

Investigating Carbon and Nitrogen Conservation in Reported CMIP6 Earth System Model Data

Gang Tang, Zebedee Nicholls, Chris Jones, Thomas Gasser, Alexander Norton, Tilo Ziehn, Alejandro Romero-Prieto, Malte Meinshausen

The authors examine mass conservation in carbon and nitrogen data reported by multiple Earth System Models involved in CMIP6. Their analysis reveals substantial accumulated mass imbalances, potentially reaching hundreds of gigatons of carbon, which introduces significant uncertainty into the models' results and conclusions. The authors attribute these imbalances to missing fluxes in reported data and inconsistent definitions. Consequently, they recommend that future CMIP reporting protocols incorporate mass conservation checks in their validation processes, include secondary variables to facilitate mass balance calculations, and standardize definitions and variable names to reduce complexity. This topic is of crucial importance, as CMIP model results are widely used to understand the dynamics of numerous variables and their effects on climate change and variability. Moreover, when evaluating the diverse and sometimes contradictory results from various models, having additional data to assess model reliability would be invaluable in determining which models are more trustworthy. While the analysis and recommendations are highly relevant, the manuscript's presentation could be improved for better comprehension. To enhance readability, I recommend simplifying the narrative by eliminating repetitive paragraphs and reducing the use of parentheses. Additionally, it would be beneficial to define variables and scenarios clearly at the outset and streamline the manuscript's organization by reducing the number of subsections. These changes would make the document more accessible and easier to follow, facilitating a better understanding of the analyses and results.

Major comments:

- I found the manuscript somewhat challenging to follow. I recommend adopting a more direct writing style and avoiding repetitive information. The presence of lengthy sentences and excessive parenthetical information complicates comprehension. Additionally, I believe the manuscript could be significantly shortened without losing essential content.
- I noticed an abundance of sections and subsections that could be streamlined by merging some of them. This would help eliminate repetitive content and enhance the manuscript's overall clarity and flow.
- The manuscript would be more accessible if all procedures and variable definitions were presented clearly from the outset. I found that the authors

provide information gradually, which adds unnecessary complexity to the manuscript.

Minor comments:

- Please replace CO2 with CO₂ throughout the manuscript.

Abstract

- Could the authors clarify what they mean by "Given that CMIP6 data is no longer being reported"?

Introduction

- L37. Please add a space before the references.
- L39-40. References are enclosed in double parentheses.
- L49. Please clarify the specific role being referenced and provide more detailed information about its significance.
- L75. What about sections 2 and 3? Please address these sections as well.
- L82. Consider removing the word "including" as there are no suggestions provided for other stakeholders.
- L85-93. Consider eliminating these reasons here, as they are addressed later in the text.

Section 2

- L101. Consider providing a brief description of the experiments.
- L103. Consider referencing a table that lists the names of the models instead of enumerating them here.
- L107. Consider clarifying what "variant_label" refers to.
- L112-114. Please rewrite this statement clearly.
- L118-119. What do the authors mean by "the monthly global pool sizes/fluxes were weighted using the model-specific calendar to calculate their annual mean"?

Section 3

- L129. What do the authors mean by "requires consideration". Please provide more specific details.
- L140. I am unclear about why the numerical errors are expected to be minimal based on the information provided in the previous sentence.

Section 4

- Eqs 2-10. Please provide definitions for each variable.
- L165. Please clarify the meaning of "The nbp and the dcLand*/dt exhibit the exact relationship as shown in Eq. 2."
- L166. There are not many details about this in Section 3, as mentioned.
- L186. Which results are you referring to? Please cross-reference the figures for each case.
- L189. Please provide a brief explanation of why it suggests that the differences are solely due to processing issues.

- L199. During which period?
- L210-211. Why does that indicate unavoidable numerical errors?
- L140. What do the authors mean by “slightly adjust the calculation of the equations mentioned earlier”? Additionally, please cross-reference those equations using their respective numbers instead of referring to them as “above.”

Section 5

- L269. Why is it stated as “should have”?
- L280. Are there differences between the reported nLand and the calculated nLand*?
- L281. There is a typo in “Notably”.
- L324. Which experimental periods are you referring to?
- L339. The value of the orders is missing.

Section 6

- Title. Consider using a more descriptive title.
- L379. Which composite variables should be removed?

Section 7

- Section 7.2. Consider highlighting the most relevant models, specifically those that demonstrate a “better conservation of mass.”
- Table 1. Consider adding the values of imbalance, with distinctions indicating whether they are lower or greater than 5 Gt, along with the effect of the variable.
- L461. Why use the term “except”? Do you mean “besides”?
- L486. What kind of complexities are you referring to? Please be more specific.

Section 8

- L539-540. Consider rewriting this paragraph in a more friendly tone. In my opinion, as it stands, it seems there is little value in sharing the data.
- L544. Please add the reference for this claim.
- L66-666. I found this important; please consider including it in the main text.

Figures

- Fig. 1. This figure is frequently cited throughout the manuscript, but it is not adequately explained anywhere.
- Fig. 2. Why is “net biosphere production” placed in quotes in the legend?
- Fig 2. Consider plotting $dcLand/dt$ in a different manner to facilitate easier differentiation.
- Figs. 5, 6, A7, A8, and A11. Please remove the CanESM models from these figures.
- Fig. A4. Please extend the y-axis to fully display the negative values.
- Fig. A6 and A9. Consider using the same colors and line types for the models in both the upper and lower panels, and include a consistent and unique legend for all graphs.