Review of: “Exploring hail and lightning diagnostics over the Alpine-Adriatic region in a km-scale climate model” by Cui et al.

Recommendation: Minor revisions

Summary

The authors compare 2.2 km COSMO simulations driven by reanalysis data with different observations of total precipitation, hail, and lightning of a number of severe weather events. Overall, the paper is written sufficiently clearly, and I only have a (relatively large number of) somewhat minor comments.

Main comments

1. It is not quite clear what the goal of the paper is: Assess the COSMO diagnostics or to explain why severe convective weather occurred on those days. I have to admit that I didn’t get much out of the short case studies (which are partially redundant with the case summaries earlier in the paper); is there anything new to be learned from these?

2. Throughout the manuscript, there are statements touting the simulations as performing “very well.” I think such statements are subjective and should be avoided unless a metric has been defined and justified, which quantifies the performance of the simulations. I suggest deleting these statements and simply report the differences between the observations and simulations, and let the reader decide to what extent the model diagnostics are useful (or maybe offer your take on the results in the conclusions section without overselling the quality of the simulations).

Specific comments

1. Line 4, and throughout: Since you are still in the km scale, the simulations are *convection allowing*, but not *convection resolving* (to resolve convective flows, a grid spacing of O(100) m is required).

2. Line 8: Write out the COSMO acronym before using it the first time.

3. Line 13: Is the topographic barrier a requirement, or is it just present? This feature does not seem to be highlighted later on, and I’m not sure it has been demonstrated that this barrier was indeed a cause of the severe weather. More analysis/discussion would be needed to grant this feature a place in the abstract.

4. Line 25: Reword thunderstorm peril region (e.g., regions at highest risk of experiencing thunderstorms).

5. Line 44: Can you be more specific about which aspects are not well-understood? Such broad statements tend to dismiss the wealth of research that has been done in this area.

6. Line 47: Wind shear also promotes the organization of quasi-linear systems and updraft intensification (e.g., Markowski and Richardson 2010).

7. Line 49: One would expect a given environment to lead to the same storm, whether or not the storm is surrounded by mountains. Can you specify what you mean by stating that the role of the ingredients is complicated? Do you mean that there is a larger suspected variability of the thermodynamics/kinematic fields compared to flat terrain, or that the environment cannot easily be sampled (because it may be rather inhomogeneous)? Or do you specifically talk about the lift ingredient (the subsequent text seems to imply such)?

8. Line 53: Grammar broken (…and subsequently to…)

9. Line 54: "The" mesoscale boundary hasn't been introduced yet (could say "a" mesoscale convergence zone).

10. Line 102: Isn't 1 h is the interval (rather than a frequency)?

11. Line 103: Add AGL/MSL after 23.5 km?

12. Line 114: Calculated instead of accumulated?

13. Line 123: Are there also non-convective thunderstorms?

14. Line 140: Maybe add that these values are much lower than what is observed in the real world because of the relatively coarse grid spacing (the convective clouds tend to be wider, and updrafts weaker than in truly convection resolving simulations).

15. Line 141: Why is CAPE negative?

16. Line 147: Delete spatial

17. Line 157: Please add the size in cm for each category

18. Line 164: Hailpad observations

19. Line 180: LINET data are not gridded (right?), so they can't really have a resolution. Perhaps say that the location error is 3 km? Why is the time known only to within 2 min? Given that the system has something like nano-second accuracy internally, 2 min seems like a huge inaccuracy.

20. Line 192: Write out the acronym before using the abbreviated form.

21. Line 198: Reword: e.g., …indicates that the model field is too spread-out/broad/diffuse. Also recommend avoiding these parenthetical constructions to shorten the sentence.

22. Line 225: "This cold front" hasn't been introduced yet

23. L. 229: Please specify what was discharged (water?)

24. Line 234: damage (instead of damages)

25. Line 265, 274: Model showing good performance: This seems like an entirely subjective statement. Either define what you consider as "good" or leave that judgement to the reader (and just report the errors).

26. Line 267: Reword: …captured relatively well

27. Line 288: Model reproducing hail "very well": Maybe the general presence of hail is predicted well, but the placement and coverage is (not surprisingly) not captured very well (e.g., 8 July 2017). Like above, I suggest reporting the differences and omit statements about the quality of the simulations.

28. Line 301: … hailstone **diameters** above 20 mm.

29. Line 315: Another instance of "good" model performance.

30. Line 321: Suggest replacing precipitation with rain or "total precipitation" (hail is also "precipitation").

31. Line 325: It is not surprising that the model is unable to produce accurate hail sizes; the updrafts are barely resolved, you are using single-moment microphysics, etc. I would probably consider omitting the hail size comparison, or at least state upfront that the model cannot be expected to produce accurate hail sizes (forecasting hail size accurately remains a holy grail of severe weather prediction).

32. Line 375: If the situation is strongly synoptically-forced, then presumably there is a good amount of flow in the troposphere, typically leading to fast storm motions. Why do the storms become quasi-stationary?

33. Line 444: True, but updrafts also suffer more from entrainment, so the net effect tends to be less evaporative cooling (James and Markowski, 2010, MWR).

Figures

Fig. 4: Consider adding labels to the rows to make it easier to identify which panels refer to the observations, and which to the simulations

Fig. 5: The panels, and especially the legends are too small; consider breaking this figure into two.