Title: Global Datasets of Leaf Photosynthetic Capacity for Ecological and Earth System Research

Author(s): Jing M. Chen, Rong Wang, Yihong Liu, Liming He, Holly Croft, Xiangzhong Luo, Han Wang, Nicholas G. Smith, Trevor F. Keenan, I. Colin Prentice, Yongguang Zhang, Weimin Ju, and Ning Dong

MS No.: essd-2022-136

Answer to Reviewer #1: Yao Zhang

Mapping the dynamics of Vcmax at global scale is important for the improvement of the model performance in predicting GPP and to understand the driving factors for its spatial and temporal variations. Recent studies have developed multiple methods to retrieve Vcmax based on satellite observations. This paper by Chen et al. summarized these approaches and provide a direct comparison between these datasets, the one predicted by optimality theory (EOT) as well as in situ observations. The satellite-based datasets generally show good consistency with the EOT and observations. The authors also evaluate the difference between the satellite observations and EOT and suggest that the difference can be explained by irrigation, soil PH, and nitrogen content. This is a solid paper and the developed datasets should be published. However, I still have some comments for the improvement of the manuscript.

Answer: Thank you for your appreciation of our work, and the critical and useful comments below that help improve our manuscript.

In the abstract, the authors mentioned that they use a data assimilation technique to combine the SIF generated Vcmax and LCC generated Vcmax to get an optimized Vcmax, I did not find the description of this data assimilation method. Later in the results, I feel that the authors are referring the TROPOMI SIF based Vcmax as the assimilated Vcmax. If this is the case, the presentation in the abstract should be revised. In the abstract, the authors suggest that the data assimilation technique is to combine "two types" of remote sensing dataset, one is SIF based, the other is LCC based. Clearly, TROPOMI SIF Vcmax, based on its names, should still be considered as SIF based. This naming system is misleading to the readers. I would suggest the authors to reconsider this naming system or revise the abstract.

We agree that the Vcmax product using the combined information of SIF and LCC was not clearly described in Abstract, although it was described in Methods (lines 123-130). We have modified Abstract in response, and TROPOMI SIF Vcmax has been changed to TROPOMI SIF+LCC Vcmax throughout the paper.

The authors suggested that irrigation may be the reason to explain the difference between satellite observed Vcmax and EOT predicted ones. I would argue that the improvement in the crop industry ("green revolution"), mostly better seeds, fertilization usages to be the plausible cause. This is based on the fact that the difference in satellite

and EOT predicted Vcmax is large over all cropland regions, no matter it is irrigated or not (e.g., irrigation cannot explain the difference in Africa and south America). Second, irrigation would provide enough water which tends to reduce Vcmax based on the optimality theory, this is different than what we see in this comparison.

This is an excellent point. The positive effect of irrigation on Vcmax found in our study should be taken as the surrogate effects of agricultural management including not only irrigation but also fertilization and genetic modification. In particular, fertilization usually accompanies with irrigation, as pointed out by Dennis Baldocchi (the other reviewer). Ecological optimality theory predicts higher Vcmax at higher vapor pressure deficit (Smith et al., 2019), which may be related to soil moisture, but the theory has not yet included soil moisture. However, leaf economics spectrum data (Wright et al., 2004; Osnas et al., 2013) show that annual precipitation and Vcmax are positively correlated. We have therefore added the text (Lines 237-240; Lines 266-279) to clarify this issue.

Wright, I. J., P. B. Reich et al., 2004. The worldwide leaf economics spectrum. Nature, 428, 821-827.

Osnas, J. L. D., J. W. Lichstein, P. B. Reich, and S. W. Pacala, 2013. Global leaf trait relationships: mass, area, and the leaf economics spectrum. Science, 340, 741-744.

The manuscript mostly focuses on the comparison of the spatial variation of Vcmax from different datasets. Based on my understanding, all three remote sensing-based Vcamx have seasonal variations. Previous studies have highlighted the importance of correctly representing the seasonal variation of Vcmax to the improvement of seasonal GPP simulations. This seems to be an advantage of the dataset. But I did not see much stress on this temporal variation throughout the manuscript, this is also no cross comparison of these datasets at temporal scales.

Indeed, in this paper we focus on the analysis of spatial variation of Vcmax without exploring its seasonal variation. There are several reasons for this focus: (1) ground-based data used in this study do not have seasonal variation, although there are a limited number of data points with seasonal variation but they are insufficient for validation purposes; (2) the ecological optimality theory can so far be used to derive the mean Vcmax values over the growing season; (3) SIF data are often not reliable over non-growing seasons; and (4) annual patterns of retrieved LCC have irregularities in some places because of inaccuracies of input LAI outside of the growing season. The current state of the art in remote sensing retrievals of Vcmax using SIF and LCC provides reliable values of seasonal maximum and mean of Vcmax, which are a solid step forward, while efforts are being made to retrieve the annual variation of Vcmax through improving the algorithms and developing new algorithms. We expect that Vcmax datasets with reliable annual variations will soon be available. We have added statements (Lines 340-350) on this issue.

Detailed comments:

L31, why three? LCC, SIF and the optimized one?

Yes, we have modified the abstract to make it clear.

L32, the link provides two SIF based Vcmax, which is not described here.

The two SIF-based products are in fact one SIF and one SIF+LCC. We have modified the description on the link.

L48, it would be good to briefly describe how Vcmax can be derived from SIF, you did this for LCC later but not here.

This line is expanded to provide the first principle of deriving Vcmax from SIF (now Lines 50-51).

L64, and SIF is quite noisy.

SIF signals are indeed small and often noisy from various sources including variations in solar observation and sensor view angles. In order not to interrupt the flow of the text, we added a sentence (Lines 101-104) to explain this.

L69, … to produce a global Vcmax time series dataset? Single time series may refer to only one vector.

We agree and have changed “time series” to “map series”, and the following sentence is also adjusted to make it consistent.

L98, the SIF-photosynthesis relationship is only linear at longer time scales (weekly or monthly), you may want to specify this. This sentence can be misleading considering you use “instantaneous”.

We agree and have changed “is approximately proportional to” to “increases with” in the sentence (now Line 100). The nonlinearity issue is actually considered in our improved Vcmax optimization method (see Lines 125-128).

L100, “sunlit leaves are the predominant sources of SIF” a reference would be helpful here.

A reference of Pinto et al. (2016) is now provided (Line 104).

L150, were these obtained from sunlit leaves only? The remote sensing datasets are for the sunlit leaves, right?

The ground measurements of Vcmax can be made on any leaves, while remote sensing of SIF signals is mostly obtained from sunlit leaves at the time of measurements. However, all leaves in a canopy have probabilities to be sunlit. There is indeed some

mismatch between remote sensing and ground data in terms of leaf sample distribution in the vertical direction in the canopy. This issue deserves further investigation with detailed ground measurements. In response to your and another reviewer's question, a paragraph (Lines 333-339) is added in Discussion regarding the use of Vcmax products for both sunlit and shaded leaves.

L165, I was expected to see the equation here.

As this equation involves many variables and constants, we feel that it is not necessary to take up much space here, but a clear source of the equation is given.

L224, also plant genetic engineering. I think this may be a more plausible reason to explain the difference between TROPOMI and EOT. Human selections are producing much more productive crops that the optimality theory cannot predict. It happens that the much of croplands have irrigation. In Fig. 4b, the different is obvious in almost all croplands across the globe.

While we are not sure if genetic engineering for productive crop species would result in higher leaf Vcmax, we certainly agree that cropland and grassland management through irrigation and fertilization would increase Vcmax. While the impact of irrigation on Vcmax is uncertain, fertilization would directly increase leaf nitrogen and hence Vcmax. Since irrigation and fertilization generally co-occur in cropland and grassland and an irrigation dataset is available at the global scale, we used the irrigation dataset as a surrogate for the cropland and grassland management. We have added a few sentences (Lines 237-240) to clarify the confusion.

L228, but the optimality theory predicts lower Vcmax at regions with abundant water resource.

The optimality theory so far has not considered the influence of soil water on Vcmax, but the water effect may be indirectly considered through air water vapor deficit. Again, the positive irrigation effect on Vcmax stated here is associated with overall cropland and grassland management as explained above.

L258, I think you mean biome level Vcmax here.

Yes, we present global mean values of Vcmax for different biomes in this section. "for different biomes" is added to the section heading.

L259, not sure if TROPOMI is the dataset obtained from data assimilation. This needs to be clarified in the method.

The TROPOMI Vcmax dataset mentioned here is obtained through data assimilation using TROMPOMI SIF data and LCC data described in Methods. To avoid confusion, we

now changed “TROPOMI dataset” to “TROPOMI SIF+LCC dataset”, and the Methods section is also slightly modified to reflect the new naming convention (Line 128).