We would like to thank the Referee for his/her helpful comments. They will allow us clarifying our discussion in the revised version of the manuscript and avoid some misinterpretation of our analysis. Please find our answer to the comments below.

The physical processes related to ice-ocean interaction in winter and summer are very different. Annual average would mix the results from those processes. For example, the top 200m would likely include both seasonal mixed layer and winter mixed layer. The annual average of heat content in the top 200m mixes results from stabilization of water column due to ice melt and vertical overturning due to ice formation. Annual surface salinity also reflects sea ice melt in summer and brine injection from ice formation in winter. What is the meaning of averaging the two processes? Which process primarily contributes to the positive feedback at interannual time scale? I would appreciate some discussion addressing the issue.

We totally agree that the major processes vary strongly between the seasons but for the majority of the variables, the trends (which are the main focus of this study) in annual mean are representative of the summer and winter trends. We have thus chosen to present only the annual means in the paper to avoid redundancy. For instance, the spatial pattern of the trend in sea ice concentration is different in winter and in summer because of the seasonal cycle of the ice extent (see Figure R1 below) but this information is included in the annual mean trends. Changes in surface salinity or surface temperature are also similar in annual mean, summer and winter. For some other variables, such as the depth reached by convection, only one season was shown in Figure 3 as it is only useful for winter. Specifically, we do not consider that it is simple to separate the role of sea ice melt in summer and brine rejection in winter as both are strongly linked and it is their combination that results in the long term trend. We, of course, do not claim that changes in melting and brine rejection are not interesting per se but this is not the framework we have chosen for our analysis.

Regarding the simple 2-layer model, it is not quite clear how the entrainment is treated between layer 2 and the layer below. When seasonal mixed layer (layer 1) is eliminated in winter, the overturning will entrain warmer and saltier water into layer 2. The negative feedback kicks in. The process will occur even there is no deep convection. What makes the system transform from a negative feedback at seasonal time scale to a positive feedback at interannual time scale? Why S1 and S2 are reduced each year? Does the result depend on the assumption of very thin layer 2 in the model? If there is a positive feedback at interannual time scale, should we see ice thickness increasing instead of decreasing from year to year (figure 10c)? Downward salt flux is overly simplified explanation since S2 does not increase over years. Considering the exchange between layer 2 and deeper ocean, the entrainment would bring salty/warmer water into layer 2 (upward salt flux). More explanations are needed for readers understand the result.

After reading this section again we recognize that the role of the exchanges between the bottom level of the model and the layers below was not well explained while it is important to understand the behaviour of the system. The two layer model gave the impression that a negative salt flux from the deep ocean was needed to stabilize the system. This is unrealistic as stated by the Referee. We will thus modify the simple model, introducing a third level, to show that the salt exported out of the first layer is stored in the pycnocline (whose depth is also changing) because of the modified salinity profile that results from the brine release and ice formation. Specifically, the fact that sea ice thickness is decreasing with time is the consequence of the negative feedback associated with mixed layer deepening. Besides, the positive feedback leads to a stabilization in a state in which sea ice is produced every year while sea ice is never produced in the unperturbed case. This interplay between positive and negative feedbacks will be discussed in more details in the revised version of the text.

Some detailed comments.

Page 7, and figure 3, Is the heat content at 200-500m affected by sea ice formation? It would be helpful to show the stratification of LOVECLIM as function of latitudes or spatial distribution of the winter mixed layer depth, so readers can see whether 200- 500m is relevant.

We first have first to apologize that we have made a mistake in the submitted version. All the plots were for the depth range 100-500m and not 200-500m. The winter mixed layer is generally close to 100 m in LOVECLIM so this depth range is the one located immediately below the mean mixed layer depth. This will be specified in the revised version. Our results are not qualitatively affected by the precise choice of the depth interval as illustrated below in Figure R2.

Page 14. Regarding LOVECLIM simulation, "The overall trend in ice extent is of -38+/- 93x103 km2 per decade". Do you really mean a negative trend (decreasing ice extent)?

Yes. This part of the discussion will be modified to make it clearer in the revised version.

Labels in figure 3 and 7 need to be enlarged.

The labels in Figure 3 and 7 will be enlarged in the revised version as suggested.

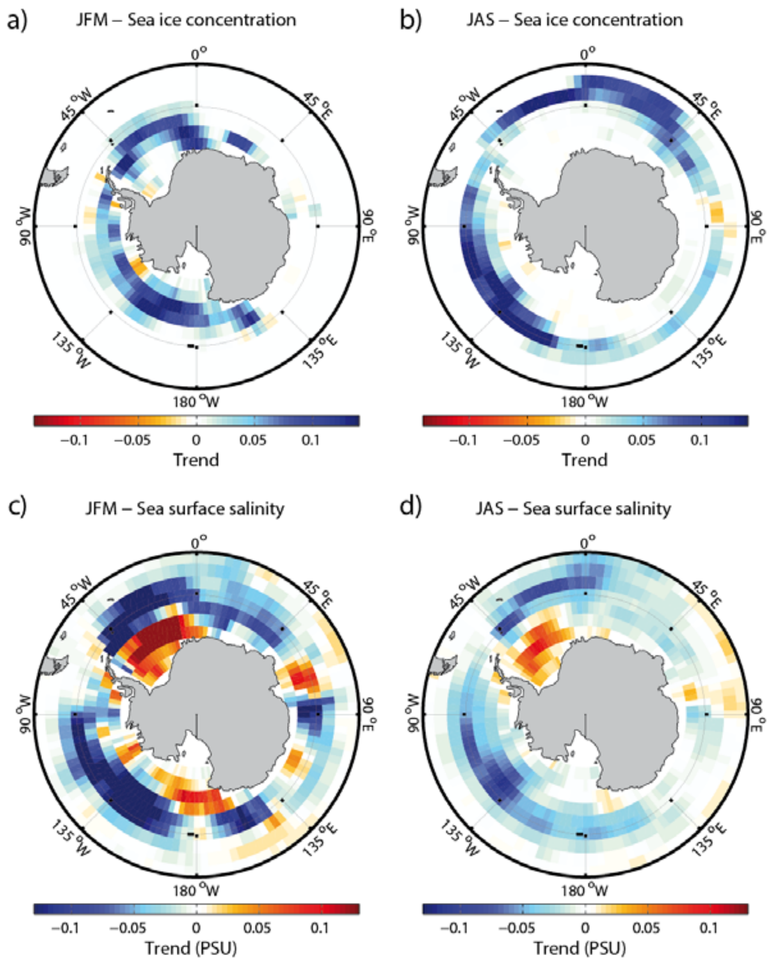

Figure R1. Trends in seasonal means averaged over the 11 periods showing a large increase in Antarctic sea ice extent scaled to represent the 30 year changes of a) ice concentration (mean over January February March), b) ice concentration (mean over July August September), c) sea surface salinity (mean over January February March), d) sea surface salinity (mean over July August September)

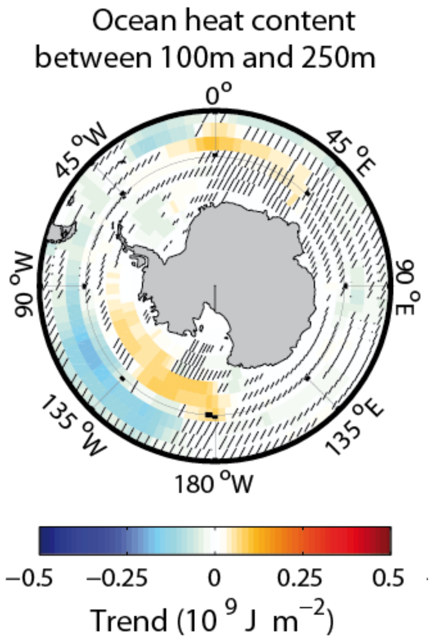

Figure R2. Trends in annual means averaged over the 11 periods showing a large increase in Antarctic sea ice extent scaled to represent the 30 year changes of ocean heat content in the layer between 100 and 250 m (Jm^{-2}).