

Biogeosciences Discuss., referee comment RC1
<https://doi.org/10.5194/bg-2022-32-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on bg-2022-32

Anonymous Referee #1

Referee comment on "Intercomparison of methods to estimate gross primary production based on CO₂ and COS flux measurements" by Kukka-Maaria Kohonen et al., Biogeosciences Discuss., <https://doi.org/10.5194/bg-2022-32-RC1>, 2022

In the manuscript "Intercomparison of methods to estimate GPP based on CO₂ and COS flux measurements", the authors compare 4 different models for calculating GPP on ecosystem level. Two of them are based on CO₂ and environmental measurements, whereas the other 2 include a dependence on carbonyl sulfide fluxes. The GPP, based on a neural network, agreed very well with classic flux partitioning. The GPP based on the LRU of chamber measurements from the top of the canopy also agreed well with the classic approach, but tended to overestimate GPP during periods of high incoming photosynthetic radiation. The second COS based approach, using a stomatal optimization model, agreed much better with classic flux partitioning and, although its implementation to other field sites might be promising, still needs to be tested.

I generally agree, that this manuscript deserves to be published, but I have some questions and suggestions to improve the document.

General comments:

- I suggest using an ANOVA and post-hoc tests to compare the results of 4 different models, daytimes and timescales instead of doing t-tests between only 2 of them. This could also end up in a nice table/plot for the reader. It's sometimes hard to grasp the differences within the text, which model results in a higher/lower GPP at different timescales and daytimes.
- (Also, if a pairwise t-test is used to compare so many samples, the p-values need to be

adjusted - see Bonferroni Holm).

- I also think the publication would profit if you put plots showing the modeled versus the measured daytime NEE (for all approaches) for interested readers into the supplement (to compare over/underestimation of the models).
- Where do the differences between the daily and monthly GPPs averages come from?
- I suggest having an English native speaker proofread the manuscript since some sentences feel off.

Specific comments:

38-40 The daytime approach of Lasslop et al. does not assume, that the respiratory processes are the same during day and nighttime, at least the base respiration is based on daytime data!

60 The carbonic anhydrase is also located within the cytoplasm. I would add this information. (see: Polishchuk, O. V. (2021). "Stress-Related Changes in the Expression and Activity of Plant Carbonic Anhydrases.")

99 Was the friction velocity threshold applied during both day and nighttime?

105 For the sake of completeness, it would also be nice to have the company/origin of pt100 sensor stated here.

112 Why did you use 50% as a threshold?

168 I think the 4th method should get a separate section including a title like the other 3.

186 Why did you use these exact values? Is there a reference for them? In which way will this influence the resulting LRU during the season regarding over and underestimation. It would be nice to see a sentence or two about this.

193 I am not sure the data is presented in a form, that will help the reader understand the data. Maybe it would be better to sort the sentences chronologically from 2013 to 2017 or group them by the environmental variables following Fig 1.

202 What does slightly higher mean? Can you give a percentage or absolute values?

209 Was the half-hourly data also statistically different?

210 How can GPP_{nlr} be negative? Shouldn't the equation make it positive in any case, and set it to 0 when the air temperature is below 0?

215-216 The GPP difference in 4d is not for daily but for 30 min data, "e" shows daily values.

232 LRU_{cap} might be higher than LRU from chamber measurements. The LRU chamber measurements might not be best in representing the whole canopy. The higher LRU of GPP_{cap} might even be closer to the true LRU value of the canopy, depending on the position of the chamber measurements. A higher LRU indicates more COS uptake per CO₂ uptake, which might happen in the lower part of the canopy since there is less PAR, but the COS uptake should continue unhindered. I would refrain from concluding that LRU_{cap} was overestimated during times of high radiation, but discuss the difference (and possible reasons) between the two COS based GPP.

234 I feel like introducing 2 new "parameters" in the result section is the wrong place, introduce them in the methods section.

239 Instead of writing "not large" can you tell if they are statistically different, and which one was higher/lower.

250-253 I don't think the comparison to a full season is needed, only state, that these cumulative measurements account for 13 weeks around the peak growing season.

274 How did you find the saturation point. Which algorithm did you use?

279 You could put a reference for Figure B2 here, showing the higher LRU at higher PAR for LRU_{cap}

283 Usually, an increase in VPD should decrease the stomatal conductance/GPP. (see page 480-481 Körner, C. (1995). *Leaf Diffusive Conductances in the Major Vegetation Types of the Globe. Ecophysiology of Photosynthesis. E.-D. Schulze and M. M. Caldwell. Berlin, Heidelberg, Springer Berlin Heidelberg: 463-490.*

and Lasslop, G., et al. (2010). "Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation." *Global Change Biology* **16**(1): 187-208.

The correlation between VPD and GPP in spring might only be caused by the correlation of the air temperature with VPD.

286-289 Since you have not observed a drought or heatwave, these sentences feel unnecessary.

303 State the difference here, "similar" feels unclear.

316 I actually dislike the term measured LRU, as LRU is a product of GPP and COS fluxes, so it can't be measured. I suggest replacing measured LRU with "LRU derived from chamber measurements" or a wording that is more representative for the LRUs calculation.

318 Why was it not comparable to the chamber measured LRU?

322 It would be nice to have the information about the position of the leaf chambers in the methods section, so that the reader knows what the basis for the LRU_{par} is.

336 I am not sure that you should conclude that the LRU_{cap} model underestimates GPP during midday compared to GPP_{par}. Due to the aforementioned issue of only having a chamber at the top of the canopy, the GPP_{par} might be overestimated and GPP_{cap} could actually be better. (The LRU at the top of the canopy might be lower compared to areas within the canopy). I propose, just writing that the GPP is lower instead of underestimated, since "underestimated" gives the impression that GPP_{par} is correct.

Fig 3 Is this figure based on half-hourly data points? Are these really average or median differences like in Fig 2?

Fig 5 Do you mean cumulative daily fluxes when you write daily flux data points? You mention, that all medians have been calculated using the same number of data points. Were these also the same data points, or could there be a bias from different days?

Fig 7 Why did you use 700 par as the threshold?

Fig B2 If measured means "chamber -measured" LRU please state so.

Technical corrections:

91 "Consisted of a Gill HS"

112 I feel like some words are missing in this sentence. The second part about monthly averages feels disconnected. Are you trying to say, that the monthly averages were also only calculated from daily means, when 50% of the half-hourly data was available?

218 To investigate further the causes for the ...

253 remove brackets from (on average 25%)

332 Do you mean noisy (scattered)?