

Review report of manuscript esd-2021-69 “Divergent historical GPP trends among state-of-the-art multi-model simulations and satellite based products”

This study investigates global and regional GPP trends during 1982-2015, based on Dynamic Global Vegetation Model (DGVM) TRENDY v6 multi-model simulations, machine learning technique based GPP products, multiple GPP data sets derived from satellite-based products, FLUXNET observed GPP products, Long-term satellite-based near-infrared radiance of vegetation (NIRv) products. DGVM ensemble is consistent with NIRv but inconsistent with satellite-based GPP products and FLUXNET observations. Significant uncertainties existed in the distribution of GPP trends among models. Most of the models used in this study have not considered Nitrogen limitation and Phosphorous availability. The manuscript re-visits all the above products and discusses the limitations. The manuscript deals with important datasets, but the results are not easy to connect.

Review Comments:

(MAJOR REVISION)

1) Abstract: “Machine Learning Technique”.....study uses the FLUXCOM product. No machine learning algorithm was used, so this sentence may be revised.

2) Abstract: “Trends after 2000 was different from the full time-series....”, Satellite-based GPP product showed a decreasing trend, and DGVM showed an increasing trend mainly because LAI was not represented well in satellite-product. Was LAI represented well in TRENDY model simulations? Also, why trends after 2000 was different from full-time series?

3) Section 2.2: FLUXCOM GPP product area percentage with no significant trend is 74.4% (Table 3), highest among all products. The spatial correlation of the annual GPP trend with NIRv is -0.26. However, the spatial correlation of climatological annual GPP trend with NIRv is +0.93. Can you explain this contrasting feature? Also, revise this section for clarity.

4) Section 2.3, last line: “....datasets are highly consistent.....”, Spatial pattern and temporal changes of these datasets are NOT highly consistent. There are

differences, e.g, ORCHIDEE shows a much stronger signal than VEGAS over South America and Africa. Revise this section.

5) Section 2.5: FLUXNET observations used in this study are distributed over Northern Hemisphere only. However, the performance of GPP products evaluated with these observations is distributed globally. How are the Southern Hemisphere GPP products evaluated?

6) Section 2.6: Message not clear. Describe the section in detail.

6) Section 2.7, line 180: "We then calculated the global and regional.....historical changes in GPP". Message not clear; revise the sentence

7) Linear trend calculated. Whether all the GPP datasets in each pixel show a linear trend?

8) NIRv should be discussed as a separate sub-section in the Datasets and Methods section

9) Fig1(a): NIRv is a unitless variable. Is it an index like NDVI or EVI? The magnitude of NIRv is in the order of the fourth decimal. Is it possible to normalize it between 0 to 1 for better representativity?

10) Fig.1(b): "DGVM ensemble mean". However In Table-3, it is represented as "TRENDY ensemble". Be consistent with naming.

11) Section 3.1, line 235: ".....consistent with NIRv in middle and high latitudes" is not valid for Southern Hemisphere.

12) Fig 2 (a,b): FLUXCOM dataset GPP trend is near zero in most latitude bands. Describe this in the text.

13) NIRv trend is in the fourth decimal number in Fig.1 (a). However, it varies between 1 to 3 in Fig.2(a). Why such discrepancy?

14) Fig.2(a,b): DGVM ensemble and NIRv products are higher in latitude bands -20:+20. However, the change of vegetated land area is greater in +30:+60 latitude bands. Why?

15) Line 250: "In comparison, the increase of GPP in satellite-based GPP productsto the NH(60%)". Valid for Trop+SH as well. Revise it accordingly.

16) Fig.3 (a,b): The highest trend has shown during the year 2000. However, it is absent in the Northern Hemisphere (NH).