

Hydrol. Earth Syst. Sci. Discuss., author comment AC2  
<https://doi.org/10.5194/hess-2021-126-AC2>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC2

Daeha Kim et al.

---

Author 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-AC2>, 2021

---

We are thankful for Dr. Fisher's positive and constructive comments. To improve the manuscript, we will conduct (1) an additional simple statistical analysis that could isolate effects of individual controls on variability of ET for each model. This would make the manuscript more informative. In addition, (2) we will show the reliability of ET<sub>wb</sub> by replacing the global precipitation product with a locally preferable one, even though the two are not very different. And, we will show predictive performance of the modeled runoff data by comparing them against available catchment runoff observations. Our responses to specific comments are following.

Comment 1: This is a good paper that discusses in depth the complementary relationship (CR) of evapotranspiration (ET), and conducts an extensive evaluation of the CR over Australia. The paper is well-motivated—the vapor pressure deficit (VPD)-soil moisture (SM) CR is widely used in ET estimation, primarily because high quality and high spatial resolution SM is not always readily available; whereas, VPD may be more readily available. Semi-arid places like in Australia are where the CR may be most important, and potentially where it may be the most uncertain.

The Introduction and Discussion are written very intelligently, taking a deep dive into the theory and formulations of the CR and ET estimation. The authors do a good job of describing each of the respective ET models and datasets. Relative to the strength of the overall writing, especially from these sections, the analysis and results were somewhat lacking in depth, however. Ultimately, the results were just a handful of maps and time series of the different products, with no real "truth" to benchmark against. Given how intelligent the authors were with their communication and writing of the theory, I was surprised to see the analysis so shallow. I would have liked to have seen that same intelligence from the writing applied to the analysis. The authors could have gone into much more analytical depth on spatial patterns, sensitivities, etc.

Response: We appreciate the positive comments. To refine our statistical analysis, we will add some analysis that could isolate effects of climatic and other controls on the interannual variability of each model, e.g., via partial correlation analyses. Such information could be beneficial for selection of grid ET products in Australia.

We want to remind that the objective of this work is a comparative evaluation of the CR method, not an absolute one. Australia include large arid areas that are almost inhabitable

for humans. Awaiting ground observations in such locations for evaluating hydrologic models is unrealistic. In this case, reliable ET estimates could become an alternative evaluation reference. Pan et al. (2020), for example, compared modeled ET products even with the ensemble of modeled ETs. We believe that such an evaluation could become informative too. We do not argue that ETwb estimates are true values, but could become an acceptable reference for evaluation.

To improve reliability of ETwb in revision, we will replace the GPCP precipitation (P) with the locally preferable SILO data as recommended by the reviewer 1. Even though the GPCP data are not considerably deviated from the SILO data (see Figure S2), the replacement would prevent prejudices in accuracy of ETwb.

In addition, we will add evaluation of the ensemble of the modeled runoffs against the available runoff observations. Since the LORA and the GRUN are validated datasets by global runoff observations, the modeled runoff would show acceptable agreement. And, we want to highlight that ETwb is mostly determined by precipitation rather than by runoff in the arid continent. Approximately, 90% of precipitation evaporates in Australia (Glenn et al., 2011), and thus quality of the modeled runoff would exert minor influences on ETwb. Please see our responses to the Referee 1's comments

Comment 2: Related to the tenuous/lack of benchmarking, I suggest editing the language for use of words like bias and under/over-estimation e.g. in the Results. These terms generally refer to a metric of truth, of which none is given here (I don't consider the water balance the "truth" given that it is also a model of models; see also comments from Reviewer 1). Better, to stick with language such as larger/smaller/etc. as the comparisons are just relative to one another.

Response: Thank you very much for these more suitable terms. We will use appropriate terminology in revision when discussing the comparisons between the models, and will highlight that ETwb is just an acceptable evaluation reference.

Comment 3: Moreover, be cognizant in attributing pattern to process relative to model run conditions, especially when it comes to relative magnitudes. Any one model can be high or low depending on the forcing dataset it used (see e.g. comments from Reviewer 1), which is not necessarily indicative of the model (or, importantly for this paper, the inferred processes therein). The closest approximation to ascertaining process from pattern would be to identify spatial and temporal patterns regardless of magnitude. For example, the patterns mentioned for AWRA-L in L251 are interesting and likely indicative of process (though they could have easily just been attributable to something unusual in the forcing used for that model).

Response: We agree. In revision, we will look into the patterns of relative magnitudes as well as those of absolute magnitudes. And, we will tabulate forcing inputs of each model so that readers could realize differences in forcing inputs of the models at a glance.

Comment 4: Abstract is written a bit, well, abstractly. It could use more take-home information/detail like what exactly where the models and what exactly was their performances.

Response: After revision, the abstract will be rewritten accordingly. We will include several

take-home lessons in the new abstract.

Comment 5: L37. See [Fisher et al., 2017]., L39. See [Polhamus et al., 2013], L47. See [Fisher et al., 2011], L54. See, for reference, [Purdy et al., 2018].

Response: Thanks for the references. We will cite them when necessary. They could improve the introduction, .

Comment 6: PT-JPL [Fisher et al., 2008] also incorporates the complementary relationship, citing Bouchet, in the soil evaporation component—e.g.,  $RH^{VPD}$ . This simple formulation tracks relative surface wetness well [Fisher et al., 2008], and has since been used in other major models of ET, e.g., PM-MOD16 [Mu et al., 2011]. Still, advection will contaminate the relationship, and replacement with direct soil moisture e.g. [Purdy et al., 2018], can eliminate that contamination. The new ECOSTRESS mission [Fisher et al., 2020] uses PT-JPL for the global ET product, but is currently being updated to incorporate the [Purdy et al., 2018] soil moisture formulation and inclusion, downscaled using the measured LST and NDVI following [Colliander et al., 2017].

Response: We will add the given attributse of the PT-JPL in the description.

Comment 7: Figure 3. I suggest making the symbols in the Taylor diagram more distinguishable.

Response: We will revise as recommended.

Comment 8: Figure 4. PT-JPL data are available from 1984 from the same link where you got the current data.

Response: We will update it as recommended in revision, and discussion will be revised accordingly

Comment 9: See [Purdy et al., 2018] for soil moisture incorporation into PT-JPL.

Response: We confirmed it. The discussion will be revised accordingly.

Comment 10: Figure 8. This seems to be redundant with Figure 4.

Response: Figure 8 updates Figure 4 with performance of the calibrated CR method. Hence, it is not redundant, but indicating differences from Figure 4. However, to make the manuscript more concise, we could consider simple explanation on performance of the CR method after calibration with ETwb.

## References

Pan, S., Pan, N., Tian, H., Friedlingstein, P., Sitch, S., Shi, H., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Lienert, S., Lombardozi, D., Nabel, J. E. M. S., Ottlé, C., Poulter, B., Zaehle, S., and Running, S. W.: Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning and land surface modeling, *Hydrol. Earth Syst. Sci.*, 24, 1485–1509, <https://doi.org/10.5194/hess-24-1485-2020>, 2020.

Please also note the supplement to this comment:

<https://hess.copernicus.org/preprints/hess-2021-126/hess-2021-126-AC2-supplement.zip>