

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
<https://doi.org/10.5194/hess-2022-281-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on AC1

Keith Beven (Referee)

Referee comment on "Why do our rainfall–runoff models keep underestimating the peak flows?" by András Bárdossy and Faizan Anwar, Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2022-281-RC2>, 2022

So to continue the discussion...

Just a couple of points arising from your comments. You say that: ***We don't use just any model, this model has precipitation inverted in order to match the observed flows.*** Inverted rainfall is not mentioned at all in the paper. Only that the gridded rainfall is produced so as to match the observed point rainfalls. Nor does it seem to be necessary as the "observed" flows are produced by SHETRAN (with its deficiencies and need for some clarifications I mentioned).

Secondly, I am not sure I mentioned parameter uncertainty anywhere in my comments (except to ask if the acceptable (or in your case calibrated) parameters might vary systematically with the input resolution). I am not sure that the hydrological community is obsessed with parameter uncertainty to the neglect of other sources, but in practical applications (where we would normally use ALL the raingauges available) it is a simple way of generating potentially behavioural models. Different model structures can easily be incorporated as well in a GLUE or Bayesian weighting framework. Input realisations are somewhat more problematic in that it is usually not obvious how to construct them (as you point out Kriging and its variance estimates depend on some rather strong assumptions, and gauges can be spaced more widely than raincells).

Your quick and temporary solution suggests a number of ideas but I am not sure they help much towards what we should do about the problem in practice. You ask for ideas – my analysis of the problem is that we know that interpolated rainfalls might be biased towards underestimation (especially where upland gauges are missing) and that traditionally model calibration (alone) or including rainfall multiplier parameters in the calibration (with a history back to the Stanford Watershed Model) has been a simple way of trying to compensate implicitly for that. That expectation of compensation extends to any errors there might be in the discharges observations (particularly at the highest and lowest flows), conditional on the model or models used. As noted above, allowing that an ensemble of behavioural models might result from that compensation is a further extension that can be used to produce some range of predicted outcomes that (again implicitly) reflect both sources of uncertainty. The analysis of runoff coefficients in the Inexact Science paper is another reflection of the joint uncertainties (there are a couple of follow-up papers in review that make use of that approach).

But, as you demonstrate, the compensation of calibration does not always produce an

underestimation of flows. BATEA etc have revealed how this is (as you note) an ill-posed problem, with resulting huge (and unrealistic?) variations in event to event rainfall multipliers when the compensation starts to allow for model structural deficiencies and consequent antecedent conditions from event to event. These are certainly all forms of epistemic uncertainties – or put it another way, even if we use all the raingauges available we cannot know what effect the interpolation will have on individual events. We can only make assumptions about what that effect might be. So to pose a hypothesis relevant to your paper: are there any systematic biases that can be detected in the reduced rainfall networks (or their effect on the discharges) that might be used to inform the model simulations (or more generally constrain their uncertainties).

This is a more challenging, but also more important, problem. I think I would approach it somewhat differently. I would eliminate the SHETRAN virtual reality – yes it is mass-conserving but that is not really relevant for practical applications for which we are trying to reproduce the observed discharge. I would try to assess the uncertainty in the observed discharges and allow for that in the model evaluations in some way. I would generate different rainfall realisations based on the samples of raingauges and compare both the rainfall biases (ie. no compensation by calibration) and the predicted discharge biases (ie. with implicit compensation by calibration) with results using the full network. For the kriging case the realisations could reflect the grid estimation variances (though there is still the issue of what minimum network numbers you need to determine a variogram).

The reduction in the variability of both inputs and simulated flows as the sample of gauges increases would still be revealed, surely. It might be considered as representative of what to expect in areas with a similar distribution of gauges with elevation (as demonstrated in the hypsometric curves in your reply). I am not sure it would apply in our Cumbrian catchments where we are much more deficient in higher elevation gauges (again there is a paper in review on this looking at different interpolation methods – though this was not extended to the type of sampling study you have done). So that might provide a range of potential outcomes for other applications with similar sampling densities but less gauges (the only problem being that we would not know where in that distribution that actual particular sample lies....though perhaps looking at the effect on the simulated outputs might help there even if after the compensatory effect of calibration – that is something that you could look at).

This might be a quick and more enduring way of going beyond demonstrating the problem towards what we might do about it..... I think we would certainly agree that there is no better solution than getting more raingauge (and discharge) observations but, certainly for historical data, we are often constrained to limited networks so we need some practical solutions.

KB