Review of WCD-2021-25

**Title**: Multi-day hail clusters and isolated hail days in Switzerland – large-scale flow conditions and precursors

**Authors**: Hélène Barras, Olivia Martius, Luca Nisi, Katharina Schroeer, Alessandro Hering, and Urs Germann

**Recommendation**: Major revision

**Summary**: The study is examining the synoptic-scale atmospheric patterns during and ahead of series of successive hail days (clusters) and isolated (one-off) hail days over Switzerland. Hail days are characterized based on a radar reflectivity-based proxy (Probability of Hail) in combination with vehicle damage data from an insurance company. The key findings of this ~18-year climatological study is that clusters preferentially occur in more amplified wave patterns with higher atmospheric instability, while isolated hail days occur in a more zonal flow pattern with less instability and stronger dynamical forcing.

**General Comments**: The study supports more rigorously existing notions that amplified and slowly progressing synoptic wave pattern allow for the building of a more significant warm sector and multi-day convection before a frontal passage drives the instability south out of the domain of interest. It combines a diverse range of tools to more objectively identify atmospheric blocks, synoptic patterns and lower-level fronts. The classification into hail days has been thorough and rigorous, with a valuable and often elusive insurance loss dataset strengthening the hail days classification scheme.

The conclusions of the study are not overly controversial, but could be made to be more insightful. In my view there is a missing connection between the larger-scale patterns and the actual storm environments that promote damaging hailstorms. The formation of those hailstorms that results in insurance claims often requires more than atmospheric instability (or some form of CAPE). The growth of larger hail (golf balls and above) may also be associated with dynamically enhanced storms where the vertical shear of the ambient horizontal flow is a key predictor. While the study comments on instability, vertical shear is not derived methodically from the reanalysis data. Of potentially significant advantage would be to reference more local characterizations of the near-storm environment in an often Alpine setting to convincingly collect all those ingredients that subsequently should be extracted from the reanalysis. Have case studies of damaging hailstorms been conducted in the Swiss Alps? What key ingredients have been identified across such studies? Has deep-layer shear been present in the more notable hailstorms, or did damaging hail form through a pulse storm mechanism alone?
A second general point worth addressing is embodied in the significant number of minor comments below. A commonality between a good number of these comments is the language of the manuscript which requires a bit more work to remove pockets of ambiguity in the expressions used (e.g., an easier description of the hail day definition employed), and missing information in a number of places so that a reader can clearly understand the argument presented or simply needs to work less hard to select between possible interpretations or clear up ambiguity. Each individual point is an easy piece of work to clarify, but in aggregate this amounts to a major comment.

In my view, the manuscript is well on its way to publication and would sit nicely in the domain where topography and severe convection interact.

**Minor Comments:**

1. p.1, l.18: The word *moreover* occurs twice
2. p.2, l.28: second Púčik citation has incorrect format
3. p.2, par.1: recommend to replace *weather situations* by *weather patterns* or *synoptic patterns*
4. p.3, l.64: Waldvogel’s dataset was compiled during or before 1979 with a radar or radars now over 40 years old. Is there a risk that the Waldvogel dataset (and therefore the POH polynomial fit) could no longer be representative of today’s radar retrievals of the 45 dBZ height due to narrower radar beam widths, improved radar calibration practices, or other factors that have changed over the 40-year period?
5. p. 3, l.64: How is the freezing level obtained for the POH calculations? Is it an operational mesoscale model, or interpolation between sounding sites? Could there be quality issues with such freezing level estimates over complex mountainous terrain?
6. p.4, l.91: What parcel is being launched for the CAPE calculation – surface, mixed layer, most unstable?
7. P.4: I am a bit unclear on what is computed off ERA-Interim, and what is computed off ERA-5 (it seems 2 out of 3 schemes run off ERA-I, one off ERA-5, and all stand-alone fields are sourced from ERA-5)
8. P.5, l.100: Presumably the POH grids are available for every radar scan during a day. Is the daily hail footprint the area (on the 1 m x 1km grid) where POH ≥ 80% for at least one radar scan during that 24-hour period?
9. P.5, first paragraph: given the daily POH footprint as well as the hail day thresholds both use "80", the paragraph confused me at first. Is the 80\textsuperscript{th} percentile taken from the daily POH distribution (excluding footprints of size zero) for the 10-yr period 2003-2012 (i.e., matching the period of the available insurance loss data)?
10. P.5, l.118: Isolated hail days appear to require three non-hail days on either side, but not a single hail day in any 5-day window (a counter example are the two isolated southern hail days on days 176 and 180 in 2005 that are separated by only 3 days).

11. P.5, l.123: total hail day numbers do not match the 18-year annual averages given in ll. 113-114. Could this be due to seasonal bounds imposed?

12. P.5, ll128-130: I am confused about how the independent cluster creation works.

13. P.8, Fig.3: I recommend presenting the annual hail climatology by calendar month in some form as most readers would better related to a discussion/presentation that refers to hail frequency peaks in July rather than days 180-199 or equivalent. I understand there are some minor inconsistencies in the mapping from DOY to months around the non-uniform duration of months (28-21 days) and the issue around the four leap years in the dataset. Perhaps a secondary x-axis for non-leap years is one option?

14. P.8, ll.166-167: This paragraph needs more explanation around the Weusthoff Weather Type Classification for interpretability. For example, is the weather type wind direction referring to the 500 hPa flow, the surface flow, some layer-mean wind direction, or other? Is there a geographical weighting on Central Europe, or all of Europe, or some other sub-region? The paragraph also needs more explanation of how the weather type across all isolated/clustered hail days was derived. Was the WTC first applied on a day-by-day bases, and then composited in some way?

15. P.8, l.170: ‘twice as frequent’ when comparing ~28% to ~19% is a tad generous, how about ‘significantly higher’?

16. P.10,ll.191-193: The sentence switches from dealing with the clustered day pattern to the isolated day pattern mid-sentence. It would be clearer to separate the two regimes through the structure in the text to avoid confusion.

17. P.10, l.197: Plot 5(b) does not show SSW 850 hPa flow over Switzerland given the flow magnitude is less than 4 m s⁻¹; could this be a case of ‘not shown’?

18. P.10, l.201: First than deserves deletion

19. P.11: The figure caption comes across as a tad organic. I would rewrite the caption by stepping through the individual panels (a), (b), ... and explain what is shaded and what is contoured. There is no need to mention against each individual panel that the sample processes consist of 135 clustered and 69 isolated hail days (mention that once only at the start or end).

20. P.12: The figure caption would benefit from a rewrite akin to the caption of Fig. 11. Details about the significance testing methodology don’t need to be repeated in the caption. The CAPE field shown requires some clarification: I assume you are showing either SBCAPE or MUCAPE as a field, and given the XXCAPE value I shown for a hail day, similar to the 2-m temperature, and I assume panels (b) and (d) show the daily maximum XXCAPE value? The frequency of fronts being present is ambiguous and requires some more detail. If a front is detected by the Schemm scheme for 1 (2, or 3) of the 6 hour outputs from ERA-I, does the hail day count as a ‘front day?’ Panel f lacks explanations of the dashed orange and blue solid contours.
21. P.13, ll.218-219: A reference that 250 hPa diffluence sustains amplification of an upstream trough would be good here.

22. P.13, l.219: What does it mean that the rough widens zonally? Given we are looking at a composite, this could mean that the meridional trough placement in the composite members becomes more diverse. Alternatively, it could mean that the trough is getting wider across the members.

23. P.13 section 4.3.2: The section is not a strong contribution to the paper, apart from the gradual amplification of the pattern ahead of both types of hail days. A lot of the statement in this section are isolated single sentences that are not linked to a meaningful wider storyline. How do these isolated statements fit into a larger view of the lead-up to the cluster and isolated hail days?

24. P.17,18: I recommend running a smoother over the significance testing sample fraction contours as these are very heard to read and interpret. Similar comments would apply to Fig. 5, 6.

25. P.22, ll. 279-281: Sentence should start with an expression that indicates you have switched to joint cluster days.

26. P.23: Figure caption – 'fronts' is the proportion of the southern cluster (joint cluster) days with a diagnosed front?

27. P.24, l.289: Better to use 'synoptic pattern' rather than 'situation'.

28. P.24, l.299: This phrase seems to contradict the previous one where downstream anticyclonic Rossby wave breaking was postulated to result in weak upper-level and surface zonal flow over central and eastern Europe. In Fig. 5 the block also seems to be sitting north of the UK and west of Scandinavia. These two sentences need a small amount of reconciliation.

29. P.24,ll.300-304: These lines are quite speculative keeping in mind that the frontal analysis is based on a reanalysis and is composited. There could be a range of initiation mechanisms at work on cluster days that would require a different data set to resolve. Are 'thermo-topographic winds' related to upslope flow forced by elevated solar ground heating?

30. P.24, ll.307-308: There is an argument pointing in the opposite direction – more meridional patterns with weaker flow and reduced horizontal pressure gradients are more difficult to predict for numerical weather prediction models. I am guessing that you are saying that slowly evolving patterns are easier to predict through persistence arguments?

31. P.24, l.311-312: This argument is not as strong as it first sounds. Moist boundary layer air can be advected over distances much larger than 600 km and thus any local evapotranspiration to boost the local dewpoints is not required. For example, moisture-hungry supercells in Nebraska over 1000 k from the Gulf of Mexico are routinely supplied by such advected moisture when a lee trough east of the Rocky Mountains drives a stream of moist air northward. Local ET may help, but it is not a necessity based on a distance-to-coast argument.

32. P.25, L.324: Blocks are not obvious in the composite – they might still occur in individual cases.
33. p.26, ll.372-375: The driver for convective initiation can not be derived from the data that have been used in this study and are supplied as an external input based on additional experience or datasets relating to a finer spatial scale.