Dear Prof. Pfahl,
Dear Reviewers,

We are pleased to submit a revised version of the manuscript "European summer weather linked to North Atlantic freshwater anomalies in preceding years". Based on the detailed feedback provided by both reviewers, we have clarified and improved the manuscript.

Both reviews were helpful, thorough, and constructive throughout. The main comments were related to the placement of important information in the appendices, an uncertainty about the physical causes of freshwater anomalies, and a lack of clarity regarding the actual role of freshwater anomalies for European summer weather. To address these comments, we have implemented the following main changes:

- We have shifted the previous Appendices A1 and A2 into the approach section, and Appendices A3 and A4 into the results section. The new appendix now only contains the calculations and the comparison with hydrographic observations. All critical information is provided in the main text.

- We have added an analysis of the causes of freshwater anomalies, which supports earlier studies. While a detailed examination of the freshwater budget and origin of the freshwater anomalies are beyond the scope of this study, the analysis provides a physical explanation for having two different types of freshwater anomalies linked to opposite atmospheric circulation patterns in the preceding summer.

- We are explicit about the involved timescales. The focus of this study is on interannual timescales, reflected in low autocorrelations of the freshwater indices and the summer NAO. Nevertheless, freshwater variations are also linked to lower-frequency variations of the North Atlantic SST. Thus, we mention that periodic freshwater releases from the Arctic may have contributed to Atlantic Multidecadal Variability. Also, we show that the freshening has a trend over the last 70 years, and we discuss the implications.

- We have added a clearer analysis of the involved air-sea coupling processes in winter and summer. Given the importance of atmospheric feedbacks in reinforcing the SST anomaly, we do not attribute the anomalies in European summer weather to freshwater variations alone. Considering the spatial and temporal characteristics of the characteristic SST pattern, and the lag of one year, we propose in the discussion that enhanced freshening can act as a trigger of subsequent, large-scale air-sea feedbacks, which in turn affect European summer weather. However, we are cautious about the wording and clearly state the involved uncertainties.

We further provided more details and explanations, removed any ambiguities, and unclear or inconsistent terminologies. Detailed changes are explained in the response letters, where the responses to the reviewers' comments are shown in blue, and the resulting changes in the manuscript are shown in red. Figure and line numbers refer to the revised manuscript without tracked changes:

- Reviewer 1: page 2
- Reviewer 2: page 15

We think the manuscript has greatly benefitted from the review, and we are confident that is now clarified, easier to read, and a worthwhile contribution to Weather and Climate Dynamics. We thank both reviewers and the editor for their efforts in reviewing and handling this manuscript.

Sincerely,

Marilena Oltmanns, on behalf of all authors

Responses to Reviewer 1

The authors use statistical analysis of observations and reanalysis data to support their hypothesis that warmer and drier summer weather in Europe can be linked to freshwater anomalies in the North Atlantic subpolar gyre region during the preceding year. The proposed mechanism for this link is a northward shift of the North Atlantic current leading to a similar deflection of the jet stream and therefore altering the advection pathway of maritime air masses. The foundation of the analysis are freshwater indices derived from a mass balance equation that are used to identify freshwater anomalies in relation to simultaneous sea surface temperature (SST) anomalies linked to the North Atlantic Oscillation (NAO).

I understand that this is a re-submission of an earlier version of the manuscript, but I was not involved in the previous review process. Therefore, I cannot assess how the manuscript has been improved, but rather provide a fresh pair of eyes.

We sincerely thank the reviewer for providing a fresh pair of eyes and reviewing our manuscript. The review was extremely thorough and detailed. Moreover, the comments and suggestions were constructive and have helped us to improve the manuscript.

I. General Comments and Suggestions:

1. One of central results of this study is the description of "a coherent, deterministic mechanism that links North Atlantic freshwater events to European summer weather" (l. 315-316). However, the actual role of the identified freshwater anomalies in the subpolar gyre remains unclear. Given the lack of salinity observations, SST anomalies in relationship to the NAO are used as proxy for freshwater anomalies. In turn, a substantial part of the described link between the freshwater anomalies and European summer weather is based on the enhanced meridional SST gradient between the subpolar and subtropical gyre, and its influence on the storm track over the North Atlantic. This raises the question to what extent the freshwater anomalies actually influence the proposed mechanism and the downstream response?

Thank you for asking about the role of the freshwater anomalies. In the previous version, we only showed that freshwater is statistically linked with subsequent European summer weather and proposed a dynamical explanation. In the revised manuscript, we are more specific about the role of freshwater and use a more cautious narrative. Therefore, we have implemented four main changes:

(1) We have included a description of the drivers of freshwater anomalies (Section 4.2). Thus, we provide a physical explanation for the start of the chain of feedbacks. Since the identified drivers (runoff for F_E events and circulation changes for F_W events) are not otherwise linked to European summer weather on the investigated timescales, and occur exactly one year in advance, they challenge the idea that an unknown third mechanism drives both freshwater anomalies and warmer European summers without the two being physically connected.

(2) We are explicit about the involved timescales of variability. While the focus is on interannual timescales (line 296), we now show that surface freshening has a trend (Fig. 9a of the revised manuscript) that is superimposed on substantial interannual variability. The interannual variability is correlated with the variability of the summer NAO and runoff from the preceding year, obtained from the Greenland climate model MAR. (Section 4.6). After considering alternative drivers acting on these timescales, we find (line 405):

"Apart from seasonal runoff- and melt-driven freshening, there are currently no conceivable, physical mechanisms in the tropics, stratosphere or outside the North Atlantic which, at the same time, have a significant trend over the last 70 years, exhibit a similarly high interannual variability, can explain the occurrence of freshwater anomalies in the subpolar region in the subsequent winter, and are significantly correlated with the characteristic summer SST pattern a year before it occurs. This suggests that seasonal freshening may not only be a predictor of the SST pattern but also a potential trigger."

(3) To assess whether runoff is a trigger of the SST pattern, rather than only a predictor, we have included an analysis of the ocean-atmosphere feedbacks that contribute to the evolution of the SST pattern in summer (Section 4.7). After investigating the surface heat and momentum fluxes and their influences on the ocean and atmosphere with ERA5, remote sensing data, in-situ hydrographic observations from the cold anomaly region, and with the models, we find that wind-driven transports may contribute to the intensification of the SST signal. Thus, we conclude (line 453):

"The atmospheric forcing contributes to the development of the SST field but is in turn, forced by it. On the one hand, the consistency of the spatial SST pattern with the wind-driven transports, the intensification of the SST signal in mid-summer, and the vertical extent of the hydrographic anomaly point to the relevance of the atmospheric forcing in driving the SST signal over the North Atlantic. On the other hand, model simulations, forced with the prescribed, observed SST, reveal the importance of the SST field for the large-scale atmospheric circulation, including for European summer weather. The underlying SST pattern covers the entire North Atlantic with the strongest signal occurring in the subpolar region (Fig. 11b). While further studies are necessary to confirm the dynamical contribution of freshwater anomalies to the large-scale SST pattern, its spatial and temporal characteristics, and the time lag of one year, indicate that enhanced surface freshening from the preceding year may have initiated the chain of air-sea feedbacks."

(4) We are more cautious about the wording. Throughout the analyses, we refer to freshwater as a predictor (e.g. line 350). In the conclusion, we discuss evidence that points to freshwater as a trigger rather than only a predictor, but we are explicit about the involved uncertainties, and we have removed phrases that may previously have caused confusion.

2. At times, it is difficult to follow the analysis which might be in part due to the overall structure of the manuscript and lack of some details in the text (often they are only mentioned in figure captions). Additionally, some of the terminology is unclear or inconsistent throughout the manuscript which is possibly an artifact from the refactoring of the previous version. Hopefully, the comments below will help to streamline the text and make it more accessible for the reader.

Thank you for providing very specific comments below on the exact places in the manuscript that have been unclear. We have now clarified each of these instances and took care to include all details in the text. In addition, we have removed all inconsistencies in terminology that you have spotted.

3. Some of the figures are hard to read as individual panels are small or details are obscured. For most maps, the colorbars and axis labels take up valuable “real estate”. I suggest to decrease their size and use the free white space to increase the maps wherever possible. Furthermore, I recommend to decrease the thickness of the coastlines since they can be quite distracting, especially on maps with vectors. It might also be worthwhile to mention differences in color scales in the caption wherever it can help guide the reader (e.g., Figure 2).

Thank you for these suggestions. We have now slightly reduced the size of the axis labels and font sizes around the colour bars. In addition, we have increased the map sizes for specific figures further to the south to ensure that all relevant signals are included in the maps. We further decreased the thickness of the coastline in all figures, so the vectors are clearly visible. We also mention differences in colour scales wherever they occur.

II. Main Comments:

1. Section 3/Appendix A: It took me a long time, including going back and forth between the main text and appendix to follow the approach. Given the importance of the freshwater indices as foundation for the subsequent analysis, I suggest to combine Sections 3.1, 3.2, with Appendices A1, and A2 to describe the derivation in the main text including the clarification of the following points:

Thank you for indicating that the derivation was difficult to follow. We have now combined Appendices A1 and A2 with Sections 3.1 and 3.2.

1. l. 126: Please state this equation.

   The equation is now fully stated.

2. l. 129: I think that M is not the same "downward mixing", but rather entrainment of water masses below the pycnocline into the surface ocean mixed layer as a result of a deepening mixed layer. In the context of this sentence, you refer to increased stratification due to large freshwater anomalies that inhibits a deepening of the mixed layer due to convection.

   We are now more precise and specify "entrainment".

3. l. 135-137: This sentence becomes only understandable after reading the appendix.

   We have removed this sentence.

4. Section 3.2/Appendix A2: The whole derivation of the freshwater indices is based on the NAO. I can't help but notice the striking similarity between the spatial pattern of the cold anomaly in the subpolar gyre for F_E and regression pattern of SST anomalies on the North Atlantic SST Index of the Atlantic Multidecadal Oscillation (AMO) in that region (see Fig. 11 in Deser et al., 2010). The AMO has been in its warm phase since the mid-1990s and thus during the time of most of the F_E years. This raises the question to what extent longer-term climate variability influences the relationship between the NAO and freshwater anomalies and if this can be utilized in the design of the freshwater indices?

   Thank you for this suggestion. We did indeed attempt to divide the full period in high and low AMO phases. However, this did not turn out to be useful. We found that the summer NAO already filters out the interannual variability of freshwater and that the targeted approach of subsampling was more effective in optimising the indices.

   In a related study, we find that the trend of the cold anomaly over the last 70 years has recently overtaken the AMO signal and now has a larger amplitude than the AMO. This has far-reaching implications for North Atlantic climate variability, including a shift of power towards interannual timescales.

   While a detailed analysis of the AMO signal is beyond the scope of this manuscript, we now mention its potential relationship to freshwater anomalies in the revised manuscript through periodic Arctic freshwater releases (line 293). Overall, we find that freshwater can explain the variability of the cold anomaly pattern on a range of timescales, including that of the AMO, and including the cold anomaly trend (Section 4.6), due to the different causes of freshwater anomalies. Thus, it may not be desirable to completely filter out these signals.

5. l. 152-153: How do you estimate the correlation between the index and freshwater anomalies and how well the index represents the initial freshwater anomalies if they "are not known a

priori" (cf. l. 148-149)? I think you are referring to SST anomalies which serve as proxies for the freshwater anomalies.

Yes, you are right. We present the correlation together with the uncertainty since both estimates belong together. We first estimate the freshwater anomalies from the SST anomalies and obtain an uncertainty of 4% and 7% respectively. We then calculate the correlation between the estimated freshwater anomalies and the NAO index.

This is now clarified in Section 4.1 (line 241) and the caption of Figure 3. We have also removed the sentence you are referring to, since it was confusing.

6. l. 156: Please define F_E and F_W explicitly. Without going through the appendix, the reader might ask themselves why there are two indices? What do the subscripts refer to? How are the two related?

Since we have removed the appendices, all information about F_E and F_W is now included in the main text (Sections 3.2, 4.1 and 4.2). Thank you for suggesting this.

7. l. 160: Please refer to Appendix A3 to show how these uncertainty estimates were obtained.

We have now shifted Appendix A3 and A4 into the results section (Sections 4.1 and 4.2). Thus, Section 4.1 now includes a detailed explanation of the freshwater anomalies, their uncertainty estimates, and how they were obtained.

8. l. 161-162 & Figure 3: I think Figure 3 deserves more prominence in the text as these are the actual freshwater anomalies your hypothesis is based on. I suggest to move this sentence into its own paragraph and add more details, e.g., by being explicit that the shown salinity anomalies are estimate based on the surface mass balance (I think?), how you obtained the relationships, and what the white areas represent. Please define SSS.

Thank you for pointing out the importance of Figure 3. We have now shifted it into its own section in the results and clearly explain it (Section 4.1). We also define SSS as the sea surface salinity (line 234).

9. Figure 2: Panels (a) and (b) should be the same as Panels (a) and (e) in Figure A1, but the structure of the largest values looks different. Is this just because of the differences in the color scale?

Yes, we had adjusted the colour scale in Figure A1e to make it the same as that of Figure A1b. However, we agree that the choice of the colour scale concealed the structure of the largest values. After combining the appendix with the approach section, we have removed this figure.

10. Equation (A2): Please define ρ_0.

We now define it as the density before the winter, for instance September or October, with dt referring to the resulting time interval of integration. We also explain in Section 4.1 that the results are not sensitive to the exact starting point if the mixed layer is still relatively shallow (as in September or October).

11. l. 350: Strictly speaking, given Equation (A2) is the result of an integral over time, h_n is the mixed layer at the end of winter.

Importantly, the time of the mixed layer depth (h_n) must be the same as that of the temperature term (T_n). While it is possible to use March only – and the results do not change appreciably – we consider the full winter period for increased robustness. In some months, the maximum mixed layer depth may already be reached in February.

You are correct that Eq. (2) [which is now Eq. (4)] is an integral over the time. Here, we integrate from an arbitrary time before the winter (for instance September or October, when the mixed layers are still shallow) to the winter period (January to March), with dt referring to the corresponding time of integration. We now clearly state these periods (line 133), motivate the choices and explain that the results are not sensitive to them (line 226, line 580, line 585).

12. l. 359-361: How realistic are these conditions and at what timescales do you expect this assumption to hold?

At this location in the text, it is only a motivation. It is an objective of the approach to achieve these conditions. We do not make any assumptions. We have now rephrased the sentence for clarification.

After evaluating the mass balances, we find that (on interannual timescales) these conditions always hold within a reasonable uncertainty range. Even without any further subsampling of the years in which $NAO_s > –0.5$, the results would still hold with an uncertainty of ~10%. Since this is an important result, we have now added (line 248):

"The result implies a close connection between freshwater and SST anomalies. A demonstration of this connection with hydrographic observations shows that, even in winters of most intense air-sea fluxes, it is possible to infer freshwater anomalies from the SST with a reasonably small uncertainty that is below that of currently available satellite products (Appendix A)."

We now also show that the results hold for longer timescales, specifically the trend over the last 70 years (Section 4.6). However, we find that the resulting uncertainties are larger for longer timescales.

13. l. 369 & l. 371: "lower NAO index" and "higher NAO index" – do these refer to the magnitude and/or phase of the NAO?

They refer to the phase. To clarify this, we have replaced all instances of "NAO index" by "NAO state" or "phase".

14. l. 375-377: It would be helpful for the reader if you add a sentence how the relationship was obtained. This is partially described in the caption of Figure A1, which makes it more difficult to follow the arguments in the text.

Thank you for pointing this out. We have now rephrased this sentence to clarify that the relationship was not obtained in that we did not take any steps or created it ourselves. Instead, we observe this relationship from visual inspection of the scatter plot.

15. l. 377: How did you determine the threshold?

We determined it from visual inspection of the relationship between the summer NAO meridional SST gradient. We have now rewritten this paragraph to clarify this (line 184).

16. l. 377-378: I think it is important to also mention the significant positive SST anomaly in the subtropical gyre/western North Atlantic (Figure A1a) which has a substantial contribution to ΔSST for NAO < -0.5.

Thank you for pointing this out. It is indeed important, and we now mention it at the location you indicated. We also explain the development of the warm anomaly through air-sea coupling in more detail in Section 4.3. Overall, we are now more explicit about the role of large-scale atmospheric feedbacks contributing to the SST signal beyond the subpolar region.

17. l. 380-400: The description of the optimization process is unclear. I understand the rationale of increasing the signal-to-noise ratio, however, the selection of included years appears very subjective. How did you choose the number of years to include in the index? How did you select the discarded years? What about the two outlier years?

Thank you for indicating that the optimisation process was insufficiently described. As you correctly point out, the objective of the process is to increase the signal-to-noise ratio.

To better describe the process, we now explain (line 196): "There is a trade-off between the number of years included and the resulting correlation. Here, we selected N=17 years as a reasonable compromise for obtaining a high correlation of 0.90 while keeping a relatively large sample size, reflected in low p-values ($p < 2.6 \times 10^{-6}$). However, the results are not sensitive to this choice."

To add further support for the subsampling, we have included a new section in the results that links the two sub-sampled indices to physical causes of freshwater anomalies, supporting the optimisation process with a physical explanation (Section 4.2).

The two outlier years correspond to years where the NAO index was not a useful indicator for the fresh and cold anomalies. However, upon investigating both years more closely and carrying out a mass balance for both years, we found that the relationship between the salinity and temperature anomalies still held. The discrepancy between the NAO index and these two freshwater anomalies can be explained by a superposition of both, enhanced runoff and changes in the circulation. Thus, the NAO index underestimated the freshwater anomalies. Since we think that these analyses did not add new information to the manuscript, we have not included them in the manuscript.

18. l. 387: Please define the SST gradient in the text.

We now define the SST gradient in the text and explain that it refers to the difference between the subtropical warm and subpolar cold anomaly regions, enclosed within the 95% confidence lines.

We further explain (line 190): "Here, we specifically used the 95% confidence regions as a means to directly inspect the robustness of the correlations and ensure that they are not due to outliers or clusters. Another advantage of using spatial differences is that we filter out any potential, spatially uniform, radiative warming signals. However, the identified relationships are not sensitive to the exact regions."

19. l. 388-389: It is unclear what you mean by “spatial gradients are more robust to local variations in the surface fluxes”. What if the spatial gradient is the result of local heat flux variations as one might expect from the response to the NAO (e.g., Cayan, 1992; Marshall et al., 2001; Deser et al., 2010).

Thank you for pointing out that the sentence was unclear. By using spatial gradients, we filter out the uniform warming effect of increasing greenhouse gases. Another advantage of using spatial gradients is that the area, used for the derivation of the indices is based on overall larger areas and therefore less sensitive to regional fluctuations in surface currents or fluxes.

We now state this in the manuscript, and we also mention that the results are not sensitive to the exact regions (line 191).

The surface fluxes were evaluated as part of the mass balance. They do not substantially contribute to the SST patterns (Figs. A1d and A2d). We now explicitly state this in the manuscript (line 336).

20. Figure A1: It would be more intuitive and consistent with the text if you wrote NAOS < -0.5 in the title of Panel (a) and in the caption. Do you include the significant cold tongue off western Africa in the calculation of ΔSST?

Thank you for pointing out that the title of the figure was confusing. We have now removed the figure.

The cold tongue can be understood as feedback since the associated large-scale atmospheric circulation anomaly induces upwelling off western Africa. However, we did not include the cold tongue off western Africa in the calculation of the SST gradient.

We have now clarified that we only use the subpolar cold anomaly in the text (line 175, line 188) and in the figure caption (Figure 2). Moreover, throughout the manuscript we are more explicit about the atmospheric feedbacks contributing to the large-scale SST pattern in winter (Section 4.3) and we have added a new description of these feedbacks in summer (Section 4.7). Thus, we also mention the enhanced upwelling off the coast of Africa (line 348).

2. Section 4.1:
   1. The circulation anomaly you describe (Figure 4b) is reminiscent of the positive NAO phase for which the atmospheric variability patterns and corresponding ocean response are known (e.g., Cayan, 1992; Marshall et al., 2001), and are in line with your findings. It would make this section stronger if you make an explicit link of your results to the winter NAO.

   Thank you for suggesting this. Following your suggestion, we now make an explicit link to the winter NAO (line 318). We also mention that this signal implies a switch of the NAO sign from more negative in summer to more positive in winter after strong F_E freshwater anomalies.

   2. Changes in the wind field associated with the NAO not only change the Ekman transport as you discuss (l. 198), but also lead to changes in latent and sensible heat fluxes. Can you elaborate to what extend these changes in air-sea heat fluxes are important for creating and maintaining the meridional SST gradient?

   We now state that we do not find any significant impact of the surface heat fluxes on maintaining the SST gradient, neither in winter (line 336, Fig. A1) nor in summer (line 426, Fig. C1). In the mass balance analysis, moreover, the surface fluxes were found to be too weak to contribute to the cold anomaly. However, wind-driven Ekman transport can, and do contribute to the large-scale SST patterns by setting up large-scale surface pressure gradients, consistent with earlier studies (e.g., Marshall et al. 2001; Zhao and Johns, 2014). This is now clarified in the description of the air-sea feedbacks, both winter (Section 4.3) and in summer (Section 4.7).

   3. l. 200: It is unclear why you bring in the second winter. A short motivation will help to keep the reader on board.

   Thank you very much for suggesting this. We now motivate the investigation of the second winter with the climatic repercussions of North Atlantic Current shifts (line 323).

   4. l. 210-213: From the presented figures, I cannot see a northward shift of the North Atlantic Current during the first winter (I think it shows up nicely int Figure 4d for the second winter). Is it possible that different timescales between heat fluxes and Ekman transport can explain the differences between the first and second winter? A SST gradient which is set up in the first

winter and shifted northward during the second year seems also more in line with the summer SST pattern that you describe in Section 4.2 (l. 227-229).

Thank you very much for pointing out that this was unclear.

The Ekman transports are an instantaneous response. They already lead to a warm anomaly in the first winter after the freshwater index. However, the northward current shift is only visible in the warm anomaly to the south of the cold anomaly (the Gulf Stream deflection point). It does not extend across to the east coast.

We now clarified (line 332): "The northward shift of the North Atlantic Current implies a large-scale warm anomaly to the south of the subpolar cold anomaly, not because the water inside the current is anomalously warm but because the current occurs at an anomalously northward location. It is seen in the first, and the second winter after freshwater anomalies (Fig. 2a and 5d). However, in the first winter, the northward shift is obscured by the wind-driven, southward expansion of the cold anomaly over the eastern North Atlantic, potentially driving enhanced mixing and erosion of the SST front."

3. Section 4.2: This section seems rather short given that it addresses one of the main results of the study. It would be helpful for the reader if you add more details and clarify the following points:

Thank you for all your suggestions! Following your suggestions below, we have now expanded the section and clarified these points.

1. l. 231: "more northerly location" compared to what?

"More northerly compared to the previous summer". This is now clarified.

2. l. 237, 238, 241: "cold anomaly" in the ocean or atmosphere?

We referred to the negative SST anomaly. This is now clarified.

3. l. 240: "over Europe" is rather vague (e.g., the warm and dry anomalies (Figures 6c and d) occur in different regions). See also next comment.

Following your comment below, we are now more specific (line 367).

4. l. 241: Is it actually true that "the overall patterns are similar after F_E and F_W freshwater anomalies"? The significant air temperature anomalies one year after F_E extend across the Iberian Peninsula all the way to northern Africa while they are more centered around over France and Great Britain after F_W. Similar the the dry anomaly occurs over the Alps and eastern Europe during the first summer after F_E, but more over Baltic region after F_W which is more similar to second summer response after F_E.

Thank you for suggesting this. We are now more specific and say that the mechanism is the same, but it occurs over a different region, consistent with the underlying SST anomalies. We now exactly specify the regions in the text (line 367).

5. It seems like that patterns after F_W are one order of magnitude smaller compared to the patterns after F_E. Is this an artifact of the smaller correlation in the construction of the freshwater indices or is it due to the stronger meridional SST gradient that exists in the F_E subset with significant positive SST anomalies in the subtropical gyre region?

Yes, you are right that the magnitudes of the obtained signals in European summer weather are carried over from the larger magnitudes in the freshwater and SST signals. The large magnitudes result from steep regression slopes. These steep regression slopes occur because the underlying changes in the F_E index are much smaller compared to the changes in the F_W index. We now clarified (line 370):

"The regressions of the SST and atmospheric circulation on F_W are weaker compared to those on F_E consistent with weaker freshwater anomalies (Fig. 3) and smaller regression slopes (Fig. 2d). Physically, the higher regressions on the F_E subsets imply a higher sensitivity of the ocean and atmospheric conditions to F_E freshwater anomalies. Once the seasonal freshening exceeded a critical threshold (corresponding to a threshold of ~0.5 in the NAO_S index), a relatively small further increase was linked to substantially warmer and drier summers."

This high sensitivity can have severe implications for European summer weather, as the associated, critical threshold in the seasonal freshening may be exceeded more frequently. Thus, we now also refer to this high sensitivity in the implications for predictability (line 486) and in the discussion (line 526).

6. Is there a reason why you show the zonal wind at 700 hPa for the F_E subset and the meridional component for the F_W subset?

To best show the link between the SST and the wind field, we selected the wind component that best matches the shape of the underlying SST pattern.

For consistency across the analyses, and since we specifically emphasise the northward deflection, we now only show the meridional component in all analyses (Figs. 6, 7, 8 and 10).

4. Sections 4.4 and 4.5.: I have to admit that I got lost here. In general, I am wondering whether the analysis of the model simulations adds any additional information that warrants its inclusion in the manuscript.

Following your suggestion, we have removed the previous model analysis. We have now replaced it by another analysis in which the link between the selected SST anomalies for the model analysis and the freshwater anomalies is clearer.

Since we have removed the analysis, the comments below no longer apply.

1. l. 271: It is not clear which pattern (F_E or F_W) you project the on.
2. l. 273: Most of the analysis in Sections 4.1 and 4.2 is focused on the first summer after both F_E and F_W years with only a brief discussion of the second summer after F_E. It is unclear why you construct a new index for the analysis of the model simulations based on the SST pattern in the second summer.
3. l. 289: cold anomalies in the ocean or atmosphere?
4. l. 294-295: Given your derivation of the freshwater indices using the surface mass balance, any cold anomaly coinicides with a freshwater anomaly, by construction. Your analysis of the observations points out the importance of the the meridional SST gradient and its influence on the position of the jetstream. This raises the question whether the freshwater anomalies are just side effect of the mechanism that sets up the SST gradient. It is unclear to me how the model simulations help to answer this question.

In the new analyses, we examine the link between the freshwater anomaly in winter and the SST anomaly in summer more thoroughly. While the freshwater anomalies can explain the initial trigger of the chain of events, the final SST anomalies in summer can only be understood as the result of large-scale air-sea coupling processes. Thus, we are more cautious in the wording in the revised manuscript.

We further clarified the meaning of the previous Section 4.5, which includes the composite analysis of warm European summers. Thus, we now explain (line 389):

"The analysis of Europe's warmest and coldest summers supports the statistical link between freshwater anomalies and European summer weather. It demonstrates that the link is robust to the analysis technique and independent of the indices. While the regressions on the freshwater indices showed that freshwater anomalies can constrain the variability of the subsequent European summer weather, the composites additionally show that, on interannual timescales, Europe's largest temperature anomalies were preceded by freshwater anomalies. This indicates that enhanced freshwater anomalies are not only a sufficient but also a necessary condition for warmer European summers."

III. Additional Comments and Suggestions:

1. l. 37, l. 83, l. 84: It would be more appropriate to use "grid spacing" instead of resolution (e.g., Grasso 2000).

Thank you very much for pointing this out. We have corrected this.

2. l. 44: It's not just cold air, but also stronger winds that increase heat fluxes.

This is true but in the preceding sentence we explain that the air is always colder than the ocean in winter. This naturally implies a mean climatological ocean heat loss. We do not exclude that stronger winds also increase the surface fluxes. They are just not relevant in the context of this paragraph.

We have now removed the sentence since it raised confusion.

3. l. 46: Please summarize the conditions here or refer to the derivation of the freshwater indices.

We now state the conditions.

4. Section 2.1: Please add details about grid spacing, temporal resolution, and any processing (e.g., calculation of anomalies, spatial interpolation, etc.). This would help make the study more reproducible. It might also be worthwhile to specify in this section which months you refer to by "summer" and "winter" throughout the text, especially since they are different from the standard definitions June-August (JJA) and December-February (DJF), respectively.

Thank you for these suggestions. Grid spacing, resolution, and processing steps are now all included in the data section.

We still state the definitions of the seasons where they occur, and we motivate these definitions. For instance, we use the NAO in July and August because runoff is largest in these months. In the analysis of European summer weather, we consider the full period from May through to August (Fig. 6) because we include the full periods over which we observe strong signal. While the exact periods may thus differ, we make sure that we always clearly state the periods.

5. l. 63: How did you combine the two datasets given their different temporal and spatial resolutions?

We did not combine them ourselves, but we downloaded the merged dataset from NCAR. This is now clarified by including the link to the datafile in the data section.

6. l. 94-97: This sentence is unclear. I do not understand why warm anomalies due to shift in the jet stream "must" be balanced by a cold anomaly elsewhere.

In contrast to greenhouse gas warming, freshwater-linked temperature anomalies do not result in a net imbalance in the Earth's surface energy budget. The warming over Europe is balanced by a cooling over the ocean since the underlying baroclinic wave activity consists of an anticyclonic anomaly on one side of the jet stream, and a cyclonic anomaly on the other side.

Thank you for pointing out that this was unclear. We have clarified this in the revised manuscript (line 109).

7. Here are a few wordings that are either inconsistent or remnants of the previous version of the manuscript:

Many thanks for spotting the below inconsistencies! We have now clarified all instances.

   1. l. 185: "in winters after stronger freshwater anomalies" – based on the construction of your freshwater indices, the anomalies should occur during winter.

   We have removed this sentence.

   2. l. 211, 271, 316, 498, 499: what are "freshwater events"?

   We have replaced "freshwater events" by "freshwater anomalies" everywhere.

   3. l. 273, l. 419: What are "melt-driven" or "melt-induced" events? How are they connected to F_E and F_W?

   Thank you for spotting this. We have removed the terms "melt-driven" and "melt-induced" at both locations. Also, we now added a section that explains the term.

   4. l. 421: What are "circulation-induced freshwater events"?

   We have removed the term.

8. Figure 4: I suggest to mask out the Ekman transport vectors over land. This would make it more intuitive that they refer to an ocean variable

Indeed, thank you very much. We have now masked them out.

9. l. 253: Please add a reference for the statement that "most current coupled global climate models have large freshwater biases".

We have now removed the statement at this location since we are already making it in the introduction (line 37) and the discussion (line 528). At both locations, we provide the associated references.

10. l. 408: Do you integrate the wind stress or resulting Ekman transport over the winter period?

We integrated the Ekman transports over the winter period, after computing it from the wind stresses. This is now clarified.

11. l. 429: In l. 412-413, you define the heat flux (Q) as positive downward. A positive surface buoyancy flux (B) anomaly means Q needs to be positive (unless it its overcompensated by the freshwater flux), i.e., the ocean gains heat.

Yes, this is true for the anomalies. However, in the mean state in winter, the subpolar ocean loses heat. Thus, we considered it more appropriate to state "the ocean loses less heat" instead of "the ocean gains heat".

We have now removed the sentence because the anomaly in the surface heat fluxes is very small, and the results are not decisive. The sign of the integrated surface heat fluxes is more sensitive to the exact region compared to the buoyancy fluxes. However, the magnitude of both types of fluxes remains negligible, independent of the exact region.

The observation that the signs of B and Q do not exactly match is not due to the P-E term but due to the dependence of B on alpha, which depends on the skin temperature.

12. l. 439-440: What is the uncertainty in the freshwater fluxes due to the constant mixed layer depth used in your analysis?

We have carefully revised the mass balance and now explicitly account for a variable mixed layer depth (Section 3.1 and 4.1). This does not affect the results.

13. Figure A4: It is unclear whether these are composites just for the winters before the warmest summers or also the difference with the coldest summers.

For consistency with the heat wave composites, the composites refer to the difference between the ten warmest and the ten coldest summers. The surface mass balance was also carried out for the differences between the ten warmest and coldest, which is now clarified.

14. l. 498: This goes back to my first general comment (I.1.): If the SST pattern drives the observed atmospheric response, what is the role of the freshwater anomalies?

Thank you for asking about the role of freshwater anomalies for European summer weather.

We have now included further analyses and are more explicit about the role of the freshwater anomalies. Given the importance of ocean-atmosphere feedbacks in the evolution and intensification of the associated SST signals, we do not attribute the anomalies in European summer weather to freshwater variations alone. Instead, considering the spatial and temporal characteristics of the characteristic SST signals, we propose in the discussion that enhanced freshening may act as a trigger of subsequent, large-scale air-sea feedbacks, which in turn affect European summer weather. Please see first our response to your first comment for more details.

IV. Typos/Wording:

l. 53, 301: “ocean atmosphere” to “ocean-atmosphere”

Thank you for spotting this. We have now corrected it.

l. 273: “over the central North Atlantic” to “in the central North Atlantic”

This sentence has been removed.

References:

Cayan, Daniel R. “Latent and Sensible Heat Flux Anomalies over the Northern Oceans: The Connection to Monthly Atmospheric Circulation.” *Journal of Climate* 5, no. 4 (April 1, 1992): 354–69. https://doi.org/10.1175/1520-0442(1992)005<0354:LASHFA>2.0.CO;2.

Deser, Clara, Michael A. Alexander, Shang-Ping Xie, and Adam S. Phillips. “Sea Surface Temperature Variability: Patterns and Mechanisms.” *Annual Review of Marine Science* 2, no. 1 (2010): 115–43. https://doi.org/10.1146/annurev-marine-120408-151453.

Grasso, Lewis D. “The Differentiation between Grid Spacing and Resolution and Their Application to Numerical Modeling.” *Bulletin of the American Meteorological Society* 81, no. 3 (March 1, 2000): 579–86. https://doi.org/10.1175/1520-0477(2000)081<0579:CAA>2.3.CO;2.

Marshall, John, Helen Johnson, and Jason Goodman. “A Study of the Interaction of the North Atlantic Oscillation with Ocean Circulation.” *Journal of Climate* 14, no. 7 (April 1, 2001): 1399–1421. https://doi.org/10.1175/1520-0442(2001)014<1399:ASOTIO>2.0.CO;2.

We strongly thank the reviewer for reading our manuscript so carefully. The level of detail was extraordinary, and the review included many helpful, and constructive comments and suggestions.

Thank you!

Responses to Reviewer 2

This paper documents an apparent impact of subpolar freshwater anomalies onto the North Atlantic sea surface temperatures (SSTs) in winter and the subsequent summer. These summer SSTs are then argued to drive changes in the atmospheric circulation that drive significant changes in summer temperatures though analysis of reanalysis data and model simulations. Overall, the authors argue that freshwater changes in the North Atlantic can drive a large amount of the variance in European summer temperatures and so argue that such a mechanism could provide significant extra skill in seasonal predictions.

Overall this is an interesting study and presents an exciting set of results. However, I have a number of issues with the manuscript in its current form, which I list below in my major comments. Therefore, I do not recommend publication at this time.

We strongly thank the reviewer for reviewing our manuscript and providing so many detailed and constructive comments and suggestions! Your comments and suggestions have helped us to clarify and improve the manuscript.

Major comments

1. My main concern with this manuscript is that the most important results are buried in the appendix - that is the definition of the freshwater budget method, and the results evaluating the impact of freshwater on the North Atlantic SSTs. I note that this is similar to the structure of Oltmanns et al, 2020, in GRL. This usage of the appendix meant that the paper was quite challenging to read as I had to keep flipping back and forth in the paper to try to understand the logic. Furthermore, I would argue that the most important result in the paper is the evaluation of impact of freshwater onto the resulting SST (e.g., section A3). Therefore, I would strongly recommend the authors to move more of the important background information into the main paper and systematically step the reader through the ideas.

   Thank you for indicating that the usage of the appendix made it difficult to follow the mass balance approach, and that the link between the freshwater and SST is also an important result.

   Following your suggestion, we have now integrated Appendix A3 and A4 into the results section. In addition, we have combined Appendix A1 and A2 with the approach section. Thus, the full derivation and results of the mass balance analysis are now included in the main manuscript.

   We still think that the link between the freshwater and SST anomalies in winter is only part of the results. The link between freshwater anomalies and the subsequent European summer weather is also an important finding. Thus, we have included additional analyses focusing on the link between freshwater anomalies in winter and the SST and European weather in subsequent summers, including the role of air-sea feedbacks. Further details are included in response to your specific comments below.

2. Although I found the results exciting, ultimately the results on the potential impact of freshwater anomalies on Europe were much more uncertain than I feel the authors described. The uncertainty spawns, I think, from the following important reasons.
   - There is little physical understanding of how the summer NAO is leading large freshwater driven SST changes across the subpolar North Atlantic. For example, there

is much discussion about how the sign of the relationship between summer NAO and the winter SST anomaly changes at ~0.5, but no reason for this is given. How can we be certain that the freshwater analysis is picking up real physical changes related to changes in freshwater?

Thank you very much for suggesting adding a physical explanation for the link between the summer NAO and the subsequent freshwater anomalies.

In the revised manuscript, we have now added an analysis of the causes of freshwater anomalies associated with both high and low NAO summers (Section 4.2).

Consistent with earlier studies, we find that a lower NAO phase is associated enhanced seasonal runoff and melting while a higher NAO phase is associated with an enhanced wind stress curl over the subpolar region, resulting in a stronger subpolar gyre circulation and enhanced advection of fresh, polar water into the subpolar region (Fig. 2 of the revised manuscript). While a detailed examination of the freshwater budget is beyond the scope of this study, our findings are consistent with earlier studies on the role of runoff (Bamber et al., 2018; Dukhovskoy et al., 2019) and the subpolar gyre circulation (Häkkinen and Rhines 2009; Häkkinen et al., 2011, 2013; Holliday et al., 2020) for freshwater anomalies in the subpolar North Atlantic.

The threshold of ~ -0.5 in the summer NAO corresponds to a critical surface freshening above which the shallower, seasonal freshwater anomalies are mixed down, such that deeper, circulation-driven signals dominate the hydrography of the subpolar gyre. We also investigated the surface freshwater fluxes (precipitation minus evaporation) but found that there were unimportant for the identified freshwater anomalies.

- Related to the above, the paper focuses on interannual changes - however, there is plenty of observed evidence that there is significant decadal time-scale variability in the subpolar North Atlantic, and so questions arise on what the freshwater analysis, which appears to assume independence between years, is picking up. For example, figure 2c shows that it is the decadal time-scales that dominate the summer NAO time-series, consistent with a southward shift of the North Atlantic Jet (e.g., Dong et al, 2013). Given the low-frequency changes, how can the authors be sure that they are seeing the interannual impact of freshwater changes as opposed to a different, or related, mechanism? For example, is the summer NAO and winter SST responding to another mechanism which is driving both (e.g., large-scale ocean circulation, external forcings etc)? As mentioned below, I think overlaying a timeseries of summer NAO with subpolar SSTs would be useful in seeing the potential importance of these longer timescales.

Thank you for pointing out that the involved timescales of freshwater variability were unclear.

In the revised manuscript, we are now more specific about the involved timescales. Thus, we state that the focus of this study is on interannual timescales, consistent with the high interannual variability of European summer weather (line 297). We also mention potential links to low-frequency variability in the causes of freshwater anomalies (line 296). Thus, in the past, decadal switches of the Arctic Ocean Oscillation have led to periodic exports of fresh and cold anomalies from the Arctic into the subpolar North Atlantic (Proshutinsky et al., 2015), which likely contributed to the decadal variability of the SST pattern in winter.

In addition, we now show the long-term freshwater trend over the last 70 years (Fig. 9, Section 4.6). Thus, we conclude that cold anomaly patterns in winter and summer are linked to freshwater anomalies over a range of timescales. While the described, dynamical feedback mechanisms act on short, sub-seasonal to interannual timescales (baroclinic wave activity in the atmosphere, Ekman transports, geostrophic adjustments), the broad range of timescales in SST and freshwater variability can be explained by the physical causes of freshwater anomalies, acting on seasonal to decadal timescales.

The autocorrelation of the summer NAO is only small at one-year lag (Fig. 9f), and the trend in the summer NAO over the last 70 years is ~–0.002 yr^{-1}. This suggests that the summer NAO mostly reflects high frequency, interannual variability of the cold anomaly pattern in the subsequent summer. Since the autocorrelations of the NAO indices are only small, we do not expect decadal variability to substantially affect the results.

- The model experiments are interesting, but they do not, of course, explore the impact of freshwater changes on European weather. Indeed, the models are feeling the influence of both SSTs and, presumably, external forcings, and the assumption is that the influence of freshwater can be isolated by focusing on the summer NAO analysis. However, and related to point b, this relies that there are no other mechanisms in play that could explain both the summer NAO and the winter SSTs.

Thank you for pointing out that the role of freshwater anomalies for European summer weather was unclear. To address your comment, we have implemented four main changes:

(1) We have included a description of the drivers of freshwater anomalies (Section 4.2). Thus, we provide a physical explanation for the start of the chain of feedbacks. Since the identified drivers (runoff for F_E events and circulation changes for F_W events) are not otherwise linked to European summer weather on the investigated timescales, and occur exactly one year in advance, they challenge the idea that an unknown third mechanism drives both freshwater anomalies and warmer European summers without the two being physically connected.

(2) We are explicit about the involved timescales of variability. While the focus is on interannual timescales (line 296), we now show that surface freshening has a trend (Fig. 9a of the revised manuscript), superimposed on substantial interannual variability that is reflected in the variability of the summer NAO and runoff from the preceding year, obtained from the Greenland climate model MAR. (Section 4.6). After considering alternative drivers acting on these timescales, we find (line 405):

"Apart from seasonal runoff- and melt-driven freshening, there are currently no conceivable, physical mechanisms in the tropics, stratosphere or outside the North Atlantic which, at the same time, have a significant trend over the last 70 years, exhibit a similarly high interannual variability, can explain the occurrence of freshwater anomalies in the subpolar region in the subsequent winter, and are significantly correlated with the characteristic summer SST pattern a year before it occurs. This suggests that seasonal freshening may not only be a predictor of the SST pattern but also a potential trigger."

(3) To assess whether runoff is a trigger of the SST pattern, rather than only a predictor, we have included an analysis of the ocean-atmosphere feedbacks that contribute to the evolution of the SST pattern in summer (Section 4.7). After investigating the surface heat and momentum fluxes and their influences on the ocean and atmosphere with ERA5, remote sensing data, in-situ hydrographic observations from the cold anomaly region, and with the models, we find that wind-driven transports may contribute to the intensification of the SST signal. Thus, we conclude (line 453):

"We conclude that the atmospheric forcing contributes to the development of the SST field but is in turn, forced by it. On the one hand, the consistency of the spatial SST pattern with the wind-driven transports, the intensification of the SST signal in mid-summer, and the vertical extent of the hydrographic anomaly point to the relevance of the atmospheric forcing in driving the SST signal over the North Atlantic. On the other hand, model simulations, forced with the prescribed, observed SST, reveal the importance of the SST field for the large-scale atmospheric circulation, including for European summer weather. The underlying SST pattern covers the entire North Atlantic with the strongest signal occurring in the subpolar region (Fig. 11b). While further studies are necessary to confirm the dynamical contribution of freshwater anomalies to the large-scale SST pattern, its spatial and temporal characteristics, and the time lag of one year, indicate that enhanced surface freshening from the preceding year may have initiated the chain of air-sea feedbacks."

(4) We are more cautious about the wording. Throughout the analyses, we refer to freshwater as a predictor (e.g. line 350). In the conclusion, we discuss evidence that points to freshwater as a trigger rather than only a predictor, but we are explicit about the involved uncertainties, and we have removed phrases that may previously have caused confusion.

> Indeed, I can't help noticing that the SST patterns being explored in the models (figure 8a and figure 9a and b) look like the cold AMV pattern (e.g., Zhang et al, 2019). Therefore, how can the authors be confident that the atmospheric circulation patterns are being driven by just the subpolar North Atlantic, rather than the tropical North Atlantic?

Thank you for indicating that large-scale ocean-atmosphere feedbacks over the full North Atlantic were insufficiently described.

In the revised manuscript, we now describe the ocean-atmosphere feedbacks more clearly in winter (Section 4.3), and we have added a new analysis on the involved feedbacks in summer (Section 4.7).

The SST in the subtropical and subpolar regions is highly coupled due to large-scale atmospheric feedbacks. For instance, the large-scale atmospheric circulation anomaly, associated with the subpolar cold anomaly in winter, leads to convergent, wind-induced surface currents in the inter-gyre region, and thus a warm anomaly that emphasises the SST gradient. At the same time, the anti-cyclonic, atmospheric circulation anomaly to the south of the SST front leads to upwelling off the coast of Africa, contributing to the characteristic tri-pole pattern of the AMV.

Moreover, we have now removed the previous model analyses and replaced it with a new analysis where the underlying SST_{FW} pattern in summer is more clearly linked to the preceding freshwater anomalies (Section 4.7). Considering that the SST anomalies are largest in the subpolar region and decrease towards the south (Fig. 11b), it can be explained by the large-scale atmospheric feedbacks associated with freshwater anomalies. This includes a range of timescales, including that of the AMV.

However, we are cautious in the wording about the role of freshwater in triggering these feedbacks. We refer to freshwater anomalies as predictor. While we discuss evidence in the conclusions that suggests freshwater may act as a trigger of the large-scale ocean-atmosphere feedbacks, rather than only a predictor, we state this as a hypothesis, not as a result (line 511).

Minor comments

Line 1 - The paper discusses here and in the introduction the possibility of long-term forced changes in the arctic having an impact on the mid-latitude weather, but then exclusively focuses on interannual variability? This left me slightly confused about what I was supposed to be taking from the analysis, so maybe it worth rethinking the motivation?

Thank you for pointing this out.

We have now added an analysis of the long-term variability of freshwater and the implications (Section 4.6). Thus, we find that there is a significant freshening trend over the last 70 years, superimposed on substantial interannual variability. This is consistent with seasonal runoff in the subpolar region, which has a high interannual variability but also a trend.

Line 63 - why not just use HadISST rather than merge the two datasets?

We did not manually merge the two datasets ourselves. The merged dataset is freely available and benefits from the optimal interpolation technique used in the NOAA data.

We have now added the link in the data description to avoid misunderstandings (line 59).

Section 2.2 - how do these simulations differ from CMIP style AMIP runs (e.g., Eyring et al, 2016)?

The simulations were designed to facilitate comparison between different models, like those in Eyring et al. (2016). Thus, they have a pre-determined forcing to facilitate the comparison, which is now clarified.

In this case, the simulations were performed with the prescribed, observed SST and sea ice cover and time varying greenhouse gases and ozone (line 90).

Section 2.3 - I am confused by the discussion of a trend, in part due to the framing of the paper on understanding the impact of long-term forced changes in the Arctic on the wider climate, but I think the point is that trends over land may alter the relationship with ocean SST anomalies? Please clarify.

We do not distinguish between trends over land and ocean. The intention of this approach was to remove the effect of Greenhouse gases, not to remove trends in general. This is now clarified (line 97).

Moreover, we have adjusted the subsequent paper to the motivation in the introduction. Thus, we have added an analysis of the trends in the freshening and SST over the last 70 years (Section 4.6).

Line 133 - You state that "we derive indices that exhibit a strong relationship to subpolar temperatures but not to the drivers of density anomalies", but then argue that the summer NAO drives changes in freshwater (which will drive the density anomalies) - can you clarify your point, please?

Thank you for pointing out that this was unclear. We referred to the active drivers of density (or in turn temperature and freshwater) anomalies. We consider freshwater as state variable that in turn has drivers, like precipitation or runoff.

We have now clarified this in the manuscript (for instance line 149). We have also removed the sentence you were referring to, since it raised confusion.

Line 140 - you mention the potential impact for the summer NAO on freshwater, but won't it also affect the heat fluxes too? How do you take account of these in your analysis?

Yes, we do consider the surface fluxes in the mass balance analysis. The surface heat fluxes are included in the buoyancy fluxes, and we also show them separately (Figs. A1d and A2d). Surface heat fluxes can be a driver of SST anomalies or a response. In this case, we find that that the surface heat fluxes in winter cannot explain the SST anomalies. They are too small, and their distribution is inconsistent with the SST anomalies.

We now explicitly state in the revised manuscript that the surface heat fluxes cannot account for the SST anomaly in winter (line 336). In addition, we now show that the surface heat fluxes in summer also cannot account for the SST pattern in summer (Fig. 1C).

Line 392 - "...we select all the years that lead to an increase in the slope of the regression line." - I'm uncomfortable about this - it sounds a bit like cherry picking to get a stronger signal (which in turn would affect your conclusions about how important freshwater changes are). Please could the authors justify this choice physically?

Thank you for suggesting this. Following your suggestion, we have now added a description of the causes of freshwater anomalies associated with high and low NAO summers (Section 4.2). Thus, we find that the subsampled indices are also more strongly linked to the physical causes of freshwater anomalies, supporting the subsampling with a physical explanation.

The objective of the subsampling is to improve the signal-to-noise ratio between the index and freshwater anomalies. The strong signal between the index and the freshwater anomalies is a pre-requisite for the index to represent freshwater anomalies. We do not use it as the conclusion.

We now explain the method more clearly and have integrated the appendices into the main text, making it easier to follow the steps (Section 3.2).

In addition, we have included an analysis of the link between European summer weather and the preceding summer NAO, without any subsampling (Section 4.7). Combined, these analyses show that the results are neither sensitive to the choice of the index (including the sampling) nor the statistical method (regressions or composites, Sections 4.4 and 4.5).

Figure A1 - I can't help but to see that there is a long-term change in the summer NAO with most negative years occurring after 2000. This is consistent with the long-term trend in summer jet. The question this raises is how important is this longer term variability in the sNAO for your results? One thing that is missing is a time series of subpolar SST, which also shows significant decadal timescale variability. Please add this to figure A1 at least. However, following up on this, previous studies have linked these low frequency variability in the jet to the changes in the subpolar temperatures (e.g., Dong et al, 2016) - therefore, how certain are you that the changes in summer NAO are not reflecting decadal time-scale changes in the subpolar gyre?

We now show the timeseries of the summer SST pattern linked to freshwater anomalies (Figs. 9 and 11) since this pattern is relevant for European summer weather. A related timeseries for winter is included in Fig. B1.

While there has been significant decadal variability in the causes of freshwater anomalies in the past, the negative NAO picks out the high-frequency variations, associated with the interannual variability of runoff (see also our response to your second comment). Since the focus of this study is on the influences of freshwater anomalies rather than the causes, we think that showing the time series for the winter SST pattern in the main text would not add substantial information.

Line 405 - It's not clear to me why the geostrophic currents are not important - there is nothing in the 95% confidence lines that are related to constant lines of density. Please can you clarify your reasoning here?

Thank you for pointing out that this was unclear. Geostrophic currents at the surface are along lines of constant pressure and density. Thus, they cannot lead to a net increase or a decrease of density. Importantly, they can, and do, drive freshwater and temperature anomalies. However, the density anomalies associated with freshwater and temperature changes need to compensate each other in geostrophically balanced flows.

We have now clarified that geostrophic surface flows cannot cross density contours (line 554).

Section A3 - it is not clear why you are regressing the JFM heat fluxes onto your summer NAO? In order to rule out the impact of surface heat fluxes on the JFM SSTs it is the integrated heat fluxes between the summer and JFM that are important?

Thank you for pointing out that this was unclear. We are examining the SST signal in winter, which is typically not influenced by the surface fluxes before the winter because the amplitude and variability of the surface fluxes in winter is much larger. In addition, the surface fluxes in autumn respond to re-emerging ocean anomalies in the mixed layer (Timlin et al., 2002), thus balancing the term M_n when integrated over longer periods (line 546). To be sure, we have repeated the surface mass balance for the period from summer to winter and found that the results did not change appreciably. In the case of the F_E subset, the overall amplitude slightly decreases, implying slightly smaller uncertainties. For the reasons stated above, however, we think that the surface fluxes in winter provide the more accurate estimate.

We now motivate the focus on the winter period by the tendency of B_n anomalies to balance anomalies in M_n when integrating over the period of rapid mixed layer deepening. Consistent with earlier studies, we also find that the amplitude and variability of the surface heat and buoyancy fluxes is largest in winter (line 546).

Moreover, we state in the revised manuscript that the results do not change appreciably if we integrate over autumn and winter or only winter (line 226).

417 - the mean mixed layer is the annual mean mixed layer?

Yes, indeed. Since we are investigating the SST anomalies in winter, we also need to use the mean mixed layer in winter.

We now clearly state the averaging and integration periods in the manuscript (line 134, line 580).

Line 440 - Please elaborate on the origin of the uncertainty numbers.

Thank you for pointing out that this was not sufficiently clear. The uncertainty originates from neglecting the terms on the righthand side of the mass balance equation. It is obtained by comparing the size of these terms to the size of the terms on the lefthand side.

We now devote a full section in the main text to the derivation of freshwater anomalies and their uncertainties (Section 4.1).

Figure 3 - what is the data used to plot the SSS anomalies, or are they implied from your mass balance analysis?

Yes, the estimates were obtained from the mass balance analysis. This is now clarified in the revised manuscript (Section 4.1, line 241, and Figure 3 caption). Thank you for letting us know that this was unclear.

Section 4.2 - I think when you mean first summer, you mean the same summer as the summer NAO event? However, it was not clear as you're talking about looking in subsequent summers which I took to be after the winter - please clarify. This is particularly important for figure 6, which is unclear what summers you are looking at.

No, we refer to the European weather in the summer one year after the NAO index. It is not the same summer. This is why the summer NAO can be a such a valuable predictor of European summer weather, one or even two years in advance. We have now clarified this (line 352, line 364, line 420, and captions of Figures 6, 7, 10 and 12). Thank you for pointing out that this was unclear since it is indeed important.

Figure 7 - Are you computing the variance explained using all years, or are you only focusing on the years that you have large changes? Either way, how important is the long-term trend in atmospheric circulation shown in figure 2c for your analysis?

The variance is computed over all the years for which the indices are defined. In addition, we now state the variances explained by the winter SST gradient and the un-subsampled summer NAO over all years (line 474).

Given the low autocorrelations of the summer NAO, we do not expect the long-term trend to substantially affect the explained variances. Also, the subsampling ensured that the high correlations are not due to outliers or clusters (Fig. 2d). Thus, the results are not sensitive to excluding individual years, implying an increased robustness of the results.

Figure 9 - Is the only difference between these figures and that shown in figure 8 (d and f) that they are split across model ensembles? Not clear to me why this is relevant - could you just include the ensemble spread in figure 8 ?

The main intention of the figure was to show the spread across the ensemble members and how they compare with the observations.

We have now removed this figure.

References:

Bamber, J. L., Tedstone, A. J., King, M. D., Howat, I. M., Enderlin, E. M., Van Den Broeke, M. R., & Noel, B. (2018). Land ice freshwater budget of the Arctic and North Atlantic Oceans: 1. Data, methods, and results. *Journal of Geophysical Research: Oceans*, *123*(3), 1827-1837.

Dukhovskoy, D. S., Yashayaev, I., Proshutinsky, A., Bamber, J. L., Bashmachnikov, I. L., Chassignet, E. P., ... & Tedstone, A. J. (2019). Role of Greenland freshwater anomaly in the recent freshening of the subpolar North Atlantic. *Journal of Geophysical Research: Oceans*, *124*(5), 3333-3360.

Hakkinen, S., & Rhines, P. B. (2009). Shifting surface currents in the northern North Atlantic Ocean. *Journal of Geophysical Research: Oceans*, *114*(C4).

Häkkinen, S., Rhines, P. B., & Worthen, D. L. (2011). Warm and saline events embedded in the meridional circulation of the northern North Atlantic. *Journal of Geophysical Research: Oceans*, *116*(C3).

Häkkinen, S., Rhines, P. B., & Worthen, D. L. (2013). Northern North Atlantic sea surface height and ocean heat content variability. *Journal of Geophysical Research: Oceans*, *118*(7), 3670-3678.

Holliday, N. P., Bersch, M., Berx, B., Chafik, L., Cunningham, S., Florindo-López, C., ... & Yashayaev, I. (2020). Ocean circulation causes the largest freshening event for 120 years in eastern subpolar North Atlantic. *Nature communications*, *11*(1), 1-15.

Proshutinsky, A., Dukhovskoy, D., Timmermans, M. L., Krishfield, R., & Bamber, J. L. (2015). Arctic circulation regimes. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *373*(2052), 20140160.

Timlin, M. S., Alexander, M. A., & Deser, C. (2002). On the reemergence of North Atlantic SST anomalies. *Journal of climate*, *15*(18), 2707-2712.

Again, we sincerely thank the reviewer for carefully reading and reviewing our manuscript and providing so many detailed and constructive comments and suggestions. Based on your comments and suggestions, we were able to clarify and improve the manuscript. Thank you!