



Machine Learning approaches to predict turbidity from multibeam echosounder data

Thomas Hermans¹, Robin Thibaut^{1,2}, Nore Praet³, and Peter Urban^{1,4}

¹Ghent University, Faculty of Sciences, Department of Geology, Ghent, Belgium

²Lawrence Berkeley National Laboratory, Geophysics Department

³VLIZ, Flanders Marine Institute

⁴Ghent University, Faculty of Bioscience Engineering, Department of Data Analysis and Mathematical Modelling

Turbidity is an essential indicator of the water quality in coastal settings as it influences the penetration of light in coastal waters. Next to natural processes, turbidity is influenced by human activities such as dredging or bottom-trawling fishing activities. If turbidity is commonly characterized locally through a range of methods (moorings, tripods, ship-based samples, ACDP), the dynamic nature of turbidity requires the development of 4D monitoring methods. Recent years have seen a growth in the use of multibeam echosounder (MBES) to characterize the water column, such as for the detection of gas bubbles and MBES is also an excellent candidate to characterize the turbidity as the backscatter value is sensitive to density.

A recent study by Praet et al. (2023) analyzed the potential of MBES data to predict the suspended particle matter concentration (SPMC). They identified a linear correlation between the average backscatter data within a sphere of predefined radius and the SPM concentration measured in-situ using a laser in-situ scattering and transmissometer (LISST). The analyzed data revealed a broad variability in the backscatter response as well as a variable correlation within the investigated SPMC range.

In this contribution, we revisit this data set using machine learning approaches to explore non-linear relationships between backscatter values and SPMC, with a special focus on uncertainty. We extended the input variables to the depth and the percentiles of the distribution of backscatter values within the predefined sphere as we anticipate they influence the uncertainty. First, we compared the ability of XGBOOST and a neural network classifier to classify MBES data into three predefined SPMC classes. Both approaches allow to identify with 90% accuracy SPMC belonging to the low value class. The accuracy for the two other classes lies around 60%, indicating the difficulty to discriminate between moderate and high concentration. Then, we used a Bayesian Probabilistic Neural Network to predict the SPMC. The latter outputs not only the estimated value but a full posterior distribution allowing uncertainty quantification. The results confirm the conclusion of the classification, with larger uncertainty observed for larger SPM concentration. Finally, preliminary tests indicate that the MBES data contain enough information to estimate the full particle size distribution within the investigated volume.

Our results reveal a complex relationship between MBES data and SPMC, requiring the use of non-linear approaches to fully exploit the information content of MBES data. The acquisition of new data should enable us to confirm and refine the machine learning models developed in this contribution and eventually use them for monitoring in real-time the turbidity of coastal waters. Particular attention should be paid to the absolute calibration of MBES data in order to use the identified relationship across multiple surveys.

Praet N., Collart T., Ollevier A., Roche M., Degrendele K., De Rijcke M., Urban P. and Vandorpe T. 2023. The potential of multibeam sonars as 3D turbidity and SPM monitoring tool in the North Sea. *Remote sensing*, 15(20), 4918.