Hydrol. Earth Syst. Sci. Discuss., 9, C4039–C4047, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C4039/2012/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Parameterizing sub-surface drainage with geology to improve modeling streamflow responses to climate in data limited environments" *by* C. L. Tague et al.

C. Luce (Referee)

cluce@fs.fed.us

Received and published: 31 August 2012

This is an interesting modeling exercise, demonstrating some useful insights about models. The experiment is sound, and it may be important to developing discussions in hydrology, but the current focus of the introduction and discussion relates more to the earlier papers of these authors than this particular modeling exercise. It does not do the experiment justice. I would recommend some rewrite to better connect to the general field of hydrologic modeling and scaling to capitalize on the information they have generated. I have three conceptually specific suggestions and two specific

technical comments.

(1) Reconciling purpose and content

The paper does not deliver on the purpose stated in the introduction, but the outcome may be of somewhat greater utility than implied. Interpretation of their results would be substantially improved with a more accurate statement of purpose and discussion of the literature relevant to the issue they addressed. This seems to be a paper on hydrologic scaling that does not discuss the literature on scaling and modeling; that does not even use the word "scaling."

The discussion of purpose in the introduction, starting from line 9 of the manuscript (p. 8667) and proceeding to line 9 of the following page, relates to the broad use of large scale hydrologic models in assessing climate change effects on water resources and ecosystems. The concern is that parameters derived from calibration to larger rivers are applied too glibly to assess responses in smaller contributing streams. They state

"In this paper we demonstrate the potential error in applying calibrated parameters across an entire watershed based solely on a larger order stream, and present a relatively simple strategy for parameter transfer based on geologic similarity."

The general idea is that one should not expect to simply scale a hydrograph (or a model) for a small stream within a basin simply from knowing the hydrograph of the larger basin in which it sits (or the calibrated model parameters of that basin). It is an idea somewhat generally accepted as a self-evident truth, and methods for downscaling information have been a subject of considerable discussion within the hydrologic community (see e.g. Blöschl and Sivapalan, 1995 and citing papers). This is not, however, the question that this paper answers. While the introduction states that the critical issue being addressed is disaggregation of streamflow estimates from coarse resolutions and sparse observations in large watersheds to the network of tributaries, the paper actually tests an approach to aggregation of parameters for a watershed as a method for hydrologic prediction in an ungaged basin.

The paper demonstrates that heterogeneity of bedrock within a basin can be successfully modeled at the basin outlet considering simply the fractional representation of two strongly contrasting lithologies. The authors did not ultimately demonstrate that the spatial mapping of geologic facies in that basin led to better predictions of tributary streamflows within the basin, which would more directly support the above purpose statement. They did, however, show that different geologies in different sub-basins of the larger McKenzie watershed were reflected in different parameters, which could be shown to be more-or-less shared in common among watersheds with similar geology. They also showed this in earlier work, which they cited.

They did not demonstrate a means to scale from knowledge of a larger watershed's parameters to the parameters of a catchment nested within it; they demonstrated how independent knowledge from basins with unadulterated end-member lithologies could be applied to a basin with a mix of those lithologies. In so doing they provide an example of a sub-element scale parameterization for geologic heterogeneity that builds off of their previous work showing how great the differences in hydrology are for these different facies. The effects of sub-element scale in strongly contrasting soil and bedrock properties is potentially a very difficult issue, and might be illustrated by Figure 2, which shows that there is some organization within the watershed with the deep-groundwater lithology placed upstream of the no-groundwater (in the model) lithology. For instance, examples have been provided before about the importance of the arrangement of soil properties on hydrology at field to plot scales (Springer and Cundy, 1987, Zhang and Cundy, 1989).

Their contribution describing a fairly simple approach to identifying the parameters of a very simple parameterization for sub-element scale heterogeneity in bedrock may well be a more important contribution than the one they claim, in part because that claim is already better supported by their earlier work. Most hydrologists looking for high resolution hydrologic information look for some basis to resolve the differences between units at the scale of interest. In earlier work, the authors demonstrated that bedrock

geology is sometimes an overriding driver for hydrologic heterogeneity. The new contribution in this paper would be much more apparent with some introductory text on how sub-element scale heterogeneity is generally applied in hydrologic modeling and specifying alternative hypothetical approaches to modeling within-watershed heterogeneity in bedrock. One of the strengths of the paper is the contrast of a simple independent end-member mixing model within the context of a self-described physically based model filled with detailed conceptualizations of physics. There is probably more generalizable learning available from this paper than is currently presented. Some broader context, such as might be provided by Sivapalan et al. (2003) and articles therein would probably strengthen the impact of this work.

There are logical crosswalks between aggregation and disaggregation approaches, whereby the logic for aggregation implies some characteristics of the disaggregation problem. In this case, the mixed-member model implies that contributions from each geologic unit could be treated independently. However, this just means that each sub-watershed could be modeled independently using the geology map to provide drainage parameters, essentially no different from any other potential modeling unit. This does kind of highlight the fact that the parameterization has only been demonstrated for one watershed. It is still important to the discussion in this field though because 1) it works very well in that one watershed without calibration (other than for total runoff and maybe precipitation) and 2) it is very simple. If it is an important idea, replication will follow.

(2) Reconciling "reductionist rhetoric" with important accomplishments in upscaling

The paper's orientation toward downscaling is to be expected in light of the authors' contributions on the importance of lithology. However, absent the context of sub-element scale heterogeneity, the argument seems to veer philosophically toward supporting reductionism, which seems like the opposite of what they end up accomplishing. The introduction asserts:

"However, calibration based on gauges from a larger order watershed does not necessarily apply to the diversity of lower order streams within that basin, or similarly, parameter transfer from neighboring watersheds may not be appropriate."

Admonitions are not constrained to the introduction, and the discussion states:

"Our results show that if predictions are needed in basins where calibrations have not been explicitly conducted, great caution needs to be exercised if these uncalibrated basins reflect different geologies than those where calibrated parameters were derived. Furthermore, in basins with mixed lithologies, which are the norm for larger watersheds, calibrated parameters need to be developed across the full range of drainage efficiencies and cannot be confidently applied simply based on basin proximity."

A concern is that this and the earlier statement of purpose could be interpreted to say that (some level of) spatially distributed modeling is necessary for accurate representation of climate change effects in large watersheds. It would be a stance that seems odd and dated considering the depth of related literature on reductionist modeling approaches in the early 1990s (reference again Blöschl and Sivapalan, 1995, and Sivapalan et al., 2003). It is also at odds with a literature on scaling that would suggest that simply increasing the spatial density of information on hydrologically important characteristics may not be necessary to improve hydrologic modeling. Despite these statements, the authors have generated insights that assist in modeling with large elements.

Ultimately, it seems like these statements are an unnecessary aside. What the modeling exercise seems to say is that if you have information about strongly contrasting geology mapped to finer resolution than the model elements you are using, you can substantially improve your models performance by modeling the geologic units independently and summing their outputs. This philosophically parallels modeling approaches described by others earlier, e.g. the Probability Distributed Approach (Moore, 1985) or even the VIC model (Liang et al., 1994).

It may be beneficial to recast the argument in terms of the value of the additional in-

formation, or in terms of how high resolution (relatively speaking) geologic information could inform coarse resolution hydrologic models. This frames it in terms of the value of added information relative to its cost (for acquisition of the information and integrating it into a model). Models relate information that we obtain at some expense to information on which we place higher value. It seems that what they demonstrate is that high resolution geologic mapping (which mostly exists) can be accounted for in larger model elements with relative ease. This framing also broadens the value of the findings in this paper to other model developers. A useful reference in this context is Wenger et al. (2010) who show that the VIC model performs poorly for some metrics in areas where deep groundwater storage and transfer is not accounted for. Perhaps a parameterization similar to that developed for RHESSys would be advisable for that situation.

(3) Provide enough information for reader to gain a more general lesson from the modeling exercise

It would be very useful to the interpretation and depth of learning that can be obtained from this experiment if the authors provide some means by which the reader can gain a better intuitive grasp of what the end member and mixed models do. One aspect of this is providing a few more specifics and mathematical formalism to the description of the RHESSys drainage model. I was not entirely clear for example what the deep drainage model is based on. Is it $dV/dt = gw2*V$ (with V as some metric of storage in the deep groundwater)? Also, to interpret the parameter values listed in the table at the end, I needed to go search for model documentation.

The addition of the GW component makes it look like a slow-water/fast-water description of the storage. This is similar to the HBV model but different in details. With the addition of the mixing model, now there are two "fast-water" components and a "slow-water" component. Does this provide a better fit ultimately? Is there a parameter set for the simple slow/fast model that would fit the SF McKenzie, or does the combination of the three different "stores" provide something fundamentally unique? Do I correctly perceive the mixed-end-member model as having 3 stores instead of two? If so, I think I'm seeing the WC component as sort of a "medium speed", but that is a bit fuzzy to grasp as well. Parameters for the surface layer in WC are K=58 m=0.8 versus K=34 and m=5.1 for HC. The K is smaller for the HC, though if units are cm/s they are both quite substantial. The decay of K with depth (the m parameter) is much greater for the HC soil. So the HC soil has a comparable and very large hydraulic conductivity near the surface, but has about 1% of the conductivity of the WC soil at 1m depth. By 2m depth, the HC soil has a very small hydraulic conductivity (< 1 mm/hr). How does this mesh with the deep groundwater conceptualization? I generally associate high m with shallow soils . . . which is not the case in the High Cascades. This contrast of the surface layer hydrology would suggest that the HC soils have generally lower conductivity and move water more slowly than the WC soils. However the 30% leaving to groundwater is pretty substantial? Are there tradeoffs in parameters, whereby the availability of the deep groundwater algorithm to handle slow flowing water allows the upper soil to behave in a way that releases water rapidly to streams? In contrast the No GW soils must moderate between the two speeds?

I had a question about how well the SF McKenzie could be calibrated using the two-store model, but on reflection of how much simpler a linear combination of outputs from two end-member is, particularly with no calibration being necessary, I'm not sure the question would be relevant to the main point of the paper. However it might be useful in describing what the effective representation of the mixed-end-member model looks like. Figures showing how the end-member parameters performed for the SF McKenzie were useful for partially understanding how the different model pieces contributed to the overall hydrograph, and seeing how they compare to a calibration of the 2-Store model would similarly help. Even if, or maybe especially if, the 2-store model does not calibrate well for SF McKenzie, it would be informative to the broad field of modeling storage effects in basins. I think in particular reference to Kirchner (2009), contrasting storage-discharge relationships from lumped versus spatially heterogeneous contexts could be very interesting.

Technical comments:

It would be acceptable, if not helpful, to drop the reference to the W2 watershed. It is not really clear how that evidence is used to help the arguments in the paper. It becomes a distraction to the reader. The difference between runoff and observation there have some similarity to the difference for the SF McKenzie as well, and leads to a bit of flipping back and forth to see consistency in how the discrepancies are dealt with. Feel free to mention that it's use was explored but don't include later discussion about it.

Pg. 8679 Lns 16-24 with reference to Figures 6-7b. In the text the discussion mentions Pearson correlation coefficients between 0.6 and 0.9, but the range seems to be closer to 0.4 to 0.9, which are the two end points, with most being in the general range of 0.5 to 0.7. In terms of fractional variance explained by climatic variation as expressed through the model, that yields 25% to 50%, which isn't really all that strong. On top of this, the slopes are not all the same, and there is notable bias. It would be interesting to see the NSE on these. Despite a statement about there being no significant difference between observed and estimated spring fractions, there is substantial bias in some of the basins. A straight up ANOVA between the two as in the comparisons shown in Figure 7a, buries the difference between modeled and observed in interannual variability in the actual values. Even if the differences are slightly non-normal, a paired t-test would be better. Alternatively, a comparison of the correlation coefficients and the Nash-Sutcliffe Efficiency is another approach to identify bias, though somewhat qualitative (see Wenger et al., 2010 for discussion) . . . or one could just draw the one-to-one line through the graphs in Figure 6, like I did. Because of the varying bias, I'm not sure that the placement, or even inclusion of the observed bar in Figures 7a is appropriate. At least Isolate observed on the left hand edge. Also, there are some differences in slopes between gages, so some of the differences between the white, yellow, and red bars are due to differences in model sensitivity in the particular basins. I did not really see so much of a difference in slope between WC and HC basins that I would guess that Figure 7b has errors, but it should be rigorously checked because of how flat the relationship is for Horse – one of only two HC streams.

References:

Blöschl, G., Sivapalan, M., 1995, Scale Issues in Hydrological Modelling: A Review, Hydrol. Process., 9, 251-290.

Kirchner, J. W., 2009, Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, Water Resour. Res, 45, W02429, doi:10.1029/2008WR006912

Liang, X., Lettenmaier, D. P., Wood, E. F., Burges, S. J., 1994, A Simple Hydrologically Based Model of Land Surface Water and Energy Fluxes for General Circulation Models, J. Geophys. Res., 99, 14415-14428.

Moore, R. J., 1985, The probability-distributed principle and runoff production at point and basin scales, Hydrol. Sci. J., 30, 273-297.

Sivapalan, M., Blöschl, G., Zhang, L., Vertessy, R., 2003, Downward approach to hydrological prediction, Hydrol. Process., 17, 2101-2111, DOI: 10.1002/hyp.1425.

Springer, E. P., Cundy, T. W., 1987, Field-scale evaluation of infiltraion parameters from soil texture for hydrologic analysis, Water Resour. Res, 23, 325-334.

Wenger, S. J., Luce, C. H., Hamlet, A. F., Isaak, D. J., Neville, H. M., 2010, Macroscale hydrologic modeling of ecologically relevant flow metrics, Water Resour. Res., 46, W09513, doi:10.1029/2009WR008839.

Zhang, W., Cundy, T. W., 1989, Modeling of two-dimensional overland flow, Water Resour. Res., 25, 2019-2035.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 8665, 2012.