Comment on wcd-2021-27

Peter Düben (Referee)

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-RC1, 2021

This is a good, well-written paper that should be of interest to the readership. However, I have a couple of minor comments that could be addressed in a revised version of the paper.

Comments:

The paper is longer than it needs to be, and some information is spread over the paper which makes it difficult to extract the relevant pieces. E.g. when introducing the vertical levels in l.117 and then mentioning that you only use 9 pressure levels in l.196. Why not describe the data-set augmentation with the data in section 2.1?

Section 3.1 and 3.2: I have tried for a while to understand why you present results for the validation AND the test dataset and gave up. Why do you need section 3.1? You are writing in l.300: "We validated our model during training using 1460 samples of data from 2017. We evaluated our trained models on 1 year of data from 2016 using an object based evaluation described as described later in this section." This does not really explain why you need the two sections. Also, section 3.1 starts with "The trained models were evaluated on test sets…" which generates ultimate confusion. Do you loose any information when removing section 3.1? Maybe I am missing something.

l.193: I do not understand this. You say you "ignore the outer 20 pixel". But then you are saying that the brighter areas can be used as input in the caption of Figure 4. Are they used as input but not predicted? But then the output domain should be smaller than the input domain in Figure 3…? And why do you crop to 128x256 pixel (l.199)? And then there is again a confusing mentioning of the 5 degree border in the caption of Table 2…

l.8: I would not call the baseline model "ETH". ETH is a very large institution.

l.21: Maybe add a reference to the Mei-Yu front?

l.22: "Determining the position and propagation of surface fronts plays an important role for weather forecasting". Well, the prediction of the position, yes. But is the same true for the automatic detection? Fronts can easily be identified in field maps by the trained eye.

Why do we need the ability to detect them automatically with ML? I do understand why, but it would be good if this would be made more explicit in the intro, otherwise it seems that you have a hammer and are searching for nails.

l.24: What are empirical guidelines?

Section 2.1: Maybe I missed it, but do you actually state the resolution of the NWS and DWD datasets somewhere (or the resolution equivalent of the PNG image)?

Figure 3: I do not understand the encode and decode blocks. Can you add some info here? Also, what are the white boxes the "copy" arrows end in?

l.198: "If both labels are available". What does this mean? At a certain point in time? Why should this matter?

Table 2: The whole caption should be reformulated. "For the global region this border is included within the mentioned range." ?

l.242: This paragraph is important but very difficult to understand. It should be rewritten.

l.279: I would not use "t" for the index of the channels as "t" is often used for time.

l.280-282: I do not understand this. "individually for each batch"? "more emphasize onto classification"? Either equation (2) holds, or not.

l.289: Why did you not evaluate the baseline at 0.25 degree? I guess there are good reasons, but please state them.

Table 3: You can as well remove the "Stationary" line.

Table 5: "The suffix "all"…" I do not understand this sentence.

l.488: I find this a bit confusing. You would not leave out a certain region in a real-world application, so why here?

Typos etc:

l.51: typically

l.253: predicted fronts

l.301: remove "described"

l.346: "be be"

l.349: "slight edge"?

l.351: "fact that training"

l.403: "most likely"

l.445: "and the European data"

Caption Figure 7: "on the for the"

l.514: "for is the lack"

l.439: "However,"