Review of "Linear Optimal Runoff Aggregate (LORA): A global gridded synthesis runoff product" by Hobeichi et al.

The paper presents a new global gridded dataset at half degree resolution based on a weighted average of global runoff and discharge estimates from global hydrological models constrained by discharge observations. It represents a valuable new dataset that may be of interest to hydroclimatologists or environmental scientists. The paper is concise and well written and deserves to be published in HESSD. There are a few issues that need to be resolved however.

1. I miss a few explicit examples explaining why runoff is at all useful, especially at 0.5 degrees. The argument is that accurate estimates of runoff are critical to inform climate change adaptation strategies, to guide appropriate water management in agriculture and to enable the assessment of the impact of anthropogenic activities on ecosystems. However, what does runoff at 50x50 km resolution even mean? It is in terms of scale too far off from being operationally relevant. Thus, a stronger justification using examples is called for.

2. I also miss some references to the earliest work on runoff fields, e.g. Fekete et al. 2002: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/1999GB001254

3. Regarding to previous work and scales, I would like to call the authors' attention to a recent publication by Barbarossa et al.(https://www.nature.com/articles/sdata201852). They provide discharge estimates at 1 km resolution. I move that these are probably more representative for local runoff than those obtained from GHMs at half degree resolution?

4. Section 2.3: I find that there is too little info on the method used to compute the weights. I don't think that it should be necessary read another paper to comprehend the essentials of the methods used. So, I would want some more explanation on how the weights are calculated.
   For instance:
   - How are correlations between models accounted for?
   - Are the weights allowed to be negative?
   - Is the sum of the weights adding up to one? If this is the case, one has to add another equation and transform a constrained optimization to an unconstrained one using Lagrange multipliers.

5. Line 8-10 page 5: This seems to assume that travel times are less than a month because it neglects routing?

6. Lines 11-13 page 5: "It provides better estimates than simply calculating the standard deviation of the involved products". Is that really the case? If your weighting method is optimal and you have bias-corrected correctly would the following estimator for each pixel not be unbiased (i and j are different products bias corrected):

$$\hat{\sigma}_r^2 = \frac{1}{N \times 11 \times 11} \sum_{i=1}^{11} \sum_{j=1}^{11} \sum_{t=1}^{T} w_i w_j (r_{i,t} - \mu_r)(r_{j,t} - \mu_r)$$

   By moving a window T over time you get your time varying variance

7. Line 20-25 page 5: transferring the weights from donors to receptors. Are one set of weights obtained jointly for the three donor catchments or are three sets of weights averaged and transferred?

8. Line 11 on page 10. Why compare only with VIC? Why not GLDAS (4 models).

9. Table 2: why not add the estimated total runoff volumes from GRDC (also globally in km^3). It would be good to see what the global runoff volume is in this product compared to other estimates.

10. Line 27 on page 10: reduced performance in dry climates. Apart from the reasons mentioned, another possible cause could be the fact that GHMs are probably less proficient in representing runoff processes in arid basins where often runoff is local and will not always be turned into streamflow