Dear Editor,

I have read the manuscript entitled "A proxy of subsurface Chlorophyll-a in shelf waters: use of density profiles and the below mixed layer depth (BMLD)" by Zampollo A. et. al. The manuscript deals with the topology of water column profiles (hydrography and Chlorophyll) aiming to relate both in large dataset at a shelf sea region. The work develops methodology for the systematic characterization of the seasonal pycnocline, pursuing objective labelling of transitional depths separating the stratified water column from mixed layers above and below. This applies to the specific circumstance of shelf seas where, below the seasonal pycnocline, there is a bottom mixed layer instead a permanent pycnocline as in Open Ocean.

I consider the subject of the manuscript interesting and I appreciate the authors’ efforts of systematic characterization of water column structure, however I find issues with the methodology and also I find the overall scope unclear. Therefore I cannot recommend its publication at this stage. My main concerns are as follows.

Starting with the methodology, the authors develop an algorithm that delimitate the stratified portion of the water column with special focus on tagging the top of the bottom mixed layer. The algorithm is complex, as can be inferred from Fig.3 diagram. Appraising the need of such iterative clustering-based approach requires time I acknowledge I could not invest for this review, and probably some trials with the code. The authors argue that their method is more accurate than simpler systems as those based in thresholds or gradients. I feel that further comparisons with outcomes produced by simpler methods would provide a stronger case for the use of their complex method. Besides thresholds...
from top and bottom, approaches based on curve segmentation may provide accurate results. To be clear in this point, I think there is likely a detailed analysis of the large profiles dataset that support the development and use of their algorithm instead of others, but at this stage this is not easily assimilated by the reader.

The main focus seems to be classifying whether the Deep Chlorophyll Maximum (DCM) is located above or below reference pycnocline levels (roughly speaking top/middle/bottom). This is strongly dependent on distinct Chla profile shapes, which are classified in 6 types following literature. If I understand properly, Chla profiles have not been classified or clustered following any systematic objective method but manually. Therefore, one of the main strengths of the work (i.e. providing automatic algorithms to process large amounts of profiles) weakens. The analysis of the dataset is therefore mostly manual and the real advantage in objectively tagging density levels to draw scientific conclusions of their dataset is unclear. To improve the manuscript, I suggest applying an automated Chla profiles classification system so the processing gets fully objective and can be applied to much larger datasets.

I am also confused about the overall scope of the paper. Conclusions are e.g. that AMLD is only correlated to Chla for certain fluorescence shapes (HCU), that there is tendency for deep DCMs in shallow waters (even below the BMLD in HCL shapes), and that DCM lies around the centre of the pycnocline for Gaussian (symmetric) Chla profiles. The discussion of Chla shapes is discussed regarding bibliography but not clearly related to the density profiles. In the end I am not sure if the authors aim to infer subsurface Chla values from BMLD in case there no Chla profiles are available. The relationship of the developed tool and/or shelf seas primary production with man-made structures, as well as possible influence of climate change, is too indirect.

The main goals stated along the paper, which are (i) providing a new analytical tool to systematically tag density profiles, (ii) helping to understand basic processes relating Chla and vertical density, and (iii) providing predictive capability for subsurface Chla at fine scales, are in my view not clearly addressed in the ms in its current status. My recommendation is that the article should be returned for major revision. I encourage the authors to focus on highlighting the improvements provided by the developed tool over other methods and describing how their results address the aforementioned main goals.

I provide some specific comments below, mainly regarding sections 1-3, I hope will help to improve further versions of the manuscript.

------

Specific Comments:
Abstract and general. The definition/selection of 8 ‘density layers’ instead of other number is not sufficiently justified. These are levels (discrete depths) instead of layers.

Specific for shelf seas with strong tides. The authors should notice that many shelf seas have small or no tides.

Bryden et al 2005 paper here is not adequate. Scale is too broad and main outcomes are superseded by further results of the rapid array and others.

There are no standard methods to MLD identification neither in shelf nor oceanic waters.

BMLD as an “indicator” of the vertical... Indicator or proxy?

BMLD is indistinctly referred to as ‘bottom mixed layer depth’ and ‘below mld”. Should address this mismatch.

‘this new level of understanding’ sounds a bit presumptuous, maybe just this new algorithm.

‘Fig.2’. It is normally requested to cite figures in order, please check.

‘standard MSS editing procedure’ requires a reference.

Not necessary to specify used functions of TEOS-10, this is too much detail.

General, I guess the authors are using Chlorophyll-fluorescence profiles (from a fluorometer) which is not the same as Chlorophyll-a. Should clarify.

Understand that smoothing/resampling refers only to undulator.
l.149 'The analyses were run in R v3.6.3...' too much detailed. Again in l.204 etc.

l.161-162. I do not understand sentence 'and three equal sections were used to divide the difference between the minimum and maximum Chl-a values into three equal sections'

l.175. Fig.2. why HCL (e) is above HCU (f)? I find this confusing.

l.191 One of the first comprehensive classifications of MLD objective methods available is provided by Thomson and Fine, JAOT, 2003, including curve segmentation aforementioned methods.

l.235 et.seq. why these ad hoc parameters? 2-delta and 90% of the entire profile.

l.240-244. I find confusing that computing the tangent of the angle phi causes issues but computing the angle does not.

l.299 again density layers vs density levels

l.320, Table.2. I miss an explanation for exploring linear regression and 'one-to-one' regression. Should intercept of regression be forced to cross zero for any reason?

Section 3.3.1. I find too many numeric details and data in the text, should be embodied in tables or figs. Same issue in 4.2.