

## Reply for Anonymous Referee #2:

Thank you for the positive appreciation of our work. We will make the following changes according to your good suggestions:

### General comments:

**1. The study presents an approach of estimating long-term time series of daily ET in China by using the SEBAL model. In the current form, the manuscript lacks the literature to justify the need for the current study and several critical information related to SEBAL processing and ET validation. For example, there is almost no study reported in the introduction section that was conducted in China. There are several studies that used SEBAL and other surface energy balance (SEB) based models to estimate ET at a field and regional scales across different land covers and climates in China.**

**Response:** We have supplemented several literatures about performance of SEB-based model in China, and analyzed the need for this study (Section 1, Introduction, Page 3, Lines 75 - 91 in new version). Thank you for your help and good suggestions in improving our manuscript.

**2. Also, the authors did not report the critical information in the methods section such as the selection of hot and cold pixels for the SEBAL processing. This is one of the main steps for the SEBAL model processing and the results may vary based on the different approaches applied (manual selection or automated selection).**

**Response:** In the process of SEBAL ET generating, the hot and cold pixels were selected automatically by following the certain rules which referred to previous studies, and these parts were supplemented (Appendix, Pages 27 - 28, Lines 580 - 588 in new version). Thank you for your help and good suggestions in improving our manuscript.

**3. In addition, for the pixel-scale validation, the authors missed to report the quality of flux tower data and any approaches (e.g. constant Bowen-ratio, residual LE closure...) applied to close the energy balance. These details are very basics and the core for any study related to SEB-based ET estimations. Without this critical information, it's difficult to warrant the validity of ET estimated from the current study.**

**Response:** Regarding the quality of flux tower data, we have filtered and corrected the flux tower measured data, first, we selected the high-quality data with Energy Balance Closure Ratio (ECR) is more than 80%, and further used Bowen Ratio energy balance method to correct the selected data, these parts were supplemented (Section 2.3.1, Page

6, Lines 144 - 154 in new version). Thank you for your help and good suggestions in improving our manuscript.

## **Specific comments:**

### **1. Section 2.2: lengthy model description...move it to appendix**

**Response:** These parts were moved to appendix (Pages 25 - 30, Lines 530 - 625 in new version). Thank you for your help and good suggestions in improving our manuscript.

### **2. Line 190-195: explain the gap-filling (spatial and temporal) process for pixels impacted with cloud**

**Response:** Regarding the MODIS data used for SEBAL input (MOD11, MOD13 and MCD43), we have filled the missed or unreliable (caused by cloud or other reasons) pixels and the methods were referred to previous studies (Section 2.2, Page 3, Lines 125 - 130 in new version). Regarding the MOD16 ET data, the missed or unreliable pixels were not used for the comparison with SEBAL ET and not filled. Thank you for your help and good suggestions in improving our manuscript.

### **3. Line 196: any modification applied to MOD11 band for Ts adjustment?**

**Response:** In this study, the daytime surface temperature in MOD11 was used for SEBAL input, the data have not been modified except dimensional conversion (scale factor is 0.02) and gap-filling (Section 2.2, Page 5, Lines 125 - 130 in new version). Thank you for your help and good suggestions in improving our manuscript.

### **4. Fig 3: could be moved to appendix.**

**Response:** This figure was moved to appendix. Thank you for your help and good suggestions in improving our manuscript. (Appendix, Page 30 in new version)

### **5. Line 223: describe the quality of flux tower data and any filtering applied to remove bad observations**

**Response:** In this study, we selected the high-quality data with Energy Balance Closure Ratio (ECR) is more than 80%, and further used Bowen Ratio energy balance method to correct the selected data, these parts were supplemented (Section 2.3.1, Page 6, Lines 144 - 154 in new version). Thank you for your help and suggestions in improving our manuscript.

### **6. Section 2.4.1: validation with flux tower and water balance would suffice.**

**Response:** The ET obtained from flux tower and water balance method could efficiently validate SEBAL ET.

Moreover, the MOD16 ET product is one of widely used evapotranspiration dataset for water resources management and global change study, which also performs accurate to some extent. In this study, the comparison of SEBAL ET and MOD16 ET was conducted to judge if the further improvement was found in SEBAL ET (Section 2.4.1, Page 8, Lines 187 - 189 in new version). Thank you for your help and good suggestions in improving our manuscript.

**7. Fig 4: add the time series plots as well...provide more information for monthly/seasonal/annual variations**

**Response:** The time series plots of ET in the flux tower stations were added as Fig .5, and we further described the ET variation characteristics in time series (Section 3.1, Pages 9 - 10, Lines 226 - 232 in new version). Thank you for your help and good suggestions in improving our manuscript.

**8. Fig 5: any obvious reason for ET underestimation for higher ET rates from SEBAL (for all land covers)?**

**Response:** The reason of underestimation of SEBAL ET at higher ET was discussed in addition (Section 4.2.1, Page 21, Lines 386 - 400 in new version), which may cause by the saturation issue of optical sensor. For example, in the dense vegetation covers, the vegetation index (e.g., NDVI) was likely underestimated and can not accurately characterize vegetation status, therefore, soil heat flux will be overestimated according to Eq. 9 (in appendix), and may further caused the sensible heat flux underestimation. Thank you for your help and good suggestions in improving our manuscript.

**9. Fig 8: discuss the seasonal overestimation/underestimation from SEBAL...what are the primary driving factors?**

**Response:** We discussed the reason of the seasonal overestimation/underestimation from SEBAL in addition (Section 4.2.1, Page 21, Lines 386 - 400 in new version). For example, the obvious overestimation in spring and summer may cause by gap-filling of unreliable pixels, spring and summer have the relatively frequent precipitation, which causes more unreliable pixels due to the cloud, and these pixels value were finally replaced by gap-filling of nearest date pixel value, therefore, the modeled ET value of these pixels was close to that of nearest date without precipitation. Actually, due to the high air humidity in rainy day, the evaporation and transpiration are relatively less than that of nearest date (Ferreira and Cunha, 2020; Li et al., 2016). Moreover, it should be noted, due to the decrease of surface temperature after precipitation, the ET (both actual and modeled value) is also in a relatively low level (Cheng et al., 2020). This may explain the reason of obvious overestimation at lower ET rates in spring, summer and other pixels affected by cloud. For underestimation of SEBAL ET at higher ET rates, which may cause by saturation issue of optical sensor. Thank you for your help and good suggestions in improving our manuscript.

**10. Section 4.2.1: this section doesn't explain the quality of input data for the current study. The QA/QC of input data is fundamental for ET modeling but this information is missing.**

**Response:** In this study, the MODIS quality control (QC) file were used to distinguish the unreliable pixels of MODIS data (MOD11, MOD13 and MOD43) and then the gap-filling method were applied for fill or replace these unreliable pixels (Section 2.2, Page 5, Lines 125 - 130 in new version). Thank you for your help and good suggestions in improving our manuscript.

**11. The reference cited in line 425 is related to GPP....not relevant to conclude that the quality of GMAO data was not accurate enough for ET modeling.**

**Response:** These parts have been removed (Section 4.2.1, Page 20, Lines 425 - 426 in old version). Thank you for your help and good suggestions in improving our manuscript.

**12. Section 4.2.2: this section is not discussing about the quality of flux tower data included in the current study...mostly literature.....not helpful to link with the results reported**

**Response:** We have rewritten this section and further discussed the errors may cause by flux tower in this study (Section 4.2.2, Pages 21 - 22, Lines 405 - 426 in new version). Thank you for your help and good suggestions in improving our manuscript.

**13. Line 432: report the error fr om the flux towers considered in this study**

**Response:** We have rewritten Section 4.2.2. In this study, the Bowen ratio method (Eq. 3), which assuming that the residual of the energy balance is attributed to sensible and latent heat flux and assigning the missing energy flux to them, was used to enforce energy closure. Actually, this assumption is not very correct, which generally led the sensible and latent heat flux overestimation, which may could explain that the SEBAL ET was generally underestimated when compared to flux tower observed ET (Fig. 9). The same issue was found in regional-scale validation, due to the ignoring of  $\Delta S$  in the water balance computation process (although it's really small), which could lead the regional ET overestimation and further caused SEBAL ET underestimation in validation (Fig. 10) (Section 4.2.2, Page 21, Lines 405 - 414 in new version). Moreover, the 1 km  $\times$  1 km area of pixel was used for matching the footprint of flux tower which was referred to the study of Velpuri et al. (2013), however, the footprint is not stable but varied with environment changed, e.g., vegetation height. Chen et al. (2012) reported that forest footprint has clear difference with grassland, the footprint of forest is much larger which is reached kilometer-scale.

In fact, forest footprint may more matching with the spatial resolution in this study. Therefore, it may explain that the SEBAL ET has the greatest performance in forest but worst performance in grassland. Compared to the study of Velpuri et al. (2013), the grassland also showed the worst remote sensing ET estimation in US when using flux tower data for validation at a kilometer-scale (Section 4.2.2, Page 22, Lines 415 - 425 in new version). Thank you for your help and good suggestions in improving our manuscript.

**14. Line 443: report the footprint of flux towers used in this study...this is critical for pointscale validation**

**Response:** In this study, the 1 km × 1 km of pixel was matched with flux footprint (Section 2.3.1, Page 6, Lines 152 - 154 in new version). And we further discussed the errors may cause by the footprint issue (Section 4.2.2, Page 22, Lines 415 - 425 in new version). Thank you for your help and good suggestions in improving our manuscript.

**15. Line 445: any explanation about overestimation during winter? Also, discuss the SEBAL overestimation at lower ET rates and underestimation at higher ET rates in Figure 7 and Figure 8**

**Response:** The contents in Section 3.2.4 showed that SEBAL ET has the highest error in winter (rRMSE = 66.92%), but the error did not show obvious underestimation or overestimation (MBE= -0.62 mm/8d). We further discussed the reason of the highest error in winter, which may cause by the low temperature and snow cover in winter (Section 4.2.1, Page 21, Lines 393 - 395 in new version). Moreover, we also discussed the possible reason that the SEBAL overestimation at lower ET rates and underestimation at higher ET rates (Section 4.2.1, Page 21, Lines 395 - 399 in new version). Thank you for your help and good suggestions in improving our manuscript.

**16. Section 4.3.2: not relevant to discuss the results from SEBAL, could be removed.**

**Response:** These parts have been removed (Section 4.3.2, Page 21, Lines 460 - 479 in old version). Thank you for your help and good suggestions in improving our manuscript.

**17. Line 490-495: report the spatial (tiles/basins) and temporal (study years) variation of hot and cold pixels....would be helpful to link with the reported results**

**Response:** The hot and cold pixels selection error is one of the main causes of SEBAL model uncertainties. We further discussed the influence of temporal variation of hot and cold pixels to SEBAL. A study reported that the cold pixel performed more stable than hot pixel in time series, especially in winter, the hot pixel was highly varied may due to the similarity of NDVI over space, it could further explain the poor performance of SEBAL ET in winter (Section 4.3.2, Page 23, Lines 451 - 458 in new version). Thank you for your help and good suggestions in improving

our manuscript.

**18. Line 490-499: the sources of errors related to H estimation can be evaluated with instantaneous H from flux tower. this would help to identify where the errors are coming from (maybe from modeled Rn and G too)....along with the quality of input data and flux tower data.**

**Response:** The source of ET estimation errors is a subject worthy of further study. In this paper, we referred to previous studies to further discussed this question: *'Besides sensible heat flux, the errors of SEBAL ET may derived from net radiation or soil heat flux as well (Li et al., 2017; Teixeira et al., 2009). For net radiation, which is computed using surface albedo and Stephen Boltzmann law (Eq. 2 in appendix), generally performed a relatively agreement with flux tower observed value, while soil heat flux, which computed using empirical formula related to net radiation and NDVI (Eq. 9 in appendix), has a poor performance (Li et al., 2017; Song et al., 2016). In the study of Li et al. (2017), soil heat flux estimation showed a clear overestimation in higher ET area, e.g., wetland, which may further cause the sensible and latent heat flux underestimation in higher ET rates. In the most SEB-based algorithms, the similar net radiation and soil heat flux estimation methods are used, and various sensible heat flux estimation methods are the main sources of the difference among the various SEB-based algorithms. However, the causes of the net radiation and soil heat flux estimation errors have not been clearly discussed, e.g., the effect of satellite transmitted time or land cover types. These issues could be the focus of our follow-up research, for example, high frequency geostationary satellite and flux tower observations may be helpful for this research.'* (Section 4.3.2, Page 23, Lines 462 - 471 in new version). However, due to the limited of flux tower data, we could not study the instantaneous energy component in this paper, and this object will be conducted in our follow-up research. Thank you for your help and good suggestions in improving our manuscript.