Referee #1

We thank the reviewer for the constructive comments and suggestions, which are in black text below. Our itemized response is followed (in red).

Background

Binghao Jia et al. investigates the effect of CO2, climate and land use change on the inter-annual variation and seasonal cycle of gross primary production (GPP) in China using 12 terrestrial biosphere models and observation driven data. Their main finding, in general, is that climate was the dominant control factor during 1981-2010 for the trends, inter-annual variations, and seasonality of China's GPP. A rise in CO2 increased GPP in China with increased inter-annual variability especially in the places where vegetation is dense.

- → I like the way authors choose to analyze the GPP data and perform the statistical tests from 12 models along with an observation-based estimate.
- → The nonparametric method is used to test statistical significance.
- → Figures are carefully chosen to communicate the essential results.
- → References are appropriately cited.

I find the paper is well written and the presentation is excellent. I thus recommend this paper for publication once the following minor issues are addressed.

Comments:

1. I am a bit surprised that China Land Use/Cover Dataset in Fig. 7 shows a decrease in cropland areas at end of the period (1982-2010) and also many satellite-based studies (you have also listed many peer-reviewed) over China shows large afforestation but Terrestrial biosphere models show an increase in cropland areas?? Then how come observation-based estimate of GPP you have shown agrees very well with the model simulated GPP? Could you please clarify this more clearly in the text?

Response: Based on the comments, we added some discussions about the effect of LULCC on the comparison between model simulated GPP and observation-based estimates.

(1) The LULCC data used in the TBMs from MsTMIP were generated by combining a static satellite-based land cover product (Jung et al., 2006) with time-varying land use harmonization version 1 (LUH1) data (Hurtt et al., 2011). The satellite-based LULCC data set, named the China Land Use/Cover Dataset (CLUD) (Liu, et al., 2003, 2005, 2010, 2014; Kuang et al., 2016), was generated using two satellite datasets: the LandsatTM/ETM+ and HJ-1A/1B images from the China Centre for Resources Satellite Data and Application (http://www.cresda.com/). In general, the LULCC data used in the MsTMIP agree well with the CLUD between 1990 and 2005, except some discrepancies in 2010. Compared to that in 2000, the CLUD showed a slight increase in forest (from 20% to 22%) and shrinking cropland (from 31% to 21%) and grassland (from 20% to 14%) in 2010 for the whole China. The decrease in cropland is mainly from R1 (Fig. 8b), R4 (Fig. 8e), R5 (Fig. 8f) and R7 (Fig. 8h). The main cause may be related with the expanding urban and forest areas. For example, rapid urbanization over eastern China induce the decrease in the cropland. In addition, due to the government-issued policies for protecting the environment, many national afforestation and reforestation projects have been implemented in China, which lead to the conversion of cropland to forest. In contrast, LULCC used in the TBMs show an increase in the cropland.

(2) The differences in the LULCC data indeed affect the model simulated GPP. For example, the MTE GPP products show a significantly increasing trend after 2005 over R4 (Fig. 4e), R5 (Fig. 4f) and R7 (Fig. 4h) while some underestimations can be found for model simulated GPP. This may be related with the discrepancies in the LULCC data sets over these areas. Please see Page 10 (Lines 2−4).

2. I suggest moving supplementary Figure S1 to main Figure 2. This is an important figure and also you are discussing this right at the start of the results section and I think this should be moved. There is an inset figure in each panel of this figure? What is it? You don't discuss. Explain what is it otherwise remove! Also, in the caption please be clear that the results shown in Fig.S1 (also many other figure captions) are the average of 28 years or what?

Response: (1) Based on the suggestions, old Figure S1 has been moved to the revised manuscript to be the new Fig. 2. Please see Page 25. (2) The inset figure in the bottom-right corner of Fig. 1, new Fig. 2, new Fig S2 represents the boundary of China. We have removed these inset figures based on your suggestions. (3) The results shown in the new Fig. 2 (old Fig. S1) are the averages of 30 years for MsTMIP models (1981−2010) and 29 years for the MTE (1982−2010), respectively. We have revised the captions of all the relative figures (Figs. 2, 3, 5, 6) according to the comments. Please see Pages 24−25, 27−28.

3. Page 5, two lines above the line 25: I suspect that Fig.2a instead of Fig.1a. Also, the GPP range you mentioned 4.9 to 9.2 PgC/yr, but I see Fig.2a maximum value in the boxplot is more than 10PgC/yr.

Response: These sentences have been revised based on the suggestions. It should be Fig. 3a (old Fig. 2a). The total China GPP ranges from 4.9 (DLEM) to 10.5 (GTEC) Pg C yr^{-1}. Please see Page 5 (Lines 25−26).

4. In general, throughout the text, it would be convenient for the readers if you mention also the abbreviation for the regions (R1, R2, etc..). For eg. at Page 5, two lines below line 25: southeastern China (1.3 PgC/yr, R7) and (1.5 PgC/yr, R9)? I suppose 1.5 PgC/yr corresponds to R9.

Response: Based on the suggestion, all the abbreviations have been added into the revised manuscript. For this example, it has been revised to be: "*The regional sum of GPP in southwestern China from the ENSEMBLE (Fig. 3b) was the highest among all nine regions (1.5 Pg C* yr^{-1}*, R9), followed by southeastern China (1.3 Pg C* yr^{-1}*, R7) and southern China (1.0 Pg C* yr^{-1}*, R8)*". Please see Page 5 (Lines 29−31). The other revisions can be found in the revised manuscript.

5. Why MTE abbreviation for the machine learning algorithm?? not sure how you have chosen MTE?

Response: Based on the comments, we added more descriptions about the MTE. The observation-based GPP product were generated using the machine-learning algorithm, model tree ensembles (MTE). Therefore, we used the "MTE" to represent this product. Please see Page 4 (Lines 18−28): "*This study used an observation-driven global monthly gridded GPP product derived from FLUXNET measurements by statistical upscaling with the machine-learning algorithm, model tree ensembles (Jung et al., 2009, 2011) (hereafter referred to as MTE). The MTE statistical model consisting of a set of regression trees was firstly trained using site-level explanatory variables and GPP estimations from eddy flux tower measurements. These explanatory variables covered climate and biophysical variables such as vegetation types, temperature, precipitation, radiation, and satellite-derived fraction of absorbed photosynthetic active radiation. Then the MTE GPP product was generated through applying the trained regression trees for global upscaling using gridded data sets of the same explanatory variables. It has a spatial resolution of 0.5° × 0.5° and is available between 1982 and 2011. The uncertainty of the MTE data is ~46 g C m^{-2} yr^{-1} (5%), which was calculated using the standard deviation of the 25 model tree ensembles (Jung et al., 2011)*".

6. In Figures 3 & 7 legends should be at the top left/right panel (eg. Fig.3a or 3b), to avoid wondering which color is what for a while. Readers usually start looking at the first panel of the figure before going to the bottom panels.

Response: Based on the suggestions, we have moved the legends to the top-left panel of the two figures. Please see the new Fig. 4 on Page 26 and new Fig. 8 on Page 30.

7. At page 7, near line 10: why some discrepancies between SG3 and MTE over northeast China, southeastern China and east parts of southwestern China? Worth explaining there!

Response: Based on this comment, we added some explanations about the differences of GPP trend between SG3 and MTE over these regions. Since MsTMIP SG3 simulations and MTE product were generated using different methods, we then compared them with another GPP data set from Yao et al. (2018) (hereafter YAO, Fig. 4a in that paper). YAO is a new GPP product for China with higher spatial resolution (0.1°) based on the same machine-learning algorithm with the MTE product, but it uses more eddy flux observations (40 flux sites). It is found that SG3 from MsTMIP shows similar trends with YAO over R1 and east parts of R7. In contrast, MTE shows the same increasing trends with YAO over east parts of R9. It suggests that both model simulations from MsTMIP and MTE GPP product shows certain uncertainties in the GPP trend over some areas of China, which needs more observations to evaluate the GPP trend in future work. Please see Page 7 (Lines 12−18).

References

[1] Hurtt, G. C., Chini, L., Frolking, S., Betts, R., Edmonds, J., Feddema, J., Fisher, G., Goldewijk, K. K., Hibbard, K., Houghton, R., Janetos, A., Jones, C., Kinderman, G., Konoshita, T., Riahi, K., Shevliakova, E., Smith, S. J., Stefest, E., Thomson, A. M., Thornton, P., van Vuuren, D., and Wang, Y.: Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Clim. Change, 109, 117–161, doi:10.1007/s10584-011-0153-2, 2011.

[2] Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling, Remote Sens. Environ., 101, 534–553, doi:10.1016/j.rse.2006.01.020, 2006.

[3] Liu, J., Kuang, W., Zhang, Z., Xu, X., Qin, Y., Ning, J., Zhou, W., Zhang, S., Li, R., Yan, C., Wu, S., Shi, X., Jiang, N., Yu, D., Pan, X., and Chi, W.: Spatiotemporal characteristics, patterns, and causes of land-use changes in China since the late 1980s, J. Geogr. Sci., 24(2), 195–210, http://dx.doi.org/10.1007/s11442-014-1082-6, 2014.

[4] Liu, J., Liu, M., Tian, H., Zhuang, D., Zhang, Z., Zhang, W., Tang, X., and Deng, X.: Spatial and temporal patterns of China cropland during 1990–2000: An analysis based on Landsat TM data, Remote Sens. Environ., 98(4),442–456, 2005.

[5] Liu, J., Liu, M., Zhuang, D., Zhang, Z., and Deng, X.: Study on spatial pattern of land-use change in China during 1995–2000, Science in China Series D: Earth Sciences, 46(4), 373–384, 2003.

[6] Liu, J., Zhang, Z., Xu, X., Kuang, W., Zhou, W., Zhang, S., Li, R., Yan, C., Yu, D., Wu, S., and Jiang, N.: Spatial patterns and driving forces of land use change in China during the early 21st century, J. Geogr. Sci., 20(4), 483–494, http://dx.doi.org/10.1007/s11442-010-0483-4, 2010.

[7] Kuang, W., Liu, J., Dong, J., Chi, W., and Zhang, C.: The rapid and massive urban and industrial land expansions in China between 1990 and 2010: A CLUD-based analysis of their trajectories, patterns, and drivers, Landscape Urban Plan., 145, 21–33, 2016.

[8] Yao, Y., Wang, X., Li, Y., Wang, T., Shen, M., Du, M., He, H., Li, Y., Luo, W., Ma, M., Ma, Y., Tang, Y., Wang, H., Zhang, X., Zhang, Y., Zhao, L., Zhou, G., and Piao, S.: Spatiotemporal pattern of gross primary productivity and its covariation with climate in China over the last thirty years, Glob. Change Biol., 24, 184–196, 2018.