Hydrol. Earth Syst. Sci. Discuss., 8, C2072–C2079, 2011
www.hydrol-earth-syst-sci-discuss.net/8/C2072/2011/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Catchment classification: hydrological analysis of catchment behavior through process-based modeling along a climate gradient" *by* G. Carrillo et al.

P. A. Troch

patroch@hwr.arizona.edu

Received and published: 9 June 2011

Reply to Interactive Comments prepared by Stefan Uhlenbrook (referee)

"The paper reports the application of a scale-aggregated, process-based model to a variety of catchments with different hydro-climatic conditions. It is an interesting, well-written paper that applies a very good step-wise model parameterization approach. The model is too a large extent a combination of existing (previously published) modules/routines that were nicely put together in a coherent framework".

We thank Dr.Uhlenbrook for his nice comments about our work. It is important to note

that the purpose of this study was not to apply a particular hydrologic model to a number of catchments, but to use a consistent modeling framework to analyze hydrologic observations in an attempt to capture the functioning of catchments across a climate gradient. The ultimate objective of our study is to use such strategy to explain apparent hydrologic similarity revealed from top-down catchment classification methods, as discussed in Sawicz et al. (this issue).

"Alternative model structures would be equally plausible (page 4601, line 10) and the model structure was not really based on detailed field investigations. However, the same model structure was applied to all catchments, assuming that all used modules/routines are suitable for all catchments (with a catchment specific parameterization, of cause). Yes, this has been done in many other inter-site comparison studies, this is a classical assumption".

We agree that applying the same model structure to different catchments for inter-site comparison is not novel and has been applied in many other studies. However, since our modeling framework is modular, not always are all available modules applied to all catchments. A good example if the snow accumulation and snow melt module that was only relevant in 6 of 12 catchments along our climate gradient. Also, the subsurface lateral flow modules are not always needed for all catchments. For example, Guadalupe catchment in Texas (GUA) does not require a deep aquifer module since all of the observed baseflow dynamics can be rproduced by the non-linear hsB module (see also discussion below).

"However, I found this in particular tricky for the subsurface flow components, where the model structure defines 2 aquifer systems: a perched and a deep system. Whereby the model was parameterized such that both contribute to the MRC and the perched aquifer (representing an interflow component) is contributing to the early part and the deep system mainly to the baseflow. But, why can the authors assume that 2 systems exist in all catchments? Is the geomorphology in all catchments (with quite different topography, geology and soils!) suggesting such a model structure and related hydrological response? I understand that no detailed process investigations were done in these catchments. One alternative model structure (many others could be suggested) would be to have only one groundwater systems with a non-linear, stronger response if the water table is higher ('transmissivity feedback'). This behavior has been described for instance in many till soil covered catchments in Scandinavia and North America (cf. papers by Kevin Bishop, Jan Seibert, Allan Rodhe etc.). Or, how about the importance of a riparian groundwater system (as found very important in many catchment in the US, see papers of McDonnell, Peters, Hooper etc.) that sustains the base flows?"

The referee is correct that we did not attempt detailed process investigations in these catchments. Since all catchments are large (> 1,000 km2) such detailed field investigations would have been very difficult to perform in a meaningful manner and was beyond the scope of the study. However, we do not a priori assume that both systems exist in all catchments. Our conceptualization of a two aquifer system was informed by streamflow observations in these catchments. From the MRC it was clear that in most cases there was a difference between the early time response and the late time response. The late time response converged in almost all cases to a linear reservoir response (straight line of MRC on semi-logarithmic plot) and suggested baseflow generation with relatively long time scales (see Table 3, the deep aquifer time scale (reservoir constant) is always larger than 2 weeks). Such long time scales suggest deeper groundwater contributions to streamflow and exclude contributions from a shallow perched system. Such perched groundwater system will likely exist only for a few days/weeks after cessation of rainfall and explains the non-linear dynamics in the early time response of the MRC. In one cacthment (GUA) the entire MRC dynamics were non-linear and could be captured accurately with the non-linear hsB module.

We agree with the referee that other conceptualizations are possible, but the same arguments against using any of the other possible conceptualizations are valid. Why would the transmissivity feedback mechanism be more universal than our conceptualization? Such question is impossible to answer without detailed field investigations,

and is a bit beyond the point of the purpose of the paper.

The riparian aquifer conceptualization is not necessarily in contradiction to our approach and may well be the reason for the observed MRC dynamics. Brustsaert (2005) clearly shows that riparian aquifers can be represented by Dupuit-Forccheimer theory that often can be further reduced to linear representation of the flow dynamics. Such theory is completely in line with our approach.

"I think the title of the paper promises too much, as this paper mainly summarizes the analysis of catchment behavior through model application in a range of catchments. However, I found the emphasis on catchment classification too strong, as this comes only out in the discussion".

The part of the title "Hydrological analysis of catchment behavior through process-based modeling across a climate gradient" completely captures the content of the paper, so I guess the referee only refers to the "Catchment classification" part of the title. The reason for this part of the title is that our work is part of a larger project on catchment classification, and it provides a bottom-up modeling approach to try to understand the reasons of regional hydrologic similarity, as discussed in the introduction and discussion sections. As a stand-alone paper, the referee is possibly right that the title suggests too much, but as part of an accompanying paper (Sawicz et al., HESSD, same special issue) the title is important to explain the context of the work.

Specific comments:

1. "The abstract is quite long and has many introduction parts. I found the last sentence too speculative based on the presented results".

We agree and will shorten the abstract by removing some of the introduction points that are also discussed later in the manuscript. The last sentence refers to the results shown in Figure 12. We find these results quite intriguing and cannot explain them without invoking some type of co-variation of climate, soils and vegetation. It is well known that soil and vegetation characteristics vary predictably across climates (see our discussion for references) and the fact that we do not take advantage of such co-variability in catchment hydrology is surprising. We hope to inspire our colleagues to begin such explorations in order to discover signatures of co-evolving structures in our landscape in the hydrological response. The results in Figure 12 are robust and cannot be explained by looking at climate only, so they are important to highlight in the abstract.

2. “Section 2: All variables should be defined with units”.

Throughout our model development we consistently use SI units, but we agree that it makes it easier to read if the reader is reminded of the units of the variables. We will modify Section 2 accordingly.

3. “I found it clearer if the units of the water balance parameters (page 4607, line 10-13) are give as fluxes (what they are!) and not as storage volumes; i.e. mm/a instead of mm. Use d as a unit instead of day”.

I think there is little room for confusion when reading the entire sentence: “The mean annual precipitation ranges from 750mm to 1500 mm, and the mean annual potential evapotranspiration ranges from 1500mm to 700 mm”, but we’re happy to accommodate this comment. We will change “day” into “d”.

4. “Model parameterization: - Why are the snow temperature threshold vales so high (1-3 C)? Did you not correct the temperature input for the mean elevation of the catchment? - I have never seen so high melt rates (degree day factor: 10-15 mm/d/C), and it not think that this is physical possible. I am aware of values in the range of 2-4 mm/d/C in forested and mountainous catchments, and know that values can go up to 8 mm/d/C in flat and open catchments, but the suggested parameters in the paper seem unrealistic”.

The referee is correct that we only use the basin-average and daily averaged tempera-

ture to force our snow melt model. Therefore, the listed threshold values are high and reflect the fact that snow most likely occurs at the highest elevations of the catchments (although none of the catchments have really strong elevation ranges) and thus the local snow melt inducing thresholds are probably closer to zero. Since this is a calibration parameter in a lumped modeling procedure this is ok. As for the high melt rates, first I’d like to observe that not all calibrated melt rates are as high as mentioned above (three of six catchments have melt rate of 0.5, 1, and 5 mm/C/d, respectively). None of these catchments develop a seasonal snow pack but rather have intermittent snow accumulation and melt periods. During snow melt it is very likely that the catchments also receive rainfall which adds energy to the melting snow layer. Since we do not model the snow melt process from an energetic point of view, such advected energy into the snow layer must be absorbed by the average snow melt rate and thus explains the high rates in three of six catchments. Also, since we only use daily averaged temperature to drive snow melt the model does not account for higher maximum temperatures at and beyond noon. Again, due to the simplicity of the snow melt model this is a calibration parameter and should be interpreted with care.

5. “Figure 2 and related discussion: I found the term Deep Aquifer Fraction confusing as it is not a fraction but a flux. I cannot believe that the R2 is 0.99 looking at the figure, considering that there is also a log-scale. How was that calculated?”

We agree with the first comment and will replace “fraction” with “component”. As for the reported R2 values, those were computed using only the tails of each of the short recession curves (colored curved lines in Figure 2), and since all tails line up to the straight line we obtain high R2 values. I don’t think the high R2 values are the point here, but the fact that the MRC analysis allows us to accurately estimate the linear reservoir constant given the convergence to a linear relationship on semi-log scale.

6. “All model results are only shown in a very aggregated way (monthly, annual values, RC matches etc.), but to really assess the model goodness 2D-plots Qobs vs. Qsim on daily scale would be more helpful. Generally, model efficiencies of about 0.5 after

extensive calibration is not really impressive assuming that the input data sets are good".

We respectfully disagree with this comment. We do report modeling results at daily time scales (which is the modeling time step and thus the shortest relevant time scale) by providing the flow duration curves (FDC) and how the model is able to reproduce those. We ourselves comment in the paper that a NSE of 0.5 is indeed not impressive but we did not attempt to optimize such statistic. The fact that the model reproduces the FDC in all cases very well (Figure 6 and reported mean AE in mm/d) indicates that all frequencies in the hydrograph are reasonably and adequately represented. The average NSE values indicate that in many cases we probably miss the exact timing of high flows (NSE is notoriously biased towards high flows), but since our modeling objective is not to reproduce the observed hydrograph but rather the different time scales of response we feel that matching the FDC is a better indicator of the appropriateness of the model to investigate hydrologic behavior of these catchments. Of course, in an ideal world we would also match the hydrographs exactly, but if optimizing NSE would compromise parameter interaction, that is a too high price to pay for the purpose of our study. That is why we developed the stepwise procedure to inform model parameter values rather than brute force optimization that would not avoid ending up in a parameter space that would misrepresent particular components of the system.

7. "Figure 8-11: Why is the number of catchments varying? What were the criteria to include/exclude a catchment? Giving a R2-value with 4 digits is really too much considering that the number of data sets (n) is often only 6".

The reason why we have different numbers in some plots (6 vs. 12) is that in many cases the snow-catchments did not show clear scaling relations, but the no-snow catchments did. In some cases the snow catchments exhibited scaling behavior with different time scales or dimensionless numbers and thus we separated them from the no-snow catchment. We have explained that in several instances in the text as well as in the captions. We agree that reporting R2 to 4 significant digits is not good practice

and we will fix this in our revised version of the paper.

Peter A. Troch June 9, 2011

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 4583, 2011.