Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-589-AC1, 2018

Hydrology and Earth System Sciences
Discussions

Interactive comment on "A geostatistical data-assimilation technique for enhancing macro-scale rainfall-runoff simulations" *by* Alessio Pugliese et al.

Alessio Pugliese et al.

alessio.pugliese3@unibo.it

Received and published: 15 January 2018

Referee comment:

The authors have presented a very well-prepared manuscript exploring the value of post-processing or "bias-correcting" daily streamflow simulations with independently derived regional information. By using an independently developed flow duration curve, the authors present an approach to customize macro-scale models to local conditions. The approach is very valuable. The manuscript is well-written and thorough. Below, I will provide some minor comments, but I see no impediment to swift publication.

Printer-friendly version

Discussion paper

Authors reply:

We thank reviewer W. H. Farmer for his valuable and precise comments. We are pleased of the appreciation for the overall study, and we hope that the replies to his comments presented below might resolve the issues arisen.

Referee comment:

Was the performance of GAE-HYPE always better than the performance of E-HYPE? At the top of page 10, the authors cite the best and worst improvements. Were there any sites that showed decreases in performance? If not, would you expect universal improvement in other regions?

Authors reply:

Yes, we obtained improvements in each one of the 11 pairs, however it is worth recalling that EHYPE was not calibrated in those sites, which results in locally poor performance of EHYPE. Overall, improvements are always to be expected when FDCs are well predicted by the geostatistical interpolator.

ACTION: In the revised version of the manuscript we will underline that all selected pairs show improvements in the performances.

Referee comment:

Was the degree of change in LNSE between GAE-HYPE and E-HYPE a function of the accuracy with which the TNDTK FDC was produced? That is, if the TNDTK FDC were poorly produced, it seems like the improvement in LNSE might be reduced or reversed.

Some exploration of this might be useful.

Authors reply:

The reviewer is right, the whole procedure relies on good prediction of regional FDCs in the study area. When FDCs are poorly predicted across the area, detriments of the geostatistical assimilation procedure are to be expected. Focusing on the 11 EHYPE prediction nodes, FDCs are actually very well predicted, with an average LNSE of 0.988 in cross-validation (with minimum and maximum values about 0.967 and 0.998, respectively) compared to the average performance in the whole study area, which is 0.898. Therefore, we cannot show here the effects of poor geostatistical interpolation of FDCs on the assimilation procedure.

ACTION: we will specifically highlight in the revised manuscript that biased regional model for predicting FDCs are expected to lead to poor performances of the proposed geostatistical assimilation procedure.

Referee comment:

What method of interpolation was used to map the residuals to simulated streamflows? On line 13 of page 9, it is stated that the TNDTK FDC is resampled to 20 points. So, if a simulated streamflow (E-HYPE) produced a duration that did not fall at one of the 20 resampled points, how was the residual estimated? That is, how was the eDC resolved between these 20 points?

Authors reply:

We thank the reviewer for letting us explain further this technical aspect. Observed FDCs are resampled to 20 equally spaced points, in the normal space, i.e. $d \simeq$ {0.0004; 0.0013; 0.0040; 0.0108; 0.0259; 0.0558; 0.1080; 0.1884; 0.2979; 0.4298; 0.5702; 0.7021; 0.8116; 0.8920; 0.9442; 0.9741; 0.9892; 0.9960; 0.9987; 0.9996}. The extreme values of the resampling scheme depends on the length of the shorter observed streamflow series in the dataset (i.e. 7 yrs), so that, those values extruding from such extremes are excluded in the resampled curves. Each EHYPE simulation covers a period of 30 yrs. with no gap, therefore we adopted the same resampling interval, with 20 points, used for the observed series. As a result, the residual duration curves are defined in the same 20 points of the resampling scheme and a linear interpolation is employed in between each data point. Finally, according to Pugliese et al. (2014), regional interpolation of TNDs, which is the weighting scheme for the curves as well, does not require any resampling scheme beforehand.

ACTION: in the revised version, we will add more information on the resampling strategy we used and we will clearly underline that a finer resolution of the curves could be necessary in specific applications, in which very high (low) duration are of particular interest.

Referee comment:

Does the simulated FDC resulting from the GAE-HYPE series match the TNDTK FDC? It seems as if it could (should?), by the nature of the method. This suggests that this method could also be considered as a re-scaling of the simulated streamflow distribution. Essentially, this means that the volumes from E-HYPE are discarded while the sequencing of E-HYPE (durations, relative values) is retained. I do not see anything wrong with this, but wonder if it is another way to think about the procedure. If accurate,

does this way of thinking provide any further insight?

Authors reply:

The reviewer rises a good point here.

ACTION: In the revised version of the manuscript we will add two panels in fig. 9 where we show FDCs of either empirical, E-HYPE, or GAE-HYPE next to their streamflow series counterparts, since we believe that the result is worth showing. We will also include in the discussion an analysis resulting from this comparison and possible insights on the aspect highlighted by the reviewer (the proposed geostatistical assimilation procedure discards simulated volumes, while retaining simulated sequencing).

Referee comment:

What was the significance of the changes in LNSE? Firstly, equation 10 can be simplified as the fractional change in root-mean-squared error of logarithms. This, can, of course, be interpreted as a percent. (Line 1 of page 11 uses percent, but the figure does not.) More importantly, the LNSE values could be compared in a pairwise test to determine if the improvement in LNSE is statistically significant (Wilcoxon). I imagine it is, but demonstrating this would provide stronger evidence.

Authors reply:

We thank the reviewer for these recommendations, which we will add in the revised version of the manuscript. We are aware that the proposed metric (eq. 10) can be simplified as the fractional change in root-mean-squared error of logarithms, therefore, we cannot see much difference in using one or the other, however, for the sake of

consistency, we would prefer to keep this assessment metric in terms of LNSE.

ACTIONS:

1. We will clearly highlight that the proposed index derives from the fractional change in root-mean-squared error of logarithms.
2. We will improve the consistency between notation and units in figures and text.
3. In addition, for each gauge density scenario, we will perform a pairwise test (Wilcoxon test) between LNSE values in order to determine whether improvements are statistically significant in terms of LNSE.

Referee comment:

How were streamflow values of zero handled? The authors measure performance in terms of the LNSE. What was the frequency of zeros? How were there logarithms taken?

Authors reply:

Although the alpine climate might be at risk for the presence of zero flows within the series, especially during the winter seasons, we observed no zero flows in the selected 11 sites with either empirical, EHYPE or GAE-HYPE.

ACTION: We will report such information in section 4.

Referee comment:

How many bins were used to discretize the variograms? Binning is described on line 27 of page 7, but it might be worthwhile to be explicit.

Authors reply:

We set the number of bins as the default setting of the "rtop" package, which is 1 bin per order of magnitude of drainage area. In our case, areas span over 4 orders of magnitude, therefore we used 3 bins for variogram pairs.

ACTION: We will add this information too.

Referee comment:

Editorial: On line 16 of page 1 the authors use "macro-scale" while line 17 uses "macroscale". Both are used throughout the manuscript; select one.

Authors reply:

Thanks.

ACTION: We will use "macro-scale" and change accordingly.

Referee comment:

Editorial: The figures seem to be out of order. Fig. 9 is mentioned after Fig. 7 and before Fig. 8.

Authors reply:

Thanks.

ACTION: We will change accordingly.

References

Pugliese, A., Castellarin, A., Brath, A., 2014. Geostatistical prediction of flow–duration curves in an index-flow framework. Hydrol Earth Syst Sci 18, 3801–3816. https://doi.org/10.5194/hess-18-3801-2014

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-589, 2017.