I would like to thank Anonymous Referee #3 for their time to evaluate my manuscript. I also appreciate their constructive evaluation style. Here are my replies, following Referee #3's suggestions in bold-italics.

The study shows which microphysics (MP) schemes perform the best for this particular event. A key point (and one which would make this paper original, and significant) would be to find the common factors among the "successful" MP schemes, in other words, Why do these particular schemes perform the best? This "Why" is not addressed in the paper. Without this cause-and-effect relationship being established, it is difficult to know whether the current findings are applicable to other extreme events.

Please see my reply to Referee #1 where I partly answer your comment. I tried to explain the reasons might affect these various results and why these particular schemes perform best leaving for future studies to be analyzed in detail. Since these differences are coming from the treatment of microphysical processes, these processes, such as number of variables, drop size distribution, terminal velocity formulations, should be discussed more detail. Moreover, according to my analyses, not shown here, precipitation amount values could be more reliable if moisture flux convergence approximation should have been used in these schemes for final precipitation, especially for extreme precipitation forecasting, because the estimation of both moisture amount and wind fields of the WRF model is much more reliable than that of microphysical processes. If you think that this approach must be added to the paper, I am willingly to do that. Thus, if I add this part to the paper, there would be no question that other extreme events can also be predicted with some accuracy level.

At the beginning of the manuscript, the model simulations are de- scribed as having 28 levels. Later in the manuscript this appears to be 41 levels. Which is true? In either case, is this number of vertical levels sufficient to model events of this type? I think a few sentences here would be useful.

The current study has 41 vertical levels and the study (Tan, 2010) that I compared to for basin-averaged values had 28 vertical levels. As a result, I showed that 41 levels would add more precipitation to the system because all MP schemes over-precipitated than that of Tan's study (2010) and the only difference with these two model structure is the vertical level number. Yes, generally 28 to 41 vertical levels would be compatible with the horizontal resolution of 3 km.

P22, L25ff. "Model uncertainties..." I think the author is rather opening a "can of worms" here. It is not evident whether differences in model results associated with differing MP schemes are comparable with, or greater than, the differences in, say, horizontal resolution. This would represent a significant undertaking, however, with many degrees of freedom. I wonder if using WRF in some idealized framework may be one way to deal with this.

I, especially, thank you for your comment on this issue. We may call this as my courage of ignorance but, actually, yes, I am trying to open *a "can of worms" here for future studies, so that we may start to think how we can initiate new approaches for MP schemes, rather than trying the combinations of current closure techniques. Precipitable water analyses that I have included to my replies to Referee #1, clearly states that effort which I would like to state in this paper that it may not work for point-wise analyses of extreme precipitation events. It is very true that model uncertainty identification studies should start with idealized cases before the real events.*