Comment on wcd-2021-27

Anonymous Referee #2

Referee comment on "Automated detection and classification of synoptic-scale fronts from atmospheric data grids" by Stefan Niebler et al., Weather Clim. Dynam. Discuss., https://doi.org/10.5194/wcd-2021-27-RC2, 2021

Review of *Automated detection and classification of synoptic scale fronts from atmospheric data grids*

Automated feature recognition has proven useful in gaining scientific knowledge of the dynamics and relationships between various atmospheric flow features such as cyclones, jets, and surface fronts. However, there are a variety of automated methods to identify the relevant feature of interest because even trained experts do not agree on how to define a feature. This applies to surface fronts for which no general definition exists. Therefore, improved methods that help to gain additional insight into the nature of fronts are important, however I am have some concerns on whether ML-based method trained on surface analysis is the next step.

General comment
1. There is no single accepted front definition and different weather centers use their own definitions based partly on physical considerations, partly on training and experience, partly on the specific local meteorology, and sometimes simply artistic. It is therefore questionable whether a front identification should be guided by manual surface maps or physical arguments. This dilemma is nicely summarized in Uccellini et al (1992) and Sanders (1999) and was lately reviewed in Schemm et al. (2018) and Thomas and Schultz (2019a). I recommend that the authors review these earlier studies; their introduction comes in its first paragraph without a single reference (and there are numerous studies that link fronts to extreme weather that could be referenced). Also, there is little historic background provided.

The front dilemma can be summarized with the following example: The UK MetOffice automated surface analysis regularly displays double fronts, while the DWD chart never shows these fronts – see Fig. 2 in Thomas and Schultz (2019a). Instead, DWD-front are Norwegian-like and hemispheric spanning, which is more art than science. The missing double fronts are however real and important to detect. They will be missed if trained on DWD charts.

Related to the definition of fronts, there is one stream of front definitions that is based on baroclinic instability and there is also a second front definition based on air-mass boundaries – see Thomas and Schultz (2019b). The two are mixed up in this study, for example, when the authors speak about fronts that are associated with the propagation of extratropical cyclones but thereafter describe fronts as air-mass boundaries. These air-mass boundaries, which provide little baroclinicity, are very interesting for research. If one wants to detect these, one cannot train a ML method on DWD charts, although it seems as if DWD uses many meteorological variables to draw their front, which is common for the air-mass boundary definition of fronts but not that based on baroclinicity. Maybe DWD excludes mesoscale fronts in general?

Overall, I therefore reject manual surface charts as ground truth, baseline method or "gold standard" for verification. The surface maps are biased, inhomogeneous, only partly based on physical reasoning and cannot be transferred between different regions. I find a tool that learns these biases, here the DWD bias, difficult to use for research purposes, though they might be useful in an automated DWD workplace environment. Even though the authors try to alleviate some of these issues with the blurring of the front position shown in their Fig. 5, I hesitate to conlcude that ML-based fronts trained on surface charts is the way forward.

2. The manuscript has a strong technical nature with only little insight into meteorology or front dynamics. I would recommend considering a transfer of this manuscript to GMD.

3. The presented comparison against a second front detection method, which is based on the thermal front parameter (TFP), is odd. First, I recommend not to call it "the ETH method", because this is not known to the community and ETH is a large institution. Maybe TFP method? Second, I recommend providing more background about the TFP, which goes back to Renard and Clarke (1965). The TFP implementation by Jenkner et al (2010), which is used here as reference method, is unique, because it places the front where TFP=0, which is at the center of the frontal zone. However, this is not where most meteorologist place the front. Most center, including DWD and ECMWF, place it where MAX(TFP)=0, which is at the leading edge of the frontal zone. So, there is a mismatch. This important difference is not explained in the current manuscript and because the width of a frontal zone can easily encompass a couple of hundred kilometers, the here used "baseline method" will be due to its design in most situations do not agree with the DWD charts. Basically, a method that was trained to reproduce DWD fronts, which it does very well, is compared against a method that was intentionally designed not to agree with DWD fronts because the front line is placed in a different location. It is thus not a meaningful comparison (and this explains much of what is found in Lines 376-4079) and I recommend that the comparison is removed.

4. More meaningful would be a comparison against another ML-based method, for example, that of Lagerquist et al. (2019), who pioneered ML-based front detection. This would be more insightful because it is not clear at this point which neural-network architecture is most suitable for front detection and why this is the case. I find Fig. 10 in Lagerquist et al. (2019) very helpful. A similar figure plus a direct comparison of these two ML-based methods would thus be of high merit.

5. It is not advisable to transfer an automated front detection from a low-resolution grid to a high-resolution grid without intensive retuning and testing. How was this retuning done? By how much was the detection thresholds for the front gradient increased? By how much was the minimum length criterion changed? Did the authors increase the minimum advection speed to separate stationary from non-stationary fronts? A method developed for a ERA-Interim 1x1 degree grid (or for a 2-km grid as in Jenkner et al. 2010) should not be transferred to another grid spacing. Further, does DWD use a minimum front length and are the authors using the same threshold? At the same time, while the front threshold presumably was increased when going from a 1x1 degree grid to a 0.5x0.5-degree grid, the number of fronts ideally should not change (see Fig. 2 in Thomas and Schultz (2019b) on the dependency to the threshold). More details on the retuning that was done when preparing the comparison is needed in this manuscript.

6. I was disappointed to see an equivalent-potential temperature based front definition purposefully applied to latitudes outside of midlatitudes. Front detection methods based on equivalent-potential temperature (called TH in the next statement) are well known to be unsafe for usage outside of the midlatitudes, for example, Schemm et al. (2015, p. 1696) noted: "… *clearly indicates that the TH method is influenced by semi-permanent convergence zones and tropical convection (although a minimum advection threshold is applied). Tropical features which, from a synoptic viewpoint, would not be regarded as a 'front', are identified as such. Accordingly, TH methods should be used with care if applied outside midlatitudes*", which is a nice way to say that it should not be done. Further they note "… *as the θe gradient can be dominated solely by moisture gradients, especially in tropical latitudes, this results in the detection of several quasi-stationary fronts (which form along mountain crests, or in association with land – sea contrasts) which must be removed in a post-processing step*" (Schemm et al. 2015, p.1687). Against these recommendations the authors decided to apply the θe method to subtropical and tropical latitudes and afterward, not too surprisingly, conclude that it detects numerous of non-cyclone related fronts. What was the intention behind this? The section between L.460-470 is therefore misleading.

7. The conclusion is short, with only a technical statement and an outlook but no conclusion related to weather and climate dynamics. Maybe you could try to conclude on how and why the ML-based method is able to distinguish mobile from stationary fronts (such as those along the coastlines or mountains), which would yield additional process understanding and it is a mayor struggle for traditional TFP-based methods.

Minor comments:
• L. 19 "much of the literature is on the larger-scale fronts" – research on mesoscale fronts is a very active field of research as well.
• L. 15 "are a vital part of the communication of weather to the public and the public perception of weather in general" – Most people use Apps; fronts are no longer a major part of modern weather communication.
• L. 27: "The former methodology goes back to the work by Hewson (1998)" – I guess it goes back to Renard and Clarke (1965).

References:
• Lagerquist, Ryan, Allen, John T., and McGovern, Amy, 2020, "Climatology and

Variability of Warm and Cold Fronts over North America from 1979 to 2018" Journal of Climate Vol. 33, No. 15, 1520-0442

• Lagerquist, R., A. McGovern, and D. Gagne II, 2019: Deep learning for spatially explicit prediction of synoptic-scale fronts. Wea. Forecasting, 34, 1137–1160.

• Sanders, F., 1999: A proposed method of surface map analysis. Mon. Wea. Rev., 127, 945–955

• Schemm, S., Sprenger, M., & Wernli, H. (2018). When during Their Life Cycle Are Extratropical Cyclones Attended by Fronts?, Bulletin of the American Meteorological Society, 99(1), 149-165.

• Thomas, Carl M. and Schultz, David M., 2019, "Global Climatologies of Fronts, Airmass Boundaries, and Airstream Boundaries: Why the Definition of "Front" Matters" Monthly Weather Review Vol. 147, No. 2, pp 691, 1520-0493

• Thomas, Carl M. and Schultz, David M., 2019, "What are the Best Thermodynamic Quantity and Function to Define a Front in Gridded Model Output?" Bulletin of the American Meteorological Society Vol. 100, No. 5, pp 873, 1520-0

• Uccellini, L. W., S. F. Corfidi, N. W. Junker, P. J. Kocin, and D. A. Olson, 1992: Report on the surface analysis workshop at the National Meteorological Center 25–28 March 1991. Bull. Amer. Meteor. Soc., 73, 459–471.