Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-595-AC2, 2021

Hydrology and Earth System Sciences Discussions

HESSD

Interactive comment

Interactive comment on "Robust historical evapotranspiration trends across climate regimes" *by* Sanaa Hobeichi et al.

Sanaa Hobeichi et al.

s.hobeichi@unsw.edu.au

Received and published: 28 April 2021

We would like to thank the referees for their constructive comments on our manuscript. This document outlines our responses to their comments. We provide a track changed version of the manuscript to highlight the changes made to the manuscript and the supplementary material. In addition to the suggested changes by the two referees, we have further improved the analysis by introducing a parallel, complementary dataset version to DOLCE V2.1, DOLCE V3, that has fewer parent datasets than V2.1, reducing the number of temporal tiers and temporal discontinuities found in DOLCE V2.1, mostly over the tropics. DOLCE V2.1 remains a more optimal dataset in many senses as it minimises bias and maximises correlation with in-situ observation, whereas V3 prioritises temporal continuity. Similar to DOLCE V2.1, the superiority of DOLCE V3 over its parents is demonstrated using an out-of-sample testing approach. DOLCE V3 is presented alongside DOLCE V2.1 throughout the manuscript and has not resulted in any new sections or qualitative change to the manuscript. The main change is in section '3.5 Changes in ET since 1980', in which DOLCE V3 was used instead of DOLCE V2.1 to carry out the analysis of trends. The new results show that trends in DOLCE V3 ET are mostly within the range of trends in available ET datasets, unlike DOLCE V2.1 whose temporal inconsistencies resulted in higher trends than the available datasets mostly over the wet ET regimes. We have amended related text, figures and tables accordingly. These updated results also help to address the concerns of the referees, as outlined below.

General Comments: Overall this work seems like a useful addition to the literature. I have some detailed comments for clarification. The results and discussion section of the paper often for large parts mostly just list what is shown in the figures, but it would make it a lot more interesting to more read about what the figures teach us. In addition, please check if small things table contents, figure axes, etc. are introduced. Often this seems somewhat lacking. We thank the referee for his comment. We have now explained the figures and the tables further throughout the text and in the captions. Detailed comments 1. L13: why "gridded"? The scale of ET observations from Eddy Covariance towers is typically less than 1000 m which does not allow to study ET at the regional scale. Therefore, gridded ET datasets are needed to understand ET at the regional scale. 2. L19: "After successful evaluation of the efficacy": a "successful evaluation" does not say anything about the efficacy, so please rephrase. We thank the referee for their suggestion. We have now changed "after successful evaluation of the efficacy of these uncertainty estimates out-of-sample" to: after demonstrating the efficacy of these uncertainty estimates out-of-sample

3. L19: coverage, rather than reach? We thank the referee for their suggestion. We have now changed 'reach' to 'coverage' 4. L33: "with different scopes" is unclear in its

meaning to me. We thank the referee for their comment. We have now clarified 'with different scopes' in the text. . . .. with different scopes (e.g. addressing key questions in ecology, hydrology, or other disciplines), ... 5. L35: "typically incorporating a range of remote sensing inputs" would benefit from some citations. We thank the referee for their comment. Citation was already included before "typically incorporating a range of remote sensing inputs", we have now moved it to the end of the sentence. 6. L36: "have been recognised for their potential to outperform single source datasets" can the strength of these methods be made in a more explicit statement that is more speficic? We thank the referee for their suggestion. We have now specified the strength of merging methods. . . . have been recognised for their potential to outperform single source datasets in reducing bias against in reducing bias against tower-based eddy-covariance ET measurements

7. L40: time resolution (rather than step)? We thank the referee for their suggestion. We believe that both 'time resolution' and 'time step' can be used here interchangeably. 8. L43: chemical seems redundant? We thank the referee for their comment. However we couldn't find any redundancy. 9. L70: physically-based We thank the referee for spotting this, we have now made the correction in the text 10. L70: which ET trends did Pan look at? Pan looked at ET trends during 1982-2011. We have now specified this in the text. 11. L142-147: it seems some references could be added here? References of these products were given in the describing paragraphs that follow. We have now added those references in L142-147. 12. Section 2.2.4. I do not suggest to redo the analysis, but why aren't weighing groups considered based on their physical similarity linked to ET (e.g. landcover) rather than these currently somewhat oddly chosen groups? Good point. We agree with the referee in that the most obvious weighting groups are land cover and climate zones. We have tried both grouping methods in a paper describing the first version of DOLCE, but this did not improve the final hybrid product. This was explained in L336 – L337 Hobeichi et al. (2018) tried to group flux tower sites based on their land cover type and computed weights for each land cover type. However, this approach did not improve the results, whether grouping by climate

zone or aridity index, with the main reason being attributed to the small number of sites in many groups. We disagree with the referee that the grouping approaches are ‘somewhat oddly chosen’. We have clarified in ‘Section 2.2.4 Weighting groups’ the motivation behind each grouping method. We have now added a new weighting group that considers both physical similarities linked to ET, and seasons. We have applied this grouping to derive the new version 3 of DOLCE which we now use to examine ET trends. We have now explained the new grouping method in Section 2.2.4. âĂć Grouping by ET regime and months: Land was classified into three distinct broad ET regimes (Fig. S4) according to two aspects of ET, mean annual total ET and within-year relative variability throughout 1980 – 2018, derived from GLEAM V3.5a, and using K-means unsupervised classification (MacQueen, 1967). We explain the classification method further in section 3.5.2. Different sets of weights were computed at each ET regime during June–November and December–May. Implementing weighting this way ensured that we account for performance differences across different physical aspects of the land and seasons. Despite that observational data was divided into six distinct groups, the observational data available in each group was still appropriate to merge the four parent datasets of DOLCE V3. However, we found this grouped weighting strategy not appropriate for merging 11 parent datasets of DOLCE V2. 13. Table 1: indicate what a (lack of) marker means. It’s somewhat obvious but it’s still good to specify. . . We are not sure what the referee is asking us to specify, but it seems from the referee’s comment, the suggested change is not that important. . ..? 14. Table 3: why are uncertainties this large for DOLCE V2? DOLCE’s uncertainty gives an accurate upper bound estimate of the likely discrepancy between the product and unseen ET measurements. Uncertainties in DOLCE V2 are large compared to uncertainties of hybrid estimates derived by different merging approaches which typically consider the spread the parent datasets. This has been clarified in Section 2.2.2

This process ensures that the computed uncertainty provides a better uncertainty estimate of the hybrid ET than simply using the spread of the parent datasets. One additional advantage of defining uncertainty in this way is that it should give an accurate upper bound estimate of the likely discrepancy between the product and unseen ET measurements at a range of spatial scales. That is, since it is based on the discrepancy of the final hybrid product and point-based flux tower estimates, which are essentially at the extremes of spatial discrepancy, the discrepancy between DOLCE and actual ET at any spatial scale greater than that of a tower footprint should be less than this uncertainty estimate (noting however that this is the estimated standard deviation of uncertainty, rather than a hard upper limit)

15. Table 5: specify unit of the trends. We thank the referee for their suggestion. We have now specified the unit of the trend as mm year-1. 16. Figure 1: is this necessary to include in the main paper, or could it be supplementary materials? We thank the referee for their suggestion. We have now moved this Figure to the supplementary material. 17. Figure 2: idem We thank the referee for their suggestion. We have now moved this Figure to the supplementary material. 18. Figure 3: can more distinguishable lines styles (i.e. color, thickness etc) be used better allow interpreting this figure? We thank the referee for their suggestions. We have now improved this figure and made the lines more distinguishable. 19. L759: reliable or robust? We have now rephrased the caption: Spatial pattern of ET climate trends in DOLCE V3 over 1980 – 2018 derived using Mann-Kendall and Sen's slope methods. Grid cells in white correspond to unreliable ET trends because (i) the confidence interval of the slope encompasses a mix of negative and positive values; or (ii) trends' slopes computed for multiple different random samples of ET within

Please also note the supplement to this comment:
https://hess.copernicus.org/preprints/hess-2020-595/hess-2020-595-AC2-supplement.pdf

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-595, 2020.

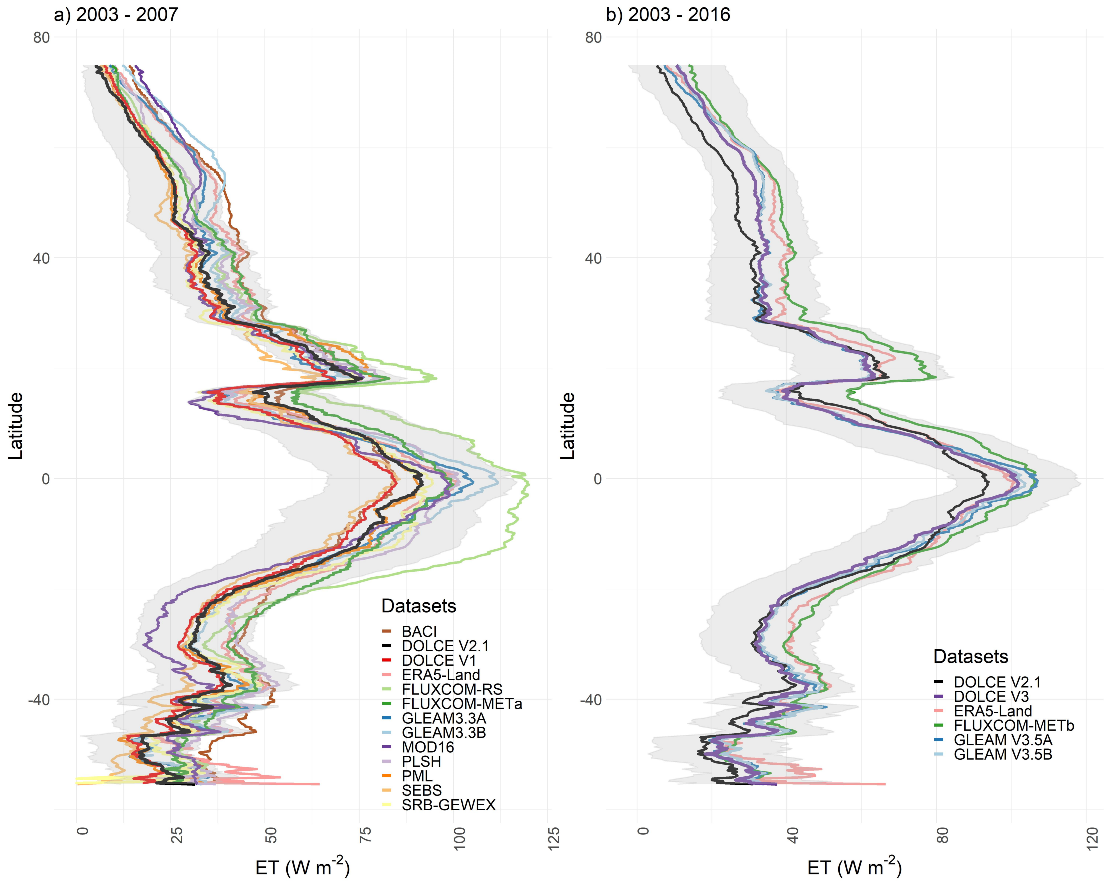
a) 2003 - 2007
b) 2003 - 2016
Latitude
ET (W m-2)
Datasets
BACI
DOLCE V2.1
DOLCE V1
ERA5-Land
FLUXCOM-RS
FLUXCOM-METa
GLEAM3.3A
GLEAM3.3B
MOD16
PLSH
PML
SEBS
SRB-GEWEX
Datasets
DOLCE V2.1
DOLCE V3
ERA5-Land
FLUXCOM-METb
GLEAM V3.5A
GLEAM V3.5B

Fig. 1.

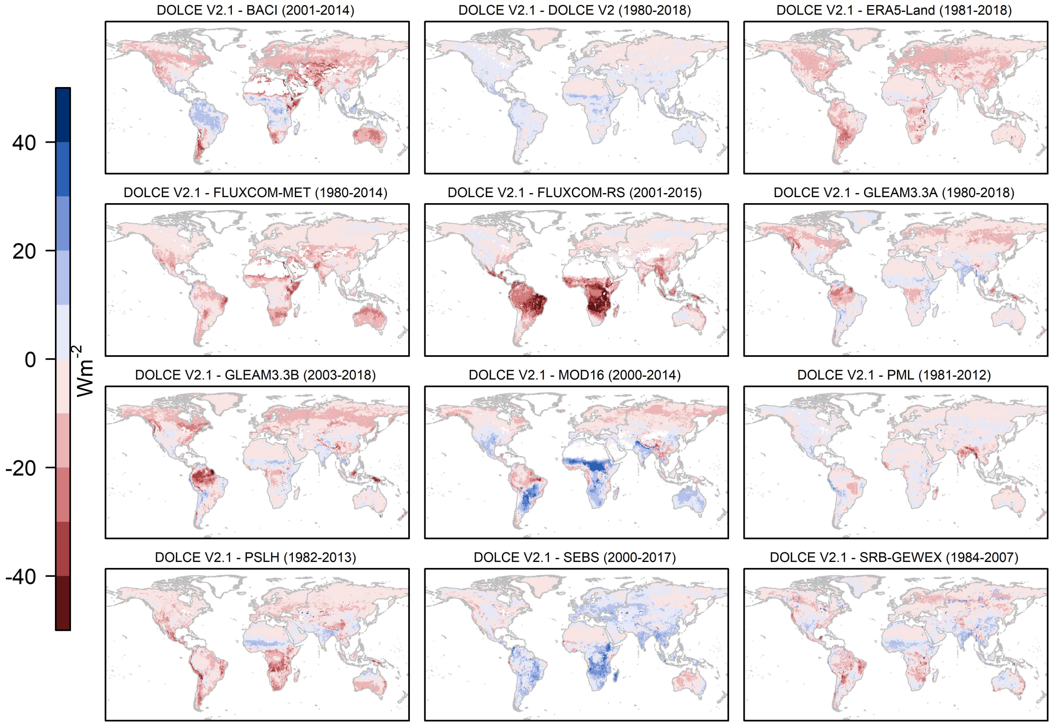
40
20
0
-20
-40
Wm^-2
DOLCE V2.1 - BACI (2001-2014)
DOLCE V2.1 - DOLCE V2 (1980-2018)
DOLCE V2.1 - ERA5-Land (1981-2018)
DOLCE V2.1 - FLUXCOM-MET (1980-2014)
DOLCE V2.1 - FLUXCOM-RS (2001-2015)
DOLCE V2.1 - GLEAM3.3A (1980-2018)
DOLCE V2.1 - GLEAM3.3B (2003-2018)
DOLCE V2.1 - MOD16 (2000-2014)
DOLCE V2.1 - PML (1981-2012)
DOLCE V2.1 - PSLH (1982-2013)
DOLCE V2.1 - SEBS (2000-2017)
DOLCE V2.1 - SRB-GEWEX (1984-2007)

Fig. 2.

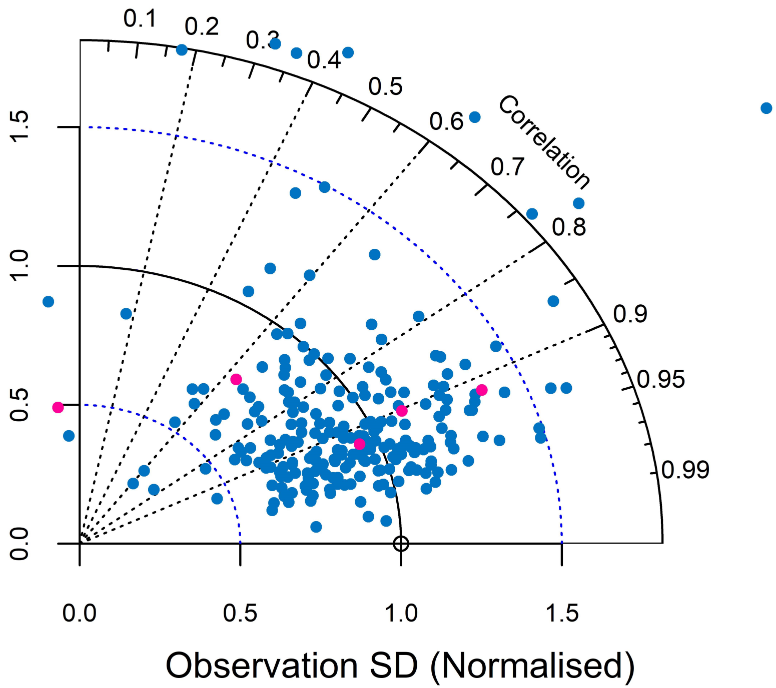
0.1
0.2
0.3
0.4
0.5
0.6
Correlation
0.7
0.8
0.9
0.95
0.99
0.0
0.5
1.0
1.5
0.0
0.5
1.0
1.5
Observation SD (Normalised)

Fig. 3.

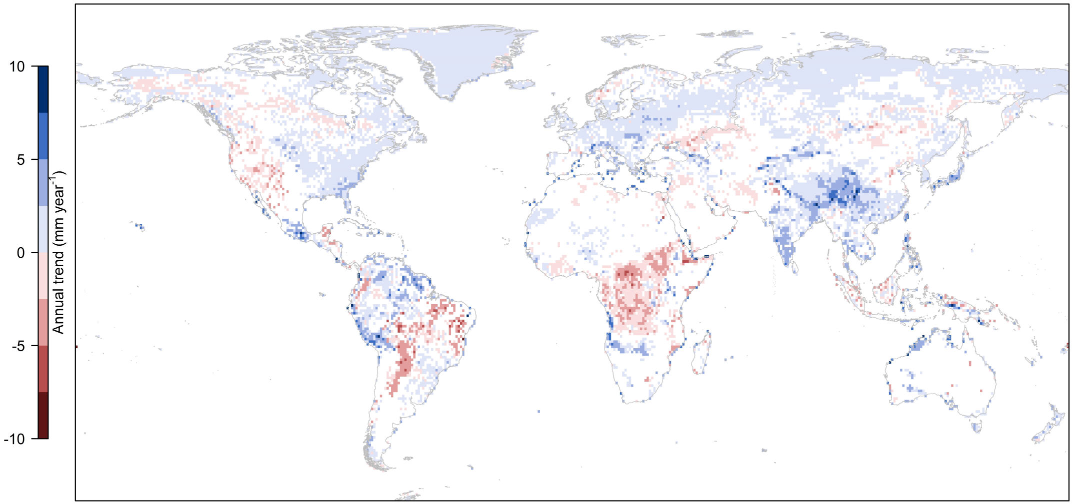

Fig. 4.

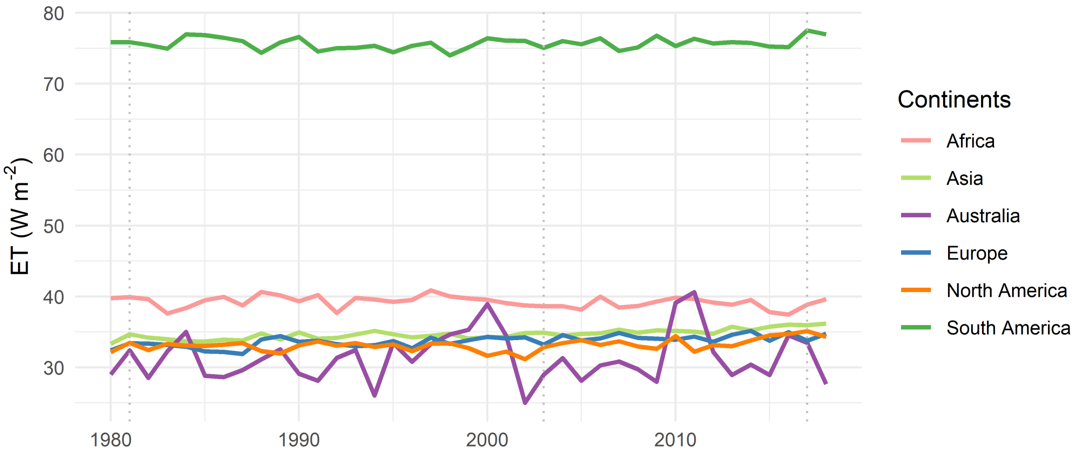
ET (W m-2)
80
70
60
50
40
30
1980
1990
2000
2010
Continents
Africa
Asia
Australia
Europe
North America
South America

Fig. 5.

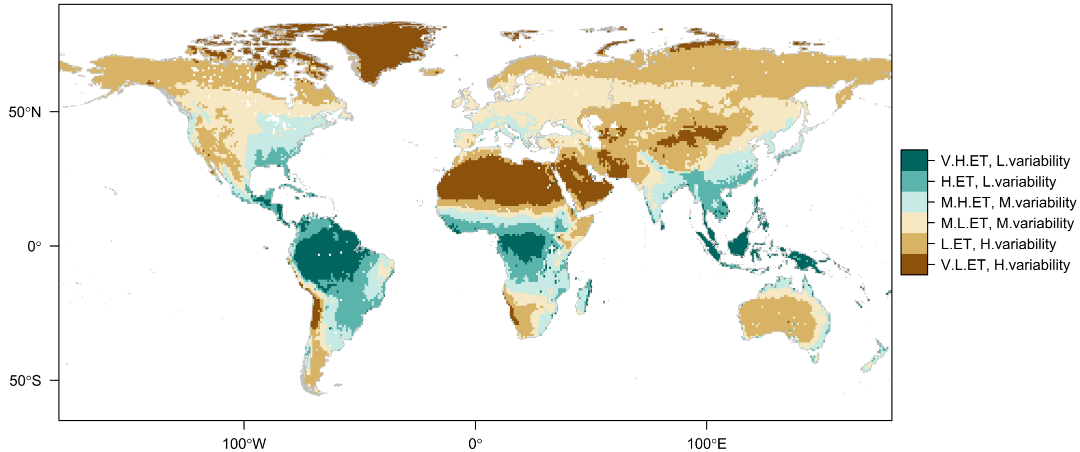
50°N
0°
50°S
100°W
0°
100°E
V.H.ET, L.variability
H.ET, L.variability
M.H.ET, M.variability
M.L.ET, M.variability
L.ET, H.variability
V.L.ET, H.variability

Fig. 6.

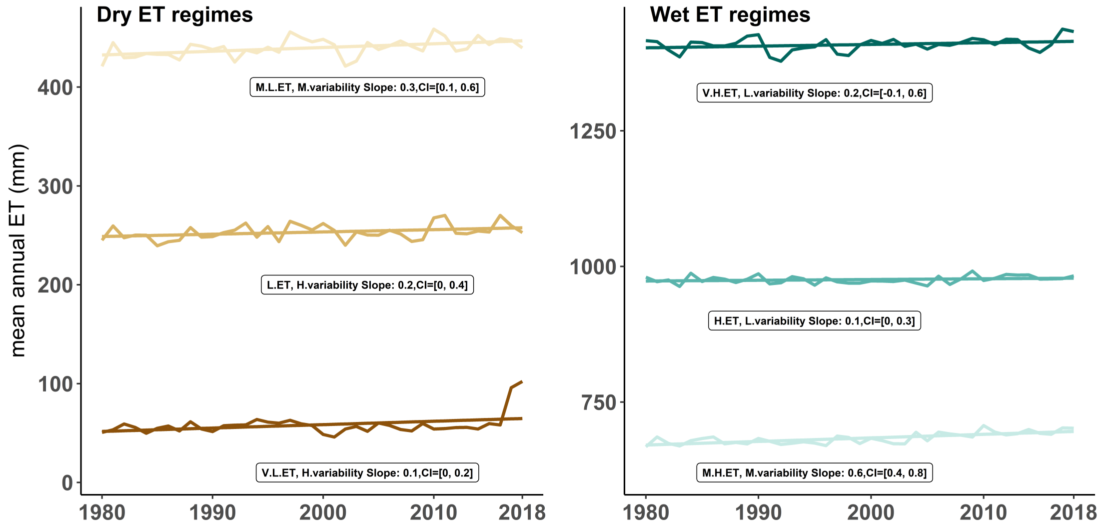
Dry ET regimes
Wet ET regimes
mean annual ET (mm)
M.L.ET, M.variability Slope: 0.3,CI=[0.1, 0.6]
L.ET, H.variability Slope: 0.2,CI=[0, 0.4]
V.L.ET, H.variability Slope: 0.1,CI=[0, 0.2]
V.H.ET, L.variability Slope: 0.2,CI=[-0.1, 0.6]
H.ET, L.variability Slope: 0.1,CI=[0, 0.3]
M.H.ET, M.variability Slope: 0.6,CI=[0.4, 0.8]
0
100
200
300
400
750
1000
1250
1980
1990
2000
2010
2018

Fig. 7.