Dear Dr. De Kauwe,

Please find below our responses to Referee comments for bg-2019-85. We agree that more references to insightful modeling papers improve our manuscript, but we also want to maintain focus on the challenges of measuring ecosystem-scale E and T so that models have adequate observations for validation and critique. We added a number of references suggested by Reviewer #3 to detail the important findings of recent modeling approaches but decided that a comprehensive analysis of modeling results would distract from the present analysis. We also prepared a response to your comments on the previous version of the manuscript and realized that this may not have been uploaded into the Copernicus system. Our original letter is pasted below for completeness.

Sincerely,

Tarek El Madany

Reviewer #3
Review of Stoy et al.

This manuscript is a review of transpiration partitioning approaches. I came in late(r) in the review process and was not able to provide a full review. In my brief review I note that modeling studies are almost completely absent from the discussion. There are several modeling approaches that have worked to understand discrepancies in transpiration partitioning and the authors do not review them. Figure 1 is a start, but I find that plots like this are not very helpful (think of the original PILPS studies where plots of an OM difference were not helpful, but groupings or tables of LSM process were). Something like a replicate of the table in Schlesinger and Jasechko, 2014 is a start, but perhaps this could be updated to go beyond simply changes in canopy fraction S 2.1), which is a bit of a naive approach in my opinion and misses actual processes in modeling transpiration. I think this manuscript would be improved with an additional section discussion the advances and modeling of T/ET in a process based manner, with a table similar to Table 1. Some suggested references to get the authors started are included below.

We to focus mostly on observational studies of transpiration and evaporation and current and emerging measurements for partitioning them. Without adequate observations, models cannot be adequately validated. That being said, the Reviewer makes an interesting point and some of the references that are noted below point to the idea that inadequate lateral flow representation in current LSMs are likely a major reason for model/measurement mismatches. This aligns with the comments of Dr. De Kauwe on a previous version of the manuscript, and we note that we may not have submitted a response letter to these comments. We have attached the original letter below.

Figure 1 is but one way of noting some of the important discrepancies that arise between ET models and measurements. It focuses on seasonal patterns, and the paper by Matheny et al. (2014) focuses on diurnal patterns. To our knowledge, no study has yet performed a comprehensive analysis of model performance across multiple time scales using the NACP Site-Level Interim Synthesis. Such an analysis would add unnecessary length to the present paper, but we agree that it would be interesting. We also feel that a comprehensive discussion of the treatment of T and E in models would make for a compelling stand-alone manuscript, but would distract from the present study, which includes 278 references before adding the interesting references that were suggested below. We thank the Referee for their comments and provide a brief discussion of all of the suggested references, including justification for not including them in two instances. Most references are now included in section 2.1.

References
Shrestha et al Effects of horizontal grid resolution on evapotranspiration partitioning using TerrSysMP JoH 2017

This study finds that larger model grid sizes result in more evaporation, suggesting that sub-grid processes including hydrologic redistribution are critical for simulating T/ET dynamics.

Chang et al Why Do Large-Scale Land Surface Models Produce a Low Ratio of Transpiration to Evapotranspiration? JGR: Atmospheres 2018

This interesting manuscript was published after we began work on the present manuscript and thank you for pointing it out to us. The authors find that realistic lateral flow simulation creates situations where soil evaporation is suppressed in favor of transpiration from deeper water sources. We now cite this reference as well as the Ji et al and Fang et al. studies in the modeling section that discusses the discrepancy between models and measurements.

Maxwell and Condon Connections between groundwater flow and transpiration partitioning Science 2016

Maxwell and Condon find that lateral flow and groundwater dynamics are critical for simulating transpiration and we now cite this important paper.

Clark et al The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism HESS 2017

This manuscript references evapotranspiration partitioning once in reference to the Maxwell and Condon (2016) paper that we now cite.

Fatichi and Pappas Constrained variability of modeled T:ET ratio across biomes GRL 2017

We cited this interesting study extensively in the manuscript.

Rogers et al A roadmap for improving the representation of photosynthesis in Earth system models New Phytologist 2017

This manuscript makes a number of key recommendations regarding photosynthesis modeling, one of which is to include the sensitivity of stomatal conductance to vapor pressure deficit. We added this reference to section 2.2 that deals with VPD sensitivity.

Ji et al Do Lateral Flows Matter for the Hyperresolution Land Surface Modeling? GRL 2017

This paper emphasizes the importance of lateral flow simulation to transpiration and evaporation partitioning, especially during dry conditions.

Fang et al Influence of landscape heterogeneity on water available to tropical forests in an Amazonian catchment and implications for modeling drought response JGR Atmospheres 2017

This manuscript studies drought in the Amazon and finds an important role of the wilting parameter in the ACME land model.

Han et al Hydroclimatic response of evapotranspiration partitioning to prolonged droughts in semiarid grassland JOH 2018

This manuscript emphasizes important differences in T/ET estimates that result in response to drought in grassland ecosystems. It follows the underlying water use efficiency (uWUE) approach that we describe in detail in section 3.1. We added it to this section.

Referee #4

This manuscript is a comprehensive and very useful review of methods for partitioning evaporation and transpiration in terrestrial ecosystems. It is well written and organized, and generally does a good job of explaining why these techniques are important, what their strengths and weaknesses are, and how their application might be improved in the future. Overall I thought it was an excellent review and will be a very useful addition to the literature.

Thank you for your support of the manuscript and for the insightful comments which improved it.

I have a few minor comments:

Page 4, line 9: I think this should read "conductance related TO soil evaporation"

This is correct, we edited the text.

Page 10, line 18: It's not clear what specifically is unprecedented here. Are the scales of current SIF measurements unprecedented with respect to previous SIF measurements? Or is SIF unprecedented with respect to other measurement techniques in its potential for high spatial and temporal resolution measurements?

We agree with you that 'unprecedented' is a qualitative superlative that should be avoided. We felt that 'multiple' was an accurate descriptor of a measurement of fluourescence, which can theoretically be measured at the scale of a single photon.

Page 11, line 10: Typos in "canopy scaling" and "they aerodynamic conductance"

We corrected the typo in 'canopy' and removed 'they'.

Page 15, line 10-20: I would make sure it is clear that the regression is not supposed to be fitting the blue dots in Figure 4, and maybe also emphasize it in the caption to the figure. At first glance it looks like the lines are very poor fits to the cloud of dots, and it took me a minute to realize that this was the wrong interpretation of the figure. Perhaps dots below the 95% level could be plotted in a lighter color to help show that the regression is not supposed to be fitting the whole cloud?

This is a good point and something that we struggled with a bit when creating the figure. The regression is meant to be a boundary line fit. We changed the color of the eddy covariance measurements to gray and now explicitly describe these in the figure legend. We also edited the legend for clarity.

Page 17, line 6-7: This sentence refers to local overpass times, but never specifies what location this is for. Overpass times would be different depending on latitude, so this must be for a specific place or a particular latitude.

This is correct; it is not possible to take a snapshot of an entire time zone at once. We removed the reference to the specific time and now just note 'morning' and 'afternoon'.

Section 4.2: This section is based on a new application of two versions of the flux-variance partitioning approach to a flux site. The technique itself is well documented in the referenced papers in this section, but there are some analysis choices and preprocessing steps associated with applying techniques like this to a new site. It might be a good idea to include a more detailed description of how the technique was applied to the site in an appendix or supplement. Alternately, making the analysis code for this section available in a public repository would allow this section to be evaluated more thoroughly and/or replicated by interested readers.

The latter analysis is based on the FluxPart algorithm available at https://pypi.org/project/fluxpart/. The innovation is the addition of a routine for closed path infrared gas analyzers, which will be released soon. In our opinion, the most critical factor when specifying how the algorithms were employed for a proof-of-concept test is the treatment of water use efficiency. Water use efficiency was estimated by the algorithm in both instances, rather than specified. We have adjusted the text accordingly.

Comments to the Author:
Dear Authors,

the reviews of your manuscript were extremely positive, I have decided that minor revisions are necessary before the manuscript can be published. I agree with the reviewers that this manuscript is very readable and will likely be well received by the community. In revising your manuscript to address the various issues highlighted by the reviewers can I ask you to consider a few points from me as well.

Thank you for your support of the manuscript and for your insightful comments, which we address below.

1. The algorithm descriptions covered in the appendix are extremely useful for the community. I note that it is referred to once in the introduction but it would be great if you could find at least one more place in the manuscript to refer to the appendix (perhaps the discussion?). I think what you've done here is very valuable and it would be a shame for a reader to miss this.

Thank you for this suggestion; we did put quite a bit of work into the Appendix but felt that it added too much length to the text. We refer to the Appendix in the Introduction, and in section 3.1 and 3.3. We added another reference in section 3 and in the Conclusions section.

2. In 2.1 where you talk about the CMIP models having a T/ET ratio of 0.22-0.58 I think it would be valuable to offer some insight into why: (a) they disagree with each other; and (b) why this ratio is noticeably below other data-based estimates. This is an optional suggestion but I do think given this is a review it would be good to inform the reader. Perhaps they might consider citing Berg and Sheffield - Evapotranspiration Partitioning in CMIP5 Models: Uncertainties and Future Projections. Similarly, you might wish to more explicitly raise the issue of discrepancies amongst how models simulate LAI and the impact this has on the water cycle / ET partitioning.

We would also like to know more about the reasons for the discrepancy but Wei et al. (2017) only note that the reason is due to methodological differences without discussing in detail why, perhaps due to the short format of Geophysical Research Letters. Additional explanations are also not available in the supplement of Wei et al. (2017). This strikes us as an important avenue of future research.

Note that we now include a brief description of modeling results in response to Referee #3.

3. Again feel free to ignore this, but this is one of the few papers I've seen raise this issue. The authors neatly raise the issue of interception. In our 2013 GCB paper on WUE (Forest water use and water use efficiency at elevated CO2: a model-data intercomparison at two contrasting temperate forest FACE sites), we found that the proportion of intercepted water varied among the models by between 2-14%. This was considerably below the field estimates for the sites (and the range you quote in 3.6). It was striking how data free the assumptions were than underpinned how interception is treated in models. I'm not suggesting you get into how models simulate interception, I just think you might consider highlighting this is a serious problem for models and

may contribute to erroneous partitioning ratios - see above.

Intercepted water is very difficult to measure and we were fortunate to have an expert (Dr. Shuguang Liu) help with a subsection on it. We added the findings of De Kauwe et al. 2013 to further emphasize its importance for models.

3. The paper didn't seem to make much of soil evaporation? I realise it is a minor component of total ET, but recently we noted how poorly this was simulated by models. In a water-limited, semi-arid ecosystem, some models thought soil evaporation was around 50-130 mm yr-1, whilst other models thought it was 2-3.5 times greater (Challenging terrestrial biosphere models with data from the long-term multifactor Prairie Heating and CO2 Enrichment experiment). I only raise this example because it suggests to me that this isn't a trivial process to model (otherwise there wouldn't be this disagreement). I was expecting to see some sub-section on soil evaporation, but this may simply be my personal bias on this issue, so ignore as you wish.

Following referee comments, we now added a section on soil evaporation following the new manuscript by Or and Lehman (2019). The reason for our very brief discussion of soil evaporation before is that other reviews have covered it. Now, with new analytical techniques for estimating it, we agree with the reviewer and added a subsection to the manuscript and cite De Kauwe et al. 2017.

4. In table 2 where the variability in the exponential term is shown across models, I feel whilst interesting - without some context or explanation, it is a bit limited in terms of insight. Could the authors group the models by their stomatal assumptions (Ball-Berry, Leuning, etc). Does this help explain why they vary? Why is the BEPS model most similar to the observations?

We are not entirely sure why BEPS is most similar to observations but find that it is interesting that it does. We did not want to pursue a long intercomparison of models versus measurements in the present manuscript, and this comment combined with the comment above made us realize that a stand-alone multi-model intercomparison of CMIP5 and other models with respect to evaporation and transpiration partitioning would be forthcoming.

5. Was there a reason the ECOSTRESS mission wasn't mentioned in section 5?

We now mention ECOSTRESS explicitly in section 5.

6. When the authors discuss partitioning via the use of GPP in the WUE approaches, it would be worth mentioning that GPP isn't strictly an observation (so any errors in GPP will propagate here).

Thank you for pointing this out, we added a passage about GPP uncertainty to section 3.

Best wishes,

Martin