Response to referee #1 on review of “Estimating catchment scale groundwater dynamics from recession analysis- enhanced constraining of hydrological models”, by T. Skaugen and Z. Mengistu

First of all, we would like to thank the referees and the editor for taking their time to read closely and comment thoughtfully on our paper. Time for such tasks is hard to find so it is very appreciated.

Referee #1

General comments

R#1: 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.

Response: There are several reasons to why it is desirable to estimate model parameters apriori calibration to streamflow, but perhaps the most important one is that when we use the precipitation-runoff relationship in model calibration, the estimated parameters will be conditioned on both inputs (precipitation and temperature) and the output (streamflow). The calibrated parameters will therefore be sensitive to biases and errors in the inputs. Consequently, the more uncertain and biased the precipitation input, the more uncertain and biased parameter estimates (e.g. Dawdy and Bergman, 1969; Kuczera & Williams, 1992; Andréassian et al., 2001; Engeland et al., 2016). Estimating model parameters uniquely to observed streamflow or its characteristics (as in our case with the distribution of recession characteristics) might therefore reduce both biases and uncertainties in their estimates. Secondly, the problem of equifinality in hydrological model parameters is reduced when a single, unique parameter value is estimated from a data set compared to simultaneously estimating a set of model parameters. Finally, it should become more straightforward to compare the quality of different precipitation products with respect to reliable streamflow simulations. These points will be emphasized in the introduction and we will include the references.

We would like to add that the streamflow characteristics used in this paper have been shown to be transferable to ungauged catchments (Skaugen et al., 2015) with good results.

R#1: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.

Response: It is correct that the hypothesis is not tested directly. On the other hand we have indirectly justified the distribution of S in the vertical, see Figure 6 and the accompanying text. In Figure 6 we have estimated S for recession events (Eq. 13) and plotted its distribution. The similarity in shape to the distribution of Λ is clear.
The basis for the model structure is, apart from the distance distributions, the *observed* distribution of Λ, the recession characteristic, from which we derive the celerities of flow. Λ is typically small for low flows (flat recession) and high for high flows (steep recession), so it is natural (and we do not think very controversial) to relate it to low and high storage, S. We have chosen to discretize the distribution of Λ (and hence the celerities) which gives the levels of the subsurface, S. These levels could be discretized much finer at the expense of computing time and little gain in precision (see Skaugen and Onof, 2014).

Catchments in which the subsurface fluctuations are investigated to such a detail as suggested by R#1 are very hard to find, especially for normally sized catchments (10-1000 km2). There is an example from Norway where a tiny catchment of 0.0075 km^2 was instrumented with over 100 manually read (intermittently) groundwater tubes for a short period of time, 1986-1989. The data (Myrabø, 1997) is, at present, not available but we hope to retrieve it and use it for the purpose suggested by R#1.

Specific comments

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) …”

   Response: A good reformulation. It will be changed.

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”

   Response: Again a good reformulation. It will be changed

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.

   Response: We agree that this comment is perhaps difficult to understand at this point in the paper. We can change it to the more straightforward, ..since overland flow is simulated more frequently.

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.

   Response: Agreed, we will change it to “rate constant”.

5. P11136 L7 “The ratio between consecutive values of runoff, Q(t + 1)/Q(t)” Do the authors mean Q(t + Δt) rather than Q(t + 1)?

   Response: Yes, it will be changed.

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) = ϑ 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.

   Response: Thank you for this. Somehow, a “Δt” was lost during the derivation. The correct relation between κ and ϑ is $\vartheta = \frac{1-\kappa}{\Delta t}$. This has to be corrected in eqs. 5-8, and 13 but is of no consequences for the other equations.

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.

   Response: The celerity is indeed not constant, but varies with storage. The runoff dynamics in DDD is taken care of by basically 4 different celerities, 5 when you include overland flow. Each of these celerites (which are determined from observed recession) are associated with a unit hydrograph. The shape of the unit hydrographs are identical, due to the common distance distribution, but the scale of the UH varies due to the associated celerity. The actual storage determines which UH are active (for example, if it is relatively dry, then no water is routed using the UH for overland flow). The actual runoff is the superposition of active UH’s. The UH’s are “triggered” for each moisture input event, so after the model has been running for a while, quite a few UH’s are at work to convey water to the river network. So, being able to relate analytically the observed distance distribution to UH’s and hence to linear reservoirs is a very important result of this study, and we believe this is a first time such a link is demonstrated. However, as suggested by R#2, this might be moved to an appendix as it is not directly a part of the deriving the new subsurface routine and may improve the readability of the paper.

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?

   Response: Yes, you are correct. The sentence will be changed

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.

   Response: We will remove the words calibration-free. We still think it is a huge difference between calibrating an ensemble of (possibly correlated) parameters to the flow record which will probably give you an equifinality problem, and estimating a single parameter apriori to observed recession characteristics. This is more elaborated upon in our first response to the general comments.

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.

    Response: This comment is similar to the second general comment. You are, of course, correct, and as soon as such data is at hand (see response to the second general comment) we will investigate how the data compares with the concepts in DDD. For normally sized catchments, (10-1000 km2), however, such a sampling approach is extremely challenging and would still involve some non-trivial upscaling (interpolation) of point values.

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?

    Response: This discussion is related to estimating model parameters from catchment characteristics. We will replace the sentence with “The estimation of θM through regression is no longer necessary.”

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?

    Response: A good point, it will be done. We have such plots and they demonstrate this point very well.

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.

    Response: This comment is similar to the first general comment and to comment 12, see response to those. In addition, we will elaborate a bit further on the PUB potential of the new structure of the DDD model. From a study of 84 Norwegian catchments we have found that the parameters of the distribution of Λ were highly correlated (r2- 0.97 and 0.98) to the parameters of λ. This is not a surprise since λ is derived from Λ (see eq. 12). The new model structure of DDD has hence effectively one parameter less (θM) to estimate from catchment characteristics for application to PUB, and Skaugen et al. (2015) show that the parameters of λ and hence Λ can be determined from catchment characteristics.

Andréassian, V., Perrin, C. Michel, C., Usart-Sanchez, I. & Lavabre, J. (2001) Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. J. Hydrol. 250(1-4), 206–223.
D.R. Dawdy, J.M. Bergmann (1969) Effect of rainfall variability on streamflow simulation Water Resour. Res., 5 (5) (1969), pp. 958–966
Kuczera, G. & Williams, B. J. (1992) Effect of rainfall errors on accuracy of design flood estimates. Water Resour. Res. 28(4), 1145–1153.
Kolbjørn Engeland, Ingelin Steinsland, Stian Solvang Johansend, Asgeir Petersen-Øverleir, Sjur Kolberg (2016) Effects of uncertainties in hydrological modelling. A case study of a mountainous catchment in Southern Norway, Journal of Hydrology, 536, Pages 147–160, doi:10.1016/j.jhydrol.2016.02.036
Myrabø, S, 1997. Temporal and spatial scale of response area and groundwater variation in till. Hydrol. Process. 11, 1861-1880.