

SOUNDLIB: AN UNDERWATER SOUND LIBRARY OF THE NORTH SEA

Cuyx, Bram*¹; Parcerisas, Clea¹; Aubach, Julia¹; Jälmy, Martin¹; Mendes, Jorge¹; Regel, Franziska¹; Debusschere, Elisabeth¹

¹Flanders Marine Institute, Ostend, Belgium

Passive Acoustic Monitoring (PAM) and soundscape analysis provide powerful tools to unlock biodiversity information embedded in underwater sound. These non-invasive techniques are gaining traction among marine bioacousticians for their effectiveness in species monitoring. We present the Soundlib project, an initiative to build a comprehensive library of acoustic signatures and long-term sound trends from the North Sea.

As sound plays an important role in aquatic communication, it can be used to assess the health of marine ecosystems by analysing the soundscape in terms of biophony, geophony, and anthrophony.

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 underwater sounds. Periodically recovered recordings form the foundation of the Soundlib dataset, enabling detailed analysis of marine soundscapes.

One of the biggest challenges of analysing the obtained data from a biological perspective is the lack of knowledge on the sounds produced by the species present in the North Sea. Identifying the source of each sound can be very time consuming and sometimes plainly not possible. However, when known sound sources are scarce, unknown sounds can provide valuable information. Creating a library of reference sounds, while allowing unknown sounds to be incorporated, has the potential to identify their producers.

Furthermore, processing these long-term recordings presents notable challenges. Background noise can obscure sound sources, and identifying contributors to the soundscape is labour 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 utilising a semi-supervised machine learning (ML) framework.

To facilitate broader access and collaboration, we have developed a Python-based software development kit (SDK) for the Soundlib dataset. This SDK will allow researchers to query the recordings and their metadata, which includes annotations for detected events. Users will also be able to contribute their own data in the future, both long-term and isolated sound events, which can be processed using the provided standardised methods. By leveraging this SDK, researchers can efficiently analyse marine soundscapes and contribute to a growing repository of acoustic data.

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