

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
<https://doi.org/10.5194/hess-2021-8-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on hess-2021-8

Anonymous Referee #2

Referee comment on "A geostatistical spatially varying coefficient model for mean annual runoff that incorporates process-based simulations and short records" by Thea Roksvåg et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-8-RC2>, 2021

This paper proposes a method for the spatial prediction of annual runoff that uses process-based model outputs (model HBV), fully gauged catchment data and partially gauged catchment data. The latter are pre-processed using a method published by the same co-authors (<https://doi.org/10.5194/hess-24-4109-2020>) in HESS and it is optional if all catchment are fully gauged.

This contribution is a significant contribution since the proposed approach allows to provide improved maps of annual runoff at the national scale, by taking into account fully gauged and and partially gauged catchments.

The proposed method is hierarchical in a Bayesian framework; it includes a Gaussian observation process, a spatial latent model and hyperpriors. The latent random field is defined on a fine 1km x 1km grid. It is a Spatially Varying Coefficient regression model to the HBV outputs that involves two GRFs characterized with stationary Matern covariance functions. The statistical analysis is performed using INLA/SPDE.

The areal nature of the data is taken into account by relating the catchment data to the sum of the latent model over the catchment.

The approach is validated using a carefully designed k-fold cross-validation. The approach shows superior performances as compared to HBV only and to Top-Kriging.

General comment:

=====

The statistical analysis is state-of-the art. It uses the most modern tools for analyzing

complex spatial data in a hierarchical Bayesian framework. The experimental set-up and the validation study are very carefully designed. The results are well discussed.

Overall, the paper is well written, maybe on the long side though considering that lots of material is common with <https://doi.org/10.5194/hess-24-4109-2020>. I spotted some rare typos and a few awkward sentences (see below). All illustrations are relevant. I believe that this paper will be a nice methodological follow-up to <https://doi.org/10.5194/hess-24-4109-2020>. I recommend publication after a minor revision that takes into account the specific comments below

Specific comments:

=====

1. A potential weakness of the method, which has been mentioned by the authors, is that the model does not prevent negative run-off predictions in some (unlikely situations). This is due to the Gaussian likelihood and Gaussian GRF. The authors mention that log-runoff could be used instead, but then linearity of Eq. (6) is lost, which is an impediment. Another way of preventing negative predictions would be to use log-Gaussian likelihood and log-Gaussian random fields for $x(u)$ and $q(u)$ in (4). This would be a marginal change, since INLA/SPDE allows for log-gaussian likelihood and LG random fields at almost no cost. As a result, predictions for x and q would always be positive. I wonder how this would work. Ideally, I'd like the authors to try this option, but I'd be happy if they only discuss this possibility.

2. The GRFs $x(u)$ and $q(u)$ are independent. This assumption is never clearly stated and it is not discussed. Is this a reasonable assumption? Is this an assumption you could check or validate? How useful/difficult would it be to relax this assumption?

3. The GRFs $x(u)$ and $q(u)$ are assumed to be stationary. Are you able to check that this assumption is supported by the data?

4. To my knowledge, the product of an exponential variogram with a fractal variogram is not a valid variogram. However, the product of an exponential covariance function with a fractal variogram *might* be a valid variogram. Please double-check and provide references if necessary.

5. Regarding the results: is it really desirable to get a correlation of 1 between measures and predictions? I would relate this to the fact that the coverage is 83%, which shows that the SVC is over-confident in the UG setting. Please comment.

Typos, etc.

=====

41: counties -> countries

51: there exist work -> there exist works
 206: remove the square bracket]
 222: Further is σ^2 -> Further, σ^2 is
 229: advice -> advise
 235 to 242: Kriging yields "predictions" not "estimations". Please change "estim*" to "predict*" everywhere. BLUE should become BLUP (lines 236, 240, 241)
 248: subcatcment -> subcatchment
 266: in 3 -> in Figure 3
 335: remove the dot in $s_i \sigma^2_y$, since it is not used in most cases later. (Also line 359 and line 406)
 358: are the scales -> the scales are
 383: I very much doubt that Norway is only 40 km wide (or less). Please double check
 430 Further is the variable x -> Further, the variable x is
 449: I did not understand the sentence. I think it needs to be rephrased.
 555: "... is defined as the probability that 90% of the observed values ..." should be "... is defined as the proportion of the observed values ..."
 584: In order to respect the order (first α , then x) I suggest to change the order between short ranged and long ranged. Hence, write "the spatial fields have picked up long range and short range processes, respectively".
 667: "only $x(u)$ or $\alpha(u)$ " -> "only either $x(u)$ or $\alpha(u)$ "
 669: "indicates that it for many study area might ..." -> "indicates that for many study area it might"
 672: "only $x(u)$ or only $\alpha(u)$ " -> one random field only.
 710: does not makes -> does not make
 734: that is -> that it is

In several places, there are sentences beginning with "Mark". This is unusual. I recommend using "Remark" or "Notice".