Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2015-533-AC1, 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Global root zone storage capacity from satellite-based evaporation" *by* L. Wang-Erlandsson et al.

L. Wang-Erlandsson et al.

lan.wang@su.se

Received and published: 22 February 2016

We are grateful for the constructive comments, and would here like to briefly respond to Referee #1's general comments. The specific comments will be addressed in a later response. Below, referee #1's general comments are in bold and our responses are in upright font. We refer to the manuscript for explanations of variables and abbreviations.

1) MAJOR: The description of the method should be improved. Is the method the same as in previous studies (e.g., Gao et al. 2014 GRL)? If yes, it should be clearly acknowledged. Is it different from the paper (under review, not available to reviewers) by Boer-Euser et al.? It should be clear to the reader if the novelty of the papers is on the method or in the satellite dataset used as input. Please clarify.

We consider this paper to be novel in several ways, not isolated to either method or satellite dataset input. The novelties include:

- showing that a mass balance based method for estimating S_{R} is suitable for global scale application,
- making use of the recent developments in satellite-based evaporation data to estimate S_{R} at the global scale,
- present a S_{R} dataset that can be used directly by the community of global modellers,
- providing significant new insights on the strategy of natural vegetation to deal with droughts and being resistant to future climate fluctuations,
- showing how evaporation simulation changes when conventional look-up table derived root zone storage capacity information is replaced by the presented S_{R} dataset, and showing how this compares to independent evaporation products, and
- investigating the differentiated drought return periods for different land use types.

While we make use of the same mass balance principle as applied by Gao et al., (2014) and de Boer-Euser et al., (2016) (now published), our algorithm is based on indirect measurements of every unique pixel that reduces various assumptions related to atmospheric exchanges between land and atmosphere. The scale of application, and input data are different as well.

Importantly, Gao et al., (2014) did not use actual evaporation data as input, and used NDVI to estimate the slope of transpiration during different seasons. The slope of transpiration was kept constant, and the mass balance was calculated using the Mass

Curve Technique, and not a running "deficit" account as in this paper. In contrast to the method used by de Boer-Euser et al. (2016), we do not a priori assume a Gumbel return period. Moreover, both Gao et al., (2014) and de Boer-Euser et al. (2016) applied the mass balance approach at the catchment scale and estimated interception in order to use transpiration with effective precipitation. In this paper, we show that it is feasible to apply the concept directly at the global scale using independent pixel measurements instead of catchments as units and simplified approximations of total evaporation and total precipitation. An overview of the differences is provided in Table 1.

Although the basics of both methods (Gao et al., 2014; de Boer-Euser et al., 2016) were covered in the Introduction, we agree that the differences could be made clearer. In the revision, we will highlight the novelties of this paper, especially attending to the differences between our method and those applied in Gao et al., (2014) and de Boer-Euser et al., (2016).

2) MODERATE: In the "Methods" section it reads several assumptions: (i) "irrigation is captured in satellite-based evaporation data", I am not sure it is true. At least, not for all satellite-based datasets, please clarify. Moreover, at page 14 it reads that the evaporation originating from irrigation water simulated by LPJmL is considered. Why irrigation is already included in the satellite-based evaporation data? (ii) "the long term average is added in order to compensate for overestimation of evaporation and underestimation of precipitation". Why? (iii) "in order to take into account of surface runoff, D never becomes negative". Again, why? I believe that the authors should better justify the assumption made in their method and why these assumptions are valid (or not). This will allow the reader to understand the strengths and the limitation of the proposed approach.

(i) Satellite-based evaporation datasets determine the latent heat flux from radiation data and thermal infrared data. The origin of evaporation cannot be determined, although it can be verified whether E is exceeding net precipitation P_{net}. Several pa-

pers have demonstrated the capability to distinguish total E into the component that originates from rainfall P and from other sources such as irrigation, inundations and groundwater dependent ecosystems (e.g. Ahmad et al., 2005; Van Eekelen et al., 2014; Bastiaanssen et al., 2014). The latent heat flux simply encompasses all these processes, hence, evaporation from irrigation is implicitly incorporated. Because of this, we must remove it from our S_{R} calculations, and therefore used irrigation evaporation from LPJmL. Other models such as GlobWat (Hoogeveen et al., 2015) and Water Footprint (Chukalla et al., 2015) could have been used alternatively. If evaporation irrigation is not removed, the S_{R} estimates in irrigated regions would be greatly overestimated, as we would have mistaken irrigation water to come from a natural soil moisture store. We wrote at P. 9 L. 16-18: "Without correction, the irrigation evaporation in the satellite evaporation data would erroneously contribute to accumulation of soil moisture deficit in our computations."

(ii) "The long-term average is added…" in certain grid cells where accumulated E-P is positive for several years in a row, in order to prevent S_{R} from growing every year. Continuously increasing E-P may be linked to wetlands, irrigated fields or natural seepage zones, where lateral inflow of water occurs. In our method this would lead to accumulation of moisture, which is not possible over longer periods of time. Some deeply rooting vegetation may accumulate moisture over more than a year to replenish moisture deficits of previous years, but positive E-P values over several years need to be compensated. For clarification, we will revise at P. 9 L. 20-21: "[the long term average]… [is added] to the inflow, in order to compensate for lateral inflow or estimation errors in evaporation or precipitation."

(iii) We assume that any excess precipitation that cannot be contained by the root zone storage reservoir (or D, which is essentially a running estimate of root zone storage reservoir size) is runoff or recharge. Since the reservoir size is calculated through D, D can never be negative. We wrote that this is a way to take into account surface runoff, but it could also be expressed as the inevitable procedure if D is defined as the running estimate of the root zone storage reservoir size.

We will make these points clearer in the revised manuscript.

3) MAJOR: The selection of the input datasets is for me a major issue. Again, it should be clarified why satellite-based data are considered for evaporation and not for precipitation. Why satellite-based datasets for precipitation are not considered (e.g., TMPA, CMORPH, PERSIANN)?

Moreover, why the average of the three evaporation datasets should be "attractive"? Are the results changing by using only one of the datasets? What is the relative impact of the evaporation and precipitation datasets on the final results? It should be clarified, too.

Why ERA-Interim data are used for temporal downscaling? Apart that it is not mentioned how the temporal downscaling is carried out, currently daily evaporation and precipitation datasets are (freely) available (actually, several datasets). Why the authors do selected monthly datasets and then performed downscaling with ERA-Interim? Why not using directly ERA-Interim data? Or other daily products (e.g., GLEAM for evaporation and TMPA for precipitation)? All these points should be clarified.

To estimate S_{R}, we need global coverage at a grid cell resolution for both evaporation and precipitation. Importantly, these products that must not be produced using assumptions on root zone storage capacity, to prevent circularity (since we are estimating root zone storage capacity). In other words, there should be no water balance type of computation process involved in the determination of S_{R}. We used satellite-based evaporation products because they are the only options available that fulfil these criteria, i.e., reanalyses and land surface model evaporation contain soil depth information. Flux net data are too sparse for acquiring consistently good quality global coverage). Conversely, precipitation data do not need to be satellite-based, but can also be ground-based. In this manuscript, we used CRU (ground-based) and CHIRPS

(satellite-based).

Inter-comparison of precipitation products show that both CRU and CHIRPS are good quality precipitation products. In particular, CHIRPS performance stands out in a comprehensive inter-comparison of 13 difference precipitation products in the Nile basin (Hessels, 2015). PERSIANN was the worst performer in this inter-comparison, whereas CMORPH performed the worst in several bias analyses (Hessels, 2015). Ongoing analyses (not yet published) in the Mekong basin, Vietnam and Colombia also show that CHIRPS performs excellently, so the CHIRPS performance in three continents are outweighing the older technologies used for PERSIANN, CMORPH and APRODITE. One paper on Vietnam is currently under review elsewhere (Simons et al., in review).

The average of the three evaporation datasets is the simplest and most transparent way to create an ensemble product. (Hofste, 2014) also showed that three different approaches towards an ensemble evaporation product (i.e., simple averaging, expert judgement, and outlier removal based on MODIS16NBI, SSEBop, CMRSET, and ALEXI7) all performed similarly better than the individual constituent evaporation datasets in the Nile basin. In their comparison analyses, the mean product outperformed the one based on expert judgement and exhibited less "worst sub-basin performances" than their outlier analysis product, resulting them to conclude that the mean ensemble without any imposed restrictions is "relatively safe" to use.

We have performed analyses of the relative impact of evaporation and precipitation, which we did not include in the paper for conciseness. However, the referee reminds us that this could be of interest for the reader, and we will therefore include these results in the Supplements in the revised version.

Daily remotely-sensed ALEXI-based E data is under production by NOAA and USDA. Daily data based on VIIRS will be made available. The new ETMonitor remote sensing algorithm from the Chinese Academy of Sciences also envisages daily fluxes at the global scale, but at this moment these products are not available yet. There are currently (February 2016) seven global remote sensing products of E available. The monthly evaporation data used in the manuscript were those available at the time of this research. We did not consider using GLEAM, since the evaporation data is modelled with a-priori assumptions on soil layer depth by land use type (Miralles et al., 2011). For the same reasons, we did not employ the ERA-Interim evaporation, since it is based on the land surface model TESSEL (van den Hurk et al., 2000).

In the temporal downscaling, we first established the ratios between daily values to the mean monthly ERA-Interim, and second, used the relationship to estimate daily values from monthly E_{SM} or E_{CSM} values. We will describe this procedure in the revised manuscript.

4) MODERATE: In most of the paper, only the $S_{\mathbf{R,CRU\text{-}SM}}$ dataset is analysed. Why two datasets are considered (CHIRPS and CRU)? The real value of considering also the CHIRPS dataset is not clear to me. Please clarify.

We present CHIRPS combined with the mean of SSEBop, CMRSET and MODIS, because CHIRPS is the lead precipitation product and has a fine resolution of 5 km pixels. The use of three evaporation datasets decrease uncertainties related to individual evaporation products because there is simply not one single preferred model. Research executed by Hofste (2014) for the Nile basin demonstrated that the performance of an ensemble E product is significantly better than using individual E products, something that was confirmed by Simons et al. (in review) in Vietnam. However, CHIRPS is unfortunately not available at the global scale, and CMRSET is not reliable in high latitudes, and the modelling community likely needs a global S_{R} product. Thus, we added $S_{\mathrm{R,CRU\text{-}SM}}$. This way, we have a global S_{R} map that can be compared to the $S_{\mathrm{R,CHIRPS\text{-}CSM}}$ for reference. We will clarify this further in the revised manuscript.

Note also that we show results of the CHIRPS analyses in the Supplementary Information. We explained (P. 19 L. 17-18) that we did not present E simulation results

using $S_{\mathrm{R,CHIRPS\text{-}CSM}}$ in the main manuscript, since it does not have global coverage, but referred to the Supplementary Information for a comparison between using $S_{\mathrm{R,CRU\text{-}SM}}$ and $S_{\mathrm{R,CHIRPS\text{-}CSM}}$.

5) MAJOR: I found the selection of the application for validating the obtained $S_{\mathbf{R}}$ maps not correct. In the paper, it is assessed the improvement in estimating evaporation with the new $S_{\mathbf{R}}$ parameterization in STEAM. It is fine for me. The problem is that the same evaporation dataset ($E_{\mathbf{SM}}$) used for computing $S_{\mathbf{R}}$ is also used for assessing the improvements due to the new $S_{\mathbf{R}}$ parameterization. It is a circular argument that is not good. I suggest performing a different validation test. Why not considering the differences in the runoff prediction with the old and new $S_{\mathbf{R}}$ parameterization? It looks to me much more relevant, and a good independent evaluation.

We consider validation using E_{SM} to be appropriate, since the algorithms for estimating S_{R}, and for estimating E in STEAM are very different. First, S_{R} is derived based on the E overshoot over P, whereas STEAM is a process-based model where evaporation originates from five different compartments, each constrained by potential evaporation and related stress functions. This means that it is impossible to reproduce E_{SM} simply by inserting S_{R} to STEAM. If S_{R} is zero because E_{SM} never overshoots precipitation, STEAM soil evaporation and transpiration would become zero. If extreme S_{R} are produced because E_{SM} is unrealistically large, STEAM evaporation will not approach E_{SM}, since it will be capped by potential evaporation. Second, consider also that the precipitation products (CRU and CHIRPS respectively) used for deriving S_{R} differ from the precipitation forcing (ERA-Interim) used in STEAM. Third, E_{SM} and STEAM are truly independent to each other as well. Whereas STEAM is process and water balance based, the ensemble E product is based on a combination of two(E_{SM})/three(E_{CSM}) well established energy balance methods. The only difference of the new STEAM simulations is the inclusion of updated information on root zone storage so that during longer periods of drought, more realistic estimations of continued evaporation processes can be expected. Last, $S_{\mathrm{R,CRU\text{-}SM}}$ is based on a single year value of E_{SM} (i.e., the year of maximum storage deficit), whereas the analyses of improvements were based on the entire available time series of 10-11 years. Thus, the fact that $S_{\mathrm{R,CRU\text{-}SM}}$ dimensioned on one year of E_{SM} nevertheless improves E simulation in STEAM with regard to 10-11 years of E_{SM} (i.e., the overall ϵ_{RMS} decreases when $S_{\mathrm{R,CRU\text{-}SM}}$ is used in STEAM) is a strong indication that the storage capacity correction was implemented for the right reason. We maintain that the comparison with E_{SM} is useful and will clarify our arguments in the revised manuscript. Note also that STEAM is not calibrated by any means.

To address the referee's concern of interdependency, we cross-check the mean monthly STEAM evaporation based on $S_{\mathrm{R,CRU\text{-}SM}}$ (2003-2013) with the mean monthly LandFlux-EVAL diagnostic ensemble evaporation (1989-2005) (Mueller et al., 2013), see comparison in Table 2. It appears that the ϵ_{RMS} improvements are even greater (mean improvement 10 mm/year instead of 4 mm/year), but with the greatest improvements in maximum monthly evaporation instead of minimum monthly evaporation. The LandFlux-EVAL diagnostic product include the evaporation products: PRUNI, MPIBGC, CSIRO, GLEAM, and AWB. Since this product includes GLEAM, which relies on water balance calculations and soil layer depth assumptions, we consider the use of this product inappropriate for our purposes and would refrain from including this comparison in the manuscript.

Runoff data represent a catchment or basin average value that is not sufficient for validating a spatial map on S_{R}. Using runoff for validation, we would for example not be able to analyse the E simulation performance with regard to climate indicators or land use types. Runoff results are also highly sensitive to the precipitation data used, especially in wet regions (Fekete et al., 2004; Materia et al., 2010). River flow data can be considerably unreliable in several large river basins such as the Congo (Tshimanga and Hughes, 2014). Nevertheless, we have compared simulation results (P-E) to annual mean GRDC runoff data, and will add it to the Supplementary Information for

readers' reference.

6) MINOR: A very recent paper with the same topic has been published in Journal of Hydrology by Campos et al. (2016, http://dx.doi.org/10.1016/j.jhydrol.2016.01.023). I suggest mentioning and analysing this study.

We thank the referee for this suggestion. We will add this to the revised manuscript.

References

de Boer-Euser, T., McMillan, H. K., Hrachowitz, M., Winsemius, H. C. and Savenije, H. H. G.: Influence of soil and climate on root zone storage capacity, Water Resour. Res., n/a–n/a, doi:10.1002/2015WR018115, 2016.

Fekete, B., Vörösmarty, C., Roads, J. and Willmott, C.: Uncertainties in precipitation and their impacts on runoff estimates, J. Clim., 17, 294–304, doi:10.1175/1520-0442(2004)017<0294:UIPATI>2.0.CO;2, 2004.

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N. and Savenije, H. H. G.: Climate controls how ecosystems size the root zone storage capacity at catchment scale, Geophys. Res. Lett., 41(22), 7916–7923, doi:10.1002/2014GL061668, 2014.

Hessels, T. M.: Comparison and Validation of Several Open Access Remotely Sensed Rainfall Products for the Nile Basin, Delft University of Technology., 2015.

Hofste, R. W.: Comparative analysis of near-operational evapotranspiration products for the Nile basin based on Earth Observations; First steps towards an ensemble ET product, Delft University of Technology. [online] Available from: http://repository.tudelft.nl/assets/uuid:16659a39-3256-4ff9-9930-

81ac4dfb4018/maindoc.pdf, 2014.

van den Hurk, B. J. J. M., Viterbo, P., Beljaars, A. C. M. and Betts, A. K.: Offline validation of the ERA40 surface scheme. [online] Available from: http://www.knmi.nl/publications/fulltexts/tm295.pdf (Accessed 13 July 2012), 2000.

Materia, S., Dirmeyer, P. a., Guo, Z., Alessandri, A. and Navarra, A.: The Sensitivity of Simulated River Discharge to Land Surface Representation and Meteorological Forcings, J. Hydrometeorol., 11(2), 334–351, doi:10.1175/2009JHM1162.1, 2010.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H. C., Meesters, A. G. C. A. and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sci., 15(2), 453–469, doi:10.5194/hess-15-453-2011, 2011.

Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., Fisher, J. B., Jung, M., Ludwig, F., Maignan, F., Miralles, D. G., McCabe, M. F., Reichstein, M., Sheffield, J., Wang, K., Wood, E. F., Zhang, Y. and Seneviratne, S. I.: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis, Hydrol. Earth Syst. Sci., 17(10), 3707–3720, doi:10.5194/hess-17-3707-2013, 2013.

Tshimanga, R. M. and Hughes, D. A.: Basin-scale performance of a semidistributed rainfall-runoff model for hydrological predictions and water resources assessment of large rivers: The Congo River, Water Resour. Res., 50(2), 1174–1188, doi:10.1002/2013WR014310, 2014.

Simons, G.W.H., W.G.M. Bastiaanssen, L.A. Ngô, C.R. Hain, M.C. Anderson and G. Senay, 2016. Integrating open-access satellite-derived global data products as a pre-analysis for hydrological modelling studies: a case study for the Red River Basin, MDPI Remote Sensing (in review)

Hofste, R.W., 2014. Comparative Analysis Among Near-Operational Evapotranspiration Products for the Nile Basin Based on Earth Observations, M.Sc. thesis, Faculty of

Civil Engineering and Geosciences, Delft University of Technology: 24 pp

Eekelen, van, M., W.G.M. Bastiaanssen, C. Jarmain, B. Jackson, F. Fereira, J. Bosch, P. Dye, P. Van der Zaag, A. Saraiva, E. Bastidas-Obando, R. Dost and W. Luxemburg, 2015. A novel approach for estimating direct and indirect water withdrawals using satellite measurements: a case study from the Incomati Basin, Agriculture, Ecosystems and Environment, 200: 126-142

Ahmad, M.D., W.G.M. Bastiaanssen and R.A. Feddes 2005. A new technique to estimate net groundwater use across large irrigated areas by combing remote sensing and water balance approaches. Hydrogeology Journal Volume 13. Number 5-6: 653-664.

Chukalla, A.D., M.S. krol and A.Y. Hoekstra, 2015. Green and blue water footprint reduction in irrigated agriculture: effect of irrigation techniques, irrigation strategies and mulching, Hhydr. Earth Syst. Sci. 19: 4877-4891

Hoogeveen, J., J.-M. Faurès, L. Peiser, J. Burke, and N. van de Giesen, 2015. GlobWat – a global water balance model to assess water use in irrigated agriculture, Hydrol. Earth Syst. Sci., 19, 3829-3844

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2015-533, 2016.

Table 1. Overview of the methods used in Gao et al., (2014), de Boer-Euser et al. (2016), and this study.

	Gao et al., (2014)	**de Boer-Euser et al., (2016)**	**This paper**
Scale/Coverage	Catchment	Catchment	Global
Unit	Catchment	Catchment	Grid cell (0.5 degree)
Water demand input in the mass balance calculation	The slope of cumulative transpiration was used as consumptive use. The slope of cumulative transpiration was derived from Normalized Difference Vegetation Index (NDVI). Interception threshold to estimate effective precipitation was assumed to be 2 mm/day.	Daily transpiration was used as consumptive use. Total evaporation estimate came from the annual water balance (i.e. P and Q data), interception was estimated through model simulation, and daily transpiration was downscaled from long term average transpiration using estimates of daily potential transpiration.	Total time-variable actual evaporation, which includes all evaporation components (e.g. transpiration, interception), was determined from the surface energy balance (not from a water balance). Interception estimation is no longer needed, as it is implicit in both P and E data, and therefore cancels out. Further, we included the effect of evaporation from irrigation.
Mass balance algorithm execution	Mass Curve Technique	Daily water balance model with interception and root zone storage reservoir. Deficit increases when transpiration exceeds effective precipitation. Any excess precipitation is assumed to runoff directly.	Daily water balance with root zone storage capacity reservoir. Deficit D increases when total evaporation exceeds total precipitation, and decreases when P>E and D>0. Excess precipitation is assumed to be runoff or recharge.
Identification of the most suitable drought return period	Identified through the best runoff simulation performance across the catchment.	Assumed 10 years across the catchment based on Gao et al. (2014).	Differentiated drought return periods are identified for different land use types using evaporation simulation performance.

Table 2. ϵ_{RMS} and ϵ_{RMS} improvements in evaporation simulation with E_{SM} and LandFlux-EVAL as benchmark respectively.

Monthly E compared	**ϵ_{RMS} with E_{SM} as benchmark (mm/year)**			**ϵ_{RMS} with LandFlux-EVAL diagnostic as benchmark (mm/year)**		
	Look-up	$S_{\text{R,CRU-SM}}$	ϵ_{RMS} improvement	Look-up	$S_{\text{R,CRU-SM}}$	ϵ_{RMS} improvement
Mean	234	230	4	136	126	10
Max	323	320	3	244	222	22
Min	189	181	8	143	129	14