

Review of the paper "Can models robustly represent aerosol–convection interactions if their cloud microphysics is uncertain?", authored by B. White, E. Gryspeerd, P. Stier, H. Morrison, and G. Thompson.

The authors have drawn a pretty grim picture of the situation with the description of microphysics in cloud resolving models using bulk parameterization schemes. Two known bulk parameterization schemes referred to as "MORR" (Morrison and Milbrandt (2011)) and "THOM" (Thompson et al. (2004, 2008)) were tested by simulation of 3 case studies with different type of convection: from shallow convection to a supercell storm. The simulations were performed at a priori given droplet concentrations of 100 cm<sup>-3</sup>, 250 cm<sup>-3</sup> and 2500 cm<sup>-3</sup>.

A dramatic difference between parameters of simulated clouds (cloudiness, simulated precipitation, etc.) and observations is demonstrated. A huge difference in cloud microphysical structure simulated by these two schemes is reported.

Both bulk schemes turned out to be insensitive to droplet concentration. So, the difference in results are related to the differences in the bulk-parameterization schemes.

My particular comments and remarks are the following.

1. The finding about high diversity of cloud microstructure and precipitation produced by different bulk schemes is not new. Any intercomparison study shows such diversity. The insensitivity of most bulk schemes to aerosols (and to droplet concentration) also well known. For instance, application of different bulk -parameterization schemes to simulation of hurricane Irene (2011), including "MORR" and "THOM", led to a TC with maximum wind varying from 30 m/s to 70 m/s (Khain et al., 2015; Khain et al. 2016-Atmospheric Research 167 (2016) 129–145). In these studies insensitivity of bulk schemes to aerosols was also demonstrated and an possible explanation of such insensitivity is presented .

2. The finding that description of the autoconversion is of crucial importance for correct simulation of cloud microphysics is also not new. For instance, Gilmore and Straka, (2008) showed that the most formulae for autoconversion used in bulk schemes are applicable to the initial stage of the first raindrop formation only, and that the rates predicted by these formulae differ by orders of magnitude. There are many other studies showed dramatic sensitivity of cloud microstructure to the scheme of autoconversion.

So, the new information in the paper is that these conclusions the authors confirmed in investigation of these two particular bulk schemes.

3. The authors illustrate vertical profiles of mass contents of different hydrometeors averaged over the entire computational area. As a result, all information concerning the microphysical structure of clouds simulated by different schemes turns out to be lost, at least for specialists in cloud microphysics. It is necessary to present vertical profiles of maximum values of the mass contents. The profiles of cloud averaged values would be also useful. These figures should be accompanied by corresponding comments and analysis.

4. The lack of physical interpretation of results is another drawback of the study. For instance, Fig. 6 and Fig. 7 show that in simulations CONGO - T250, T-2500 maximum cloud water content (i.e. small cloud droplets) is located at the surface. One gets the impression

that the entire boundary layer is filled with tiny droplets (but not with raindrops). What can be physical mechanisms leading to this very strange effect?

The comment concerning the lack of interpretation is related to most figures.

5. The section of Conclusions is weak. The authors stress that their key findings are

- a) "A key finding is that the simulated hydrometeor classes differ significantly between microphysics schemes"
- b) "Another key finding is that the difference between the hydrometeor classes simulated by each microphysics scheme varies between cases of convection."
- c) Another key finding is that the cloud morphological difference and the difference in the hydrometeors between different schemes is significantly larger than that due to CDNC perturbations.

As it was said above, all these findings are not new.

I would recommend to rewrite conclusions by adding more detailed analysis of results and recommendations of the ways to improve the schemes.

I recommend *major revision* before the decision on publication will be discussed.