Passive Acoustic Data Guidance

Passive acoustic platforms used by investigators include bottom-moored Environmental Acoustic Recording System (EARS) buoys, deep-diving sea gliders, and autonomous surface vehicles (ASVs).

Summary of Challenges

- Raw data generated and acquired from passive acoustic platforms can be very large generating tens to hundreds of terabytes of raw data from a single deployment.
- Some passive acoustic platforms record data in proprietary formats that are not usable by other researchers without additional processing or documentation, while others record data directly to a nonproprietary format such as WAV.
- Investigators do not always convert the raw proprietary data files into a non-proprietary format for analysis.
- In some cases, investigators transform raw proprietary acoustic data files into acoustic files in WAV file format. These files may also be very large (terabytes).
- Investigators have committed to storing the raw passive acoustic data files.
- Investigators analyze raw data from passive acoustic platforms using algorithms to detect and measure ambient noise of interest as detection events and further analyze detection events to generate parameters of interest.
- Raw passive acoustic data files will include ambient noises that are not of interest to the original investigator and are not captured in the detection event data; however, other investigators may be interested in these ambient noises.
- Detection event data files are smaller than raw proprietary and processed acoustic data files and can easily be shared via the GRIIDC system.
- Detection events are further processed to generate parameters of interest to the investigator, such as species abundance and distribution, and these data can be shared easily via the GRIIDC system.

Data Sharing Requirements for Passive Acoustic Data

Data sharing requirements are framed according to the level of processing. Acquisition data includes the information describing the deployment of acoustic sampling devices, such as times and locations, platforms, equipment, and methods. Level 0 data is raw audio data typically in a proprietary format as recorded by an acoustic instrument. Level 1 data is Level 0 data transformed into a nonproprietary format, such as WAV.

Level 2 data are detections or measurements of ambient noise. For detection events, parameters describing the acoustic detections, such as time, pitch track, tallies, average call amplitude, and classified call type, are documented in a data file. Level 3 data are derived or interpreted from the acoustic data (level 2), such as population characteristics or other information used in a journal article.

Acquisition information and levels 2, and 3 data should be submitted to GRIIDC. They may be submitted as one dataset or multiple datasets depending on the progress of the research and timing of publications. The level 0 or level 1 data to which they pertain should be referenced in the metadata.

Because of the size of Level 0 and Level 1 data, these data must be stored by researchers at the institution with which they are associated. GRIIDC is currently exploring options for bringing Level 0, or 1 data into stable repositories; therefore, this guidance may change.

Following is a summary of the levels of processing of passive acoustic data and guidelines for data submission.

	Acquisition information	Level 0	Level 1	Level 2	Level 3
Data processing level	Locations, times, equipment, methods of acquisition	Unprocessed audio files at full resolution in proprietary format	Processed audio files in nonproprietary format (e.g., WAV)	Data files of detection events in a standard format	Derived data such as population characteristics and other high-level data products
Submission guidance	Submitted to GRIIDC	Stored with researchers	Stored with researchers if generated	Submitted to GRIIDC	Submitted to GRIIDC