

Reviewer 1

The authors investigated the future trajectories, patterns and uncertainties of northern ecosystem carbon fluxes using an ensemble of CMIP6 climate models. They found that under future warmer climates, the CMIP6 models project an overall enhanced net carbon uptake by the ecosystems, resulted from a tradeoff between NPP and RH increases. The spread of estimates across individual models is larger than that of the global average. The paper is methodologically sound, well written, and the results are nicely presented. I have a number of line comments on the manuscript, as given below. I would like to recommend its publication with the following possible revisions or clarifications.

Line comments:

Lines 113-114: It's better to use "ten models" directly, rather than "nine models with ten datasets".

**Response: we used 'ten models' instead in the revised manuscript.**

Line 123: I don't understand why land surface temperature, rather than 2-m temperature, is used in this analysis. When people say ecosystem response to temperature, they often refer to air temperature.

**Response: Thanks for pointing out this. In the revision, we replaced the land surface temperature with 2-m air temperature in our analyses and reported the new results in the updated manuscript. We found no essential changes to the results and findings with the new variable. But it is indeed true that the use of 2-m air temperature will potentially make our report more useful.**

Line 129: Do you mean the original annual outputs from models, or the annual values aggregated from original monthly outputs?

**Response: We meant the annual values aggregated from original monthly outputs. We clarified this in the revised manuscript.**

Line 130: If the model data are resampled to 1-degree global grids, this is supposed to be 360x180? Please clarify.

Response: The ‘1-degree grid’ was an approximation. In practice we resampled the model data based on the model grids of BCC-CSM2-MR whose resolution is in the middle (globally 320x160). We clarified it in the revised manuscript.

Lines 145-151: The model ensemble mean of global NEP is strongly higher than the estimate by the Global Carbon Project. What’s the implications for the NHL NEP and future projections of NEP changes?

Response: Thanks for pointing out that. We focused on NEP in our analyses while Global Carbon Project (GCP) reported Net Biosphere Productivity (NBP) which took disturbance-induced carbon fluxes from NEP and thus was lower. To make a meaningful comparison between CMIP6 and GCP data, we added historical CMIP6 NBP into analysis in the revision. According to the updated Fig.1a, the CMIP6 estimated lower global NBP than GCP data, and the CMIP6 NBP is closer to GCP data than NEP,

Lines 167-172: It’s interesting to look at how the NHL mean NEP compares with the global mean, and whether this difference is contributed more by NPP or RH.

Response: Thanks for the suggestion. We looked into the mean carbon fluxes and added the following sentences:

“Except SSP126, similarly positive but generally smaller trends were found for RH at global scales (Figure S4, Table 2) with the rates as 87.15, 173.39, 254.43 and 318.31 Tg C/year<sup>2</sup> under the four scenarios. The NHL RH trends are 18.64, 36.27, 55.39 and 72.56 Tg C/year<sup>2</sup>, normalized by the area, the growth rates are 0.44, 1.33, 1.99 and 2.62 g C/year<sup>2</sup> for global NPP over the four scenarios respectively. The area-normalized growth rates in the NHL NPP are 0.54, 1.37, 2.03 and 2.63 g C/year<sup>2</sup>, respectively. Area-normalized global RH growth rates are 0.59, 1.17, 1.72 and 2.15 g C/year<sup>2</sup> while the area-normalized NHL RH growth rates are 0.62, 1.20, 1.84 and 2.41 g C/year<sup>2</sup> under the four scenarios, respectively. These results indicate that a faster average growing NPP and RH in the NHL than the global average. The fast-growing RH cancelled a large part of the NPP growth and resulted in small growing NEPs.”

Figure 2 and 3: Please also include the multi-model ensemble mean result.

Response: we added multi-model ensemble mean results in the updated Figure 2 and 3.