Reviewer #1

#RC1.1 Based on my personal reading, the presented analysis is of interest for hydrological modelling of ungauged catchments. I believe that the manuscript is well written and organized, and it deserves to be published in NHESS, after minor revisions listed below.

We thank the reviewer for this positive feedback and the different comments that will help to improve the manuscript.

#RC1.2 Line 56: “less than 200 km2” … you mention 300km2 at line 69

Thank you for this comment. We will replace 200 by 300 at this line. The majority of the catchments have an area smaller than 200 km^2 but indeed a few of them are in the range 200-300 km^2.

#RC1.3 Line 66. I suggest to add (in this section 2.1 or at least in the supplemental) a table with the list of the catchments with their main features (numbering, identification name, area, elevation range or average value, …). If possible, consider to report the number of each catchment in the map in figure 1, and in the explanations of results and discussion together with the catchment name. This could help the reader in identify the catchments and their main features when reading the following results and discussion section. I was sometimes lost with the different names of the catchments.

Thank you for this suggestion that would help to follow the interpretations of the results. We will try to add numbers associated to each catchment in Figure 1, along with their names on a right panel. It is also interesting to add a table with the features of the catchments and it will be added in the manuscript.

#RC1.4 Line 225. In all the equations in section 4.1, are you using absolute differences?

In all the equations in Section 4.1, we are using absolute differences as opposed to square differences in order to avoid inflated influences of the largest values. For some criteria (quantile relative error, volume error), these absolute differences are divided by the observed criteria and are in the end relative differences.

#RC1.5 Line 251-258. You define some errors (let’s say E, where E = PFE, TPE, VE) using absolute differences, then you transform them as 1- E. I have to comments here: i) why not just considering E with sign, for having a measure of the direction of the error (under- or overestimation)?; ii) I found confusing the use of both E and 1-E in the explanations of the results (for example, lines 279-289). I suggest to use just one.

Thank you for this comment. Concerning the first comment i), the idea was to obtain criteria that are all positively oriented, as a mixture of positively and negatively oriented criteria can also be confusing. However, considering the second comment ii), we understand that the explanations are not clear with the interpretations of both raw and transformed criteria. We propose to keep only the raw criteria in the revised manuscript and indicate clearly if they are positively or negatively oriented in Figure 3.

#RC1.6 Line 270-275. i) For better compare the results, I suggest to use same y-axis limits for the plots referring to the same type of index. For example, for mNSE in panels a-d, for QRE in panels e-h, … ii) maybe a comment is needed about SMASH model in panel d), showing huge range of mNSE compared to MORDOR-SD.

We thank the reviewer for these constructive comments. Following comment i), we will modify figure 3 to have the same y-axis limits for similar criteria. We agree with the reviewer that it will help the comparison of the performances for different hydrological signatures. Concerning comment ii), we comment on the differences obtained with SMASH later in the discussion, at lines 449-452, but we agree that a comment could be added at this stage of the manuscript.

#RC1.7 Line 439: You mention some problems in high-elevation catchments, and (line 468) that COMEPHORE underestimates precipitation in high elevation areas. Consider to add a plot of mNSE vs elevation, for example as panels in figure4, to synthesis the maps and made more evident (if there is) a relation of mNSE with elevation.

Thank you for this suggestion. Please note that even if COMEPHORE underestimates precipitation in high elevation areas, this does not translate into weaker mNSE values since these underestimations can be compensated by applying correction parameters, as explained at l. 472-474 of the manuscript. The figure R1 below shows the relationships between the different mNSE criteria and the elevation. There are no clear positive or negative relationships. In winter and spring, mNSE(Q) seems to exhibit a positive trend due to large variations of the criteria for low elevations. In summer and spring, the scatterplot seems to indicate a negative trend for elevation greater than 1000 m but this impression is mainly due to weak mNSE values obtained with SMASH. For the mNSE applied to the 10 largest floods (Fig. R1c), it is clearly highlighted that the floods are not well reproduced for some catchments located at high elevations, e.g. for the catchment L'Arve@Chamonix-Mont-Blanc. As these results do not seem to add more evidence of these relationships, they will not be included in the manuscript.

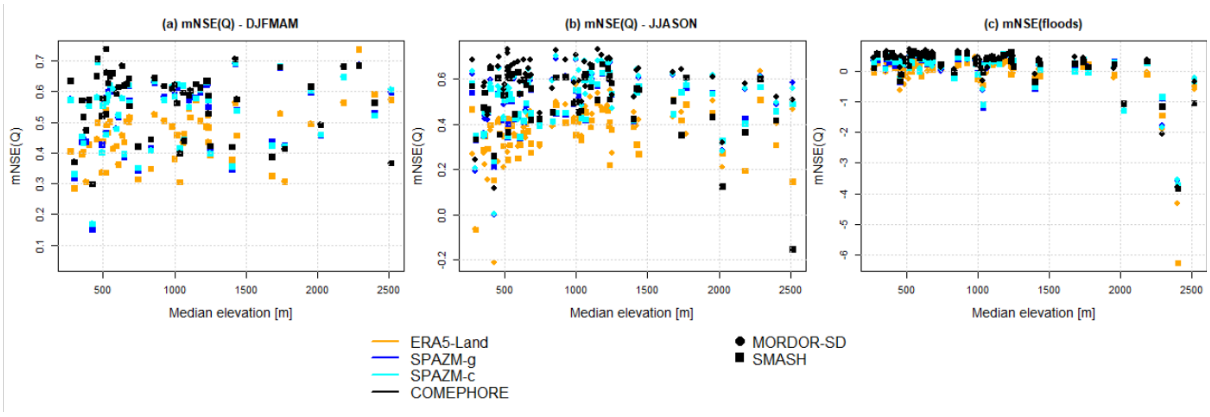

Figure R1. mNSE criteria as a function of the median elevation of the catchments, for the four different precipitation reanalysis and the two hydrological models. (a) *mNSE* of the streamflows *Q* in winter and spring. (b) *mNSE* of the streamflows *Q* in summer and autumn. (c) mNSE for the 10 largest floods.

#RC1.8 Line 457: "small mountainous" maybe is "small catchments".

Thank you for noticing this missing word, this will be corrected.

#RC1.9 Supplemental, Figure S36-37. Same color scales in the panels could help in the comparisons.

Thank you for this suggestion, we agree that it would help the comparisons between these two figures and this will be modified.

#RC1.10 Supplemental, line 16. “than with COMEPHORE” is maybe “MORDOR-SD”.

Thank you very much for your careful reading of the supplements and for noticing this mistake, this will be corrected.