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

Formatting as follows:

Reviewer comments
Reply to comments
Changes to the manuscript

Anonymous Referee (#1)
Received: 7 November 2021

This is my second round of reviewing. The author's responses confirmed my concerns, and this study or this dataset includes several important flaws which may substantially mislead the future studies in this filed. Therefore, I strongly suggest the authors seriously consider the method and the GPP dataset, and take the efforts to develop the reliable method.

Response: Dear Referee,

Many thanks for reviewing our manuscript again. We appreciate your feedback, and we are sorry that we could not settle your doubts regarding the VODCA2GPP dataset.

First, the authors confirmed their estimate of global GPP reaches to 200 Pg C yr-1, which is almost double of the current estimates. Although the authors argued that their estimates are close to the estimates from Welp et al (2011) (150-175 Pg C yr-1) and Koffi et al. (2012) (146 Pg C yr-1). However, the estimate in this study also is higher than these two studies about higher 30%-50%. Besides, these two studies are based on atmospheric inversion methods to indirectly estimate GPP, and ecosystem respiration may highly impact their estimates to GPP. As the Welp et al (2011) claimed "best guess of 150-175" of GPP. On contrary, MODIS and FLUXCOM used site-based GPP observations to constrain their estimates, and which provide the robust estimates of GPP compared to Welp and Koffi.

Yes, global annual VODCA2GPP reaches 200 Pg C yr-1, which is indeed higher than observation-based GPP derived from MODIS or FLUXCOM. We agree that this positive bias suggests an overestimation of VODCA2GPP at a global scale, and we understand the concerns regarding this overestimation. We devoted a separate section (sect. 5.3) to this issue and discussed the drivers for this positive bias in respect to the optical remote sensing-based references in detail. Thus, we are aware of this issue and transparently inform about it. By no means we claim that the global estimates from our product are closer to the truth than other products.

What we want to highlight again, however, is the fact that when comparing VODCA2GPP with in-situ GPP from FLUXNET we only find substantial overestimation in water-limited regions (e.g., open shrublands and savannas). For most other biomes, the comparison with in-situ GPP does not suggest an overestimation of VODCA2GPP. Furthermore, we want to emphasize that there is no consensus in estimates of global annual GPP among existing datasets (Anav et al. (2015)) which complicates a fair validation of global annual GPP estimates. Thus, we agree that VODCA2GPP derived global annual GPP are too high in some biomes (e.g., arid regions), which presumably also leads to an overestimation at a global scale, but the magnitude of this overestimation cannot be quantified reliably as estimates for global annual GPP are contradictory.

The authors validated their GPP estimates at eddy covariance towers. VODCA2GPP are comparable to tower-based GPP as Fig. 1A shown. I am wondering that there are large differences over the global estimates. The method may have significant flaw that make it impossible to apply over global scale. Therefore, I strongly suggest the authors investigate the reliability of the method before producing global GPP dataset. As I pointed out that there are several unclear items in the model algorithms, which may induce large uncertainties for GPP estimates. For example, the response #5, the authors changed the definition of mdn(VOD) from landcover to vegetation density. It is totally confused what vegetation density means? By my knowledge, there is no concept of vegetation density, instead that we say Species Density, which is obvious different with the authors' idea. It is my largest concern the authors failed to propose the robust physiological principle for using VOD to estimate GPP at all.

Indeed, VODCA2GPP compares well with in-situ GPP from FLUXNET, which also indicates that VODCA2GPP can in principle be trusted. Also, the underlying method (the VOD2GPP model) is generally reliable, has been peer-reviewed by several experts, and was published in reputable journals (Teubner et al., 2018, 2019, 2021).

As outlined in our revised manuscript, the mdn(VOD) term in the model formulations represents a static vegetation biomass component which helps the model to subtract larger structural vegetation elements (e.g., stems and branches) and thus makes the VOD2GPP model more sensitive to photosynthetically active parts of the vegetation. So to say, the additional term mdn(VOD) allows the model to differentiate between canopy types that have similar VOD dynamics but different above-ground biomass. This is also visible in the partial dependency plot, where we assessed the influence of each input variable on the GPP estimates (Fig. 1 c). Regarding the nomenclature, we believe that vegetation density is a good term for intuitively describing the role of mdn(VOD) in the VOD2GPP model but we are open for other suggestions. As correctly suggested by the reviewer, species density is a different concept.

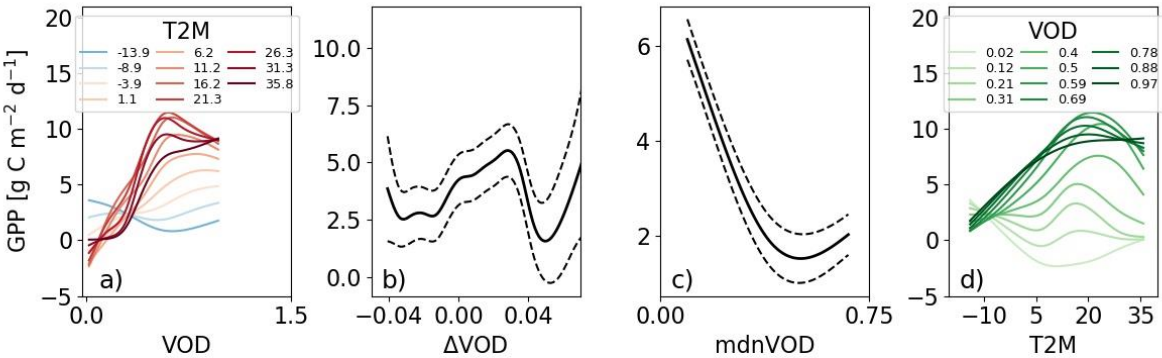

Fig 1: Partial dependency plot for GPP and the input variables: VOD (a), $\Delta(VOD)$ (b), mdn(VOD) (c), T2M (d). Dashed lines denote the 95% confidence interval. The interaction term between VOD and T2M is depicted as 3D surface which is bin-wise projected onto a 2D plane for visualization.

In addition, the authors examined MODIS and FLUXCOM dataset against eddy covariance-based GPP. However, the results in this manuscript look quite different with previous reports. Especially, FLUXCOM is data-driven dataset, which should be compared with site-based GPP. However, the authors showed the underestimated GPP by FLUXCOM, which is quite different with previous studies and also difficulty to understand.

The results of FLUXCOM GPP and MODIS GPP that we present may indeed differ from results presented in other studies, because of a different selection of ground data, and spatial and temporal subsetting. In general, however, our results are in line with what has been reported for FLUXCOM and MODIS GPP (e.g., underestimation of MODIS and FLUXCOM GPP in highly productive regions (Joiner et al., 2018; Anav et al., 2015; Turner et al., 2006)). Nevertheless, differences between our results and those of previous studies can arise for multiple reasons. First, we spatially aggregated

FLUXCOM GPP and MODIS GPP to 0.25° to match VODCA2GPP's resolution. This is most likely not the case for other studies and can thus lead to differences. Furthermore, we only considered pixels that are available in all three datasets (VODCA2GPP, MODIS GPP, and FLUXCOM GPP). Consequently, a certain share of available MODIS/FLUXCOM pixels, which are not available in VODCA2GPP, is not used in our comparisons but is included in other studies. Also, the observed time periods might differ between the studies, which further complicates a direct comparison. Considering these aspects, our results are not directly comparable with results from other studies.

Second, the VODCA2GPP dataset showed the low performance both over spatial and temporal scales. The authors added the validations on model performance for reproducing interannual variability of GPP (response #9). However, the performance is quite low, and mean R2 value is only 0.2 or even lower. By this low performance, I can not trust the capability of VODCA2GPP, and will not use it to conduct any further analyses. So, I still doubted why we still need VODCA2GPP dataset. The authors argued that we need other satellite data source besides optical data, but it is not a reason for accepting its low performance.

We do not share the reviewer's opinion that VODCA2GPP shows an overall low performance. The VODCA2GPP model performs reasonably well for reproducing 8-daily and monthly variations of GPP (Pearson's r: 0.53 and 0.6). It is true that median Pearson's r drops substantially at yearly sampling. This decrease, however, is explicable with extremely low availability of (FLUXNET) data points at this temporal scale. The average time span of FLUXNET time-series are only 7 years. When only considering significant correlations (p-value < 0.1) we find that median Pearson's r reaches much higher values. The number of significant values for Pearson's r at yearly sampling, however, drops to only 8.

We updated Figure 5 by only including significant correlations and we added the number of significant Pearson's r values in the caption to remind the reader about the relatively low expressiveness of this value. Furthermore, we added the following text:

(Page: 13, Line 326*): (…) It is to be noted that there are only 8 significant Pearson's r values for yearly sampling which decreases the expressiveness of this value. This is explicable with the short observation period of most FLUXNET sites which might not exhibit interannual variability. (…)*

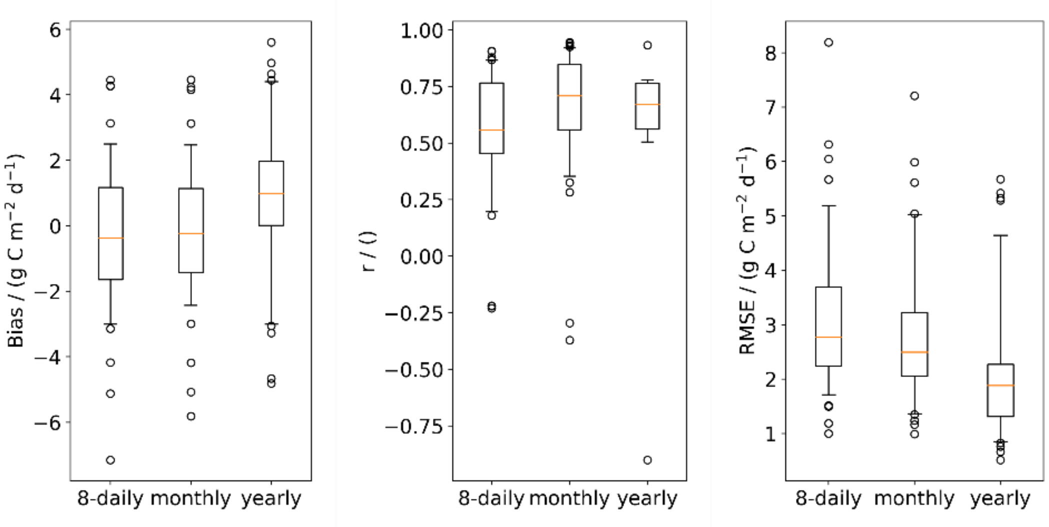

Figure 5: Site-based cross-validation for 8-daily, monthly, and yearly sampling of GPP from VODCA2GPP and FLUXNET. RMSE, Bias and Pearson's r were computed at each of the 10% of FLUXNET sites that were omitted during the respective training run. Non-significant Pearson correlation (p-value < 0.1) were ignored. The boxplots for

Pearson's r are based on the 71 (8-daily), 66 (monthly) and 8 (yearly sampling) significant values for Pearson's r values. The whiskers of the boxplots extend to the 0.05/0.95 percentiles.

References

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones,C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., and Zhao, M.: Spatiotemporal patterns of terrestrial gross primary production: A review, Rev. Geophys., 53, 785–818, doi:10.1002/2015RG000483, 2015.

Joiner, J., Yoshida, Y., Zhang, Y., Duveiller, G., Jung, M., Lyapustin, A., Wang, Y., and Tucker, J. C.: Estimation of Terrestrial Global Gross Primary Production (GPP) with Satellite Data-Driven Models and Eddy Covariance Flux Data, Remote Sens., 10, 1346, https://doi.org/10.3390/rs10091346, 2018.

Teubner, I. E., Forkel, M., Jung, M., Liu, Y. Y., Miralles, D. G., Parinussa, R., van der Schalie, R., Vreugdenhil, M., Schwalm, C. R., Tramontana, G., Camps-Valls, G., and Dorigo, W. A.: Assessing the relationship between microwave vegetation optical 725 depth and gross primary production, Int. J. Appl. Earth Obs., 65, 79–91, https://doi.org/10.1016/j.jag.2017.10.006, 2018.

Teubner, I. E., Forkel, M., Camps-Valls, G., Jung, M., Miralles, D. G., Tramontana, G., van der Schalie, R., Vreugdenhil, M., Mösinger, L., and Dorigo, W. A.: A carbon sink-driven approach to estimate gross primary production from microwave satellite observations, Remote Sens. Environ., 229, 100–113, https://doi.org/10.1016/j.rse.2019.04.022, 2019.730

Teubner, I. E., Forkel, M., Wild, B., Mösinger, L., and Dorigo, W.: Impact of temperature and water availability on microwave-derived gross primary production, Biogeosciences, 18, 3285–3308, https://doi.org/10.5194/bg-18-3285-2021, 2021.

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.

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

Formatting as follows:

Reviewer comments
Reply to comments
Changes to the manuscript

Anonymous Referee (#2)
Received: 25 November 2021

Compared to the last version, the manuscript is clearer now and fixes some problems. I still have a few concerns before considering the acceptance.

Response: Dear Referee,

Thank you very much for reviewing our manuscript again. We appreciate your feedback, and we are happy to address your concerns.

Minor comments:

1 Why does r drop with increase in time scale but RMSE does the opposite? You can add explanation in the end of line 383.

We assume that this comment refers to the cross-validation results depicted in Fig. 5. It is true that Pearson's r drops substantially at yearly sampling. The low Pearson's r values at this timescale, however, are more likely caused by relatively short observation periods of FLUXNET GPP than by weak performance of VODCA2GPP. The average FLUXNET GPP time-series only covers approximately 7 years and thus might only show little to no interannual variability resulting in a high number of non-significant correlations. In fact, the number of significant correlations (p-value < 0.1) is only 8 for yearly sampling. When removing all non-significant values from this analysis, median Pearson's r increases substantially for yearly sampling (from ca. 0.17 to 0.69), while the other time scales (8-daily and monthly) are only slightly affected from the exclusion of non-significant correlations (cf. Fig 5). Excluding non-significant correlations also removes the contrary behavior of Pearson's r and RMSE.

We updated Figure 5 by only including significant correlations and we added the number of significant Pearson's r values in the caption to remind the reader about the relatively low expressiveness of this value. Furthermore, we added the following text:

(Page: 13, Line 326*): (…) It is to be noted that there are only 8 significant Pearson's r values for yearly sampling which decreases the expressiveness of this value. This is explicable with the short observation period of most FLUXNET sites which might not exhibit interannual variability. (…)*

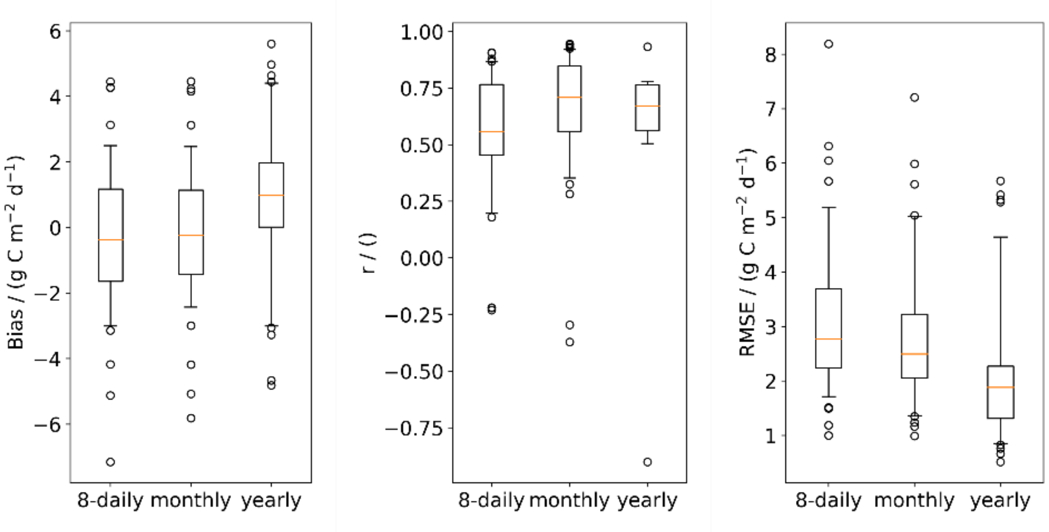

Figure 5: Site-based cross-validation for 8-daily, monthly and yearly sampling of GPP from VODCA2GPP and FLUXNET. RMSE, Bias and Pearson's r were computed at each of the 10% of FLUXNET sites that were omitted during the respective training run. Non-significant Pearson correlation (p-value < 0.1) were ignored. The boxplots for Pearson's r are based on the 71 (8-daily), 66 (monthly) and 8 (yearly sampling) significant values for Pearson's r values. The whiskers of the boxplots extend to the 0.05/0.95 percentiles.

2 Since the uncertainty metric is derived from the 10 VODCA2GPP models, you can also mention the source of such uncertainty. Is it can be regarded as extrapolation as well?

The uncertainty analysis is based on 10 VODCA2GPP models in which 10% percent of the station data was retained during each run. The results of this uncertainty analysis are depicted in Fig 1d and Fig 2b. Fig. 1d suggests that the choice of stations influences the resulting GPP estimates. We find the lowest spread in the 10 models (i.e., lowest uncertainty) north of 20° N where also the majority of FLUXNET GPP stations are located. The Southern hemisphere, where only few in-situ stations are located, generally exhibits a higher spread (larger uncertainty) indicating a considerable sensitivity to the choice of stations. This emphasizes the need for a globally well distributed network of in-situ flux towers.

Thank you for bringing up the extrapolation capabilities of the VODCA2GPP model which were tested through the site-based cross-validation. This analysis revealed that the VODCA2GPP model performs reasonably well at all time scales (Fig 5). The fact that higher correlations are found with global GPP from FLUXCOM and MODIS (median Pearson's r: 0.75 and 077) indicates that FLUXNET stations might not be always representative for the 0.25° VOD pixels.

We added the following text in the revised manuscript:

Page: 11, Line: 286: *(…) The lowest spread in the 10 models (i.e., the lowest uncertainty) is found north of 20°N where also the majority of FLUXNET GPP stations is located. The Southern hemisphere, where only few in-situ stations are located, generally exhibits a larger spread (higher uncertainty) indicating a considerable sensitivity of the model to the choice of stations. This emphasizes the need for a well distributed network of in-situ flux towers across all biomes. (…)*

3 You can also make comparison between your GPP uncertainty and that of other GPP datasets. Indeed, such VOD-GPP dataset is independent of optical-based one, but you really need remind the uncertainty in its application.

Thank you for this suggestion. It is indeed important to also discuss and compare the uncertainties of other global GPP datasets in respect to uncertainties of VODCA2GPP. Uncertainties in optical remote sensing-based GPP are mostly associated with the used wavelength. Optical remote sensing is heavily influenced by weather and illumination conditions. Clouds often contaminate or prevent the observations which is presumably the main reason why the largest uncertainties for FLUXCOM and MODIS are found in the tropics where GPP is known to be underestimated (de Almeida et al., 2018; Jung et al., 2020). In contrast to this, VODCA2GPP shows very good correspondence for densely vegetated areas (e.g., Broadleaf evergreen forests) and is hardly affected by weather conditions. However, we do find comparatively high uncertainties in water-limited areas (e.g., savannas and open shrublands) which presumably originate from multiple sources (which are discussed in the manuscript). The site-based uncertainty analysis also revealed that VODCA2GPP exhibits large uncertainties in mountainous regions with high topographic complexity which has also been reported to decrease reliability of GPP estimates in other GPP products (Xie et al., 2021).

Page: 18, Line: 418: *(…) A comparison of uncertainties between VODCA2GPP and optical remote sensing based GPP (Xie et al., 2021) shows that in both cases topographic complexity decreases the reliability. Furthermore, the reliability of GPP estimates based on optical remote sensing is highly dependent on weather and illumination conditions. Clouds often contaminate or prevent the observations which is presumably the main reason why the largest uncertainties for FLUXCOM and MODIS are found in the wet tropics where GPP is known to be underestimated (de Almeida et al., 2018; Jung et al., 2020). In contrast, VODCA2GPP shows good skill for densely vegetated areas, including broadleaf evergreen forests. On the other hand, the relatively high uncertainties of VODCA2GPP in water-limited regions have not been reported for FLUXCOM or MODIS GPP, indicating that these are VOD-specific and presumably caused by the abovementioned isohydric behavior of plants in arid regions.*

4 And a similar question to the last round, the process-based model cannot be treated as ground truth. I don’t think the similarity between VODCA2GPP and TRENDY models can be an advantage.

It is true, that process-based models cannot be treated as “true” reference or “ground truth”. The same applies to observation-based estimation approaches such as FLUXCOM or MODIS GPP. The source of data that comes the closest to actual “ground truth” is data from eddy covariance flux measurements. Therefore, our analysis and validations are first and foremost based on the comparison with in-situ estimates from FLUXNET. Due to the sparse and uneven distribution of FLUXNET stations, however, the comparison with other state-of-the-art GPP datasets such as TRENDY GPP is important to assess the validity of GPP estimates at a global scale. In various other GPP related studies, TRENDY GPP has served as reference data set for global patterns in GPP (e.g., O’Sullivan et al., 2020). Especially the evaluation of FLUXCOM GPP at global scale was largely based on the comparison with TRENDY GPP (e.g., Jung et al., 2020). We decided to add TRENDY GPP in our analysis not only to assess the validity of VODCAGPP trends and anomalies at global scale but also to highlight the diversity of GPP estimates that are currently available.

References

de Almeida, C. T., Delgado, R. C., Galvao, L. S., de Oliveira Cruz e Aragao, L. E., and Concepcion Ramos, M.: Improvements of the MODIS Gross Primary Productivity model based on a comprehensive uncertainty assessment over the Brazilian Amazonia, ISPRS J. Photogramm. Remote Sens., 145, 268–283, https://doi.org/10.1016/j.isprsjprs.2018.07.016, 2018.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., Chevallier, F., Gans, F., Goll, D. S., Haverd, V., Köhler, P., Ichii, K., Jain, A. K., Liu, J., Lombardozzi, D., Nabel, J. E. M. S., Nelson, J. A., O'Sullivan, M., Pallandt, M., Papale, D., Peters, W., Pongratz, J., Rödenbeck, C., Sitch, S., Tramontana, G., Walker, A., Weber, U., and Reichstein, M.: Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach, Biogeosciences, 17, 1343–1365, https://doi.org/10.5194/bg-17-1343-2020, 2020.

O'Sullivan, M., Smith, W. K., Sitch, S., Friedlingstein, P., Arora, V. K., Haverd, V., Jain, A., Kato, E., Kautz, M., Lombardozzi, D., Nabel, J, Tian, H., Vuichard, N., Wiltshire, A., Zhu, D., and Buermann, W.:Climate-driven variability and trends in plant productivity over recent decades based on three global products. Global Biogeochem Cy, 34, e2020GB006613. https://doi.org/10.1029/2020GB006613, 2020.

Xie, X., Li, A., Jin, H., Bian, J., Zhang, Z., & Nan, X.: Comparing Three Remotely Sensed Approaches for Simulating Gross Primary Productivity over Mountainous Watersheds: A Case Study in the Wanglang National Nature Reserve, China. Remote Sensing, 13(18), 3567, https://doi.org/10.3390/rs13183567, 2021.

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

Formatting as follows:

Reviewer comments
Reply to comments
Changes to the manuscript

Anonymous Referee (#3)
Received: 26 November 2021

This a is an overall good manuscript, well written. The interest and complementary information of microwave data with respect to optical data for vegetation studies, in particular for the estimation of the Gross Primary Production is clear.

However, I have a number of concerns and questions for the authors.

Response: Dear Referee,

Thank you very much for reviewing our manuscript and your overall positive feedback. We are happy to answer your questions and we believe that we can dispel your concerns.

This GPP product comes from a VOD archive that it is produced as a "sub-product" of the ESA soil moisture Climate Change Initiative data set. In the CCI data set dense forest regions as masked because retrievals are not considered to be reliable, in particular those of AMSR-E, AMSR-2. Soil moisture and VOD are retrieved simultaneously, therefore why the retrievals are considered to be good for VODCA over those regions but not for SM CCI ?

Thank you for this important remark and question. First of all, we want to mention that although soil moisture and VOD are retrieved simultaneously, the retrieval algorithm distinguishes between emitted radiance coming from the soil surface and that coming from vegetation. Thus, VOD and SM, although retrieved simultaneously, can and should be regarded as separate products. It is correct that CCI soil moisture data is masked over densely vegetated regions. This is done because in these regions the largest part of microwave emission is caused by vegetation and thus almost the entire signal (i.e. the measured brightness temperature) comes from the vegetation, making soil moisture retrievals unreliable, but not VOD retrievals. Yet, VOD tends to saturate for very dense vegetation making it less likely to distinguish variability. We also observe this tendency for saturation in VODCA2GPP in our analysis (e.g., Fig A6). Nevertheless, our landcover-based analysis of mean yearly GPP (Fig. A3) suggests that VODCA2GPP compares very well with FLUXNET in-situ GPP in densely vegetated regions (Pearson's r: 0.89 for Evergreen Broadleaf Forest) which led us to the conclusion that VODCA2GPP is reliable also in regions with dense vegetation.

This is indeed an important point, thus we added the following paragraph to the manuscript:
Page: 19, Line: 440: *(…) Furthermore, VOD retrievals exhibit a tendency for saturation in regions with very dense vegetation making it less likely to distinguish variability. A slight tendency for saturation was also observed for VODCA2GPP but the landcover based analysis exhibited a very high agreement between VODCA2GPP and in-situ GPP indicating high reliability of VODCA2GPP over densely vegetated regions. (…)*

VOD depends on frequency. What is the physical meaning of a rescaling by CDF matching? It is not rescaling apples (VOD at C band, for isntance) and oranges (VOD at K band, for instance)? Can the authors justify this approach and evaluate the impacts of this assumption into applications such as GPP estimation?

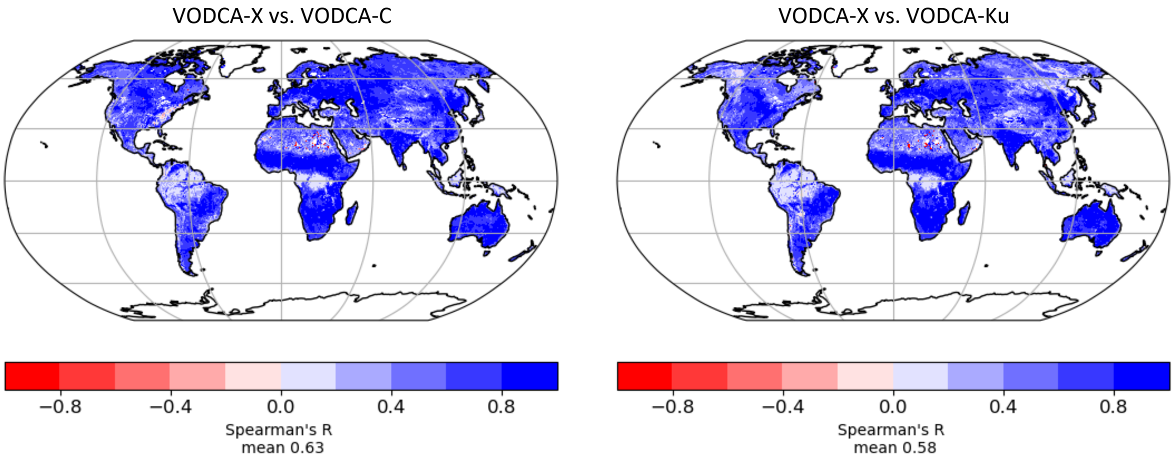

Figure 2a: Temporal correlation between X-band- and C-band VODCA (left) and between X-band and Ku-band VODCA (right). Only spatially and temporally collocated data has been used. The correlations are based on the overlapping observation period (2002-01 – 2018-08).

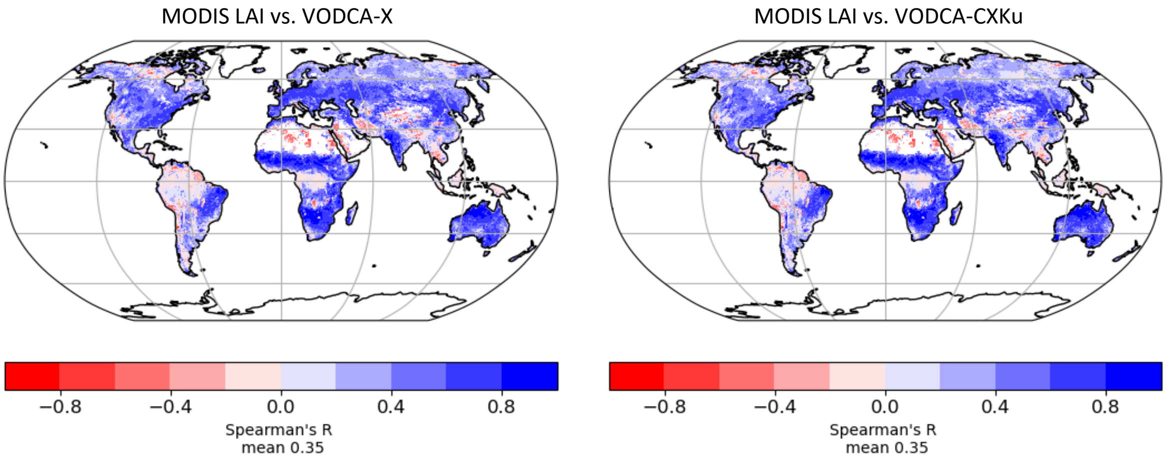

Figure 2b: Temporal correlation between MODIS LAI and X-band VODCA (left) and between MODIS LAI and VODCA CXKu (right). Only spatially and temporally collocated data has been used. The correlations are based on the overlapping observation period (2002-01 – 2018-08).

Thank you for this critical question. Indeed, VOD depends on the observed wavelength domain. Low-frequency observations such as those from L-band are sensitive to the water content in the whole vegetation, including the woody components, while high-frequency observations, such as C-, X-, and Ku-band, are more sensitive to the water content of the upper canopy layer (Li et al., 2021).

VODCA CXKu incorporates only high-frequency VOD products, namely C-, X-, and Ku-band, all of which indicate upper canopy dynamics and are highly correlated with each other. This is shown in Figure 2a, where for all biomes but those with little inter- and intra-annual variability (deserts and humid tropics) correlations are very high. In this figure, we show the correlation of C- (left) and Ku-band (right) with X-band VODCA, which has been used as scaling reference in the CDF-matching procedure. More so, in Figure 2b, we look at the agreement with MODIS LAI, which is an independent vegetation dataset related to leaf biomass (Tian et al., 2018). We can observe that the Spearman's R of MODIS LAI with X-band (left) and VODCA CXKu (right) are almost identical in all regions. This indicates that VODCA CXKu is very similar to the product used as scaling reference during CDF-matching, which is X-band.

However, VODCA CXKu exceeds the temporal length of the three single-frequency products, covering over 30 years of observation (1987 - 2020) and exhibits lower random error levels due to the merging approach employed. These features have led to an improved VODCA2GPP.

Line 138: ERA-5 Land resolution is not 8 km but 9 km

Thank you for making us aware of this. We revised this.

(Page: 5, Line 138): *(…) ERA5-Land is produced at a spatial resolution of 9 km (…)*

Line 187: I reckon that the dependency on time should be explicit in this equation or at least that the time scales at which those different VOD quantities are estimated should be explicit. Since VOD is approximately mdn(VOD)+delta(VOD) what is the real interest of adding a third term on VOD?

Thank you for this comment. The time scale at which delta(VOD) is derived is indeed crucial. We already incorporated this information in chapter 3.3 ("Preprocessing") but we agree that this term should already be explained at its first occurrence and in more detail. Thus, we added the following text (below Eq. 3.3) and removed the, now redundant, information from chapter 3.3:

(Page: 7, Line: 191): *Δ(VOD) is derived for each pixel (x_i) by computing the difference between two consecutive VOD observations of the smoothed and 8-daily aggregated VOD Signal (Teubner et al. 2019):*

$$\Delta(VOD) = VOD_{x_i,t_j} - VOD_{x_i,t_{j-1}}$$

The smoothing was performed in order to increase the robustness of the derivation and implemented using a Savitzky-Golay filter with a window size of 11 data points as suggested by Teubner et al. (2021).

Thank you for the question regarding the VOD term. For the answer to this question, we would primarily like to refer you to Teubner et al. (2019) who provide a detailed derivation of the theoretical background of the VOD2GPP model where also the relationship between GPP and VOD, delta(VOD), and mdn(VOD) is discussed in detail. Here, we provide a summary of the VOD2GPP theory which is based on Teubner et al. (2019):

For deriving the relationship between VOD and GPP we start with the relationship between GPP and ecosystem net uptake of carbon (NPP) and autotrophic respiration (R_a) (Bonan, 2008):

$$GPP = R_a + NPP, \qquad (1)$$

The VOD2GPP-model is essentially based on the assumption that R_α can be expressed as differential equation (Ryan, 1990):

$$R_\alpha = a_0 \left(\frac{dB}{dt}\right) + b_0\, B, \qquad (2)$$

The terms $\frac{dB}{dt}$ and B denote biomass (B) and temporal changes in biomass ($\frac{dB}{dt}$) and they are proportional to the two constituents of R_α, growth and maintenance respiration, respectively.

NPP can be approximately written as:

$$NPP \approx \left(\frac{dB}{dt}\right) + loss\ terms, \qquad (3)$$

As the $loss\ terms$ only make a small fraction of NPP and are not directly reflected in VOD, they are neglected in the VOD2GPP-model (Teubner et al. 2019). By combining Eq. 1-3, we can express GPP via the following differential equation:

$$GPP = NPP + R_\alpha \approx a\left(\frac{dB}{dt}\right) + b\,B, \tag{4}$$

Another assumption is that AGB can be expressed as a function of VOD:

$$AGB = f(VOD) = \widetilde{VOD} \tag{5}$$

Assuming that Biomass B can be expressed as AGB, we can rewrite Eq. 4 and find the theoretical relationship between GPP and VOD:

$$GPP = a\left(\frac{d\widetilde{VOD}}{dt}\right) + b\,\widetilde{VOD} + c \tag{6}$$

Eq. 2 shows that temporal changes in VOD ($\sim\frac{dB}{dt}$) are needed to represent growth respiration and NPP while the bulk VOD signal ($\sim B$) is needed for representing the maintenance part of R_α. As R_α exhibits a high sensitivity to temperature (Ryan et al., 1997) we included 2m surface temperature in an interaction term with VOD (Teubner et al., 2021). $Mdn(VOD)$ on the other hand corresponds approximately to the time-invariant offset c (Eq. 6) which is a static term and aids to convert VOD to GPP if the offset is not already represented in $\widetilde{VOD}$. In other words, mdn(VOD) helps to make the VOD2GPP model more closely related to photosynthetically active parts of the vegetation by subtracting larger structural vegetation components (e.g., stems) which is also visible in Fig 1 c.

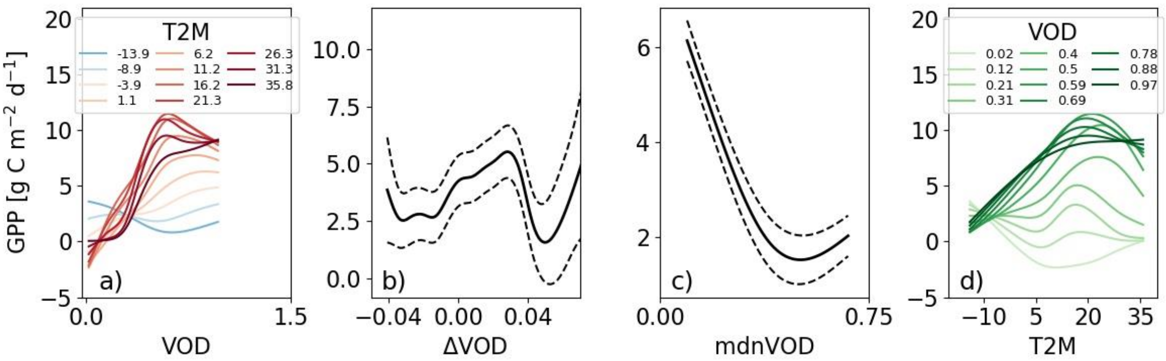

Fig 1: Partial dependency plot for GPP and the input variables: VOD (a), $\Delta(VOD)$ (b), mdn(VOD) (c), T2M (d). Dashed lines denote the 95% confidence interval. The interaction term between VOD and T2M is depicted as 3D surface which is projected bin-wise onto a 2D plane for visualization.

Figure 4: it is atypical to show the reference data in the y-axis. It is confusing for the reader. I strongly suggest inverting the axis. In addition, the linear regression equation should be shown or at least the slope and the intercept should be given in addition to R, RMSE and bias.

Thank you very much for making us aware of this. Having the reference data in the y-axis is indeed counter intuitive which is why we adapted our plots accordingly:

We swapped the x/y-axis in each scatter plot where the reference data (i.e., FLUXNET GPP) was in the y-axis (Fig. 4, Fig. A1, Fig. A2, Fig. A3). We added the linear regression line and equation in Fig. 4 and Fig. A2:

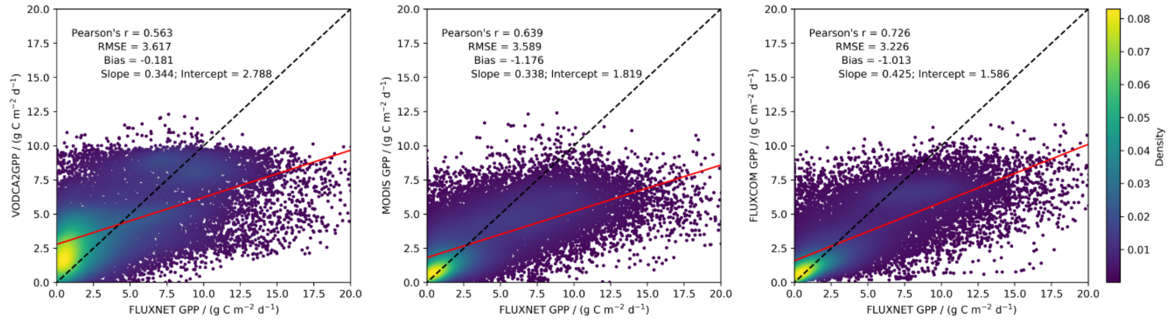

Figure 4: GPP from FLUXNET plotted against GPP from VODCA2GPP, MODIS and FLUXCOM for the period 2002-2016 with 8-daily sampling.

References

Bonan, G.: Ecological climatology: Concepts and applications, 2ed., Cambridge University Press, Cambridge, UK, New York, 550 pp., 2008.

Li, X., Wigneron, J. P., Frappart, F., Fan, L., Ciais, P., Fensholt, R., ... & Moisy, C. (2021). Global-scale assessment and inter-comparison of recently developed/reprocessed microwave satellite vegetation optical depth products. Remote Sensing of Environment, 253, 112208.

Ryan, M. G.: Growth and maintenance respiration in stems of Pinus contorta and Picea engelmannii, Can. J. Forest Res., 20, 48–57, doi:10.1139/x90-008, 1990.

Ryan, M. G., Lavigne, M. B., and Gower, S. T.: Annual carbon cost of autotrophic respiration in boreal forest ecosystems in relation to species and climate, J. Geophys. Res.-Atmos., 102, 28871–28883, 1997.

Teubner, I. E., Forkel, M., Camps-Valls, G., Jung, M., Miralles, D. G., Tramontana, G., van der Schalie, R., Vreugdenhil, M., Mösinger, L., and Dorigo, W. A.: A carbon sink-driven approach to estimate gross primary production from microwave satellite observations, Remote Sens. Environ., 229, 100–113, https://doi.org/10.1016/j.rse.2019.04.022, 2019.730

Teubner, I. E., Forkel, M., Wild, B., Mösinger, L., and Dorigo, W.: Impact of temperature and water availability on microwave-derived gross primary production, Biogeosciences, 18, 3285–3308, https://doi.org/10.5194/bg-18-3285-2021, 2021.

Tian, F., Wigneron, J.-P., Ciais, P., Chave, J., Ogée, J., Peñuelas, J., Ræbild, A., Domec, J.-C., Tong, X., Brandt, M., Mialon, A., Rodriguez-Fernandez, N., Tagesson, T., Al-Yaari, A., Kerr, Y., Chen, C., Myneni, R. B., Zhang, W., Ardö, J., and Fensholt, R.: Coupling of ecosystem-scale plant water storage and leaf phenology observed by satellite, Nat. Ecol. Evol., 2, 1428–1435, https://doi.org/10.1038/s41559-018-0630-3, 2018.