Hydrol. Earth Syst. Sci. Discuss., 9, C4454–C4458, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C4454/2012/


Hydrology and Earth System Sciences Discussions

# *Interactive comment on* "Weather model performance on extreme rainfall events simulation's over Western Iberian Peninsula" *by* S. C. Pereira et al.

**Anonymous Referee #1**

Received and published: 29 September 2012

The authors of this paper use the WRF-ARW model to create three different month-long runs to evaluate them against a single month of precipitation observations from December 2009 over the Iberian Peninsula. One of these runs is used as a base run that is forced with a global circulation model, the second run includes nudging from a single observation station (which the authors refer to as "data assimilation"), and the third run includes nudging on all points on the grid. Through various statistical error techniques, the model runs are compared to the observations and to each other to assess the ability of the model to accurately reproduce the observations during December 2009. The authors conclude that the model is "reliable and consistent" (pg



9175, L25) in terms of how it reproduces the precipitation among the different runs. They also claim that the model does a good job reproducing the individual precipitation events throughout the month of December and does a reasonable job reproducing the spatial patterns of precipitation; however, it does not do so well with reproducing the proper precipitation intensities, especially in complex terrain.

The main motivation for this paper is to evaluate the ability of the WRF model to accurately supply precipitation fields of an extreme nature for future use in hydologic models. While all three runs of the model produce similar results to each other in all forms of error calculation (Figures 4, 6), I question if this is enough to actually declare that the "extreme episode was successfully reproduced" (pg 9164, L18), especially since the amount of precipitation is not reproduced well (pg 9164, L16 and pg 9171, L10-11)). Additionally, the correlations between observed and modeled time series are low, as described at the end of section 3.1 but not shown (pg 9172, L1-6). The authors claim that it is due to timing and space issues (pg 9172, L5), even though they claim that the timing of precipitation is well represented by the model (pg 9164, L15-16). A point correlation for the time series at each observation station with the time series at each grid point for each of the model runs would tell you exactly how well each observation station correlated in space with the nearest grid points. Additionally, if you repeated the analysis done in Figure 3, looking at anomalies from observations rather than absolute quantities, you would be able to better assess the timing errors. The combination of these two results would allow you to tell which factor has a larger contribution to low correlations, the timing error or the spatial error. Other analysis might include splitting the stations into multiple groups according to elevation. It appears that you credit much of the error signal in space to high terrain stations not lining up appropriately with the model grid, yet the highest elevation stations (S14, S24) appear to have some of the smallest errors (combining information in Table 3 with Figure 2). The largest errors occur at stations at intermediate elevations. A final major concern with this study is the nature of the month of December 2009 being described as extreme. An overall extreme month does not equal a month containing extreme events as described in

your title. The authors even mention that the high levels of rainfall for the month could be attributed to a greater number of rainfall days rather than intense events (pg 9167, L10-11). The analysis of return time displayed in Table 2 shows that the maximums like those observed in December 2009 occur roughly every two years. So, although several of these events occurred in the same month, how extreme are the individual events really? Has any of this same analysis been done on other months with heavy precipitation? Has this analysis been done on other months that contain very extreme events (like 99th percentile type events?

Finally, the conclusions section (really more of a summary section rather than conclusions) could do a much better job synthesizing the analysis and talking about the implications of the results on applications for hydrologic models. The authors do not mention the motivation in the conclusions section, or anywhere else, after the original reference at the end of the Introduction section.

In addition to the concerns mentioned above, the manuscript contains several grammatical errors and technical problems, such as, a mislabeling of the last several figures, making it quite difficult to read and comprehend. Specifics of these issues are detailed below:

1. The title is misleading. This study does not evaluate model performance on extreme events. It evaluates the models ability to reproduce observed precipitation fields over the course of a single month, a month in which higher than average precipitation fell, not necessarily as few extreme events but probably as many moderate events. 2. (pg 9164, L12) Simulation 2 is referred to as data assimilation but it is not mentioned this way again in the entire manuscript. It was initially confusing to me and took me some time to realize that it was just a different kind of nudging from experiment 3, since L13 describes simulation 3 as nudging. 3. (pg 9166, L23) 2300mm must actually be 2300 meters. 4. (pg 9167, L5) Please define INAG. Is it the Portuguese Water Resources Institute? I had to look this up on the web. 5. (pg 9167, L10-11) This is a key statement leading me to believe that the individual events during this



month aren't actually that extreme and led to comment 1 above. 6. (pg 9169, L17) Please define SNIRH and describe the dataset. There are also several references I was able to find online, although it does not appear that any of them were published. (http://archive.rec.org/REC/Programs/Telematics/DETERMINE/WaterSession/JRibeiraDaCo
7. (pg 9169, L18) Figure 1 does not include units on the elevation profile. Kilometers along the x-axis in the same panel would also be helpful. 8. (pg 9169, L23-26) It might be helpful to linearly interpolate the model data to the observation location and compare that time series with the observations. This is another way of verifying just how much of the error is coming from the spatial issue or the timing issue. In areas of complex terrain this could make a huge difference. 9. (pg 9170, L9-14) The figure talked about in this paragraph does not exist and is the source of the mismatched figure numbers. Figures 3-8 are all off by one. 10. (pg 9170, L20-25) I feel like all of the statistical calculations show basically the same thing. Is it necessary to include all of them? The same goes for Figure 6, where panels a-e all show basically the same thing, the model run performance relative to each other. 11. (pg 9171, L14) Do you mean to say overestimate instead of underestimate? 12. (pg 9171, L16) Fig 3 should be Fig 4. Fig 4 should have units on the y-axis and perhaps a legend. Also, there is no mention of the colored plus symbols on the plot. Either remove these symbols or address their meaning. 13. (pg 9172, L7) Fig 4 should be Fig 5. Please add units to the plot and the caption for Fig 5. 14. (pg 9172, L10) Fig 5 should be Fig 6. Please add units to the x- or y-axes. The plot is described as grey-scale in the caption, yet it is a color plot. 15. (Section 3.2) Again, I think it is unnecessary to detail all of the various statistical methods. They appear to all be saying the same thing. Choosing one (or a couple) is sufficient to assess how the various runs compared to each other and the observations. A brief mention that you tried several other methods and they produced similar results would be sufficient. 16. (pg 9173, L28) Do you mean climatology instead of persistence? 17. (pg 9174, L1-2) The fact that the skill score indicates that some stations are better off with climatology instead of modeled precipitation is worrisome. What does this mean for the accuracy of representation of

precipitation, etc., in the hydrologic models these runs are meant to be used for? This is never discussed. 18. (pg 9174, L18) "...the average magnitude of the error is the same: it is on average larger in areas of high slope and small in the lowlands."

(pg 9174, L21-22) "...magnitude error higher at the majority of stations and lower in the areas of rugged terrain (top right corner).

These two statements appear to contradict each other. Especially since, in Fig 7c, it appears that the highest values on the plot are over the rugged terrain in the top right corner. The contour there says 2, which is the highest in the domain.

19. (pg 9174, Section 3.3) Figure 7d is never referenced. Is it necessary? 20. (pg 9176, L16) I think you mean overestimates, not underestimates.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 9163, 2012.