Response to reviewer comments by Igor Polyakov

Thank you very much for your very insightful review and very constructive comments. We repeat your comments (using a dark red font color) below. Our responses are given in black font. Please excuse the delay of our reply. We realize that additional work will be required to prepare a revised manuscript.

The manuscript proposes a new method for estimating the depth of the Cold Halocline Layer (CHL). The topic is important and warrants a lot of attention in the published paper. Thus, I am very positive that the authors have the potential to produce a nice publication.

However, at the current stage, the manuscript suffers from several major shortcomings:

- I did not find a convincing and satisfactory comparison of the three methods defining the halocline. The majority of the materials presented in the manuscript are about the illustration of the application of the methods and not their comparison. This comparison should include an evaluation of each method's performance and an explanation of the benefits of using each method. Right now, the way it is done is not satisfactory. I would like to see, for example, individual temperature (T) and salinity (S) profiles where the authors show what each method provides and explain the physical reasons for that. I found Fig. 3, which is devoted to method comparison, to be hard to read and not informative.

  We have analyzed individual profiles from ITP74 and ITP33. In response to a comment below, we have also re-processed the data at the native vertical resolution. Reprocessing the data at the native vertical resolution affected individual results, but our overall conclusion regarding the robustness of the methods still holds.

  Figure RP1 shows a revised version of Fig. 2. Figure RP2 shows individual profiles before and after winter convection. In Figure RP1, we excluded profiles that started below 15 m. This strongly reduces the effect of noise in determining the SML base without affecting the overall result.

  We find that investigating individual profiles does help to better understand the differences between the methods. Figure RP2 shows profiles

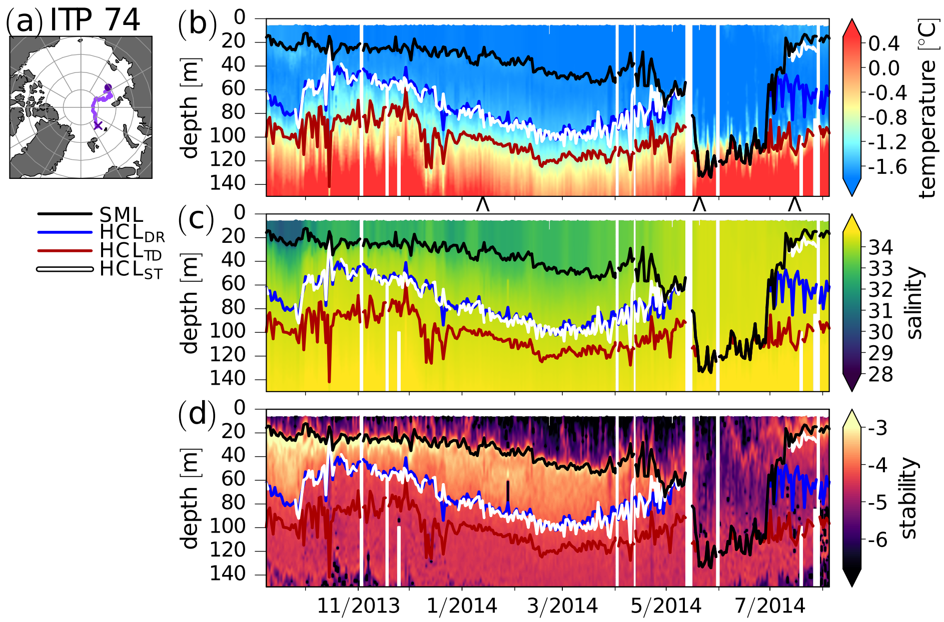

Figure RP1: Modified version of Fig. 1 based on data at the native vertical resolution. Individual profiles at the location of the wedge symbols (∧) below the x-axis in **b** are shown in Fig. RP2. HCL stands for halocline.

before winter convection, during winter convection, and after winter convection from ITP74. Figures RP2a and c for 13 January 2014 show that the TD method misinterprets the halocline (HCL) base. During winter convection (Figs RP2e–h), no HCL was identified. This is because the threshold for identifying the HCL base was already exceeded at the SML base for the DR and the TD method (Fig. RP2f and g), while the threshold was not reached for the ST method (Fig. RP2h). Fig. RP2i for 15 July 2014 (after winter convection) shows freshening and warming near the surface. This indicates that relatively fresh melt or shelf water may have been advected above a colder and saline layer, which had been preconditioned by winter convection. The freshening near the surface leads to a salinity gradient below, and also a stability maximum, which is captured by the ST method (Fig. RP2h). This appears to be consistent with the mechanisms for halocline formation described by Rudels et al. (1996). Rudels et al. (1996) found new halocline formation taking place when relatively fresh shelf waters near the surface were advected above

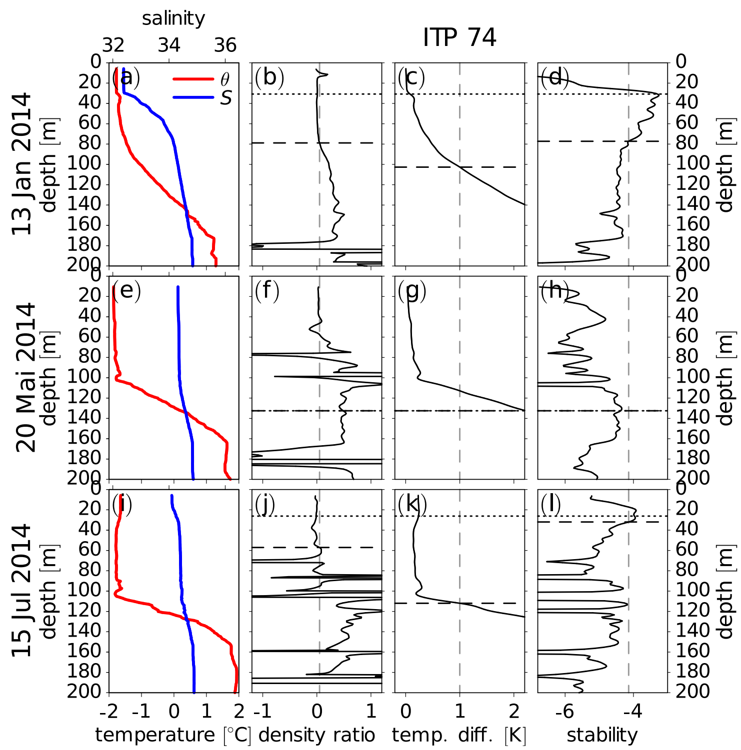

Figure RP2: Potential temperature θ and salinity S, density ratio, temperature difference, and stability from ITP74 before winter convection on 13 January 2014 (a–d), during winter convection on 20 May 2014 (e–h) and after winter convection on 15 July 2014 (i–l). Vertical dashed lines indicated threshold values. Horizontal dashed lines indicate the halocline (HCL) base determined by the three different methods. Dotted lines indicate the SML base. Dash-dotted lines in (f) and (g) indicate that a threshold for identifying the HCL base was exceeded at the SML base.

denser and saltier water below, limiting winter convection. Rudels et al. (2004) and Alkire et al. (2017) stressed the role of melt water in the warm Atlantic inflow through the Fram Strait and via the Barents Sea for halocline formation via this type of capping mechanism. Fig. RP2h for 15 July 2014 (after winter convection) suggests that the ST method might indeed be useful for identifying the beginning new halocline formation via the Rudels et al. (1996, 2004) mechanism.

Figure RP3 for ITP33, which is a revised version of Fig. 7a (now including

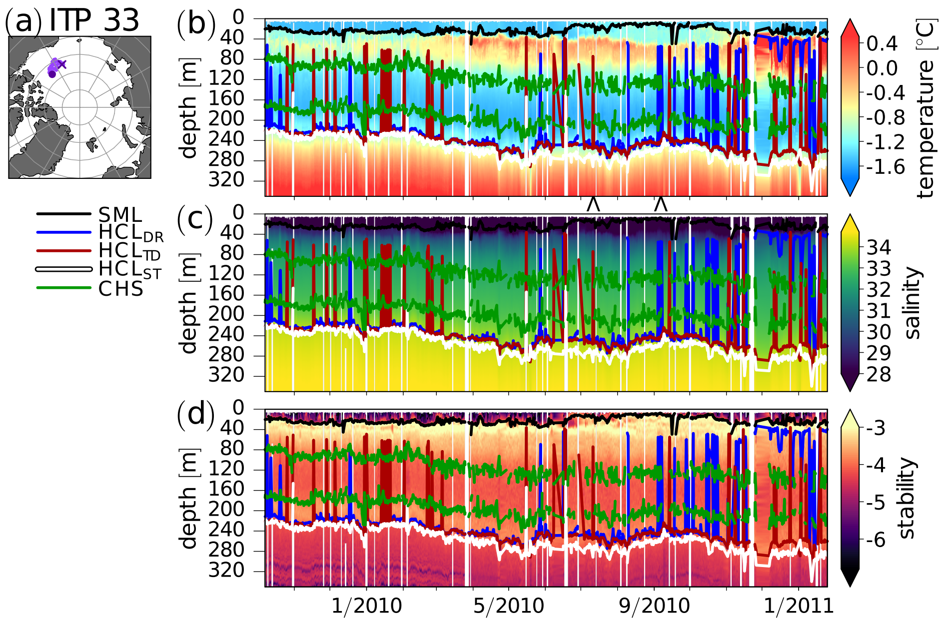

Figure RP3: As Figure RP1 but for ITP33. In addition, the results of the cold halostad (CHS) bound estimation are shown in dark green .

salinity and stability) shows isolated HCL base depth minima for the DR and the TD method, but not for the stability (ST) method.

Figure RP4a–d shows a case from ITP33 on 11 July 2010 in which the TD method produces an isolated HCL base depth minimum and Fig. RP4e–f shows a case on 6 September 2010 in which the DR method produces an isolated HCL base depth minimum. In both cases, the isolated minima are related to the presence of a layer of warm Pacific Summer Water (PSW) around ∼80 m (Figs. RP4a and e). For 11 July 2010, the best estimate of the HCL base depth is provided by the DR method (Fig. RP4b). The DR method correctly places the base of the halocline water at a depth, where the salinity gradient (Fig. RP4a) changes. Stability (Fig. RP4d) also increases markedly at this depth, although the ST method identifies the base of the halocline base about 20 m below this location (Fig. RP4d). The TD method (Fig. RP4c) places the base of the HCL at about 80 m in a layer of warm water, although the salinity gradient below this layer still indicates the presence of a halocline and

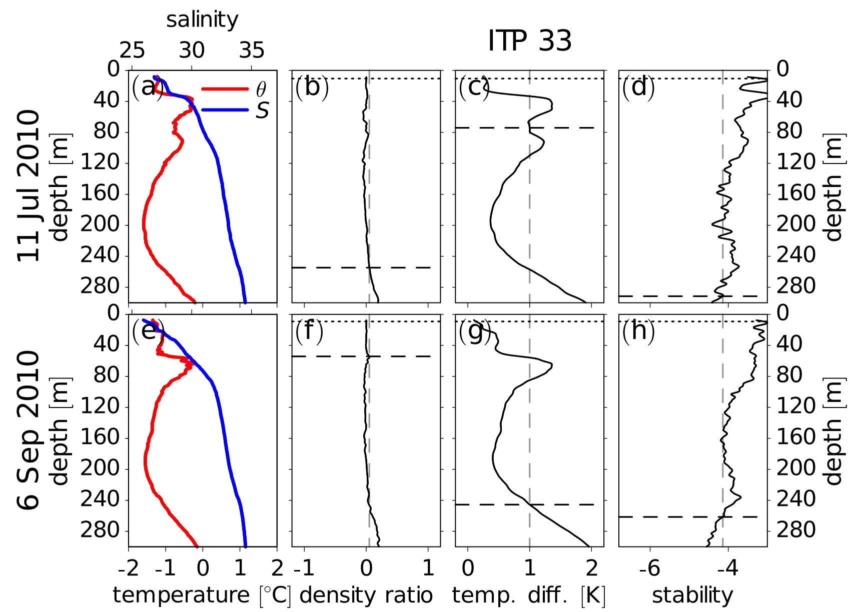

Figure RP4: As Figure RP2, but for two profiles from ITP33. For 12 July 2010 (a–d) the TD method shows an isolated minimum of the HCL base depth, and for 6 September 2010 (a–d) the TD method shows an isolated minimum of the HCL base depth.

temperature decreases below this layer, indicating the presence of Pacific Winter Water (PWW). For 6 September 2010, the DR threshold is exceeded at a local maximum which is related to a very steep temperature gradient (Fig. RP4f) at the base of the PSW. While all three methods rely on finding a threshold, the search direction differs. Because the ST method searches upward, the warm PSW layer does not result in isolated depth minima (Figs. RP4d and h).

Overall, the analysis suggests that the search direction matters. With the DR and the TD method, we search downward, while with the ST method we search upward. This may help to explain why the ST method yields more robust results. We would very much like to further emphasize this point in a revised version of our manuscript. We are planning to remove the discussion of the seasonal cycle and instead to focus more on comparing the methods.

- It looks like the authors misinterpret the water structure of the Arctic Ocean, which has direct implications for their comparative analysis of the three methods used for the definition of the halocline base. In the Eurasian Basin, CHL is the layer where T is close to the freezing point (that is why it is called cold) and S rapidly increases. However, below the CHL, there is the second portion of halocline—the lower halocline water—in which T and S increase with depth. The authors should define what they investigate. In the Canadian Basin, this structure is further complicated by the presence of two other halocline water varieties: Pacific summer and winter waters. Showing the Nordic Seas and Siberian shelves should be excluded from their analysis, for example.

  We erroneously equated the CHL with halocline (HCL). In a revised version of our introduction, we explain the HCL, CHL, lower halocline waters (LHW), PSW and PWW and switched from CHL to HCL unless we are explicitly referring to the CHL. We are planning to revise the subsequent discussion of the results and the title of our manuscript accordingly. Please refer to our response to the reviewer comments by M. Athanase for the revised version of our introduction.

  In Fig. RP5, which represents a revised version of Figs. 5 and 6, we excluded regions shallower than 100 m. Please let us know if you suggest to further adjust this threshold in case we are eventually granted a second round of reviews.

- By the way, it looks to me like the authors miss two golden opportunities to give a beautiful explanation to their methods. a) Fig. 2 from a single ITP record shows a mismatch of results between the authors' stability method (ST) and the density ratio (DR) method after winter convection. I may be wrong, but it is worth checking whether the ST method captures the beginning of the new halocline formation, which is not captured by the two other methods. The halocline formation is a very important topic of Arctic oceanography. b) When the authors analyzed halostad, a possible interpretation may be that they analyzed the boundaries of a variety (winter or summer) of Pacific water. It is worth checking, and if it is true, this may be a nice touch to "sell" the method.

  Thank you very much for pointing this out.

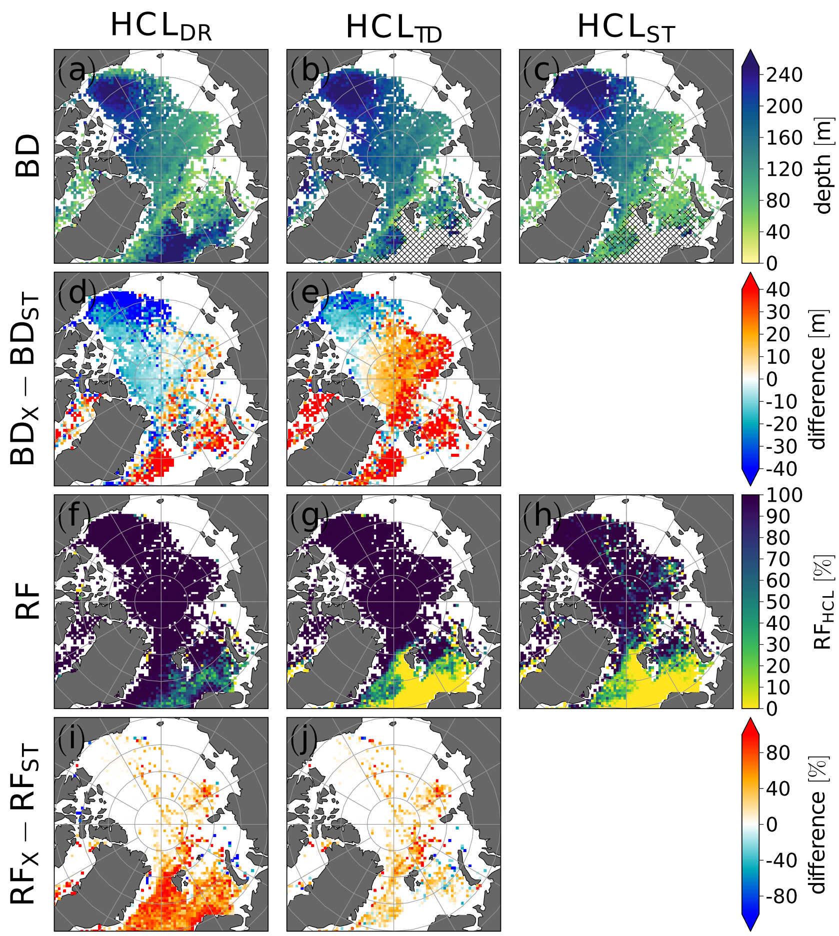

Figure RP5: Revised version of Figures 5 and 6 showing HCL base depth and occurrence frequency, excluding plots for SML depths and kriging results.

a) In our response to your first major comment above, we argue that the ST method may indeed capture the beginning of the new halocline formation based on Figs. RP2i and l.

b) Shimada et al. (2005) explained that the cold halostad (CHS) is formed by Pacific Winter Water (PWW). In the revised introduction, we first introduce PWW and PSW and then explain the link between CHS and PWW as follows:

The PWW could be referred to as a type of cold halocline water (Zhong *et al., 2019), although compared to the CHL in the Eurasian Basin, in the PWW, the salinity is lower and the salinity gradient is smaller. This is why Shimada et al. 2005 called the layer that is formed by PWW a cold halostad (CHS). Similarly, interaction between glacial melt water and Arctic water north east of Greenland forms an intermediate low salinity layer with small salinity gradient which is also called a cold halostad (Dmitrenko et al., 2017). Below, we argue that a lower salinity and a smaller salinity gradient in the CHS compared to the LHW below results in two distinct local stability maxima between the base of the LHW and the SML base: The upper stability maximum is associated with an increase of salinity in the upper PWW. The lower stability maximum is associated with another increase of salinity in the LHW. The lower one of these two stability maxima is absent in the presence of a CHL in the Eurasian Basin (except in regions off the eastern coast of Greenland and also Svalbard where melt water also forms a CHS).*

- The way materials presented is often not good enough – see my detailed comments provided to the authors.

  Thank you very much for the suggestions. Please find a revised version of our introduction based on comments by both reviewers in our response to the reviewer comment by M. Athanase. We are planning to revise the rest of the manuscript as well.

Despite these deficiencies, I believe that the authors can improve the manuscript to the level suitable for publications. That is why I give a "major revision" to this manuscript.

We appreciate your patience.

Below are my specific comments.

Comments:

1. Line 16: As defined, it is not CHL, but CHL+Lower Halocline Water in the Eurasian Basin.

Yes, thank you very much for pointing this out. The cited values from Polyakov et al. (2018) refer to the base of the halocline (HCL) and not the CHL. We revised the introduction based on comments by both reviewers as described in our response to the reviewer comment by M. Athanase. In the revised version, we introduce HCL, CHL, PWW, PSW, and LHW, and describe the structure of the HCL for the Eurasian Basin and the Amerasian Basin.

2. Line 17: 300m may be correct for the Canadian Basin but not for the Eurasian Basin.

   We now write that the Atlantic water is centered near 300 to 500 m in the Eurasian Basin and somewhat deeper in the Canada Basin, citing Macdonald et al. (2015) in addition to Aagaard et al. (1981).

3. Line 43: The halocline has a complex structure indeed, but this was first described long time ago – see papers by e.g. Rudels, Aagaard, Carmack, Steele.

   We revised the introduction based on suggestions by both reviewers. We now first explain the structure of the halocline with CHL, PHW, and LHW, in the first paragraph of the introduction. We cite Bertosio et al. (2020) and Bertosio et al. (2022) for their methods used to determine the base of the LHW in the third paragraph. Please refer to our reply to the reviewer comment by M. Athanase for details.

4. Line 57: Same as for line 43, plus add here Pacific waters.

   In the revised introduction, we explain the structure of the halcoline in the Amerasian Basin at the end of the first paragraph, mentioning the origins of PSW, PWW, and LHW. We explain that according to Shimada et al. (2005) CHS is formed by PWW.

5. Data description.

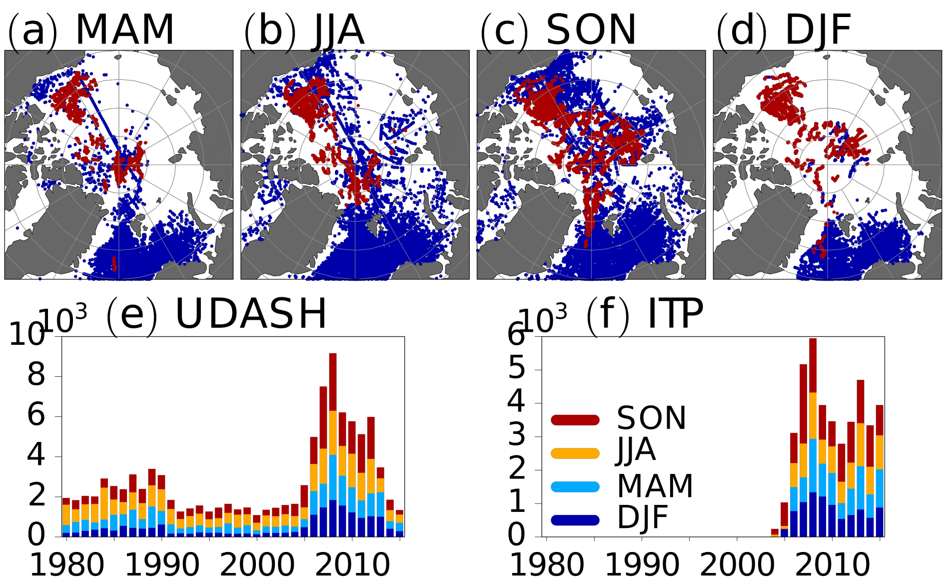

Figure RP6: Map of observations for each season (starting with MAM for March, April, and May) with blue dots for UDASH profiles and red dots for ITP profiles (a–d). Temporal coverage for UDASH (e) and ITP (f) observations.

- Please provide vertical resolution of original data, time covered by observations, show distribution of data coverage in time and space for annual and seasonal coverage, show separately spatiotemporal coverage provided by ITP and other sources. Provide accuracy of observations.

  Figure RP6 shows the spatiotemporal coverage provided by UDASH and ITP data. The vertical resolution for 45% of the UDASH data is finer than 2 dbar. However, ∼25% of the profiles have a resolution coarser than 20 dbar for the deepest five points. We will further investigate this issue and most likely set a minimum requirement for the vertical resolution before recomputing HCL and CHS depths. The vertical resolution for ITP level III data is 1±0.1 dbar. The accuracy of the sensors used for the ITP observatations is 0.002°C for temperature and 0.002 for salinity (Polyakov et al., 2017). The UDASH dataset (Behrendt et al., 2018) was assembled from a number different sources and platforms.

- I do not understand the need of this complex vertical data interpolation (lines 89-90). I suspect that the original (raw) data have 1 or 2m resolution which should be sufficient for the purposes of the study.

  We re-processed the data without prior vertical interpolation.

- Exclude all points from the shelves where there is no typically halocline found. The same is true for the Nordic Seas.

  We excluded regions shallower than 100 m (Fig. RP5). Please let us know if you suggest to further adjust this threshold in case we are eventually granted a second round of reviews.

- Regions are shown in Fig 1 but not used. What is the purpose for that?

  Thank you for pointing this out. We are planning to remove the regions from Fig. 1 in a revised version.

1. Line 101: Why T is given in K, not in C?

   This will be changed to °C in a revised version.

2. Line 113: Shallower, not smaller.

   Yes, thank you. This will be changed.

3. Section 2.2.4 and further discussion: All these methods should be illustrated using individual profiles where everything is clearly marked and visible. Same section: You may use definition of the upper Atlantic Water layer by using 0°C isotherm.

Yes, please refer to Figs. RP2 and RP4 above. We replaced the doubled DT-threshold by the 0°C isotherm.

4. Section 2.2.5: The authors may want to clearly define halostad by giving some physical interpretation for this layer.

   In the revised version of our introduction, we associate the cold halostad (CHS) with PWW based on Shimada et al. (2005) and explain the origin of the term cold halostad (see above). We also explain that in the presence of a CHS, the underlying LHW accounts for a second stability maximum (compare Fig. RP4a and e)

5. Line 144: Taking 0.001m threshold seems misleading since the original data have a 1m vertical resolution (or even coarser).

   We use a vertical interpolation between points, and the 0.001m accounts for rounding errors. We are planning to revise our methods section for clarity.

6. Section 2.3. I found the use of kriging more misleading than helpful in this study.Actually there is no need in that at all since the authors used another much simpler method which serves for the purposes.

   We removed the maps with the kriging results (Fig. RP5).

7. Line 168: Filtration of outliers using such a severe threshold of 25/75% (which is less than 1.5 standard deviation) needs an explanation. Sensitivity study where the authors show how sensitive their estimates to different thresholds may be helpful.

   Outliers were defined as values outside the interval $[Q1 - 1.5\,IQR, Q3 + 1.5\,IQR]$, where $Q1$ is the first quartile, $Q3$ the third quartile, and $IQR = Q3 - Q1$ is the inter-quartile range. The section on kriging will be omitted.

8. Line 198: An example where the authors misinterpret water masses: I think they found that the CHL disappeared but the lower halocline water (not AW) is below the SML.

   We are planning to replace "warm Atlantic water" by "large temperature gradients". The sentence would read: *However, for the model data analyzed in Metzner et al. (2020), large temperature gradients were indeed often found directly underneath the SML during halocline thinning events.*

9. Line 202: Please provide specific for the cases when the criteria were not met: what criteria, why, etc.

   Figure 3 will be removed. Instead, we will discuss individual profiles as suggested. The corresponding figures show thresholds used as criteria (Figs. RP2, RP4).

10. Line 204: I did not understand what is written there.

    The sentence will be removed. Sometimes, thresholds are exceeded already at the SML base. For these profiles, the HCL base depth is initially computed to be very close to zero (instead of exactly zero because of small rounding errors due to floating point operations). These near-zero values are excluded in the computation of average HCL base depth. When computing the occurrence frequency, these profiles are treated as "no halocline detected". The occurrence frequency is defined as number of profiles in which a cold halocline was detected divided by number of profiles. We will try to explain this better in a revised methods section, focusing more on content and less on technical detail.

11. Fig 2: I think this, plus Fig 7, is a good plot to be used for the method interpretation. I suggest to move line definition from the panel - the authors have space below the map. This case would be nice to illustrate further by using individual profiles of T & S for different regimes.

    Thank you very much. Figures 2 and 7 were, of course, inspired by Polyakov et al. (2017). We moved the legend defining the lines and we followed your suggestion to analyze individual profiles (see above).

Similar plot for the Canadian Basin, plus individual profiles from there, would be a good illustration for regional halocline differences.

Please refer to Figures RP3 and RP4, which we discussed above.

1. 3. Not a good figure. I would completely eliminate it since I found no information in the current version.

   The figure will be eliminated.

2. Section 3.2. What does “occurrence” mean in this context?

   The occurrence frequency is defined as number of profiles in which a cold halocline was detected divided by number of profiles. When a threshold was either exceeded at the SML base or not exceeded at all, the profile was treated as “no halocline detected”.

3. Fig 4: Please eliminate point in the Nordic Seas and over the shelves. I suggest to eliminate panels with kriging. The point here is not to show the differences between spatial interpolation methods, but between halocline definition. So, I suggest to show additional panels for differences between methods: e.g. ST-DR and ST-TD. Again, please explain “occurrence”. I did not understand the ”nearest-neighbor” method: If the nearest data point to the grid point is used, why there was averaging then?

   In the revised version of Fig. 4 (now included in Fig. RP5), we excluded points in regions shallower than 100 m, eliminated the panels with kriging, and included additional panels for differences between methods. Regarding the nearest neighbor method, we first define a grid. Then, all the profiles which are closest to a given grid point are assigned to that grid point. We will explicitly explain this point in the revised methods section.

4. 5: The same as for Fig 4: Please eliminate estimates from kriging and compare methods of halocline definition and not the methods of spatial interpolation. I.e. show CHL-DR minus CHL-ST and CHL-TD minus CHL_ST.

   We combined Figs. 4 and 5 in Fig. RP5 and followed your suggestions.

5. Lines 245-247: Often, some analysis of SML is given (like in this paragraph). This is not the topic of the study, first. SML parameters may be given, if they serve the purpose of illustrating the methods. Otherwise, please skip these places. As they are now, they raise questions about newness of these results (e.g. asking for comparison with the previous papers on the subject)

   We eliminated the maps showing the SML depths and we are planning to also eliminate the corresponding discussion.

6. Fig 6: The caption is not clear – the difference is not defined. "(e) to (f)" is not clear. Cos fit is not illustrated.

   The plot will be eliminated. Difference referred to the difference between seasonal and annual mean.

7. Fig 7. Is good, but a) add T &S (like in Fig 2) b) move line definition from the panel.

   Done.

8. Fig 8: What is the point of giving fraction of grid points? What is halostad depth? Give a physical interpretation (Pacific water?).

   We reported the fraction of points because the spatial coverage varies with season. CHS depth was defined here as the depth of the center of the CHS. The center was defined as the mean of the upper and the lower bound of the CHS. The upper and the lower bound of the CHS were defined by a threshold stability. This threshold stability was defined as the mean of the stability minimum in the CHS and the stability maximum in the LHW. We are planning to revise the methods section for clarity.

9. Line 314: What is “semi-saline”?

   We replaced the expression “semi-saline” by “low salinity” based on a suggestion in the reviewer comment by A. Athanase.

10. Line 321: How can we see the point made for the Barents Sea?

    This discussion will be removed from the revised manuscript. We should have written Nansen Basin.

11. Line 322: This discussion and comparison with Steele work needs much more thourough analysis – not just a single line.

    We are planning to revise this discussion based in part on our revised introduction.

12. Line 325 – What does this “lack: mean?

    In a revised manuscript, the word “lack” would by replaced by the word “absence”.

Thank you very much again and please excuse our delay. We very much appreciate your comments. Should the Editor decide to encourage a re-submission, we will be happy to provide more complete answers to some of your comments and a version containing track changes together with the revised manuscript.

References

Macdonald, R. W., Kuzyk, Z. A., and Johannessen, S. C.: It is not just about the ice: a geochemical perspective on the changing Arctic Ocean, J. Environ. Stud. Sci., 5, 288–301, https://doi.org/10.1007/s13412-015-0302-4, 2015.

Zhong, W., Steele, M., Zhang, J., and Cole, S. T.: Circulation of Pacific Winter Water in the Western Arctic Ocean, J. of Geophys. Res. Oceans, 124, 863–881, https://doi.org/10.1029/2018jc014604, 2019.