Hess-2012-552

Title: Simulation of a Persistent Medium-term Precipitation event over the Western Iberian Peninsula

Authors: S. C. Pereira, A. C. Carvalho, J. Ferreira, J. P. Nunes, J. J. Keizer, and A. Rocha

Recommendation: Significant Revisions

General Comments:

This manuscript investigated how suitable the Weather and Research Forecast (WRF) numerical weather prediction (NWP) model is to simulate rainfall event over a complex orographic region in north-central Portugal in order to fill the gaps of ground observation. For the purpose, three experiments using different nudging method were conducted on resolution of 1km in the innermost domain for a single month of December 2009. The results were quantitatively evaluated compared to the 27 rain-gauge data and showed that the control WRF run with no nudging technique (RunRef) could simulate reasonably the periods of precipitation occurrence. Model performance was more improved by applying grid-nudging technique (i.e., RunGridN run) in terms of continuous statistics, and local nudging technique (RunObsN run) also showed slight improvement.

While this manuscript is well written and the results are of some interest, I believe further work needs to be done before it would merit publication, and therefore I recommend it to be published after the following comments are addressed. My specific comments are included below.

Major Comments:

1. Is the model configuration first used for that region and period? If yes, the evaluation of control experiment (i.e., RunRef) should precede before elaborate analysis in terms of precipitation. In addition to the D03 (1 km) domain, D01 (25 km) and D02 (5km) domains need to be evaluated in terms of synoptic feature as well as precipitation, compared to the reanalysis and satellite-derived precipitation data. Please refer to Koo and Hong (2010). Regarding synoptic feature, the author stated "*The heavy to extreme rainfall periods were caused by several low surface pressure systems associated with frontal surfaces*." in Abstract, but this conclusion was not addressed in the manuscript at all. This can be confirmed by synoptic analysis from both observation and simulation.

2. Throughout the manuscript, Figures and Tables provide excess information. For example, Figs. 6 and 7 exhibit statistics for ALL stations but average value of classes would be enough to identify altitude dependency of model simulation if outliers are eliminated (S27MOSC2, S02BCBC2 and S25CASC3). Moreover, line plot or vertical bar chart would be better to compare the results than Table.

3. In categorical measures of Table 9 (B, PC, POD, F and ETC), skill scores seem to be almost the same among three experiments. Is it meaningful to describe them in Table 9? I think that, in Section 3.2, the statement "…, *the RunGridN experiment slight outperforms* ...

with increasing threshold value." is misleading as written.

4. Is cumulus parameterization (CP) scheme (in this study, Grell Devenyi ensemble convective parameterization scheme) used for D03 domain, i.e. 1 km horizontal resolution, as well (P1430L29)? As the author mentioned in Introduction, CP may be avoided for better model performance in terms of precipitation when horizontal resolution is less than 3 km. Therefore, sensitivity of precipitation simulation to CP should be checked for D03 domain.

5. (P1436L14) How was the model output on regular grid interpolated onto station location? Typical method is to average four grid-point values neighboring a station location (see Koo et al., 2009), which may be different from the nearest grid-point value. Please exhibit their difference.

6. (P1437L4) Does the spatial correlation between a station value and model output at different location have any particular significance? I think that spatial correlation may be low even between station values estimated at different location. Moreover, is the correlation value of 0.018 really significant? (Fig. 2)

7. (P1437L17) In Table 9, categorical verification measures (B, PC, POD, F and ETS) are almost same among three experiments irrespective of threshold. Are the statistics valid for evaluating them? How about the thresholds above 3 mm per hour that provides information for heavy rainfall?

Minor comments:

1. Please specify full name when an abbreviation is first used.
- P1428L25: IDF
- P1429L18: MM5
- P1438L8: IQR
- P1442L3: HIRLAM
- Table1: PP (Its full name should be described in caption)

2. Specific comments
- P1424L2: numerical weather model → numerical weather prediction model
  L5: What is the basis of "*The heavy to extreme rainfall periods were caused by several low surface pressure systems associated with frontal surfaces.*" ? I couldn't find any relevant discussion throughout the manuscript.
  L16: Is root-mean-squared error (RMSE) be the only representative of model accuracy?

- P1425L28: numerical weather prediction → NWP

- P1426L7: domain **horizontal** resolution?
  L8: Remove "*and*".
  L8: domain what?
  L19: Who are the same authors? Luna et al. (2011)? or Heikkila et al. (2011) and Luna et al. (2011)?

- P1428L5: What is the criterion of exceptional amount?
  L13: S and E part → southern and eastern parts
  L17: S and SE → south and southeast
  L21: NW → northwest

- P1429L26: (NW) → i.e. northwestern region

- P1430L22: WRF Single Moment 6 class scheme microphysics → WRF Single Moment 6 (WSM6) microphysics scheme
  L25: Remove ";".
  L26: Change the reference "Noh et al., 2003" to "Hong et al., 2006".

- P1432L22: the station location → the ***i*-th** station location

- P1436L21: mean absolute error (MD) → MD
  L22: "*The results among experiments are identical*" is not correct. The results among RunRef, RunObsN and RunGridN are different one another in Table 4.
  L22: The MD values ... → For RunRef run, the MD values ...
  L23: 0.31 → 0.31 (S17PARC3)
  L24: For S25CASC3 in Table 4, MD is not 1.49 but 1.50 mm h^{-1}.

- P1437L7: lag → time lag
  L11-12: This statement is different from the results of Table 5-7.
  Moreover, line plot would be better to compare one another than Table 5-7.
  L13-17: Is Table 8 positively necessary? In this statement, median and mode are not described.
  L26: as well → as well as

- P1439L27: This paragraph is not quantitative but qualitative. How small (large) is the mean error (MSE and RMSE) compared to what?

- P1440L3: In Fig. 7, specify the statistics' name at the top of the table.

- P1442L25: RUNGridN → RunGridN

3. Table and Figures
- Table1: PP ?
  return period (yrs) → return period (yr)

- Table 5: Lag correlation → Time (hour) lag correlation

- Table 7: RunObsN → RunGridN

- Figure 6: For RunRef, MSEs were colored by red although they should be positive values.

- Figure 7: Captions are needed in the top of Table. (for experimental names and statistics)

Reference:

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318-2341.

Koo, M.-S., and S.-Y. Hong, 2010: Diurnal variations of simulated precipitation over East Asia in two regional climate models. *J. Geophys. Res*., **115**, D05105, doi:10.1029/2009JD012574.

Koo, M.-S., S.-Y. Hong, and J. Kim, 2009: An evaluation of the Tropical Rainfall Measuring Mission (TRMM) Multi-Satellite Precipitation Analysis (TMPA) data over south Korea. Asia-Pacific J. Atmos. Sci., 45(3), 265-282.