Hydrology and Earth System Sciences

Discussions

HESSD

Interactive
comment

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

W.H. Farmer (Referee)

wfarmer@usgs.gov

Received and published: 17 October 2017

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.

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

[Printer-friendly version]

[Discussion paper]

any sites that showed decreases in performance? If not, would you expect universal improvement in other regions?

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.

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?

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?

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.

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?

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.

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.

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

Finally, I sincerely thank the authors for their work and their well-written manuscript. It was easy to read, which made it easy to consider its technical merit. I look forward to their considered response and hope to have ignited some useful thought and discussion. If you would like any clarification on my comments, I strongly encourage you to reach out to me.

Thanks, William Farmer

Printer-friendly version

Discussion paper