Responses to Referee #1 :

This is an interesting study in which the authors use a suite of ESMs to evaluate future changes in NPP and their associated drivers in the EBUSs. The study highlights that the traditionally assumed key mechanism, upwelling-favorable winds, can only explain about 25% of NPP changes in the EBUS, while the remaining three-quarters are influenced by diverse drivers. However, I do not recommend this manuscript for publication at this stage, as major revision is required. In my view, the current analysis of the relationship between upwelling-favorable winds and future NPP changes is somewhat simplistic and would benefit from a deeper investigation of model biases, inter-model differences, and regional variability. In addition, the discussion of the "diverse drivers" is uneven: analyses of physical processes such as geostrophic transport and wind-stress curl are extremely limited, whereas stratification and biogeochemical control of subsurface nutrient reservoirs are discussed in much greater detail. I therefore suggest that the authors thoroughly reconsider and improve the structure and balance of the manuscript.

We thank the reviewer for the valuable feedback and have made significant revisions to improve the manuscript. We have conducted a more thorough investigation into inter-model differences by explicitly evaluating individual model behavior in both historical simulations and future projections, including a specific assessment of the CanESM5 model. We expanded the Results section to better describe the sources of NPP uncertainty at the EBUS scale and highlight the role of spatial heterogeneity. Moreover, we have reorganized the Discussion to strengthen the geostrophic transport analysis. Finally, we rearranged the Introduction and figure labeling to enhance overall readability and clarity.
Below are the authors' point-by-point responses to the comments, with reviewer's comments in bold and author's responses in normal font.

Major Comments

1. Rearrangement of the Introduction section. In L53, the authors state "Alongside projected changes in vertical velocity," which may not be fully appropriate, as both changes in vertical velocity and in nutrient supply to the surface are discussed previously. Moreover, the subsequent paragraph focuses on the timing of the upwelling season, which could also be classified as changes in vertical velocity. I therefore suggest reorganizing the Introduction between L32-58. Specifically, the authors could first introduce the projected changes in different mechanisms that modulate the upwelling intensity and duration, and then present additional sources influencing nutrient supply to the surface beyond the physical drivers. L59 will follow just after this, summarizing how variations in upwelling itself and changes in biogeochemical properties of source waters together modulate nutrient supply in EBUSs. such a reorganization could improve the logical flow of the intro.

We agree, we will modify the manuscript according to your indications.

2. As the manuscript highlights regional differences in Fig. 3, additional evaluation would be valuable. For example, as shown in Fig. A2, most models project a weak positive change in the CalCS, while fewer models project a strong negative change, particularly CanESM5. It would be useful to clarify whether CanESM5 exhibits mean NPP anomalies comparable to those of other models, or whether there are substantial spatial differences. Moreover, could these differences be related to model biases? For instance,

does CanESM5 already exhibit biases in simulating NPP in the CalCS during the historical period? In addition, the model agreement showed in Fig. 3 raises further questions. In the CalCS, model agreement is weak and the projections remain highly variable, which is acceptable; and in the CanCS, model agreement is also modest, yet the projections show relatively high consensus.
The high uncertainty in the BCS can be explained by Fig. A3, where the poleward and equatorward regions show relatively consistent changes but in opposite directions. I believe adding this information to this section would be better. Therefore, the diversity of NPP projections warrants further evaluation to identify and discuss the potential sources of uncertainty that lead to this strong sensitivity to model choice, such as model biases, inter-model differences, or spatial heterogeneity.

We agree that a deeper investigation of inter-model differences, spatial patterns, and potential model biases would be valuable. While a model-by-model evaluation of biases would be beyond the scope of the manuscript, the goal of which is to assess consensus in NPP projections and drivers across CMIP6 models, we have taken several steps to strengthen the analysis in this respect:

- We have expanded the Model Evaluation section and moved it to the Results section . We have assessed the ability of CMIP6 models to reproduce both the historical mean state and trends in NPP, using five observationally based products from Ryan-Keogh et al. (2023). This will provide a more systematic evaluation of model performance;
- To better illustrate the diversity of projections, we propose to include an additional figure in the Supplementary Information showing time series for each individual model, rather than only intermodel spread (Figure A of this document). This will make the number of models projecting increases versus decreases more explicit. In addition, we are considering moving the table currently shown in Fig. A2 into the main text to better highlight individual model behavior and inter-model differences;
- In response to your specific question on CanESM5, we have analyzed NPP anomalies in the CalCS for CanESM5 and compared them with the CMIP6 ensemble mean. We find that the spatial pattern of anomalies is broadly consistent with the ensemble mean, although the magnitude of both positive and negative anomalies is larger in CanESM5 (Fig. B of this document).
- Regarding the model agreement in the CanCS, we argue the signal is more robust than it first appears. While strict thresholds suggest lower consensus, the projected changes in CalCS-P align with the consistent, large-scale decreasing patterns observed across CMIP6. Furthermore, sensitivity tests indicate that lowering the agreement threshold to 70% yields strong consensus, and at 60%, nearly the entire CanCS region exhibits significant robustness (indicated by stippling on Fig. C of this document).
- We agree that the contrasting responses between poleward and equatorward regions, as shown in Fig. A3, help explain the high uncertainty in the BenCS. We will improve the clarity of section 3.1 to better highlight the role of spatial heterogeneity in driving inter-model spread.

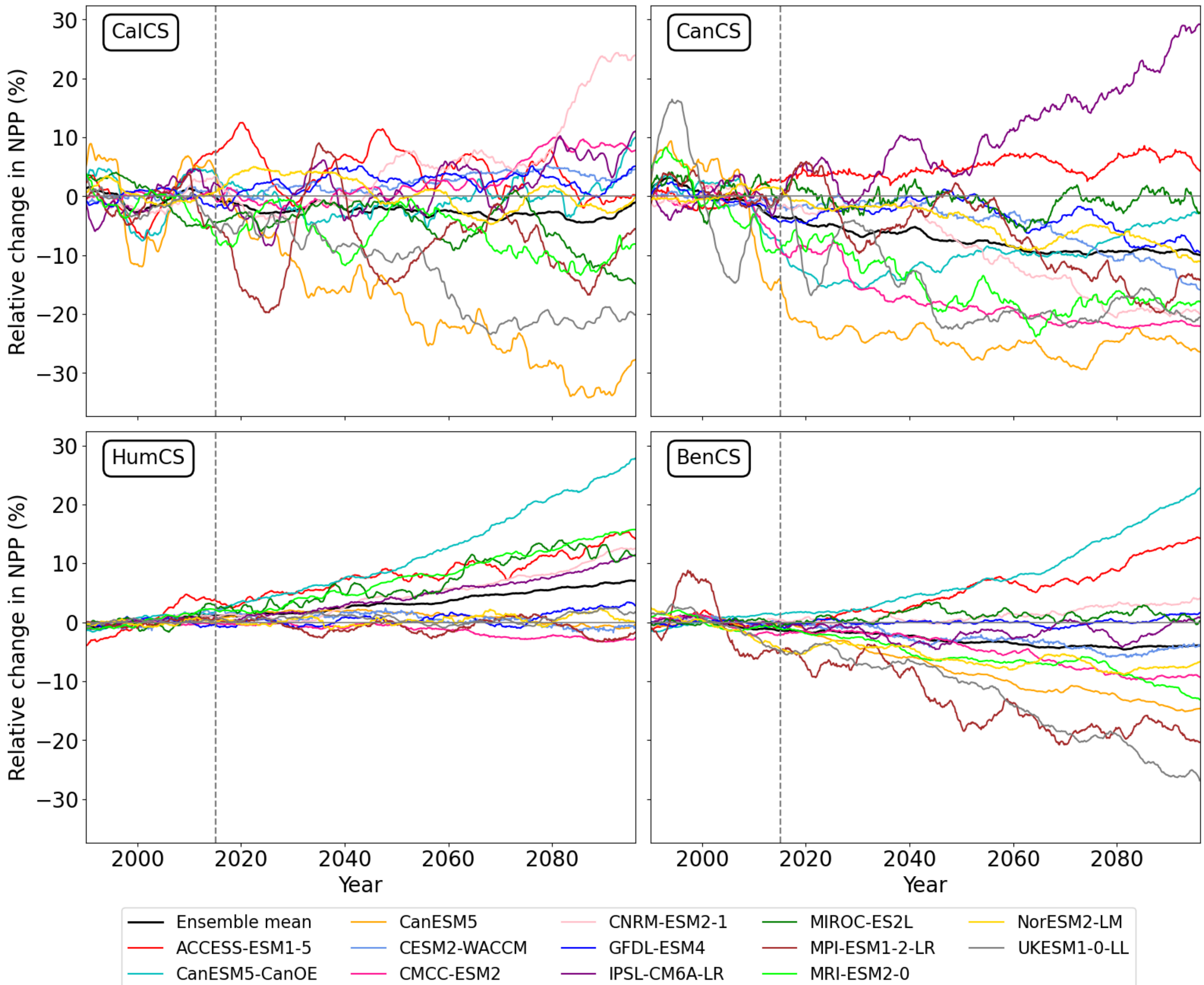


Fig. A) Historical and SSP5-8.5 projected time-series of relative change in NPP for CMIP6 models in the EBUSs.

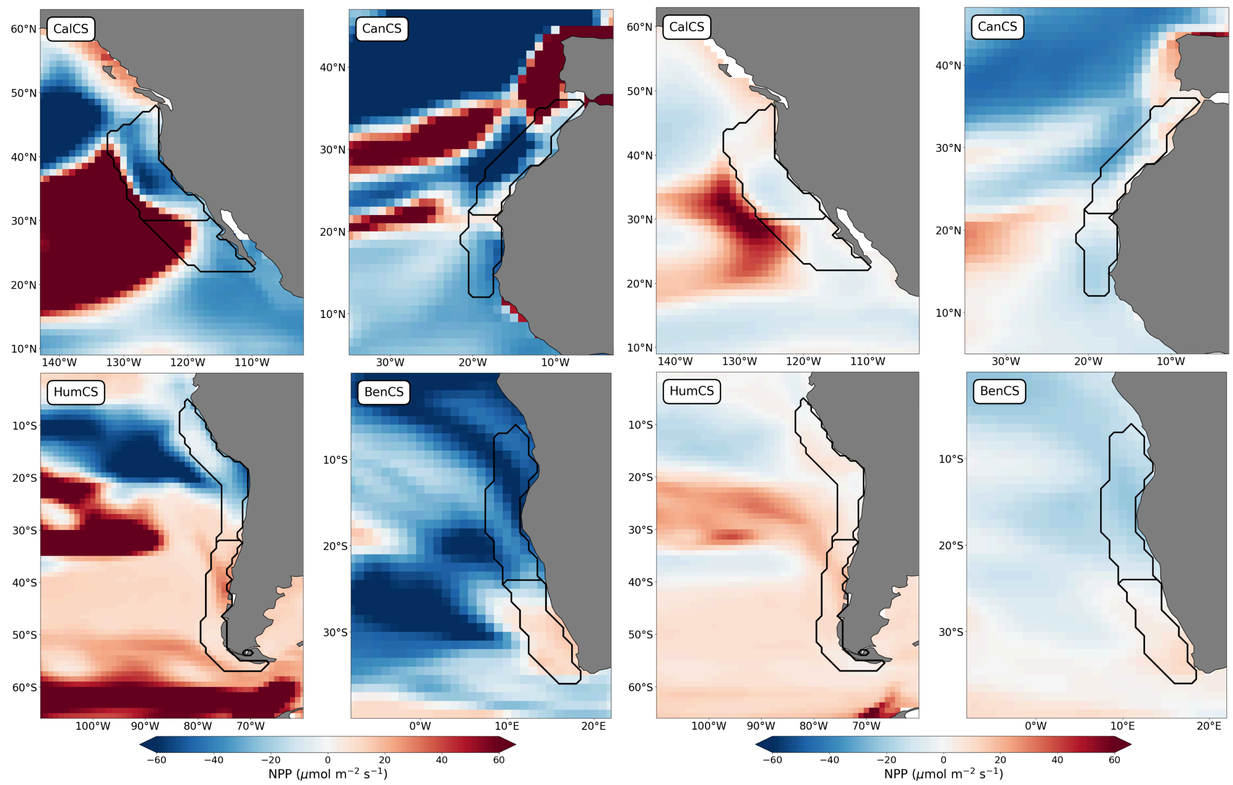
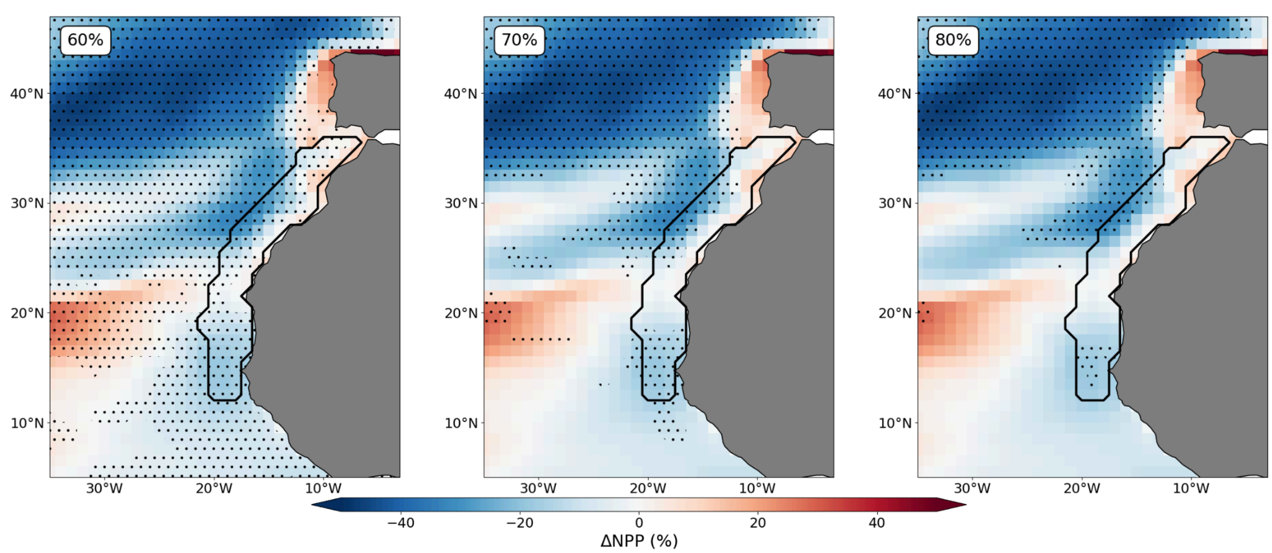


Fig. B) CanESM5 (left) vs. ensemble mean (right, 13 models) NPP anomalies (2071–2100 relative to 1985–2014) under SSP5-8.5 in the four eastern boundary upwelling systems.



Fic. C) Ensemble mean (13 models) NPP anomalies (2071–2100 relative to 1985–2014) under SSP5-8.5 in the CanCS. Areas of high model agreement are indicated by stippling, based on different thresholds of agreement on the sign of change: at least 60% (left), 70% (middle), and 80% (right).

3. Some rearrangement of the Discussion section is recommended. I found that Section 4.1 is somewhat repetitive, covering material similar to the Introduction. Some of the content could be interspersed into the preceding text. This section also does not present any additional analysis. I therefore suggest deleting Section 4.1 and incorporating any necessary statements into Sections 3 and 5. Section 4.2 is somewhat unclear and abrupt. It appears to discuss the mismatch between wo60 and tau, possibly attributing it to contributions from wind stress curl, but geostrophic transport is also mentioned without clear connection. I suggest integrating the discussion of wind stress curl, together with Figure A10, into Section 3.2 for a clearer and more coherent presentation. Additionally, the role of geostrophic transport requires further explanation.

We will revise the Discussion section to improve both clarity and completeness following your suggestions. In particular:

- We will remove Section 4.1 and incorporate its relevant content into Sections 3 and 5;
- We will expand Section 4.2 to improve its integration within the overall analysis. Specifically:
    - We will provide a more detailed explanation of the processes already illustrated in Fig. A12 (Appendix), the current description is perhaps too concise.
    - To strengthen the analysis, we will include an additional "four-color" EBUS map (similar to Fig. 4 in the manuscript), showing the relationships between projected anomalies in Ekman and geostrophic transports (Fig. D of this document). This additional analysis shows that in the central HumCS and BenCS-P regions, Ekman and geostrophic transports tend to vary in opposite directions, likely contributing to the weak correspondence between wind changes and vertical velocity in these regions. This helps explain the projected changes in wo_{60} in the BenCS-P region, but it provides limited additional insight for the HumCS, where vertical velocity remains highly variable.

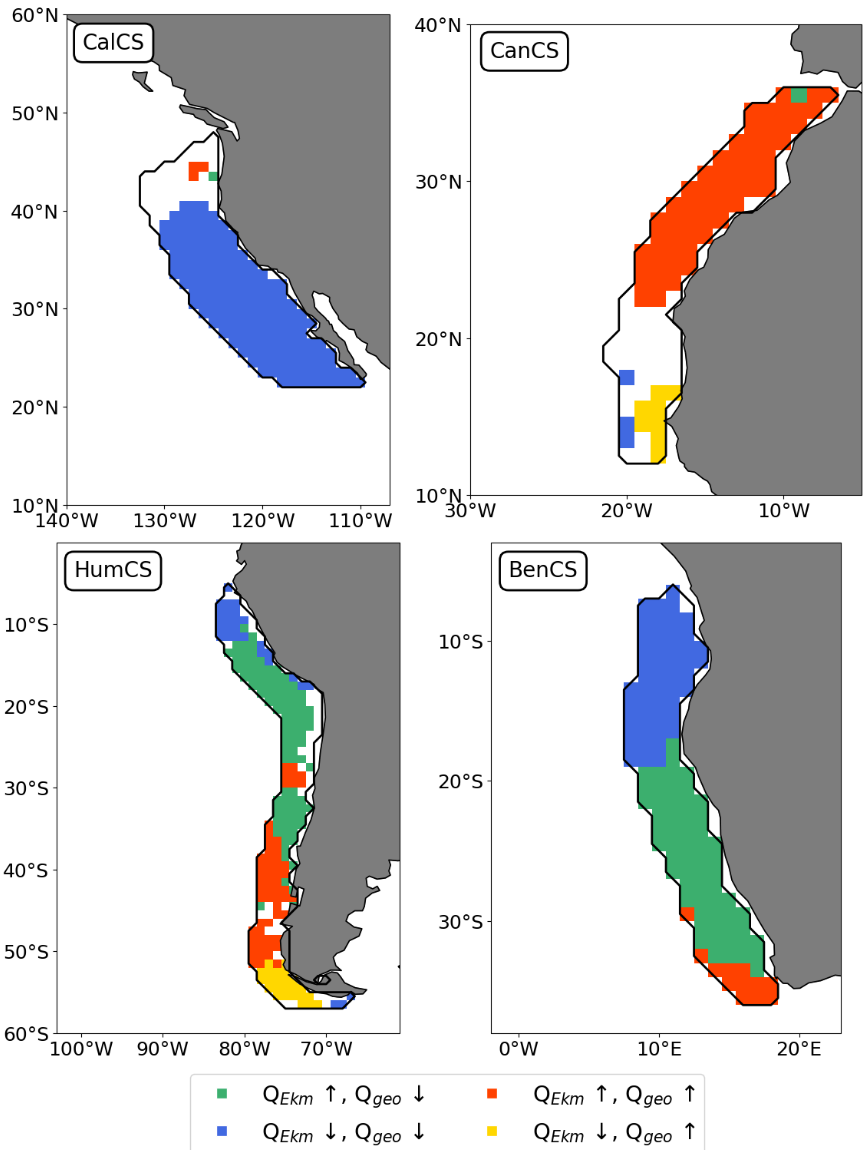


Fig. D) Ekman and Geostrophic transport changes are reinforcing each other in some regions (blue and red), offsetting in others (green and yellow). Signs of twenty-first century colocalized anomalies in Ekman and Geostrophic transport under SSP5-8.5 (2071-2100). Grid cells are colored if there is ≥ 50 % model agreement.

4. The manuscript includes many different regions and numerous subplots that are frequently cross-referenced in the text. However, in several places the relevant figures or panels are not explicitly specified, making it difficult for readers to follow. I recommend labeling multi-panel figures with clear identifiers (such as Fig. 5; e.g., a, b, c, d) and explicitly referring to these labels in the text to improve clarity and readability.

Yes, we will modify the manuscript.

Detail Comments:

1. There are some errors in the line numbers on page 6 and 7. A large section of text between lines 115 and 120 is missing line numbers.

Yes, we will modify the manuscript.

2. L25-29 should include appropriate references not only for the "other physical processes", but also for the broader statements made in this part of the text.

Yes, we will modify the manuscript.

3. In L67-74, the influence of decadal and multidecadal climate variability on NPP should be introduced in greater depth to reinforce the importance of the limited duration of the observational records.

Yes, we will modify the manuscript.

4. L105: The authors use the variables "tauu/tauv" to represent the surface wind stress. These variables originate from the atmospheric model on its native grid, which differs from the ocean model grid. Could the authors please clarify whether these fields were first interpolated onto the ocean grid and then regridded to the common 1-degree resolution, or whether they were directly regridded from the atmospheric grid to the 1-degree grid? This distinction is important because wind stress is interpolated to the ocean grid to drive coastal upwelling, and CMIP also provides the "tauuo/tauvo" variables on the ocean grid.

In our analysis, we used the tauu/tauv as direct model outputs from the atmospheric grid, and we have directly regridded to the 1-degree resolution grid. To assess the robustness of this choice, we compared results obtained using tauu/tauv with those derived from the ocean-grid variables tauuo/tauvo, and we found generally very small differences. In particular, the spatial patterns, as well as the sign and magnitude of projected anomalies, are highly consistent between the two datasets. Thus, we opted to use tauu/tauv, as these variables were for us more readily available across the full ensemble considered.

5. What is the relative width (200 km?) of the LME mask?

Masks identifying these domains were obtained from the ISIMIP repository, and have variable extents in longitude (widths from 4 to 9°) and latitude (lengths from 25 to 50°).

6. L157 Stating the exact number of models projecting positive versus negative changes would provide useful quantitative information.

Yes, we will modify the manuscript.

7. L207-208 Please specify which figure is being referred to, the second left column in Fig 5?

We are referring to the second column of Fig.5, we will modify the manuscript to clarify better.

8. L272 "While nitrates are projected to increase around 200 m in this region (Appendix Fig. A15), these waters do not reach the typical source depth (60 m)" can this be attributed to enhanced stratification?

Yes this could be an explanation, we will add this hypothesis in the manuscript.

<u>References</u>:

Ryan-Keogh, T. J., Thomalla, S. J., Chang, N., & Moalusi, T. (2023). A new global oceanic multi-model net primary productivity data product. Earth System Science Data, 15(11), 4829-4848.