General Comments

The manuscript by Qiu and colleagues analyzes future projections from models included in the Coupled Model Intercomparison Project Phase 6 (CMIP6), with a focus on carbon fluxes in the northern high latitudes (NHL) compared to rest of the globe. The authors find that the CMIP6 models project terrestrial ecosystems to remain carbon sinks, with net sequestration rates increasing with global temperature changes, and that NHL ecosystems are a relatively minor fraction of the net sink but considerably more uncertain than the global mean.

Overall, this is a straightforward and relatively simple paper documenting CMIP6 output in the NHL domain. It reads somewhat like a report and contains little novelty, using familiar tools and approaches with largely unsurprising results, especially for readers following the evolution of these models and how they represent the carbon cycle. Nonetheless, the data being reported is important to the scientific community. The manuscript would benefit from additional context and discussion, grammar checks, and potentially additional analyses. It’s also unclear to me if the authors are actually analyzing the net carbon sink (Net Ecosystem Productivity vs. Net Biome Productivity).

Specific comments
My main technical question is why the authors chose to analyze Net Ecosystem Productivity, which is the balance between photosynthesis and ecosystem respiration, versus Net Biome Productivity, which includes other fluxes such as fire and harvest disturbance and is a land model’s attempt at the complete carbon cycle (it would also be helpful to include lateral fluxes if possible). I realize the participating models represent different aspects of the carbon cycle, which may make comparisons more challenging, but the manuscript essentially equates NEP to net carbon balance, which is incorrect (although the difference may be minor). It may be worth adding Supplementary analyses to address this: does the inclusion of these other fluxes change the results shown, even for a subset of the models? Furthermore, is direct comparison of CMIP6 NEP to values from the Global Carbon Project valid given these differences? Does that partially account for the large differences shown between the two in Figure 1? The authors do not provide details on how they made this comparison, so it’s not possible for the reader to tell.

Personally, I also believe a manuscript such as this that discusses future carbon budgets would benefit from a historical comparison to data. This may be beyond the scope, but how should readers interpret the large differences between CMIP6 models and the GCP, and what does this ultimately mean for the future projections? Do models that tend to represent historical properties better (e.g., upscaled carbon fluxes from FLUXCOM, global biomass, LAI, etc.) tend to project higher or lower carbon sinks? These are important questions, some of which could be addressed in the Discussion and bringing in past literature. Along similar lines, the Discussion could benefit from more context and interpretation. Why haven’t uncertainties been reduced throughout CMIP versions, how would including permafrost carbon and disturbances change the assessment, what are the major recommendations on ways forward considering this?

Why did the authors analyze land surface temperature as opposed to a property such as 2 m air temperature, which is much more commonly used as a metric and benchmark? Land surface temperature accounts for not only the climate changes but also land surface responses, and in that way seems to add unnecessary complexity.

Finally, the manuscript would benefit from a thorough grammar check throughout. I addressed a minor fraction of grammatical errors in my comments below, but many more remain.
**Technical corrections**

L 16: What's the meaning of the word 'extent' here - spatial distribution, magnitude, or other?

L 20: do you mean to say CMIP5 here?

L20: NHL was defined and used previously as a plural noun, but here singular. Please remain consistent.

- Maybe define what domain NHL is referring to in the abstract?

L 32: The land carbon sink changes by a large amount interannually owing to annual climate oscillations, disturbances, etc.

L 45: this reference is almost a decade old now, and the literature it cites is over a decade; consider adding newer references for warming rates

L 76: What does the word 'devoted' mean here?

L 77: missing parenthesis

L 86: Can the authors expand on the 'newly updated data' they're referring to?

L 114: Incorrect use of “i.e.”. Could say “...in this study, including...”

L 126: Did the authors account for non-land fractions of grid cells in their area-weighting?

L 138. The readers would benefit from more details on how these sensitivity analyses were conducted
Figure 1: 2095-2100 is a short period to use to calculate standard deviations. More typical would be something like a 20 year time period.

L 189-190: Change 'huge' to something like 'large'.

Fig 1: Difficult to see the GCP values with all the light blue bars. Possible to change that into a shaded time series as well.

L 273: ‘Minimization’? Do the authors mean mineralization?

L 281: What does the word ‘special’ mean here?

L 287: The carbon balance will also significantly be impacted by disturbances, mentioned in the introduction but mostly not included in the CMIP6 models. This point should be emphasized.

L 291: Particularly poor grammar

L 302: “plant functional types”

L 305: What is meant by ‘compensation’ here?

L 332: Hyphen after Northern not necessary