Transformer for Passive Acoustic Distance Estimation of Cetaceans

Marie-Lou Buisson(1)(2) and Hervé Glotin(1)(2)

(1) Université de Toulon, Aix Marseille Univ, CNRS, DYNI, LIS, Toulon, France (2) Centre International d'Intelligence Artificielle en Acoustique Naturelle

author: marie-lou.buisson@lis-lab.fr other contact: glotin@univ-tln.fr

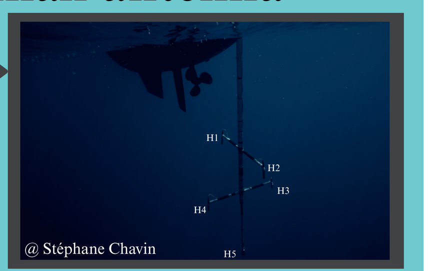
Introduction

- The underwater behaviour of sperm whales *(Physeter macrocephalus)* remains poorly understood, including their trajectories
- Passive acoustics readily estimate their directions, but range estimation remains challenging in case of small hydrophone arrays
- Existing methods based on time differences of arrival [1] or echoes [2,3] require large hydrophone arrays or lack precision

Can deep learning provide a way to estimate distances on a small hydrophone array by integrating various observables?

Materials and Methods

- To enable comparison with a reference track [1, 2], the deep learning method was first applied in the case of a large-aperture hydrophone array, and then to a small antenna
- Small, near-surface hydrophone array →
  Large case: Five bottom-mounted hydrophones spaced from 3.5km to 7.5km



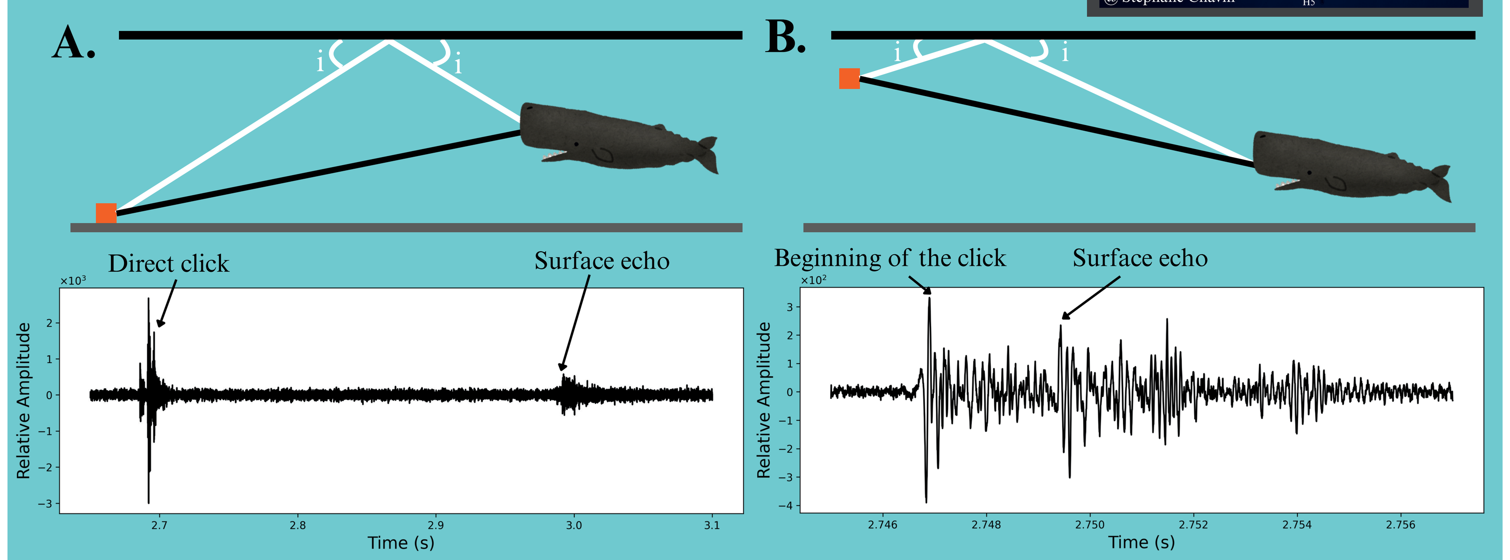
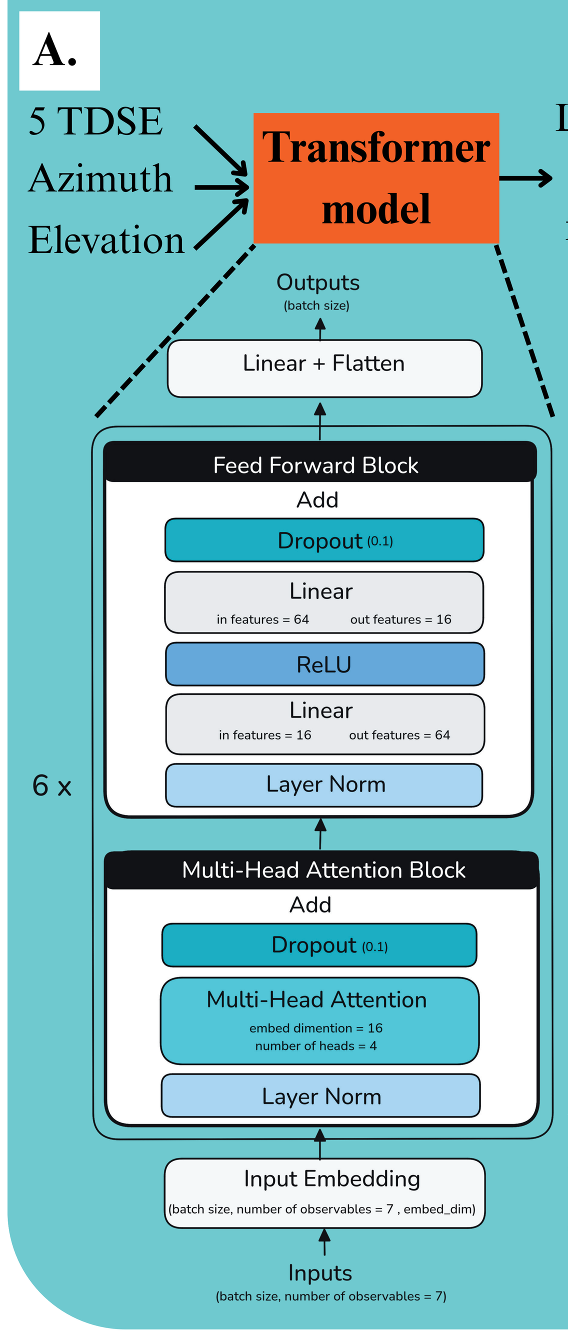


Fig.1: Surface echoes in case of (A) large and (B) small-aperture hydrophone array

- Due to the lack of labeled trajectories for supervised learning, the Transformer models [4] were trained on simulated sources, and then tested on both simulated and field datasets



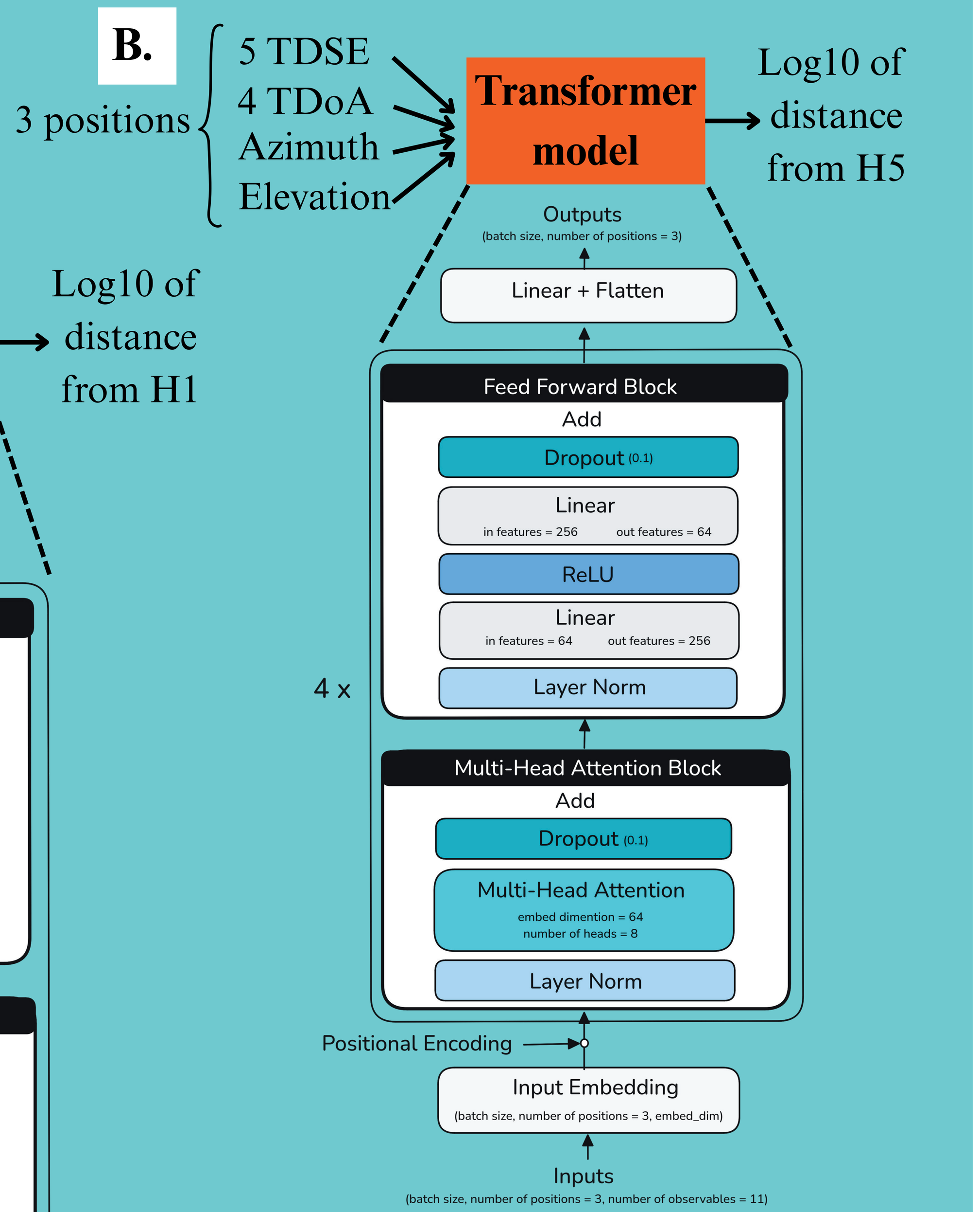


Fig.2: Selected Transformer training setup and architecture for (A) a large and (B) a small-aperture hydrophone array. TDoA: Time Difference of Arrival; TDSE: Time Difference Surface Echo

Results

- <u>Large-aperture hydrophone array:</u>
  - Simulated dataset : **MAPE = 1.41%** **R^2 = 99.8%**

The best performances are around 700m, due to the weights applied during model training

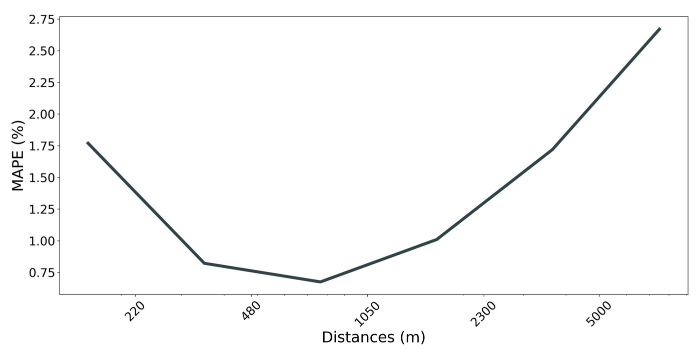
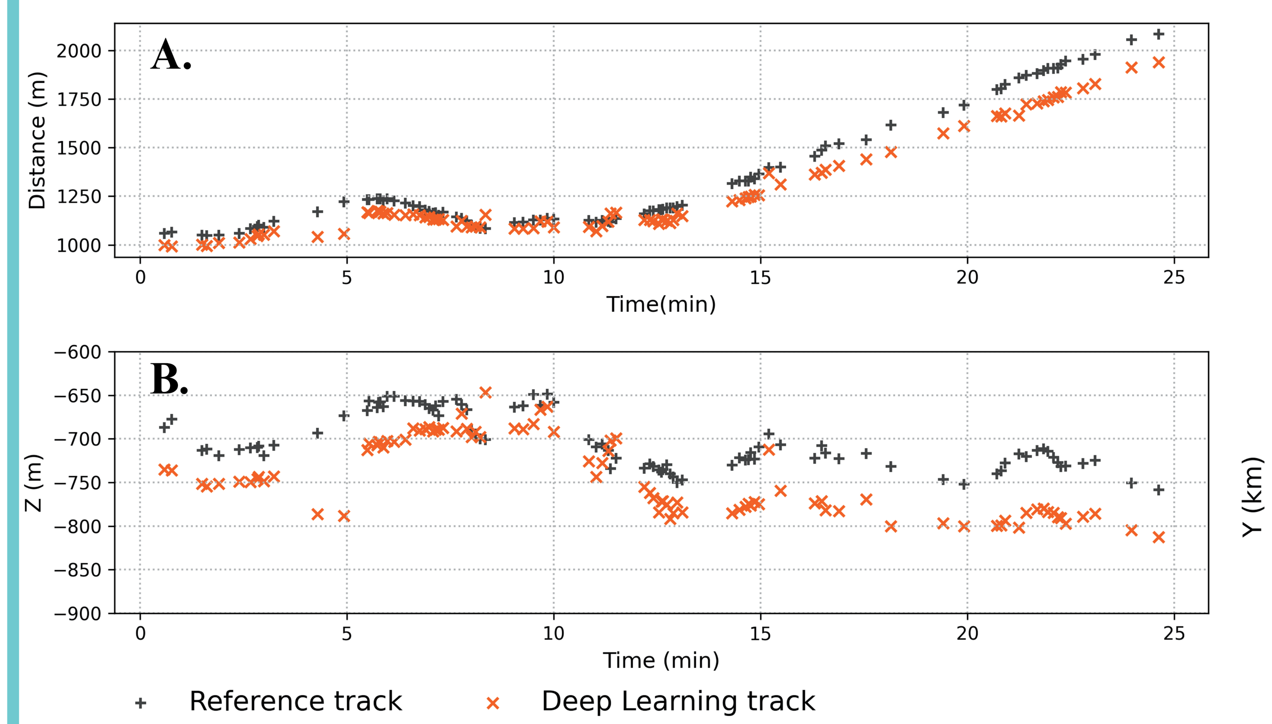


Fig. 3: MAPE on the simulated test dataset for the large-aperture antenna across different distance ranges

- Field dataset:

Comparison with reference track [1,2]: **MAPE = 5.29%**



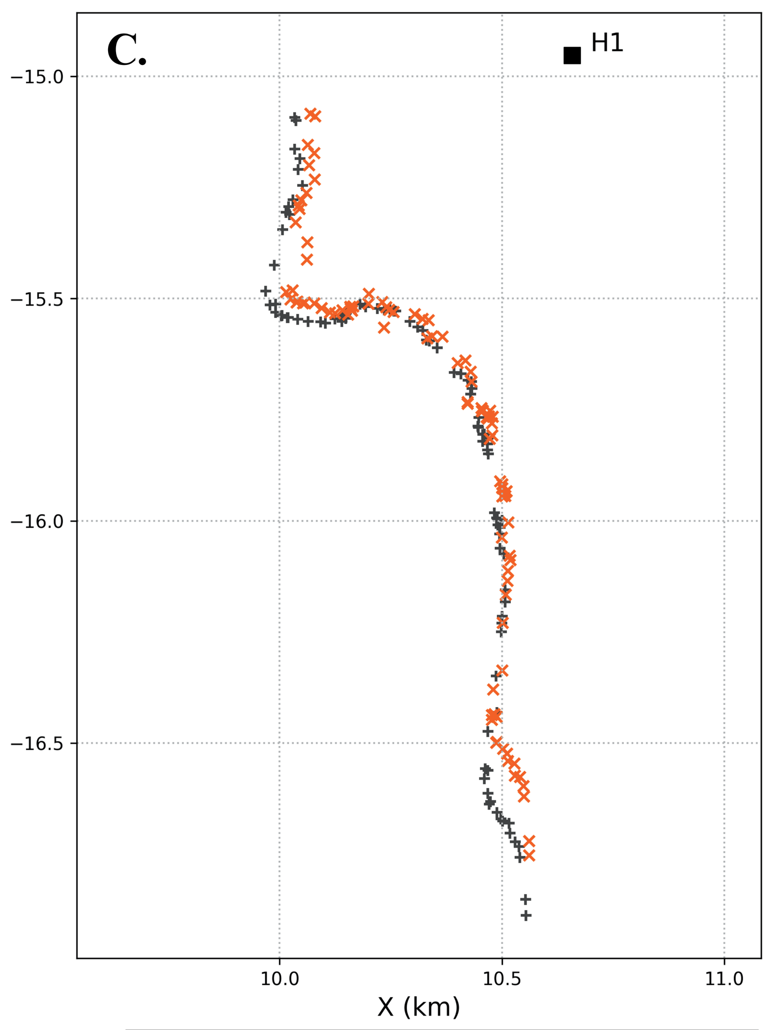


Fig. 4: Comparison between the sperm whale track estimated by our deep learning method (orange) and the multi-validated reference track (black). (A) distance from the first hydrophone, (B) depth profile and (C) 2D horizontal track of a single sperm whale

- <u>Small-aperture hydrophone array:</u>
  - Simulated dataset : **MAPE = 6.14%** **R^2 = 98%**

The best performances are from 200m to 1000m, due to the weights applied during model training

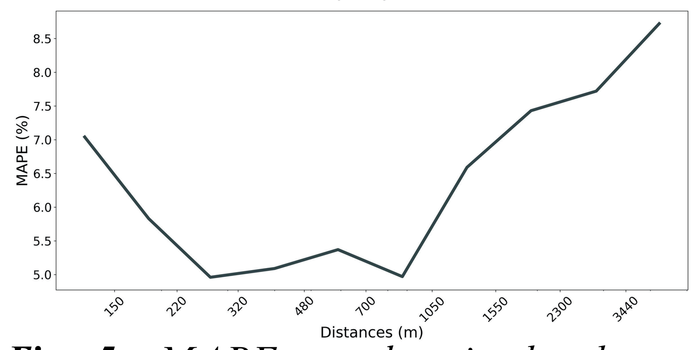


Fig. 5: MAPE on the simulated test dataset for the small-aperture antenna across different distance ranges

- Field dataset:

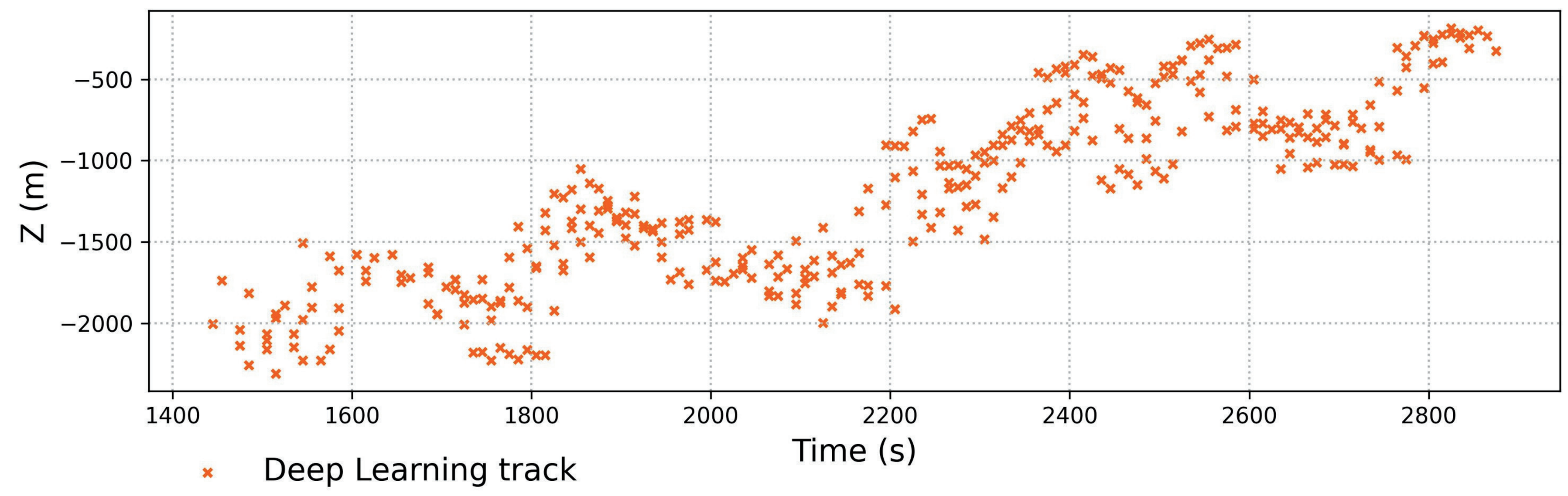


Fig. 6: Depth profile of a single sperm whale during the ascent phase, obtained with our deep learning method. Data recorded during Whale Way missions in the Mediterranean Sea.

Conclusion and Discussion

- With a large-aperture hydrophone array, the obtained track closely matches the multi-validated track and shows a similar temporal evolution
- These promising results, coupled with a MAPE below 10%, support the use of deep learning methods for range estimation in complex scenarios
- On simulated sources and small antenna, the method achieves a MAPE below 10% and an R^2 exceeding 95%, demonstrating its relevance
- On field data, although the estimated depth profile is broadly consistent, significant uncertainties remain in depth estimation
- This deep learning method for range assessment requires testing on more robust datasets, to be collected during future missions

Acknowledgements

We thank AUTEC and NUWC for recording and distributing the test dataset. We also thank Longitude 181, partner of the Whale Way missions, from which the second dataset was extracted. We acknowledge Pascale Giraudet and Stéphane Chavin for sharing code that enabled the extraction of surface echoes and TDoAs. We would also like to thank Franck Zibi for providing the initial Transformer model architecture. H.G. thanks ANR for grants ULPCochlea ANR-21-CE04-0020 and DGA IAD for AI Chair ADSIL ANR-20-CHIA-0014. We also acknowledge Région Sud PACA (ETHAC project) and EUROPAM Biodiversa for funding.

References

[1] Giraudet, P. and Glotin, H. (2006). Real-time 3D tracking of whales by echo-robust precise TDOA estimates with a widely-spaced hydrophone array. Applied Acoustics, 67(11–12):1106–1117.

[2] Nosal, E. M. and Frazer, L. N. (2006). Track of a sperm whale from delays between direct and surface-reflected clicks. Applied Acoustics, 67(11–12):1187–1201.

[3] Aubauer, R., et al. (2000). One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water. Journal of the Acoustical Society of America, 107(5):2744–2749.

[4] Vaswani, A., et al. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30:5998–6008.