

We thank Review 2 for the very helpful and constructive comments that we feel have greatly improved our manuscript. All comments have been addressed below (please see the blue coloured text). The line numbers below refer to the tracked changes document, with all revisions showing.

Overall manuscript revisions:

The correlation and trend analysis were repeated for the 925 mb temperatures (ERA5) and downwelling longwave radiation at the surface (APP-x), though only the longwave radiation is included in the final version of the manuscript.

Trend sentences are all group together now, followed by the correlations between the parameters at the pan arctic scale. Correlations between pan-arctic phenology and air temperature / downwelling longwave radiation are focussed on in the regional section. Additional information regarding the phenology timing has been added to each region to enhance the discussion and provide context for the correlations. All correlation numbers have been removed to improve readability, and only trend numbers remain.

Summary

This paper presents trends in the time of sea ice, snow, and lake ice presence based on the IMS product. Correlations are done with ERA5 2 m temperature data. The trends show earlier sea ice retreat and later freeze-up, in agreement with previous studies. The snow trends indicate earlier snow melt but also earlier first snow, so that trends in snow length are small. There is regional variability with the Bering and Chukchi Seas showing the greatest warming.

General Comment:

Overall, this is a competent manuscript. The data used are generally good (though see some comments below on this aspect) and the analysis is solid. However, it is lacking a coherent “story”. The goal of the paper is to simultaneously highlight changes in sea ice, lake ice, and snow, but each is largely treated in parallel – as if there are three different papers just merged together. I’d like to see more discussion of connections and interactions, covariance in parameters, etc. For example, Figure 6 shows snow/ice-off trends and I see a distinct difference in Lake Ladoga versus the surrounding snow on land with the lake showing a much stronger trend toward earlier onset. I don’t see a similar pattern in the Canadian lakes for snow/ice-off, but in Figure 7, some lakes – e.g., Great Bear Lake - do show a unique trend compared to the surrounding land. Digging into these types of things would be pretty interesting, but they aren’t discussed in the paper. In short, based on the title and the abstract, I was expecting more of a co-variance analysis – how are lakes and nearby snow on land and sea ice correlated and how might the forcing mechanisms affect them similarly or differently.

The LISA maps were meant to indicate where the trends are significantly clustered between snow/sea ice/lake ice and highlight the links - but they were not discussed adequately in the manuscript. More sentences have been added highlighting similarities and differences in the trends

with some more detail, with extra focus on the large lakes. Lake specific comments added listed below. Other sentences on links are throughout the manuscript.

313-321: Lake ice-off timing spans from April in the south to July in the northern islands (very few lake pixels experience ice-off in August, other than Lake Hazen). Both first and final lake ice-off regional mean dates are significantly correlated with May and June 2 m air temperatures, while the bulk of the ice-off dates in this region are through June and July; it is established that there is lag in air temperatures crossing the 0 °C isotherm and the timing of lake ice-off; up to about a month on average for lakes across Canada (Duguay et al., 2006). Regional mean downwelling longwave radiation through this region shows significant (negative) correlations between first and final open water in both June and July (as well as May WCIL, Table 5). A strong example of the lag between lake ice-off and snow-off can be seen using Great Bear and Great Slave Lakes compared to their surrounding areas (Figure 3): snow-free timing occurs here in May, while the ice remains on the large lakes until June/early-July due to the extra energy required to melt ice vs. snow.

324-326: The significant clustering between water and land pixels in the northern CAA (Figure 6c, d), with predominantly later trends for both open water and snow-off (no lakes are large enough in this area to be detected by IMS), indicates ice and snow are responding similarly in this region.

337-338: Significant clustering is shown in the ice and snow-off trends along the southern stretches of Hudson Bay and near-onshore regions (Figure 6c, d), indicating links in the response of ice and snow in those regions.

348-354: While the snow trends surrounding these large lakes are also earlier, they are not significantly clustered with the lake ice trends, with the exception of the eastern portion of Great Bear Lake, and a small portion of the western edge. The western portion of Great Slave Lake shows mostly earlier ice-off trends, while the surrounding snow trends show very slight tendency towards later snow-off, which aligns with both cooling temperature trends and less downwelling longwave radiation in that region (e.g. June, Figure 8f, 9f). Interestingly, the temperature and radiation trends here clearly show the effect of the lakes on their surroundings with contrasting trends for the lakes compared to land to the west.

387-388: Local clustering in this western region shows that the on-land and sea ice parameters have statistically significant clustered trends and are changing similarly.

398: Little to no significant clustering between the snow and ice trends are identified in this region, other than a limited swath of near-shore region, showing the sea ice trends here are stronger than the onshore snow/ice trends.

436-443: The eastern portion of Lake Ladoga shows very slight trends towards later WCIL, which is in sync with later snow off trends in that region, while the south/western portions shows stronger earlier WCIL off trends, likely related to delayed freeze (resulting in thinner, more easily melted ice) and predominant location of the ice over the season (as the ice forms in the southern shallower portions of the lake first) (Karetnikov et al., 2017). FOWL trends were not calculated for Lake Ladoga as the intermittent and moving ice cover through the season presents limitation to the

current search algorithm. There is only an 80% chance that Ladoga will experience a full ice cover during the ice season (Karetnikov et al., 2017).

553-556: Mean snow-on and lake ice-on do not show correlations to temperature or downwelling longwave radiation in this broad region, with the exception of final lake ice freeze in February – though this should be interpreted with caution as only Lake Ladoga experiences final freeze that late in the season and the intermittent and partial ice cover on this lake may affect the mean values.

576 – 585: The contrasting trends for delays in freeze onset compared to the earlier snow onset surrounding Lake Ladoga to the north and east also highlight the differences affecting snow-on and ice-on in the freeze season. While the air temperature may be cold enough to sustain a first snowfall (first snow-on), the lake water remains warmer through the fall, releasing the stored energy that was absorbed through the extended open water season as trends here are towards earlier WCI_L through much of the lake. In fact, for final freeze, although no trends were calculated for Lake Ladoga due to the intermittent ice, the region surrounding the lake shows a delay in final snow-on, and the 2 m air temperature for January (Figure 8a) shows a region of localized warming in the vicinity east of Lake Ladoga and Onega. Whether these localized trends are a direct result of temperature moderation from the lakes, or potentially from feedback processes in the ERA5 data is unknown, but as this is in a common downwind direction of the lakes (Kondratyev and Filatov, 1999), it would suggest lake induced temperature moderation as these lakes tend to freeze through January.

The correlation with temperature is also rather superficial, simply noting correlation coefficients. I don't think the paper needs to be an in-depth attribution study. But simply noting correlation doesn't really provide much insight. I note below a couple examples where the paper could dig in a bit more – e.g., what is the cause of the correlations? The way it is presented, it suggests that higher temperatures result in earlier ice/snow loss and later ice/snow formation, but it's not necessarily that simple. For one thing obviously, there is more than just SAT going on – there are winds, SLP, ocean temperatures, sea ice advection. Just looking at SAT may not tell one much, especially in specific regions.

Downwelling longwave radiation has now also been added to the analysis to represent the combined effects of the temperature and moisture fluxes that influence sea ice citing literature highlighting the importance of downwelling longwave and comments from Reviewer #1. Better links between the correlations and the ice information have been added throughout the manuscript to improve the discussion, though no further detailed analysis was done as the idea was present an overview more so than a detailed examination that would really need to be done an even more regional scale.

I also think that Section 3 is very wordy and somewhat hard to follow – it is one number after another. The data are all in the figures and tables, right? So, I don't think you need to worry so much about putting the numbers in the text. Focus on describing the main characteristics – significantly positive and negative trends, regional variations, etc. I also found the mixing of trends and correlation a little odd. I can see the rationale to relate those in the discussion. However, you

are showing trends, but then correlations with de-trended data. So, it feels a little inconsistent, though I understand you're trying to relate temperatures to the snow/ice changes. You might consider splitting the two into separate sections: first the trends, then the correlations and use the correlation to discuss the relationship between temperature and snow/ice on-off dates. Maybe that won't work well, but something to consider trying.

Thank you for this suggestion, we tried a few options for rearranging and have settled on separating the correlations into an introduction paragraph for each region, along with the mean phenology timing. Trends then follow with links back to the temperature/radiation data where suitable. Removing the numbers from the correlations and only mentioning the months improved the readability considerably.

Finally, more discussion of the data is needed (including proper citation as noted in the comments below), particularly data quality. IMS is described reasonably, but because IMS is an analyzed product and has used differing amounts and quality of data over the years, one should use caution when applying it towards trend analysis. I think it generally is okay in this context, looking at overall trends, but I think a comparison (at least for snow and sea ice) with passive microwave products would be beneficial to demonstrate consistency; at the very least, a discussion of the limitations of the dataset is needed.

Thank you for pointing out this omission on our part. A detailed comparison of IMS data for sea ice with microwave data was done by Brown et al 2014 – we showed that not only was IMS consistent with the overall patterns it was better than the traditional passive microwave as it can resolve the near coastal regions and leads when using the 4km product. Brown and Duguay, 2012, also highlight some limitations of IMS data when used for lake ice, and several other studies have commented on the utility of IMS for snow cover. These points have now been added to the data description.

Lines 128 – 141: IMS has been shown to outperform data from traditional passive microwave products (AMRS-E, SSM/I, SSMI-SSMIS) for both the timing and extent of first open water in the Arctic (Brown et al., 2014). For example, through the Barrow Strait in the CAA, the ability of the 4 km IMS data to resolve narrow channels lead to 17% more open water detected than with SSM/I, and 35% more open water than detected with AMSR-E, validated with RADARSAT-1 (Brown et al., 2014). Overall, most pixels compared between IMS and the two passive microwave datasets for first open water were within ± 5 days, with a greater percentage of the pixels in the categories beyond the ± 5 days identifying open water earlier with IMS than the other products (Brown et al., 2014). IMS has been shown to map higher snow cover fractions during the spring melt period than other snow products (Brown et al. 2010; Frei and Lee, 2010), but is reported to have mostly between 80-90% agreement with other snow products during the winter season of non-arctic North America, with better agreement in the later part of the winter season when deeper and more extensive snow cover is present (Chen et al., 2012). For lake ice, the 4 km IMS product occasionally identifies earlier lake ice-on dates in regions of prolonged cloud cover (e.g., northern Quebec, Canada), though both ice-on and -off timing detected using IMS are significantly correlated with, and comparable to, phenology dates extracted from the MODIS Snow Cover product (Brown and Duguay, 2012).

This goes doubly so for the temperature data. There needs to be more than a simple two-sentence paragraph. For one thing, reanalysis products are notorious for issues in polar regions because the data sparsity, particularly over sea ice. I note in comments below a couple artifacts that jumped out at me in the figures. And 2 m temperatures can be particularly problematic because the reanalysis models don't necessarily capture the surface boundary layer accurately. A now somewhat outdated paper (Lindsay et al., J. Climate, 2014) discusses some of the limitations and biases of different reanalysis products in the Arctic. There is nothing wrong with using ERA5 – it is one of the more updated and better products – but I think it is important to discuss the uncertainties and limitations. Along these lines, it might be interesting to correlation with 925 mb temperatures in addition to SAT (2 m) fields as that gets above the boundary layer and the reanalyses may be more reliable/consistent there.

ERA-Interim uses the observed 2 m temperatures in post processing procedures leading to smallest biases amongst other models (Dee et al., 2011, from Lindsay et al., 2014), presumably ERA5 does as well – though not confirmed in our literature search. ERA-Interim indicates a cold bias in Lindsay et al 2014 over land, however Wang et al 2019 show a warm bias over Arctic sea ice. Unfortunately, no recent 2 m temperature tests over arctic land areas could be located in our literature search - other useful and relevant information on ERA5 biases has been added instead:

Added lines 147-159: Temperature data near the surface (1000 mb) were reported to have a 0.89 K difference from radiosonde observations, and the ensemble spread is quite low at ~ 0.4 K or less, from 1979 – 2018, which can be used as an indicator of uncertainty (Hersbach et al., 2020). Compared to radiosonde temperature profiles in the Fram Strait, ERA5 had the smallest bias ($\leq 0.3^\circ$) and RMSE ($\leq 1.0^\circ$), and highest correlation coefficients (≥ 0.96) over four other reanalysis datasets tested (ERA-Interim, JRA-55, MERRA-2, CFSv2) (Graham et al., 2019). The 2 m air temperature in ERA5 has improved fit to observations in the Arctic compared to its predecessor ERA-Interim (Hersbach et al., 2020), though Wang et al. (2019) show ERA5 has a warm bias over sea ice compared to observation data from buoys. The identified warm bias is stronger in the cold season, particularly when the 2 m air temperature is below -25°C (daily mean value of 5.4°C), however monthly mean differences between ERA5 and buoys are $\sim 2^\circ\text{C}$ or less through all months other than March, April, and May (Wang et al., 2019). Regionally, ERA5 performs best in the Central Arctic, followed by the Pacific Sector; the Atlantic sector shows good agreement only while the 2 m temperatures are above -25°C (Wang et al., 2019). We acknowledge that some small potentially spurious regions of opposite trend directions appear in some months of the temperature trend maps (e.g., February: Eastern Siberia, March and October: Arctic Ocean) however these data are not used in a quantitative comparison and therefore do not affect the overall discussion.

We redid the trend analysis and correlations using the 925 mb temperatures as suggested – although the overall trends maps are not overly different, the locations of some trend patterns did shift. Some potential artifacts are still evident at even that high above the surface, though not as strong and in different months than t2m. The correlations did not provide any extra insight over the 2m air temperatures (there were less significant correlations for sea ice and snow, but interestingly more for lake ice, which is something we may explore further in a separate ongoing lake ice project). Based on the readings regarding the handling of 2 m temperature in ERA-Interim, and the bias information discussed above, we feel 2 m temperature is more appropriate for representing the melt and freeze discussed here. A positive aspect of surface climate representation

ERA5 is the incorporation of improved lake parameterization – using the Flake model to couple with the atmosphere should be producing even better surface fluxes than previously – which have strong importance for melt/freeze. The 925 mb data was incorporated, however during the editing phase we opted to remove it as it did not add anything extra to the analysis other than length. Figure included below out of interest.

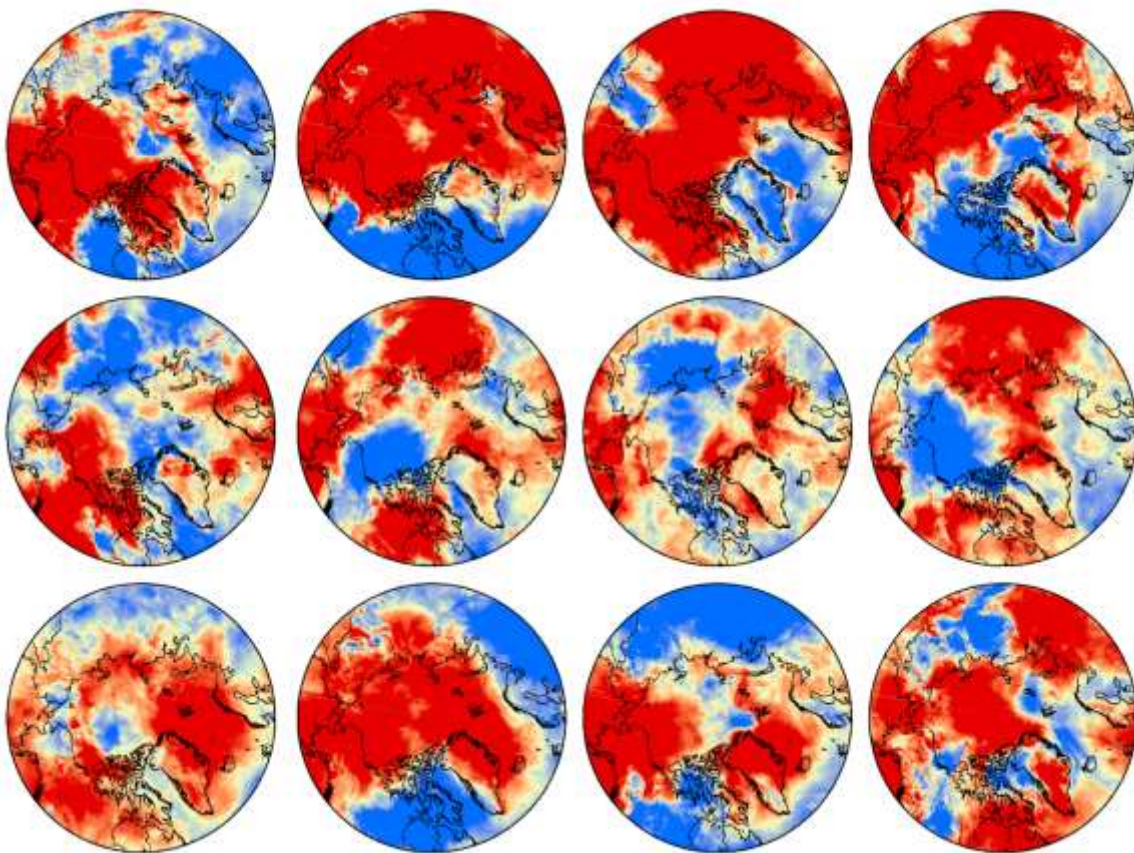


Figure 8b –Trends in monthly 925 mb temperature from 2004 – 2019 in (a) January , (b) February, (c) March, (d) April € May, (f) June, (g) July, (h) August, (i) September, (j) October, (k) November, and (l) December.

There are several fairly major comments here, along with the specific comments below. Thus, I recommend major revisions. However, I think the comments can be addressed with some thought and some rewriting and additional discussion, which I don't think should be too onerous. So, I think these revisions can be done in a reasonable amount of time.

Specific Comments (by line number):

56: I understand there are a lot of ice melt/freeze papers, and there is no need to be comprehensive, but it might be best to focus on more recent papers. E.g., Markus et al. (2009) has been updated by Stroeve et al., GRL, 2014, and even more recently by Bliss et al., Remote Sensing, 2017, and Bliss et al. Env. Res. Letters, 2019.

Noted, thank you, the missing recent papers have been added as suggested. Line: 56

102-109: I find the region definitions somewhat arbitrary and each encompasses quite different conditions. The Canadian Arctic spans the CAA, which has landfast and MYI, along with Baffin and Hudson Bays, which are nearly all FYI; and then it extends to a seemingly arbitrary longitude, east of the Canadian/Alaska border – so it not all of the Canadian Arctic, but most of it. The Alaska/Far East similar includes the Chukchi and some of the Beaufort, which have MYI and FYI, plus the Bering, which is exclusively FYI. Eurasia, seems like a catch-all region for everything else. And looking at ice/snow on-off dates is going to be at least partly latitude dependent, so splitting into regions with a wide range of latitude within them would seem to muddle any relationships. I agree this simplifies the analysis having only three regions, but I think explaining better the rationale would be helpful.

The snow and lake ice were grouped into the two continental regions to align with previous studies comparing Eurasia to North America, as they are affected by different weather patterns and storm tracks. The sea ice regions were reduced to three regions based on trend similarities and driving weather. We know the Canadian region is not changing the same way as the other Arctic regions in terms of trends. Eurasia is seeing earlier ice-off but the largest changes are in the US regions. Since there are 3 phenology parameters being compared the goal was to reduce the regions to encompass the broad patterns. As an aside, investigating the trends based on FYI or MYI would certainly be interesting, but not possible with IMS data alone. Added the following for further explanation:

Lines 112-113: These regions were grouped based on similar trends in phenology parameters and differences in climate and weather characteristics at the hemispheric scale.

And similarly, is there a rationale for using 56° N latitude as the southern boundary? It also seems somewhat arbitrary – it cuts off the southern Hudson Bay, for example. I guess it maybe encompasses the southern boundary of ice within the Bering, and it includes the large Arctic lakes. Thus, I can see a potential rationale (if I'm correct), but it should be explained within the text.

Correct, we were trying to maximize the ice included while minimizing the non-arctic land areas included. Added the following sentence for clarity:

Lines 104-106: In this study, regions north of 56° were considered when examining pan-Arctic snow and ice phenology (Figure 1) to include much of the southern limits of the sea ice in the Bering Strait and large Arctic lakes that can be resolved using IMS.

111-114: Make sure you are clear on the source for the data. You mention NIC, but it has only recent data I believe. The historical archive is at NSIDC: <https://nsidc.org/data/G02156/versions/1>. This is referenced a few lines later for more information, but I think it is important to be clear upfront that the data were obtained from NSIDC (assuming they were). Also, the appropriate citation should be provided for the data set, as noted on the NSIDC page:

Most of the data was from NSIDC, but a few years were retrieved from NIC as we had trouble accessing the FTP at NSIDC. Text has been edited to put the emphasis on the NSIDC archive since the bulk of the data was obtained there. Dates included in each archive also added for clarity.

Lines 115-117: Snow and ice data were obtained from the Interactive Multisensor Snow and Ice Mapping System archived at the National Snow and Ice Data Centre (2004 – present, <https://nsidc.org/data/G02156/versions/1>) as well as from the National Ice Centre (2014 – present).

The NIC web address has recently changed, so it needs to be updated to: <https://usicecenter.gov>, or more specifically: <https://usicecenter.gov/Products/ImsHome>

Updated, thank you.

123: Similarly, for temperature data, more specificity is needed. Is there a website from where the data were obtained? And, if available, a full citation should be provided or at least the data set DOI (if one exists).

Thank you for pointing this omission out. The citations are in the reference list, but we inadvertently removed them from the data section of the manuscript. Added back in correctly now (CS3, 2017; Hersbach et al., 2020)

Lines 144: Temperature data are from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 global reanalysis (CS3, 2017 DOI: [10.24381/cds.f17050d7](https://doi.org/10.24381/cds.f17050d7); Hersbach et al., 2020) and ...

128: Maybe not directly relevant here, but could be worth mentioning that there is now a 1 km product starting in 2014. I wouldn't expect these to be used in the study as the time frame is too short to be of use, but readers may be interested to know that.

Clarified in the data section that the 1km product is not used here due to the limited time series available:

Line 119: ... 1 km (2014 – present; not used in this study due to the limited time series),...

145: Were the data re-gridded/re-projected onto a consistent grid for comparison?

Data was analysed using centre-point locations (all points within the regions analysed), with ERA5 and APP-X reprojected into polar stereographic to match the IMS data. Added the following to be explicit:

Lines 195-196: All data was projected to match the IMS data, and the centre point of all grid cells was used for the analysis.

198: Here is an example of where the paper could go farther. The authors note that freeze onset shows significant correlations with October air temperature. That's not surprising and it is not necessarily due to warmer air temperatures delaying freeze-up. As other studies have noted, it's actually the reverse: warm SSTs in the ice-free region give heat to the atmosphere and warm the atmosphere. So, it is the lack of sea ice that causes the October air temperatures.

This section of text has been removed in editing, but this is a good point that was not explicitly stated before – it has been worked into the freeze section with respect to the feedbacks also causing the downwelling longwave to increase during the fall

217-220: Not discussed here is the possible role of ice dynamics. How ice moves into and through the CAA can have a significant impact on the presence or lack of ice in the region during break-up and particularly freeze-up.

Line 278: This is one of the main findings from Dauginis and Brown, 2020 that is cited here, but we failed to actually highlight the dynamics aspect. Added in parentheses: (due, at least in part, to increased ice dynamics through the CAA)

343-347: Another example where more analysis/discussion is needed. First snow-on dates are positively correlated with August, September, and October temperatures. But as noted above, the Aug, Sep, Oct temperatures may be the result of lack of sea ice causing the warming temperatures. And the lack of sea ice also provides a moisture source for earlier potential first snow. There is nothing groundbreaking in my analysis here, but it is not acknowledged at all in the manuscript.

Thank you for pointing this out, while we were inferring this in places where we mention increased atmospheric moisture you are quite correct that we never actually explicitly stated the open water / more moisture link. We have added several discussion sentences about the increased moisture in the atmosphere and resulting increase of snow, in particular for earlier snow-on in Euraisa.

Figure 1: Not to be too pedantic, but Greenland is shaded as Eurasia, but it is technically part of the North American plate. Just changing the color key to "Eurasia and Greenland" is reasonable, though it might look better shade Greenland with the North America color, or consider a separate shade altogether.

Greenland is now shaded the same as North America, thank you.

Figure 3: This is pretty obvious, but for total clarity, it would be worth adding a note in the caption that white indicates locations where there is no snow/ice on date (either because snow/ice never occurs there or snow/ice is present all year – e.g., GrIS and multi-year sea ice). Also, in the color legend, it would be helpful to label the center date (e.g., the pale yellow color). For snow/ice on date, it is okay because the middle is near day 365; but for snow/ice off labeling the yellow would be helpful for discerning colors between the extremes of 40 (red) and 275 (blue).

Centre data added to the colour legend, and addition line in the caption added: White regions indicate where either no snow/ice forms, or snow/ice remains all year, in 14 or more years of the dataset.

Figure 8: While listed in the captions, it would be more helpful to label each map with the month (or appropriate abbreviation, e.g., “Jan”, “Feb”, etc.), either in addition to or instead of the figure letter. This makes it easier to read the figure – I can just look at “Sep” instead of looking at “i” and having to connect it to the month in the caption text.

Agreed, revised figure now included.

Figure 8j: This demonstrates some of my skepticism of SAT trends over sea ice. The “bullseye” of blue in the middle of the Arctic Ocean doesn’t look realistic to me – probably an artifact of some interpolation (November shows a similar, albeit larger and more dispersed).

The October bull eye is an oddity, you can still see it in the 925mb layer though much less strong and the same direction as the surrounding ocean. However, this region does not have sea ice trends mapped as no consistent break up occurs here. We considered masking out the regions with no ice phenology in the temperature and longwave radiation maps, however decided against it as some interesting information would be lost, particularly through the Atlantic sector.

A sentence was added in the manuscript to acknowledge the potential artifacts in the data (both ERA5 and App-x) and the decision was made to continue with the 2m air temperature.

(ERA5) Lines 157-159: We acknowledge that some small spurious regions of opposite trend directions trends appear in some months the temperature trend data (e.g. February: Eastern Siberia; March: Arctic Ocean and October: Arctic Ocean) however these data are not used in a quantitative comparison so these small regions do not affect the overall discussion.

(APP-x) Lines 167-169: Some artificial patterns are evident in the data (e.g. March, Figure 9c, near the pole), however for the purpose of regional comparisons this is not limiting as this regions is not used in any quantitative comparisons.

Figure captions: Just a general comment on style – putting the figure letter - (a), (b), (c), etc. – at the end of the corresponding description is somewhat confusing and makes it more difficult to read. I think it would be easier to read if the letter is placed before the description for that letter.

Agreed. This has been changed in the captions now, thank you.