Comment on gi-2021-27
Carine van der Boog (Referee)


General remarks:
This manuscript improves a previously developed algorithm to detect thermohaline staircases from temperature and salinity profiles. As the developer of the first algorithm, I see that the improvements suggested in this manuscript are highly valuable. Especially the possibility to operate on dataset with any regular vertical spacing is a significant improvement to the original algorithm. Besides improving the algorithm, the authors apply their algorithm on glider profiles in the North Atlantic Ocean and conclude that the vertical gradients in temperature and salinity determine the height of the mixed layers in the thermohaline staircases. Overall, I think this manuscript is of substantial interest to the scientific community, but there are a few points that need to be addressed before it is ready for publication.

Major comment:
One of the suggested changes is that the algorithm can now use suboptimal salinity data. As thermohaline staircases have a staircase structure in both the temperature and salinity profiles, I wonder how this adjustment affected the performance of the algorithm. (how do you distinguish thermohaline staircases from thermohaline intrusions, etc.) Because the impact of this change is only discussed qualitatively, it remains unclear whether this simplification can be applied. So therefore, a quantitative discussion on this topic is necessary to justify this adjustment.

Line-by-line comments:

line 33-36: These two sentences suggest here that double-diffusive mixing affects the meridional overturning circulation. However, this effect seemed to be negligible small when considering the contribution of the mixing by thermohaline staircases (https://doi.org/10.1038/s43247-021-00113-x). Therefore, this suggestion needs to be either weakened or given more context.

line 40: What do you mean with 'steps'? Is that the mixed layer thickness? Or do you refer to the height of the temperature and salinity steps between the mixed layers?

line 50: 'Firstly, the Turner angle must fall in the regime favourable to double diffusive
processes of diffusion-convection (−90° ≤ Tu ≤ −45°) or salt finger (45° ≤ Tu ≤ 90°). Secondly, the density ratio must be within a critical range.’

The Turner angle and the density ratio are both indicators of the stratification, and can be transformed into each other with the equation: \( R_\rho = -\tan(Tu+45°) \). So, this means that the Turner angle and density ratio are in essence two different variables to describe the same thing. Therefore, it is not clear to me why the Turner angle and density ratio are described as two different criteria here.

line 72: Does the accuracy of the observations differ between ascending and descending profiles, due to the positioning of the sensors and the turbulent wake induced by the glider? And if so, how does that affect your results?

line 77: what is a shallow dive slope?

line 88: you mentioned that the glides have a vertical resolution of approximately 0.5m (line 77), and mention here that you binned the data into 1m depth bins using the median value of the samples within the bin. Does this mean there are usually only 1 or 2 points in each bin? If so, what do you define as the median of 2 values? Do you take the upper or lower value, or is this random?

line 101: interfaces is more widely used than gradient layers. You can consider using interfaces instead.

line 108: ‘diffusive-convective regime’ should be ‘double-diffusive regime’.

line 110: Which hard-coded aspects do you refer to here? I would assume that, because you binned the data into 1m depth bins (line 88), that you should be able to do a direct comparison with the original algorithm.

line 115: Can you add an explanation why you added a ‘maximum mixed layer height’?

line 120-122: It might be helpful to add your documentation to the supplementary information (or appendix) for reference.

line 135: does the total thickness of the mixed layers also include the interfaces. If so, isn’t it then more a measure for the total height of the staircase?
Where do you show this? You could consider adding a table with these numbers (also numbers from 159-164) to clearly summarize your findings.

'73% of the steps identified were in the salt-finger regime'. Does this mean that the other 27% are in the diffusive-convective regime? Or are they not identified as either regime?

No steps were observed at R<1. Doesn’t this directly follow from the requirement that the Turner angle of the steps should be in double-diffusive regimes?

Because the haline contraction coefficient and thermal expansion coefficient vary over depth, the upper limits of the temperature and salinity gradients vary over depth in terms of their density contribution. Why did you use the temperature and salinity directly instead of their density components?

I do not completely understand why the latter explanation is more likely. Can you elaborate on that?

You miss a reference here.

Looking at 200-250m in Fig. 9, it seems that the masks of the Turner angle (Fig. 9a), the density ratio (Fig. 9b) and the salinity gradient (Fig. 9d) are all valid, while the temperature gradients (Fig. 9c) changes in time. Can you explain why you use all 4 criteria, instead of just the temperature gradient, as that appears to be the governing one.

It is not entirely clear to me on which data the results are based in different sections. For example, the glider data is obtained in the North Atlantic, while they are here compared to thermohaline staircases in the Mediterranean Sea. This is confusing, because Fig. 1 and Fig. 5 contain observations from the Mediterranean Sea.

In the Arctic, the raw data of the Ice-Tethered Profilers can be used to analyze the thermohaline staircases. The vertical resolution of this data is much higher than the Argo floats. I think it is worth mentioning that there has been studies that analyzed the Arctic staircases (for example: https://doi.org/10.1002/2016JC012419), before discussing how gliders can be used in that region as well.
line 226: ‘varying vertical spacing’. Is this varying within a profile?

line 274: What is a ‘step height ratio’?

line 279: ‘We used the parameter set demonstrated in profile iii for this study’. What are exact numbers / settings that you used? And does that mean that you only used temperature profiles to detect the staircases throughout this study? How did that affect your results?

line 301: If your algorithm works on any regular vertical spacing, could you then apply it as well to the dataset of VDB? If so, you can make a comparison between the two algorithms and more quantitatively (compared to Figure 5) discuss how and where the detection of staircases differ. Such an analysis would also clarify whether using only temperature profiles give significant different results.

line 305: can you elaborate more on how you can use your study to improve model sub-grid parameterizations?

Other comments:

the figures do not appear in chronological order.