

Interactive comment on “Topological and canonical kriging for design-flood prediction in ungauged catchments: an improvement over a traditional regional regression approach?” by S. A. Archfield et al.

Anonymous Referee #1 - *referee comments in italics*

GENERAL COMMENTS The comparison of predictions from different models, or model blends, is very useful for practical applications in ungauged basins; however, I would include in this paper also some considerations concerning the inherent uncertainties of the estimation methods.

Let's first consider the sample uncertainty, that is implicitly taken into account by both the GLS and the kriging procedures, although no information about it is provided in the paper. I first suggest to include in the paper a summary of records length as a first approximation of such uncertainties, or directly a summary of sample variances (here, only the minimum record is reported at page 12198 line 8). This information is very important because, for short records it is likely to have a regional estimate more accurate than the empirical one, especially for large return period. For instance, this effect is considered by Gotvald et al (2009) who suggest, for practical applications at gauged sites, to compute the design peak flow weighting the empirical and the regional estimates, according to their variances. But, this effect can also affect the comparison of the residuals (e.g. fig. 3) where large error could be due to large uncertainty in the empirical value instead of bad predictions. It would be interesting to check the results of fig. 3 in light of the sample uncertainty.

Response:

This is an interesting point and worth investigating this effect. We will address this in the following ways: 1) Adding to figure 1 a box plot showing the variability of record lengths for the streamgauges in the study region, and 2) create scatter plots of the residual versus record length for the streamgauges. If there is a relation between the residual and record length, we will add this figure and accompanying text to the discussion.

The second point is the comparison of prediction variances of the different models at ungauged sites. The authors could check them and evaluate if better results are possible weighting different approaches (e.g. GLS and TK).

Response:

We do not examine the theoretical prediction variances that arise from each method, choosing to focus on the prediction residuals determined from the leave-one-out cross validation approach. Weighting of the approaches is a very interesting thought and, although beyond the scope of this manuscript, it is worth investigating a possible extension of this work. We will add this point to the discussion section of the manuscript.

SPECIFIC COMMENT page12204 line15-16: it is not clear why flood quantiles have been scaled by the factor $DrA^{0.65}$. I would add a short explanation of this point.

Response:

We will supply additional explanation on the motivation and reasons for the scaling of the flood quantiles with appropriate references.