Interactive comment on ‚VODCA2GPP – A new global, long-term (1988-2020) GPP dataset from microwave remote sensing' by Benjamin Wild et al.

Anonymous Referee

Received and published: 11 August 2021

Reviewer comments
Reply to comments

General comments:

This study introduced a new global GPP dataset using satellite-derived vegetation optical depth (VOD) dataset spanning 1988-2020 in spatial resolution of 0.25o. The method for relating GPP to VOD was proposed in Teubner et al (2019, 2021). This study also presented validation against FLUXNET observation and other global GPP datasets (MODIS, FLUXCOM, TRENDY models). The difference of spatial and temporal GPP dynamics from other GPP reference datasets were shown. However, given the lower VOD-GPP model performance and extremely high global amount of VODCA2GPP, I do have concern about the quality of this dataset and its advantage over other datasets. The writing also needs a lot of improvements since there are both phrasing mistakes and technical mistakes.

Response: Dear Referee,

Thank you very much for reviewing our paper and giving very constructive feedback. We understand the concerns regarding the quality of the VODCA2GPP dataset but we strongly believe that although VODCA2GPP shows a slightly weaker performance than current state-of-the-art GPP products when comparing mean annual GPP, VODCA2GPP can still be a very valuable data source for studying the global productivity.

It is true that VODCA2GPP GPP estimates are relatively high compared to estimates from current state-of-the-art products. On the other hand, MODIS and FLUXCOM are known to substantially underestimate GPP especially in very productive regions (Fig 3 (new), Fig A1 (new); Turner et al. 2006; Wang et al., 2017) which is why large differences between MODIS/FLUXCOM and VODCA2GPP do not necessarily implicate substantial overestimation of VODCA2GPP. In fact, Welp at al. (2011) suggested that current estimates of GPP are generally too low and they estimated the actual yearly GPP to be between 150-175 Pg C yr^{-1}. Koffi et al. (2012) came to similar conclusions (146 ± 19 Pg C yr^{-1}). The comparison of mean annual GPP between VODCA2GPP and FLUXNET GPP does not suggest an overall overestimation of VODCA2GPP but exhibits lowest bias for VODCA2GPP among all analyzed datasets (Fig 3 (new)). Nevertheless, we agree that VODCA2GPP has a tendency for overestimation for arid regions (e.g., (woody) savannas and open shrubland; Fig A1 (new)). However, due to the missing consensus among existing datasets (Anav et al., 2015) we do not see this as an inherent flaw of VODCA2GPP but rather as a systematic problem in the observation of GPP originating from generally low in-situ data availability.

Regarding the style of writing we are very grateful for your detailed feedback and we will thoroughly revise the manuscript and improve the writing and other technical mistakes that are pointed out in your review.

Specific comments:

1 The global total amount of VODCA2GPP datasets is not given.

The global total GPP as derived from VODCA2GPP amounts on average to 200 ± 2.2 Pg C yr^{-1}. We will add this number in a revised version of the manuscript.

2 line 19, line31, it should be ‘at global scale’.

Thank you for this suggestion, we have revised this in the latest version.

3 line 32, it should be ‘different’.

Thank you for this suggestion, we have revised this in the latest version.

4 line 32-33, it should be ‘similar significant increase’. Please check the usage of two ‘for’.

Thank you for this suggestion, we have revised this in the latest version.

5 line 43-44, please reformulate it.

Thank you for this suggestion, we have revised this in the latest version.

6 line 46-47, please reformulate it.

Thank you for this suggestion, we have revised this in the latest version.

7 line 51, please use clause.

Thank you for this suggestion, we have revised this in the latest version.

8 line 53, what did you mean by ‘measurement bias’? Is it affected by the distribution of flux towers or just measurement techniques?

Here, we refer to the uneven distribution of flux towers which leads to the under/overrepresentation of certain landcovers/biomes in the training data. We have revised this in the latest version.

9 line 236, please reformulate it.

Thank you for this suggestion, we have revised this in the latest version.

10 line 245, for site-level validation, you should also use cross validation method.

We will include a 10-fold cross-validation analysis in addition to the uncertainty analysis.

11 line 264, the subtitle for section 4 is missing.

Thank you for making us aware of the missing subtitle. We have added the subtitle in the latest version.

12 In figure 2, the uncertainty is very high in desert area (more than ten time of the GPP itself, why?), for example, northern Africa, which is blank in figure 1. Please use single-hue color series and check to ensure the consistency with figure 1.

It is true that the uncertainties are very high in some very arid regions such as the Sahara. The exact reason for the low robustness of the model in these regions is unclear but we assume that these instabilities are partly explicable with missing training data from desert areas.

The reason for the missing consistency between Figure 1a and Figure 2 (old) is that in Figure 1a we only used data that is available in all three products (MODIS, FLUXCOM, VODCA2GPP) in order to allow comparability. We agree that this is confusing and added an additional figure including all available VODCA2GPP data which is consistent with Figure 2 (old).

Furthermore, we updated Figure 2 by using a single-hue color scale and merged it with the added figure of VODCA2GPP mean annual GPP.

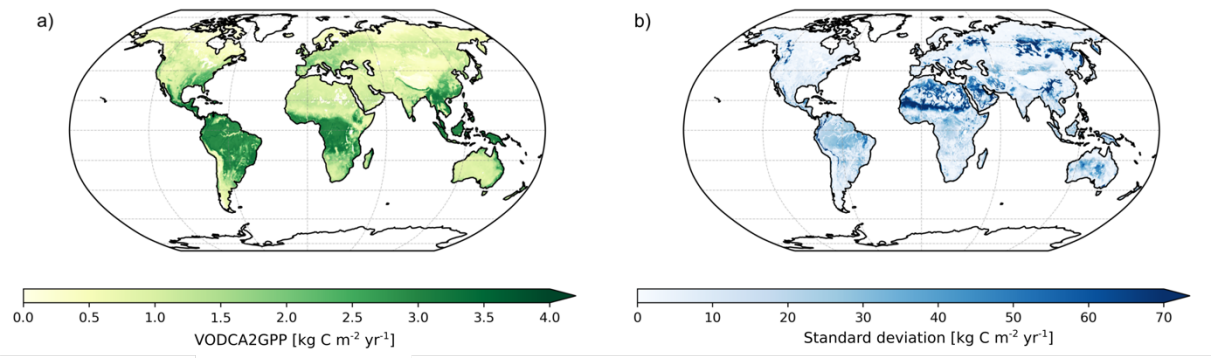

Figure 2 (new): a) Mean annual GPP as derived from VODCA2GPP for the period 1988-2019. b) Standard deviation of mean yearly annual GPP (1988-2019) as obtained by the uncertainty analysis.

13 In figure 3, please check the unit in both x and y axis. Please use dashed line for 1:1 line. Here, you presented the site-level evaluation for mean annual GPP. What about the annual, seasonal (or monthly) evaluation? In fact, Pearson value of around 0.5 seems to be relatively lower for GPP estimation. Did you try 10-fold cross validation for model performance?

In Figure 3, we adapted the units, replaced the 1:1 line and added the bias in the performance metrics:

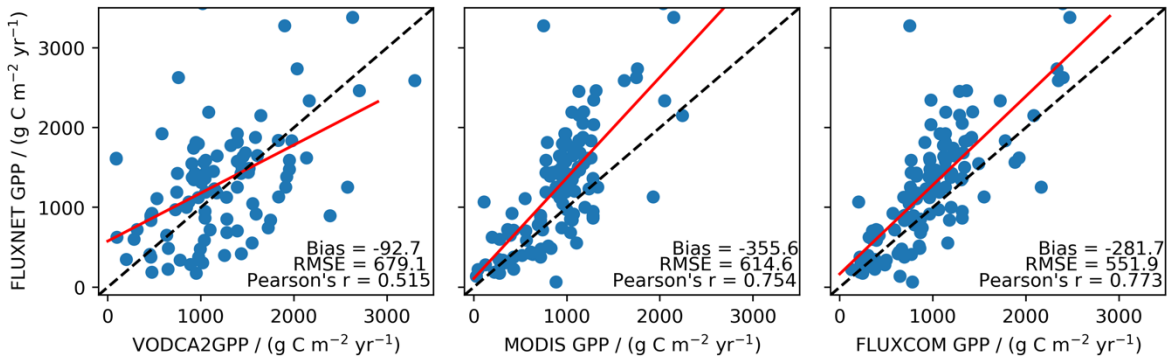

Figure 3 (new): Mean annual in-situ GPP (FLUXNET) plotted against mean annual GPP from VODCA2GPP, FLUXCOM and MODIS for the respective grid cells. Mean annual GPP was computed from all available overlapping years and thus each station is represented by one dot. Red lines indicate the best linear fits determined by ordinary linear regression and the black lines represent the 1:1 lines.

We agree that it would also make sense to use annual, seasonal or even monthly values for the comparison. A slight advantage of using mean annual values is that it can be visualized without a density map while still including the entire information from the data.

VODCA2GPP underperforms mainly for sparsely vegetated biomes, while performance is good for densely vegetated regions (Fig A1). In these regions VODCA2GPP's performance metrics (Pearson's r and RMSE) are comparable or even better than those from MODIS and FLUXCOM. Therefore, we suggest that VODCA2GPP can be used as a reliable source of GPP data as long as the users are aware of the performance differences depending on the landcover. We understand that this can be seen as drawback but considering the fact that VODCA2GPP performs well for most regions and is largely independent from current state-of-the-art products, we believe that VODCA2GPP can provide very valuable input for the scientific community. Due to the independence of VODCA2GPP, it also allows

users to gain insights on global GPP trends and anomalies that are partly hidden in optical remote sensing based GPP datasets.

We will include 10-fold cross validation in a revised version of the manuscript.

14 line 308, when you mention 'good temporal agreement', you need to specify the region, since the correlation is even negative in tropical America.

We agree that this needs to be reformulated more precisely. We will revise this.

15 line 308-309, it should be 'the highest' and 'the lowest'.

Thank you for this suggestion, we have revised this in the latest version.

16 line 311, please rephrase it.

Thank you for this suggestion, we have revised this in the latest version.

17 line 313, why monthly GPP anomalies are compared? Why you did not use monthly GPP instead?

Here, we wanted to evaluate VODCA2GPP's ability to capture extreme events and anomalies in general.

18 line 334, here you used 2002-2016. Why did you use 2003-2015 in next paragraph?

Thank you for bringing this to our attention. The period 2003-2015 is correct. We will revise this.

19 line 335, there are only three sentences in this paragraph. It would be better to merge it with next paragraph.

We will merge the paragraphs.

20 In figure 6, at global scale, the VODCA2GPP is two times of the MODIS and FLUXCOM. So the global GPP can be more than 200 PgC yr-1, which is the highest estimation as far as I know. You should give reasonable evidence for such value, or it won't be convincing.

It is true that VODCA2GPP estimates are substantially higher than MODIS and FLUXCOM estimates but as already outlined above, optical products are known to significantly underestimate GPP, especially in very productive regions (Fig 3 (new), Fig A1 (new); Turner et al., 2006; Wang et al., 2017). Thus, differences in global annual GPP between VODCA2GPP and MODIS/FLUXCOM need to be interpreted with caution. In fact, the comparison of VODCA2GPP with in-situ GPP does not suggest a systematic overestimation (Fig 3 (new)). VODCA2GPP only shows strong overestimation for (woody) savannas and open shrublands (Fig A1 (new)). For all FLUXNET sites located in other landcovers our analysis suggests very mild overestimation or even underestimation of VODCA2GPP. Furthermore, the comparison of VODCA2GPP with the mean of all TRENDY models (Fig. 6) indicates that the overestimation is by far not as severe as implied by MODIS and FLUXCOM.

21 In figure 7, it should be 2002-2016 or 2003-2015?

2003-2015 is correct.

22 In section 5.2, TRENDY ensemble is indeed independent from VODCA2GPP. But I couldn't get the meaning of this paragraph. The process-based estimation without accounting for nutrient constraints is also reported to overestimate GPP.

The objective of this section is to emphasize that the validation of global GPP datasets generally needs to be interpreted with caution due to a) the generally low availability and uneven distribution

of FLUXNET stations and b) due to the often shared in-situ data-source (i.e. FLUXNET). MODIS, FLUXCOM and VODCA2GPP use the same reference measurements for calibration and validation. As a result MODIS and FLUXCOM cannot be viewed as completely independent from VODCA2GPP although they are using fundamentally differing modelling approaches. Thus, this paragraph specifically aims to transparently inform potential users about these inevitable validation problems. We will merge this paragraph with section 5.4.

23 In section 5.6, in terms of the VODCA2GPP datasets, more similarities with TRENDY than that of other observational products were mentioned. Is it really a strong point for this GPP dataset?

As already mentioned in response 13, one strong argument for VODCA2GPP is its large independence from other (optical) observational products. This independence is evident especially when analyzing long-term trends in GPP where process-based and optical remote sensing products only exhibit very weak correspondence. VODCA2GPP on the other hand exhibits patterns that are visible in at least one other dataset which led us to the conclusion that VODCA2GPP is able to largely bridge the obvious gap between process-based and optical remote-sensing based products. Thus, not the good correspondence with TRENDY alone is a strong point for VODCA2GPP but more importantly its ability to complement existing products, both process-based and observational, is advantageous.

24 In figure A1, please check the unit in both x and y axis. You can give annual, seasonal (or monthly) evaluation as well.

We adapted the units in the axis labelling and added the bias in the performance metrics:

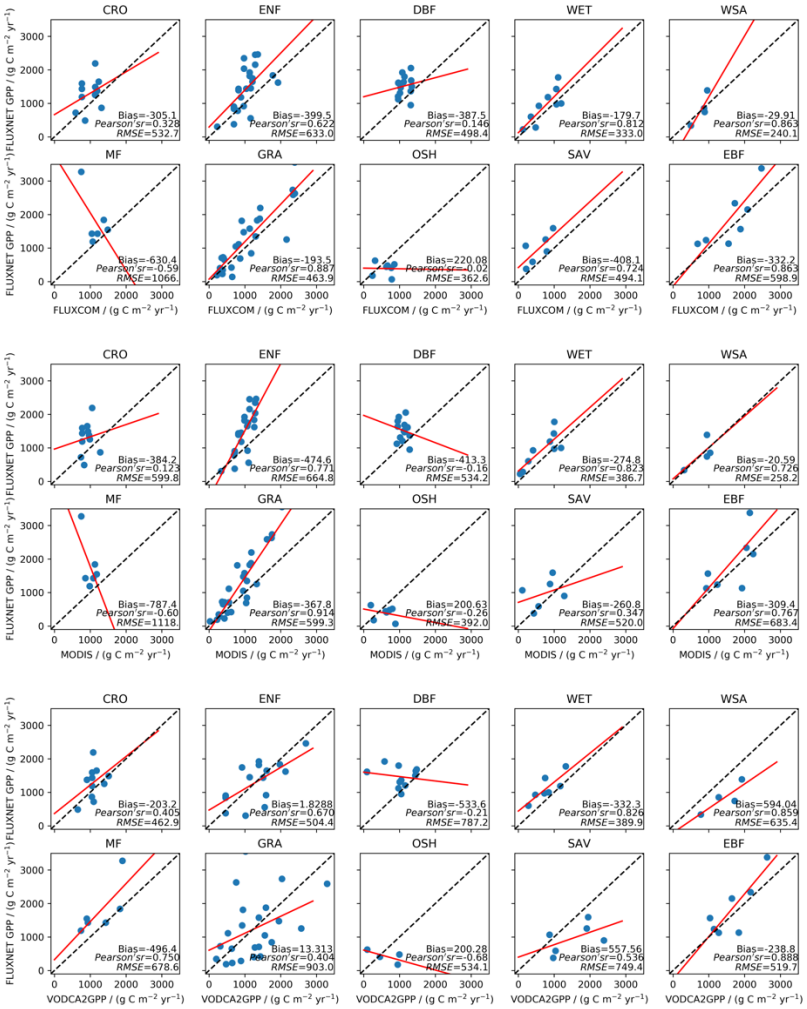

Figure A1 (new): Scatterplots of mean annual GPP for the period 2002-2016 per vegetation type. Vegetation types indicate the predominant IGBP-vegetation type at the respective FLUXNET station. Abbreviations: CRO: Croplands; ENF: Evergreen Needleleaf Forests; DBF: Deciduous Broadleaf Forests; WET: Permanent Wetlands; WSA: Woody Savannas; MF: Mixed Forests; GRA: Grasslands; OSH: Open Shrublands; SAV: Savannas; EBF: Evergreen Broadleaf Forests.

25 The code for producing the data and figures should also be publicly available in a repository for transparency reason.

We generally embrace the idea of open access science which is why we made the VODCA2GPP dataset publicly and freely available. However, at the moment we do not consider our work finished. We still strive for improvements of the VODA2GPP dataset which we will then of course also share with the scientific (and non-scientific) community. Therefore, we are currently not planning to make the VODCA2GPP code publicly available but we are happy to share the code with anyone who approaches us with ideas for collaboration. Regarding the transparency, we described the model algorithms and the data that is depicted in the figures in detail which is why we do not see a necessity to explicitly share the underlying code.

References

Anav, A., et al. (2015), Spatiotemporal patterns of terrestrial gross primary production: A review, Rev. Geophys., 53, 785– 818, doi:10.1002/2015RG000483.

Koffi, E. N., Rayner, P. J., Scholze, M., and Beer, C. (2012), Atmospheric constraints on gross primary productivity and net ecosystem productivity: Results from a carbon-cycle data assimilation system, Global Biogeochem. Cycles, 26, GB1024, doi:10.1029/2010GB003900.

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running,S. W., Zhao M., Costa, M. H., Kirschbaum, A. A., Ham, J. M.,Saleska, S. R., and Ahl, D. E.: Evaluation of MODIS NPP and GPP products across multiple biomes, Remote Sens. Environ.,102, 282–292, https://doi.org/10.1016/j.rse.2006.02.017, 2006.

Wang, L., Zhu, H., Lin, A., Zou, L., Qin, W., and Du, Q.: Evaluation of the Latest MODIS GPP Products across Multiple 770Biomes Using Global Eddy Covariance Flux Data, Remote Sensing, 9, 5, https://doi.org/10.3390/rs9050418, 2017.

Welp, L. R., Keeling, R. F., Meijer, H. A., Bollenbacher, A. F., Piper, S. C., Yoshimura,K., Francey, R. J., Allison, C. E., and Wahlen, M.:Interannual variability in the oxygen isotopes of atmospheric CO2 driven by El Niño. Nature, 477(7366):579–582, https://doi.org/10.1038/nature10421, 2011.