Thank you very much for your comments. Please find below our responses to each of your comments.

"It is not clear why a dryer area (less than 320 mm rainfall) was chosen for the validation. Especially, the title specifically talks about the east African highlands which are generally wetter than the validation site. Also, not clear what each of the 16 or so data points represent in time. Are they daily averages of some period? If data was collected on a daily time step from July to December, not clear why not use most of the data. How about other years?"

- Flux sites are sparse in Africa, especially in the East African highlands. Among the CarboAfrica flux network, we only found the SD-Dem flux site in the East African highlands. So our site selection was dictated by data availability.
- We have added new quantitative analysis that reveals the error contributions from various error sources (net radiation, EF, and assumption of constant EF during the daytime). We believe that these results are also indicative of the results in the higher elevation areas. For example, our results show that cloudy sky conditions lead to underestimation of net radiation, and we expect this to hold true in higher-elevation areas as well. So the diagnostic nature of our study provides lessons for satellite-based ET estimates in the whole East African highlands.
- We have cases where we have actual ET values but missing corresponding reference ET values. In the original manuscript, we removed these cases from the evaluation of the actual ET values to maintain the same sample size between the evaluation of actual ET and evaluation of reference ET. We agree with you that we need to include all available data. In the revised manuscript, we have included all available data for evaluation of actual ET. We have modified Figure 5 (and related results and discussion) accordingly.
- The data points shown are daily values. We have added a new Figure (Fig. 4) that shows the time series of daily ET estimates and observations.
- Data period was limited by data availability.

"If flux tower is not available in the highlands, the use of a water balance approach should be attempted so one can evaluate the performance of the model in more complex regions."

- The focus of our study is to evaluate satellite-based ET estimates through **direct comparison** with ground-based measurements. Despite the short period of the data, our diagnostic analysis reveals several interesting features on errors (and sources of errors) in satellite-based ET estimates.
- The water balance approach is an indirect and a completely different approach, which is beyond the scope of our study. Our opinion on this issue is that it is difficult to quantify the errors in the satellite-based ET estimates based on the water balance approach, because this approach involves errors of various sources (e.g., precipitation, soil moisture, hydrologic model) and it is difficult to disentangle the various error sources.

"Figure 3: Also, not clear why the actual ET is only posted on a scale of 0 to 4 mm while the potential ET is up to 7 mm. One would think some well-watered areas would also have ET at a potential rate such as the Sudd; it could even be more depending on the reference crop type. An explanation is requested on the absence of pixels that have as high ET rate as the ETo. Is the reference ET standardized to some crop, clipped grass or alfalfa?"

• Yes, you are right. There may be three reasons. One is what you gave, the absence of pixels that have as high ET rate as the ET₀. The second is the underestimation of daily actual ET, which has been well addressed in the discussion section of this paper. The third is that pixels

in dry regions usually have large ET_0 values, but low accrual ET values; pixels in wet regions may have as high ET rates as ET_0 , but ET_0 values usually less than those in dry regions. As shown in dryness index maps in Fig. 3, there are many pixels whose values are less than 20%.

• We have added the definition of the reference ET in the first paragraph of Section 3.2 as follows: "The reference ET is standardized to ideal grass, a hypothetical surface of 8 to 15 cm tall green grass cover of uniform height, actively growing, completely shading the ground and not short of water (de Bruin and Lablans, 1998)."

"It would be great to see monthly and seasonal total ET for the region. With all the problem of cloud, it is hard to read the spatial pattern on a daily time step. Is there a plan to fill cloudy pixels, especially if we can assume 8-day ET fractions are stable over a certain period?"

We do not have currently an operational method to fill in the cloudy cases. Although the assumption that 8-day EF is stable during a period of 8 days can significantly reduce the cloudy pixels in EF images, it is hard to find net radiation values for the cloudy pixels because the net radiation values themselves are affected by the cloudiness conditions.

"The authors should also refer to similar indices by other researchers such as the evaporative stress index (ESI) from ALEXI model and also the water requirement satisfaction index (WRSI, originally developed by FAO)." Done

"There are also other similar models that combine MODIS-based ET fractions with weather data sets, implemented for drought monitoring in east Africa. Although not the scope of this paper, a comparison between their approach and one based on reference ET generated from data assimilation models such as GDSAS should be encouraged." This is a good suggestion, but out of the scope of this study as you pointed out.

"I also wonder if assuming a constant lapse rate is valid. Did they check if this rate behaves similarly from month to month? Other researchers have found that it varies from time to time and so a temporal varying lapse rate may be more appropriate."

For our site, a constant lapse rate (i.e. 0.67 °C/100 m) has been used. In the dry season this might be 1 °C/100 m. The error is less than 5% when taking 1 °C/100m instead of 0.67 °C/100 m. This is justified only in the dry season, but then during the dry season, actual ET is small and ET₀ is not really used.

"Page 6290, last line: are the 0.1 and 0.4 applicable for the region? Agricultural bare soil vs desert bare soil"

The values of 0.1 and 0.4 are used in the Sim-ReSET model. We have tested them in an agricultural region in China (Sun et al., 2009), but not in this study because of lack of appropriate data at the SD-Dem site. In some prior studies, soil heat flux is ignored in daily ET estimation (e.g., Price, 1982). We may assume that uncertainties in soil heat flux estimation have minor effect on daily ET estimation.

"Why use 16-day NDVI and 8-day LST. Why not use the 8-day NDVI? Or the NDVI is not that sensitive."

MODIS land science team only provides 16-day NDVI data product. This 16-day NDVI represents a 16-day composite NDVI, so one 16-day NDVI image can be used to generate two ET images along with two 8-day LST images.

"Not clear on the statement that the Dryness Index is operational. Is it running operationally and the results posted on the web?"

We have clarified the statement as follows: "...the new dryness index can be used for drought monitoring."

Reference:

Price, J.C.: On the use of satellite data to infer surface fluxes at meteorological scales. Journal of Applied Meteorology, 21, 1111-1122, 1982.