Response to RC1

Below, please find the reviewer's comments in blue and our replies in black.

This manuscript applies a method of unsupervised machine learning called a profile classification model (PCM) to ocean profile observations in the Weddell Gyre region with the goal to identify and classify areas, or sub-regions, within the Weddell Gyre that share similar temperature and salinity characteristics.

This manuscript clearly highlights how unsupervised classification schemes, such as PCM, can be powerful tools when applied to poorly sampled regions as they are able to identify patterns within highly complex data with no user input. Importantly, the manuscript stresses that PCM is *a complementary technique* in addition to other types of analysis techniques and confirms previously known thermohaline structures as well as sheds new light on more subtle thermohaline patterns within the Weddell Gyre. The authors use PCM to identify and analyze four categories of ocean profiles within the Weddell Gyre as follows: i) the circumpolar class, ii) a transition class, iii) a gyre edge class, and iv) a gyre core class.

This manuscript is clearly written and highlights a powerful, yet under-utilized technique in oceanography and climate science. I think this work is very interesting and I recommend it is accepted for publication after the following concerns are addressed and several modifications are made.

We are glad that you found the manuscript to be clearly written and interesting; thank you for your thoughtful, detailed, and constructive feedback. We have attempted to incorporate your detailed suggestions into our manuscript, and we hope that you find the revised work to be suitable for publication.

Specific Comments

-The authors stress one of the benefits of the PCM method is that it identifies "*both expected and underappreciated* structures" (line 499). However, the results are not clear about which structures are the novel, underappreciated, or previously unknown ones. The manuscript would benefit from a short discussion or clarification on which individual results from the PCM technique are the most critical or important to the research community and which results simply confirm already known patterns.

Yes, we agree that this was unclear in the original draft. On further reflection, we removed the somewhat subjective phrase "expected and underappreciated structures" from the manuscript, since expectations and appreciation may vary between researchers, depending on their particular expertise. However, we have attempted to better emphasize how our results compare with current understanding (see the revised results and discussion sections). In particular, there is now a summary of key refinements to current understanding near the start of the discussion section.

-The description of the training process for the PC model would benefit from more details. The spatial bias is carefully considering in the training process, however the data contains significant temporal biases as well. Summer months are more heavily observed, as well as a general pattern of increasing observations through time (with spikes in recent years as well as around year 2010). How are the seasonal and annual temporal biases accounted for in the training process? What impact may this have on the results?

We agree that the original description of the training was lacking detail. We have expanded this section, which is now in Appendix B2. Also, yes, we agree that the seasonal and annual bias in training is a limitation of our current method. We discuss this in subsection 4.6 in the new draft. Overall, because we are trying to build a climatological picture of the Weddell Gyre region, we do not expect this bias to drastically alter our results. However, this is clearly an area where further investigation would be useful.

Additionally, it is not mentioned how large the training data set is, or what the 'training' process looks like for the PC model. Are the final PCM conclusions sensitive to how the PC model is trained?

We have expanded this section - see what is now Appendix A. Our sensitivity tests suggest that the PCM is robust with respect to how the training is done, considering both the principal component step and for the GMM step. We tried increasing the number of principal components to 10 and saw no appreciable difference in the results - this is likely because the existing six-component PC captures 95% of the variability. We also saw little difference between using the spatially biased dataset and the spatially unbiased dataset, except for some shifts around the Prime Meridian, which is co-located with two GO-SHIP repeat sections and is therefore overrepresented in the spatially biased dataset. Furthermore, the size of the training dataset did not substantially alter our results when varied between using 50% and 100% of the full profile dataset.

-Section 3.2-3.4 Figure 6/7/8 – Are these figures showing the mean of all individual profiles assigned each class at every spatial grid box? Why do you show the mean for these metrics, yet describe the profile classifications with the median? How sensitive are these metrics or the profile classifications to outliers?

Yes, this was due to a technical limitation with xarray plotting features. The xarray package supports calculating the mean at each spatial grid box, but calculating other quantities is more difficult at present. We do not expect our results to be affected by using both of these metrics, since we are only attempting to give indicative values for each class, as opposed to rigid values, i.e. we don't define the classes by these statistics.

Furthermore, did you look at the seasonal variations in space for the depth of mixed layer depth/minimum/ maximum temperature? You suggest the patterns represent deep winter convection in the shelf waters, yet the data is averaged over time for each grid box, and the observations contain more observations during summer months. Figure 10 is helpful to understand the seasonal variability of the classes – but it would be interesting to also analyze the spatial distribution of the depth metrics by season.

Thank you for this suggestion. We agree that the spatial variability of the seasonal cycle is important to consider. To address this, we have expanded the "4_seasonal_analysis" Jupyter notebook to include seasonal plots. We found that the seasonal cycle for each class displayed relatively little spatial variability, with a few key exceptions:

Quantity	Class	Note
MLD	Circumpolar	Summertime values shallower in western part of domain
All	Transition	CDW intrusion most apparent in summer and autumn, partly due to seasonal bias in sampling
MLD	Gyre core	Winter and spring values especially deep in the south
Depth of PT max	Circumpolar	Summer and autumn values are shallower in western part of domain away from the Antarctic Peninsula, corresponding to lower latitudes
Depth of PT min	Circumpolar	North-south contrast in depth is especially apparent in winter and spring, i.e. shallower depths at higher latitudes as expected. Winter water is shallower at higher latitudes and deeper at lower latitudes
Depth of PT min	Transition and Gyre edge	Especially deep summertime and springtime values in the western part of the domain

Overall, the spatial structure of the seasonal cycle mostly highlights the gradients already seen in the climatological average. We have added a paragraph and a new figure on this topic in section 3.5.

(cont'd) For example, are the MLD and min/max temperature depths distinct in wintertime vs summer over the near coast shelf?

There are seasonal contrasts in the transition and gyre edge classes in the near-coastal shelf areas just off the Antarctic Peninsula. We see this for MLD, the depth of the minimum temperature, and the depth of the maximum temperature.

Do we even have observations in those regions in the wintertime to identify known wintertime signatures with this method?

Although there is a seasonal sampling bias, with associated spatial sampling biases, the distribution of profiles covers the regions examined in this study (Figure R1). There are some gaps in wintertime open ocean sampling just east of the Prime Meridian, as well as a persistent gap just east of 40°E. More sampling in these regions, and an eventual revision of our clusters based on improved data, would of course be beneficial.

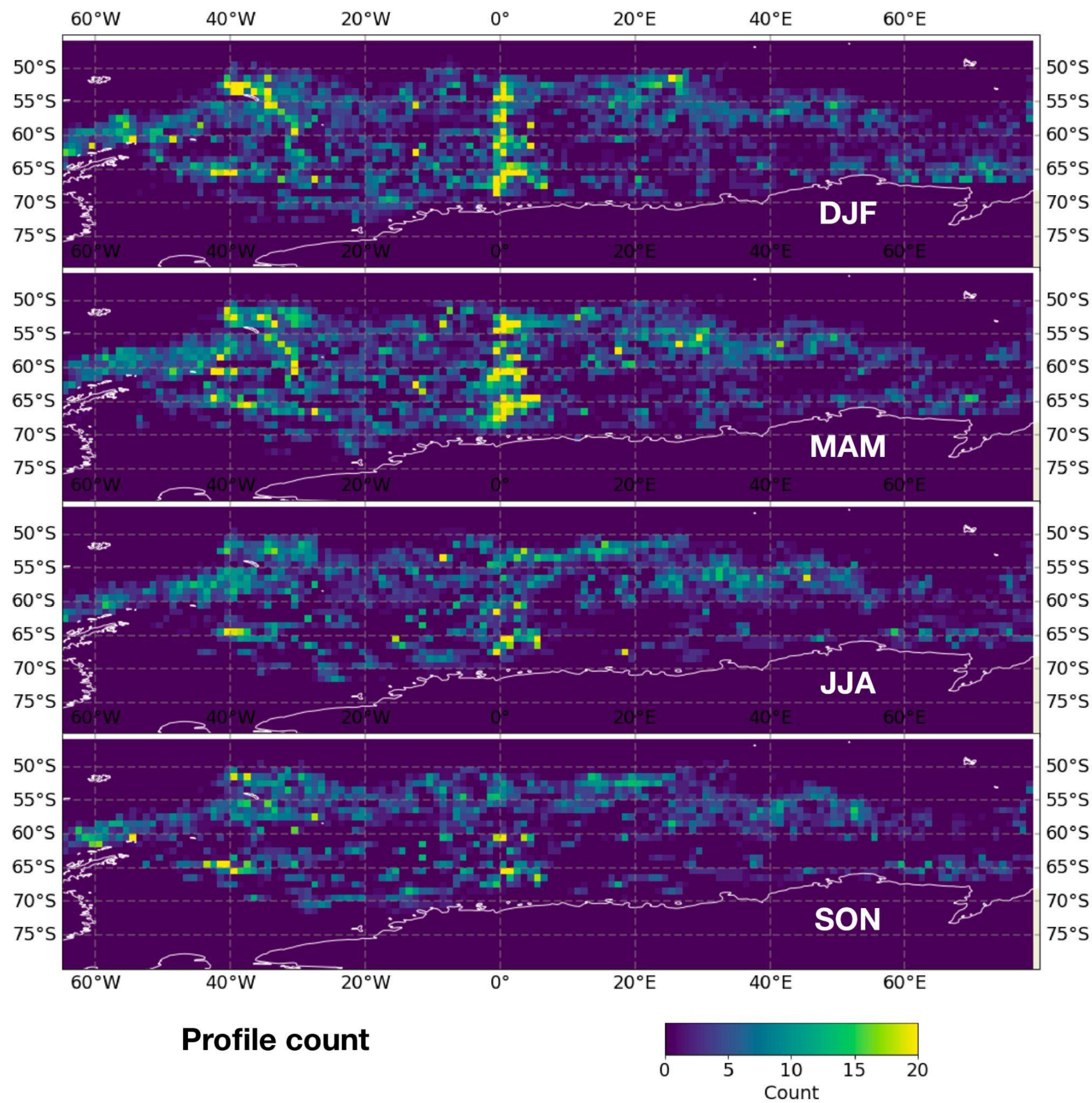

Figure R1. Profile counts in 1° latitude-longitude bins, grouped by season.

-Line 351-351/Figure 12: Part A: The strongest upwelling does appear to be co-located with the circumpolar and transition classes in most of the domain, however there is some strong upwelling in the far western part of the domain (between 40-60W, south of the SBDY) that do not seem to overlap with the circumpolar or transition class profiles, and seems to overlap more with gyre edge profiles, yet the upwelling here is stronger than the general large-scale gyre class upwelling. Do you have an explanation why this region seems to be unique?

Thank you for your thorough investigation of the upwelling region. We believe that the region that the reviewer is referring to is in the northwest Weddell Sea, off the Powell Basin/south to South Orkney (highlighted in Figure R2(a) with the green box). It is interesting that the upwelling signature extends south of SBDY into the gyre edge. And we believe that it is likely due to the distribution of sea ice, sea ice drift velocity, and the resultant surface stress exerted onto the sea surface.

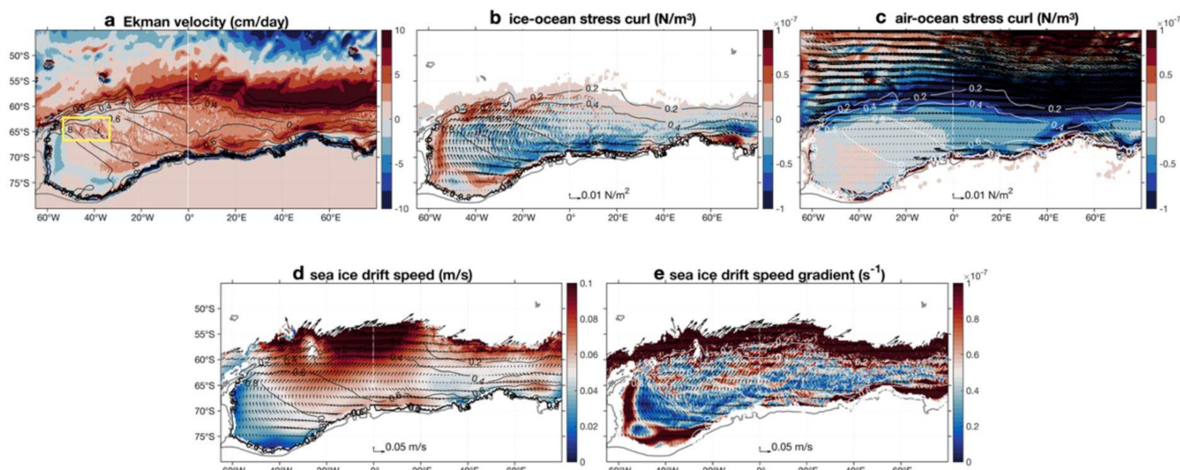

Figure R2. (a) Ekman velocity. (b) ice-ocean stress curl overlaid with ice stress vector. (c). air-ocean stress curl overlaid with air-ocean stress vector. Note that the stress curl and stress vector are all incorporated with the sea ice distribution, hence the weak air-ocean stress over the region with sea ice coverage. (d) ice drift speed overlaid with ice drift vector. (e) ice drift speed gradient over laid with ice drift vector. Contours in all subpanels are climatological sea ice concentration distribution. Each field is climatological.

The Ekman upwelling velocity at the base of Ekman layer is,

$$w_{ek} = \frac{1}{\rho_o}\left[\nabla \times \left(\frac{\tau}{f}\right)\right].$$

Here, τ is the ocean surface stress, which is determined by both the sea ice drift and wind (Dotto et al. 2018):

$$\tau = \tau_{aw} * (1 - SIC) + \tau_{iw} SIC,$$

where τ_{aw} is the wind stress and τ_{iw} is the ice stress. We compared the ocean surface stress curl component associated with ice drift and wind (Figure R2b, c), and it is clearly shown that the ocean surface stress curl is dominated by the ice stress in the gyre edge and gyre core. The highlighted region is characterized by strong negative curl (cyclonic) that leads to a strong localized upwelling pattern. This region is on the edge of a region of year-round sea ice coverage and also a region where the sea ice drift starts to accelerate along with the gyre circulation. This ice drift speed acceleration signature can be seen in climatological ice drift speed map (Figure R2d) and also the ice speed gradient map (Figure R2e), which can contribute to the upwelling signal highlighted by the reviewer. Overall, the strong upwelling signal mostly falls in the transition class, but this particular pattern is of the same magnitude as the rest of gyre edge/core classes.

- Line 351-351/Figure 12: Part B) If more observations existed in the near coast downwelling region, would you expect to identify an additional ‘near coastal’ class in this region? Would the exceptionally large seasonal cycle in vertical mixing in the near coast shelf region impact the results?

At present, the northern and southern branches of the gyre edge class do not separate into separate classes when *K* is increased. This suggests that there are enough structural similarities between the two regions, and perhaps even an advective/dynamical connection, to warrant them being placed in the same class. However, it is of course possible that including more measurements along the near-coastal downwelling region may identify this region as a separate class. One might consider targeting a shallower pressure/depth range and perhaps including seal-tagged profiles to differentiate this structure from the other classes.

- Section 4.6: Figure 14 shows several profiles which lie separately from both the transition class and gyre core class (in PC space) – this grouping is comprised primarily of both transition class and gyre core. Is this separation in PC space meaningful? Does this grouping have some traits in common that results in grouping them together in PC space? For example, are they co-located in space or time within the Weddell gyre region? Or do they have certain temperature/salinity traits that can be attributed to specific PC’s in common so that they are clustered and isolated in this PC space, yet are categorized in different classes? Does increasing the number of classes used in the PCM change how these ‘isolated’ profiles are categorized?

We apologize for the confusion - this figure shows the data in t-SNE space, not the data in PC space. We have revised and expanded this section, including a revised figure that displays the t-SNE transformation using four different parameter values. We hope that this has improved the clarity of this section.

The grouping that you pointed out is still present in our revised t-SNE, but it is not separate from the wider distribution. So, the separation was neither robust nor meaningful. Upon further investigation, we found that this small grouping consists of profiles from the high-latitude, far-eastern portion of the domain, where the transition class and the southern extent of the gyre edge class overlap (in the GitHub repository associated with this paper, see the Jupyter notebook "8_examine_tSNE" for details). Increasing the number of classes does not change how the small grouping is classified (in the GitHub repository associated with this paper, see Jupyter notebook "2d_classify_antarctic-K8" for an example).

- Line 584: Please clarify. What process is applied 20 times for each value of K? It seems that the process of applying PCM can produce multiple realistic results (a different answer for any given iteration). Why were 20 iterations chosen? Are the results sensitive to the number of iterations?

Thank you for correctly pointing out that there was not nearly enough detail in this section. We have added extra details about this process. Specifically, in what is now Appendix A, we have added the following text explaining the process and why we used 20 iterations:

"For each value of *K*, we fit 20 different GMMs using randomly drawn 1,000-profile subsets of the training dataset. We used 20 different subsets because this empirically gave us stable statistics, i.e. doubling this to 40 made no appreciable difference on the distributions of BIC, AIC, and the silhouette score."

To answer your question specifically, no, the results do not appear to be sensitive to the number of iterations.

On PCM producing multiple realistic results: it is true that using different training subsets can produce different PCMs. By using different subsets of the spatially-unbiased training dataset as described in the text added above, we are examining the variability that is associated with the distribution of the training dataset. The resulting statistics help us to evaluate the generality of our GMM. As we argued in this section, *K*=4 is a generalisable and robust choice given this statistical distribution of BIC, AIC, and silhouette score values. Increasing *K* up to *K*=8 may be defensible if greater detail is needed, but this will come at the expense of interpretability and generalisability.

Technical corrections

- The grey profiles on the figures (for example: Fig 3; Fig 10) are very difficult to see. Recommend a darker shade of grey.

Yes, thank you for this suggestion. This was due to our approach of plotting 18% of the profiles for each class, which of course meant that some plots had more lines than others, leading to uneven visibility across the subplots. We have revised our approach to include a random 1000 profiles (Fig. 3) or a random 400 profiles (Fig. 10) instead, producing more consistent visibility across the subpanels.

- Line 89: define ENSO

Done.

- line 135 – either define PF or spell it out.

Changed.

- Figure 2a: add units/label to color bar.

Done.

-Figure 2b: There are only 11 bars plotted in the monthly chart.

Thank you so much for pointing out this oversight - well spotted! We took the opportunity to review and adjust both distribution plots accordingly. The updated monthly plot now more clearly displays the bias towards observations in austral spring and summer. Additionally, the new distribution by year utilizes five-year bins, providing a more concise visualization. Thanks again for catching this - very helpful.

- Line 189: define WW

Done.

- line 245-255- Figure 5 is never referenced

Fixed. It is now referenced in section 3.2.

- line 251- the text specifies the ‘mean’ yet, the metric given is the median.

Fixed.

- line 290: typo – ^{o}C?

Yes, good catch. Changed to °E.

- Figure 11: make color bar labels+unit text larger

The revised Fig. 11 now features:

- Larger x-axis and y-axis labels
- Larger x-axis and y-axis tick labels
- A single enlarged colorbar for the entire figure, instead of several smaller, identical colorbars
- Bigger colorbar tick labels and a bigger colorbar label
- Grey contours have been replaced with white contours for better contrast

- line 428 – change ‘lighter’ to ‘less dense’

Agreed, changed.

- Figure A1 caption: typo. quantity shown *"is"*

Fixed.

- The contents of Appendix A are very difficult to interpret since the methods are not presented until Appendix B. Add references to Eqns B1 and B2 for the AIC and BIC, and refer the reader to Appendix B.

Yes, this is a good point. We have switched the order of Appendix A and Appendix B, such that the methods are discussed first. We have also added the references to equations A1 and A2.

- Line 567 and Eqns B1, B2: K is never defined

Fixed, thank you.