Review of ACPD-2018-636

"Heat Transport Pathways into the Arctic and their Connections to Surface Air Temperatures"

by D. Mewes and C. Jacobi

This paper addresses the wintertime transport of heat into the Arctic and aims at identifying the major pathways and trends in those. For that purpose the authors apply the method of Self-Organizing-Maps to vertically integrated horizontal heat transport derived from the ERA-Interim reanalysis. They identify three preferential large-scale patterns of heat transport over the Northern Hemisphere mid- and high latitudes. Trends in the frequency of these transport patterns are then related to warming trends in the Arctic. Specifically, their analysis suggests that the transport through the North Atlantic sector has become stronger throughout the analysis period.

Given the amplified warming of the Arctic, especially in the European high Arctic, the topic of this paper is highly relevant. Furthermore, the approach to use Self-Organizing-Maps for identifying the dominant atmospheric patterns of heat transport – though not novel to the stratification of atmospheric data – is certainly innovative. The study suffers, however, from several methodological issues as outlined below. These issues need to be resolved before this paper can be considered publishable. Furthermore, the overview of the existing literature is somewhat on the short side and should be extended. Overall, my recommendation is major revisions.

Specific comments:

1. I am puzzled that there is such a strong heat transport over Greenland in many of the clusters (most pronounced in the cluster 2.2, 3.1 and 3.2). The transport clearly seems to increase there. I would expect the opposite since the vertical integral of the heat transport is proportional to the thickness of the layer from the surface to the 200 hPa level such that over high topography it should decrease. The same concerns intense transport over the Tibetean plateau and across the Rockies in some clusters.

   I strongly suspect that the authors used data on the default pressure levels for computing the vertically integrated heat transport and they did not take into account that some pressure levels intersect the surface. Over topography, where surface pressure can be below the pressure of, for example the 1000 hPa level, the ECMWF provides extrapolated fields. The heat flux computed from such extrapolated fields is unphysical. This suspicion is further fuelled by the fact that there are no substantial T2m anomalies over Greenland in clusters with a strong heat transport there.

   The method section does not provide evidence of how the computation is done and what type of levels are used. It only states that the data is available on 37 levels. The default set of pressure levels comprises 37 levels, whereof only 23 are located between 1000 hPa and 200 hPa. Anyway, the above should be clarified and if my suspicion is true, the computation of the vertically integrated heat flux needs to be redone properly. Otherwise the reason for the high heat transports above high topography must be explained.

2. The authors decide to perform a SOM analysis using 4x3 patterns. The number of patterns needs to be prescribed and is a subjective choice required by this method, which as such is fine. What I am worried about, however, is that they subsequently lump several clusters together according to the heat transport at 80°N, resulting in three major patterns, which are then considered for the rest of the paper. The only argument provided to justify this approach is that "this is common practice in SOM analysis". In my view, the SOM analysis on 2d fields is an overkill to obtain just these three patterns that are based on 1d fluxes at 80°N. If mainly the transport at 80°N is of interest for stratifying the patterns, why not perform an EOF or clustering analysis on the transport at 80°N and start from there?

   To make the effort put into the SOM analysis more valuable, I'd suggest to present some additional analyses that make use of the nuances displayed in the 12 SOM clusters. For example, the authors could display other composited fields for these clusters. These fields could, for example, be T2m anomalies, and frequencies of atmospheric weather systems (cyclones, blockings). The latter would allow to relate the heat transports patterns to the dynamics, which I think would be a valuable addition and strengthen the paper.

3. There is a misleading use of statistical significance testing at several occasions in the paper. Statistical significance testing provides information about the likelihood of e.g. observing a certain trend under the assumption of the null hypothesis (the trend is not real and just the result of the sampling). It does not provide any evidence about whether the trend is real. A large p-value simply implies that the data is consistent with the null hypothesis and there is no evidence for a trend. It does also not rule out that there is a trend, but given the data we cannot tell. A low p-value in turn does not indicate a high likelihood for the hypothesis to be true, it just tells us that the data is unlikely to be observed under the assumption that the null hypothesis is true (for a thorough discussion see Ambaum 2010: Significance test in climate science, J. Climate, 23, 5927 – 5932).

   Specifically in this paper the following misleading use of significance testing occurs:

   - P8 L6: "…the group of the Siberian pathway does not provide a significant trend…" This is a misleading statement, as it suggests that we should trust the weak trend in the frequency of the Siberian pathway less than the other trends, which we should trust because they are significant. This is of course absurd because the sum of the trends needs to balance. Hence, if we trust the trends of the other two patterns, we have to trust the weak trend of the third by as much. The significance test is therefore not helpful.

It would be more insightful to provide confidence intervals that illustrate the robustness of the trends to rule out that the trend is strongly influenced by a few data points.

- P9 L11: “For the regions that correspond to lower temperatures with an increased occurrence of the North Atlantic Pathway and a decreased occurrence of the North Pacific Pathway no significant temperature trend can be found on the left panel of Fig. 7. This suggests that the temperature anomalies due to the transport changes are counteracted by other processes.”

  The reasoning is wrong here. It could well be that there is a cooling trend at these locations, but the trend is hidden because of one or two much warmer winters, which may be outliers. The statistical significance test does not provide us any information either way. A Monte Carlo resampling assessing the robustness of the local trends would give more insight.

- In Fig. 7 (left panel): The significance test here does again not provide any information about whether the trends in some regions are robust. Again confidence intervals would be more insightful. Furthermore, in multiple testing scenarios, if any significance test is done at all, a field significance test should be done to take spatial correlations and erronous rejections of the null hypothesis into account (cf. Ventura et al. 2004: Controlling the Proportion of Falsely Rejected Hypotheses when Conducting Multiple Tests with Climatological Data, J. Climate, 17, 4343 - 4356)

4. What is the reason for not considering moisture fluxes as well? These are arguably highly important for Arctic heat anomalies because of their impact on the radiation balance. See for example Woods et al. (2013) and Messori et al. (2018; cited in the paper).

   Woods, C., R. Caballero, and G. Svensson (2013), Large-scale circulation associated with moisture intrusions into the Arctic during winter, Geophys. Res. Lett.,40, 4717–4721, doi:10.1002/grl.50912.

5. For displaying the differences between individual clusters more clearly, it could help to show heat transport anomalies instead of the full transports. In many patterns the differences in the heat transport are rather nuanced and hard to see.

6. How large is the within cluster variance for the SOM clusters and the three main clusters?

7. P5 L10ff: T2m anomalies: Why don’t you consider vertically averaged / integrated (potential) temperature anomalies? These would be more clearly related to the heat flux divergence than T2m anomalies, which are strongly influenced by surface heat fluxes. This is especially true in regions with a rapidly declining wintertime sea ice cover

(Barents and Kara Seas), where the temperature trends are to a large extent due to surface heat fluxes.

8. Related to the above I would be interested in seeing the divergence of the heat flux. Warming at a certain location will be more related to the heat flux divergence than the flux itself.

9. P5 L10ff: How are anomalies computed? Are they taken from the period (DJF 1979 – 2016) mean or is a running mean used to account for intra-seasonal variations?

10. Fig. 4: There seems to be a large compensation between the poleward and the equatorward transports, which I find surprising, especially concerning the strong southward heat transport at -120°E, which must be associated with very cold air (with low heat content).

    And does the standard deviation depict the inter-annual variability? That is, is it computed from the means of each winter? Or is it the standard deviation computed from daily data?

11. Fig. 7 (right panel): The caption should state that for the North Pacific pathway the inverse of the temperature anomaly was taken (as described in the text).

12. P1 L16: To first order the much stronger warming in the Arctic compared to lower latitudes is caused by the loss of sea ice, exposing major areas of the Arctic ocean to the atmosphere, leading to subsequent warming of the lower troposphere, and not the other way round. Additional melting of sea ice because of Arctic amplification would require additional transport of heat into the Arctic.

13. P2 L1: "To summarize, there is a clear indication that Arctic Amplification alters the circulation and heat transport patterns in the Arctic." I'd suggest to tone this statement down a bit. The causality is not fully clear in my view. See also Screen et al. 2018: Consistency and discrepancy in the atmospheric response to Arctic sea-ice loss across climate models, Nat. Geosci., 11, 155 - 164

14. P2 L6: "… that have been emerged …" → "… that have emerged …"
15. P2 L9: pattern should be patterns
16. P2 L9: either high or strong, not both
17. P2 L26: reanalyses → reanalysis
18. P2 L28: Please rephrase "This is used to obtain informations from the whole tropospheric column."
19. P3 L11: "… an average picture …" → the heat transport throughout the entire troposphere (?)
20. P4 L4: Fig. → Figs.
21. P4L6: This is likely an artefact from the vertical averaging

22. P5 L1: … are directed → … is directed …
23. P5 L4: zonally → zonal
24. P5 L7: Awkward formulation " … with two cyclone motions …", please rephrase
25. P5 L9: "… an anti-cyclone motion…" dito
26. P6 L12: Why focus on 75°N when SOM clusters are grouped together according to the heat flux at 80°N? Generally, I think 75°N is better suited because 80°N lies largely in the interior Arctic (except for the European sector).
27. P6 L33: Since you integrate H vertically, you could simply state that you consider the meridional component of the heat flux Eq. (1).
28. P7 L4: Remove "Generally, the meridional transports of the three groups fit well to the described pathways." - of course they have to be consistent as you look at the same quantity (the vertically integrated heat transport).
29. P8 L2: occur → occurs
30. P9 L14: remove "can not"
31. P10 L6: favors → favor
32. P10 L20: measurement → measurements
33. P10 L25: Awkward phrasing, please rephrase.
34. P10 L28: the presented work here → the work presented here
35. P10 L31: a increase → an increase
36. P11 L3: generally → general
37. P11 L13: that at region → that in regions
38. P11 L16: changing of → changes in
39. P11 L20ff: awkward phrasing until and including "… whole picture."
40. P11L27: an guide → a guide