A Public Dataset of Annotated Killer Whale Calls and Detections for Species Detection and Ecotype Classification Models

Authors: (in no particular order)

Kaitlin J Palmer, Fabio Frazao, Jenn Waldichuck3, April Houweling3,4, James Pilkington, Harald Yurk, Holger Klinck, Lucy Quale, Holly LeBlond, Scott Viers, Val Viers, Jasper Kanes, Krista Trounce, Lauren Laturnus, Olivia Murphy, Amalis Riera Vuibert, Alex Harris, Emma Cummings, Kait Frasier, Amanda A. Leu, Ruth Joy

This list is not comprehensive nor in any particular order. Please add any missing names

# Affiliations

Please fill in your own affiliations

# Abstract

# Background and Summary

Passive acoustic methods are critical tools used to monitor vocally active biota within the marine environment. PAM methodologies are used in both real time conservation studies and longitudinal studies spanning decades (Parijs et al., 2009). All methods rely on automated detectors to identify relatively rare sounds of animals and discriminate between species, and in the case of killer whales, between different ecotypes and lineages.

Killer whales (*Orcinus orca*) are found in all oceans of the world and are arguably one of the best studied cetacean species. Salish Sea and adjacent coastal waters, there are genetically and culturally distinct lineages of killer whales that overlap in distribution. These lineages presently include three ecotypes are resident, transient or Biggs, and offshore (Baird & Stacey, 1988; Balcomb III & Bigg, 1986; Ford et al., 1998). For consistency, we refer to these lineages as populations but readers should note that the nomenclature and taxonomy are in flux (Morin et al., 2024). These populations are known to occupy the same regions but do not interbreed. Populations within the resident ecotype include Southern Resident killer whale (SRKW), Northern Resident killer whale (NRKW), West Coast Bigg's (Transient) killer whales (Bigg’s), and Offshore killer whales (OKW). Each killer whale population is ecologically specialized through their seasonal distributions, social structure, and behavior which may be a consequence of differences in their preferred prey. Southern and Northern Resident killer whales are obligate teleost fish consumers, West Coast Bigg’s killer whales feed exclusively on marine mammals, and Offshore killer whales specialize on sharks. Each population faces different environmental stressors, with Southern Resident killer whales especially vulnerable to extinction due to lack of food, pollution in their environment, and acoustic masking from transiting vessels which hinder their ability to find food. Presently there are 75 SRKWs left in the world. Due to these critically low numbers, there are significant and sustained efforts to improve the outcome for the SRKW population including habitat improvement and noise reduction efforts on both sides of the border. These efforts are based on critical habitat as determined by visual and acoustic detections of the population. Acoustic data are most valuable for times/locations where visual detections are not possible.

Killer whale occupancy monitoring involves both visual and acoustic surveys. While visual surveys are restricted to daylight hours, acoustic surveys collect data continuously and in areas inaccessible to visual observers. Acoustic surveys, however, generate large volumes of data which require automated processing to produce results within reasonable timeframes. A variety of generalized detection algorithms (Gillespie et al., 2013; Helble et al., 2012) are available that work reasonably well on killer whale calls and at least one neural network based killer whale detector has been published (Kirsebom et al., 2022). Several groups have independently been working to build Killer whale detection and classification algorithms but there is a need to combine efforts to build an ecologically representative dataset.

Each of the four populations of killer whale has a distinct vocal dialect or repertoire of stereotyped sounds that are unique to their population, but can be broadly grouped three broad categories, echolocation clicks, whistles, and pulsed calls. Echolocation clicks are broadband impulsive sounds with the majority of the energy between 20 and 100 kHz. Whistles are tonal calls typically used for social communication among individuals within a pod. These whistles have a broad frequency range, generally spanning from 0.5 to 25 kHz, and are involved in coordinating movements and maintaining group cohesion. Pulsed calls are distinct, complex vocalizations characterized by a series of discrete, pulsed sounds varying in frequency and amplitude. Unlike echolocation clicks and whistles, these calls are primarily used for social communication within pods, serving functions in group coordination, individual identification, and conveying social and behavioral cues. These broad sound classes contribute to the distinct vocal repertoires associated with different whale social groups and form the motivation for harnessing the power of modern classification methodologies to make the most of acoustic surveys both in archived or near real-time settings.

These data were principally being compiled as a challenge dataset for the 2026 conference and workshop on the Detection Classification, Localization, and Density Estimation of marine mammals using Passive Acoustics (DCLDE). The goal of this workshop dataset is to encourage researchers to build tools for detecting killer whales in the large datasets and, where possible, classify the calls to ecotype. The following sections describe the methods for collecting and collating each of the datasets that have been provided by the collaborators.

# Methods

Machine learning models are only as good as the data used to train them. For acoustic ecology, data used to train the algorithm must be representative of the animals’ repertoire to be effective (Shiu et al., 2020). Furthermore, many machine learning applications in conservation are targeted at longitudinal, or backwards looking studies in order to assess changes on the scale of years or decades. In species capable of cultural adaptation of their repertoires including humpback and killer whales, data for machine learning algorithms must represent signals that were previously heard in the environment (e.g. antiquated song, and killer whale calls from now diseased animals). Furthermore, environmental factors including but not limited to background noise, instrument parameters, sound propagation conditions can all influence how robust a detection and classification algorithm is.

We refer to the ideal dataset as “ecologically representative” indicating that annotated audio signals encompass the range breadth of the target species repertoire. The dataset must also contain recordings from a variety of locations, and recording conditions. Equally as important, the detection and classification algorithm must be able to discriminate between target sounds and those by other animals in the survey area. Notably, this includes humpback whales. Anthropogenic sounds, and other odontocetes can also be easily confused with the acoustic signals of killer whales and examples of these should be in the final dataset.

The goal of this dataset is to facilitate the construction and evaluation of detectors that are capable of 1) discriminating killer whale calls from other acoustically similar species and 2) determine the which ecotype, or species, of killer whale are present as conservation status and regulations vary between each. There are a few challenges in doing so. First, data donated to this effort were amassed from several independent projects each with different goals, using different methods, and annotated to different levels. For example, some DFO and JASCO/VFPA audio were all processed with energy detectors and analysts validated each detection. Other efforts focused on annotating all signals of interest in a small subset of data, such as the Scripps files. Second it is not always possible for even expert analysts to discriminate between killer whale and other species in the frequency range, notably pacific white-sided dolphins and humpback whales. Within killer whale acoustics it is also possible to discriminate between ecotypes, clans, and even individuals. However, the resolution varies. Often, it is possible to discriminate ecotypes if numerous calls are detected within an encounter. In some instances, it’s possible to determine matriline or individual identification. However, it’s also as likely that no ecotype specific information is present within the acoustic encounter. Thus, the annotation scheme should retain classification uncertainty where present.

Building such a dataset is challenging and often cost prohibitive for a single organization. Thus, in this effort we have combined smaller annotated datasets from multiple commercial, non-commercial, and governmental organizations to build a an ecologically representative annotation dataset. Much of the annotation effort was provided through the Humans and Algorithms Listening for Orcas (HALLO) project which used a standardized annotation procedure included in supplemental information 1. The following sections provide detailed information on the 1) Deployment 2) Processing and 4) Annotation procedure for each of the projects. Metadata, where available, is outlined in Table 2.

## Data Records

The overall challenge dataset contains audio and annotations provided by a collaboration of businesses, not-for-profits and governmental organizations (Table 2). These include Orcasound, Ocean Networks Canada (ONC), the Canadian Department of Fisheries and Oceans (DFO), JASCO Applied Sciences (JASCO) and Malahat First Nations, and Saturna Island Marine Research & Education Society (SIMRES). Data were collected using a variety of instruments deployed in the Pacific Northwest including AAMRs ([https://www.jasco.com](https://www.jasco.com/amar-g4)), Ocean Sonics icListen hydrophones (https://oceansonics.com/products/iclisten-sj9/), and Ocean Instruments soundtrap recorders (<https://www.oceaninstruments.co.nz/>), and in depths ranging from 8-253m. Data coverage varied, as time and funding allowed but covered a 9-year period between May 2013 and June 2022. Deployment, processing, and annotation details for each dataset are provided in the following sections.

To address consistency issues across multiple annotation schemes and annotators we take a two-step approach. First, we provide the raw annotations with explanations from each organization regarding how the data were processed. The original annotations often contain considerable information that is beyond the scope of the challenge including a variety of different labels for biologic and anthropogenic sounds and finer resolution on killer whale calls (e.g. matriline or call type). These annotations details should be of interest to those knowledgeable in the field of killer whale acoustics. We also provide a collated dataset across all providers and deployment locations to get detection and classification algorithms running quickly. The collated annotation table (Annotations.csv) includes standardized annotations from across all datasets with labels described in Technical Validation section (Table 1). Finally, we include the code used to homogenize the datasets that produces the Annoations.csv file.

### Orcasound

Orcasound is a cooperative hydrophone network and an open-source software & hardware project. Orcasound audio and annotations were compiled from multiple recording efforts spanning from 2017 to 2020. This public dataset includes nine labeling efforts with the 'Pod.Cast' annotation tool, an open-source web app developed by Microsoft Hackathon volunteers to efficiently analyze audio data to detect the presence of killer whale calls (https://ai4orcas.net/portfolio/pod-cast-annotation-system/). Origional audio recordings and annotations are accessible via Orcasound's open labeled data bucket. The dataset is organized into annotation rounds that used audio data from various Orcasound locations with a range of SRKW call signal to noise ratios and background noise characteristics. Full details of orcasound data are available on the gitub account for these projects (*Orcasound*, n.d.)

**Deployment**

The Orcasound data were gathered from three sites in Washington State, USA: the Orcasound Lab on San Juan Island (Haro Strait), Bush Point on Whidbey Island, and Port Townsend (the latter two within Admiralty Inlet). At each location, low-cost hydrophones were deployed: LabCore-40 or CRT26-08 elements were utilized in Admiralty Inlet, whereas Orcasound Lab tested a wide variety of elements, including HTI 99-MIN, Aquarian AS-1, and ITC1032 models. These hydrophones were deployed in shallow waters (less than 10 meters at low tide) using bespoke, affordable live-streaming equipment (Raspberry Pi with the Pisound ADC HAT [24 bit, stereo, max 192 kHz]) and the orcanode open source code that generates compressed, lossy audio segments in HLS format and uploads it to an open S3 bucket sponsored by Amazon. Hydrophones and recording systems for these projects have not been calibrated.

**Processing**

Audio data were collected in a variety of formats and at multiple sample rates. The majority of the audio data were sampled at 48khz but a strong lowpass filter with a steep rolloff at 16.5khz was applied rendering frequencies above this filter unusable. All audio files were pre-processed with an anti-aliasing filtration diminishing sound intensities at frequencies above 12kHz. Potential southern resident killer whale calls were initially detected by citizen scientists who have access to live-streamed audio recordings. Citizen scientists indicate periods of likely killer whale activity, and those audio files are reviewed by expert analysists and annotated. These files were archived and noted as ‘candidates’ for further analysis (*OrcaHello*, n.d.).

**Annotation**

A sub-set of Orcasound's open labeled data includes archives that were prepared via the Pod.Cast system. Audio and annotations consist of 9 of the 10 ‘rounds’ of pod.cast datasets, each being part of a Google Summer of Code competition. For each 'Round' of data, candidate audio data for SRKW pre-labeled by running an existing classifier with a threshold tuned for high-recall, and validated by crowd-sourcing the predictions.

This project's annotations specifically aimed at identifying Southern Resident Killer Whales (SRKW), categorizing detections into two classes: SRKW and False Positive. The annotation granularity varied between these classes; for confirmed SRKW calls, the start and end times were documented. For non-detections (i.e., 'false positives'), the files were marked as 'FP,' without specifying time or frequency boundaries. Citizen scientist-flagged files underwent expert review to confirm the presence of SRKW calls, noting the call's start and end times. Frequency bounds were not recorded, hence listed as 'NAN' in the frequency columns of the annotation files. Files lacking identifiable SRKW calls were tagged as 'noise,' and all noise labels were reclassified as 'Abiotic' in the SpeciesClass column.

### Ocean Networks Canada

Ocean Networks Canada (ONC) operates cabled underwater observatories in Canadian waters collecting continuous oceanographic data for the benefit of science, society, and industry. Many of their nodes are equipped with calibrated hydrophones (Biffard et al., 2022) to record long term data on changing ocean soundscapes and support research on noise and soniferous animals. Calibration information and other metadata are available on the Ocean Data Portal (*Ocean Networks Canada - Oceans 3.0*, n.d.).

**Deployment**

Acoustic data were collected using an Ocean Sonics SC2 (<https://oceansonics.com/>) recording system deployed on the Barkley Canyon Upper Slope platform of ONC’s North-East Pacific Time-series Underwater Networked Experiments observatory. The hydrophone was mounted 1 m above the sea floor and sampled continuously at 64 kHz. data that did not contain classified signals were archived after review by regional navies.

**Processing**

The hydrophone sampled at 64kHz but uses a 25.6kHz anti-aliasing filter during data collection and digitization, yielding a 32kHz bandwidth with reduced apparent sound intensities above 25.6kHz. Data were evaluated for the presence of killer whales and other species in three separate efforts with varying protocols for each. All manual annotation was completed initially using JASCO’s PAMLab software. Annotations were produced using a logarithmic spectrogram display with different spectral settings in 4 different bands, enabling multi-species identification across the full bandwidth in a single pass. All visible signals were considered for annotation, and no signal-to-noise ratio threshold was used. Annotations initially made in PAMLab were reviewed for accuracy, signal diversity, and completeness using Raven Pro v 1.6. No automatic detection algorithms were applied during any part of the analysis.

**Annotation**

Two groups independently annotated these data for different purposes using different approaches. The original annotation effort focused on categorizing marine mammal presence as well as producing diverse call-level annotations for classifier development. Here, every second file for each of the first four days of the month in 2014 was reviewed for the presence of marine mammal signals. If a marine mammal call or signal was found, the signal was annotated on one of three levels. If the file contained a killer whale signal, all pulsed signals within the file were annotated if present. Some whistles were annotated but echolocation clicks were not annotated. Due to their infrequent use of the area, files previously identified as containing killer whales in 2013 during opportunistic data checks were also annotated and included in the data set. For all other species, only one signal (e.g. fin whale classic call) was annotated per file with exceptions made to include an extra call to capture diverse signals. Thus, some of the audio files containing biological sounds contain more biological signals than were annotated.

A latter effort sought to incorporate inter-observer variability into the analysis and thus hired two outside experts to re-annotate a section of data between May 20 and June 23, 2023. Thus, two outside experts from JASCO analyzed a subset of the annotations from A Raven Pro. In the secondary analysis all SRKW and Bigg’s communication signals such as pulsed calls, whistles, buzzes, and rasps were annotated with bounding boxes demarcating the start and end time of the signal as well as the low and high frequencies. When possible, pulsed calls were further classified into specific call types, following the call types demarcated by Ford (1987). Although not all echolocation clicks were recorded, a singular instance or burst of clicks was marked in each audio file to indicate the presence of echolocation in the file. Each signal was assigned a confidence rating of either ‘low’, ‘medium’, or ‘high’ to specify the level of certainty provided by the annotator. If the call could not be identified, it was left as ‘unknown’. The two annotation streams have been combined for the final dataset. The two annotation sets have not been thoroughly cross referenced in order to allow for inter-observer analysis.

For the purposes of the detection and classification dataset, all annotations indicating the possible presence of killer whales were categorized as ‘KW’ regardless of certainty. Annotations that indicated uncertainty to the species by indicating either possible alternative species or were demarcated with medium or low certainty were defined in the KW\_certain category. Because click annotation varied between groups, click annotations in the ONC data were classified as ‘undetermined biological’ sounds. Killer whale annotations that were paired with other potential species, e.g. “killer whale/white sided dolphin” were similarly classed as undetermined biological sounds. All killer whale annotations containing a “?” were tagged as uncertain as

There were over a hundred different tags for specie data in these annotations including all variation of possible confounding species (e.g. killer whale, humpback, or pacific-white sided dolphin, or unidentified biological sounds). For this dataset any label that contained killer whale possibility was tagged as KW and if other species were listed as alternative possibilities the KW certainty column was set to 0. Calls that were identified as possible humpback whale calls were added to the humpback category The ClassSpecies label for killer whale clicks and buzzes was set to ‘UndBio’ as this was the only dataset that labeled impulsive calls. Should users wish to include clicks in classifiers, they should refer to the original annotations

### Department of Fisheries and Oceans Canada

Two groups within DFO provided datasets to the challenge, the Cetacean Research Program and Whale Detection and Localization Program. Data processing methods were consistent across projects within each lab but varied slightly between labs. Exact hydrophone locations are not publicly available for any DFO hydrophone dataset. Instead, general location descriptors are provided.

*Cetacean Research Program*

Data from the Cetacean Research Program (CRP) lab consisted of two deployments, one on the continental shelf edge off the west coast of Vancouver Island and another from an instrument deployed on the northern mainland coast of British Columbia. Data were based on approximately 375 days of recording off Vancouver Island and 116 days of recording in northern BC. With the former targeting the winter months. The focus of the original analysis effort that resulted in these datasets was simply to identify which of the recording files contained killer whales calls for use in various habitat studies. The analysis was conducted by using an automated detector and manually identifying all of the resulting detections. Such a manually annotated dataset may be useful to detector/classifier development efforts.

**Deployment**

Data were collected using AURAL-M2 and SM2M off of Vancouver Island and Northern BC, respectively. As with all DFO data exact locations are not publicly available. Deployment depth off of Vancouver Island was approximately 114m and 35m at the Northern BC deployment site. The AURAL-M2 sampled audio at 16.384 kHz and the SM2M sampled at 16khz.

**Processing**

The raw audio recordings (WAV) were post-processed using the Whistle and Moan Detector in PAMGuard version 1.12.08 (Gillespie et al., 2013) with an FFT length of 512 and hop size of 50% (256). The detector was user- configured with a high-pass filter of 800Hz to limit the number of humpback whale detections and lessen the manual validation burden. The SNR detection threshold was set to 6dB. All detections in the first two seconds of each file were excluded because the detection algorithm produces several false detections within this period.

PAMGuard detections were subsequently processed into calls using a simple grouping algorithm. The algorithm first searched for temporally overlapping detections. Detection start, end, low and high frequencies were combined if the labels were consistent across overlapping detections. This was done with finer resolution labels (e.g. call type, clan, subclan, or pod) than are presented in the final dataset.

**Annotation**

All detections including whistles and pulsed calls were aurally and visually reviewed using PAMGuard and identified to species (for biotic) and sound type (for abiotic). Where applicable and as time allowed, detections were also acoustically identified to ecotype. Note that files may contain more identifiable calls than the annotations indicate due to false negatives that are inherent when using automated detectors. These manual reviews were conducted by trained and experienced analysts using the HALLO protocol (HALLO Annotation Guidelines).

Note that individual detections may be separate components of the same discrete call (i.e. harmonics or sidebands), thus, not every detection represents a unique vocalization. The PAMGuard Whistle and Moan detector detects individual contours, so all individual harmonics within a call would constitute separate detections if they meet the detector’s criteria (this happens quite frequently). Also, the settings of the detector mean that independent tones (like from multiple individuals) that cross or overlap in frequency and time may be detected as a single detection. Of the 158k annotations provided, the start time of 348 annotations coincided (e.g. overlapped) with the time start time and duration of other annotations.

*Whale Detection and Localization Program*

Whale Detection and Localization Program (WDL) provided data from four deployment locations in Canadian waters including, Carmanah Point, Swanson Channel, and two locations in the Northern and Southern ends of the Strait of Georgia. The annotated dataset spanned 298 days from September 2021 through June 2022.

**Deployment**

Four locations were chosen for the study area. Carmanah Point, Swanson Channel and the Northern and Southern Ends of the Strait of Georgia. Exact locations are not disclosed. A SoundTrap (www.oceaninstruments.co.nz) was used for the Carmanah point location and AMARS (https://www.jasco.com) were used for all other deployments. All deployments were between about 3- 5 months long. Audio data were continuously sampled at either 192 kHz for the SoundTrap or 256 kHz for the AMARs.

**Processing**

Audio recordings were processed with the PAMGuard (Gillespie et al. 2013) Whistle and Moan Detector (v. 2.02.03) for the presence of potential killer whale calls. Audio files were downsampled within PAMGuard to 48 kHz, and a weak IIR Butterworth high-pass filter with a threshold of 2 kHz and an order of 1 was applied to reduce background noise in the lower frequency bands. Nominal sensitivities of -164.1 dB and -176.2 were used for AMARs and the SoundTraps, respectively. The Whistle and Moan Detector used a minimum frequency threshold of 200 Hz, a maximum frequency threshold of 24000 Hz (the Nyquist rate), and a minimum contour length of 15 time slices (about 341 milliseconds); otherwise, all other detection settings were kept at their defaults. In the detector's noise and thresholding tab, all boxes except "Run Gaussian Kernel Smoothing" were checked and any input values were kept at their defaults as well. The FFT engine used with the detector used an FFT length of 2048, a hop size of 1024, and a Hann window function, with the same noise parameters as those in the detector.

**Annotation**

All PAMGuard detections were evaluated for the presence of killer whales by expert analysts and annotated as such. Annotations included whistles and pulsed calls but echolocation clicks were not included as they were rarely discovered by the whistle and moan detector. As with the Pilkington dataset that was similarly processed by PAMGuard, multiple annotations could represent a single call. In this case, 27% of the detections overlapped in time and could represent call partitions.

### JASCO and Vancouver Frasier Port Authority

The Vancouver Frasier Port Authority (VFPA) in collaboration with JASCO Applied Sciences, collected data from two locations in Haro Strait and Boundary Pass. These data were part of the Enhancing Cetacean Habitat Observation program which aims to improve killer whale acoustic habitat through voluntary vessel speed reductions (Joy et al., 2019).

**Deployment**

AMAR recorders were deployed directly adjacent to the southbound and northbound shipping lanes in Haro Strait (Table 2, Figure 1). Instruments at both locations were deployed and recovered twice. The first deployment extended between July 6th and September 8th 2017. Instruments were deployed and refurbished AMAR’s were re-deployed at the same locations on September 8th and recovered October 26th of the same year. Data from the boundary pass location were collected over the period between September 2018 and April 2019. Deployment depths ranged between 193m to 251m across the three regions.

**Processing**

For all deployments, data were sampled at 256 kHz and killer whale encounters were identified with a proprietary detection algorithm developed by JASCO Applied Science.

**Annotation**

Encounters were manually annotated by expert analysts for the presence of killer whale calls following the HALLO protocol. Expert annotators used Raven Pro to identify killer whale calls and, where possible, classify calls to call type. Annotators also noted the presence of a variety of non-target calls and abiotic sounds including unknown signals, background noise, fish, and potential pacific-white-sided dolphins.

### JASCO, Vancouver Frasier Port Authority, Ocean Networks Canada

**Strait of Georgia**

The Strait of Georgia underwater listening station (ULS) is a collaborative project between the Vancouver Fraser Port Authority, Transport Canada, Fisheries and Oceans Canada, Ocean Networks Canada and JASCO Applied Sciences. Data from this hydrophone have been used in evaluating changes in noise levels associated with voluntary vessel slowdowns (Joy et al., 2019).This listening station has been in place since September 2015 and is now in its third year of operation. A small section of the data were manually validated for the presence of killer whales, humpback whales and other signals of interest.

**Deployment**

The ULS is situated on the seabed at approximately 170 m depth approximately 30km west of Vancouver, Canada. The location aims to monitor noise in association with the northbound shipping lane. Synchronized data from four hydrophones are streamed to shore in near real-time via the Victoria Experimental Network Under the Sea (VENUS) Observatory operated by Ocean Networks Canada.

**Processing**

For all deployments, data were sampled at 256 kHz and killer whale encounters were identified with a proprietary detection algorithm developed by JASCO Applied Science.

**Annotation**

Encounters were manually annotated by expert analysts for the presence of killer whale calls following the HALLO protocol. Annotators used Raven Pro to identify killer whale calls and, where possible, classify calls to call type. Annotators also noted the presence of a variety of non-target calls and abiotic sounds including unknown signals, background noise, fish, sonar, and potential pacific-white-sided dolphins.

### SIMRES

The Saturna Island Marine Research and Education Society (SIMRES) maintains several hydrophones along the BC coast as part of the Whale Sound Network. This network collaborates to enable scientific quantification of how the ocean soundscape is changing. Since 2015, hydrophones have been placed in Boundary Pass at two different locations, East Point and Monarch Head (SIMRES, 2020). The annotated data provided were from the East Point Hydrophone located off the southeasternmost point of Saturna Island. From June through October 2022, Southern Resident killer whales (SRKW) were both acoustically detected and visually sighted 13 times by nearby citizen scientists from the Southern Gulf Islands Whale Sighting Network (SGIWSN) and student researchers from Simon Fraser University.

The goal of this annotation effort was to determine how vessel noise impacts SRKW acoustic behavior in Boundary Pass. The annotated dataset includes 13 hours and 40 minutes of SRKW acoustic activity and call types were identified from all three SRKW pods, J, K, and L. The duration of acoustic events varied and ranged from 5 minutes to 190 minutes.

**Deployment**

An Ocean Sonic’s icListen high-frequency (HF) smart hydrophone (RB9-ETH) with ethernet (<https://oceansonics.com/products/iclisten-sj9/>) was used to collect audio recordings. The shore-cabled hydrophone is located at a depth of 18 m, approximately 120 m from shore, near the commercial shipping channel in Boundary Pass (48.780° N, 123.052° W). Data are continuously sampled at 128 kHz but down sampled to 64 kHz files are provided.

**Processing**

Audio data were not pre-processed with any detection algorithms for this for this study.

**Annotation**

Audio files were manually annotated in Raven Pro v 1.6. All SRKW communication signals including pulsed calls, whistles, buzzes, and rasps were annotated with bounding boxes demarcating the start and end time of the signal as well as the low and high frequencies. When possible, pulsed calls were further classified into the specific call types outlined by Ford (1987). A singular instance or burst of clicks was marked in each audio file to indicate the presence of echolocation but echolocation clicks were not otherwise annotated. The original selection tables also contain annotators comments which may be useful in selecting data to build finer level detection and classification algorithms than outlined here.

Annotated signals were assigned a confidence rating of either ‘low’, ‘medium’, or ‘high’ to specify the level of certainty provided by the annotator. If the call could not be identified, it was left as ‘unknown’. A small number of potential humpback whale calls were also annotated with uncertainty, these have been included in the ‘humpback’ SpeciesClass.

All killer whale annotations were included in the combined annotation dataset regardless of quality. Annotiation indicating a low or medium quality were noted as ‘Uncertain” or 0 in the KW\_certain colum.

### Scripps Institute of Oceanography

Data from two locations spanning 2008-2013 years were provided by the Scripts Institute of Oceanography. Data were part of a long-term monitoring project off the Washington Coast and consist of encounters included in previously published work (Leu et al., 2022; Rice et al., 2017).

**Deployment**

High-frequency acoustic recording packages (HARP; Wiggens et al.(2007)) packages were deployed in a nearshore (Cape Elizabth) and offshore (Quinault Canyon) location. HARPs sampled continuously at 200 kHz. Data from this project represents the most southern locations studied as well as the deepest deployment (1400 m).

**Processing**

Audio data were not pre-processed with any detection algorithms for this for this study.

**Annotation**

Original pulsed annotations ad described in Rice et al. (2017) were not available. As such, data were re-analyzed for the presence of killer whale calls using Raven Pro v1.6. Only calls that could be confidently identified as killer whales were included in the final annotations. Humpback whale calls were added opportunistically and examples of self-noise, tagged as abiotic signals, were included as these signals show structural similarities to biological signals and should be trained against. Killer whale ecotype classes were defined off of the original encounter labels (Leu et al., 2022, 2022). Though present in the encounters, echolocation clicks were not labeled during the annotation effort.

### SMRU Consulting

SMRU Consulting in collaboration with the Whale Museum have maintained a cabled hydrophone within SRKW core habitat for nearly two decades. These data have also been involved in evaluating then potential benefits of voluntary ship slowdowns (Joy et al., 2019). Data are routinely evaluated for the presence of killer whales and humpback whales. The hydrophone location is also within visual range of the Lime Kiln lighthouse which houses volunteers trained for whale and dolphin identification. Audio files associated with visually confirmed acoustic encounters over several periods from 2016-2021 were provided for the challenge dataset.

**Deployment**

The recording setup consists of a cabled hydrophone ~30m from shore mounted to the seafloor at 23m depth. Data are streamed continuously onshore, uploaded to a cloud-based system.

**Processing**

Data are processed for the presence of biological sounds with the PAMGuard whistle and moan detector (Gillespie et al., 2009). Signals of interest are flagged for later review.

**Annotation**

Annotation procedures followed the HALLO protocol.

### University of Alaska Fairbanks

Data contributed by the University of Alaska Fairbanks and North Gulf Oceanic Society are part of a long-term killer whale monitoring project in the Gulf of Alaska. This includes recordings of the southern Alaska resident, Gulf of Alaska transient, AT1 transient, and offshore killer whale populations.

**Deployment**

Recordings of southern Alaska residents were taken with a dipping hydrophone during vessel survey encounters in Prince William Sound and Kenai Fjords (Figure 1) between May and October in 2019, 2020, and 2021. When killer whales were encountered, we photographically identified as many individuals present as possible. We then maneuvered the vessel approximately 500 m in front of the animals, shut off the engine, and collected a field recording using either an Ocean Instruments SoundTrap ST300 hydrophone or a High-Tech, Inc. HTI-96-Min hydrophone with a TASCAM DR100 portable digital recorder (sampling rate 24 kHz). Dipping hydrophones were deployed at approximately 8 – 10 m depth (Table 2).

We rarely encountered transient or offshore killer whales during vessel surveys, and transient killer whales also vocalize less often than residents (Deecke et al., 2005; Saulitis et al., 2005) making field recordings difficult to obtain. We therefore contributed recordings from moored hydrophones in which we detected Gulf of Alaska transient, AT1 transient, or offshore killer whales. Moored hydrophones were deployed in Hinchinbrook Entrance, Montague Strait, Resurrection Bay, and Kachemak Bay (Sup. Fig XXX,Table 2) beginning in 2016, though for this analysis we included transient recordings from June 2019 to May 2021 and offshore recordings from May 2022 and April 2023. Hydrophones were deployed at depths of 25 – 42 m on primarily gravel and sand substrate and were moored approximately 2 m above the seafloor. Moored hydrophones recorded at a 24 kHz sampling rate and on a predominantly 5 min on, 10 min off duty cycle, although some deployments recorded 4 min on, 11 min off or 4 min on, 16 min off to conserve battery.

**Processing**

All acoustic data from moored hydrophones were processed using the Whistle and Moan Detector in the open-source software package PAMGuard v.1.15.17 (Gillespie et al., 2009). Spectrograms were created with a 1024 Hz Fast Fourier transform length and 50% overlap. The Whistle and Moan detector identified tonal signals in the 700 – 12,000 Hz frequency range with a minimum length of 15 time slices, minimum size of 30 pixels, and that met an 8 dB signal-to-noise ratio threshold. Recordings with at least three detections were manually checked visually and aurally by H. Myers and classified to the population level. Gulf of Alaska transients and AT1 transients were identified using published call catalogues (Myers et al., 2021a; Saulitis et al., 2005). Offshore killer whale detections were confirmed by J. Pilkington. A small minority of recordings included multiple killer whale populations or killer whale and humpback whale (*Megaptera novaeangliae*) vocalizations; these recordings were not included in the dataset.

**Annotation**

Discrete pulsed calls were manually annotated by H. Myers in Raven Pro v.1.6.5. A bounding box was drawn around each call, and the call start time, end time, low frequency, high frequency, and length were saved in selection tables.

## Technical Validation

All potential killer whale annotations were created by expert analysts at their respective institutes based on a canonical catalogue of killer whale calls (Ford 1987). As with all biological signals, the quality varied considerably based on the background noise, distance between the animal and the hydrophone, and propagation considerations.

Low SNR detections, as indicated by the reviewing analyst, were not included in the dataset or tagged as uncertain. Calls associated with SIMRES were linked with concurrent sightings of SRKW animals.

Collated annotations covered an approximately 11-year span from May 2011 through June 2022 and were recorded on a variety of instruments including JASCO AMARs, Soundtraps, IC listening devices, and custom-built hydrophones. Sample rates ranged from 16-125khz but were downsampled to 16khz.

An annotation file is provided as a CSV that includes links to audio files. The following describes each column in the annotation file.

Table 1 Annotation file descriptors

|  |  |  |
| --- | --- | --- |
| CSV Headings | Contents | Structure |
| **SoundFile** | Name of the audio file from which the annotation was derived | Character string |
| **FilePath** | Full file path to the to the audio file above | Character string |
| **FileBeginSec** | Seconds into the audio file representing the start of the bounding box or detection | Double |
| **FileEndSec** | Seconds into the audio file representing the end of the call annotation | Double |
| **LowFreqHz** | Lower limit of the bounding box or detection, in Hz | Double |
| **HighFreqHz** | Upper limit of the bounding box or detection, in Hz | Double |
| **UTC** | UTC time at the beginning of each annotation (**FileBeginSec**) | Character string ISO formatted date/time |
| **ClassSpecies** | Species or class description with the following options: Killer Whale (**KW**), Humpback Whale (**HW**), Abiotic (**AB**), and Undetermined Biological sound (**UndBio**). | Character string |
| **AnnotationLevel** | Caracter string representing whether the annotation represented a validated **detection, call,** or **file** | Character string |
| **KW** | Indicator of whether or not the annotation denotated that the annotation represented a killer whale call | Bool (0,1) |
| **KW\_certain** | Indicator of whether or not the annotator was certain that the annotation was a KW. This is often represented by a question mark in the annotations. For ONC data, annotators listed all potential species that the thought the call could come from. | Bool (NA,0,1) |
| **Ecotype** | Killer whale ecotype or population represented by the KW annotation, if known. **SRKW**- Southern Resident Killer Whale, **BKW**- Biggs killer whale, **NRKW**- Northern Resident Killer Whale, or **OKW**- Offshore Killer Whale or blank when no ecotype could be determined | Character string or NA |
| **Data Provider** | Group providing the data | Character string |
| **Dep** | Shorthand for the deployment location | Character string |

## Usage Notes

The intended purpose of these data is to build classifiers for detecting killer whales and classifying signals to population or ecotype in the Eastern Pacific. In building detection and classification algorithms users should consider both their intended applications and potential limitations. For instance, users will immediately note that sample rates differ between each of the contributed datasets and analysis of the annotations will show that down sampling the higher frequency data will limit exclude some of the higher frequency annotations. Conversely, excluding the lower frequency annotations will result in a much-reduced dataset. The biological implications of the sample rate are also worth noting. Much of the effort in classifying killer whale ecotypes and populations has utilized lower frequency sound <12 kHz (Ford et al. 2022). However, as seen here, killer whale vocalizations may have fundamental frequencies at or above 20khz. Whether or not the features present at higher frequencies represent useful information for population or ecotype type discrimination is yet to be determined.

The audio files presented here are done so in their raw state. They have not been normalized to account for different gain and calibration settings between the various instruments and individual project goals. Researchers wishing to measure received levels should reach out individual data providers directly to to determine appropriate calibration offsets. It is also important to note that the sample rate is not always indicative of the useful frequencies. Some groups have applied low-pass filters with cutoff frequencies considerably below Nyquist.

Note also that echolocation clicks have not been not annotated consistently and are thus not included in the final dataset. However, the presence of echolocation clicks has been noted in some files from JASCO and ONC. See original files for those annotations. As echolocation clicks can be diagnostic of species and potentially ecotype (cite Maddie/April?), further annotation of this dataset could feed into training or validation based on echolocation parameters.

Data for this project represent a large collaboration of groups and institutions and each dataset was processed in accordance with each groups project goals. Post processing of the annotations was done to provide a uniform system for machine learning algorithms. However, users should consider details from each deployment carefully to determine whether they wish to do any additional post-processing. For example, multiple annotations from the DFO datasets may represent different harmonics of the same call. Alternatively, data derived from ONC projects considered only pulsed calls. Thus, unannotated whistles and echolocation clicks may be present in some files. See individual datasets above for details.

## Code Availability

Data and annotations were collated using R and available here. https://github.com/JPalmerK/DCLDE2026

# Acknowledgements

# Author Contributions

Holger Klinck, Lucy Quale, Holly LeBlond, Scott Viers, Val Viers, Jasper Kanes, Krista Trounce, Lauren Laturnus, Olivia Murphy, Amalis Riera Vuibert, Alex Harris, Emma Cummings, Kait Frasier, Amanda A. Leu, Ruth Joy

KJ Palmer collated the final dataset, managed data sharing agreements, produced the collated annotation files, and annotated the SCRIPPS dataset. She also drafted, edited and reviewed the manuscript a

F. Frazao devised the HALLO annotation procedure and participated in data curation throughout the process. He also participated in editing the manuscript.

J Waldichuckwas the lead expert analyst on the HALLO datasets, managed JASCO datasets, participated in managing data sharing agreements, facilitated data transfers, and participated in writing and editing the manuscript.

A. Houweling was part of the expert annotation team, participated in writing and editing the manuscript, and facilitated data transfers, and participated in editing the manuscript.

J. Pilkington provided data and annotations from the Department of Fisheries and Oceans Cetacean Research Program, participated in writing and editing the manuscript.

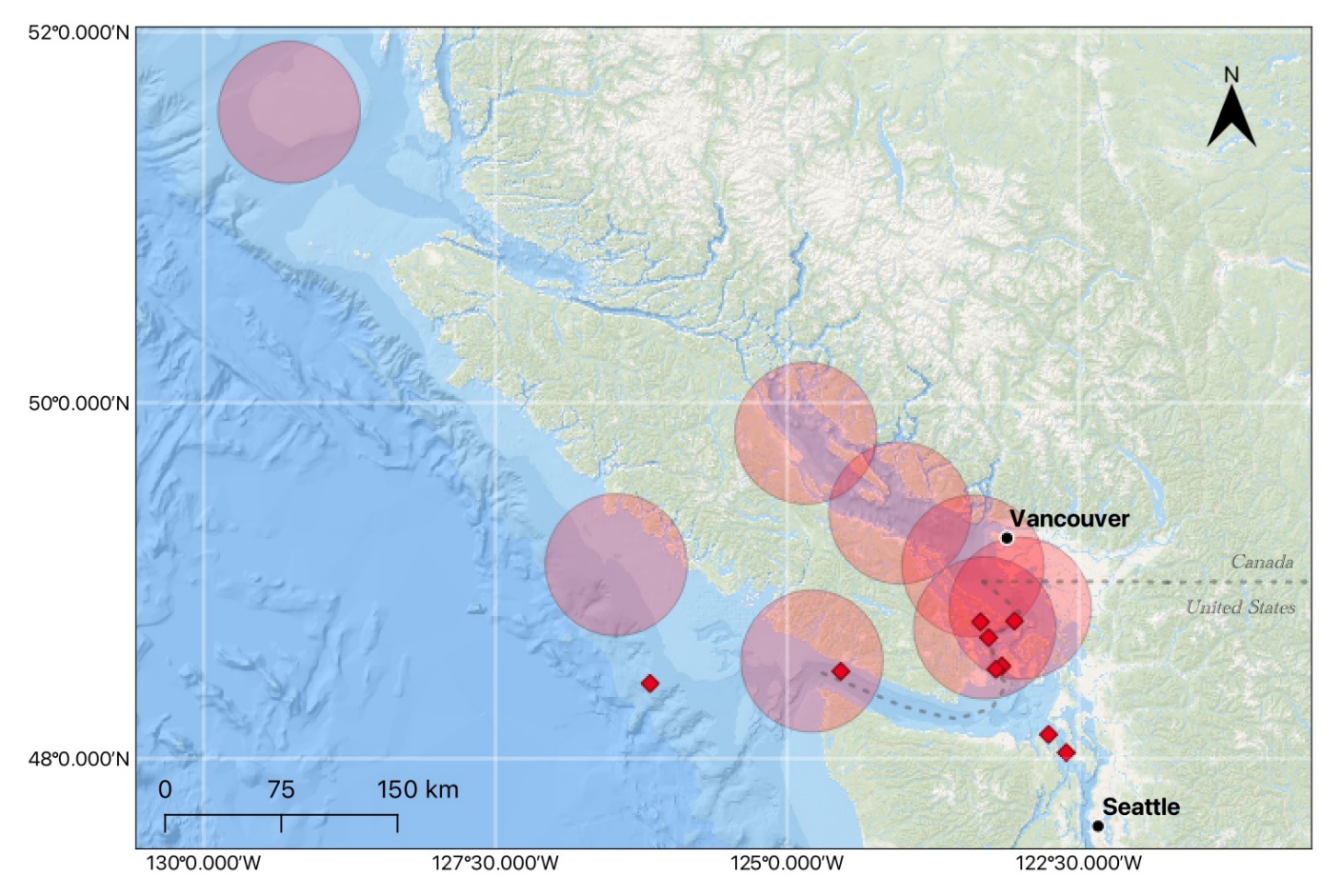
H. Yerk provided data and annotations from the Department of Fisheries and Oceans Whale and Dolphin Listening Program, participated in writing and editing the manuscript.

H. Klinck provided manuscript and data storage funding, was involved in the project conception and participated in

# Competing Interests

The authors declare no competing interests.

# Figures



# Figure Legends

Figure 1 Map of study area and hydrophone locations. Points represent data collection sites and transparent bubbles indicate approximate regions for DFO data collection sites (exact coordinates are not provided).

# Tables

Table 2 Deployment summary for the data included in the detection and classification dataset. Annotation start and finish dates represent first and last annotation included in the dataset. Deployment is the name of the deployment location used in the annotations table.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset Provider** | **Location Description** | **Deployment** | **Latitude** | **Longitude** | **Depth (m)** | **Sample Rate (kHz)** | **Hydrophone/Recorder** | **Annotations Start** | **Annotations Finish** |
| Orca Sound | Orcasound Lab | podcast\_round2 | 48.525 | -123.159 | 8 | 32 |  | 2020-09-27 | 2019-07-05 |
|  | Orcasound Lab | podcast\_round3 | 48.525 | -123.159 | 8 | 32 |  | 2020-09-07 | 2017-09-27 |
|  | Orcasound Lab | podcast\_round5 | 48.525 | -123.159 | 8 | 32 |  | 2020-09-08 | 2020-07-25 |
|  | Orcasound Lab | podcast\_round6 | 48.525 | -123.159 | 8 | 32 |  | 2019-07-05 | 2020-09-01 |
|  | Orcasound Lab | podcast\_round7 | 48.525 | -123.159 | 8 | 32 |  | 2017-09-27 | 2020-09-07 |
|  | Port Townsand | podcast\_round9 | 48.135 | -122.760 | 8 | 32 |  | 2020-07-25 | 2020-09-29 |
|  | Bush Point | podcast\_round10 | 48.031 | -122.608 | 12.5 | 32 |  | 2020-09-01 | 2020-09-28 |
|  | Bush Point | podcast\_round11 | 48.031 | -122.608 | 12.5 | 32 |  | 2020-09-05 | 2020-10-18 |
|  | Port Townsand | podcast\_round12 | 48.135 | -122.760 | 8 | 32 |  | 2020-09-29 | 2020-10-08 |
| ONC | Berkley Canyon | Berkley Canyon | 48.426 | 126.174 | 40 | 64 | Ocean Sonics SC2 | 2013-05-20 | 2014-12-04 |
| DFO CRP | West Vancouver Island | WVanIsl | NA | NA | 114 | 16.384 | AURAL-M2 | 2011-05-18 | 2012-05-24 |
|  | Northern Mainland British Colombia | NorthBc | NA | NA | 35 | 16 | SM2M | 2013-10-10 | 2014-02-03 |
| DFO WDLP | Carmanah Point | CarmanahPt | NA | NA | 55 | 192 | SoundTrap 6249 | 2022-03-08 | 2022-06-29 |
|  | Strait of Georgia North 1 | StrGeoN1 | NA | NA | 72 | 256 | AMAR | 2021-09-05 | 2021-10-01 |
|  | Strait of Georgia North 2 | StrGeoN1 | NA | NA | 72 | 256 | AMAR | 2021-11-27 | 2021-11-28 |
|  | Strait of Georgia South 1 | StrGeoS1 | NA | NA | 193 | 256 | AMAR | 2021-11-11 | 2021-11-18 |
|  | Strait of Georgia South 2 | StrGeosS2 | NA | NA | 193 | 256 | AMAR | 2021-09-04 | 2021-09-16 |
|  | Swansen Channel | SwanChan | NA | NA | 245 | 256 | AMAR | 2021-11-13 | 2022-01-09 |
| JASCO/Malahat | Station 3 | Stn\_3 | 48.686 | 123.274 | 237 | 64 | AMAR | 2015-10-14 | 2016-10-08 |
|  | Station 4 | Stn\_4 | 48.507 | 123.211 | 188 | 64 | AMAR | 2015-10-15 | 2017-02-11 |
|  | Station 5 | Stn\_5 | 48.495 | 124.540 | 213 | 64 | AMAR | 2015-10-14 | 2016-02-16 |
|  | Station 6 | Stn\_6 | 48.775 | 123.343 | 74 | 64 | AMAR | 2016-02-16 | 2017-02-10 |
| SIMRES | Tekteksen (East Point), Saturna Island, BC | Tekteksen | 48.780 | 123.052 | 27 | 128 | Ocean Soncics IC Listen | 2022-06-24 | 2022-06-24 |

Table 3 Summary of annotations for each contributor’s detection and classification dataset. CRP indicateds Cetacean Research Program and WDLP indicates Whald and Detection and Localization Program. Detection dataset annotations are divided into killer whale, other or undetermined biological sounds, abiotic sounds, and humpback whales. Population/Ecotype classification task includes annotations for southern resident killer whales (SRKW), Biggs, northern resident killer whales (NRKW) and offshore killer whales (OKW)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  |  | |  | **Species/Class Annotations** | | | | | | | **Ecotype/Population Annotations** | | | | | | |
| **Dataset Provider** | **Dataset name** | | **Detection** | **Annotation Level** | | **KW** | **Und. Bio** | | **Abiotic** | | **HW** | | **SRKW** | **Biggs** | | **NRKW** | | **OKW** | |
| Orca Sound | | Podcast 2 | None | | Call (KW)/ File | 435 | | 0 | | 11 | | 0 | 435 | | 0 | | 0 | | 0 |
|  | | Podcast 3 | None | | Call (KW)/ File | 436 | | 0 | | 60 | | 0 | 436 | | 0 | | 0 | | 0 |
|  | | Podcast 5 | None | | Call (KW)/ File | 0 | | 0 | | 31 | | 0 | 0 | | 0 | | 0 | | 0 |
|  | | Podcast 6 | None | | Call (KW)/ File | 151 | | 0 | | 25 | | 0 | 151 | | 0 | | 0 | | 0 |
|  | | Podcast 7 | None | | Call (KW)/ File | 0 | | 0 | | 167 | | 0 | 0 | | 0 | | 0 | | 0 |
|  | | Podcast 9 | None | | Call (KW)/ File | 198 | | 0 | | 89 | | 0 | 198 | | 0 | | 0 | | 0 |
|  | | podcast10 | None | | Call (KW)/ File | 510 | | 0 | | 113 | | 0 | 510 | | 0 | | 0 | | 0 |
|  | | Podcast 11 | None | | Call (KW)/ File | 0 | | 0 | | 118 | | 0 | 0 | | 0 | | 0 | | 0 |
|  | | Podcast 12 | None | | Call (KW)/ File | 54 | | 0 | | 0 | | 0 | 54 | | 0 | | 0 | | 0 |
| ONC | | Berkley Canyon | None | | Call | 1626 | | 9392 | | 156 | | 2946 | 130 | | 834 | | 0 | | 418 |
| DFO CRP | | WVanIsl | Pamguard WM | | Call | 10384 | | 2757 | | 5054 | | 95861 | 48 | | 5336 | | 4558 | | 258 |
|  | | NorthBc | Pamguard WM | | Call | 6886 | | 10696 | | 1178 | | 26058 | 0 | | 2309 | | 3501 | | 947 |
| DFO WDLP | | CarmanahPt | Pamguard WM | | Detection | 2668 | | 33 | | 297 | | 0 | 1610 | | 694 | | 364 | | 0 |
|  | | StrGeoN1 | Pamguard WM | | Detection | 4777 | | 0 | | 190 | | 131 | 4184 | | 593 | | 0 | | 0 |
|  | | StrGeoN1 | Pamguard WM | | Detection | 324 | | 0 | | 1 | | 42 | 0 | | 324 | | 0 | | 0 |
|  | | StrGeoS1 | Pamguard WM | | Detection | 350 | | 0 | | 3 | | 221 | 159 | | 191 | | 0 | | 0 |
|  | | StrGeosS2 | Pamguard WM | | Detection | 2141 | | 0 | | 152 | | 114 | 2070 | | 71 | | 0 | | 0 |
|  | | SwanChan | Pamguard WM | | Detection | 5655 | | 0 | | 383 | | 1660 | 5630 | | 25 | | 0 | | 0 |
| JASCO/Malahat | | Stn\_3 | JASCO | | Call | 7446 | | 3 | | 852 | | 39 | 2574 | | 937 | | 0 | | 0 |
|  | | Stn\_4 | JASCO | | Call | 15788 | | 0 | | 1177 | | 231 | 7647 | | 318 | | 0 | | 0 |
|  | | Stn\_5 | JASCO | | Call | 3039 | | 0 | | 3311 | | 1084 | 395 | | 994 | | 0 | | 0 |
|  | | Stn\_6 | JASCO | | Call | 1333 | | 0 | | 607 | | 26 | 519 | | 242 | | 0 | | 0 |
| SIMRES | | Tekteksen | None | | Call | 3578 | | 21 | | 0 | | 0 | 3418 | | 0 | | 0 | | 0 |

# References

Baird, R. W., & Stacey, P. J. (1988). Variation in saddle patch pigmentation in populations of killer whales (Orcinus orca) from British Columbia, Alaska, and Washington State. *Canadian Journal of Zoology*, *66*(11), 2582–2585. https://doi.org/10.1139/z88-380

Balcomb III, K. C., & Bigg, M. A. (1986). Population biology of the three resident killer whale pods in Puget Sound and off southern Vancouver Island. *Behavioral Biology of Killer Whales. Alan R. Liss, New York, New York*, 85–95.

Biffard, B., Morgan, M., Muzi, L., Dakin, T., & Buren, P. V. (2022). An Integrated Hydrophone Calibration System for Ocean Observing: ONC HydroCal. *OCEANS 2022, Hampton Roads*, 1–5. https://doi.org/10.1109/OCEANS47191.2022.9976955

Deecke, V. B., Ford, J. K. B., & Slater, P. J. B. (2005). The vocal behaviour of mammal-eating killer whales: Communicating with costly calls. *Animal Behaviour*, *69*(2), 395–405. https://doi.org/10.1016/j.anbehav.2004.04.014

Ford, J. K. (1987). *A catalogue of underwater calls produced by killer whales (Orcinus orca) in British Columbia* (Canadian Data Report of Fisheries and Aquatic Sciences 633; p. 165). Department of Fisheries and Oceans,. https://www.researchgate.net/publication/285709635\_A\_catalogue\_of\_underwater\_calls\_produced\_by\_killer\_whales\_Orcinus\_orca\_in\_British\_Columbia

Ford, J. K., Ellis, G. M., Barrett-Lennard, L. G., Morton, A. B., Palm, R. S., & Balcomb III, K. C. (1998). Dietary specialization in two sympatric populations of killer whales (Orcinus orca) in coastal British Columbia and adjacent waters. *Canadian Journal of Zoology*, *76*(8), 1456–1471. https://doi.org/10.1139/z98-089

Gillespie, D., Caillat, M., Gordon, J., & White, P. (2013). Automatic detection and classification of odontocete whistlesa). *The Journal of the Acoustical Society of America*, *134*(3), 2427–2437. https://doi.org/10.1121/1.4816555

Gillespie, D., Mellinger, D. K., Gordon, J., McLaren, D., Redmond, P., McHugh, R., Trinder, P., Deng, X., & Thode, A. (2009). PAMGUARD: Semiautomated, open source software for real‐time acoustic detection and localization of cetaceans. *The Journal of the Acoustical Society of America*, *125*(4\_Supplement), 2547. https://doi.org/10.1121/1.4808713

Helble, T. A., Ierley, G. R., D’Spain, G. L., Roch, M. A., & Hildebrand, J. A. (2012). A generalized power-law detection algorithm for humpback whale vocalizations. *The Journal of the Acoustical Society of America*, *131*(4), 2682–2699. https://doi.org/10.1121/1.3685790

Joy, R., Tollit, D., Wood, J., MacGillivray, A., Li, Z., Trounce, K., & Robinson, O. (2019). Potential Benefits of Vessel Slowdowns on Endangered Southern Resident Killer Whales. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00344

Kirsebom, O. S., Frazao, F., Padovese, B., Sakib, S., Su, Y., & Matwin, S. (2022). MERIDIAN open-source software for deep learning-based acoustic data analysis. *The Journal of the Acoustical Society of America*, *151*(4\_Supplement), A27. https://doi.org/10.1121/10.0010545

Leu, A. A., Hildebrand, J. A., Rice, A., Baumann-Pickering, S., & Frasier, K. E. (2022). Echolocation click discrimination for three killer whale ecotypes in the Northeastern Pacific. *The Journal of the Acoustical Society of America*, *151*(5), 3197–3206.

Morin, P. A., McCarthy, M. L., Fung, C. W., Durban, J. W., Parsons, K. M., Perrin, W. F., Taylor, B. L., Jefferson, T. A., & Archer, F. I. (2024). Revised taxonomy of eastern North Pacific killer whales (Orcinus orca): Bigg’s and resident ecotypes deserve species status. *Royal Society Open Science*, *11*(3), 231368. https://doi.org/10.1098/rsos.231368

Myers, H. J., Olsen, D. W., Matkin, C. O., Horstmann, L. A., & Konar, B. (2021a). Passive acoustic monitoring of killer whales (Orcinus orca) reveals year-round distribution and residency patterns in the Gulf of Alaska. *Scientific Reports*, *11*(1), 20284. https://doi.org/10.1038/s41598-021-99668-0

Myers, H. J., Olsen, D. W., Matkin, C. O., Horstmann, L. A., & Konar, B. (2021b). Passive acoustic monitoring of killer whales (*Orcinus orca*) reveals year-round distribution and residency patterns in the Gulf of Alaska. *Scientific Reports*, *11*(1), 20284. https://doi.org/10.1038/s41598-021-99668-0

*Ocean Networks Canada—Oceans 3.0*. (n.d.). Retrieved June 12, 2024, from https://data.oceannetworks.ca/home

*OrcaHello*. (n.d.). Retrieved June 19, 2024, from https://aifororcas.azurewebsites.net/

*Orcasound*. (n.d.). Retrieved June 19, 2024, from https://github.com/orcasound

Parijs, S. M. V., Clark, C. W., Sousa-Lima, R. S., Parks, S. E., Rankin, S., Risch, D., & Opzeeland, I. C. V. (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

*Raven Pro* (1.6.5). (2022). [Computer software]. Cornell Lab of Ornithology. https://www.ravensoundsoftware.com

Rice, A., Deecke, V. B., Ford, J. K., Pilkington, J. F., Oleson, E. M., & Hildebrand, J. A. (2017). Spatial and temporal occurrence of killer whale ecotypes off the outer coast of Washington State, USA. *Marine Ecology Progress Series*, *572*, 255–268.

Saulitis, E. L., Matkin, C. O., & Fay, F. H. (2005). Vocal repertoire and acoustic behavior of the isolated AT1 killer whale subpopulation in southern Alaska. *Canadian Journal of Zoology*, *83*(8), 1015–1029. https://doi.org/10.1139/z05-089

Shiu, Y., Palmer, K. J., Roch, M. A., Fleishman, E., Liu, X., Nosal, E.-M., Helble, T., Cholewiak, D., Gillespie, D., & Klinck, H. (2020). Deep neural networks for automated detection of marine mammal species. *Scientific Reports*, *10*(1), 607. https://doi.org/10.1038/s41598-020-57549-y

Wiggins, S. M., & Hildebrand, J. A. (2007). High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring. *2007 Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies*, 551–557. https://ieeexplore.ieee.org/abstract/document/4231090/

Baird RW, Hanson MB, Dill LM (2005) Factors influencing the diving behaviour of fish-eating killer whales: sex differences and diel and interannual variation in diving rates. Can J Zool 83: 257−267

Balcomb KC, Bigg MA (1986) Population biology of three resident killer whale pods in Puget Sound and off southern Vancouver Island. In:Kirkevold BC, Lockard JS (eds) Behavioural biology of killer whales. Alan R. Liss, New York, NY, p 85−95

de Bruyn, P.J.N., Tosh, C.A. and Terauds, A. (2013), Killer whale ecotypes: is there a global model?. Biological Reviews, 88: 62-80. <https://doi.org/10.1111/j.1469-185X.2012.00239.x>

Ford J.K.B (1987) A catalogue of underwater calls produced by killer whales (Orcinus orca) in British Columbia. Canadian Data report of Fisheries and Aquatic Sciences. No 633. Department of Fisheries and Oceans, Nanaimo, British Columbia, Canada. <https://publications.gc.ca/collections/collection_2007/dfo-mpo/Fs97-13-633E.pdf> (Accessed March 31, 2024)

Ford, J. K. B., G. M. Ellis, L. G. Barrett-Lennard, A. B. Morton, R. S. Palm, and K. C. Balcomb III. 1998. “Dietary Specialization in Two Sympatric Populations of Killer Whales (Orcinus Orca) in Coastal British Columbia and Adjacent Waters.” Canadian Journal of Zoology 76: 1456–1471. <https://doi.org/10.1139/z98-089>.

Ford, J.K.B, Southern Resident Killer Whale Call Catelogue. https://orca.research.sfu.ca/ (Accessed March 31, 2024)

Ford, John K.B., Volker B. Deecke, and James F. Pilkington. 2023. Dialects of Killer Whales in Coastal Northeastern Pacific Waters: A Catalogue of Call Types. https://orca.research.sfu.ca/call-library. Accessed 26 Feb. 2024 (or whatever date you accessed it).

Gillespie, D. M., Gordon, J., McHugh, R., Mclaren, D., Mellinger, D., Redmond, P., Thode, A., Trinder, P., & Deng, X. Y. (2008). PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localisation of cetaceans.

Kirsebom, O.S., Frazão, F., Padovese, B., Sakib, S.M., Su, Y., & Matwin, S. (2022). MERIDIAN open-source software for deep learning-based acoustic data analysis. The Journal of the Acoustical Society of America. <https://doi.org/10.1121/10.0010545>

Murphy O (2023) Analysing marine vessel noise impacts on endangered Southern Resident killer whale acoustic behaviour during transits of Boundary Pass in the Salish Sea. MSc Thesis, Simon Fraser University <https://www.sfu.ca/~rjoy/OMurphy_ARP.pdf>

Shiu Y, Palmer KJ, Roch MA, et al. Deep neural networks for automated detection of marine mammal species [published correction appears in Sci Rep. 2020 Jun 30;10(1):11000] [published correction appears in Sci Rep. 2021 Oct 21;11(1):21189]. Sci Rep. 2020;10(1):607. Published 2020 Jan 17. doi:10.1038/s41598-020-57549-y

SIMRES (2020) Hydrophones. <https://simres.ca/projects/hydrophones/> (Accessed February 14, 2024)

Van Parijs, Sofie M., Chris W. Clark, Renata S. Sousa-Lima, Susan E. Parks, Shannon Rankin, Denise Risch, and Ilse C. Van Opzeeland. "Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales." Marine Ecology Progress Series 395 (2009): 21-36.

# Supplemental Information

## HALLO Annotation Guidelines

The following section contains the annotation guidelines that were provided by expert annotators on the HALLO project.

**Annotation Guidelines**

* Draw a box specifying the time and frequency boundaries that contain the call/sound
* Assign labels to the box in the 4 label columns (or as many as possible)
* Although the table below contains values for a variety of species for the ‘Sound.ID.Species’ field, our focus remains on Killer whales. Don’t go out of your way to annotate every single dolphin and vessel noise. All labels other than KW are basically there to add a bit of extra information on difficult cases. So, if there are humpback calls sneaking amidst killer whale calls and you think ‘Oh, this could be confusing! I think it’ll be helpful to mark this right here as a humpback’, please go ahead and label it! But you don’t need to chase down every single signal that is not a KW.

|  |  |  |
| --- | --- | --- |
| Label field | Possible values |  |
| Sound\_ID\_Species | KW | Killer whale |
|  | KW? | potential killer whale (if it was unknown but had the potential to be KW, it fell into this category) |
|  | HW | Humpback whale |
|  | HW? | potential humpback whale (certainly not a KW, possibly a HW) |
|  | HW/KW? | either HW or KW, cannot determine |
|  | PWSD | Pacific White Sided Dolphin |
|  | PWSD? | potential Pacific White Sided Dolphin |
|  | KW/PWSD? | either KW or PWSD, too faint or in descript to determine |
|  | GW | Grey Whale |
|  | GW? | potential Grey Whale |
|  | HW/GW? | either HW or GW, cannot determine without further review |
|  | Odontocete | vocalizations from a small unidentified odontocete, not PWSD |
|  | Echolocation | Echolocation clicks that can’t be safely identified as KW |
|  | Odontocete? | potential vocalizations from a small unidentified odontocete, not PWSD |
|  | Rissos | Risso’s Dolphin |
|  | SPW | Sperm Whale |
|  | Vessel Noise | Vessel Noise |
|  | Clang | some metallic-like anthropogenic clang with unknown source |
|  | Mooring | noise likely derived from the instrument’s mooring equipment |
|  | Sonar | Noise likely due to sonar activity |
|  | Unknown | unable to identify or attribute sound to a definitive class |
| KW\_ecotype  (For Killer whales only) | KWSR | Southern Resident Killer Whale |
|  | KWNR | Northern Resident Killer Whale |
|  | KWT | Transient (Bigg’s) Killer Whale |
|  | KWU | Outercoast Transient (Bigg’s) Killer Whale |
|  | KWO | Offshore Killer Whale |
|  | Unknown | unable to identify or attribute sound to a definitive class |
| Pod  (for KWSR only) | J | J pod |
|  | K | K pod |
|  | L | L pod |
|  | Unknown | unable to identify or attribute sound to a definitive class |
|  | K/L | K or L pod (because of the call types they have in common) |
| Call\_Type | S1 | S1 call |
|  | S2 | S2 call |
|  | ... |  |
|  | EL | Echolocation clicks |
|  | Unknown | unable to identify or attribute sound to a definitive class |
| Confidence | High  Medium  Low | Refers to the entry in the field of highest resolution (eg call\_type if there is something in there, otherwise in KW\_ecotype if that’s the last field that has info, etc).  Confidence to be indicated only if the field has no question mark. If there is a question mark, it is assumed that the confidence is very low.  If there is no question mark and confidence level field is blank, it will be assumed to be high. |

# Notes to Authors

Group Discussion Points

Clarification: calls, whistles, pulsed calls. I would classify pulsed calls and whistles as ‘calls’. Is the typical jargon ‘pulsed calls’ simplified to ‘calls’? Need consistency across labelling or a better understanding

Each section in the methods should contain the following sections

What other figures/tables should we include?

ONC double annotation effort – I’m still vague here. Please someone fill in the blanks. If April and Jenn annotated only some of what Jasper did, then we need to define these as two different datasets – which is fine but we need to know what that is.

**Deployment –** sample rate, deployment depth, hydrophone, recorder, are all files included or only files with annotations?

**Processing –** How ere the data processed? Filters? Detectors (.e.g Pamguard version, whistle moan detector including settings parameters)

**Annotation-** Which type of signals were annotated? KW pulsed calls only? KW whistles and pulsed calls? What was annotated, the detections? The file only?