Towards representation learning of radar altimeter waveforms for sea ice surface classification

Satellite radar altimeters provide crucial insights into polar oceans and their sea ice cover, enabling the estimation of sea level, sea ice freeboard, and thickness. These retrieval algorithms depend on accurate discrimination between radar altimeter waveforms from sea ice and ocean surfaces in heterogeneous and dynamic surface conditions. A further and less mature step is classifying different sea ice types in addition to the ice/ocean discrimination. We aim to develop new methods for a novel multi-category sea ice and ocean surface classification directly from satellite radar altimeter data to improve sea ice climate data records. Traditional waveform representations are limited to a small set of parameters, leading to information loss. Moreover, machine learning models for sea ice classification often depend on supervised training, which is vulnerable to uncertainties in labeled data, especially in polar regions.

To address these limitations, we explore self-supervised learning methods to optimize waveform representations, which can capture more detailed information for a classification with finer granularity. Furthermore, they do not require labeled data, which is not available at the spatial coverage and resolution of radar altimeter waveforms. We apply these techniques to SRAL data from the Sentinel-3 mission to produce a representation that is optimized for the down stream task of sea ice surface classifiaction. We show that the information preserved in the latent space of an auto-encoder enhances the feature space of traditional waveform parameters, improving the subsequent classification process, when comparing our results to available sea ice charts and other remote sensing products. Our results demonstrate better generalization compared to supervised approaches. However, training only on the reconstruction loss of the auto-encoder lacks a mechanism to explicitly separate dissimilar waveforms. This is directly addressed by the loss of contrastive learning. To apply contrastive learning in a self-supervised manner the positive and negative pairs have to be defined not based on class labels but on information inherent to the input space, information directly associated with the radar altimeter waveforms. Consequently, our adaptation of contrastive learning approaches to radar altimeter data in polar ocean regions is focused on this data engineering part, incorporating expert knowledge from the sea ice domain.