Creating training data sets for the automatic identification of wildlife sounds: a narwhal example

Marques, C. S.[1], Dufourq E.[2], Donovan C.[3,4], Marcoux M.[5], Marques, T. A.[1,6,7]

[1] Centro de Estatística e Aplicações, Universidade de Lisboa, Lisboa, Portugal

[2] African Institute for Mathematical Sciences, Muizenberg, South Africa

[3] DMP Statistical Solutions UK Ltd.

[4] School of Mathematics and Statistics, University of St Andrews, St Andrews, UK

[5] Arctic Aquatic Research Division, Fisheries and Oceans Canada, Winnipeg, MB R3T 2N6, Canada

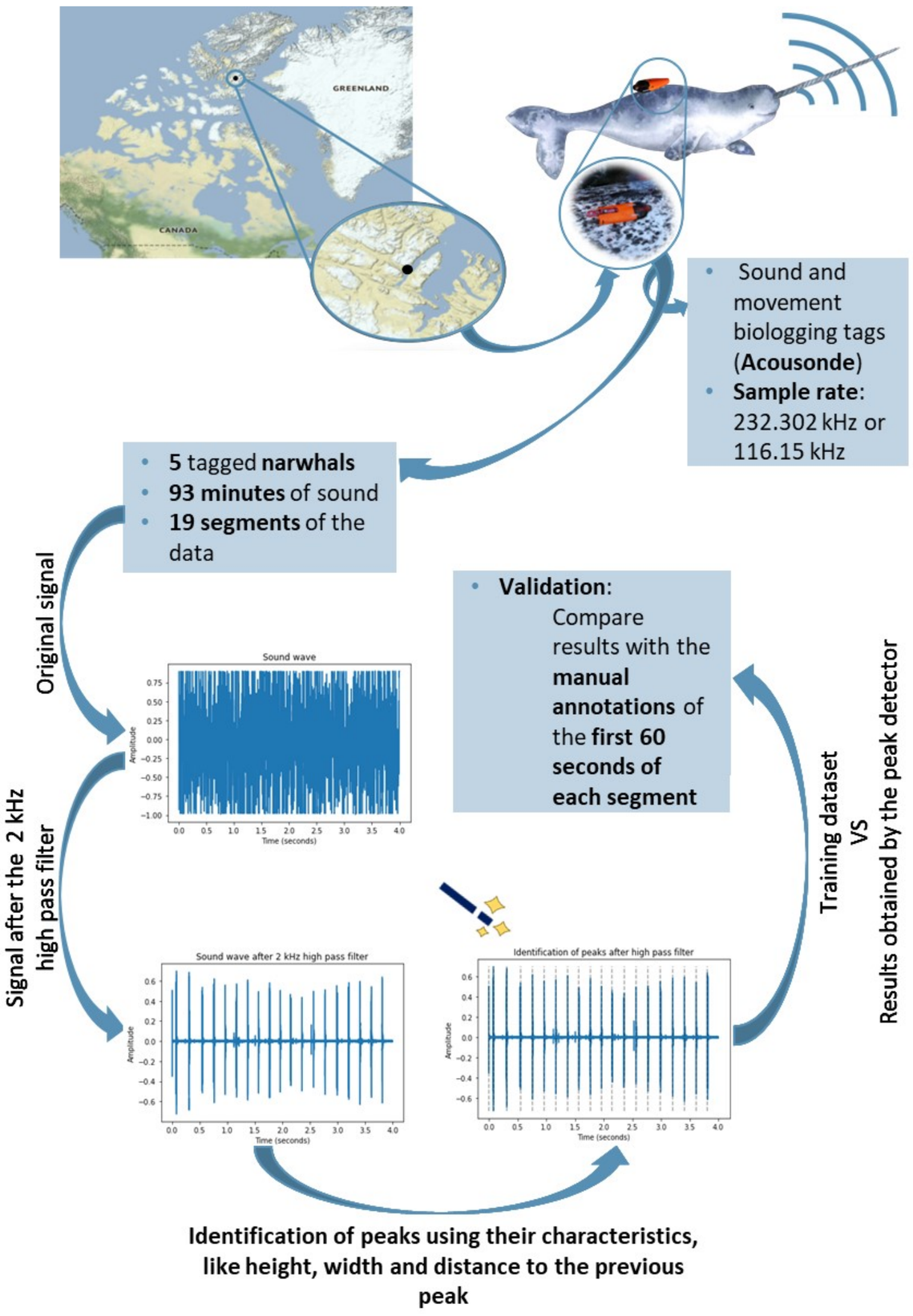
[6] Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, UK

[7] Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

Introduction

- ↑ Capacity for **automated data collection** → **Large amount of data** → Search for **automatic methods** for data processing;
- **For example:** in bioacoustics it is possible to obtain **acoustic data for long periods** leading to **terabytes of data**;
- **Animal vocalizations** are often **embedded in background noise** that can make their **detection challenging** (Huynh et al. (1998));
- **High pass filters remove** low-frequency noise that is often associated with **environmental sounds** while **preserving** the higher frequency sounds that are typically associated with **animal vocalizations** (Xie et al. (2021));
- As a result, **high pass filters** have become an **essential tool** in acoustic ecology and conservation biology, **helping** researchers to **monitor and understand the vocal behavior** of a wide range of animal species.

Methodology



Results

- **31899** clicks identified **automatically**
- **3401** clicks identified **manually** from the **first 60 seconds** of each segment

- **Signal-to-noise ratio:** -11.124 ➡ -0.1691
- **Mean Squared Error:** 0.0531 ➡ 3.350e-03
- **Precision** $\left(\frac{TP}{FP+TP}\right)$: 0.339
- **Sensitivity** $\left(\frac{TP}{FN+TP}\right)$: 0.671
- **False negative rate** $\left(\frac{FN}{FN+TP}\right)$: 0.329
- **Balanced accuracy** $\left(\frac{\frac{TP}{TP+FN}+\frac{TN}{TN+FP}}{2}\right)$: 0.835
- **F-1 score** $\left(\frac{2TP}{2TP+FP+FN}\right)$: 0.451

Where TP= true positives, FP = false positives, FN= false negatives, TN= true negatives

Discussion and Conclusions

The **high pass filter** shows a significant **increases** in the **signal-to-noise ratio**. This leads to an **easier identification of the click sounds** produced by the narwhals. The **automatic detection** was done by **finding the peaks** in the denoised signal, with certain characteristics, like height, width and distance to the previous peak. The comparison between the automatic detection and the manually annotated datasets shows that the **high pass filter generated peaks that are easily identified as vocalizations**. With the use of the high pass filter method we get a **precision of 0.656**, a **sensitivity of 0.707** and an **F-1 score of 0.680**. Therefore, these results show that the **high pass filter** is an easy and **straightforward way to denoise the sound**. This will **allow easier identification of vocalizations** through other methods that are more robust like neural networks.

Acknowledgements

This research was supported by **FCT** through funding of Centro de Estatística e Aplicações (**CEAUL**), ref. **UI/BD/154258/2022** and it was only possible via a sub-award from the **ACCURATE project**, funded by the **US Navy Living Marine Resources program** (contract no. N3943019C2176). We also thank the **Mittimatalik Hunter and Trapper Organization**, and the **community of Pond Inlet** for supporting this research.

References

Huynh, Q., L. Cooper, N. Intrator, and H. Shouval (1998). Classification of underwater mammals using feature extraction based on time-frequency analysis and bcm theory. *IEEE Transactions on Signal Processing 46*(5), 1202–1207.

Xie, J., J. G. Colonna, and J. Zhang (2021). Bioacoustic signal denoising: a review. *Artificial Intelligence Review 54*, 3575–3597.