An underwater sound library for the North Sea

Cuyx Bram, Parcerisas Clea, Aubach Ratcliffe Julia, Jälmby Martin, Mendes Jorge, Regel Franziska and Debusschere Elisabeth

Flanders Marine Institute (VLIZ), Jacobsenstraat 1, 8400 Oostende, Belgium
E-mail: bram.cuyx@vliz.be

Passive Acoustic Monitoring (PAM) and soundscape analysis provide powerful tools to unlock biodiversity information embedded in underwater sound. In this presentation, we introduce the Soundlib project, an initiative to build a comprehensive library of acoustic signatures from the North Sea. This work highlights the motivation behind the project, the methodologies employed for data collection and curation, and its potential for advancing marine ecological research.

Sound plays an important role in aquatic communication, making it a powerful tool for monitoring marine biodiversity. The combination of all active sounds in an environment is referred to as the soundscape. In the case of underwater acoustics, the contributing sound sources can be categorized as biophony, originating from biological processes, anthrophony, originating from human activity, and geophony, originating from geological processes. One can assess the occurrence of these three categories to help assess the health of marine ecosystems.

To contribute to this understanding, we employ passive acoustic monitoring (PAM) in the Belgian part of the North Sea (BPNS). Hydrophones deployed on seabed moorings continuously record ambient underwater sounds. This non-invasive technique is gaining traction among marine ecologists for its effectiveness in biodiversity monitoring. Periodically recovered recordings form the foundation of the Soundlib dataset, enabling detailed analysis of marine soundscapes.
Processing these long-term recordings presents notable challenges. Background noise can obscure sound sources, and identifying contributors to the soundscape is labor-intensive. To address these issues, we propose a two-step approach: (1) a pre-processing phase to reduce noise and enhance signal clarity, and (2) an event detection and classification phase utilizing a semi-supervised machine learning (ML) framework.

In our workflow, expert annotators label a subset of the data, which is then used to train a ML algorithm. The trained model processes unseen data, with its results reviewed and corrected by experts to iteratively refine its accuracy. This approach significantly reduces the manual workload while enhancing the utility of the dataset.

By advancing methods for efficient soundscape analysis, this work lays the groundwork for more comprehensive monitoring of biodiversity and human impact in the North Sea, contributing to broader ecological and conservation efforts.

Keywords
Machine Learning; Deep Learning; Acoustics; Dataset; Marine Ecology; Signal Processing; Soundscape Analysis