

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC1  
<https://doi.org/10.5194/nhess-2021-373-RC1>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## **Comment on nhess-2021-373**

Anonymous Referee #1

---

Referee comment on "How well are hazards associated with derechos reproduced in regional climate simulations?" by Tristan J. Shepherd et al., Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2021-373-RC1>, 2022

---

Review of "How well are hazards associated with derechos reproduced in regional climate simulations?" by Shepherd et al.

This is a model evaluation study to evaluate WRF simulations downscaling to 1.3 km grid spacing with changes of cloud microphysics schemes, lateral boundary conditions (LBC), start date, and nudging. The focus is derecho induced from a mesoscale convective system (MCS). Since derechos cause significant infrastructure damages and economic loss, it is interesting to see if models can capture such extreme events and how the simulations are sensitive to different model setups and physical parameterizations. So, I advocate such studies. However, after looking at the results, I had to doubt whether the simulations were carried out correctly or the simulations were produced from a stable supercomputer/cluster. The model results swift from a convective system simulated with one microphysics scheme to the disappearance of the system with another microphysics scheme is something I never experienced as a senior cloud modeler. Particularly, switching graupel to hail in Morrison scheme also caused the disappearance of the MCS, which is not likely to occur since the change from graupel to hail renders minor changes relative to the entire scheme (mainly in fall speed and density). The hail option is recommended to use for continental deep convective cloud cases by the model developer but it does not simulate the MCS at all. I tested both options in several studies before and this never happened (the simulated convective cloud systems were generally very similar in morphology). In addition, there are so many literature studies with different microphysics schemes for a variety of cases that do not show such a result. As the authors stated this is expected result.

Also, none of the simulations can simulate both derecho and front stages of the observed system, then if the study have focused on why the simulations fail like this, it would still be useful. Furthermore, the sensitivity to two lateral boundary data (ERA5 and ERA-Interim) is also opposite compared with the previous studies (many literature studies showed ERA5 data is improved on ERA-Interim). With all these considered, it is very difficult for me to trust these model simulations thus I would recommend the rejection of the manuscript at this time.

Below I have some specific comments including the appropriate way to calculate the maximum hail size to compare with observations. Hope these comments will be useful for authors to improve the study.

Abstract,

"We also examine the degree to which each ensemble member differs with respect to key mesoscale drivers of convective systems (e.g. convective available potential energy and vertical wind shear) and critical manifestations of deep convection; e.g. vertical velocities, cold pool generation, and how those properties relate to correct characterization of the associated atmospheric hazards (wind gusts and hail)."

-The sentence is near the end of the abstract about it is still about the scientific approach. Suggest changing to phrasing it from the angle of describing your key findings, which is more appropriate for a scientific paper.

Introduction,

- "deep convection disproportionately contributes....", disproportionately does not deliver a good meaning here. Suggest rewording.
- "...and three events caused more than 60% of a utilities' customers power outage; a derecho, an ice storm and a hurricane (Shield, 2021)", not three events, should be three types of events.
- Section 1.2: this section is for case description. There is too much text in describing the societal and economic impacts but lacking a description of the large-scale and mesoscale metrological environments in which the event developed, which is the key information for the study. Also, there should be observational analysis of this event on storm and wind properties which should be discussed to provide a better background of the events.

Section 2,

- Section 2.1: (a) the description of model simulations needs some clarification. I am confused by the descriptions at line 160-165. First you described 3 nested domains were used but then said "A single domain configuration and inner nest grid spacing is used in all members of the ensemble...". Are you using two types of domain settings (3

nest domains and single 1.3 km domain)? If so, please clearly describe which domain setup is used for each simulation listed in Table 2? (b) Nudging simulations are not even mentioned here but they are discussed a lot in the Result section.

- Section 2.2. There are better precipitation data than retrieved precipitation rate using Z-R relationship, which in general has a large uncertainty, such as rain gauge data and Stage IV data from NOAA which combines radar and rain gauge measurements.
- Section 2.3, "Grid cells in d03 are classified as containing 'significant hail' in the WRF simulations if there is > 1 mm of hail and/or graupel accumulation, and in RADAR observations for MESH > 5mm. First, how the simulation and observations can be compared since different calculations of significant severe hail (SSH) are used? Second, What are you based on to define SSH with "> 1 mm of hail and/or graupel accumulation" in the simulations? Accumulation over what time period? In literature there are a few methods to calculate the maximum hail size based on model predicted hail/graupel size distribution such as Thompson (J. A. Milbrandt, M. K. Yau, A Multimoment Bulk Microphysics Parameterization. Part III: Control Simulation of a Hailstorm. J. Atmos. Sci. 63, 3114-3136, 2006) and Snook (N. Snook, et al. Prediction and Ensemble Forecast Verification of Hail in the Supercell Storms of 20 May 2013. Weather Forecasting 31, 811-825,2016) methods. These methods make the model-observation comparison more physically consistent. BTW, "significant hail" should be "significant severe hail" based on the conventional terminology from literature. Another comment on this section is that the metrics description for evaluating models takes too much text and can be tightened up.

### Section 3,

- Figure 4, the simulations with changes of different microphysics only (Goddard, Morrison, Morrison+hail, Thompson, NSSL, Milbrandt-626) totally fails to simulate the convective systems except Morrison captured the linear system. Switching to another microphysics scheme in WRF usually does not make such a large storm system totally disappear (never experienced or saw this in literature). The coupling of those microphysics scheme with WRF should have no problem so it is strange this happened. What is especially suspicious is when switching the graupel option to hail in Morrison scheme, the large linear MCS system disappeared. This would not be possible in a few days of forecasting simulations since slight differences in microphysics should not have such a huge upscale effect on mesoscale systems in a few days.

Furthermore, based on Figures 4-7, none of the simulations can simulate both derecho and front stages of the observed system, then why the simulations fail like this should be the first priority of the study.

- Figure 8, why not plot the radar measurements?
- Figures 9 -11, because the observed MCS was not captured in most of the simulations, one can see the simulated wind speeds are also way off. I do not see a point to intercompare convective winds, cold pools, or other storm related properties since the mesoscale system does not even exist in the model simulations or the basic mesoscale storm is totally wrong as shown in Figures 4-7. Therefore, I did not review Section 3.2

“Linking fidelity to metrics of CAPE, downbursts, and cold pool generation”.

- Since the sensitivity to two lateral boundary data (ERA5 and ERA-Interim) is opposite to previous studies (many studies showed ERA5 data is improved on ERA-Interim), the authors need to evaluate both datasets with observations to show this is an exception that ERA5 performs worse than ERA-Interim. Otherwise, this will make people doubt if the simulations are done correctly.