

Reviewer: Konings et al. provide globally distributed estimates of heterotrophic respiration (R_h), both from satellite based observations and from a bottom-up scaling of an empirical model. The satellite based estimates in particular are very novel, and are obtained by combining atmospheric inverse estimates of net ecosystem production (NEP) with global photosynthesis (GPP) estimates informed by solar induced fluorescence data, multiple vegetation models, and empirically upscaled estimates. To combine NEP and GPP to get an estimate of R_h , the authors need to estimate global variations in carbon use efficiency (CUE). The use estimates provided by CARDAMOM, a simple empirical model of carbon fluxes with parameters constrained by global observations. The final estimate of R_h is then taken as [sic] $\delta^{13}C_{SE} = \delta^{13}C_{\sim} [\dots]$ The authors are to be commended for their heavily data-informed approach, which highlights the potential to use disparate observations to inform global estimates, not just test predictions. The results are interesting and the manuscript is very clearly written and no doubt will be of interest to the readers of *Biogeosciences*.

Response: We are glad the reviewer believes our new satellite-based approach to estimating R_h is very novel, and that the manuscript will be of interest to *Biogeosciences* readers.

Reviewer: There are several limitations to the approach, however, and the estimates of R_h should be taken as a first pass of a promising approach rather than a reliable and informative quantification of the global distribution of R_h . As the authors note, the R_h estimates should not be used as a benchmark for other estimates, as their global quantified uncertainty is 50% of the mean flux. There are simply too many uncertainties, some quantified in this manuscript, and some not. Great caution should also be taken in using the approach to quantify trends over time.

Response: We agree that the results of our approach remain uncertain, and tried to be clear about this. We are glad the reviewer also agrees that further work building on this first manuscript could help to improve the robustness of the method. In the revised version, we will add additional language to highlight the uncertainties of the approach, and its limitations for trend analysis and global benchmarking.

Reviewer: The uncertainties stem primarily from the fact that both GPP and CUE are not known, but must be estimated themselves. Global GPP estimates vary a lot between approaches, and although the authors use an approach that combines SIF with DGVMs and upscaled GPP estimates, the relationship between SIF and GPP is poorly understood, and even the magnitude and spatial distribution of GPP has considerable uncertainty. The cited paper on which the GPP estimates are based, Parazoo et al. (2014) does a good job of assessing some of those uncertainties, but important sources of bias persist. For example, Parazoo et al. (2014) used the empirically upscaled GPP from Jung et al. 2012 to constrain the magnitude of GPP to roughly 120 PgC, but

recent results of more updated empirical upscaling approaches from the FLUXCOM project (<https://www.bgc-jena.mpg.de/geodb/projects/Data.php>) show global GPP estimates vary from 108PgC (neural net based) to 125 PgC (Random forest based), each with an associated uncertainty of ± 8 PgC (standard deviation). It would be worth including an assessment of the contribution of this uncertainty to the global estimates reported here. The global estimates of CUE are also subject to large uncertainty, though the authors do a great job of assessing the impact on their results. In the absence of a global database of CUE and its seasonal variability, however, the uncertainty is difficult to quantify accurately. It would be worth highlighting in the abstract what these uncertainties indicate regard research needs to improve this approach.

Response: We agree that both the global CUE and GPP variability remain poorly understood, and that the Parazoo et al. (2014) approach is influenced by the quality of the Jung et al. (2012) estimates and TRENDY models. In the revised manuscript, we will include a sensitivity analysis that uses GPP estimated from FLUXCOM (the median across the three estimation methods) instead of the Parazoo et al. (2014) estimate used for the main calculations. We will also add text to the abstract regarding the dominant sources of uncertainties across the three fluxes and how our approach can be improved.

Reviewer: Detailed comments:

The annual totals for GPP and NEP should be given in the methods section to allow the reader to assess their relationship with the annual total Rh.

Response: Good point. We'll add these in the revised manuscript.

Reviewer: Page 2, line 20 Although the authors are correct that heterotrophic respiration is relatively unconstrained, the same cannot be said for ecosystem respiration, particularly at night. Eddy-covariance observations provide a direct observation of ecosystem respiration at night at 100's of sites around the world. Consider rewording.

Response: We will rephrase in the revised manuscript to mention eddy covariance sites as a widespread constraint on NEE (and nighttime respiration). However, we note that even the observational network of eddy covariance sites is not fully representative, with particularly strong biases in among others, regions of high topography and the Southern Hemisphere. For example, of the 225 sites used to constrain the FLUXCOM effort, only 17 are in the Southern Hemisphere.

Reviewer: Page 3, line 20: 'is calculated as'

Response: We'll fix this in the revised manuscript, thanks.

Reviewer: Page 6, line 16: 'To reduce error, all visual maps are presented after applying a 3 pixel by 3 pixel moving average smoother.' This does not reduce error, unless the error is randomly distributed around zero. Do you have evidence that there is no systematic bias in spatial distribution of the CMS-Flux predictions?

Response: In the revised version of the abstract, we will clarify that this moving average smoother was applied to reduce the random component of the error only. This can be accomplished simply by changing the quoted text to "to reduce the random component of the error, all visual maps are applied...". This 3x3 moving average window was also used by Liu et al, 2017.

Reference:

Liu et al. (2017): Detecting drought impact on terrestrial biosphere carbon fluxes over contiguous US with satellite observations. Biogeosciences Discussions, in review. <https://doi.org/10.5194/bg-2019-37>

Reviewer: Page 6, line 17: “The NEP is a small number that is the balance of many larger components, so small errors in NEP could lead to large compensating errors in Rh.” This is not clear. A small error in NEP should have little effect on the derived Rh, as NEP itself has a small role in the calculation especially relative to GPP, which is a very large number, and CUE.

Response: This was indeed unclear – we meant small errors in NEP in combination with errors in GPP. We’ll remove this in the revised manuscript, thanks for pointing it out.

Reviewer: Page 6, line 25: the uncertainty in total annual GPP from the Parazoo et al (2014) paper does not consider methodological uncertainty (see differences between methods in FLUXCOM). How would this affect the results presented here.

Response: A full assessment of the methodological uncertainty in any given GPP estimate is difficult without perfect knowledge of the GPP. For example, even the different methods in FLUXCOM do not capture all methodological uncertainty as they do not, for example, capture uncertainty related to possible missing input information (which may be part of the reason the uncertainty of the FLUXCOM products is actually lower than that of the Parazoo et al. (2014) estimates). Insofar as there is non-captured uncertainty in GPP, this would have a non-negligible effect on the Rh, as shown in Figure 5. We will further illustrate this in the new manuscript with a new sensitivity analysis that uses FLUXCOM GPP.

Reviewer: Page 12, line 24: Most plant traits can not be estimated from space, and it is difficult if not impossible to properly characterize the uncertainty associated with estimates of photosynthesis from space as there are no observations of ecosystem photosynthesis. The authors should show some restraint when trying to argue that estimates of photosynthesis, plant traits and Rh from space contain significantly lower sampling errors than bottom-up estimates. Also please clarify what you mean by sampling errors here and how sampling errors relate to total uncertainty.

Response: By sampling errors, we mean errors associated with the fact that in situ observations may not be in locations that are representative of environmental conditions across the globe – for example, because they are underrepresented in the tropics, or because they occur more frequently in disturbed areas than ecosystems as a whole, or because they undersample regions of high topography, etc. Because they sample across the globe, spaceborne remote sensing estimates would be expected to have significantly lower representativeness errors, although we agree with the reviewer that some errors may remain due to e.g. reductions in accuracy correlated with cloud cover. We also agree that remote sensing of photosynthesis (and net carbon fluxes) still has significant sources of error. We did not mean to imply in this section that any top-down estimate of an environmental variable will always have lower overall uncertainties. We will rephrase the text in the revised manuscript to clarify all of these issues, including clarifying our use of the term ‘sampling error’ and removing the phrase ‘plant traits’.

Reviewer: Page 14, line 10: measurements of SIF and estimates of GPP. GPP is not measured by TROPOMI.

Response: Thanks, we’ll fix this.