

## ***Interactive comment on “Evaluation of the WRF model with different domain configurations and spin-up time in reproducing a sub-daily extreme rainfall event in Beijing, China” by Qi Chu et al.***

### **Anonymous Referee #2**

Received and published: 15 October 2017

In this contribution, the authors evaluate the performance of the WRF model in different configurations for a single heavy rainfall event centered over Beijing, China. The evaluation differs from other studies in that field in the sense that no physics parameterization evaluation is attempted. Instead, the model setup (domain configuration, number of vertical levels = vertical resolution, nesting ratio = horizontal resolution, forecasting lead time) are explored. In the order of the above, the best configuration is chosen in each step to perform several experiments in the next step. Several verification measures are employed for precipitation and precipitable water.

The design of these experiments are convincing despite a few weak points listed below.

C1

The use of English language, however, needs improvement. Grammatical mistakes and strange wordings render some parts of the text unclear. I did not make any attempt to correct this, but highlight a few common issues below. With improvements to the language and several changes to the contents, the contribution may be suitable for publication.

#### General remarks

- Dependence of parameters varied: Although discussed in the introduction, the dependence of optimal lead forecasting time on domain extent (and vice versa) is not considered in the study. Instead, based on a standard lead time of 12h, the "best" domain configuration is derived as C1, based on which an optimal lead time of 60h is inferred later on as C11. In my understanding, these two lead times should match if one really found the "best" combination of these two parameters
- The authors choose an adaptive time step to conduct their modeling experiments. This introduces another free parameter, since the actual time step adopted in each simulation may vary and as such influence the results
- The performance evaluation of the WRF model is performed over the intermediate domain D02 and not over the highest-resolution domain D01. It is argued that the two-way nesting approach does inform D02 about the results in the innermost domain D03, but interpolation to the coarser D02 grid and the (possible) difference in setup of the model physics (see next point) may influence the conclusions drawn. I would like to encourage the authors to conduct simulations without D03 for their best setup at least
- Model physics: It is unclear to me whether the GD cumulus parameterization is also employed in D03 at convection-permitting resolution (<5-10km)
- The forcing data is obtained from ERA-Interim on pressure levels (28 levels). This is not ideal, in particular since the authors are trying to assess the added value of a higher vertical resolution and since ERA-Interim is also available on 38 model levels.

C2

I would like to encourage the authors to repeat experiments with their optimal setup C11 using ERA-Interim model-level data and varying the vertical resolution as in S2, for example.

- Several of the abbreviations in the text or the figure captions are not introduced before they are used (or not at all), please check and correct

- The statistical measures used here have different directions of "good" (i.e. RMSE is good if low, R is good if high). This is not mentioned anywhere in the text and in the figures, which makes the interpretation confusing, also because the statistics are rescaled. I would encourage to state explicitly what value implies a better model performance for which statistical measure, and possibly encode this (in color or differently) in the figures and the tables

- Beneath dynamical downscaling explored here, also statistical downscaling methods and new global modeling approaches on irregular grids (e.g. MPAS) have been used and show promising results to forecast extreme precipitation events. This should be discussed briefly in the introduction or discussion section

#### Specific remarks

- Language-specific: - the word "occurred" is often missing a leading "that" or "which" - the expression "demonstrated true" seems odd to me - several times in the text, "grid" is used whereas "grid point" should be - singular and plural, as well as the use of articles need to be checked carefully

- Page 2, lines 24-26: WRF being used at resolutions >10km only is not true. Many leading operational NWP centers are employing WRF at convection-resolving resolution (NCEP: HRRR, 3km over CONUS; MeteoGroup: 3km over Europe; New Zealand MetService: <4km over New Zealand) operationally

- Page 4, line 10: "coaster-scale" -> "coarser-scale"

- Page 4, line 21: "Earth-system system"

C3

- Page 4, line 27: Isn't ERA-Interim available from 1979 (not 1989)?

- Page 4, line 29: "WRFV3.7.2" or "WRFV3.7.1" ?

- Page 5, line 6: the correct reference should be Skamarock et al. (2008)

- Page 5, line 14: "They two together" -> "The two together"

- Page 7, line 31: RRTMG schemes (not RRTM)? Or is it "RRTM" for lw, and "Dhudia" for sw?

- Page 13, line 11: "less sensitivty" -> "less sensitive"

- Page 14, lines 2-15: the discussion is confusing for the reader as he/she is expected to translate nesting ratios into effective horizontal resolution. In this paragraph, as well as in Table 1, the effective horizontal resolution should be stated explicitly, alongside with the nesting ratios

- Page 14, line 31 to page 15, line 1: the authors state that the positive bias in precipw depends on the initialization time, with largest biases for initialization times with highest amounts of precipw. This, in my opinion, is an important finding and should be highlighted and possibly discussed further

- Page 17, lines 8-11: The authors briefly discuss the dependence on the quality of the forcing data. This highlights the importance to conduct additional experiments with ERA-Interim model level data as described above, and at least discuss (if not evaluate) potential effects when using ECMWF high-res forecasts on 137 model levels and approx. 9km horizontal resolution

- Example for a necessary rewording: caption of figure 1: "Location and topography of the study area. Left panel: three levels of nested domains adopted in most experiments, with D03 covering the Beijing area; right panel: zoom-in on the topography of the Beijing area"

C4

