Hydrology and
Earth System
Sciences

Open Access

Discussions

HESSD

Interactive
comment

Interactive comment on "Which rainfall metric is more informative about the flood simulation performance? A comprehensive assessment on 1318 basins over Europe" *by* Stefania Camici et al.

Stefania Camici et al.

s.camici@irpi.cnr.it

Received and published: 8 June 2020

The paper addresses the relevant scientific question of what are the most important metrics to assess the goodness of a SRPs product for hydrological applications. The question as well as the motivation of this work are stated clearly in the context of a comprehensive literature review. The methodology is appropriate to answer the question and the extensive analysis over 1318 basins across Europe defines the main novelty of this paper. Substantial conclusions about the most relevant indexes for assessing the quality of SRPs product for hydrological applications are reached, so overall this is good contribution for the scientific community. However, there are a number of issues

that the authors need to address before the paper is accepted for publication.

R: We thank the reviewer for his/her supportive review. In the revised version of the manuscript the following changes will be implemented:

1. the title will be changed in "Which rainfall performance score is more informative about the river discharge simulation? A comprehensive assessment on 1318 basins over Europe".

2. Any reference to flood will be removed and modified with river discharge to highlight that the purpose of the study is to investigate the performances between rainfall and discharge time series (without specific focus on high and/or low flows);

3. a discussion about the quality of the E-OBS rainfall data and the impact of its density network on river discharge simulation will be added to the revised manuscript;

4. to avoid misunderstanding between KGE of rainfall and discharge throughout the manuscript (and in the figures) KGE will be replaced by KGE-P and KGE-Q, to indicate the KGE index referred to precipitation and the one referred to discharge respectively;

5. more information about why only the KGE-Q index has been selected for the analysis will be added to the revised manuscript;

6. Tables and figures will be modified according to the reviewer's suggestions.

Specific comments

1. Line 158-169: the E-OBS dataset is built on a station network with an average station density of 1 in 4000 km2 and the basin areas range from 200 to 136'000 km2. Is the E-OBS dataset a reliable benchmark for the smaller basins? Maybe it is worth to discuss this in your discussion section.

R: In the submitted version of the paper, an error occurred in the definition of station used within the E-OBS datasets. Indeed, 2316 stations (i.e., equivalent on average to a density of 1 station every 4000 km2) is referred to the first versions of E-OBS whereas

in the version 17 (used in the manuscript) the number of stations increased up to 9618 (equivalent on average to a density of 1 station every 1000 km2). However, as correctly raised by the reviewer even the E-OBS density network referred to version 17 could be too low to correctly represent the rainfall spatial variability over small basins. This, in turn could affect the discharge simulation. To consider this aspect, it has been verified that 1) for all the analysed basins the KGE-Q values obtained by the calibration of the model by using E-OBS dataset as input were greater than -0.41, i.e., the model improves upon the mean flow benchmark 2) no relationship between basin area and KGE-Q exists (see Figure 1). As these conditions were satisfied and as the purpose of the study was to investigate the performances between rainfall and discharge time series (without specific focus on high and/or low flows), the limitations about the E-OBS station density can be assumed to have a negligible impact for the analysis purpose.

Figure 1. Relationship between basin area and KGE-Q for the analysed catchments.

Accordingly, two sentences will be added in the revised version of the manuscript (on section 5.2): "The results of this calibration, carried out for the entire observation period (2007-2016), are good as illustrated in Figure 3a: for all the analysed basins the KGE-Q values are greater than -0.41, i.e., the model improves upon the mean flow benchmark and the median KGE-Q value obtained for the European area is equal to 0.768 (0.770 over the TMPA area). In addition, to take into account that due to the density network, E-OBS rainfall data could be not reliable for smaller basins (area<1'000 km2) the relationship between basin area and KGE-Q has been investigated (not shown). As no relationship was found, and as the purpose of the study is to investigate the performances between rainfall and discharge time series (without specific focus on high and/or low flows), the limitations about the E-OBS station density can be assumed to have a negligible impact on the analysis results and QE-OBS data can be assumed as a good benchmark for the successive analysis."

2. Line 331: for the discharge assessment you used only one performance score, the KGE. Can you provide more information about why you selected this score?

R: We select only the KGE score to evaluate the hydrological model performances for three main reasons: 1) due to inherent limitations recognized for NSE (e.g., Schaefli and Gupta 2007; Gupta et al., 2009), KGE is today the criterion most commonly recommended and applied to evaluate the performance of hydrological models and therefore its use allows meaningful comparisons with other studies. 2) the purpose of the of analysis was to investigate the relationship between rainfall score and discharge simulation, without specific focus on high and/or low flows. In this respect, it is known that KGE assign a relatively more importance to discharge variability with respect to other scores (e.g., NSE or RMSE) generally found to be highly sensitive to high discharge values (Gupta et al., 2009); 3) for a practical reason, i.e., it was a decision of the author to limit the number of investigated performance scores to communicate in the most efficient way the results of the work. However, as stated in the conclusion section, in the future a more comprehensive study could consider a larger set of discharge scores metrics to better address the SRP selection. The reasons of why we selected the KGE score will be added in the revised manuscript in the section "performance scores" as on the following: "To evaluate the suitability of rainfall products for river discharge modelling, the KGE index between observed and simulated river discharge data has been computed. In particular, we selected only this score for three main reasons: 1) due to inherent limitations recognized for other indices (e.g., Nash-Sutcliffe Efficiency index, Schaefli and Gupta 2007; Gupta et al., 2009), KGE is today the criterion most commonly recommended and applied to evaluate the performance of hydrological models and therefore its use allows meaningful comparisons with other studies; 2) the purpose of the of analysis was to investigate the relationship between rainfall score and river discharge simulation, without specific focus on high and/or low flows. In this respect, it is known that KGE assign a relatively more importance to discharge variability with respect to other scores (e.g., NSE or RMSE) generally found to be highly sensitive to high discharge values (Gupta et al., 2009); 3) for a practical reason, i.e., it was a decision of the author to limit the number of investigated performance scores metrics to communicate in the most efficient way the results of the work."

3. Line 397: Can you explain better why KGE of rainfall is not relevant? From figure 4 the increasing trends of KGE-Q with rBIAS and KGE of rainfall look quite similar.

R: The authors verified the increasing trend both for KGE-Q vs rBIAS and KGE-Q vs KGE-P. Although a difference in the magnitude and correlation of the relationship between KGE-Q vs rBIAS and KGE-Q vs KGE-P can be noted, i.e., the slope coefficient is equal to 1.07 (R2= 0.98) and 0.80 (R2= 0.81) for KGE-Q vs rBIAS and for KGE-Q vs KGE-P, respectively, the sentence in the revised manuscript will be smoothed as: "SRP hydrological performances decrease by increasing the absolute value of rBIAS, |rBIAS|, and the RRMSE values (Figure 4a and b) whereas KGE-Q increase with R and KGE-P (Figure 4c and d)."

4. Line 411: How do assess that R and KGE ranges are large?

R: In Line 411 it has been observed that "R and KGE-P seem to have a small impact on KGE-Q as for a large range of R and KGE-P values (from 0.5 to 0.8 and from 0.4 to 0.8, respectively), it is possible to obtain high KGE-Q values." The assessment about the "large ranges" for R and KGE-P values has been carried out by considering that, even if the two scores potentially range from -1 to 1 and from $-\infty$ to 1, respectively, reliable performances are obtained for R and KGE-P values constrained between 0 to 1 and between -0.41 to 1, respectively. Therefore, a range of 0.3 and 0.4 can be considered "large" with respect to the variability range for which the rainfall scores suggest reliable rainfall data. To better explain this aspect in the revised manuscript a sentence will be added in the discussion section as follows: "In particular, it has been noted that R and KGE-P rainfall scores seem to have a small impact on KGE-Q as for R ranging from 0.5 to 0.8 and for KGE-P ranging from 0.4 to 0.8, it is possible to obtain high (>0.5) KGE-Q values. If these two variability ranges are compared against the variability range for which the same rainfall scores identify reliable rainfall data (i.e., for R and KGE-P values constrained between 0 to 1 and between -0.41 to 1, respectively) it can be concluded that R and KGE-P are not suitable scores to define a criterion able to discern between good/bad hydrological simulations."

Technical corrections

1. Line 124: State all the questions here. I can see that you have more questions later (e.g. lines 416, 417, 418).

R: According to the reviewer suggestion, all the questions will be moved at the end of the introduction.

2. Line 167: add spatial resolution of the product in the text.

R: The resolution of the E-OBS dataset will be added to the revised manuscript.

3. Line 215: it is a bit confusing when you say below TMPA area because I guess you mean the TMPA area. Change accordingly also in the other paragraphs and tables.

R: The reviewer is right. With "below TMPA area" the authors were referring to the TMPA area. The sentence will be modified with "TMPA area" throughout the manuscript.

4. Line 262-263: swap the two lines because in the plots you present first rBIAS.

R: According to the review suggestion, the two lines will be swapped in the new version of the manuscript.

5. Line 262: remove "x".

R: Accordingly, the "x" will be removed in the formula.

6. Line 263: I think the numerator shouldn't be squared.

R: The reviewer is right; the numerator shouldn't be squared. In the revised version of the manuscript the rBIAS formula will be modified, accordingly.

7. Line 265: in the second bracket under the square root I think there is a mistake (see Gupta et al., 2009). The ratio in the bracket should be just between standard deviation of the SRP and of the E-OBS.

R: The reviewer is right. The KGE formula will be modified in the revised version of the manuscript.

8. Line 300-Figure 2: you are talking about "patterns" so I assume you are referring to Figure 2, but then the values at line 302 are the ones reported in Table 3, so for the TMPA area. It is a bit complicated to follow, maybe you can just condense the most relevant information in figure 2 and put table 3 in supplementary material, since it doesn't provide much more information.

R: The reviewer is right; this part is difficult to follow. Therefore, in the revised version of the manuscript it will be modified as: "Already at first glance of Figure 2, it is possible to note that the three products show similar patterns in terms of R and RRMSE whereas the same does not hold for the rBIAS and KGE-P. The rBIAS is small for TMPA and SM2RASCAT, with median values equal to -0.127 and 0.047, respectively, whereas CMOR show a clear underestimation of the daily rainfall data over the entire European area. Higher/lower R/RRMSE values are obtained in Central Europe; the opposite is observed in the Mediterranean area. In terms of KGE-P, TMPA presents higher values with respect to the other two products above all over the basins whose outlet section is located between 40° and 50° latitude. Median KGE-P value for TMPA is equal to 0.516; this value reduces of about 24However, to be consistent with Table 3, Table 2 will not be removed from the main manuscript.

9. Line 328: I would put Figure 1.b which belong to result section in a separate figure from Fig.1.a which belong to the dataset section.

R: Accordingly, Figure 1b will be merged in Figure 3.

10. Line 342: the 39

R: We thank the reviewer. The values in the manuscript will be modified in accordance with Table 4.

11. Line 379: higher absolute values of rBIAS.

R: We thank the reviewer. The sentence will be modified accordingly.

12. Line 389: maybe name the KGE as KGE-Q otherwise it can be confused with KGE of rainfall

R: We thank the reviewer for this suggestion. The KGE will be modified as KGE-P and KGE-Q to refer to KGE of rainfall and discharge, respectively. A sentence to clarify this distinction will be added to the revised manuscript (section 4.5): "To distinguish between the KGE of rainfall and discharge, hereinafter, the symbols KGE-P and KGE-Q will be used."

13. Figure 2-3: you can add the name of each SRP product at the top of each column.

R: Figure 2 and 3 will be modified according to the reviewer suggestion.

14. Line 734: there are no figures d), e), f). Change CMORPH to CMOR to be consistent.

R: The caption of figure 3 will be modified accordingly.

References Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.
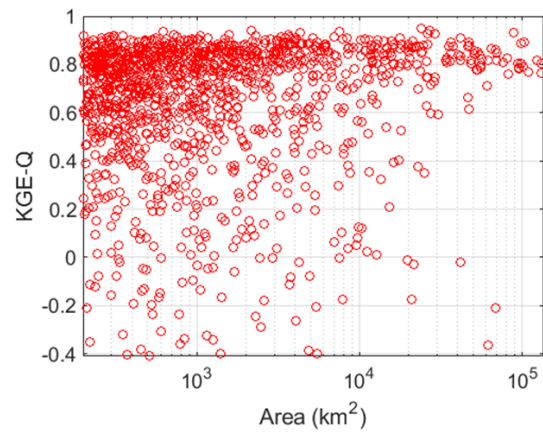

Fig. 1.