Review 1

General comments –

The paper provides an assessment of the mass budgets in the MIZ/non-MIZ regions in Arctic sea ice simulations. It uses a comprehensive sea ice model which has been well utilized. It acknowledges that the lack of active atmosphere and dynamic ocean components affects feedbacks in the system, although some deeper discussion of how this might affect the results could be helpful. Overall, I believe that the experiment design is reasonable and can provide some valuable insights on changing Arctic sea ice mass budgets and their projection into the future. However, as noted below, I believe that there is a need for a somewhat deeper analysis and more interpretation of the results.

For example, the study explores differences between simulations with different atmospheric forcing. However, it provides little information on why the ***NCEP forced vs HadGEM forced runs*** *simulate different behavior. Additionally, the comparison of MIZ versus non-MIZ mass budgets is given but again there is limited information on* ***why the mass budgets differ across these different regions****. I do believe that there can be value in delimiting the analysis into MIZ/non-MIZ/changing to MIZ regions. However, I felt that there was a missed opportunity within the manuscript to better articulate* ***why this approach was useful and what new insight*** *it provided relative to previous studies on sea ice mass budgets. Additionally, it would have been useful to discuss the* ***implications from these new insights on broader questions*** *such as the future evolution of the sea ice and/or discrepancies across models, for example. Overall, I believe that some deeper analysis (including possible new analysis) and deeper discussion of the study implications are needed. With changes in these areas, I believe that the originality and significance of this work would be clearer and this would result in a more impactful study.*

Thank you for your suggestions. We have incorporated the answers to these general comments and questions into an added Discussion section which we feel strengthens the impact and clarity of the study. We have addressed each point you have raised separately below and linked it back to text we have added or changed from the revised manuscript.

NCEP vs HadGEM2-ES behaviour: We have included a section showing the differences in the NCEP and HadGEM2-ES forcing fields and have tried to come back to this in the results and the discussion sections. As would be expected there are differences in the atmospheric fields,

from Section 2.2, second paragraph:

“The surface air temperatures are significantly higher in the NCEP reanalysis than HadGEM2-ES between December to early April in both the 1980s and 2010s period…

The shortwave and longwave radiation values are very different between the two data sets. NCEP has much higher summer shortwave radiation values, whilst HadGEM2-ES has much higher year-round longwave radiation values, but particularly summer values. It is likely that this dramatic difference is due to differences in cloud cover."

Then later we link the differences in melt rates to the difference in the surface air temperature field in Section 3.2:

"Melting occurs first in the outer regions (*always MIZ* regions) and progresses inwards across the sea ice cover to the *always pack ice* region. This is more pronounced in the NCEP simulation, likely a reflection that the NCEP atmospheric forcing is warmer in summer (see Figure 1a)."

"Peak melting rates increase, particularly in the NCEP simulation where they increase by 49%, compared to a 17% increase in the peak melting rate in the HadGEM2-ES simulation. The difference is likely due to the greater summer surface air temperatures (see Figure 1a) which drives a larger seasonal sea ice cycle (see Figure 4a)."

We then come back to discuss the differences between the NCEP and HadGEM2-ES forced simulations in the added discussion section:

"We chose to run simulations with the NCEP and HadGEM2-ES forcing so that the NCEP forced simulation could act as a check on the HadGEM2-ES forced simulation, which is projected to 2050. The two simulations were relatively similar in terms of sea ice extent, MIZ extent (see Section 2.4) and MIZ fraction (see Figure 7). Largely the overall results and proportions of growth and melt were similar in the two simulations, however the changes between the 1980s to the 2010s were generally larger in the NCEP forced simulation. This includes the changes in volume fluxes in both regions (see Figures 8a-f), reflecting the larger reduction in summer sea ice volume between the two periods (see Figure 5a). The differences between the NCEP and HadGEM2-ES forced simulations volume changes are largely a reflection of HadGEM2-ES having a much lower sea ice volume in the 1980s, however as the simulations become closer over time, it is difficult to assess whether the HadGEM2-ES simulation is underestimating the change from the 2010s to the 2040s."

MIZ vs non-MIZ mass budgets and reasons for the differences: We have tried to incorporate the answer to this in several places throughout the manuscript. Starting with in the introduction we have included paragraphs detailing MIZ processes and a qualitative description of contrasts we might expect to see based on the definition and conditions in the MIZ:

"The strength of sea ice is strongly dependent on the SIC. For 80% SIC (the upper MIZ boundary), we can estimate (Hibler 1979) that ice strength is less than 2% of its maximum. In the MIZ internal stresses in the ice play only a small role and sea ice is essentially in free drift. The sea ice in the MIZ behaves distinctly to pack ice as it can be more easily advected. This has implications for those wanting to cross the Arctic: a larger Arctic MIZ would be easier to send ships across.

The larger concentration of smaller floes and lower sea ice concentration in the MIZ has a number of consequences for the sea ice interactions with the ocean and atmosphere. Lateral melting will be enhanced due to the increased perimeter to surface area ratio, creating open water more efficiently than top or basal melt, and potentially fuelling the ice-albedo feedback. The lower the ice concentration, the more the surface ocean is warmed due to the lower albedo of open ocean, further enhancing ice melt and leading to the positive ice-albedo feedback. The increased open water fraction can also mean an increase in wind mixing in the mixed layer and will affect Arctic Ocean spin up (e.g. (Martin et al 2016)). There is a wave-floe size feedback that means the smaller the floes, the larger the impact of the waves, so a positive feedback loop exists that can act to increase the action of waves on the sea ice floes and further increase the concentration of smaller floes. The location and volume of sea ice melt has implications for stratification and so how deeply solar heat is mixed down. More sea ice melt means the mixed layer is shoaled and solar heat is concentrated in the upper water column.

Meanwhile there are other important sea ice processes, such as top melting where it is less clear that we would expect there to be a contrast between the MIZ and the pack ice, for example in the formation of melt ponds. In the Arctic, the snow thickness is generally modest compared to that on Antarctic sea ice, and the location of top melting and the formation of surface melt ponds is primarily driven by atmospheric conditions. Projections suggest that the MIZ will increasingly dominate the Arctic sea ice cover, especially in summer. It seems likely, therefore, that MIZ-focussed processes will play an increasing role in controlling the mass budget of Arctic sea ice."

We have then revisited these points in the Discussion section:

"The CICE model set up we used is relatively physics rich, which we believe is needed to represent the contrast between the pack ice and MIZ, as well as some of the changes over time. The differences in lateral melting was very likely caused by the inclusion of the FSTD model (Roach et al., 2018, 2019; Bateson et al., 2022), although we did not directly test this within this study. It is possible that lateral melt might be increased by the inclusion of a full wave model, though Bateson et al. (2022) show brittle fracture is likely just as important, and more so in the pack ice. The increase in top melting in the 2040s in the projection supports the importance of the topological melt pond scheme (Flocco et al., 2010, 2012) that we use, and the increasing role they play in the sea ice mass balance and evolution. As an increasing fraction of the summer sea ice cover becomes MIZ we expect that the representation of FSD-wave interactions and the melt processes themselves is likely crucial to realistically representing Arctic sea ice and the transition to sea ice free summers. The representation, or lack of representation, of such processes can contribute to discrepancies of Arctic sea ice (Diamond et al., 2021)."

Approach of splitting the ice cover into MIZ and non-MIZ regions and new insights from this: The comments about processes in the MIZ in the introduction included in the previous comment help to introduce why we would want to use this approach why we might expect differences between the regions through the description of sea ice processes in the MIZ. We have then added a paragraph in the discussion to add some of our thoughts on the use of the method:

"Our results show that sea ice processes do have a dependence on ice concentration, as would be expected, supporting the separation of the sea ice cover into MIZ and non-MIZ regions for analysis of the volume budget. Our results also indicate that if we separated the MIZ (and the pack ice) into more ice concentration based categories we would see distinct behaviour in the balances of processes, particularly in the type of melting. The MIZ and pack ice divide we have used differentiates between where internal stresses becomes important (SIC>80%), and where they become small in the MIZ, and the sea ice is in free drift. Our approach also has the advantage of simplicity, the more categories the MIZ is split into, the more complex the analysis becomes and the less clear the results. We believe we have struck a balance between the complexity required and keeping the analysis as simple as possible to understand. Although the ice is more dynamic in the MIZ, it was shown in this study to be a decreasing sea ice sink term due to the reduction in sea ice volume, meaning there is less sea ice to transport. Additionally. there was a relative increase in the melt terms, see Figure 8."

Insights on broader questions: We have added some comments on the importance of the inclusion and development of processes that are more dominant in the MIZ to the Discussion and Conclusion. Lines added to the discussion:

"The CICE model set up we used is relatively physics rich, which we believe is needed to represent the contrast between the pack ice and MIZ, as well as some of the changes over time. The differences in lateral melting was very likely caused by the inclusion of the FSTD model (Roach et al., 2018, 2019; Bateson et al., 2022), although we did not directly test this within this study... As an increasing fraction of the summer sea ice cover becomes MIZ we expect that the representation of FSD-wave interactions and the melt processes themselves is likely crucial to realistically representing Arctic sea ice and the transition to sea ice free summers. The representation, or lack of representation, of such processes can contribute to discrepancies of Arctic sea ice (Diamond et al., 2021)."

Lines added to the end of the conclusion: "Our analysis demonstrates a different balance of processes control the volume budget of the MIZ versus the pack ice. They are understandable in terms of the physical processes that dependent on the ice concentration, such as wave-ice interaction and lateral melt, which we are able to account for in our relatively physics rich sea ice model. We suggest that representation of such processes, in models such as climate models, requires more attention as a greater fraction of the sea ice cover becomes MIZ."

Specific comments.

1) *Please provide more information throughout the manuscript on the value of separating analysis into MIZ/non-MIZ regions. How do you expect processes to differ in these regions? Why do things differ across these regions both in their mean state and in their response over time? What value is added by looking at the system from this perspective?* A section has been added to introduction discussing processes that might be expected to differ between the MIZ and pack ice, see answer to General comments, **MIZ vs non-MIZ processes** section and **Approach to splitting the sea ice cover in MIZ and non-MIZ regions**. We have also added a Section on the atmospheric forcing used, to show the warming being applied, and a Discussion where we come back to the approach we have used.

2) *Please provide more information on what new information/insights were learned from this study relative to previous work.* Whilst the Keen et al study looked at mass budget between different climate models in the mean state and over time, we carried out the first study to our knowledge studying the MIZ/non-MIZ. We believe this is important to studies interested in the MIZ itself and those studying the ice cover over periods when the MIZ is a dominant part of the summer sea ice cover. In particular, when thinking about which model use to study this period. We believe it also highlights future work which could be done to investigate the feedbacks between waves-FSD-lateral melting further in the MIZ. We have discussed some of these points in a discussion, see the manuscript excerpt in the response to the General comments, **Insights on broader questions** section.

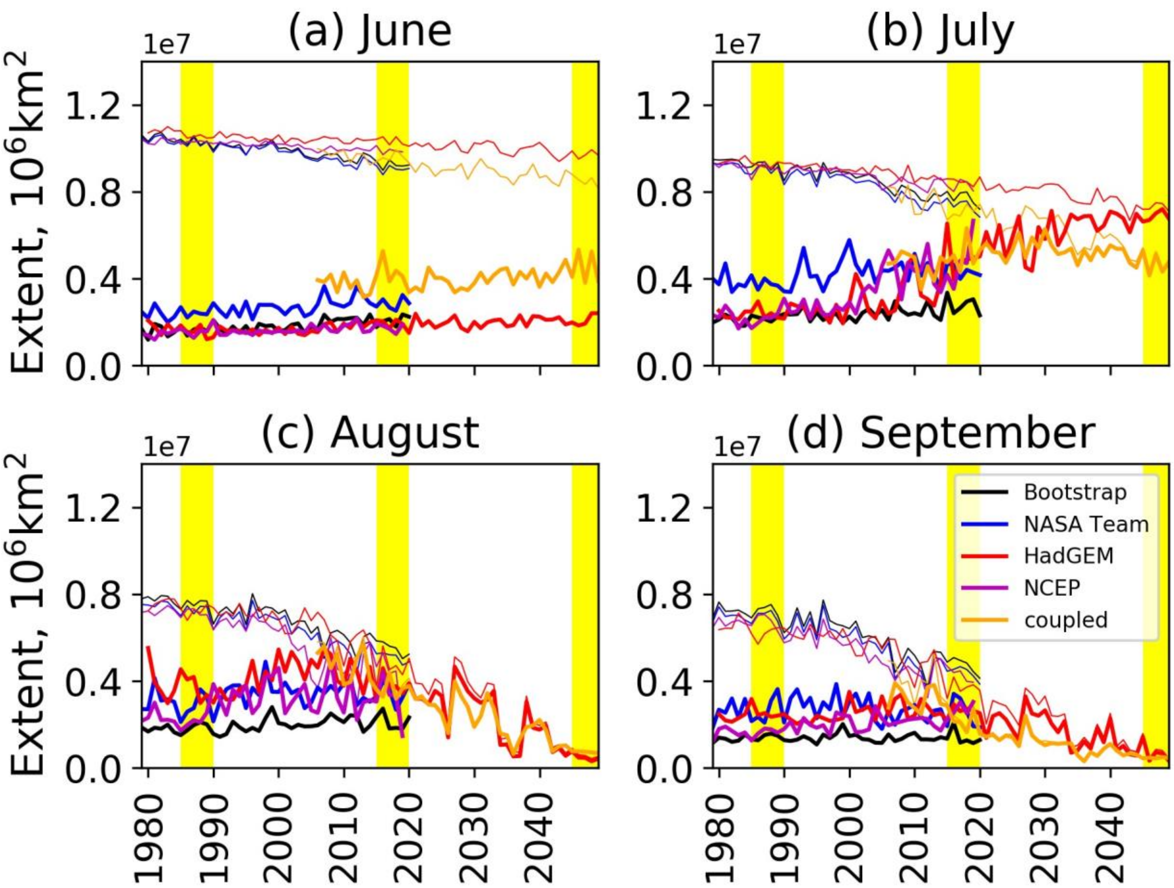
3) *Please provide more information on the implications of these new findings for bigger questions on the future evolution of the ice cover and/or discrepancies across models (for example).* I have added a Discussion section and some discussion of the importance of representing different processes, see answer to General comments on **insights on broader questions**.

4) *Line 25, "if" should be "is"* Corrected, see line 25.

5) *Lines 52-53: I appreciate that you acknowledge the limitations of the forced model framework. I think that it would be useful to return to this in the results or conclusions section to discuss how these limitations may specifically affect the results from this study.* We have added this to a discussion to revisit the implications. Lines 285-326:

“Our forced sea ice-mixed layer model receives no trend in subsurface ocean properties, such as the "Atlantification" of the Arctic as the subsurface Atlantic Water layer becomes warmer and thicker (Grabon et al., 2021), which has the potential to cause sea ice loss if the heat reaches the surface (Polyakov et al., 2013; Onarheim et al., 2014; Carmack et al., 2015). It is possible that some of the relative increase in top melting could be due to the constant ocean forcing, which may lack some ocean warming that we might expect to see in the 2040s. Although how much of this is mixed into the upper layer that interacts with the sea ice is an open question. Additionally, field observations indicate that the majority of the ocean heat needed to explain basal ice melt rates can be explained from solar radiation (Perovich et al., 2011), something our model does capture. Note that using coupled and climate models introduces its own set of problems, e.g. CMIP6 models fail to simulate a realistic seasonal cycle of sea ice area (Notz et al., 2020). Using a forced sea ice model allows us to simulate a more realistic sea ice state, which has been shown to affect the balance of sea ice processes (Holland et al., 2010; Keen et al., 2021).”

6) *Line 90-92: “The ocean temperature and salinity below the mixed layer …” Does this lead to a weaker sea ice response, especially near the ice edge? How do the results from this study compare to the sea ice simulated in the HadGEM2-ES runs?* The restoring is required to account for horizontal ocean heat transport. A 20-day time scale does not affect the ocean-ice interactions on shorter time scales. The sea ice response is not weaker, it is the mixed layer properties that impact the sea ice. This modelling approach for the mixed layer has been used in several other studies, see lines 87-89 in the revised manuscript. In the original coupled HadGEM2-ES simulation, the MIZ is typically smaller, also not in the right places, as it is simulated all the way around the ice cover at a relatively constant width, including at all coastlines. The difference in coastline and sea ice model, means that comparing our simulation to the original HadGEM2-ES simulation does not show us the impact of using a fixed T/S below the mixed layer on ice loss. We have included a Figure below showing how sea ice and MIZ extent compare in the summer months, using only the regions that CICE also simulates sea ice. The HadGEM2-ES simulation (orange line, thin lines are sie and thicker and MIZ extent) does have similar sea ice extent in August and September but is lower than our simulation (red line) in June and July. The MIZ extent is much higher than observations and the simulations in June (for the reasons listed above), but is more similar for the other months, although the MIZ being compared is not in the same places.



7) *Line 99-100: "and a prognostic floe size distribution model" What is the wave forcing used to drive this model? Are there any feedbacks from the model onto the wave forcing? I assume not since this coupling is not mentioned. The discussion of the limitations of this, particularly for studying things in an MIZ/non-MIZ perspective should be discussed.*

We have now added some description of the wave model and wave forcing used, see lines 96-105:

"We used the same wave forcing set up as Bateson et al., (2022), note that this is a different set up to Roach et al., (2019) where a separate wave model is coupled to the sea ice model to calculate the wave properties in the grid cells that contain sea ice. Instead, we use an extrapolation method as used and documented in (Roach et al., 2018), where ERA-I wave forcing (Dee et al., 2011) is used to calculate the necessary in-ice wave properties. The wave forcing consists of the significant wave height and peak wave period for the ocean surface waves. These fields are updated every 6 hours in the grid cells that contain less than 1% sea ice. Crucially for this study, despite not having a coupled wave model our set up still enables wave induced fracture causing enhanced lateral melting and wave-dependent new ice formation, as outlined in Roach et al., (2019). After 2017 we repeat the wave forcing, which does mean there is no trend in the wave forcing. Sensitivity studies varying the wave forcing using this model have demonstrated limited sensitivity to the wave forcing (Bateson. 2021) and comparisons to future 2056-2060 climatology from a global RCP8.5 wave simulation shows no significant change in significant wave height or interannual variability in significant wave height (Bateson. 2021). Although the wave forcing fields do not have any trends, the propagation of the waves into the ice field does respond to the changes in the ice cover over time. The simulations were initialised with a 6 year spin up period, this is a similar length to previous studies using the same model setup (Rolph et al., 2020; Bateson et al., 2022). As we are using a standalone sea ice model, the amount of spin-up required is much shorter than a climate simulation, or a coupled sea ice-ocean model."

8) *Line 102-103: "The HadGEM2-ES product …" Could you say a bit more about this product? Is it from a single ensemble member? How does this forcing (for example, Arctic mean surface air temperature, precipitation, etc.) compare to observations?*

There were three members of the ensemble, we used the first one. We have now added a section on the forcing sets where we compare it to the NCEP reanalysis, see Section 2.2 in the amended manuscript. Keen and Blockley (2010) includes a short summary of a comparison of the main atmospheric fields and biases present, but it is considered to be reasonably realistic.

From Keen and Blockley 2010:

"HadGEM2-ES is a coupled atmosphere–ocean model that was submitted to CMIP5. The model includes interactive atmosphere and ocean carbon cycles, dynamic vegetation, and tropospheric chemistry (Martin et al., 2011; Collins et al., 2011). It is considered to have a good depiction of present day global cloud characteristics (Jiang et al., 2012) and the best model depiction of Arctic cloud and surface radiative forcing (English et al., 2015). The mean Arctic ice extent lies within 20 % of observed values at all times of the year, although the September extent has low bias and the magnitude of the seasonal cycle is too large, consistent with biases in winter net surface long wave (LW) and summer net surface short wave (SW) radiation (West et al., 2018)."

9) *Line 116-118: "The maximum sea ice extent …": The winter ice edge (and maximum ice extent) are particularly sensitive to ocean conditions. How does the use of fixed T/S below the mixed layer influence ice loss to 2040? How do the results in this forced model framework differ from the HadGEM2-ES simulations and what does that tell you about the role of coupling? Similar questions are also relevant for the summer ice loss in the HadGEM2-ES forced runs and I'd suggest that you compare those also to the coupled HadGEM2-ES simulation that was used to obtain the atmospheric forcing data.*

See answer to point 6. It is difficult to make any conclusions from our study about the role of coupling on simulated future sea ice loss as there are differences in the sea ice model.

10) *Line 118-119. "The NCEP forced simulation shows a stronger declining summer sea ice extent trend …". Why is this the case? Is it consistent with internal variability? Is it due to biases in HadGEM2-ES forcing?*

The NCEP SATs are considerably warmer in Jan-April in all decades, and a larger warming between these periods in these months, which is likely the cause of the stronger trend in the NCEP forced simulation. See Figures 1 and 2, which have now been added to the manuscript.

11) *Table 1. For the 1980s and 2010s, it would be useful to also include observations on this table. Adding the range of values for the individual years in the 5-year periods that are analyzed would also be helpful for putting the differences between runs into context. It would also be helpful to show this information visually with a bar chart or something similar.*

We have removed this table and replaced it with a figure of the average annual cycle for each period of sea ice extent and MIZ extent. The new figure includes the observations and shows the differences between the study periods in each product/simulation. See Figure 4 in the revised manuscript.

12) *Figure 1. The quality of this figure should be improved. It is hard to distinguish the MIZ lines on the plots (particularly for Aug and Sept). What do the vertical bars on the observed extent signify?*

The vertical bars were error bars (+/-10% of sie), following the same method as Rolph et al.,(2020). However, to make the plot clearer we have removed them and changed the sea ice extent line to half line thickness and the MIZ line to solid, which is hopefully clearer to read. See the revised Figure 3.

13) *Line 121. "NASA Team and NASA Bootstrap" – please provide a reference for these datasets.* References added, see lines 131 and 132 in Section 2.3.
14) *Lines 133-134. "By the 2010s the MIZ becomes the dominant part of …" Please clarify for what months this is true.* This is now clarified in the text, it is true for July, August and September. See line 159.
15) *Line 134-135. "the summer sea ice cover is almost entirely MIZ …" Please be more specific on what months this is for.* This has been clarified in the text, see lines 160-161.
16) *Line 138. "PIOMAS" – please provide a reference and brief information on the PIOMAS product. In particular, it should be noted that PIOMAS is a model product and not*

direct observations. We have now added a section on model validation data, see Section 2.3, where we have noted that PIOMAS is a model product.

17) *Line 165. "By the 2010s the MIZ" – please provide a comparison to the observations here*. The comparison to observations has been added in a new figure (replacement to Table 2 suggested by reviewer 2), see Figure 7.

18) *Line 175. "basal growth makes up the largest proportion of melting" - Do you mean "basal melt" here?* Thank you, this was a typo, changed to 'basal melt', see line 246.

19) *Line 177: "significantly more" – what is the significance level used to determine this?* "Significantly" was not being used in a statistical way here, we have changed it to "substantially" to avoid confusion, see line 221.

20) *Line 186. "This means that it is likely that MIZ closer to the pack ice has a different balance of processes to the outer MIZ that has lower ice concentrations." Given this, what is the value in separating things into a MIZ/non-MIZ framework? Is it useful? What are the limitations? Would it be beneficial to look at the mass budgets as a function of ice concentration instead?* We have now added some comments on this to a new Discussion section, see Section 4 and the response to General Comments, **Approach of splitting the ice cover into MIZ and non-MIZ regions** section.

21) *Line 192-194. "A possible explanation for the increase in top melting in the future MIZ…" Is this inconsistent with the results from Keen et al. that did include active ocean models? Are there other factors that could explain increased top melt? Do factors like the change in seasonality, regionality play a role? Could you provide more analysis (for example of heat budgets) to gain a better understanding of the factors at play?*

   The increase in top melting is very likely driven by the increase in surface air temperature, see added Figures 1 and 2 in the new manuscript. The increase in top melting is consistent with the Keen et al paper, which also found an increase in top melting in the near future. We have now made this clearer in the text, see lines 227-231 and also in the Discussion section in Section 4.

22) *Line 194-197. "This means that the results here could be seen as a lower estimate on the ice loss…" I would encourage you to also assess the HadGEM2-ES runs to determine if those exhibit different ice loss characteristics. I appreciate that the ice models are different in these runs which will affect the results but it would nonetheless provide a useful point of comparison on the sea ice response with coupled ocean feedbacks.*

   See response to point 6 and attached figure.

23) *Line 198. "convection and ridging". Isn't it just advection of the ice that can cause a mass flux? Won't ridging affect the distribution of ice but conserve ice volume/mass?*
Yes, this is correct we have amended this in the manuscript to just advection.

24) *Line 229-230. "Peak total melting rates …" Why are there these differences between the NCEP and HadGEM2-ES runs?*

We have added some comments on differences between the simulations in the Discussion section, linking to the differences in the atmospheric forcing fields.

25) *Line 234-236. "In the NCEP simulation there is a 17% increase, whilst in the HadGEM2-ES simulation there is a 6% decrease, likely due to the larger melt rates in the NCEP simulation." Why "due to the larger melt rates"? How do the differences in melt rates drive the different ice growth response?*

We have rewritten this to remove the implied causality and added some NCEP-HadGEM2-ES comparison analysis in the discussion, see **NCEP vs HadGEM behaviour** in the answer to the General comments.

26) *Line 268-270. "Top melt was twice as important …" Please provide information on why these differences are present.*

It is likely these differences are driven by a combination of the differing FSTD in the MIZ vs pack ice, and the variation of warming over time spatially. The regions where pack ice is located have warmed more, driving more top melting. Meanwhile the presence of more smaller floes drives more lateral melting in the MIZ. We discuss some of these processes now in the Introduction, see lines 32-49 and also in the added Discussion section.

27) *Conclusions section. Please provide information on what new insights were gained in this study relative to previous work and the broader implications of this study.*

We have revised the Conclusion section and that has included making some comments on broader insights: "Our analysis demonstrates a different balance of processes control the volume budget of the MIZ versus the pack ice. They are understandable in terms of the physical processes that dependent on the ice concentration, such as wave-ice interaction and lateral melt, which we are able to account for in our relatively physics rich sea ice model. We suggest that representation of such processes, in models such as climate models, requires more attention as a greater fraction of the sea ice cover becomes MIZ."