Review of manuscript 2021-020 submitted to Weather and Climate Dynamics

Interactive 3-D visual analysis of ERA 5 data: improving diagnostic indices for marine cold air outbreaks

by Meyer et al.

This study investigates the usefulness of conventional marine cold air outbreak (MCAO) indices, which are typically based on the sea-air potential temperature difference, to identify regions where polar lows (PLs) develop and seeks to improve them in terms of their skill in predicting favourable regions for PL occurrence. For that purpose, the study advocates the use of 3D visualizations of MCAOs and PLs, informing the formulation of alternative indices. Based on case study analyses, two indices are then presented that perform particularly well in terms of predicting PL occurrence as demonstrated for a set of PLs from the STARS database and randomly chosen reference samples. The first index measures the depth of the unstable MCAO air mass by identifying the uppermost convectively unstable level, the MCAO top height, thereby following other studies (e.g., Terpstra et al. 2021). The second index also takes the role of upper-level forcing into account by considering the distance of the MCAO top height and the tropopause, which to my knowledge is novel.

While the the presentation of figures is of high quality and, especially the 3D ones, are simply beautiful to look at, the writing of the manuscript and, in particular, also its organization need improvement. Also I am a bit unsure about novelty and scope of the study, as detailed below. Despite of that I think this study has the potential to become a valuable contribution to WCD if the manuscript is appropriately revised.

General comments:

- One of the difficulties I have with this manuscript is whether it fits into the scope of WCD. On one hand, it introduces a workflow for 3D visualization of meteorological fields using the Met.3D software. While I have no doubt that 3D exploration of meteorological phenomena is highly useful and inspiring for formulating hypotheses, introducing such a workflow does not fall into the scope of WCD and other journals might be more appropriate (e.g., Geoscientific Model Development). On the other hand, the specific application, namely identifying better indices for capturing regions favourable for PL development, definitely falls into the scope of WCD. The way it is presented now, the paper tends towards both directions. I strongly suggest the authors focus the paper more towards the meteorological aspectes and less towards a 3D visualization workflow to make it more appropriate for WCD (of course keep the beatiful figures!).

- I also have some reservations regarding the novelty of the presented work. Other studies (see Papritz and Spengler 2017 and Terpstra et al. 2021) have considered the depth of the unstable layer for identifying MCAOs before. In particular, Terpstra et al. (2021) also explore its usefulness for predicting PL occurrence. In addition, while the second index, which also takes the tropopause level into account, is – to my knowledge – novel, it bears some resemblance to the criterion developed by Kolstad (2011). What are the differences and advantages of your index? Why another index?

- The manuscript is partly not very clearly organized, this applies especially to the methods section, which at great length explains the workflow including many technical and often irrelevant details (such as the supercomputer used) and unnecessary statements à la "We have implemented a functionality in Met.3D…" but omits important information about the methodology. Some unclear points are:
  - Definition of the MCAO indices: these indices should be defined in the methodology section and not in the main part of the paper.
  - Definition of "pseudo-events": How are the spatial regions chosen?
  - For assessing the skill of the MCAO indices it is critical that in the reference sample no PLs occurred. How did you check that during the random time steps there are really no PLs? As I understand the STARS database contains only selected PL cases in the region but by far not all. For example, Michel et al. (2018) find about 243 PLs per winter season, compared to only about 13 PLs per winter in STARS.
  - How are the 95th percentiles computed? Are they spatial or temporal? If they are spatial, what regions are considered?

- The new MCAO index is evidently better suited for identifying regions favourable for PL development than conventional MCAO indices. However, it also has a clear disadvantage: it does no longer take the coldness of the air mass relative to the sea surface into account. Thus, it looses the useful property of conventional indices that they are directly proportional to the sea-air sensible heat flux (cf. Papritz et al. 2015). Hence, it is a matter of application which index is better suited. I would wish the authors include a discussion of this aspect.

Specific comments:

- L57: please explain the term potential skin temperature, this may be unfamiliar to non-experts
- L64ff, discussion of different MCAO indices: I miss two important aspects here. First of all, CAOs are convectively unstable air masses subject to strong vertical mixing. This becomes also quite evident in the selected profiels that you show in Fig 4, where the variation of potential temperature with height is small within the CAO air compared to the tropospheric lapse-rate further aloft. Accordingly, the potential temperature is fairly uniform in the vertical, strongly reducing the sensitivity to the choice of characteristic pressure level. Second, there is a clear preference for choosing a lower level because CAO air masses tend to be quite shallow initially and their depth only grows with fetch from the ice edge due to convective overturning and the associated entrainment of air from aloft. Hence, CAOs remain undected near the ice edge if a high level is chosen.
- L111: What do you mean by the "climatological analysis"?
- L112: Please clarify which MCAO index you are refering to, i.e., indicate the level.

- L148: What do you mean by rasterize?
- L165ff: Please properly define the MCAO indices m_i
- Section 3.1: I really like the 3D figures! However, I believe the benefit of these does not yet come out clearly in the text. What specific features can only be made visible in 3D? I am sure there are, but the text remains vague about this. Also in principle, the "topography" of the isentropic surface used to delimit the MCAO air can easily be visualized in 2D maps. Please be specific about what can be gained from the 3D analysis that is not possible with more conventional charts?
- L303: Definition of new MCAO index: I don't understand the rationale behind choosing a fixed surface pressure p_0. You state that the new index "measures the vertical extent". In that case you should choose the actual surface pressure instead of a fixed value (see also Terpstra et al. 2021). This will make a big difference in regions where the MCAO air-mass is still shallow (i.e., near the ice edge).
- L330ff: Please clarify how you define the areas inside and outside the PLs. I understand from what has been said earlier that the area inside is the area within the PLs' radius. How then is the area outside of PLs delimted?
- L350: How are the pseudo-events defined spatially? Are you using the same areas as for the actual events?
- L366: 95th percentile with respect to space or time?
- L381: Are these the same pseudo-events as the ones before and why then do you explain their definition again? Or are they different? If so, I don't see how come. Please clarify.
- Figure 7: What is the step-size for the critical thresholds used?
- L457: Honestly, I am lost in this paragraph. I understand the authors want to find an alternative metric in order to avoid vertical interpolation to find the critical level where potential skin temperature equals potential temperature. The description of the methodology how to obtain this alternative index is more than confusing. I strongly suggest the authors rewrite this section entirely for better clarity. Furthermore, I believe that also some evidence needs to be shown that this approach is equivalent to the previous.

Technical corrections:

L9: remove comma after that

L22: … representing conditions conducive for …

L62: here and throughout the manuscript: please fix references by removing parentheses around year, should be (REF1 YYYY1; REF2 YYYY2…)

L84: remove full stop

L147: Barents and Nordic Seas

L157: than → as

L188: here and elsewhere: section should not be capitalized if it is not refering to a specific section

L202: here and elsewhere: add a comma before "e.g."

L263: remove sea

L286: Potential Vorticity **Unit**

L326: please remove parentheses around refs
L367: approximately
L384: Pseudo-events
L416: … has an area under the ROC curve (AUC score) …
L474: ROC analysis
L521: Terpstra et al. (2020) suggest
L551: front shear → forward shear
L560: (EDI; Wulff and Domeisen, 2019)

Caption Fig. 1: No need to repeat the detailed steps (1-3) here since you do that in the main text.

References:

Terpstra, A., Renfrew, I. A., & Sergeev, D. E. (2021). Characteristics of Cold-Air Outbreak Events and Associated Polar Mesoscale Cyclogenesis over the North Atlantic Region, *Journal of Climate*, *34*(11), 4567-4584