Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/hess-2021-126-RC1, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on hess-2021-126

Anonymous Referee #1

Referee comment on "A comparative evaluation of the calibration-free complementary relationship with physical, machine-learning, and land-surface models" by Daeha Kim et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-126-RC1, 2021

MODEL FORCING:

The most questionable issue in this MS is that the authors still use ERA-Interim forcing to drive the CR in the MS submitted in 2021 rather than 2011, the latter is the year when the widely-known ERA-Interim paper was published in Quarterly Journal of the Royal Meteorological Society. In 2021, this is an obviously out-of-date meteorological forcing which should not be used (and also ERA-Interim's ET output). Its successor, ERA5, with a spatial resolution of 0.25 degree, should be considered as a significant improvement, evidenced by a few recent papers when both ERA-Interim and ERA5 were used to drive the ET model (Martens et al., 2020) and hydrological model (Tarek et al., 2020). In general, ERA-Interim has larger errors than ERA5, which would propagate into the CR-model simulated ET values.

It is very strange that the authors did not use locally-preferable meteorological forcing when they focus only on Australia. In Australia, the SILO forcing (https://www.longpaddock.qld.gov.au/silo/), produced by interpolation of ground observations from the Australian Bureau of Meteorology and also other sources, should be much better than those forcing developed for a global coverage. While SILO has no net radiation nor wind speed, its air temperature, vapor pressure, air pressure, and solar radiation should be much more reliable than those from ERA-Interim/ERA5 for Australia. For this reason, the authors should try to drive the models using these.

Net radiation is often regarded as the most sensitive input for most ET models (see e.g., Figure 3 in Fisher et al. 2017). For this reason, I have to remind that net radiation from any atmospheric reanalysis dataset is essentially from the model simulations (for upward short- and long-wave radiation), which may have greater uncertainties than satellite observations. While satellite-based net radiation often has relatively coarse spatial resolution (one degree), the authors used 0.5 degree for their simulations, which is not far from it. Therefore, I wonder if the CR model could be improved with satellite-based net radiation data driving it. I suggest trying to use net radiation from CERES

(https://ceres.larc.nasa.gov/) and/or GEWEX SRB (https://www.gewex.org/data-sets-surface-radiation-budget-srb/).

Refs:

Martens, B., et al., 2020. Evaluating the land-surface energy partitioning in ERA5. Geosci. Model Dev., 13, 4159–4181

Tarek, M., et al. 2020. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. Hydrol. Earth Syst. Sci., 24, 2527–2544.

Fisher, J., et al., 2017. The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. Water Resources Research, 53, 2618-2626

MODEL VALIDATION:

As the water-balance-based ET is key for assessing the ET models, the weakest point in ET_{wb} of this MS is the grid-based runoff data (GRUN and LORA) which has much larger uncertainties than the station-measured ones at the outlet of a basin. In Australia, the most popular runoff data is "Zhang et al., 2013. Collation of Australian modeller's streamflow dataset for 780 unregulated Australian catchments. CSIRO Land and Water, 115 pp." (available at: https://publications.csiro.au/rpr/pub?pid=csiro:EP113194). Note also that it is more appropriate to involve only the unregulated basins with minimum human activities for validation purpose. While the authors did compare their continent-averaged ET_{wb} with previous similar studies in Line 230-239, this does not necessarily mean that ET_{wb} is accurate at the grid or basin scales.

Another key deficiency is the precipitation data for calculating ET_{wb}. For the same reason explained above, the authors should use precipitation data with a regional/continental focus rather than the one developed for a global coverage. The BILO precipitation data is often regarded as the most reliable one for Australia (https://data.gov.au/data/dataset/67749ef0-7223-437a-851a-573edde09567), which should be used to replace GPCC for a more accurate ET_{wb}.

I do understand that the authors want to use grid-based ET_{wb} data for evaluations, but the authors should also test its reliability with basin-scale ET_{wb}. The latter could be derived using measured runoff data from the above-mentioned 780 basins.

I would further argue that the authors' ET_{wb} data are at least 10% smaller than the real values most likely are due to the suboptimal choice of the precipitation as well as the gridded runoff data. The reason I am saying this is that FLUXCOM and CR ET yields pretty similar values seen in Fig. 4. It should be noted that FLUXCOM must provide (and it does) one of the most accurate ET data available today since it is based on actual ET measurements by eddy-covariance, even though its inter-annual variance is somewhat subdued as seen in Fig. 3 in comparison with the CR ET values, but this temporal smoothing feature of FLUXCOM is well known from previous studies.

So all in all, it is argued here that most of the difference between CR ET and ET_{wb} is most likely due to unsatisfactory choices in model and water-balance forcing rather than to the need of spatially changing the alpha value of the CR model. (This is not to say that the CR would not overestimate ET rates near the sea since there the air moisture is significantly decoupled form the underlying land surface). As another choice, the authors could also apply several different sources for the forcing in the CR model as well as in the water-balance and see how they affect the outcomes (this would also serve as a sensitivity analysis). Chances are they do in a significant way.

So before one just replaces a unique calibration-free model with one needing calibration, one must make sure that the original model was evaluated correctly and exhaustively. I do not feel at all this is the case in this study.

Note also that there is a significant difference between Brutsaert's alpha and the alpha value of the CR employed in this study. The Priestley-Taylor (PT) equation is evaluated at the measured air temperature in the former case, while in the latter at the required (but mostly unknown) wet-environment air temperature (estimated via T_{ws}). Without the latter, the PT equation naturally overestimates the wet-environment ET rates (and thus the actual ET rates as well) the more significantly, the drier and hotter the environment has become, therefore a correction (typically based on some measure of aridity) in the alpha value is necessary in the Brutsaert model, but not in the CR model employed in this study. So in this study the alpha value is meant to be the best available estimate of the real PT alpha value and not some weak analog of it, as in the Brutsaert model (i.e., Brutsaert et al., 2020), the latter taking up values much below the physically still interpretable value of one.

Another model application issue is that in the CR model of this study T_{ws} is estimated only one way, while the original authors of this CR model also described another method for the T_{ws} estimation, yielding somewhat smaller T_{ws} values (as mentioned in this MS, and therefore potentially resulting in a higher alpha-value estimate, most probably bringing the derived alpha value into the often quoted, typically observed 1.1 – 1.32 interval). In fact, most of this CR model's applications use the latter, so it would be worth to check how it affects model outcomes and the constant alpha value estimation.

Line 238: I do not think GPCC is a reanalysis precipitation data.

Please check the text for the numerous typographical errors. For example: correctly ‘Priestley’.