

Dear Prof. Chien,

We sincerely thank you for taking the time to thoroughly review our manuscript and for providing such insightful and constructive comments. Your valuable suggestions have significantly helped us refine the physical mechanisms, clarify our regional classification criteria, and improve the overall clarity and rigor of the manuscript.

Below, we have addressed all of your comments point-by-point and outlined the corresponding revisions made to the manuscript.

Prof. Chien's comments are formatted in italics, and the authors' responses are presented in regular font.

Notation guide: "RC1.P#" refers to Reviewer 1's Comment, Paragraph #.

[RC1.P1] *Some statements remain too strong relative to the evidence presented. For example, the claim that POP2-waves shows "superior performance and statistical robustness" is too strong, given that you also acknowledge regions where the wave-coupled model does not outperform B-CTL, both in Sect. 3.1 and in the conclusion. The text should use more cautious language, such as "improves performance in most selected regions," rather than implying universal improvement. Another example is the phrase "pervasive and substantial reduction in systematic biases" on line 336.*

Response:

We completely agree with the reviewer that absolute terms like "superior performance" and "statistical robustness" overestimate the results, especially given the noted regional exceptions (e.g., the Northwest Pacific, Equatorial Atlantic, and specific mid-latitude sub-basins) that we identified in our text. We appreciate this insightful note and have thoroughly revised the manuscript (including the Abstract, Section 3.1, and the Conclusions) to ensure our tone remains strictly objective, balanced, and aligned with the presented data.

Specifically, we have removed all overstating language (such as "superior performance", "statistical robustness", "pervasive", and "directly amplify") and replaced them with more measured, regionally constrained expressions.

[RC1.P2] *In addition, the conclusion states that "neglecting this coupling by treating $dpCO_2$ and wave-induced ($k_{w,660}$) independently leads to a substantial overestimation of global fluxes." However, the manuscript later states that "POP2-waves enhances regional oceanic CO_2 uptake and outgassing by 11.8% and 41.6%, respectively, yielding a 1.8% increase in the global net ocean CO_2 sink compared to the B-CTL simulation." A 1.8% increase does not seem to justify the term "substantial" for the global net flux. Moreover, if POP2-waves leads to an increase relative to B-CTL, then the wording should be checked carefully, as "underestimation" rather than "overestimation" may be more appropriate depending on the specific comparison being made.*

Response:

We thank the reviewer for pointing out this potential confusion. We would like to clarify that these two statements refer to two completely distinct mechanisms and comparisons:

1. The "substantial overestimation" regarding independent treatment:

This statement addresses a methodological comparison. When calculating the air–sea CO₂ flux using independent datasets (e.g., combining observational or offline $\Delta p\text{CO}_2$ with wave-induced $k_{w,660}$ without considering their concurrent covariance), the mathematical neglect of their negative correlation leads to a substantial overestimation of the calculated global flux. This is a common pitfall when cross-multiplying averaged or independent fields of gas transfer velocity and partial pressure gradients.

2. The 1.8% increase in global net sink in POP2–waves vs. B–CTL:

This statement refers to the actual dynamic simulation results between our coupled model and the uncoupled control run. The incorporation of wave effects via the Deike and Melville (2018) parameterization inherently enhances $k_{w,660}$ under high wind conditions, which drives up both regional uptake and outgassing.

However, because our model applies an online coupling, a strong $\Delta p\text{CO}_2$ -driven negative feedback mechanism is activated: as the gas transfer enhances, the local $\Delta p\text{CO}_2$ gradient is rapidly depleted, which in turn suppresses further air–sea CO₂ flux exchange. It is precisely this negative feedback that dampens the global net increase to a modest 1.8%, preventing the runaway "overestimation" that would occur if the two terms were treated independently.

To avoid any ambiguity, we will revise the wording in the conclusion of the revised manuscript to clearly distinguish between the methodological overestimation and the simulated model sensitivities.

[RC1.P3] *The criteria used to assign regions to Figs. 3 and 4 are also not fully clear. In particular, the exact threshold or objective rule for the classification described on lines 315–316 should be clarified.*

Response:

We entirely agree with the reviewer that the classification criteria should be strictly objective and transparent. We apologize for the lack of explicit threshold values in the original draft. To resolve this ambiguity, we have redefined the sorting

mechanism using explicit statistical thresholds in the revised manuscript (Lines 315–316). Following this rigorous re-evaluation, the South Atlantic Ocean (SAO) region, which exhibits a correlation of $r = 0.33$, has been correctly reallocated from the highly correlated group to the low-correlation group. This adjustment changes the respective group sizes to 8 and 4 regions. The explicit criteria applied in the revised draft are as follows:

1. Seasonal phasing consistency (temporal alignment): Regulated by the Pearson correlation coefficient (r) between the simulated monthly flux cycles and NOAA CT2022. Regions assigned to Fig. 3 exhibit consistent seasonal phasing with significantly positive correlations ($r \geq 0.50$, now totaling 8 regions). Conversely, regions in Fig. 4 display prominent phase mismatches, weak correlations, or anti-correlated trends ($r < 0.50$ or statistically non-significant r , now totaling 4 regions, including the SAO).
2. Relative magnitudes of model–data biases (amplitude deviancy): Defined as the monthly averaged absolute deviation normalized by the overall peak-to-peak seasonal range (95th minus 5th percentile) of the NOAA CT2022 baseline. A threshold of 25% was applied to further distinguish regions where structural amplitude offsets dominate over temporal phase alignment.

These definitions have been fully integrated into the revised text to ensure reproducibility.

[RC1.P4] *The figure quality and consistency still need improvement. Several figures remain difficult to read at manuscript scale, especially the multi-panel time-series and map figures. Font sizes, legends, color bars, contour labels, and panel titles should be enlarged and standardized. For example:*

Fig. 2, Fig. 7, Fig. 8, Fig. 9, and Fig. 10: The figure panel titles are generally too small.

Fig. 2: Why are the numbers in the color keys for panels a–d bold, whereas those for panels e–f are not?

Fig. 7: There is a similar issue as in Fig. 2. The color-key numbers are bold only for panel b.

Fig. 10: The panel title is bold only for panel a.

Response:

Thank you for this helpful suggestion. We have increased the font size of the panel titles in Figs. 2, 7, 8, 9, and 10 to improve readability. Additionally, we rearranged Figs. 3 and 4 into a 2-column layout (2 panels per row) so that the figure

details are clearer. To maintain visual consistency across all figures, we also bolded the axes and legend text in Fig. 6.

Thank you for pointing out these formatting inconsistencies. We have thoroughly checked and revised Figs. 2, 7, and 10 in the revised manuscript. All color-key numbers across all panels are now consistently formatted in bold. Additionally, we have unified the font sizes and styles of all panel titles to ensure complete visual consistency throughout these figures.

[RC1.P5] *The English still requires careful editing. There are many remaining grammatical errors, incomplete sentence structures, punctuation problems, and awkward expressions throughout the manuscript. I have listed some examples below, but the authors should carefully check the entire manuscript rather than correcting only these specific cases.*

Line 47: “Additionally,,” → “Additionally,”

Line 241: What does “t” mean in “and t sea-ice fractions”?

Line 271: “to providing” → “to provide”

Response:

We sincerely apologize for the grammatical oversights, punctuation errors, and awkward phrasing in the original draft. We highly appreciate the reviewer's careful reading and constructive feedback.

[RC1.P6] *Line 248: Please check the grammar around “constrains.”*

Response:

We appreciate the reviewer's careful eye. We realize that the phrasing in the original draft contained a grammatical/typographical error and that the verb usage could be further refined to better reflect the atmospheric inversion process.

In the revised manuscript, we have rephrased this sentence to improve grammatical correctness and replaced the verb with a more accurate and standard terminology (“optimizes”). The sentence now reads as follows:

“Unlike our forward, ocean-centric modeling frameworks, CT2022 optimizes the system from an atmospheric perspective, utilizing a top-down approach to infer net surface CO₂ fluxes by assimilating observed atmospheric CO₂ mole fractions.”

[RC1.P7] *Lines 296–297: Why are EEP and NEP underlined? Please also spell out these abbreviations.*

Response:

We thank the reviewer for identifying this formatting issue. In the original draft, the underlining was intended to visually highlight these two key regions on the corresponding figure. However, we agree that this formatting is unconventional and lacks clarity.

In the revised manuscript, we have removed the underlines, fully spelled out both abbreviations upon their first appearance, and rephrased the sentence for professional clarity.

EEP has been spelled out as Equatorial Eastern Pacific, and NEP has been spelled out as Northeastern Pacific.

[RC1.P8] Line 561: “Several” appears to be in a different font.

Response:

Thank you for bringing this formatting issue to our attention. We apologize for the inconsistent typography. In the revised manuscript, the font for the word “Several” has been corrected and fully standardized to match the rest of the text.

[RC1.P9] Line 593: Please check the use of “constrains.”

Response:

We apologize for the grammatical error (subject-verb misalignment) and the awkward wording in the original text.

In the revised manuscript, we have corrected this sentence by replacing the improper form with a more mathematically precise verb (confines), and rephrased the passage for clarity.

[RC1.P10] Lines 718–719: “this wave-induced divergence results in a systematically air–sea CO₂ flux under 10-m wind speed and significant wave height.” This sentence is unclear. Do you mean “a systematically enhanced air–sea CO₂ flux under conditions of high 10-m wind speed and significant wave height”?

Response:

Yes, the reviewer's interpretation is completely correct and precisely captures our intended meaning. We apologize for the awkward phrasing and the missing word in the original draft.

The text has been updated as follows:

“This wave-induced divergence results in a systematically enhanced air–sea CO₂ flux under high 10-m wind speed and elevated significant wave height.”

[RC1.P11] *Line 721: “exceed” → “exceeds”*
Line 723: “elevate” → “elevated”
Line 729: Please add a space between “approximately” and the plus-minus sign.
Line 747: “framework ,” → “framework,”

Response:

We thank the reviewer for the meticulous attention to these technical and formatting details. We apologize for these grammatical and typographical oversights. Following your precise suggestions, we have corrected all of these listed issues in the revised manuscript (including fixing the subject-verb agreements, adding the missing space, and cleaning up the punctuation typo). The entire section has been carefully re-vetted to ensure complete consistency.

[RC1.P12] *Line 1327: “version 4.0:” Please check why a colon is used here instead of a semicolon, as in the surrounding entries.*

Response:

We appreciate the reviewer's meticulous check. The use of a colon was an oversight that disrupted the formatting consistency of our bibliography/entries. In the revised manuscript, we have corrected this by replacing the colon with a semicolon (“version 4.0;”) to match the surrounding entries perfectly.

[RC1.P13] *In addition, as I noted at the beginning of the review process, please include a justification in the manuscript for why an older version of CESM, CESM1.2.2, was used instead of CESM2, which includes important infrastructure improvements.*

Response:

In this study, CESM version 1.2.2 was employed as the modeling framework. This choice builds upon our previous successful integration and validation of a high-resolution one-dimensional ocean model (SIT) coupled with CAM5 (namely CAM5-SIT) under the CESM1.2.2 infrastructure (e.g., Lan et al., 2022, 2024), providing a consistent and well-tested baseline for isolating air–sea flux dynamics. While newer model versions (e.g., CESM2) offer enhanced infrastructure, CESM1.2.2 remains fully capable of accurately representing the physical and biogeochemical processes essential to this study.

We hope these revisions successfully address all the concerns raised. Thank you again for your time and efforts in reviewing our manuscript.

Sincerely,

Yung-Yao Lan, Huang-Hsiung Hsu, Wei-Liang Lee and Simon Chou