Hydrol. Earth Syst. Sci. Discuss., 10, C38–C42, 2013
www.hydrol-earth-syst-sci-discuss.net/10/C38/2013/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Comparative assessment of predictions in ungauged basins – Part 3: Runoff signatures in Austria" *by* A. Viglione et al.

Anonymous Referee #1

Received and published: 31 January 2013

The manuscript "Comparative assessment of predictions in ungauged basins – Part 3: Runoff signatures in Austria" by Viglione et al. compares how a process-based and a statistical modelling approach can predict various runoff signatures, characteristic for different time scales in a cross-validation approach for more than 200 catchments, spanning a significant gradient of hydroclimatic and geologic/pedologic conditions. It is a very interesting topic which is quite timely and illustrates the importance of not only getting one aspect of the runoff right in models, but ideally many more simultaneously to make sure that the entire characteristics of the system are well captured. The manuscript is very well written and structured, however, although the some of the methods used were common practice for several decades now, I am not sure if some details of these methods can, in the light of recent advances in this area, still be called "sound" in the strictest sense. Furthermore, although the manuscript starts off very

promising it does in the end not quite meet the expectations, as unfortunately a wider and deeper discussion of the results is not provided, i.e. what are the wider implications of the results? How do the results help to improve modelling strategies? Do the results contribute anything to how we can better predict? This makes the paper a nice standalone comparison of the predictive power of two modelling approaches in a certain environmental setting but the authors unfortunately do not try to generalize the results in some way or another. Having said that, I think the manuscript should definitely be published eventually, as it will make a nice contribution to literature if the authors gave some considerations to the detailed comments below.

(1) p.451, l.5: please provide a references for the theory behind cross-validation and possibly one example of successful application

(2) p..454, l.15-16: should maybe better read:"were the central part of the flow duration curves are particularly flat"

(3) p.456, l.5ff: this part raises some questions and it is in fact my biggest concern. Unfortunately, the reader is not given any detailed information on the calibration strategy and on how the parameterizations used in the end were chosen. Were the models calibrated on the hydrograph or on all signatures? Which objective function(s) were used – the same as for the rest of the analysis? If only one objective function was used (e.g. Nash-Sutcliffe Efficiency), how can it be expected that the predicted catchments perform well with respect to the other signatures? If we want to ensure space-transferability, shouldn't we first make sure that our parametrerizations are transferable in time, i.e. to make sure that the chosen parameterizations give adequate results in an independent test (validation) period as stressed by Klemes (1986), Andreassian et al. (2009) and in an innovative approach recently addressed by Gharari et al. (2013)? If such tests are not carried out, we run the risk of choosing the "optimal" parameterization as a result of a mere mathematical fitting exercise ("mathematical marionettes" as termed by Kirchner (2006)), rather than on basis of an adequate process representation, thus leading to limited predictive power. Similarly, Beven (2006) and Andreassian

et al. (2012) noted that frequently the most suitable parameterization for a model in a given catchment is “sub-optimal”! Of course, it has a long tradition just to calibrate models and declare the parameterization with the highest performance the most suitable one – BUT: we could do SO MUCH BETTER! This is in my opinion especially true in an analyses as the presented one, were prediction is at its core. Maybe this has all be done by the authors, but then I think it should be prominently commented on as these are crucial details. Further, although there is, for practical reasons, in principle no strong argument against one-fits-all modeling approaches in studies like this one, I could, however, imagine that predictions could be improved by at least introducing 6-7 different model classes depending on the dominant runoff pattern (similar to what was recently shown by Ye et al., 2012). Having said that, I do not necessarily want the authors to redo their entire analysis, although this would a fantastic effort to increase the relevance of the results, but I would at least like to encourage them to give the above mentioned concerns some consideration and discuss the limitations of their methodology accordingly.

(4) p.457, l.17 & l.23: please adjust table numbering in the correct order

(5) p.457, l.23, Table 1: maybe more instructive to show as boxplots rather than in a table

(6) p.460, l.15: why 0.368? please explain where this value comes from

(7) p.464, l.10ff: comparing the results to catchment characteristics is a very important and instructive part of the paper. However, I was surprised that the analysis was not done with some more depth, for example also including some measures of catchment organization (e.g. drainage density, or the the inverse of that, the average flow path lengths; average flow path gradients; some indicator of soil types; or height above nearest drainage – HAND (Renno et al., 2008) as proxy for the hydraulic head). This could have given some insight in which types of catchments predictions work better or worse.

(8) p.467, l.5: maybe better “the bars contain the interquartile range of the values...”

(9) p.468, l.16-17: may be add Euser et al. (2012) as example of how this could be done

(10) Please comment in detail on the wider implications of the results: How do the results help to improve modelling strategies? Do the results contribute anything to how we can better predict in general?

(11) Figure 2: ellipse and arrows not quite clear. Please explain the connections. For example, why winter in (d) is connected with snow in (b) but not with snow in (f)?

References

Andreassian, V., Perrin, C., Berthet, L.,LeMoine, N., Lerat, J., Loumagne, C., Oudin, L., Mathevet, T., Ramos, M., Valery, A., Crash tests for a standardized evaluation of hydrological models, Hydrol. Earth Syst. Sci. 13, 1757-1764, 2009

Andreassian, V., LeMoine, N., Perrin, C., Ramos, M., Oudin, L., Mathevet, T., Lerat, J., Berthet, L., All that glitters is not gold: the case of calibrating hydrological models, Hydrol. Process. 26, 2206-2210, 2012.

Beven, K., A manifesto for the equifinality thesis, J. Hydrol. 320, 18-36, 2006

Euser, T., Winsemius, H.C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., Savenije, H.H.G., A framework to assess the realism of model structures using hydrological signatures, Hydrol. Earth Syst. Sci. Discuss. 9, 12989-13036, 2012

Gharari, S., Hrachowitz, M., Fenicia, F., Savenije, H.H.G., An approach to identify time consistent model parametersa: sub-period calibration, Hydrol. Earth Syst. Sci. 17, 149-161, 2013.

Klemes, V., Operational tresting of hydrological simulation models, Hydrol. Sci. J. 31, 13-24, 1986

Renno, C.D., Nobre, A.D.,Cuartas, L.A., Soares, J.V., Hodnett, M.G., Tomasella, J, Waterloo, M.J., HAND, a new terrain descriptor using SRTM-DEM: Mapping terra firme rain forest environments in Amazonia, Remote Sens. Environ. 112, 3469-3481, 2008

Ye, S., Yaeger, M.A., Coopersmith, E., Cheng, L., Sivapalan, M., Exploring the physical controls of regional patterns of flow duration curves – part 2: role of seasonality and associated process controls, Hydrol. Earth Syst. Sci. 16, 4447-4465, 2012

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 449, 2013.