Response to Interactive discussion

Hydrology and Earth System Sciences (HESS)

Title: A Nonparametric Statistical Technique for Combining Global Precipitation Datasets: Development and Hydrological Evaluation over the Iberian Peninsula

Md Abul Ehsan Bhuiyan,[1] Efthymios. I. Nikolopoulos,[1,2] Emmanouil. N. Anagnostou,[1] Pere Quintana-Seguí,[3] Anaïs Barella-Ortiz,[3,4]

We would like to thank Dr. V. Maggioni for her insightful discussion and constructive suggestions. Below we provide a point-by-point response to her comments. Dr. Maggioni's comments are in red and our responses in black font.

Sometime in the text, there is confusion in the way the words "ensemble" and "members" are used. An ensemble is made of several members. Therefore, "ensembles" would refer to multiple ensembles made of several members

Page 1, Line 12: replace "generated ensembles to force" with "to generate ensembles that force"

Thank you. It will be corrected in the revised paper.

Page 2, Line 18: rephrase as "Satellite rainfall error models, such as SREM2D (Hossain et al., 2006), have been used to. . ."

Thank you. It will be corrected in the revised paper.

Page 2, Line 22: replace "the error characteristics" with "errors and uncertainties"

Thank you. It will be corrected in the revised paper.

Page 2, Line 24: replace "allow for efficient combining of" with "efficiently combine"

Thank you. It will be corrected in the revised paper.

Figure 2: a "d" is missing in the word "and" in the top center block. This methodological framework scheme could be improved: as it is, it is unclear what each block does.

Thank you. It will be corrected in the revised paper. The methodological framework scheme will be updated by improving content and clarity.

Figure 6: can the authors discuss why QRF performs so poorly for lower rain percentiles (<25%) in term of bias?

Ans:

Thank you for the comment. Generally, the QRF model is expected not to capture well very low and extremely high values due to the weakness of the empirical distribution function to model

probabilities close to 0 or 1. Moreover, studies have shown that QRF can perform better in generating one-sided prediction intervals, which is the cases in Juban et al. (2007), Francke et al. (2008) and Zimmermann et al. (2012). Nevertheless, we would like to note that these very low rain rates that the model cannot capture well represent a small fraction of the precipitation accumulation.

Major comments:
A better explanation on why the authors picked those predictors is needed. For instance, why including soil moisture, but not a vegetation indicator? Why including three satellite precipitation products instead of two? Was the dataset combination that produced the best results picked? Are all the predictors really needed?
What is the impact of merging datasets that are not totally independent? For instance CMORPH and TMPA 3B42 use the same MW overpasses in their algorithms, as PERSIANN, CMORPH, and TMPA use the same IR observations. I am wondering whether there is any chance that too much weight is given to this information in the merging model.

Ans:

Thank you for asking about the contribution of the different predictors to the final combined precipitation product.

We selected all the predictors in our error analysis based on the variable importance methodology (Breiman, 2001), which indicated the level of influence of variables in the model prediction. We are going to include an explanation for choosing all those predictors in the revised paper.

Soil moisture and vegetation are dependent on each other. If vegetation increases, soil moisture increases. For different vegetation, moisture holding capacity is also different. As we had a quality controlled soil moisture dataset available, it was pretty straight forward to use that data, instead of using any vegetation indicator.

Below, we present results for the variable importance test only for one of the groups included in our methodology (for warm period-high elevation when rainfall is greater than zero for all products). The figure shows the variable importance of the seven predictors. According to these results, soil moisture, reanalysis and three satellite precipitation (CMORPH, PERSIANN and 3B42 (V7) were ranked as most important predictors. This is the reason why we choose three satellite precipitation products instead of two to produce best results.

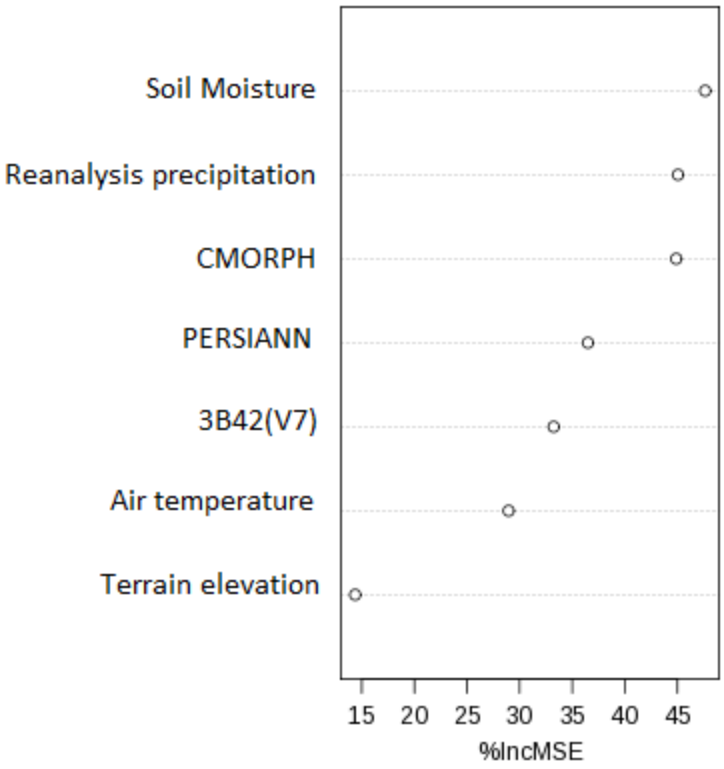

Figure: Variable importance plot, where %IncMSE is the percentage increase in mean square error. Higher values of %IncMSE indicate higher importance of the predictor variable.

The variable importance test for the all predictors showed that the model results closely sensitive to all those variables. Therefore, we select all the dataset combination to produce blending precipitation product.

From the variable importance test, it is clearly shown that the individual satellite precipitation product has a strong impact in model prediction. So, better prediction is expected after merging all those products although they are not totally independent.

We are going to include a complete analysis in the revised paper.