



Classifying ocean profiles with machine learning algorithms

Kimmo Tikka, Antti Westerlund, Pekka Alenius, and Laura Tuomi
Finnish Meteorological Institute, Marine Research, Finland (kimmo.tikka@fmi.fi)

Automatic floats, profilers, and gliders provide oceanographers vast opportunities and challenges. These instruments measure hundreds of thousands of new profiles of varied quantities every year and that data should be available to the scientist as quality controlled data in a manageable form.

Analysis and quality control of such vast amounts of data is a demanding task. Automatic methods for outlier recognition are, consequently, important in process analysis, and also in real time piloting of gliders and or Argo-floats.

The variation of the density of the seawater has an effect on how the glider will fly and how it will react to commands. On a mission, we, therefore, try to forecast possible changes in conditions. Also, we seek to detect if there were some problems with the instruments (QC). Likewise, we attempt to recognize any interesting phenomena ongoing such as anomalies, fronts, and eddies in order to change the observation frequency of a float or to change the route of a glider to measure them.

In stratified sea areas, one has to consider the depth and sharpness of the thermocline while piloting the glider. In several cases, the glider's altimeter may discover a false bottom signal from sharp thermoclines or other non-linearities of the water column. Accordingly, the pilot has to adjust the diving parameters to overcome these problems.

If the accurate automatic classification of the profiles were possible, the data interpretation, the depth of thermocline estimation and also, the piloting system may be partly automatized.

In this work we used clustering and machine learning algorithms developed for time series analysis to classify vertical profiles on stratified brackish water sea area. We endeavored to classify profiles into classes with similar shapes. Then, we defined whether it was possible to define the depth of the upper mixed layer and possible halocline depth in each of these classes. These results may be implemented in the glider and Argo piloting routines.