Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-690-RC1, 2018

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Understanding terrestrial water storage variations in northern latitudes across scales" *by* Tina Trautmann et al.

V. Humphrey (Referee)

vincent.humphrey@env.ethz.ch

Received and published: 22 February 2018

This study aims to evaluate the relative contribution of snow versus liquid water in (total) terrestrial water storage changes in northern latitudes. In order to investigate this question, the authors construct a simple bucket-type hydrological model with 10 free parameters which they calibrate against four different datasets (satellite observations of terrestrial water storage from GRACE satellites, snow water equivalent from the GlobSnow product which combines satellite and ground observations, evapotranspiration from FLUXCOM which is based on an ensemble of machine learning methods calibrated with in-situ observations, and E-RUN estimates of gridded runoff also based on a machine learning model calibrated with in-situ observations). Following a short evaluation of the performance of the presented model (and a comparison with the

Earth2Observe ensemble of hydrological models), the main results are focused on distinguishing the respective contributions of snow and liquid storages to terrestrial water storage. Their analysis contrasts 1) local scale effects versus a spatially integrated average and 2) the mean seasonal cycle versus inter-annual variability. Consistent with previous studies, the authors find that the seasonal cycle is dominated by the snow component. The main finding of the paper is that liquid water storage clearly dominates inter-annual variability both at local scale and when considering a spatially integrated time series. They also find that the relative contribution of liquid water is weaker for the spatial integral compared to the local scale analysis. The authors argue that because snowpack evolution is primarily dependent on temperature (which has high spatial coherence (fig. 10)), this explains why the relative contribution of snowpack to large-scale inter-annual TWS variability is higher. In their conclusions, the authors comment on the usefulness of a simple hydrological modeling approach informed by multiple observational constraints. They suggest that long-term changes in water availability in northern latitudes might be driven by soil moisture rather than by snow dynamics.

This is a really good and well conducted paper. I find the results very interesting and worthy of publication in HESS. One can see that a lot of effort was invested in developing a custom hydrological model and this is reflected by the relatively important share of the methods and model evaluation sections in the paper. However, the authors manage to keep the results and discussion focused around the primary objective of quantifying the relative contribution of snow and non-snow storages to overall water storage variability.

I have four major comments/suggestions which I would like the authors to consider as well as some minor comments that are listed below.

Major comments

In their results, the authors find that the modeled seasonal cycle of TWS has a systematic lag compared to observations (model TWS preceding observed TWS). This lag is also present in other models from the Earth2Observe ensemble. The analysis of the authors convincingly shows that their modeled snow storage seems to have a correct phase and is therefore not responsible for this lag between modeled and observed TWS. They also mention that adding delayed storage responses (as e.g. with a groundwater module) could not correct this effect either. I find this a major finding for the research community (which could be made more prominent in the conclusions) since it is often supposed that such model errors mainly stem from the lack of long-memory water storages and poor representation of snow dynamics. Here the authors conclude that neither of these seem responsible and that the origin of the phase lag in TWS must reside elsewhere, which brings me to my main suggestion below. One important limitation that the authors fail to mention is that there is no consideration of permafrost and liquid/solid phase transitions of soil moisture content. In the proposed model, soil moisture does not have temperature neither does it store energy. In reality, it is well known that freeze/thaw dynamics are also a dominant factor for water and energy fluxes in high latitudes. Freeze/thaw is the on and off switch for evapotranspiration and vegetation growth. However, a phase lag between the availability of energy and the ET response cannot be modeled with the current model setup (the alpha parameter only conditions ET amplitude). Potentially, a lot of ground heat flux might be required before ET can actually take place. In addition, from my understanding of the equations presented in the supplementary material, actual ET is not reduced in the case of snow cover neither is it dependent on vegetation growth. This might introduce a too early response of ET to net radiation compared to reality, leading to a fast rise of soil moisture depletion already in early spring. Later, soil moisture would become limiting already in mid-summer and ET would peak in June and start to reduce already in July (Fig S1). The reference below suggests a peak of vegetation growth in August for a boreal forest (from one FLUXNET site). The authors might consider exploring this direction and maybe check whether there is some evidence that FLUXCOM ET itself (the observational constraint) already contains such a phase lag. As this would require some additional work, it would also be fine if the authors prefer to simply mention this

as a hypothesis to explore.

Brown, S. M., Petrone, R. M., Chasmer, L., Mendoza, C., Lazerjan, M. S., Landhäusser, S. M., ... & Devito, K. J. (2014). Atmospheric and soil moisture controls on evapotranspiration from above and within a Western Boreal Plain aspen forest. Hydrological processes, 28(15), 4449-4462.

This leads me to my second main comment: The separation of TWS into liquid and snow water seems a bit misleading since the liquid phase might implicitly include some frozen water as well (frozen soil moisture). As mentioned by the authors, there is a mismatch between explicitly represented processes and observed processes (TWS includes frozen water) that may be compensated by adjustments in model parameters. The expression “liquid phase” is hence misused in my opinion and might very well lead to confusion. It might be more accurate to refer to snow versus non-snow changes as done for example in page 27 line 30. I think this terminology should be extended to the rest of the manuscript.

Third comment: Figure S7 is quite pre-occupying because it suggests a dependency of your results on the forcing dataset. For instance, the difference might be related to your partitioning between snowfall and rainfall (which was not applied when using WFDEI). One possibility to check if this comes from uncertainty in the precipitation data might be to compare the regional mean time series of the two products and look for large differences in 2005 and 2010. This would also indicate whether GPCP-1DD appears superior to WFDEI. In relation to this -> Line 11-12 page 26: this is a rather unsubstantiated statement. Please give it more weight, for instance by replicating key figures (e.g. Fig 9) in the supplementary material.

Fourth comment: You could make lines 24-29 of your abstract clearer. Upon first reading, I understood that snow dynamics dominate IAV on a large scale, which is not the case. It should be clearly said that “liquid water” dominates IAV at all spatial scales while snow dominates MSC at all spatial scales (Fig. 9). In addition, for IAV, the relative influence of snow increases with spatial aggregation due to the spatial coherence of T, a main driver of snowfall and snow melt. The wording “liquid water storages, comprising mainly of soil moisture” is also a bit misleading. It is not really clear what is implicitly incorporated in the soil moisture reservoir in order to fully reproduce TWS (as mentioned in the third comment). Güntner et al. 2007 provides a similar analysis based on WaterGAP. This would be an interesting point of comparison since they indicate a contribution for IAV of 33% snow, 27% soil and 12% groundwater and 28% surface water (!) for cold climates (their table 5). I think this reference should be discussed and compared with your results.

Güntner, A., Stuck, J., Werth, S., Döll, P., Verzano, K., & Merz, B. (2007). A global analysis of temporal and spatial variations in continental water storage. Water Resources Research, 43(5).

Minor comments

Your work is very new and promising in the sense that multiple remote sensing or machine-learning observation-derived products are used simultaneously for calibrating a hydrological model. This is not easy to do and a research direction worth to explore. The overall modeling framework however still relies on a very standard land surface model structure. One missed opportunity may be to have used these observational datasets not only to calibrate model parameters, but also to identify functional relationships directly from the data (as opposed to fitting the parameters of a pre-defined equation to the data). Such research might be suggested as one possible future direction in the discussion. Finally, the paper does not emphasize on the added value of using remote sensing products to constrain the model (except for a lower RMSE against observations, which is somewhat expected since other models were not calibrated with these observations). Could similar results have been obtained with the Earth2Observe ensemble ? (especially on IAV?) If not, this would better show the merit and relevance of the presented approach.

Liquid water is explicitly modeled as soil moisture + runoff routing but also likely includes river storage, lakes and wetlands implicitly (e.g. large water holding capacity mentioned on page 12, lines 7 and 19). This could also be made a bit clearer already in the model description in order to avoid some confusion later. Using a snow/non-snow terminology would also help resolving this.

It could be made clearer that runoff is currently only generated from infilitration limitation (e.g. no baseflow in Eq. S10). Also mention that this is partially compensated by the recession time scale parameter that delays runoff generated on a specific day. Likely because the model is evaluated at monthly scale, this only has a limited impact on model performance and this model parameter is the least constrained by observations.

Methods: You could make a better distinction between purely observational products, and observation-based upscaled products such as Tramontana et al. or Gudmundson et al. which also rely on the quality of the underlying forcing data.

Line 29-30, page 10: this assumption seems a bit dangerous given figure 6. Could you please document the degree to which this assumption is correct and if this might affect the results qualitatively (possibly in supplementary information)?

** in-text comments **

Line 29, page 2: and in addition there can be no retrieval of SM in snow-covered or highly vegetated regions.

Line 3, page 4: for clarity, maybe you could add an introductory sentence indicating that this whole section is meant to give an overview of the model setup.

Line 10, page 4: E-RUN, based on E-OBS

Line 10-14, page 7: I thought FLUXCOM was based on an ensemble of machine learning algorithms (e.g. not only random forest). Could you also briefly comment on the performance of FLUXCOM in snow regions and high latitudes? Any idea if FLUXCOM is already accounting for sublimation?

Page 9: Is there any reference for this cost function ?

Lines 17-20 page 9, I think this is indeed a very good idea!

Line 22, page 9, can you indicate where these commonly reported values can be found? (It's also fine if you decide to assume 10%)

Line 8, page 10 : typo

Line 22-23 page 11: interestingly however, this also contradicts Behrangi et al. 2017 for mountainous regions..

Behrangi, A., Gardner, A. S., Reager, J. T., & Fisher, J. B. (2017). Using GRACE to constrain precipitation amount over cold mountainous basins. Geophysical Research Letters, 44(1), 219-227.

Line 15 page 12 : maybe this rather small value is in relation with the relatively large soil water holding capacity.

Line 12-13 page 15: For instance, Humphrey et al. 2016 figure 6 shows that the central North America and Eastern Eurasia is rather dominated by IAV (which appears more difficult to model according to your figure 5).

Humphrey, V., Gudmundsson, L., & Seneviratne, S. I. (2016). Assessing global water storage variability from GRACE: Trends, seasonal cycle, subseasonal anomalies and extremes. Surveys in geophysics, 37(2), 357-395.

Line 16 page 12: the sentence is inaccurate : a recession time scale of x days does not mean that only runoff of the preceeding x days contributes to "total runoff" (check Orth et al. 2013).

Line 13, page 13 : maybe not necessary to say that these approaches are not commonly accepted as this might be a subjective statement in my opinion. The arguments you give just before (on overfitting) and the continental-scale focus of your study might be sufficient arguments. Another argument you could mention is that allowing locally

varying parameters would contaminate your conclusions: with locally dependent parameters, the differences in local-scale / large-scale contribution to IAV might due to the spatial dependency of parameters. But with your current setting, they can only be attributed to climate forcing. This is also why it makes a very clean experiment setup. This last point also calls for one caveat in the conclusion : your picture of the partitioning and scale-dependency of liquid versus snow might also change once you introduce spatial variability of the model parameters (e.g. snow melt factor might be very dependent on the vegetation cover, contrasting the responses of tundra versus boreal forests).

Page 14, lines 3-6 : essentially repeats page 13 line 10.

Figure 3. If values were truncated (e.g. Fig3d) this should be indicated in the legend and in labels.

Line 1 page 19 : TWSmod ?

Line 7 page 19 : typo in earth2observe

Line 5 page 10 : replace "grids" with "grid cells" idem on lines 5-6 page 20

Line 5 page 20 : was the use of a subset mentioned also in the methods ?

Line 6 page 21 : coincides

Figure 7: It would be nice to add units to the colorbars (in addition to qualitative labels), same in Figure 8.

Line 18 page 22: is "received" the adequate word?

Line 7 page 23: frozen soil is not modelled

Line 11 page 25: On first read I could not follow since you cannot invoke geographic characteristics when you have spatially constant model parameters. The only source of spatial variability is in the model forcing. This is mentioned but only later on page 26 line 5, maybe you could reformulate this in a way that avoids such a misunderstanding.

Line 7-9 page 26: note that ET is also influenced by Rnet which might also be less spatially coherent.

Line 11-12 page 27: and there can be no SM retrieval in snow-covered regions.

Line 15-18 page 27: solid/liquid phase transitions in soil moisture layers are another type of neglected effect relevant in the study domain as mentioned in the main comment.

Line 10 page 28: the fact that snowpack anomalies are "erased" each summer and partially transferred to soil moisture through snow melt also largely explains this pattern. Hence, soil moisture also by construction allows for a longer memory than snowpack. This could be made more prominent in the discussion as well.

Line 11-12 page 28: I would not qualify this as "diverging" since the sign is still the same (Figure 9). The non-snow storage only becomes less dominant when spatially integrated.

Line 23-25 page 28: adding this to the abstract would really explain better what you mean with the cryptic ending on line 32-34 page 1.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-690, 2017.