Review for HESS of

Estimating catchment scale groundwater dynamics

Thomas Skaugen and Zelalem Menghistu

25 January 2016

In this paper, the authors compare the suitability of two flow estimates from a conceptual hydrological model, taking two approaches to parameter estimation (and introducing some changes to a previously-published the structure of the DDD model). In one approach, the subsurface parameter value is obtained using calibration to flow records, and in the other, the subsurface parameter value is instead obtained from hydrograph recession analysis (other parameters were held fixed at values found in a previous calibration). Using 73 catchments across Norway, they find that there is no loss of performance when using recession analysis for parameter estimation, instead of typical calibration. The authors interpret this as evidence that the model structure is suitable for their task.

Overall, I did not find this to be convincing. Replacing some of the calibration by recession analysis effectively breaks the parameter estimation into two steps, but it does not reduce reliance on extracting information from the hydrograph. The main justification for the authors' conceptual model is that it produces good hydrographs for their study area without having many parameters. It is well-known that in relatively humid areas, simple hydrological models are often sufficient.

There are many low-dimensional time-stepping conceptual hydrological models, each different in the details of how they conceptualise hydrology, but in the end most include (i) evaporation which depends on soil moisture (ii) runoff which comes from a set of one or more linear or nonlinear reservoirs of various timescales, in parallel and in series. The authors have a particular way of deriving their model structure, including the way subsurface flows are drawn from different levels in the soil (e.g. Fig 12), with a particular distribution in the vertical. However this remains a hypothesis which is not directly tested. If there is evidence to support the authors' hypothesis, that would be of more interest, e.g. spatially-distributed monitoring of lateral water flow through different soil horizons, water tracer data implying the that contributions to river flow come from particular levels in the soil column.

I think the authors need to make a stronger case why their particular conceptualisation is strongly supported by the available evidence. Without this, it is not clear why it is important to test changes to one particular component of the model (which I see as the authors' main objective in the paper).

<u>Specific Comments</u>

1. P11134 L1 "and an unsaturated zone with volume D (mm), called the soil water zone. The actual water volume present in the unsaturated zone, D, is called Z (mm)." It is hard to understand the difference between D and Z from this text. I think it would be clearer to start this phrase "and an unsaturated zone with capacity D (mm) …"
2. P11134 L20 "Experience using the DDD model shows that the subsurface water reservoir M largely controls the variability of the hydrograph." I think it clearer to say "the subsurface water capacity parameter M"
3. P11134 L21 "Low values of M increase the amplitude of the hydrograph, since the entire range of celerities is engaged, and vice versa." This sentence is impossible to interpret without a description of the role that celerities play in the model.

4. P11136 L6 “according to a linear reservoir in recession with runoff coefficient ϑ” It seems confusing to call this a runoff coefficient. That term is generally reserved for a ratio of runoff to precipitation. This parameter seems more like a rate constant, since it presumably has units of 1/time.
5. P11136 L7 “The ratio between consecutive values of runoff, Q(t + 1)/Q(t)” Do the authors mean
   $Q(t + \Delta t)$ rather than Q(t + 1)?
6. P11136 L14 Equation 6 indicates that ϑ is dimensionless, but since Q is presumably a flux (mm/day) and S is a storage (mm), the linear reservoir equation $Q(t) = \vartheta S(t)$ indicates that ϑ has units of 1/day. This inconsistency needs to be resolved, as ϑ is closely linked to Λ and κ, which are pure ratios.
7. P11136 L20 “This brief discussion on the distance distribution and linear reservoirs is relevant because it suggest that if a catchment exhibits an exponential distance distribution the linear reservoir comes as a natural choice for modelling the interaction between hillslopes and the river network.” This is true only if hillslope celerity is effectively constant. A rather strong assumption given the nonlinearity of some soil water processes! If I understand correctly, the DDD model has a celerity which varies with water storage, i.e. the effective celerity is not constant. Thus I am unclear why the discussion on linear reservoirs is seen as especially relevant.
8. P11137 L15 “The parameter Λ is thus the slope per Δt of the recession in the log-log space”. This is not correct. If one plots $\log(Q(t+\Delta t))$ against $\log(Q(t))$ (the only log-log space I can see in the paper), the slope of the line is unity, and the offset is Λ. Isn’t Λ the recession slope when log flow is plotted against (linear) time?
9. P11143 L8 “We will test the performance of the new calibration-free formulation for the subsurface.” It seems overstated to call the new approach calibration-free, because calibration-free is often interpreted to mean that no flow record is required to estimate the parameter. Parameter estimation by recession analysis still requires measured streamflow. The new approach differs in that parameter estimation does not use traditional hydrograph-matching using a time-stepping model, but instead uses recession analysis.
10. P11146 L16 “as we have no way of actually knowing the true empirical distribution of storage at the catchment scale” It would be entirely possible to install a spatially distributed monitoring network which measured the changes in unsaturated and saturated storage at multiple locations. If a stratified sampling approach was taken when selecting sites, then this could be used to estimate catchment-scale storage. This may not be practical for the authors’ specific situation, but it is possible, and has been done in other situations.
11. P11147L7 “The estimation of θM is, however, no longer needed.” But surely the calibration has merely been replaced by recession analysis to determine the parameter?
12. P11148 L21 “Figure 13 shows simulated storage S, plottet against observed runoff Q, for two catchments of different size (50 and 1833 km2). It is quite clear that the relationship between Q and S is not single valued.” Some of the reason for the scatter could just be that the model is not well correlated with the observations? Why not plot simulated storage against simulated runoff?
13. P11150 L9 “An important contribution of the new formulation is that its parameters are estimated solely from observed recession data and the mean annual runoff (i.e. not through calibration).” To me, it is still calibration (albeit multi-stage calibration), if it is necessary to use measured flow to estimate parameter values. If instead the parameters could be reliably estimated from catchment and climate characteristics, that would be of great interest.