Review of Northern Greenland transect stacked ice cores as a proxy for winter extreme events in Europe

In this manuscript, Gagliardi et al combine a stack of ice core records and a paleoclimate reanalysis to study past regimes of atmospheric circulation, and in particular conditions of atmospheric blocking suitable to general extreme events in Europe. The study uses both datasets from observational period (1920 to present) and from a long term perspective (1602 to 2011) and shows that the atmospheric circulation patterns over the two periods are relatively similar.

The manuscript is well written, the analyses are sound and appropriate to study the dynamical systems. From a paleoclimate point of view, I feel that the manuscript is barely scratching the surface and that the manuscript as it is now is missing a discussion: the results are clear, but their consequences is not discussed: what does it mean that similar atmospheric blocking conditions can be found in both the 1602-2011 and the observational period? Considering that $\delta^{18}O$ is also a temperature proxy (Hörhold et al., 2023), is it possible to disentangle the blocking conditions (characterised with $\delta^{18}O$ below -1σ) happening less/more often with colder/warmer conditions from the thermodynamical response of water isotopes? i.e. with a warmer average conditions, should the -1σ threshold also evolve? And finally, and maybe the more importantly, can you conclude anything on the impact of the anthropogenic climate change on the frequency of extreme events in Europe from the stack and the reconstruction? I believe that this manuscript would be a great addition to Climate of the Past once these questions are answered. I include general and specific comments below.

We would like to thank the reviewer for the appreciation/suggestions/comments/feedback that will help us improve our manuscript, and for taking the time to read and review our paper.

General comments

As discussed above, I feel like the manuscript is missing a discussion that would put the results into a larger context, as well as to discuss the limitations of the datasets used. In particular, I believe that answering the following questions would be beneficial to the study:

1. The long term period analysis is based on a paleoreanalysis, which rely on a dataset of proxy records and instrumental measurements. Before 1850, the reanalysis is constructed using almost solely tree ring records. How do the limitations of using tree rings for the paleoreanalysis affect your results? Typically, in order to reconstruct fields of atmospheric circulation patterns, temperature, and precipitation, the EFK v2 makes use of ECHAM5.4 and these datasets, but this is still a reanalysis based on a limited set of proxies which themselves have some well-known biases in term of reconstructing variability, including the change of growth rate for the different life stage of the trees leading to non-linearity in the relationship between the tree ring growth and isotopic composition and the local climatic conditions.

   The EKF v2 reconstruction is indeed mainly based on tree rings before 1850. However,

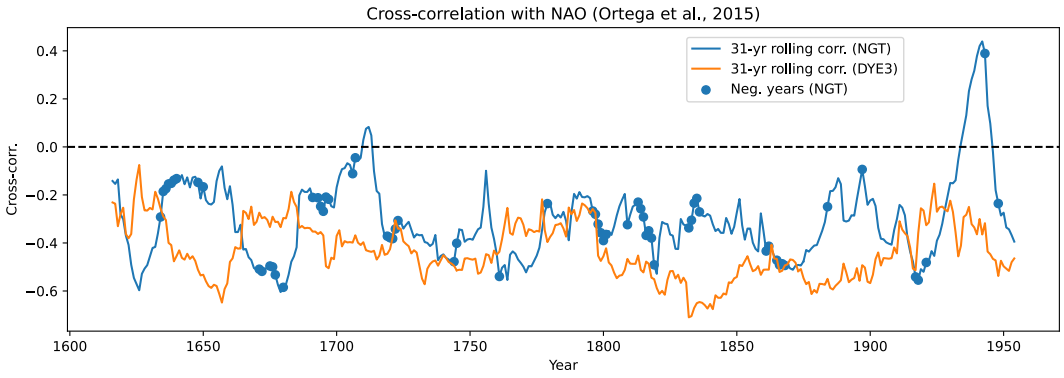
the average geopotential height pattern during negative years in the NGT stack is consistent with the pattern observed in the instrumental period, where the reanalysis does not rely on proxies. In addition, over Greenland, the negative anomaly in geopotential height is also captured, despite the absence of tree-ring data in this region. It is also important to note that tree rings are generally better proxies for summer-based reconstructions. Therefore, the biases introduced by tree-ring proxies are mostly evident in the summer season. Moreover, the EKF re-analysis assimilates also other types of proxies (e.g., corals) as well as old documentary evidence.

2. It also raises the question of the weight of the reconstruction from tree rings in the paleoreanalaysis. The consistency between the results from the observational period and the long term perspective periods (lines 179-180) could also be linked with the tree ring reconstruction might share some of the variance of the ice core stack, are just representative the same mode of variability. While it's beyond the scope of the manuscript to compare the tree ring variability with the ice core variability, I think that a critical discussion of the impact of the tree ring reconstructions on the paleoreanalysis in the framework of comparing it with another paleoclimate reconstruction could be valuable.

   As stated by the reviewer, the interesting question on how tree ring reconstructions impact paleoreanalysis is beyond the scope of the paper. However, we are going to try to make it clearer in the manuscript about the limitations of the tree ring in the paleoreanalysis.

3. The NAO is mentioned once in the introduction and then not a single time in the manuscript before the conclusion where an entire paragraph discusses the link between the atmospheric patterns described here and the NAO. The conclusion should not include new information, and this highlight the lack of discussion section in the manuscript. In term of content, how does the NGT stacked δ18O compare with NAO indices (Ortega et al., 2015)? The paragraph in the conclusion doesn't appear convincing: all the ice cores from the stack should be under the same influence of NAO patterns considering the relatively small area in which they were found (Casado et al., 2013).

   The plot below shows the 31-year rolling cross-correlation between the NAO reconstruction by Ortega et al. (2025) and the $\delta^{18}O$ records from the NGT stack and DYE3 ice cores (Rasmussen et al., 2022). Negative years are marked as dots, which are plotted over the extreme negative years (-1σ) in the NGT stack. Overall, the cross-correlation between the reconstructed NAO index and the NGT stack is, for most periods, lower than that between the reconstructed NAO and the DYE3 core. The reviewer's comment allows us to rephrase the manuscript to clarify that the NGT stack is not completely unaffected by the NAO. Our point is that the stacked signal from northern Greenland ice cores reflects not only the NAO but also the increase in atmospheric blocking events occurring over Europe. In the revised version of the manuscript we will add more information about NAO in relationship with the NGT timeseries.



4. Since you have a 400-year reconstructions, how are the blocking conditions changing over time? Is there a link between the temperature (which also affects the NGT stacked δ18O) and the blocking conditions?

   The atmospheric blocking events appear to be quite stable, as shown in Supplementary Figure A.3. The only period where there is less agreement with other intervals is the transition phase following the end of the Little Ice Age (LIA).

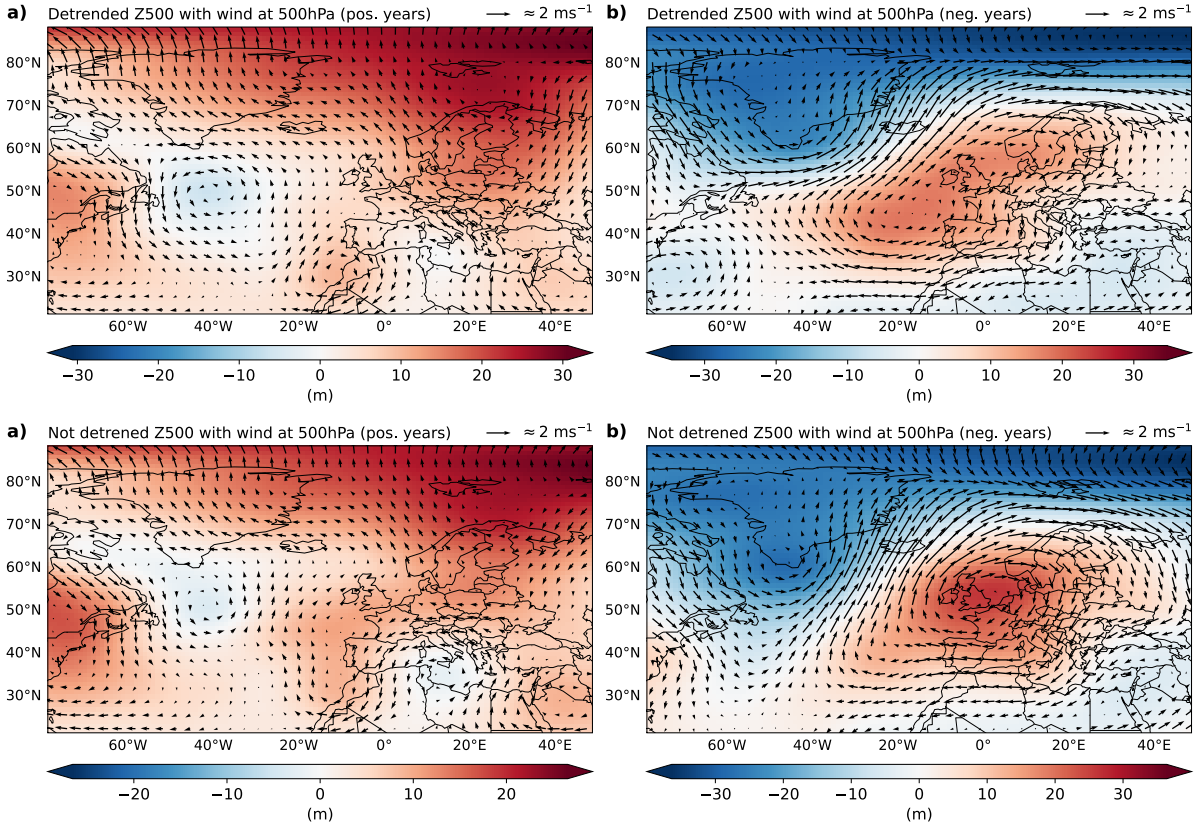
   The connection between temperature and blocking events in Europe arises from the fact that, during such events, less moisture and relatively warmer air masses from the mid-latitudes are able to reach Greenland. Consequently, the relationship between temperature in northern Greenland and temperature in Europe is also mediated through atmospheric blocking events. We will also add this information in the revised version of the manuscript.

5. (Hörhold et al., 2023) identified a large warming in Greenland, with an increase of δ18O. Here, you are using detrended δ18O, which remove this effect. Nonetheless, it should be discussed that the minima, in particular the recent ones in the 1980s are actually associated with values close to 0‰. In particular, two aspects are key to be mentioned: (i) (Hörhold et al., 2023) shows that there is a regime change with a trend changing around 1800, so the detrending from 1602 – 2011 is not necessarily physically based, how does the window used for the reference trend is affecting your results? and (ii) how does a warmer baseline affect your results? Overall, it's not clear to me after reading the article if the detrended NGT variability is a direct signal from the atmospheric circulation, or temperature variability in Greenland that happens to be, at least partly, correlated to atmospheric circulation.

   (i) We agree that a linear detrending may not fully capture the change around 1800. At the same time, we acknowledge a shift in the NGT stack during this period and tested some data-driven change detection methods. However, applying such approaches would introduce further subjectivity, since the outcome depends on the choice of method. In our tests, using a data-driven method for the period 1750–2011 (https://pypi.org/project/pwlf/), the detected change in trend occurs around 1991/1992. This mainly affects the last 20 years of the series, where only one year

shows an extremely low value. Therefore, we consider the effect of using a linear detrend on the results to be negligible.

(ii) We run again the composite analysis for not detrended NGT stack series. The results are essentailly the same, therefore the baseline is not affecting much the dynamics.



Specific comments

Lines 28 to 29:	When such extreme events persist over a region for extended periods, they can be classified as extreme climate events." While I feel the goal here is to distinguish between extreme weather events and extreme climate events, the sentence is not very clear. We are going to rephrase the sentence and make it clearer the concept between extreme weather and climate events.
Line 31:	"However, the lack of high temporal resolution in proxies data makes a challenge reconstructing weather extreme events."

	I don't think proxies data is a clear concept. Paleoclimate records maybe? We are going to rephrase it.
Lines 31 to 35:	Overall, this paragrap seems a bit weaker than the rest of the introduction, because it seems you're not saying what you want. Since it's not clear what you are studying here, it's not clear to see why tree ring reconstructions are limited. In the abstract, you mention 1602 to 2011, but there are multiple reconstructions from tree ring covering this time span, for instance, Freund et al., 2023 covers exactly this window. Tree rings are valuable proxies, but they mainly reflect summer conditions. In contrast, ice cores are more suitable for reconstructing winter climate, especially in the case of Greenland. We will rephrase the text to make this clearer.
Line 36:	"Ice cores records can be used for multidecadal and longer time scale reconstructions (Rimbu and Lohmann, 2010b)." Yes, but this is not maybe the most relevant citations for this, and seems to promote self citation quite a lot. Clearly, papers ranging from Vinther et al, 2010 to GRIP/NEEM papers would be more relevant here. We are going to modify including more relavant citations.
Lines 36 to 38:	"The growing number of high temporal resolution ice cores from the Greenland ice sheet gives valuable information on climate variations from seasonal to multidecadal time scales." This sentence should be justified, but it's not clear to me that obtaining high resolution ice cores is new. We are going to justify and rephrase the sentence.
Lines 40 to 42:	"Recent studies, though, have identified strong links between Greenland δ18O variability and atmospheric weather regimes (Rimbu and Lohmann, 2010a; Ortega et al., 2014) and relationship with atmospheric blocking during boreal winter months (Rimbu et al., 2007, 2017, 2021)." 10 to 15 years old studies cannot be really that recent. Overall, the introduction does not need to emphasize so much on how recent records are, but should focus on giving readers information about the important aspects of what can and cannot be done with ice cores.

	We will include more details about the possibilities and limitations of using ice cores.
Lines 44 to 45:	"To this end, this paper assesses the validity of the δ18O variability in the Northern Greenland Transect (NGT) stacked ice cores (Hörhold et al., 2023) is a proxy for extreme climate events." You mention reconstruction from 1600's to 2020's, while the NGT stack goes all the way back to 1000 AD. Why are you stopping there? It seems peculiar that you put so much value on the NGT stack, and not so much on the EKF paleoreanalysis which is as important if not more important to your analysis than the NGT stack. We limited the investigation period to 1600 due to the availability of the EKF paleoreanalysis dataset. However, we acknowledge the added value of discussing the NGT stack values before 1600, and we will include such a discussion.
Lines 128 to 130:	"The average pattern in negative years features a high-pressure system extending from the Azores Islands to the Baltic Sea and low-pressure system over Greenland, whereas the average circulation pattern in positive years is not close to be the opposite of that in negative years." It's difficult not to think of the link with NAO here. We agree that, as stated, it may resemble a purely NAO+ signal. However, our point is that the NGT stack reflects more than just NAO influence, unlike the southern Greenland ice cores. We will rephrase to clarify that our argument is not that NAO+ has little or no influence, but rather that the signal also reflects atmospheric blocking events.
Figure 2:	Shouldn't there be a figure, at least in supplement that show the reference against which the anomalies have been plotted ? Here, it's difficult to know for instance if the changes are equivalent to less strong winds toward Europe in negative years, or actually an opposite wind direction. We did not compute anomalies relative to a reference period because our goal was not to assess the impact of global warming. By applying a linear detrending over the entire periods, the observational period (1920–2011) and the long-term period (1602–2003), the fitted values (detrended series) naturally have a mean zero value.
Lines 141 to 144:	"Given the clear atmospheric circulation pattern observed during the negative years of the NGT stacked $\delta^{18}O$ series, the blocking pattern highlighted by the two atmospheric blocking indices and the role of atmospheric blocking in favoring extreme weather events (Rex, 1950),

	the following analyses of temperature and precipitation effects will focus exclusively on negative years." I'm not sure that this is a very sound argument, yes it peaks around 5% for the negative years in Fig 3a versus 1% for the positive years, but 1% is still quite a large number of occurence. Since it's over the ocean mostly, the effects aren't crucial, and you are more interested about Europe? We focus on the negative years in the geopotential pattern shown in Figure 2b, since these years display nearly four times more atmospheric blocking events compared to positive years. For this reason, we considered it more relevant to emphasize the negative years.
Lines 176 to 177:	"The observed patterns in temperature and precipitation results to be more regular than the observational period due to the use, in this case, of a reanalysis product." This sentence is unclear, are you talking about the long term perspective or something else, and also because the datasets used are reanalysis for both the observational period (20thcentury reanalaysis) and long term perspective (EFK v2 paleoreanalaysis). We agree on the fact that the sentece is not clear and we are going to rephrase it.