployment locations, depths and durations.

#	Location	Depth (m)	Array type	Deployment date	Deployment duration
1	Ogden Pt.	10	Large	2019-05-04	7 days
2	Ogden Pt.	10	Large	2019-06-15	7 days
3	Mill Bay	9	Mini	2019-07-29	7 days
4	Mill Bay	9	Large	2019-08-18	14 days
5	Hornby Isl.	8	Large	2019-09-15	7 days
6	Hornby Isl.	3–20	Mobile	2019-09-16	2 h
7	Hornby Isl.	3–20	Mobile	2019-09-16	2 h
8	Hornby Isl.	3–20	Mobile	2019-09-18	1 h
9	Hornby Isl.	3–20	Mobile	2019-09-18	2 h
10	Hornby Isl.	3–20	Mobile	2019-09-19	2.5 h
11	Hornby Isl.	3–20	Mobile	2019-09-20	1.5 h
12	Hornby Isl.	3–20	Mobile	2019-09-20	2 h
13	Snake Isl.	11	Large	2019-11-28	8 days
14	Macaulay Pt.	3	Mobile	2020-09-10	1.5 h

fishing site located near a breakwater with over 30 different fish species (iNaturalist research-grade observations, INaturalist, 2021). The Mill Bay site is located near the ship wreck of the *M/V Lord Jim* which attracts a number of fish species. Hornby Island and Snake Island were selected because previous passive acoustic surveys reported that unknown fish sounds were recorded there (Nikolich et al., 2016; Dana Haggarty pers. comm.). Additionally, the Snake Island site is located inside the Northumberland Channel Rockfish Conservation Area (Thornborough et al., 2020). The Macaulay Point site was only used for testing the acoustic localisation of the mobile array (see Supporting Information). The Trident underwater ROV was used prior to each deployment to inspect the seafloor and select a suitable location for the audio-video array.

The large array was deployed five times at four sites between May and December 2019 for durations of 1 to 2 weeks. For each deployment, the array was assembled on shore, loaded on a 7 m Lifetimer, aluminium hull boat (*M/V Liber Ero*) and, once on site, lowered to the seafloor with two temporary lines attached to lifting thimbles at the top of the array frame. Divers secured each leg of the PVC frame with sand bags (total ballast: 90 kg on land) and attached the side camera (C2 in Figure 1) to the main frame. Deployment operations for the large array required a boat driver, a dive tender, a deck-hand and three SCUBA divers.

The mini array was deployed once at Mill Bay, in August 2019, for 14 days. The array was assembled on shore and deployed from a 3.3 m × 1.5 m inflatable boat (Intex Mariner 4). Divers lowered the mini array to the seafloor by hand and secured it in place with sand bags (total ballast: 40 kg on land). Deployment operations for the mini array required a boat driver, a dive tender and two SCUBA divers.

The mobile array was deployed seven times between September 16 and 20, 2019, off Hornby Island and once at Macaulay Point in September 2020. The duration of each deployment was between 1 and 2.5 h and was dictated by the battery life of the Trident underwater ROV. The mobile array was deployed and piloted from a 3.3 m × 1.5 m inflatable boat (Intex Mariner 4). Deployments typically consisted of (1) anchoring the boat with a small weight, (2) deploying the mobile array in the water, (3) piloting the ROV to explore the area to find fish and (4) land the array on the seafloor or on rocks to quietly record fish sounds (see video SuppInfo_Video_MobileArrayDeployment.mp4 in the Supporting Information). This sequence was repeated several times for each deployment. The choice of locations to land the array was made based on the video footage only as the acoustic data are not accessible in real-time. The deployment at Macaulay Point was performed from a public dock and therefore did not require a boat. Deployment operations for the mobile array required a single person to drive the boat (when the measurements were not done from shore) and pilot the underwater ROV.

For all three arrays, data were downloaded after recovery of the instruments and processed in post analysis. Fieldwork operations did not require any licences or permits.

3 | RESULTS

This section shows examples of fish sounds that were identified in the field using each platform. Videos corresponding to Figures 9–14 can be found in the Supporting Information.

3.1 | Identification of fish sounds using the large array

Figure 9 shows the acoustic localisation of fish sounds over a 10-s period when a lingcod *Ophiodon elongatus* swam inside the array from the left (Figure 9d) and stopped at the centre of the array below the PVC tubes holding hydrophone 4 (Figure 9e). The detected fish

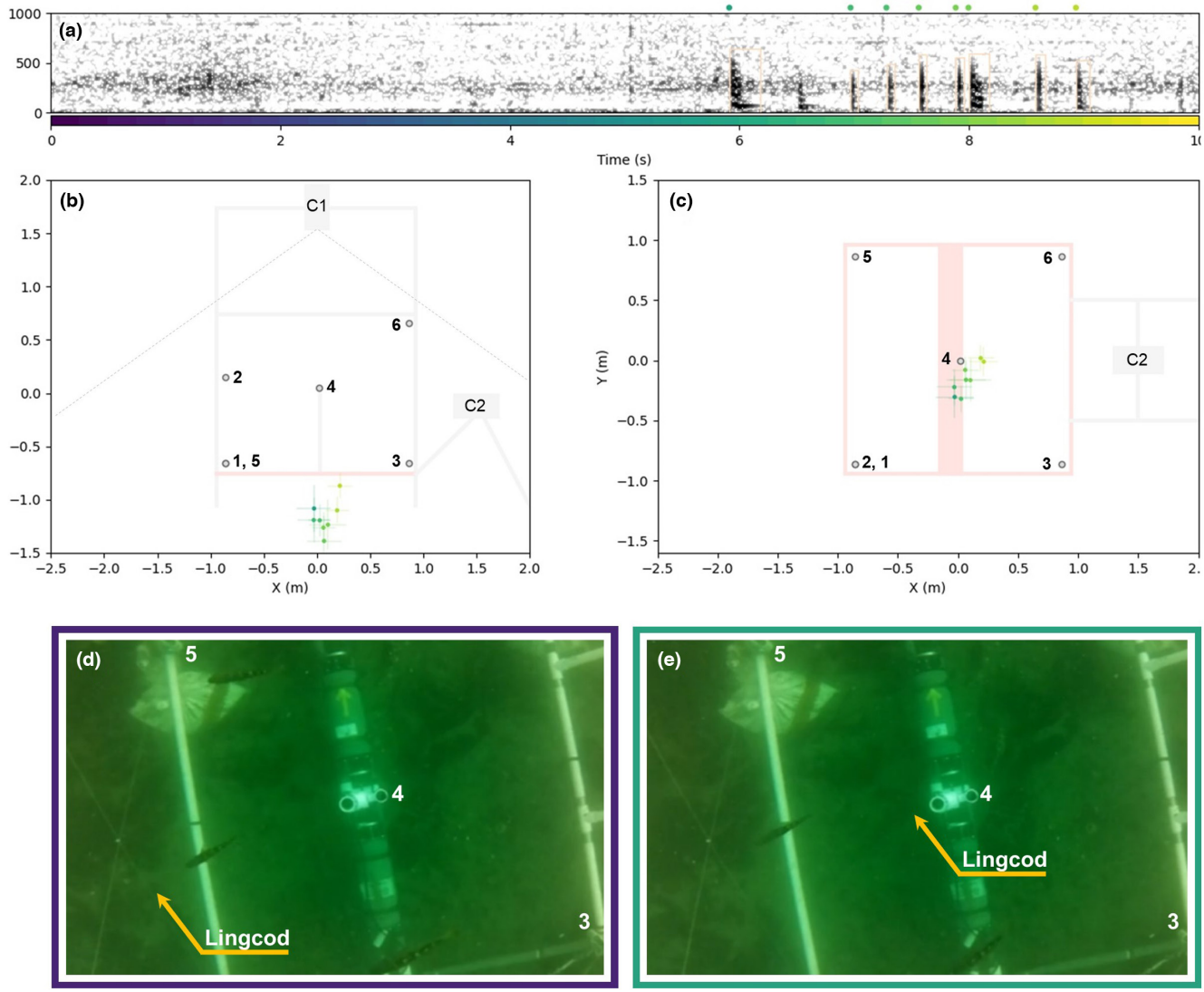

FIGURE 9 Identification of sounds from lingcod using the large array deployed at Ogden Point (17 Jun. 2019). (a) Spectrogram of the acoustic recording acquired by Hydrophone 4 (frame: 0.0624 s, FFT: 0.0853 s, step size: 0.01 s, Hanning window). Beige boxes indicate the time and frequency limits of the fish sounds that were automatically detected. Dots at the top of the spectrogram indicate the colours associated to the start time of each detection (see colour scale on the *x*-axis) and used for the localisation. Coloured camera icons indicate the time of the camera frames shown in (d) and (e). (b) Side and (c) top view of the large array. Coloured dots and lines represent the coordinates and uncertainty (standard deviation) of the acoustic localisations, respectively. (d) Image taken by video camera C1 at $t = 0.5$ s, and (e) $t = 6$ s, showing a lingcod entering the array from the left and stopping at the centre of the array on the seafloor below hydrophone 4. Bold numbers in (b), (c), (d) and (e) correspond to the hydrophone identification numbers. Dashed grey lines in panel (b) indicate the field of view of camera C1. Note that images in panels (d) and (e) have been cropped from the original to focus on the centre of the array. See video SuppInfo_Video_Figure9.mp4 in the Supporting Information.

sounds were localised on the seafloor near the centre of the array (i.e. below hydrophone 4), with a localisation uncertainty less than 20 cm for all dimensions and corresponded to the location of the lingcod (Figure 9b,c). We conclude that fish sounds recorded were therefore emitted by the lingcod.

Figure 10 shows the acoustic localisation of fish sounds over a 17-s period when a lingcod and three quillback rockfish *Sebastes maliger* are in the field of view of video camera C1. The sequence of actions for this recording was as follows. The lingcod, initially located on top of the acoustic recorder, started to move ($t = 2$ s) and relocated itself on top of the battery pack ($t = 10$ s). In response to the lincod's movement, quillback rockfish # 1, which was located near the centre of the array, abruptly swam to the left ($t = 1$ s), went outside the array frame and slowly swam towards Hydrophone 5 near quillback rockfish # 3 ($t = 14$ s). Quillback rockfish # 2 entered the top of the array ($t = 1$ s) and steadily swam to the right towards video camera C2. The first sequence of fish sounds (purple dots in Figure 10) were localised at the location of quillback rockfish # 1 as it swam away from the lingcod. The second sequence of fish sounds (green and yellow dots) were localised on the left side of the array where quillback rockfish # 1 and # 3 were

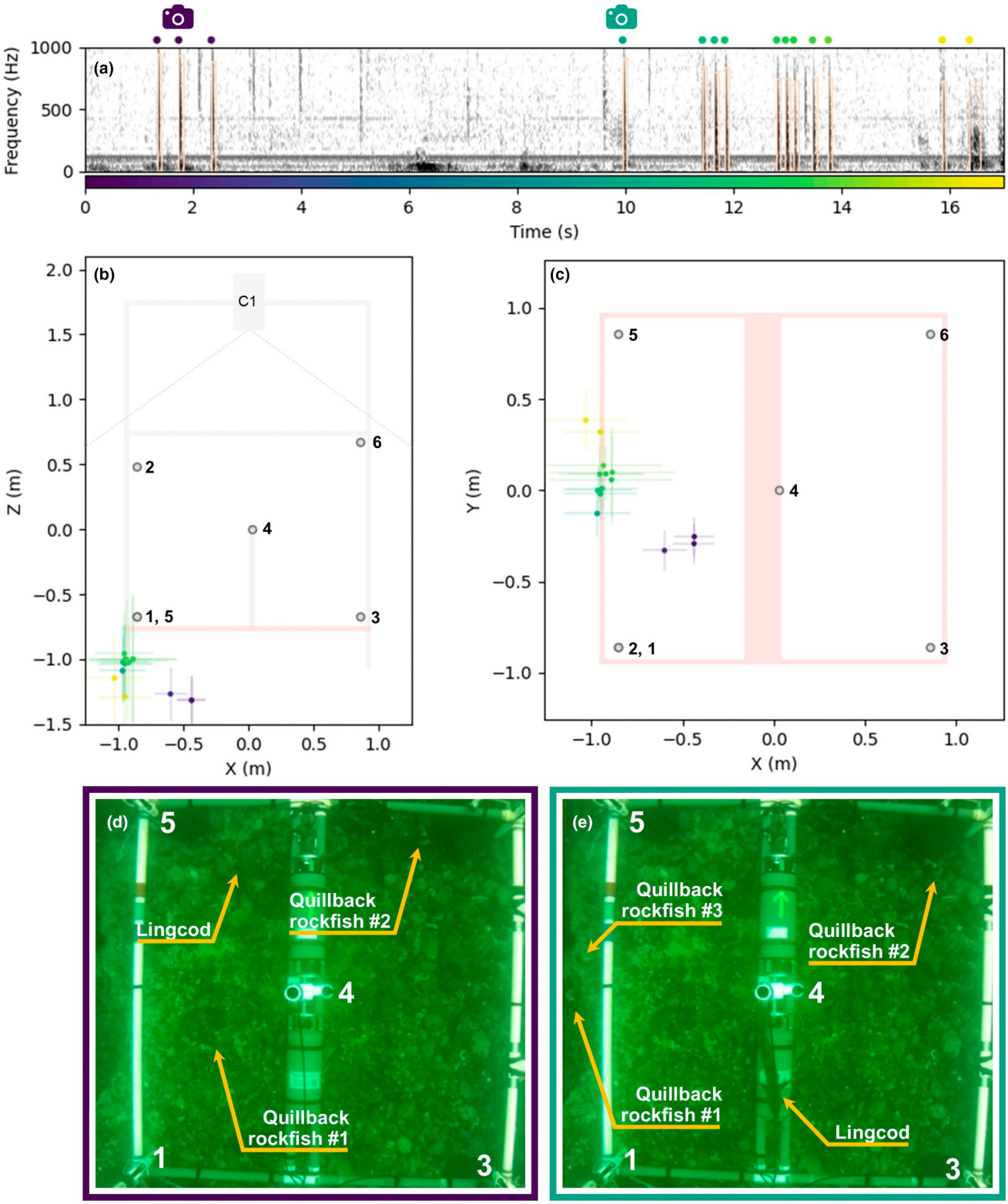

FIGURE 10 Identification of sounds from quillback rockfish using the large array deployed at Hornby Island (20 Sep. 2019). (a) Spectrogram of the acoustic recording acquired by hydrophone 4 (frame: 0.0624 s, FFT: 0.0853 s, step size: 0.01 s, Hanning window). Beige boxes indicate the time and frequency limits of the fish sounds that were automatically detected. Dots at the top of the spectrogram indicate the colours associated to the start time of each detection (see colour scale on the x-axis) and used for the localisation. Coloured camera icons indicate the time of the camera frames showed in panels (d) and (e). (b) Side and (c) top view of the large array. Coloured dots and lines represent the coordinates and uncertainty (standard deviation) of the acoustic localisations, respectively. (d) Image taken by video camera C1 at $t = 2$ s, and (e) $t = 10$ s, showing a lingcod and three quillback rockfish near or inside the array. Bold numbers in (b–e) correspond to the hydrophone identification numbers. Dashed grey lines in (b) indicate the field of view of camera C1. Note that images in (d) and (e) have been cropped from the original to focus on the centre of the array. See video SuppInfo_Video_Figure10.mp4 in the Supporting Information.

located. Localisation uncertainties were less than 20 cm inside the array and less than 40 cm outside the array, which leaves no ambiguity that the fish sounds were produced by the quillback rockfish.

3.2 | Identification of fish sounds using the mini array

Figure 11 shows the acoustic localisation of two fish grunts while a copper rockfish *Sebastes caurinus* was sitting on top of the mini array. While only the tail of the fish was visible in the video when the fish sounds were emitted (Figure 11e), the manual inspection of the video data several minutes prior to the sounds confirmed the fish was a copper rockfish (Figure 11d). Both fish grunts (at $t = 0.5$ s and $t = 0.8$ s) are localised at about 30 cm of the top right of the camera with uncertainties in each dimension less than 4 cm. From these localisation results, and given an approximate fish length of 30 cm, we conclude that the fish grunts were emitted by the copper rockfish sitting on top of the array.

3.3 | Identification of fish sounds using the mobile array

Figure 12 shows the acoustic localisation of five impulsive fish sounds while two copper rockfish were located in front of the mobile array. All sounds were localised at the front right of the array, near the seafloor (Figure 12b,c). Localisation uncertainties were less than 15 cm along the *x* axis, and 10 cm along the *z* axis. Despite the greater localisation uncertainties in range (i.e., > 30 cm along the *y* axis), the absence of other fish in the video within the boundaries of the localisation uncertainties confirms that the fish sounds in Figure 12a were produced by the copper rockfish in front of the mobile array. These impulsive sounds seemed to be associated with an agonistic behaviour.

Figure 13 shows the acoustic localisation of six impulsive fish sounds while a copper rockfish and a blackeye goby *Rhinogobiops nicholsii* were located in front of the mobile array. All sounds were localised at the front of the array, near the seafloor, and had localisation uncertainties less than 5 cm and 10 cm along the *x* and *z* axes, respectively, and up to 20 cm along the *y* axis (Figure 12b,c). Given the proximity of the two fish and the larger localisation uncertainties in range, it is not possible to identify with certainty which fish produced the sounds.

Figure 14 shows the acoustic localisation of fish sounds while a blackeye goby was in the front of the mobile array and was chasing away other fish intruders from its shelter under a rock. All fish sounds were localised to the rear left of the mobile array (Figure 12b,c). Localisation uncertainties were very large in all dimensions (up to 3 m), but did not spatially overlap with the area covered visually by the video camera. Fish sounds recorded by the mobile array were therefore not emitted by the blackeye goby in the front of the array.

3.4 | Description of identified fish sounds

Table 2 indicates the pulse frequency, pulse repetition rate, duration and source levels of the lingcod, quillback rockfish and copper rockfish sounds identified by the audio-video arrays. Source levels were only calculated for sounds localised by the large array (using the hydrophone closest to the sound source).

3.5 | Estimated detection ranges

Table 3 shows the estimated detection ranges at the Mill Bay and Hornby Island locations where the large audio-video array was deployed (Figure 8, Table 1). The calculation was performed using a source level value of 113 dB re 1 μPa (as measured in Table 2) and noise levels measured at the middle hydrophone of the large array between 20 and 1000 Hz (i.e. the frequency band of fish sounds) using the software PAMGuide (Merchant et al., 2015). Given that noise levels constantly fluctuate in time, the detection range was calculated for the minimum (L_{min}) and maximum (L_{max}) noise levels, as well as for the 5th, 50th and 95th percentile levels (L_5, L_{50} and L_{95}, respectively). Detection range values in Table 3 show that at Hornby Island, fish sounds can be detected at up to 33 m under the quietest conditions (L_{min}), but only up to 8 m for half of the time (L_{50}). At Mill Bay, the detection range is less than 11 m under the quietest conditions and less than 2 m for half of the time. During the noisiest events (L_{max}), typically when a boat is passing close by, fish sounds are only detectable over a few centimetres for both locations. These results illustrate how important source level measurements are to assess spatial coverage and define a monitoring plan. The Mill Bay site is near a marina and is noisier than Hornby Island, which results in smaller detection ranges. Consequently, developing a passive acoustic monitoring program for fish would likely require deploying more acoustic recorders at Mill Bay than at Hornby Island. Such analysis may also help define areas where passive acoustics is simply not suitable for conducting fish monitoring due to high noise conditions. Statistics describing the noise levels recorded at Hornby Island and Mill Bay are included in Supporting Information (Supporting_Information.pdf).

4 | DISCUSSION

Our results show that all three audio-video arrays can successfully identify fish sounds in the wild. Field tests using a controlled sound source for each platform ground-truthed the accuracy of the localisation results and confirmed that the instrumentation and analysis process are working correctly (see Supporting Information).

The large array provides the most accurate acoustic localisations (Figure 6) and, with its two video cameras, has the largest field of view. Hydrophone placement for this array was optimised using simulated annealing to minimise localisation uncertainties. This optimisation resulted in a hydrophone geometry that was different than the

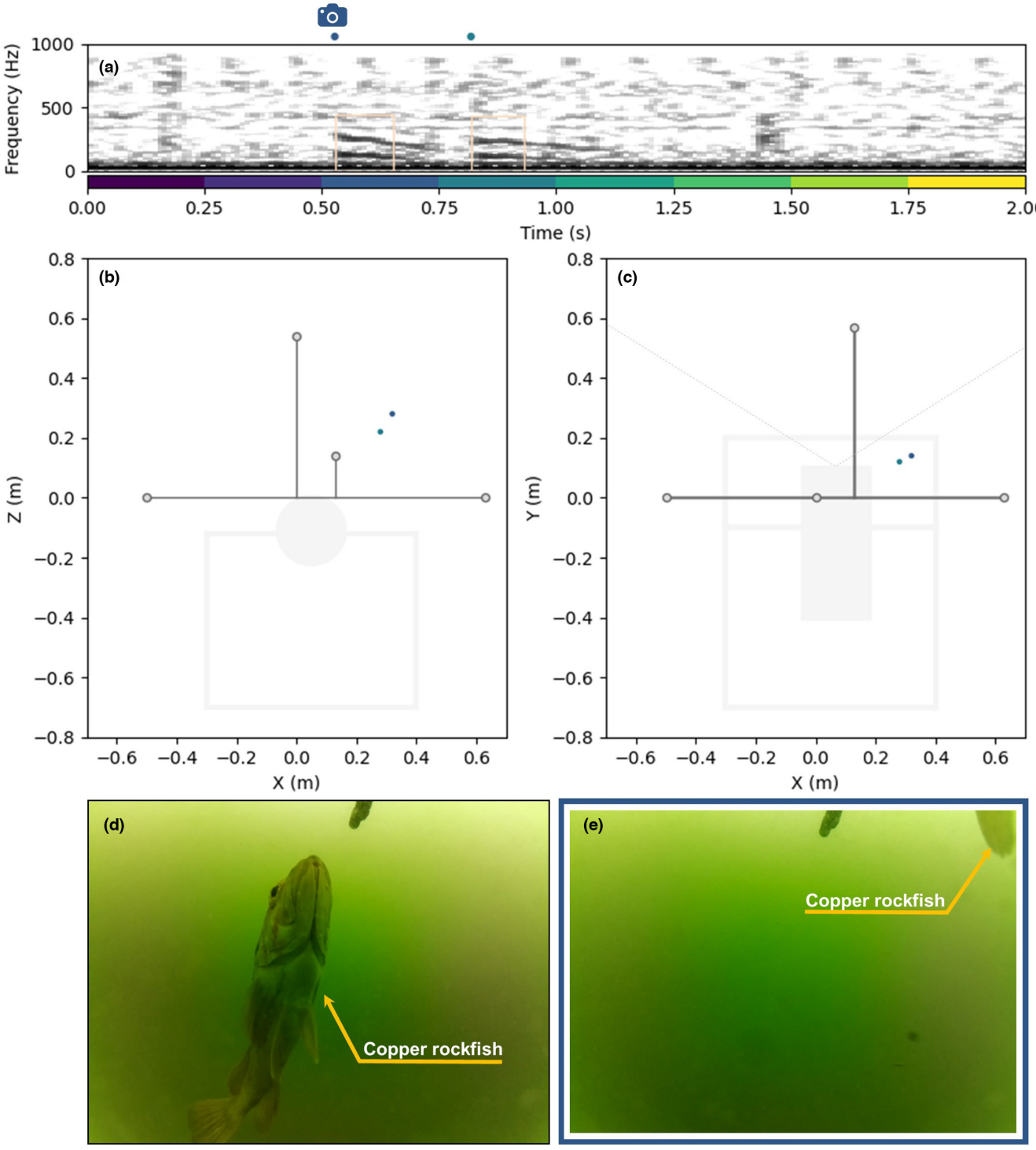

FIGURE 11 Identification of sounds from copper rockfish using the mini array deployed at Mill Bay (1 August 2019). (a) Spectrogram of the acoustic recording acquired by Hydrophone 2 (frame: 0.0624 s, FFT: 0.0853 s, step size: 0.01 s, Hanning window). Beige boxes indicate the time and frequency limits of the fish sounds that were automatically detected. Dots at the top of the spectrogram indicate the colours associated to the start time of each detection (see colour scale on the *x*-axis) and used for the localisation. The blue camera icon indicates the time of the camera frame showed in (e). (b) Rear and (c) top view of the mini array. Coloured dots and lines represent the coordinates and uncertainty (68% credibility interval) of the acoustic localisations, respectively. Dashed grey lines indicate the field of view of the camera. (d) Image from the video camera, taken several minutes before the fish sounds were emitted, showing the copper rockfish before it sat on top of the mini array. (e) Image taken from the video camera of the mini array at $t = 0.5$ s, showing the tail of the copper rockfish sitting on top of the mini array. See video SuppInfo_Video_Figure11.mp4 in the Supporting Information.

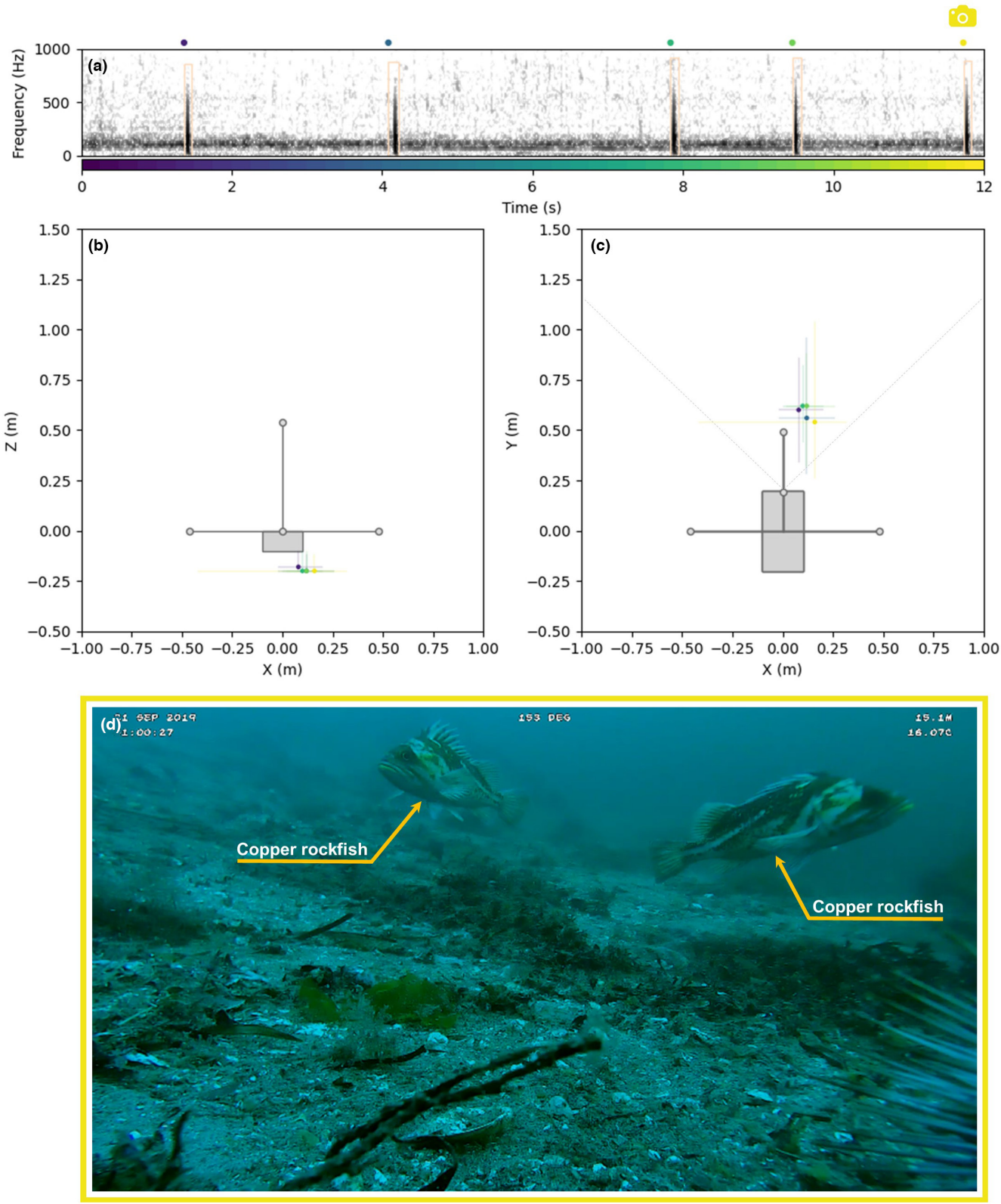

FIGURE 12 Identification of sounds from copper rockfish using the mobile array deployed at Hornby Island (21 September 2019). (a) Spectrogram of the acoustic recording acquired by hydrophone 2 (frame: 0.0624 s, FFT: 0.0853 s, step size: 0.01 s, Hanning window). Beige boxes indicate the time and frequency limits of the fish sounds that were automatically detected. Dots at the top of the spectrogram indicate the colours associated to the start time of each detection (see colour scale on the *x*-axis) and used for the localisation. The yellow camera icon indicates the time of the camera frame showed in panel (d). (b) Rear and (c) top view of the mobile array. Dashed grey lines in (c) indicate the field of view of the camera. Coloured dots and lines represent the coordinates and uncertainty (68% credibility interval) of the acoustic localisations, respectively. (d) Image from the underwater ROV's video camera and taken at $t = 11.9$ s, showing two copper rockfish at the front of the mobile array. See video SuppInfo_Video_Figure12.mp4 in the Supporting Information.

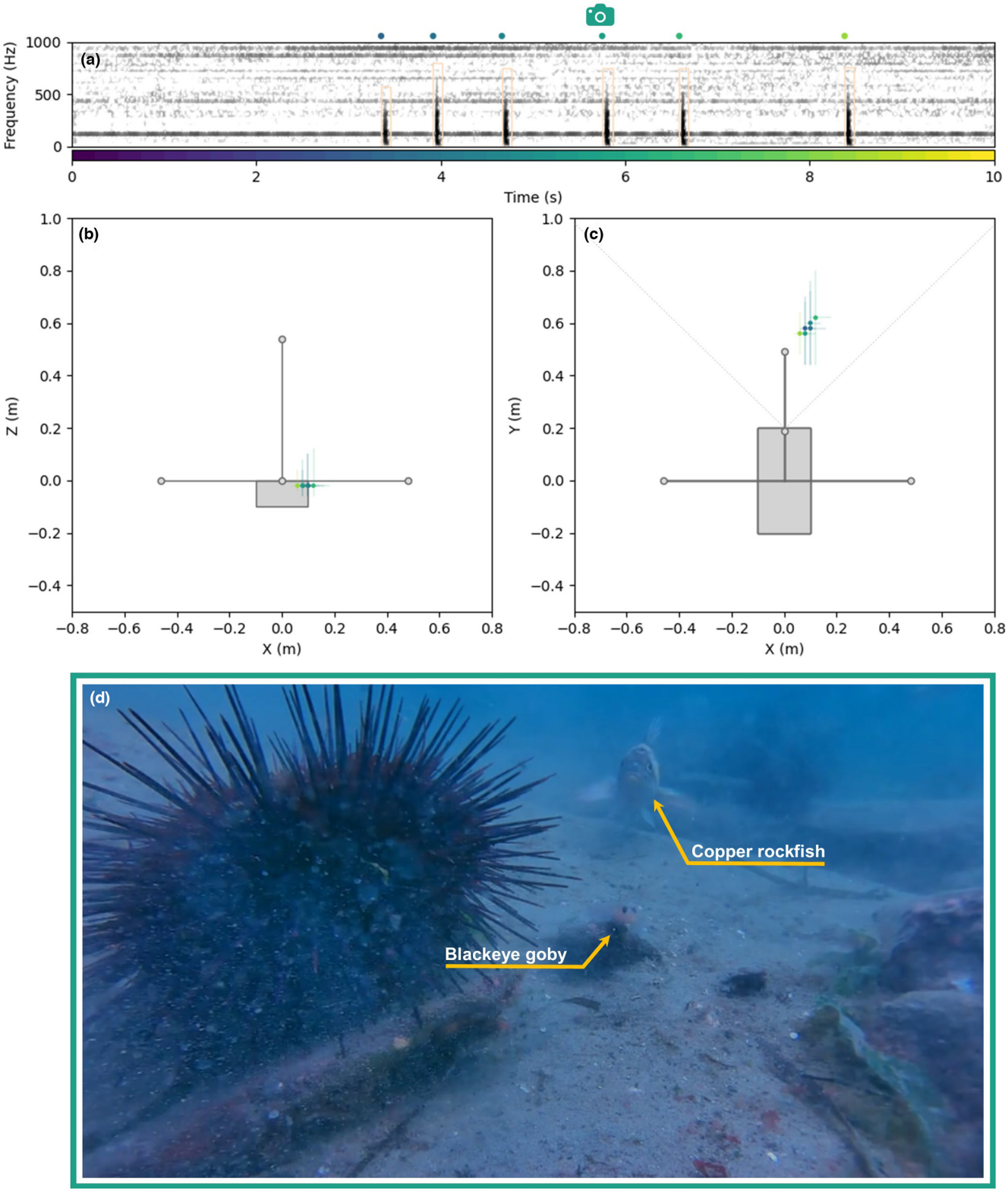

FIGURE 13 Localisation of unknown fish sounds using the mobile array deployed at Hornby Island (18 September 2019). (a) Spectrogram of the acoustic recording acquired by hydrophone 2 (frame: 0.0624 s, FFT: 0.0853 s, step size: 0.01 s, Hanning window). Beige boxes indicate the time and frequency limits of the fish sounds that were automatically detected. Dots at the top of the spectrogram indicate the colours associated to the start time of each detection (see colour scale on the *x*-axis) and used for the localisation. The green camera icon indicates the time of the camera frame showed in panel (d). (b) Rear and (c) top view of the mobile array. Dashed grey lines in panel (c) indicate the field of view of camera. Coloured dots and lines represent the coordinates and uncertainty (68% credibility interval) of the acoustic localisations, respectively. (d) Image from the underwater ROV's video camera taken at $t = 5.9$ s, showing a blackeye goby and a copper rockfish in front of the mobile array. See video SuppInfo_Video_Figure13.mp4 in the Supporting Information.

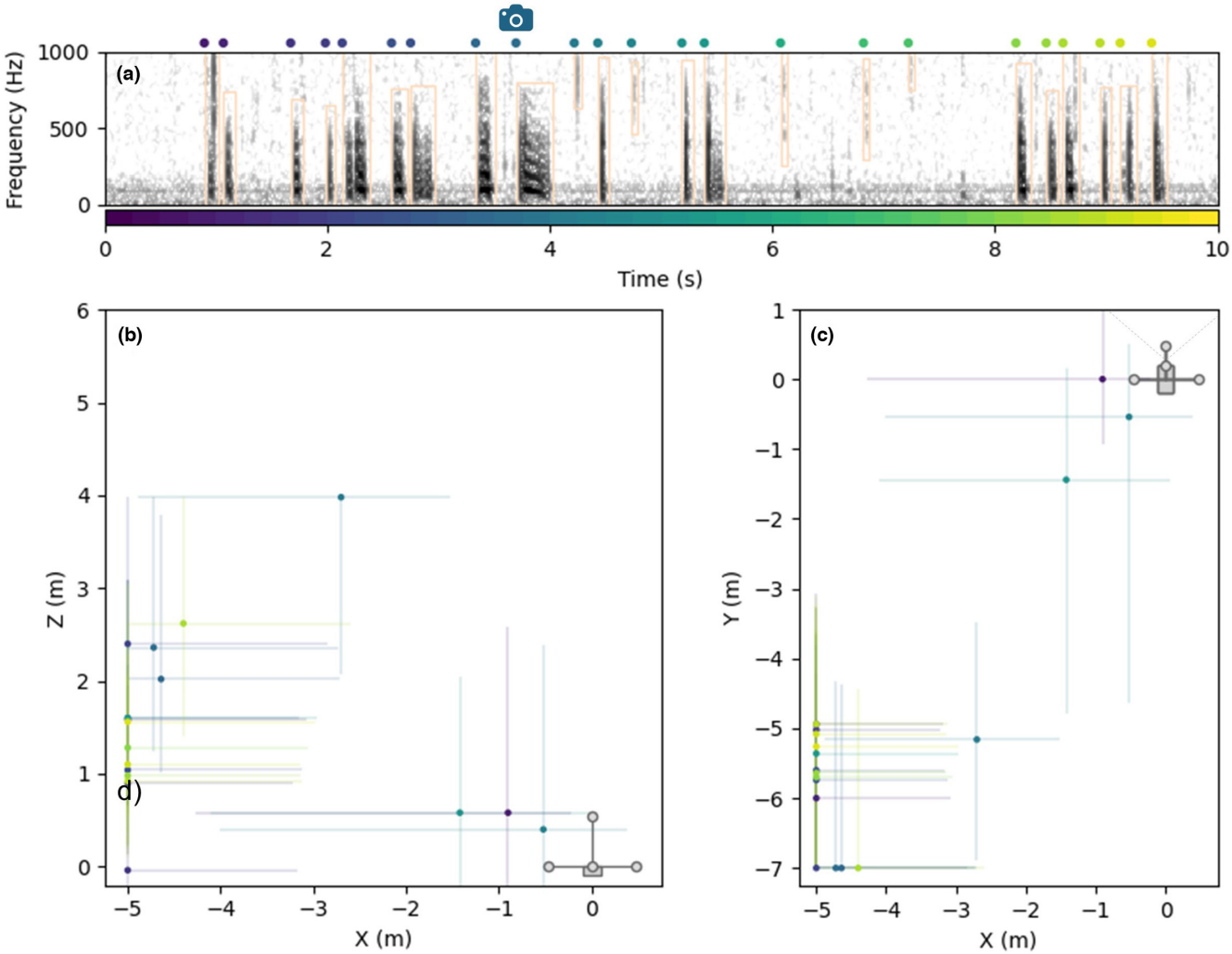

FIGURE 14 Localisation of unknown fish sounds using the mobile array deployed at Hornby Island (16 September 2019). (a) Spectrogram of the acoustic recording acquired by hydrophone 2 (frame: 0.0624 s, FFT: 0.0853 s, step size: 0.01 s, Hanning window). Beige boxes indicate the time and frequency limits of the fish sounds that were automatically detected. Dots at the top of the spectrogram indicate the colours associated to the start time of each detection (see colour scale on the *x*-axis) and used for the localisation. The turquoise camera icon indicates the time of the camera frame shown in (d). (b) Rear and (c) top view of the mobile array. Dashed grey lines in (c) indicate the field of view of the camera. Coloured dots and lines represent the coordinates and uncertainty (68% credibility interval) of the acoustic localisations, respectively. (d) Image from the underwater ROV's video camera taken at $t = 3.8$ s, showing a blackeye goby in front of the mobile array, defending its territory. See video SuppInfo_Video_Figure14.mp4 in the Supporting Information.

TABLE 2 Characteristics of the fish sounds identified. Duration, pulse frequency, pulse repetition rate, and source level are reported by their mean ± standard deviation. Asterisks (*) indicate measurements for which there were not enough samples of fish sounds (*n*) to estimate the standard deviation.

Species	Sound type	Dur. (ms)	Pulse freq. (Hz)	Pulse rep. Rate (pulses/s)	Source level (dB re 1 μPa)	*n*
Lingcod	Grunt	170*	301 ± 32	84.5 ± 10.5	115.4*	2
Lingcod	Pulse	19 ± 4	318 ± 13	N/A	113.0 ± 3.5	6
Quillback rockfish	Pulse	9 ± 1	387 ± 77	N/A	113.5 ± 2.0	13
Copper rockfish	Pulse	23 ± 2	163 ± 2	N/A	-	5

one used by Mouy et al. (2018) to localise fish sounds off Cape Cod. Using this hydrophone configuration increased by a factor of about seven the spatial localisation capacity of the array over that used by Mouy et al. (2018) (for the same array volume) and allowed fish to be localised accurately (i.e., localisation uncertainty < 50 cm) at up to 3 m from the centre of the array (Figure 6). The mini and mobile arrays have much smaller footprints and most of the sound sources to localise are outside the array. Consequently, small errors in the measurement of the time-difference of arrival lead to large errors in the localisation results. Nevertheless, these two arrays are capable of determining the bearing and elevation of the sound source, which in many cases, is enough to confirm that the sounds recorded are emitted by the fish in front of the camera (e.g. Figure 12). In some circumstances, when several fish are located along the same bearing angle from the mini or mobile arrays, the larger localisation uncertainties in range do not allow definitive identification of the fish producing the sound (e.g., Figure 13). This is typically not an issue if the fish are from the same species.

Although attempting to identify fish sounds in the wild using a single hydrophone and a single camera is relatively inexpensive and logistically easy, our study shows the importance of having several hydrophones (or directional sensors) for performing acoustic localisation. Inferring which individual fish produces the sounds, based only on visual observations from video footage, is prone to errors and can lead to assigning sounds to the wrong fish species. The case presented in Figure 10 is a good example of this, where four fish are interacting together and for which the identification of the individual fish emitting the sounds, only based on the video data, could be ambiguous. Based on behaviour only (i.e. without the acoustic localisation), the emitted sounds could have been interpreted either as a territorial display from the lingcod, or as a reaction/defence display from the quillback rockfish. Another example illustrating the usefulness of the acoustic localisation is the recording with the blackeye goby in Figure 14. Given, the clear territorial display of the blackeye goby observed in the video, and the high signal-to-noise ratio of the fish sounds received by the hydrophones, it would have seemed relatively evident to assign the fish sounds to this individual. Performing the acoustic localisation proved that this was not the case. It is therefore critical to perform acoustic localisation when cataloguing fish sounds in a natural setting.

TABLE 3 Estimated detection range (R_{max}) of fish sounds at Mill Bay and Hornby Island. Noise levels (*NL*) were computed in the 20–1000 Hz frequency band every minute by averaging 120 1-s long Hann-windowed fast Fourier transforms overlapped by 0.5 s.

	Mill Bay		Horny Island	
	NL (dB re 1 μPa)	R_{max} (m)	NL (dB re 1 μPa)	R_{max} (m)
L_{min}	92.6	10.5	82.6	33.0
L_5	102.2	3.5	87.2	19.6
L_{50}	107.9	1.8	95.0	7.9
L_{95}	116.1	0.7	105.1	2.5
L_{max}	146.9	0.0	130.1	0.1

In 2006, Rountree et al. (2006) had already identified the need to combine passive acoustic localisation and underwater video to identify fish sounds in the wild. Our work provides three different hardware and software solutions that attempt to fill this research gap and help make passive acoustics a viable technique to monitor fish. Each of the audio-video platforms that we developed have different constraints, strengths and weaknesses. Given their different sizes, all arrays cannot be used in every habitat. The large array requires a relatively flat bottom (at least a 4 m^2 surface) and a water depth of at least 6 m. The mini array can be deployed in more complex habitats with rougher terrain and steeper slopes, and in shallower water (minimum of 3 m depth). The mobile array can cover most types of habitat, from rough to flat seafloor, from very shallow (2 m) to deep (100 m) water depths, and can explore small crevasses where fish can hide in. However, it cannot be used in areas with strong currents due to reduced manoeuvrability. With proper ballast, the large and mini

arrays are less affected by currents. From a logistics perspective, the large array is the most complex to deploy because of its size. The mini array is smaller and therefore much easier to deploy. The mobile array is the easiest platform to deploy as it only requires a single person piloting the underwater ROV. Cost wise, the large array uses six hydrophones and a high-quality multichannel acoustic recorder, which makes it the most expensive platform (~ ϖ40,000 USD). The mini and mobile arrays are less costly (~ ϖ8,000 and ϖ11,000 USD, respectively). In terms of the sampling, both the large and the mini arrays are static platforms that are deployed over several weeks at a time. This long-term duration allows non-intrusive observation and measurement of fish sounds related to a variety of behaviours.

The short home-range of some fish species (Tolimieri et al., 2009), and the static nature of the large and mini arrays, mean these platforms may only sample sounds from a small set of individuals. If this is an issue, carrying out short deployments at different locations may be preferable to performing a series of longer deployments at the same location. The mobile array can sample several individuals over a larger spatial area but can only record for a few hours. The mobile array is more intrusive than the two other arrays, which tends to more often elicit aggressive behaviours. Consequently, the mobile array may sample a more restricted set of acoustic behaviours. As demonstrated in Section 3, all arrays can successfully attribute sounds to individual fish and therefore measure the temporal and frequency characteristics of sound emitted by specific fish species. The large array provides accurate localisation over greater distances (Figure 6), captures a large field of view, and is consequently the preferred platform for estimating source levels. Source levels can also be estimated using the mini and mobile arrays when the fish is close to or inside the array (e.g. Figure 11). However, the larger localisation uncertainty in range for fish located farther away (e.g. Figure 14) or on the side of the array (see section 4.3 in the Supporting Information document) may not allow source levels to be estimated accurately. Figure 15 illustrates the constraints, strengths and weaknesses of each audio-video array. When used in unison, these platforms can cover many different habitats, species, and logistical constraints.

All three audio-video arrays presented in this study can be further developed to broaden the variety of habitats they can be deployed in and to collect longer-term datasets. The deployment duration of the large and mini arrays was limited by the memory limitations of the video camera. Recent field tests using larger memory storage in the FishCam (i.e. 400 GB SD card) show that the autonomy of the array could be extended up to 19 days. For even longer-term observations, it is possible to integrate multichannel acoustic recorders and video cameras to existing cabled observatories where storage and power are not a limitation (Aguzzi et al., 2019; Rountree et al., 2020). Since the beginning of

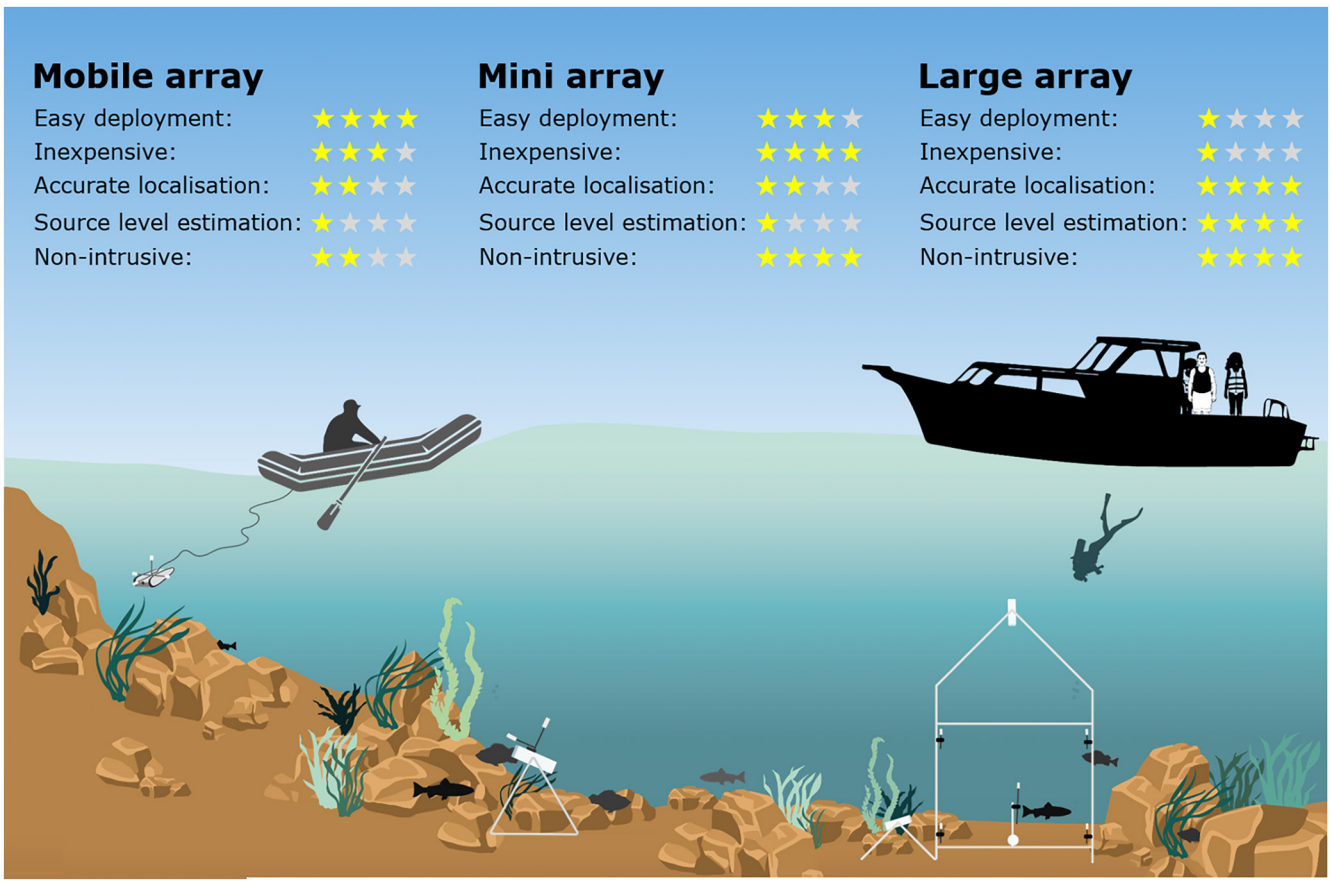

FIGURE 15 Illustration of the different constraints, strengths and weaknesses of each audio-video array proposed in this study.

this study, production of the Trident underwater ROV, which was used for the mobile array, was discontinued. A number of other capable low-cost remotely operated vehicles, such as the open source BlueROV2 (Blue Robotics), are available and could be used for the mobile array. While there is already considerable interest in using passive acoustics for monitoring fish in coastal areas, there is also a growing interest in exploring sounds produced by fish in deeper habitats (Bolgan & Parmentier, 2020; Mann & Jarvis, 2004). Due to the low light conditions in the deep ocean, the design of our arrays would need to be modified to sample in such environments. External LED lights could be added to the FishCam on the array and could be controlled via its onboard single board computer (Mouy et al., 2020). Alternatively, low-light cameras could be used (Pagniello et al., 2021). There is also a strong interest in using passive acoustics for monitoring fish in tropical waters (e.g., coral reefs) where the density of fish and fish sounds is typically much greater than in British Columbia (Looby et al., 2022). In these environments, our audio-video arrays may not be as effective at associating specific sounds to individual fish (due to their close proximity), but they should allow the association of sounds to species. Audio-video arrays with more hydrophones would be recommended in such environments to maximise localisation accuracy. These arrays could also be further developed by improving the data processing workflow. Currently, the audio and video data are only weakly linked. This could be improved by projecting the bearings and associated uncertainties of the localised sounds onto the video data, which would further help in associating sounds to individual fish. In very shallow environments, fish sounds could be received with multi-path reflections from the surface and bottom boundaries. In such case, the measurement of the TDOAs using the full waveform (as we do here) may become inaccurate and degrade the accuracy of the localisation. Consequently, in these environments, it would be preferable to measure TDOAs using just the first amplitude peak of the waveform which is more likely to capture the direct path of the sound.

Lingcod have not previously been documented to produce sound. Kelp greenling *Hexagrammos decagrammus*, which belong to the same family as lingcod *Hexagrammidae* and also lack a swim bladder, have been reported to have muscles possibly responsible for sound production (Hallacher, 1974), but their sounds have not been recorded. A number of rockfish species have been reported to have sonic muscles (Hallacher, 1974) and some have been documented to produce sounds (Nichols, 2005; Širović et al., 2009; Širović & Demer, 2009). Sounds from quillback and copper rockfish have not been described in the peer-reviewed literature. The source levels we measured (i.e. 113 dB re 1 μPa) are consistent with the source levels measured for other species of rockfish (i.e. 103–113 dB re 1 μPa; Širović & Demer, 2009). Note that the sound measurements presented in this study aim at illustrating the type of information that can be measured with the audio-video arrays. A comprehensive description of the variability of the sounds from each species based on the analysis of the entire dataset we collected will be the object of a future study.

In this paper, we proposed three audio-video array designs and demonstrated that they can be used to successfully identify fish sounds in the wild in a variety of coastal habitats. We also provided detailed building instructions and processing scripts that allow this work to be easily replicated. Our contribution fills a current research gap and will help expand the worldwide fish sound catalogue and therefore make passive acoustics a more viable tool to monitor fish populations.

AUTHOR CONTRIBUTIONS

Xavier Mouy conceived the ideas, designed the methodology, constructed the hardware, analysed the data, and led the writing of this manuscript. Morgan Black, Kieran Cox and Jessica Qualley led the diving operations for all the field measurements and helped with the fish identification. Francis Juanes and Stan Dosso supported this work and provided critical feedback throughout the development of the project. All authors contributed critically to draft versions of this paper and gave final approval for publication.

ACKNOWLEDGEMENTS

Data collection in the field was made possible thanks to all the volunteers from the Juanes Lab at the University of Victoria, Dr. Heloise Frouin-Mouy (NOAA), Dr. Dana Haggarty (DFO), Dr. Sarah Dudas (DFO), and Scott Trivers. We thank James Pilkington (DFO) for providing the acoustic projector, Marie Taureau (Manekala) for producing the illustration in Figure 15, Hugo Mouy for helping with editing the videos in Supporting Information, and Philippe Frouin for helping with the design of the PVC frame for the large array. We also thank Dr. Rodney Rountree for sharing his experience and providing invaluable feedback during this project. National Geographic and Sofar Ocean Technologies provided the Trident ROV through the Science Exploration Education Initiative. All SCUBA divers involved in the deployment and recovery operations were scientific divers certified by the Canadian Association for Underwater Science (CAUS). This research is supported by the Natural Sciences and Engineering Research Council (NSERC) Canadian Healthy Oceans Network (CHONe) and its partners: DFO and INREST (representing the Port of Sept-Iles and City of Sept-Iles), JASCO Applied Sciences, and NSERC. Xavier Mouy was also funded by an NSERC Postgraduate Scholarship and a Mitacs Accelerate Fellowship. Field expenses and equipment costs were funded by an NSERC Discovery grant, the Liber Ero Foundation and CFI/BCKDF Funding.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14095.

DATA AVAILABILITY STATEMENT

Data and processing scripts used in this paper are available on these online repositories:

- **GitHub repository**: https://github.com/xaviermouy/XAV-arrays. Contains python scripts and Jupyter notebooks for reproducing the data processing performed in this study. Also contains examples of field logs and checklists used in the field. A snapshot of the repository is also archived on Zenodo (Mouy, 2023). Licence: BSD-3-Clause.
- **OSF data repository**: https://osf.io/q8dz4. Contains the acoustic data, metadata, and configuration files needed to reproduce the detection and localisation results in Figures 9–14. Also contains higher definition versions of the the videos in the Supporting Information. Licence: Creative Commons Attribution-ShareAlike 4.0.

ORCID

Xavier Mouy https://orcid.org/0000-0003-4938-1214
Kieran Cox https://orcid.org/0000-0001-5626-1048
Stan Dosso https://orcid.org/0000-0003-2384-7370
Francis Juanes https://orcid.org/0000-0001-7397-0014

REFERENCES

Adam, O., & Samaran, F. (2013). *Detection, classification and localization of marine mammals using passive acoustics, 2003–2013: 10 years of international research*. Dirac NGO.

Aguzzi, J., Chatzievangelou, D., Marini, S., Fanelli, E., Danovaro, R., Flögel, S., Lebris, N., Juanes, F., De Leo, F. C., Del Rio, J., Thomsen, L., Costa, C., Riccobene, G., Tamburini, C., Lefevre, D., Gojak, C., Poulain, P.-M., Favali, P., Griffa, A., … Company, J. B. (2019). New high-tech flexible networks for the monitoring of Deep-Sea ecosystems. *Environmental Science & Technology*, *53*(12), 6616–6631. https://doi.org/10.1021/acs.est.9b00409

Amorim, M. C. P., Stratoudakis, Y., & Hawkins, A. D. (2004). Sound production during competitive feeding in the grey gurnard. *Journal of Fish Biology*, *65*(1), 182–194. https://doi.org/10.1111/j.0022-1112.2004.00443.x

Amorim, M. C. P., Vasconcelos, R. O., & Fonseca, P. J. (2015). Fish sounds and mate choice. In F. Ladich (Ed.), *Sound commun. Fishes* (pp. 1–33). Springer. https://doi.org/10.1007/978-3-7091-1846-7_1

Bass, A. H., & Ladich, F. (2008). Vocal–acoustic communication: From neurons to behavior. In J. F. Webb, R. R. Fay, & A. N. Popper (Eds.), *Fish bioacoustics* (pp. 253–278). Springer. https://doi.org/10.1007/978-0-387-73029-5_8

Bolgan, M., O'Brien, J., Chorazyczewska, E., Winfield, I. J., McCullough, P., & Gammell, M. (2017). The soundscape of Arctic Charr spawning grounds in lotic and lentic environments: Can passive acoustic monitoring be used to detect spawning activities? *Bioacoustics*, *27*(1), 57–85. https://doi.org/10.1080/09524622.2017.1286262

Bolgan, M., & Parmentier, E. (2020). The unexploited potential of listening to deep-sea fish. *Fish and Fisheries*, *21*(6), 1238–1252. https://doi.org/10.1111/faf.12493

Casaretto, L., Picciulin, M., & Hawkins, A. D. (2015). Seasonal patterns and individual differences in the calls of male haddock *Melanogrammus aeglefinus*. *Journal of Fish Biology*, *87*(3), 579–603. https://doi.org/10.1111/jfb.12740

Cott, P. A., Hawkins, A. D., Zeddies, D., Martin, B., Johnston, T. A., Reist, J. D., Gunn, J. M., & Higgs, D. M. (2014). Song of the burbot: Under-ice acoustic signaling by a freshwater gadoid fish. *Journal of Great Lakes Research*, *40*(2), 435–440. https://doi.org/10.1016/j.jglr.2014.02.017

Dask Development Team. (2016). Dask: Library for dynamic task scheduling. https://dask.org

Davis, G. E., Baumgartner, M. F., Bonnell, J. M., Bell, J., Berchok, C., Bort Thornton, J., Brault, S., Buchanan, G., Charif, R. A., Cholewiak, D., Clark, C. W., Corkeron, P., Delarue, J., Dudzinski, K., Hatch, L., Hildebrand, J., Hodge, L., Klinck, H., Kraus, S., … Van Parijs, S. M. (2017). Longterm passive acoustic recordings track the changing distribution of North Atlantic right whales (*Eubalaena glacialis*) from 2004 to 2014. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-13359-3

Di Iorio, L., Raick, X., Parmentier, E., Boissery, P., Valentini-Poirier, C. A., & Gervaise, C. (2018). 'Posidonia meadows calling': A ubiquitous fish sound with monitoring potential. *Remote Sensing in Ecology and Conservation*, *4*, 248–263. https://doi.org/10.1002/rse2.72

Dosso, S. E., & Sotirin, B. J. (1999). Optimal array element localization. *The Journal of the Acoustical Society of America*, *106*(6), 3445–3459. https://doi.org/10.1121/1.428198

D'Spain, G. L., & Batchelor, H. H. (2006). Observations of biological choruses in the Southern California bight: A chorus at midfrequencies. *The Journal of the Acoustical Society of America*, *120*(4), 1942–1955. https://doi.org/10.1121/1.2338802

Ferguson, B. G., & Cleary, J. L. (2001). *In situ* source level and source position estimates of biological transient signals produced by snapping shrimp in an underwater environment. *The Journal of the Acoustical Society of America*, *109*(6), 3031–3037. https://doi.org/10.1121/1.1339823

Gannon, D. P. (2008). Passive acoustic techniques in fisheries science: A review and prospectus. *Transactions of the American Fisheries Society*, *137*, 638–656. https://doi.org/10.1577/T04-142.1

Gervaise, C., Lossent, J., Valentini-Poirier, C. A., Boissery, P., Noel, C., & Di Iorio, L. (2019). Three-dimensional mapping of the benthic invertebrates biophony with a compact four-hydrophones array. *Applied Acoustics*, *148*, 175–193. https://doi.org/10.1016/j.apacoust.2018.12.025

Godøl, O. R., Handegard, N. O., Browman, H. I., Macaulay, G. J., Kaartvedt, S., Giske, J., Ona, E., Huse, G., & Johnsen, E. (2014). Marine ecosystem acoustics (MEA): Quantifying processes in the sea at the spatio-temporal scales on which they occur. *ICES Journal of Marine Science*, *71*(8), 2357–2369. https://doi.org/10.1093/icesjms/fsu116

Hallacher, L. E. (1974). The comparative morphology of extrinsic gasbladder musculature in the scorpionfish genus *Sebastes* (Pisces: Scorpaenidae). *Proceedings of the California Academy of Sciences*, *4th Series*, *40*, 59–86 Retrieved from https://www.biodiversitylibrary.org/part/52869

Harris, C. R., Millman, K. J., van derWalt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hoyer, S., & Hamman, J. J. (2017). Xarray: N-D labeled arrays and datasets in python. *Journal of Open Research Software*, *5*, 1–10. https://doi.org/10.5334/jors.148

INaturalist. (2021). Research-grade fish observations at Ogden point. www.inaturalist.org; https://www.inaturalist.org/observations?captive=false&nelat=48.415119924587685&nelng=-123.38085285967878&place_id=any&quality_grade=research&subview=map&swlat=48.41076166252949&swlng=-123.396066349211125&verifiable=any&iconic_taxa=Actinopterygii

Kaatz, I. (2002). Multiple sound-producing mechanisms in teleost fishes and hypotheses regarding their behavioural significance. *Bioacoustics*, *12*(2–3), 230–233. https://doi.org/10.1080/09524622.2002.9753705

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science (80-.)*, *220*(4598), 671–680. https://doi.org/10.1126/science.220.4598.671

Ladich, F., & Myrberg, A. A. (2006). Agonistic behaviour and acoustic communication. In F. Ladich, S. Collin, P. Moller, & B. Kapoor (Eds.), *Communication in fishes* (pp. 122–148). Science Publishers.

Lobel, P. S. (1992). Sounds produced by spawning fishes. *Environmental Biology of Fishes*, *33*(4), 351–358. https://doi.org/10.1007/BF00010947

Locascio, J. V., & Burton, M. L. (2015). A passive acoustic survey of fish sound production at Riley's hump within Tortugas south ecological reserve; implications regarding spawning and habitat use. *Fishery Bulletin*, *114*(1), 103–116. https://doi.org/10.7755/FB.114.1.9

Locascio, J. V., & Mann, D. A. (2011). Localization and source level estimates of black drum (*Pogonias cromis*) calls. *The Journal of the Acoustical Society of America*, *130*(4), 1868–1879. https://doi.org/10.1121/1.3621514

Looby, A., Cox, K., Bravo, S., Rountree, R., Juanes, F., Reynolds, L. K., & Martin, C. W. (2022). A quantitative inventory of global soniferous fish diversity. *Reviews in Fish Biology and Fisheries*, *32*, 581–595. https://doi.org/10.1007/s11160-022-09702-1

Looby, A., Riera, A., Vela, S., Cox, K., Bravo, S., Rountree, R., Juanes, F., Reynolds, L. K., & Martin, C. W. (2021). FishSounds, version 1.0. https://fishsounds.net

Lowerre-Barbieri, S. K., Ganias, K., Saborido-Rey, F., Murua, H., & Hunter, J. R. (2011). Reproductive timing in marine fishes: Variability, temporal scales, and methods. *Marine and Coastal Fisheries*, *3*(1), 71–91. https://doi.org/10.1080/19425120.2011.556932

Luczkovich, J. J., Pullinger, R. C., Johnson, S. E., & Sprague, M. W. (2008). Identifying sciaenid critical spawning habitats by the use of passive acoustics. *Transactions of the American Fisheries Society*, *137*(2), 576–605. https://doi.org/10.1577/T05-290.1

Mann, D. A., & Jarvis, S. M. (2004). Potential sound production by a deep-sea fish. *The Journal of the Acoustical Society of America*, *115*(5), 2331–2333. https://doi.org/10.1121/1.1694992

McKinney, W., & Team, P. D. (2015). Pandas - powerful python data analysis toolkit. https://pandas.pydata.org

Merchant, N. D., Fristrup, K. M., Johnson, M. P., Tyack, P. L., Witt, M. J., Blondel, P., & Parks, S. E. (2015). Measuring acoustic habitats. *Methods in Ecology and Evolution*, *6*(3), 257–265. https://doi.org/10.1111/2041-210X.12330

Montie, E. W., Kehrer, C., Yost, J., Brenkert, K., O'Donnell, T., & Denson, M. R. (2016). Long-term monitoring of captive red drum Sciaenops ocellatus reveals that calling incidence and structure correlate with egg deposition. *Journal of Fish Biology*, *88*(5), 1776–1795. https://doi.org/10.1111/jfb.12938

Moulton, J. M. (1960). Swimming sounds and the schooling of fishes. *The Biological Bulletin*, *119*(2), 210–223. https://doi.org/10.2307/1538923

Mouy, X. (2021). Ecosound bioacoustic toolkit. https://ecosound.readthedocs.io

Mouy, X. (2023). Xaviermouy/xav-arrays: 1.0 (version MEE_paper). https://doi.org/10.5281/zenodo.7694025

Mouy, X., Black, M., Cox, K., Qualley, J., Mireault, C., Dosso, S., & Juanes, F. (2020). FishCam: A low-cost open source autonomous camera for aquatic research. *HardwareX*, *8*, e00110. https://doi.org/10.1016/j.ohx.2020.e00110

Mouy, X., Rountree, R., Juanes, F., & Dosso, S. (2018). Cataloging fish sounds in the wild using combined acoustic and video recordings. *The Journal of the Acoustical Society of America*, *143*(5), EL333. https://doi.org/10.1121/1.5037359

Nichols, B. (2005). *Characterizing sound production in nearshore rockfishes* (*Sebastes* spp.) MSc thesis (pp. 1–55). University of South Florida.

Nikolich, K., Frouin-Mouy, H., & Acevedo-Gutiérrez, A. (2016). Quantitative classification of harbor seal breeding calls in Georgia Strait, Canada. *Journal of the Acoustical Society of America*, *140*(2), 1300–1308. https://doi.org/10.1121/1.4961008

Pagniello, C. M., Butler, J., Rosen, A., Sherwood, A., Roberts, P. L., Parnell, P. E., Jaffe, J. S., & Širović, A. (2021). An optical imaging system for capturing images in low-light aquatic habitats using only ambient light. *Oceanography*, *34*(3), 71–77. https://doi.org/10.5670/oceanog.2021.305

Parmentier, E., Lagardère, J. P., Vandewalle, P., & Fine, M. L. (2005). Geographical variation in sound production in the anemonefish *Amphiprion akallopisos*. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1573), 1697–1703. https://doi.org/10.1098/rspb.2005.3146

Parmentier, E., Scalbert, R., Raick, X., Gache, C., Frédérich, B., Bertucci, F., & Lecchini, D. (2021). First use of acoustic calls to distinguish cryptic members of a fish species complex. *Zoological Journal of the Linnean Society*, *195*(3), 964–975. https://doi.org/10.1093/zoolinnean/zlab056

Parsons, M., McCauley, R., Mackie, M., & Duncan, A. (2010). A comparison of techniques for ranging close-proximity mulloway (Argyrosomus Japonicus) calls with a single hydrophone. *Acoustics Australia*, *38*(3), 145–151.

Parsons, M., McCauley, R. D., Mackie, M. C., Siwabessy, P., & Duncan, A. J. (2009). Localization of individual mulloway (*Argyrosomus japonicus*) within a spawning aggregation and their behaviour throughout a diel spawning period. *ICES Journal of Marine Science*, *66*(6), 1007–1014. https://doi.org/10.1093/icesjms/fsp016

Parsons, M., Salgado Kent, C. P., Recalde-Salas, A., & McCauley, R. D. (2017). Fish choruses off Port Hedland, Western Australia. *Bioacoustics*, *4622*(September), 1–18. https://doi.org/10.1080/09524622.2016.1227940

Parsons, M., Salgado-Kent, C., Marley, S., Gavrilov, A., & McCauley, R. (2016). Characterizing diversity and variation in fish choruses in Darwin harbour. *ICES Journal of Marine Science*, *73*(8), 2058–2074. https://doi.org/10.1093/icesjms/fsw037

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pittman, S. J., Monaco, M. E., Friedlander, A. M., Legare, B., Nemeth, R. S., Kendall, M. S., Poti, M., Clark, R. D., Wedding, L. M., & Caldow, C. (2014). Fish with chips: Tracking reef fish movements to evaluate size and connectivity of Caribbean marine protected areas. *PLoS ONE*, *9*(5), 1–11. https://doi.org/10.1371/journal.pone.0096028

Portt, C. B., Coker, G. A., Ming, D. L., & Randall, R. G. (2006). A review of fish sampling methods commonly used in Canadian freshwater habitats. *Canadian Technical Report of Fisheries and Aquatic Sciences*, 1–2604.

Putland, R. L., Mackiewicz, A. G., & Mensinger, A. F. (2018). Localizing individual soniferous fish using passive acoustic monitoring. *Ecological Informatics*, *48*(June), 60–68. https://doi.org/10.1016/j.ecoinf.2018.08.004

Radford, C. A., Ghazali, S., Jeffs, A. G., & Montgomery, J. C. (2015). Vocalisations of the bigeye *Pempheris adspersa*: Characteristics, source level and active space. *The Journal of Experimental Biology*, *218*(6), 940–948. https://doi.org/10.1242/jeb.115295

Riera, A., Rountree, R. A., Agagnier, L., & Juanes, F. (2020). Sablefish (*Anoplopoma fimbria*) produce high frequency rasp sounds with frequency modulation. *The Journal of the Acoustical Society of America*, *147*(4), 2295–2301. https://doi.org/10.1121/10.0001071

Riera, A., Rountree, R. A., Pine, M. K., & Juanes, F. (2018). Sounds of Arctic cod (*Boreogadus saida*) in captivity: A preliminary description. *The Journal of the Acoustical Society of America*, *143*, 1–6. https://doi.org/10.1121/1.5035162

Rountree, R. A., Aguzzi, J., Marini, S., Fanelli, E., De Leo, F. C., Del Rio, J., & Juanes, F. (2020). Towards an optimal design for ecosystem-level

ocean observatories. *Oceanography and Marine Biology: An Annual Review*, *58*, 79–106.
Rountree, R. A., Gilmore, G., Goudey, C. A., Hawkins, A. D., Luczkovich, J. J., & Mann, D. A. (2006). Listening to fish: Applications of passive acoustics to fisheries science. *Fisheries*, *31*(9), 433–446. https://doi.org/10.1577/1548-8446(2006)31[433:LTF]2.0.CO;2
Rountree, R. A., & Juanes, F. (2010). First attempt to use a remotely operated vehicle to observe soniferous fish behavior in the Gulf of Maine, Western Atlantic Ocean. *Current Zoology*, *56*(1), 90–99. https://doi.org/10.1093/czoolo/56.1.90
Rountree, R. A., & Juanes, F. (2017). Potential of passive acoustic recording for monitoring invasive species: Freshwater drum invasion of the Hudson River via the New York canal system. *Biological Invasions*, *19*(7), 2075–2088. https://doi.org/10.1007/s10530-017-1419-z
Sánchez-Gendriz, I., & Padovese, L. R. (2017). A methodology for analyzing biological choruses from long-term passive acoustic monitoring in natural areas. *Ecological Informatics*, *41*(January), 1–10. https://doi.org/10.1016/j.ecoinf.2017.07.001
Sirovic, A., Brandstatter, S., & Hildebrand, J. A. (2012). Fish recordings from NEPTUNE Canada. *The Journal of the Acoustical Society of America*, *132*(3), 1916. https://doi.org/10.1121/1.4755031
Širović, A., Cutter, G., Butler, L., & Demer, D. (2009). Rockfish sounds and their potential use for population monitoring in the Southern California bight. *ICES Journal of Marine Science*, *66*, 981–990. https://doi.org/10.1093/icesjms/fsp064
Širović, A., & Demer, D. A. (2009). Sounds of captive rockfishes. *Copeia*, *2009*(3), 502–509. https://doi.org/10.1643/CP-08-121
Spiesberger, J. L., & Fristrup, K. M. (1990). Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography. *The American Naturalist*, *135*(1), 107–153. https://doi.org/10.1086/285035
Sprague, M. W., & Luczkovich, J. J. (2004). Measurement of an individual silver perch *Bairdiella chrysoura* sound pressure level in a field recording. *The Journal of the Acoustical Society of America*, *116*(5), 3186–3191. https://doi.org/10.1121/1.1802651
Tavolga, W. N. (1977). Mechanisms for directional hearing in the sea catfish (*Arius felis*). *The Journal of Experimental Biology*, *67*, 97–115. https://doi.org/10.1242/jeb.67.1.97
Thornborough, K. J., Lancaster, D., Dunham, J. S., Yu, F., Ladell, N., Deleys, N., & Yamanaka, L. (2020). Risk assessment of permitted human activities in rockfish conservation areas in British Columbia. In *DFO Canadian Science Advisory Secretariat (CSAS) Research Document*. Fisheries and Oceans Canada, Ottawa. Retrieved from https://waves-vagues.dfo-mpo.gc.ca/Library/40888460.pdf
Tolimieri, N., Andrews, K., Williams, G., Katz, S., & Levin, P. S. (2009). Home range size and patterns of space use by lingcod, copper rockfish and quillback rockfish in relation to diel and tidal cycles. *Marine Ecology Progress Series*, *380*(February 2015), 229–243. https://doi.org/10.3354/meps07930
Too, Y. M., Chitre, M., Barbastathis, G., & Pallayil, V. (2019). Localizing Snapping Shrimp noise using a small-aperture Array. *IEEE Journal of Oceanic Engineering*, *44*(1), 207–219. https://doi.org/10.1109/JOE.2017.2777718
Urick, R. J. (1983). *Principles of underwater sound*. McGraw-Hill Companies.
Van Parijs, S. M., Clark, C. W., Sousa-Lima, R. S., Parks, S. E., Rankin, S., Risch, D., & Van Opzeeland, I. C. (2009). Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales. *Marine Ecology Progress Series*, *395*, 21–36. https://doi.org/10.3354/meps08123
Wall, C. C., Rountree, R. A., Pomerleau, C., & Juanes, F. (2014). An exploration for deep-sea fish sounds off Vancouver Island from the NEPTUNE Canada ocean observing system. *Deep Sea Research Part I*, *83*, 57–64. https://doi.org/10.1016/j.dsr.2013.09.004
Wilson, K. C., Semmens, B. X., Gittings, S. R., Pattengill-Semmens, C., & Sirovic, A. (2019). Development and evaluation of a passive acoustic localization method to monitor fish spawning aggregations and measure source levels. *Ocean. 2019 MTS/IEEE Seattle, Ocean*, 1–6. https://doi.org/10.23919/OCEANS40490.2019.8962663
Zimmer, W. M. X. (2011). *Passive acoustic monitoring of cetaceans*. Cambridge University Press.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Supporting_Information.pdf. PDF document including (1) the building instructions and deployment procedures of the three audio-video arrays, (2) detailed description of the automatic detection and localisation process, (3) detailed description of the simulated annealing process to optimise the hydrophone placement, (4) localisation results using controlled sound sources and (5) the description of sound levels measured at Hornby Island and Mill Bay.

SuppInfo_Video_Figure9.mp4: Video of the localisation results from Figure 9.

SuppInfo_Video_Figure10.mp4: Video of the localisation results from Figure 10.

SuppInfo_Video_Figure11.mp4: Video of the localisation results from Figure 11.

SuppInfo_Video_Figure12.mp4: Video of the localisation results from Figure 12.

SuppInfo_Video_Figure13.mp4: Video of the localisation results from Figure 13.

SuppInfo_Video_Figure14.mp4: Video of the localisation results from Figure 14.

SuppInfo_Video_LargeArrayAssembly.mp4: Video showing the assembly of the large array.

SuppInfo_Video_LargeArrayDeployment.mp4: Video showing the deployment of the large array.

SuppInfo_Video_MiniArrayDeployment.mp4: Video showing the deployment of the mini array.

SuppInfo_Video_MobileArrayDeployment.mp4: Video showing an example of survey with the mobile array.

SuppInfo_Video_SimulatedAnnealingOptimisation.mp4: Animation showing the optimisation of the hydrophone placement for the large array using simulated annealing.

How to cite this article: Mouy, X., Black, M., Cox, K., Qualley, J., Dosso, S., & Juanes, F. (2023). Identification of fish sounds in the wild using a set of portable audio-video arrays. *Methods in Ecology and Evolution*, *00*, 1–22. https://doi.org/10.1111/2041-210X.14095