2022/7/25

Reply to Reviewers

Dear Reviewers:

We would like to thank you for reviewing our manuscript. We appreciate your time and effort in considering this manuscript for publication.

We address all comments carefully point-by-point, and corresponding contexts have been incorporated into the revised manuscript. Our answers to every question/comment are provided below.

NOTE: In our answers to every question/comment, the newly added content is in blue, and the revised content is in red.

Reviewer 1

General Comments:

This manuscript addresses a useful but challenging topic. Although there are many remotely sensed global or regional evapotranspiration (ET) datasets, their performances varied across different biomes or regions due to high uncertainty exist in ET estimates. The current manuscript provides a good try to retrieve a land ET product in 2001-2020 using a three-temperature model without resistance and parameter calibration, which is different from the available ET products generated by methods including Penman-Monteith equation-based and surface-energy-balance-residual-based methods. The validation performed at different scales sound good. The intent of the manuscript is worthy and significant, and the topic generally fits the scope of the Earth System Science Data. The manuscript is well-written, and the methods, results, and discussion are clearly presented. Seeing the potential of this, I am in general supportive of publication after minor revision.

Responses:

Thank you for appreciating our work and considering that the studied topic is a much-needed idea. With the help of your constructive comments and suggestions, we believe that our manuscript will be improved substantially.

Specific Comments:

1. It may be better to clearly state the temporal resolution and duration of the product.

Responses:

Thank you very much for the valuable suggestion. The specific temporal resolution and duration of the product have been stated in the abstract as follows:

Abstract. Accurate global terrestrial evapotranspiration (ET) estimation is essential to better understand Earth's energy and water cycles. Although several global ET products exist, recent studies indicate that ET estimates exhibit high uncertainty. With the increasing trend of extreme climate hazards (e.g., droughts and heat waves), accurate ET estimation under extreme conditions remains challenging. To overcome these challenges, we used 3-hour and 0.25° Global Land Data Assimilation System (GLDAS) datasets (net radiation, land surface temperature (LST), and air temperature) and a three-temperature (3T) model, without resistance and parameter calibration, in global terrestrial ET product development. The results demonstrated that the 3T model-based ET product agreed well with both global eddy covariance (EC) observations at daily (root mean square error (RMSE) = 1.1 mm day^{-1}, N=294058) and monthly scales

(RMSE = 24.9 mm month^{-1}, N=9632) and basin-scale water balance observations (RMSE =116.0 mm yr^{-1}, N=34). The 3T model-based global terrestrial ET product was comparable to other common ET products, i.e., MOD16, P-LSH, PML, GLEAM, GLDAS, and Fluxcom, retrieved from various models, but the 3T model performed better under extreme weather conditions in croplands than did the GLDAS, attaining 9.0–20% RMSE reduction. The proposed daily and 0.25° ET product covering the period of 2001-2020, could provide periodic and large-scale information to support water cycle-related studies. The dataset is freely available at the Science Data Bank (http://doi.org/10.57760/sciencedb.o00014.00001, Xiong et al., 2022).

2. Please add the data points used for validation.

Responses:

Thank you. The missing information has been added to the abstract as well as to related figures. Please see details in our previous reply to comment 1.

3. "the energy balance product" is better changed to "the ET product".

Responses:

The words have been revised as suggested:

Several global ET estimates have been developed over the past two decades based on various theories, including 1) surface energy balance residual methods, e.g., the ET product based on the Surface Energy Balance System (SEBS) (EB) (Chen et al., 2021);

4. PMLv2 → PML. Missing a PT-based ET product, GLASS.

Responses:

The "PMLv2" has been changed to "PML" in the revised manuscript. The ET product, GLASS, produced using multiple methods, including the PT equation, has been added in this revision as follows:

2) Penman–Monteith (PM) and Priestley–Taylor (PT) equation-based methods, e.g., MOD16 (Mu et al., 2011), P-LSH (Zhang et al., 2015), PML (Zhang et al., 2019), and GLEAM (Martens et al., 2017; Miralles et al., 2011); 3) land surface models, e.g., the Global Land Data Assimilation System (GLDAS) (Rodell et al., 2004); 4) multimodel ensemble approach, e.g., GLASS (Yao et al., 2014), Hi-GLASS (Yao et al., 2017), and a synthesized ET product (Elnashar et al., 2021); and 5) empirical methods, e.g., Fluxcom (Jung et al., 2019).

Reference:

Yao, Y., Liang, S., Li, X., Hong, Y., Fisher, J. B., Zhang, N., Chen, J., Cheng, J., Zhao, S., and Zhang, X.: Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations, J. Geophys. Res.: Atmos., 119, 4521-4545, https://doi.org/10.1002/2013JD020864, 2014.

Yao, Y., Liang, S., Li, X., Zhang, Y., Chen, J., Jia, K., Zhang, X., Fisher, J. B., Wang, X., and Zhang, L.: Estimation of high-resolution terrestrial evapotranspiration from Landsat data using a simple Taylor skill fusion method, J. Hydrol., 553, 508-526, https://doi.org/10.1016/j.jhydrol.2017.08.013, 2017.

5. Fluxcom also has no value in some arid regions.

Responses:

Thank you very much for your suggestion. The reference for the ET product Fluxcom has been added to the sentence as follows:

Consequently, ET in these arid regions has usually been assumed as zero in certain ET products (Mu et al., 2011; Jung et al., 2019).

6. It is better to provide a journal article as the thesis may not be available.

Responses:

Thank you very much for your suggestion. The thesis has been replaced with a journal article as follows:

This model mainly utilizes net solar radiation, surface temperature and air temperature as model inputs. In this model, the resistance terms in the energy balance equation are eliminated via the introduction of a dry surface without evaporation or transpiration, as detailed in Qiu et al., (1999).

Reference:

Qiu, G. Y., Ben-Asher, J., Yano, T., and Momii, K.: Estimation of soil evaporation using the differential temperature method, Soil Sci. Soc. Am. J., 63, 1608-1614, https://doi.org/10.2136/sssaj1999.6361608x, 1999.

7. "G equals to 0.315Rn" may be misleading.

Responses:

Thank you very much for your suggestion. The sentence "which equals $0.315\times R_n$." has been revised and moved to Section 2.2 as follows:

The ground heat flux (G) can be directly extracted from net radiation R_n according to Su (2002).

Reference:

Su, Z.: The surface energy balance system (SEBS) for estimation of the turbulent heat fluxes, Hydrol. Earth Syst. Sci., 6 (1), 85–99, https://doi.org/ 10.5194/hess-6-85-2002, 2002.

8. The subscript l is not consist with those in equations 10 and 11.

Responses:

Thank you very much for your careful reading. The inconsistency has been revised as follows:

$$R_{n,sr} = \overline{R_{n,s}\left[T_{s,upper5\%}\right]} = \text{mean}\left(R_{n,s1}, R_{n,s2}, \ldots, R_{n,sj}\right) , \tag{10}$$

$$R_{n,cr} = \overline{R_{n,c}\left[T_{c,upper5\%}\right]} = \text{mean}\left(R_{n,c1}, R_{n,c2}, \ldots, R_{n,cj}\right) , \tag{11}$$

where $R_{n,sj}$ and $R_{n,cj}$ denote the soil and vegetation net radiation values, respectively, corresponding to pixel j (j=1, 2, 3…) of the upper 5% T_s and T_c values, respectively, within the same subregion.

9. Order of the section title was wrong as well as the following section. The authors should proof read the manuscript to avoid such mistakes.

Responses:

Thank you very much. We carefully checked our manuscript, and the order of the section title has been corrected as follows:

2.3.2 Evaluation considering the water budget in global main catchments

2.3.3 Evaluation via comparison to other commonly used global ET products

10. TWSC is better replaced with ΔS. In L171, it is better to use annual TWSC.

Responses:

Thank you very much for your suggestion. The TWSC has been replaced with ΔS as suggested, as follows:

$$ET_{wb} = P - R - \Delta S \quad , \tag{12}$$

where P, R and ΔS are the precipitation (mm yr^{-1}), runoff (mm yr^{-1}), and terrestrial water storage change (mm yr^{-1}), respectively, in a given catchment. Annual ΔS can be calculated as the terrestrial water storage anomaly (TWSA) difference between Decembers of the target year and its previous year.

11. Typos (the unit). The authors should proof read the manuscript to avoid such mistakes.

Responses:

Thank you very much for your careful reading. The unit "mm yr-1" has been corrected as follows:

The estimated mean ET value was 514.5 mm yr^{-1}, with a standard deviation of 211 mm yr^{-1}, whereas the mean ET_{wb} value reached 476.5 ± 280 mm yr^{-1}.

12. The value 133 mm/yr was from what data?

Responses:

The value was derived from the water balance equation. The sentence has been modified for clarity as follows:

These river basins were mainly located at high latitudes (approximately 60° North) with relatively low ET_{wb} values (133 ± 50 mm yr^{-1}).

13. It seems the ET should be removed.

Responses:

We carefully checked the manuscript, and the comment may be related to "Nonetheless, the above results generally suggest that the 3T model performance was comparable to that of the water balance ET equation". We deleted ET in this revision as follows:

Nonetheless, the above results generally suggest that the 3T model performance was comparable to that of the water balance equation.

14. PMLv2 looks curious. "v2" may be the version number, suggesting delete it across the entire manuscript but remain a statement somewhere.

Responses:

Thank you very much for your suggestion. The version of ET products PML, "v2", has been removed across the entire manuscript, except in section 2.3.3, where we introduced the ET products as follows:

Among the selected ET products, three products were based on the PM model with varying resistance parameterization schemes, i.e., MOD16 (version 6, Mu et al., 2011), P-LSH (Zhang et al., 2015), and PML (version 2, Zhang et al., 2019), while the remaining three products were based on the PT model (GLEAM version 3.5a; Miralles et al., 2011; Martens et al., 2017), land surface models (GLDAS-Noah version 2.1; Beaudoing and Rodell, 2020; Rodell et al., 2004), and machine learning (Fluxcom; Jung et al., 2019).

15. It is better to clearly state that the EC datasets are the same as those used in figure 2.

Responses:

Thank you very much for your suggestion. The caption of Figure 4 (Figure 5 in this revision) has been improved as follows:

Figure 5: Validation of 6 commonly used ET products (GLDAS, PML, P-LSH, GLEAM, Fluxcom, and MOD16) against EC tower observations. The data are monthly average ET values over the 2003–2013 period and are the same as those used in Fig. 3a.

16. Texts in the figures are too small to read. I suggest the authors enlarge these texts to improve their quality and readability.

Responses:

We tried our best to improve the quality of the figures, including the enlarged text in Figures 2, 4, and 5. Figure 2 (Figure 3 in this revision) shows an example as follows:

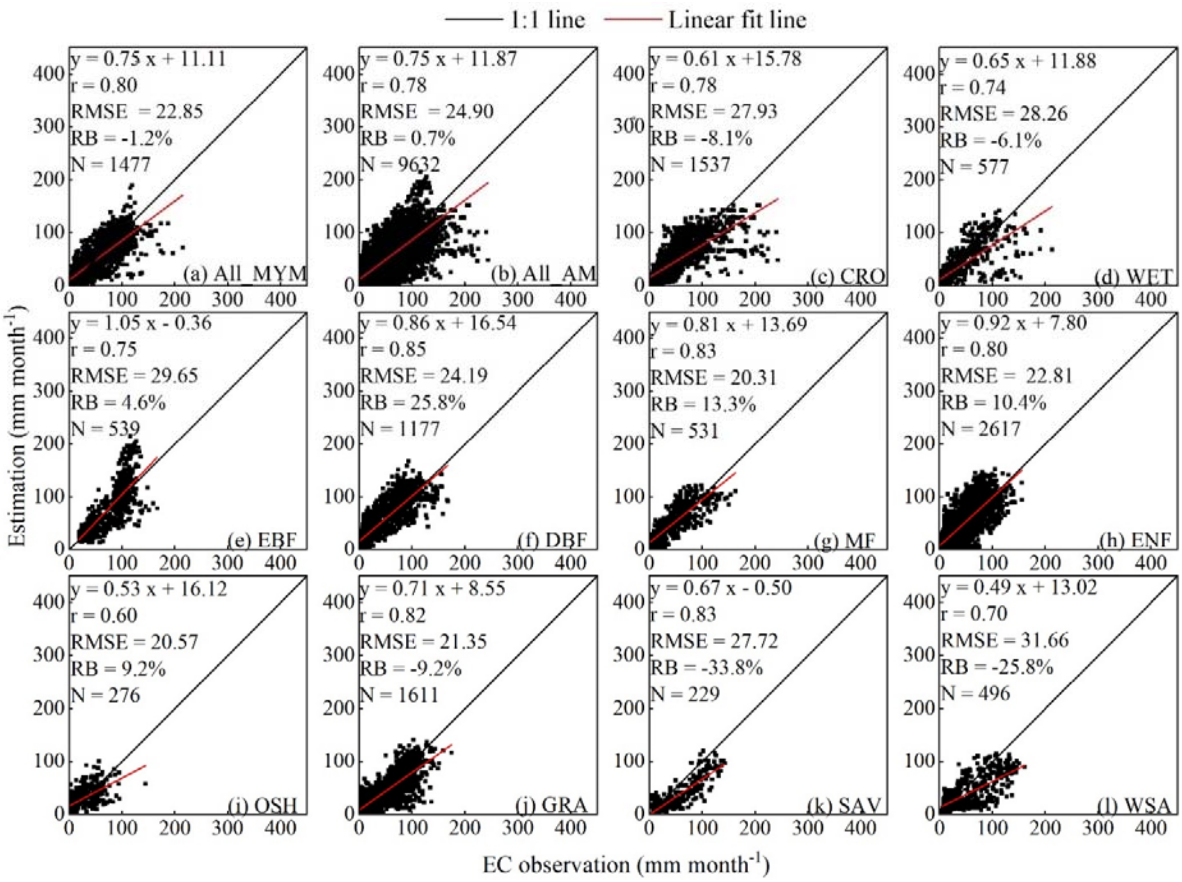

Figure 3: Comparison of the estimated (3T model) and measured (EC tower) monthly ET values from 2003–2013, where (a) shows the data for all 126 sites on a multi-year monthly mean (MYM) scale and (b) shows the data for all sites on an annual mean (AM) monthly scale. (c)-(l) show all land use/land cover types on an annual monthly scale. The abbreviations in (c)-(l) are the same as those in Fig. 1.

17. What does “PFT” mean? Please consider define such abbreviation.

Responses:

“PFT” in Figures 11-13 has been removed and replaced by “Land use land cover” in the revised manuscript.

18. “the whiskers indicate the extreme values” should be “the whiskers indicate the outlier values”.

Responses:

The “extreme” in the caption of Figure 11 has been replaced by “outlier” in the revised manuscript.

Figure 11: Monitoring performance of the 3T model-based terrestrial ET product under extreme heat conditions in the different biomes. The daily ET is shown in energy units. In the box plot (a), the black point indicates the mean, while central line in the box indicates the median value. The edges of the box indicate the 25th and 75th percentiles, and the whiskers indicate the outlier values. In the violin plot (b), the white point indicates the median value, and a wider violin plot indicates denser data for the same RMSE value. N denotes the number of data points.

19. Area of the Antarctica should be wrong.

Responses:

The area of Antarctica is 0.14×10^8 km^2 and has been corrected as follows:

Table 3: Multi-year (2003–2013) average ET values considering the water depth (mm yr^{-1}) and volume (km^3 yr^{-1}) of the different products used in this study for the global land surface.

ET products	ET rate (mm yr^{-1})	ET volume ($\times10^3$ km^3 yr^{-1})
3T	546 ± 22	73.8 ± 3.0
Fluxcom	549 ± 3	74.2 ± 0.4
GLDAS	551 ± 10	74.5 ± 1.3
GLEAM	544 ± 6	73.6 ± 0.7
MOD16	468 ± 6	63.3 ± 0.8
P-LSH	551 ± 8	74.5 ± 1.0
PML	542 ± 12	73.2 ± 1.7

Note: global land surface has an area of 1.35×10^8 km^2, excluding Antarctica. Fluxcom and MOD16 do not provide ET values in Greenland and desert areas.

Reviewer 2

General Comments:

This manuscript uses the classic 3T model to generate global terrestrial evapotranspiration (ET) product at 0.25 degree resolution and 3hour temporal resolution. I am glad to see the 3T model has been expended to a global scale. The authors also tested their estimates against flux observations at a monthly scale and water balance ET estimates at an annual scale. Overall, the paper is well written, and the key message is get crossed. Having said that, I have some major concerns on the method applied to the global scale, and validation processes. Follows are the key concerns.

Responses:

We thank the reviewer for the positive evaluation and consideration that generating a global ET product is a much-needed idea. Although there are numerous ET products that have been rigorously evaluated, notable disagreement exists among these products, indicating that high uncertainties remain in ET estimates and products. One highlight of our study is that we retrieve a land ET product using a three-temperature model without resistance and parameter calibration, which is different from the available ET products generated by methods including Penman–Monteith equation-based and surface-energy-balance-residual-based methods. Our validation results indicated that the proposed ET product has reasonable accuracy. However, some misunderstandings occurred due to our unclear description; for example, validation of the ET product was verified at the daily scale under extreme conditions. Nonetheless, we further evaluated the ET product at 3-hour and daily scales in this revision as suggested. Please see the following point-by-point replies for details.

Comments 1:

The author grouped the climate regimes into 5 groups according to Koppen-geiger climate classification, then assumes the means of reference net radiation values of the soil and vegetation components are similar within the same group of samples. I am afraid that the classification is too coarse, which will result in large uncertainty in estimating the two important parameters. I think the more details should be exhibited and the uncertainty should be quantified.

Responses:

The Köppen-Geiger climate regimes consist of 5 groups (equatorial, arid, warm temperate, snow, and polar), but they were subdivided into 31 climatic regions according to different conditions of precipitation and temperature (Kottek et al., 2006).

In fact, in our study, we further classified these 31 climatic regions of the Köppen-Geiger climate regimes into more detailed subregions via principal component analysis (PCA) and K-means clustering methods. Specifically, PCA was used to select major variables to describe regional characteristics from meteorological factors (i.e., net radiation, air temperature, humidity, wind speed, precipitation, and air pressure) and land surface conditions (i.e., albedo, land surface temperature, NDVI, soil moisture, and soil temperature). Thereafter, these variables were used to classify the 31 climate regions through the K-means clustering method. Because of meteorological and land surface variations, the subregions varied from 90 to 110 in different months. As such, each classified subregion has an area of approximately 1.35×10^6 km^2 (considering 100 subregions) and may be good enough for the reference assumption of the 3T model at the global scale.

To test the impact of region size on the 3T model-based ET estimation, we performed a comparison under two different classification methods with different subregions and sizes. In the first method, Köppen-Geiger climate regimes with 31 subregions were used, whereas detailed subregions with numbers of 90-110 were adopted as the second method. The daily ET estimates in 2011 were used as examples. In general, the two groups of daily ET estimates showed little difference, with mean values of 47 and 42 W m^{-2}, respectively, and were close to the EC observation, with RMSE values of 32 and 33 W m^{-2} (see the following text for details). At the yearly scale, however, the 3T model-based ET estimates from 90-110 subregions were much closer to the water balance ET than those estimates from 31 subregions (see the following text for details). The results indicate that the smaller the region where the reference parameters were obtained, the more accurate the 3T model, which is consistent with our previous findings (Xiong et al., 2015, 2019).

The uncertainty of the 3T model at the global scale caused by region size have been incorporated into section 4.1 in the revised manuscript, as follows:

4 Discussion

4.1 Characteristics of the global terrestrial ET product based on the 3T model

As indicated in section 3, the 3T model-based global terrestrial ET product agreed well with ground observations and was comparable to other commonly used ET products. Particularly, the determined global terrestrial (excluding Antarctica) ET volume (in units of 10^3 km^3 yr^{-1}) based on 3T model-based estimates reached 73.8 from 2003-2013, which is not only consistent with that determined based on the other ET products (excluding MOD16), as indicated in section 3.3, ranging from 73.2 to 74.5 (Table 3), but also consistent with values reported in other studies, e.g., 72.3±0.9, as obtained with a complementary relationship-based ET product from 1982–2016 (Ma et al., 2021), and 71.1, as determined with a water balance and machine learning-based ET product from 1982–2009 (Zeng et al., 2014).

It should be noted that the 3T model differed from the methods used to estimate ET in the adopted ET products, as listed in Table 2. …

Although the 3T model-based ET estimates suffer from the domain size when determining the reference site, the uncertainty may be limited. We tested the difference between the 3T model-based ET values estimated using two different regimes with different subregions and sizes. Specifically, Köppen-Geiger climate regimes with 31 subregions and detailed subregions with numbers of 90-110 were used. In general, the two groups of daily ET estimates in 2011 showed little difference, with mean ET values of 47 and 42 W m^{-2}, respectively and were close to the EC observation, with RMSE values of 32 and 33 W m^{-2} (Figs. 10a and 10b, respectively). At a yearly scale, however, the 3T model-based ET estimates from 90-110 subregions (Fig. 10d) were much closer to the water balance ET than those estimates using 31 sub-regions (Fig. 10c). The results indicate that the smaller the domain size of a region where the reference parameters were obtained, the more accurate the 3T model, which is consist with our previous findings (Xiong et al., 2015, 2019).

In addition, the 3T model-based global terrestrial ET product required less data in terms of model inputs than those required by the adopted ET products, as listed in Table 2. …

Reference:

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World map of the Köppen-Geiger climate classification updated, Meteorol. Z., 15, 259-263, https://doi.org/10.1127/0941-2948/2006/0130, 2006.

Xiong, Y., Zhao, S., Tian, F., and Qiu, G. Y.: An evapotranspiration product for arid regions based on the three-temperature model and thermal remote sensing, J. Hydrol., 530, 392-404, https://doi.org/10.1016/j.jhydrol.2015.09.050, 2015.

Xiong, Y., Zhao, W., Wang, P., Paw U, K. T., and Qiu, G. Y.: Simple and applicable method for estimating evapotranspiration and its components in arid regions, J. Geophys. Res.: Atmos., 124, 9963-9982, https://doi.org/10.1029/2019JD030774, 2019.

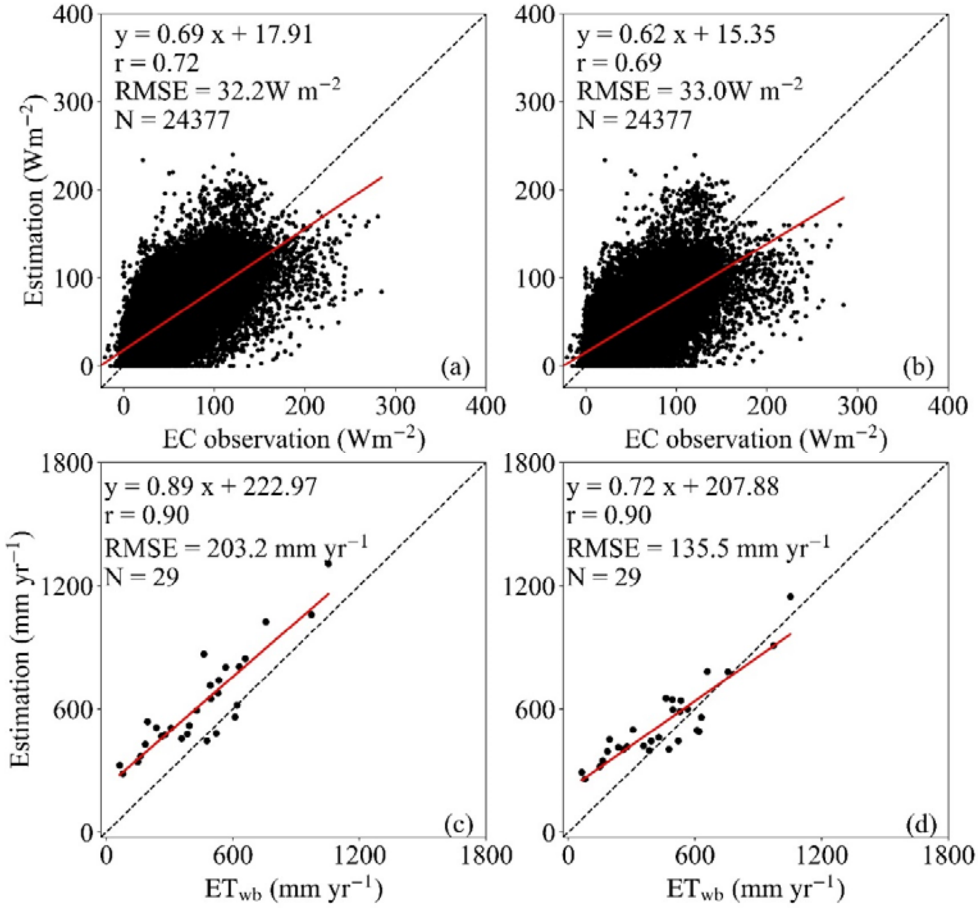

Figure 10 Comparison of the estimated (3T model) and measured ET values in 2011 (daily ET from EC tower and annual ET from water balance equation). The left panel shows ET estimates using Köppen-Geiger climate regimes with 31 subregions at daily (a) and annual (c) scales, respectively, whereas the right panel is the same but with ET estimates using 90-110 subregions.

Comments 2:

The validation of the 3T products is a bit weird. The benefit to use the 3-T model is to detect ET variation in a short period of time. Considering that the 3-T products has been run at 3-hour temporal scale, its robustness should be demonstrated at 3-hour or daily scale at least. Currently, the authors focused their validations against flux measurements at a monthly scale, which is too coarse to be acceptable.

Responses:

In fact, validation of the 3T model is performed not only at a monthly scale but also at a daily scale. In particular, the discussion that the 3T model-based ET product could accurately capture the low ET values under extreme conditions in section 4.2 used daily ET estimates. We apologize for our unclear description.

We further tested the performance of the 3T model with all daily EC observations (because the results in section 4.2 only contain extreme conditions). The results showed that the 3T model-based ET estimates agreed well with the observations (N=294058), with an RMSE of 33 W m^{-2} (or 1.1 mm day^{-1}) (see the following text for details), which was also comparable to other ET products, such as GLDAS (RMSE: 32 W m^{-2} or 1.1 mm day^{-1}) (see the following text for details), PML (RMSE: 0.7 mm day^{-1}) (Zhang et al., 2021), and SEBS (RMSE: 1.6 mm day^{-1}) (Chen et al., 2021).

At a 3-hour temporal scale, the data are too large to perform an entire comparison at a global scale. Moreover, as you may know, high temporal data even at the daily scale, especially remote sensing data, may encounter missing values for several reasons, such as clouds and precipitation (the MODIS land surface temperature product is a good example), which complicates the comparison and contains more uncertainty. In fact, the performance of the 3T model over a short period of time (30 minutes to 1 hour) has been tested at the point scale, and the results from both our studies (Qiu and Zhao, 2010; Tian et al., 2014; Qiu et al., 2015; Xiong et al., 2019) and others (Zhou et al., 2014; Zhang et al., 2020) show that the 3T model generally performed well. Nonetheless, we further tested the performance of the 3T model at the 3-hour scale across the world using EC observations in 2011 (6278 data points) (see the following text for details). Data were selected from the 15th day of each month in 2011. Although the RMSE (74 W m^{-2}) was slightly greater than that at the daily scale compared to the EC observations, the 3T model-based ET estimates at the 3-hour scale agreed well with the GLDAS ET, with an r of 0.89 and RMSE of 21 W m^{-2} (see the following text for details).

These results indicate that the 3T model is robust at different temporal scales. These results have been incorporated into section 3.1 in the revised manuscript, as follows:

3 Results

3.1 Performance of the 3T product versus the global EC network

At the daily scale, the 3T model-based ET estimates agreed well with the observation (N=294058), with RMSE of 32 W m^{-2} (or 1.1 mm day^{-1}) (Fig. 2a), which was comparable to other ET products, such as GLDAS (RMSE: 32 W m^{-2} or 1.1 mm day^{-1}) (Fig. S1), PML (RMSE: 0.7 mm day^{-1}) (Zhang et al., 2021), and SEBS (RMSE: 1.6 mm day^{-1}) (Chen et al., 2021). Moreover, the 3T model could capture the change trend of daily ET because comparison results at 10 EC sites covering various biomes in both the Southern (Figs. 2b-2c) and Northern Hemispheres (Figs. 2d-2k) indicate that interannual variability of the estimates were close to that of the observed ET. A comparison at an instantaneous 3-hour scale was also performed to test the ET estimates. EC observations across the world for the 15th day of each month in 2011 (N=6278) were compared because the data are too large to perform an entire comparison at a global scale. Although the RMSE (74 W m^{-2}) was a slightly greater than that at daily scale (Fig. S2a), the 3T model-based ET estimates at the 3-hour scale agree well with the GLDAS ET, with an r of 0.89 and RMSE of 21 W m^{-2} (Fig. S2b). The explanation is likely that high temporal data may encounter missing values, which complicates the comparison.

At the monthly scale, the paired ET values between the 3T model and EC observations were generally distributed on both sides of the 1:1 line, revealing relatively large differences at a few points for ET values higher than 100 mm month^{-1}, and resulting in regression line slope and r values of 0.75 and 0.80, respectively (Fig. 3a). …

Reference:

Chen, X., Su, Z., Ma, Y., Trigo, I., and Gentine, P.: Remote Sensing of Global Daily Evapotranspiration based on a Surface Energy Balance Method and Reanalysis Data, J. Geophys. Res.: Atmos., 126, e2020JD032873, https://doi.org/10.1029/2020JD032873, 2021.

Qiu, G. Y. and Zhao, M.: Remotely monitoring evaporation rate and soil water status using thermal imaging and "three-temperatures model (3T Model)" under field-scale conditions, J. Environ. Monit., 12, 716-723, https://doi.org/10.1039/B919887C, 2010.

Qiu, G. Y., Li, C., and Yan, C.: Characteristics of soil evaporation, plant transpiration and water budget of Nitraria dune in the arid Northwest China, Agric. For. Meteorol., 203, 107-117, https://doi.org/10.1016/j.agrformet.2015.01.006, 2015.

Tian, F., Qiu, G. Y., Lü, Y., Yang, Y., and Xiong, Y.: Use of high-resolution thermal infrared remote sensing and "three-temperature model" for transpiration monitoring in arid inland river catchment, J. Hydrol., 515, 307-315, https://doi.org/10.1016/j.jhydrol.2014.04.056, 2014.

Xiong, Y., Zhao, W., Wang, P., Paw U, K. T., and Qiu, G. Y.: Simple and applicable method for estimating evapotranspiration and its components in arid regions, J. Geophys. Res.: Atmos., 124, 9963-9982, https://doi.org/10.1029/2019JD030774, 2019.

Zhang, Y., Kong, D., Gan, R., Chiew, F. H., McVicar, T. R., Zhang, Q., and Yang, Y.: Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017, Remote Sens. Environ., 222, 165-182, https://doi.org/10.1016/j.rse.2018.12.031, 2019.

Zhang, Y., Qin, H., Zhang, J., and Hu, Y.: An in-situ measurement method of evapotranspiration from typical LID facilities based on the three-temperature model, J. Hydrol., 588, 125105, https://doi.org/10.1016/j.jhydrol.2020.125105, 2020.

Zhou, X., Bi, S., Yang, Y., Tian, F., and Ren, D.: Comparison of ET estimations by the three-temperature model, SEBAL model and eddy covariance observations, J. Hydrol., 519, 769-776, https://doi.org/10.1016/j.jhydrol.2014.08.004, 2014.

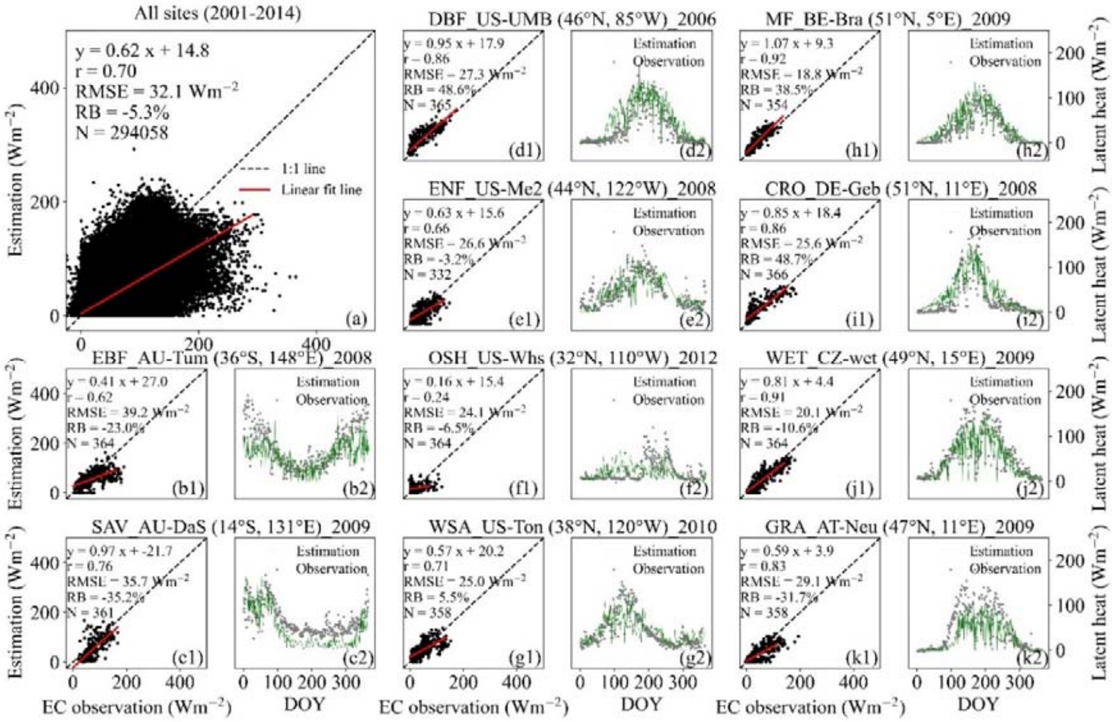

Figure 2. The temporal variations in daily ET estimated from the 3T model (green line) using EC observations (gray dot). The scatter plot between EC observations and ET estimates for one selected year (with RMSE values at average level) at ten EC sites covering various biomes: (a) all 126 sites, (b1-2) EBF_AU-Tum (36°S, 148°E) in 2008, (c1-2) SAV_AU-DaS (14°S, 131°E) in 2009, (d1-2) DBF_US-UMB (46°N, 85°W) in 2006, (e1-2) ENF_US-Me2 (44°N, 122°W) in 2008, (f1-2) OSH_US-Whs (32°N, 110°W) in 2012, (g1-2) WSA_US-Ton (38°N, 120°W) in 2010, (h1-2) MF_BE-Bra (51°N, 5°E) in 2009, (i1-2) CRO_DE-Geb (51°N, 11°E) in 2008, (j1-2) WET_CZ-wet (49°N, 15°E) in 2009, (k1-2) GRA_AT-Neu (47°N, 11°E) in 2009.

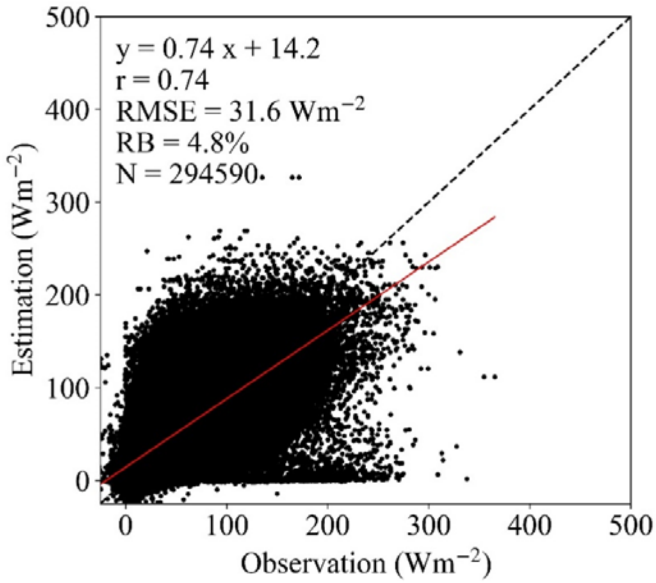

Figure S1 Comparison of the estimated (GLDAS) and measured (EC tower) daily ET values from 2001–2014.

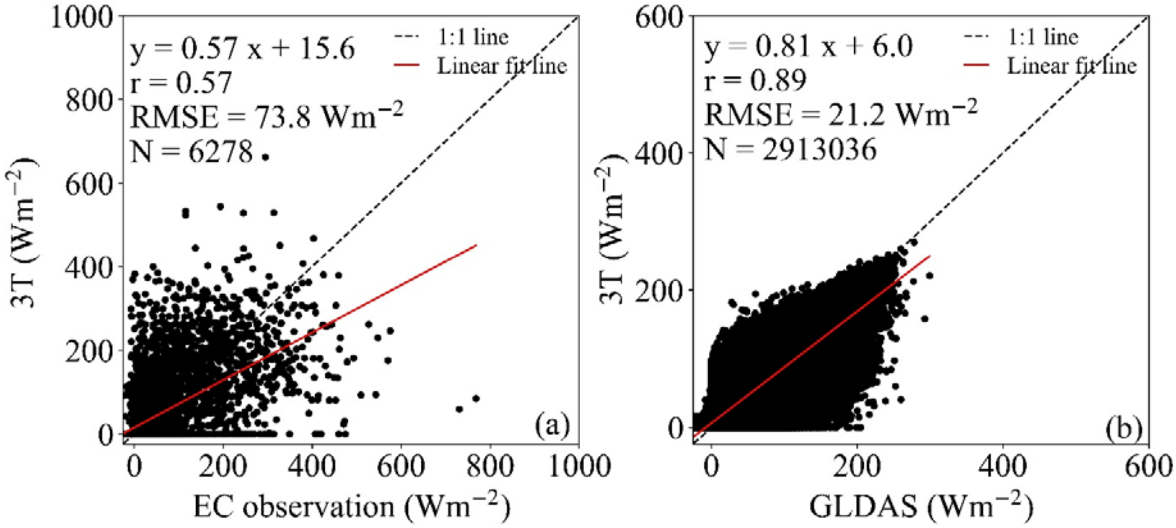

Figure S2. Validation of the 3T model-based ET estimates at the 3-hour temporal scale: (a) comparison with the EC observation and (b) comparison with GLDAS ET. Data were selected from the 15th day of each month in 2011.

Comments 3:

Calculation of water balance ET can be improved. The authors used GLDAS forcing data to drive the 3T model. However, the calculation of water balance ET relies on another precipitation product GPCC. I think the authors need to test the consistency between GLDAS precipitation and GPCC precipitation. In addition, I encourage the authors not only test its performance at a mean annual scale, but also need to test its interannual variability. This will demonstrate its full strength.

Responses:

In fact, the water balance ET was independent of the estimates from the 3T model because inputs of the 3T model only consist of net radiation (R_n), air temperature (T_a), and land surface temperature (LST) from GLDAS forcing data. Precipitation is not required in the inputs of the 3T model. Therefore, consistency between GLDAS precipitation and GPCC precipitation is unnecessary.

We agree with the comment that the performance of the 3T model should be tested at the interannual scale to enhance its robustness. We selected 10 EC sites covering different biomes across the world to test the interannual variability of the 3T model-based ET estimates at a daily scale. The variations in the 3T product generally fit the observation (see the following text for details). We also compared multiyear (2003–2013) mean monthly ET values between several ET products, and the results further indicate that the interannual variability of the 3T model was similar to the other ET products (see the following text for details). These results have been incorporated into section 3.1 (see details in the response to comment 2) and section 3.3 in the revised manuscript as follows:

3.3 Comparison of the 3T product to other global ET products

To further assess the performance of the 3T model across the various terrestrial land types, 3T model-based ET estimates were cross validated against six global ET products during the 2003–2013 period.

When EC observation data were adopted as a reference, 3T model-based ET estimates were comparable to GLDAS, GLEAM, and MOD16 data in terms of r and RMSE, with values of 0.8 and 22 mm month^{-1}, respectively (Fig. 3a and Figs. 5a to 5c). …

When ET_{wb} values were adopted as a reference, although the 3T model performance was slightly lower than that of the PML, GLEAM, and P-LSH products in terms of RMSE, with a value of 116 mm month^{-1} versus values of 96, 111, and 115 mm month^{-1}, respectively (Fig. 4a and Figs. 6a to 6c), the 3T model performed better than did the GLDAS (RMSE=120 mm month^{-1}), Fluxcom (RMSE=149 mm month^{-1}), and MOD16 products (RMSE=182 mm month^{-1}) (Figs. 6d to 6f). …

Via comparison of the terrestrial (excluding Antarctica) ET values retrieved from the various ET products, the mean ET value of the 3T model reached 546 mm yr^{-1} during the 2003–2013 period, whereas the mean ET values obtained with the MOD16, PML, GLEAM, Fluxcom, GLDAS, and P-LSH products reached 468, 542, 544, 549, 551 and 551 mm yr^{-1}, respectively.

In terms of interannual variation (excluding Antarctica, Greenland, and desert areas according to Jung et al. (2019)), the 3T model-based estimates were similar to the other six ET products with an increasing trend from January (approximately 40 mm month^{-1}) to July (approximately 65 mm month^{-1}) and then a decreasing trend through the following months (Fig. 7a). The latitudinal distribution of the values obtained with each ET product was also determined, and the changing trend of the 3T model-based ET values was similar to that of the values obtained with the six considered ET products (Fig. 7b). …

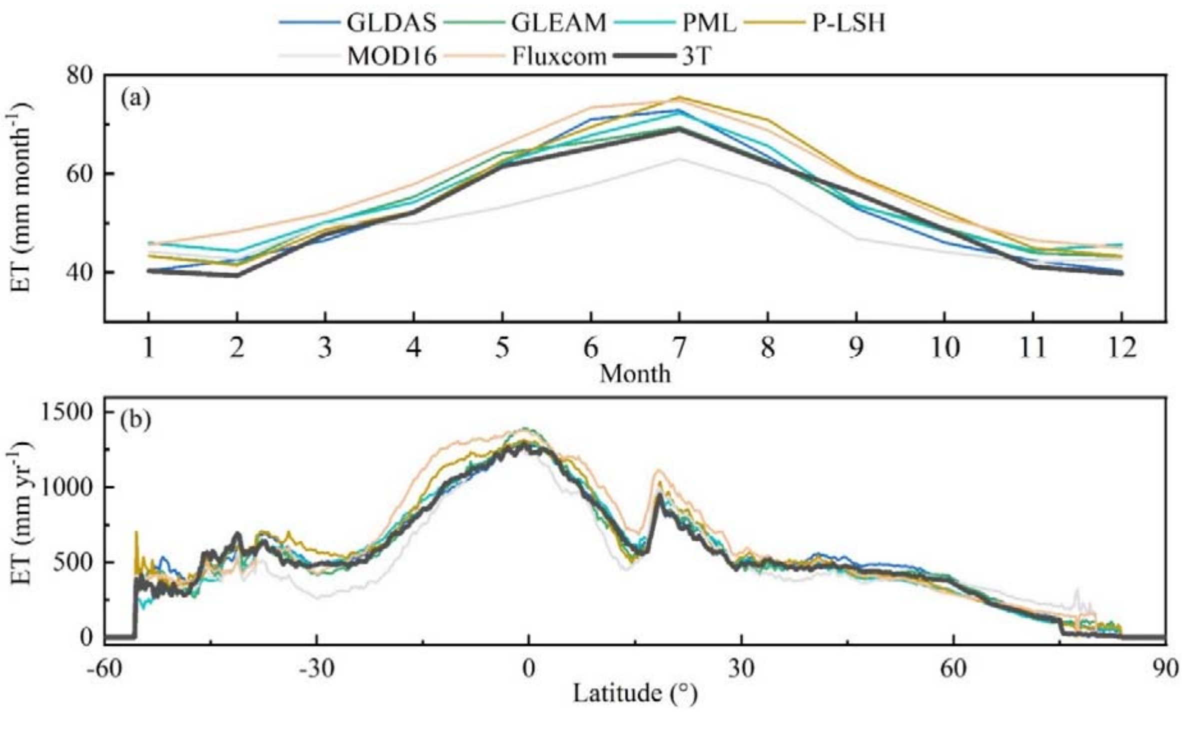

Figure 7: (a) Monthly variation and (b) annual latitudinal distributions of the multi-year (2003–2013) mean ET value estimated with the 3T model (black line) and 6 ET products in vegetated areas (mainly excluding Greenland, Antarctica and desert areas, according to Jung et al. (2019)).

Community Reviewer 1

Comments:

This manuscript describes an actual evapotranspiration dataset with high spatial resolution, no parameter calibration, and good accuracy, which is particularly important in the context of increased extreme climate hazards. However, this paper needs to further verify the accuracy of its product, which will be more appropriate.

1. Based on the Conclusion part, it seems the authors produced ET with a resolution of 3-hour. However, the current validations against the observed ET and ET from water balance are mostly at the monthly or annual scales. For this reason, it is not clear the accuracy of the 3-hour ET data. Therefore, it is suggested that the new product have validations against the observed ET at the 3-hour scale.
2. In addition, the authors claim that the new ET product has a good accuracy compared with other ET products. Yet, such a conclusion is also mostly based on the validations at the monthly or annual scales. It is appropriate to compare with other ET products at the 3-hour scale, such as GLDAS and ERA5.

Therefore, it will be more rigorous to add content for validations against observation data of 3-hour and comparison with other ET products at the 3-hour scale.

Responses:

Thank you for your interest in our work. In fact, validation of the 3T model was performed not only at a monthly scale but also at a daily scale in our previous manuscript. In particular, the discussion that the 3T model-based ET product could accurately capture the low ET values under extreme conditions in section 4.2 used daily ET estimates.

Your comments are similar to those of the second reviewer. We further tested the performance of the 3T model at the daily scale with all EC observations (because the results in Section 4.2 only contain extreme conditions) as well as at the 3-hour temporal scale. The results indicate that the 3T model is robust at different temporal scales. Please see our reply to reviewer 2 for details.