

901502 - Ocean Soundscape

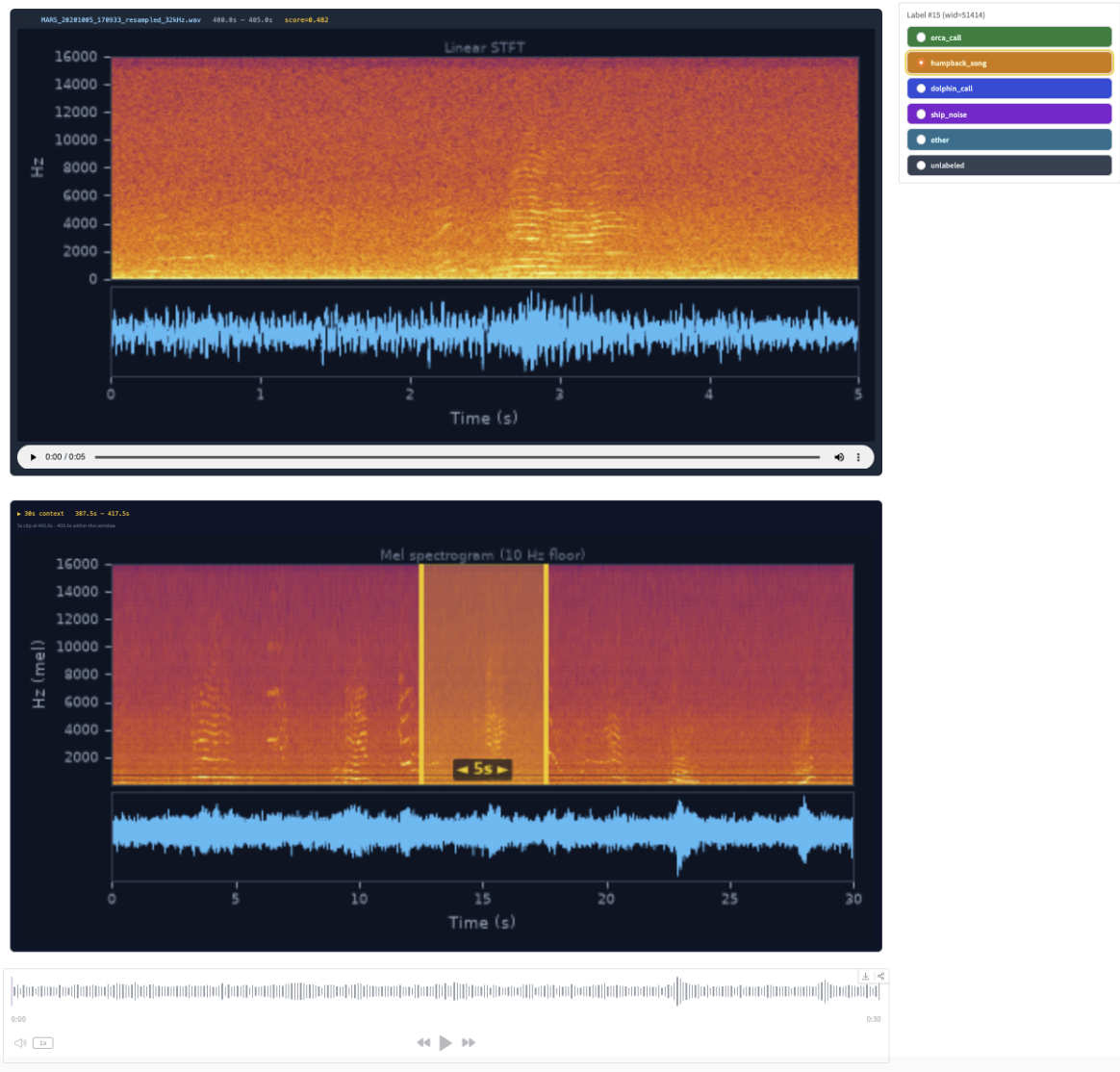
Week 29

1. Danelle Cline
 - a. Team meeting. STM optimization with panacetacea dataset. Minor maintenance on s3 cron. Drafting SMM workshop abstract.
2. Duane Edgington
 - a. Submitted poster proposal to Pytorch 2026 conference (October, San Jose). *"Perch 2.0 and Perch-Hoplite in Pure PyTorch: An End-to-End TensorFlow-Free Bioacoustics Pipeline"*
 - b. Present progress, status, and next steps to the Soundscape team July 15.
 - c. Good progress on analyzing results and ideas to address linear model's classification confusion.
3. Carlos Rueda
 - a. Our intern, Ariel Raizman, continues making good progress on his project.
 - b. Group meeting, with external collaborators, and updates.

Week 28

1. Danelle Cline
 - a. STM model sweeps with panacetacea dataset
 - b. All three soundscape submissions for Oceans were accepted. Many thanks to Carlos, Duane, and John for their good work on this.
2. Duane Edgington
 - a. Great progress in perch2 model and perch-hoplite framework project. Over the 4th holiday I ported perch2 from tensorflow to pure pytorch so I could run native (and fast) on my NVIDIA DGX Spark systems. Great joy. Hard tedious work, but very successful. I hope to present at the Pytorch conference as a poster in October.
 - b. Using "my" perch2 model and perch-hoplite agile modeling system (also ported to pure pytorch and python, no tensorflow), harvesting 5 second annotations of different sound sources, hunting for orcas. Getting great separation of different sounds within the perch2 embeddings
 - c. Enhanced Gradio audio custom annotation tool to show a 30 second context window for each 5 second audio clip to be annotated, which is critical in annotating humpback calls (which go on and on) where a 5 second clip is easily confused with other species. Motivated by John's comments while using my Gradio tool to annotate humpbacks (which I did

not recognize, then).



3. Carlos Rueda
 - a. Intern mentoring, discussions, etc.

Week 27

1. Danelle Cline
 - a. Oceans abstract on PBP accepted. Thanks to the team for their help with the submission.
 - b. STM sweeps with Perch2 model optimizing ARI are in place. Thanks to collaborator Franny for cleaning-up the sound annotations. The single 6 and a half minute humpback recording used for testing had 32 unique sound units !
2. Brent Jones
 - a. Wire Walker deployed

Week 26

1. Carlos Rueda
 - a. Group meetings. Intern mentoring. New VM+GPU in place now (thanks IS!).

Week 25

1. Danelle Cline
 - a. Organized team discussion for proposal input
 - b. Labeled segmented units with Voxel51/Perch2 model. Refining model sweeps to optimize ARI.
2. Duane Edgington
 - a. Proposal work
3. Carlos Rueda
 - a. Our intern, Ariel Raizman, is making great progress with his project.

Week 24

1. Danelle Cline
 - a. Labeling single segmented sound file for optimized STM hyperparameter sweeps. Perch2 model + Voxel51 + Raven Annotation proving useful for this task. ARI with a simple tonal benchmark reached 0.8 confirms the sweeps are good.
 - b. Discussion with our new summer intern, Arie Raizman, Master CS student of UCSC on STM approach, hop-lite, and our overall goals.
2. Duane Edgington
 - a. Great progress and success doing agile modeling on Google perch2 model using perch-hoplite vector database system. Labeled ~200 examples of orca/not orca (roughly 50/50), trained 2 linear classifier models against embeddings from one day of data, and inference one day, with some 220 detects. Will screen detects in RavenPro to assess accuracy.
 - b. Next steps – get embeddings from a month of data, expand the training set, and inference over multiple days.
3. Brent Jones
 - a. WW prep
 - b. WW telemetry package not communicating with inductive modem. Diagnosing

Week 23

1. Danelle Cline
 - a. Continued work on topic model sweeps and summer intern meetings
2. Duane Edgington

- a. Good progress on adapting google perch-hoplite to MARS PAM dataset. First trial running One selected day of data containing known Orca calls near completion.
 - b. Ported google perch-hoplite to dgx spark system.
 - c. Presented status at project team meeting.
 - d. Plan team meeting on 2027 proposal June 4.
- 3. Brent Jones
 - a. Wire walker prep. Swapping in 250m cable
- 4. Carlos Rueda
 - a. Added PBP to <https://deepwiki.com/mbari-org/pbp> (but minimal exploration so far).
 - b. Group meeting

Week 22

- 1. Duane Edgington
 - a. Working on porting perch2.0 hoplite to dgx spark system
 - b. Team meeting

Week 21

- 1. Danelle Cline
 - a. Team meeting. Discussed proposal feedback which was very positive.
 - b. CSUMB student Gabriel completed his work on the Athena database.
- 2. Duane Edgington
 - a. Continued work reviewing 10 minute sound segments flagged by OrcAI with many “orca” detects.
 - b. Meet with John R. to review some of those sound segments – what are they? Orca? Humpback? Other??
 - c. Team meeting May 20.

Week 20

- 1. Danelle Cline
 - a. CSUMB student project complete. AWS Athena database with 99% coverage of soundscape metadata to provide basis for sql queries.
 - b. Working on testing Perch2 model with Humpback groundtruth data.
- 2. Duane Edgington
 - a. Processed one month of PAM data through OrcAI model mining for orca and dolphin vocalizations. To correlate with sightings posted on [California Killer Whale Project](#)

Week 19

- 1. Duane Edgington

- a. Processing November 2021 through Google whale-multispecies model, and through OrcAI model, mining for orca calls.
 - b. Submitted abstract to IEEE Oceans 2026 conference
- 2. Carlos Rueda
 - a. Poster abstract submitted to OCEANS 2026
 - b. Clustering pipeline exercised on “olive” datasets; no definitive conclusions yet.
 - c. Meetings

Week 18

- 1. Duane Edgington
 - a. Completing abstract on orca call detection with ML for IEEE Oceans conference 2026
 - b. Identifying and collecting annotations / clips of orca and dolphin calls recorded in pacific ocean sound recordings.
 - c. Ran job resampling 256 kHz audio recordings 2020 October to 24kHz, 322kHz, 48kHz as required by different models (Google multi-species whale, orcAI, Perch2.0)
 - d. Plan to resample March 2026, March 2019, late December 2024 when pods of orcas were reported by California Killer Whale Project to see if models can detect their vocalizations.
- 2. Carlos Rueda
 - a. OCEANS 2026 abstract/coordination with co-authors
 - b. Looking into exercising the clustering pipeline I recently put together also on various “olive” datasets provided by Franny O. and Jim D.

Week 17

- 1. Danelle Cline
 - a. Running parameter stm sweep optimization for Perch2 model. Embeddings look promising with this model. Evaluating both [perplexity](#) and [ARI](#) metrics.
- 2. Brent Jones
 - a. mtg
- 3. Carlos Rueda
 - a. Group discussion and prepping Oceans 2026 abstracts.
 - b. A fix to the daily humpback whale detection cronjob.

Week 16

- 1. Duane Edgington
 - a. Continuing evaluation of Perch-hoplite Perch2.0 model embeddings for few-shot transfer learning.

- b. Resampling 2018 April month of Pacific Sound data to 32kHz for Perch2.0 . Stored in smb://thalassa.shore.mbari.org/PAM_Analysis/GoogleMultiSpeciesWhaleModel2/resampled_32kHz/2018/04/
- 2. Carlos Rueda
 - a. Experimentation with AVES, UMAP, and clustering toward characterization of humpback whale song units. [Presented results](#) to the team last week, and here's the [code repo](#).
 - b. Using [sonus](#) to continue the experimentation while taking advantage of GPU.

Week 15

- 1. Duane Edgington
 - a. Team meeting
 - b. Reviewing update to [OrcAI](#) repository. Plan to update my clone and run tests.
 - c. Running test of Amrit's [whalu](#) Pacific Sound data fetch and run inference tool. Looking to integrate with Perch2.0 foundation model.

Week 14

- 1. Danelle Cline
 - a. Setup Tator project config in preparation for loading training spectrograms
 - b. Meetings to discuss realtime data use for whale monitoring(pacific coast) and interactive sound exhibits(international)
 - c. Moving AWS upload to near-realtime cron instead of nightly cron
- 2. Duane Edgington
 - a. Team meeting
 - b. Continuing investigation of [perch2](#) - [hoplite](#) for agile modeling of orca / dolphin calls.
- 3. Carlos Rueda
 - a. Internship candidate interviews
 - b. Experimenting with [AVES](#) for clustering humpback whale song detections (as given by the NOAA/Google model via [our tool](#)) into *unit types*.

Week 13

- 1. Duane Edgington
 - a. Continued evaluation/exploration of perch 2.0 - hoplite applied to MBARI soundscape recordings. I am posed to do transfer learning using some of our orca and dolphin recordings harvested from earlier model investigation.
 - b. Team meeting
- 2. Carlos Rueda
 - a. Internship applications review; scheduling interviews.

- b. LLM related literature review
- c. Group meeting

Week 12

- 1. Duane Edgington
 - a. Orcas in the bay reported March 17. Did we hear them in the MARS pacific soundscape recordings? And if we did, can we tell which population of orcas are they, and can we guess why they are in the bay?
 - b. Meeting with Yanwu on his efforts to separate orca and dolphin calls.

Week 11

- 1. Danelle Cline
 - a. Mentoring CSUMB student Gabriel. He is making good progress.
- 2. Duane Edgington
 - a. Team meeting
 - b. Exploring Perch 2.0-hoplite bioacoustics foundation model
- 3. Carlos Rueda
 - a. Meetings. Check point with John to talk about possible next tasks for the project.

Week 10

- 1. Danelle Cline
 - a. CSUMB student weekly meeting. Good progress so far. Gabriel demonstrated Athena queries on PBP generated metadata and presented a general system workflow.
 - b. Upgrades to Shoutcast stream and fixes to email notifications of recording failures
- 2. Duane Edgington
 - a. Reading “Perch 2.0 transfers ‘whale’ to underwater tasks” Dec 2025 – supervised bioacoustics foundation model applied to marine mammal (orca) audio after few-shot transfer learning. Plan to test against MBARI MARS audio data

Week 9

- 1. Danelle Cline
 - a. AWS setup for CSUMB student project and project planning for 10 week program
 - b. Updated recording JSON metadata in bitbucket
 - c. Helped with collaborator, Franny with ROST docker build
- 2. Duane Edgington
 - a. Team meeting

Week 8

1. Danelle Cline
 - a. CSUMB student, Gabriel Meyers, planning for service learning project on AWS
 - b. Fixed issue in the [syncing of backlogged data](#) and revised the wav file size check per the latest recorder. Many thanks to Carlos Rueda for flagging problems with downstream 16K audio processing.
2. Carlos Rueda
 - a. A certain cronjob failing for the last couple days has revealed a more general issue.

Week 7

1. Danelle Cline
 - a. Reprocessed select dates with the realtime processing code [soundpipe](#) for JRyan to review. Blue, fin, and humpback model outputs agree with daily and Matlab processed results. Next step is discussion with stakeholders on outputs.
 - b. CSUMB student Zoom session
 - c. Thalassa outage caused some data loss
2. Duane Edgington
 - a. Team meeting
3. Carlos Rueda
 - a. Group meeting.
 - b. New PBP release (previous one was in Nov 2025).

Week 6

1. Danelle Cline
 - a. STM (Sound Topic Model) bug fix for misaligned documents
 - b. CSUMB student review
 - c. Starting evaluation of Whisper model embeddings
 - d. Finished [soundpipe](#) which will run a series of processors on incoming MARS audio files to monitor for blue/fin/humpback presence.
2. Duane Edgington
 - a. Met with Yanwu. Dolphin vs Orca confusion.

Week 5

1. Danelle Cline
 - a. Revised realtime MARS pipeline with calibrated spectrogram and improved plots. Code walkthrough with J. Ryan to verify correct output per expected 5 second time resolution. Next steps will be feedback from collaborators. [Triton](#) server is working very well for this application.
 - b. Bi-weekly meeting

2. Duane Edgington
 - a. Team meeting
 - b. Explored confidence threshold on OrcAI model. A low setting (0.1) lets in a lot of false positives. The default setting of 0.5 seems about right.

Week 4

1. Danelle Cline
 - a. Realtime-time workflow review with John. Triton inference server deployed and working well on new VM; 10 minute audio processes in less than 40 seconds.
2. Brent Jones
 - a. Recover, clean, de mob wire walker
 - b. Wire Walker data retrieval

Week 3

1. Duane Edgington
 - a. Team meeting Jan 14.
2. Carlos Rueda
 - a. Migrated [humpback-whale-song-detection](#) to using [uv](#) as project manager (from poetry).
 - b. Completed Google/NOAA humpback whale song detection scores for all our time series. Moving forward with a regular update, put in place a daily cronjob to generate the score product for the previous day.
 - c. PBP:
 - i. Added a [Release Notes](#) page to the PBP documentation.
 - ii. Migrated type checker from mypy to [ty](#), which already proves its power by uncovering several issues. A future modernization update is to also migrate to [uv](#).
 - d. (From EC note being submitted by JR). The Ocean Soundscape team met with personnel from NOAA Fisheries and the NOAA Office of National Marine Sanctuaries this week. We are supporting their adoption of [open source software](#) co-developed by MBARI and a European colleague. The software processes voluminous passive acoustic monitoring data to compact standard data products, which in turn enable research for scientific and management purposes.

Week 2

1. Danelle Cline
 - a. Realtime workflow modifications to handle OOM errors for hourly processing.
 - b. Deployed Triton inference server on new VM sonus.shore.mbari.org
2. Duane Edgington

- a. Analyzing output of OrcAI model for MARS soundscape archive month of April 2018. Trying to resolve confusion of Orca sounds with sounds from Dolphins and Baleen whales.
- 3. Carlos Rueda
 - a. Jupyter@gizo now uses a persistent access token.
 - b. Running Google/NOAA humpback whale song detection model to complete the regular scores for 2025. Also doing some maintenance of the code and providing assistance to run the core scoring logic in the realtime workflow that Danelle is setting up.