

Peer Review File

Distinct sources of uncertainty in simulations of the ocean biological carbon pump at different depths



Open Access This file is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. In the cases where the authors are anonymous, such as is the case for the reports of anonymous peer reviewers, author attribution should be to 'Anonymous Referee' followed by a clear attribution to the source work. The images or other third party material in this file are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Web links to the author's journal account have been redacted from the decision letters as indicated to maintain confidentiality.

Decision letter and referee reports: first round

19th Jan 24

Dear Dr Wang,

Your manuscript titled "Organic matter transfer efficiency is the largest uncertainty for biological carbon pump in CMIP6 models" has now been seen by 3 reviewers, and we include their comments at the end of this message. They find your work of interest, but some important points are raised. In particular: 1) it is important that methods and terminology are clearly described so that readers know precisely what is meant when uncertainty is discussed, 2) the justification for, and limitations of, a Martin Curve approximation should be discussed, and 3) the models should be compared to observations to give readers context for the relative accuracy or inaccuracy of different modeling approaches. We are interested in the possibility of publishing your study in Communications Earth & Environment, but would like to consider your responses to these concerns and assess a revised manuscript before we make a final decision on publication.

We therefore invite you to revise and resubmit your manuscript, along with a point-by-point response that takes into account the points raised. Please highlight all changes in the manuscript text file.

We are committed to providing a fair and constructive peer-review process. Please don't hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the referees' comments (which should be in a separate document to any cover letter), a tracked-changes version of the manuscript (as a PDF file) and the completed checklist:

[redacted]

** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to receive your revised paper within six weeks; please let us know if you aren't able to submit it within this time so that we can discuss how best to proceed. If we don't hear from you, and the revision process takes significantly longer, we may close your file. In this event, we will still be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

Please do not hesitate to contact us if you have any questions or would like to discuss these revisions further. We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Best regards,

Michael Stukel, PhD
Editorial Board Member
Communications Earth & Environment
orcid.org/0000-0002-7696-6739

Clare Davis, PhD
Senior Editor
Communications Earth & Environment

EDITORIAL POLICIES AND FORMATTING

We ask that you ensure your manuscript complies with our editorial policies. Please ensure that the following formatting requirements are met, and any checklist relevant to your research is completed and uploaded as a Related Manuscript file type with the revised article.

Editorial Policy: [Policy requirements](#) (Download the link to your computer as a PDF.)

Furthermore, please align your manuscript with our format requirements, which are summarized on the following checklist:

[Communications Earth & Environment formatting checklist](#)

and also in our style and formatting guide [Communications Earth & Environment formatting guide](#) .

*** DATA: Communications Earth & Environment endorses the principles of the Enabling FAIR data project (<http://www.copdess.org/enabling-fair-data-project/>). We ask authors to make the data that support their conclusions available in permanent, publically accessible data repositories. (Please contact the editor if you are unable to make your data available).

All Communications Earth & Environment manuscripts must include a section titled "Data Availability" at the end of the Methods section or main text (if no Methods). More information on this policy, is available at <http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf>.

In particular, the Data availability statement should include:

- Unique identifiers (such as DOIs and hyperlinks for datasets in public repositories)
- Accession codes where appropriate
- If applicable, a statement regarding data available with restrictions
- If a dataset has a Digital Object Identifier (DOI) as its unique identifier, we strongly encourage including this in the Reference list and citing the dataset in the Data Availability Statement.

DATA SOURCES: All new data associated with the paper should be placed in a persistent repository where they can be freely and enduringly accessed. We recommend submitting the data to discipline-specific, community-recognized repositories, where possible and a list of recommended repositories is provided at <http://www.nature.com/sdata/policies/repositories>.

If a community resource is unavailable, data can be submitted to generalist repositories such as [figshare](#) or [Dryad Digital Repository](#). Please provide a unique identifier for the data (for example a DOI or a permanent URL) in the data availability statement, if possible. If the repository does not provide identifiers, we encourage authors to supply the search terms that will return the data. For data that have been obtained from publically available sources, please provide a URL and the specific data product name in the data availability statement. Data with a DOI should be further cited in the methods reference section.

Please refer to our data policies at <http://www.nature.com/authors/policies/availability.html>.

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

Review of Wang and Fennel, "Organic matter transfer efficiency is the largest uncertainty for biological carbon pump in CMIP6 models"

This manuscript analyses the suite of CMIP6 models that report data on particulate organic carbon export and 3-D fluxes to identify sources of uncertainty in the modern-day modelled biological carbon pump estimates. The manuscript presents an interesting approach to considering uncertainty and provides useful results to those working in the field of biogeochemical oceanography.

The approach to estimating structural uncertainty is not described in the Methods section. In equation 13 and 14, the TE_{correction} term isn't defined. This made it a bit confusing to understand the results.

The use of the Martin curve as the "true" attenuation curve against which to assess parameterisation uncertainty makes the assumption that the Martin curve correctly describes the flux attenuation. However, there are many possible mathematical forms to describe the flux attenuation curve (Cael & Bisson, 2018, *Frontiers Mar Sci*). Although I understand the rationale behind choosing the Martin curve, the authors need to include a discussion of the implications of this assumption for their results.

Please include a table in supplementary material with the key references describing the biogeochemical model structure for each of the CMIP6 models used in the analysis.

Section 2.1: I found this section quite heavy-going as there are lots of numbers being cited. The authors should consider whether the detailed info on the fluxes and multiplication factors for the different biomes can be tabulated (or graphically presented), thus leaving the text to describe the key features.

Line 124: how do these values of b compare to observations-based estimates?

Line 125: do none of the models used here have temperature-dependent remineralisation?

Line 142: it's not clear what "neglecting the protections by opal" means here.

Lin 162-163: It would be interesting to discuss the observation that the models used here have contrasting latitudinal patterns of transfer efficiency in the context of the ongoing discrepancy in observations-based global patterns of transfer efficiency (e.g. Henson et al., 2012 vs Marsay et al., 2015).

Line 167: include a reference for "oligotrophic regions tend to export fecal pellets and refractory organic particles". Also, as written this sentence implies that fecal pellets aren't exported at high latitudes.

Figure 5a: Is the x-axis percentage (/100), as used in lines 198-220? What do negative contributions to uncertainty mean? Define U_f , U_{ep} etc. in the caption.

Line 239-241: the new paper by Ricour (2023, *Nature Geosci*) may provide some useful information on sequestration horizon.

Figure 7: Define U_p , U_s etc. in the caption.

Line 289: "more similar" than what?

Section 2.5: this section was interesting, but rather an aside – it didn't seem to address the main point of the manuscript at all. I got to the end of the section and couldn't see how it was more than tangentially relevant. I suggest the authors rewrite this section. For example, what are the implications of the results to our understanding of future project changes to carbon storage? This section gives the impression that the only issue is paucity of observations and increasing the size of in situ datasets will make the models better, but there are also constraints on model parameterisation, complexity, structure etc. too that could be discussed.

Figure 8b: how are the match-ups between the models and observations done?

Line 339: why do the 14 models produce 2 contrasting patterns of transfer efficiency?

Conclusion section: the manuscript ends rather abruptly! What are the implications of the results shown for model development?

Reviewer #2 (Remarks to the Author):

The manuscript "Organic matter transfer efficiency is the largest uncertainty for biological carbon pump in CMIP6 models" by Wang and Fennel analyzes CMIP6 models for their ocean biogeochemical diversity. The manuscript provides a highly useful comparison of the CMIP6 models in their diversity in representing carbon export at 100 and transfer efficiency. However, I have three criticisms of the manuscript in its present form: 1) the authors are overly vague on their definition of "uncertainty", 2) the authors do not compare the models in a quantitative way to available estimates of transfer efficiency to assess whether any of the models are reproducing observed patterns (with the exception of the qualitative statement comparing to Guidi et al., 2015) and 3) the separation between "structural" and "Parametric" uncertainty is unclear. All these concerns are described below and should be easily rectified in a revised version.

7, 8, 9, and 10 - The general word "uncertainty" is used where I believe the more specific "model structural diversity" is intended. As the CMIP6 models are all derived from a similar set of scientific literature and often the same algorithms, they cannot be expected to span the true uncertainty range. As such, the analysis of CMIP6 models can only sample the realized structural diversity. On the flip side, it may be that some of this diversity can be falsified by observations such that the differences from observations should be characterized as "bias" rather than "uncertainty". However, given that the Hawkins and Sutton use of "structural uncertainty" has become so pervasive, I think it is fine to use "structural uncertainty" throughout as long as the authors add a contextualization statement on the above limitations.

9-10 – "but by uncertainty in" should be "and"

41-43 – The sentence justifying a distinction between "structural" and "parametric" uncertainty - "The former arises from the choice of model structure while the latter is due to the assignment of poorly known parameters and the varying degrees of parameter tuning across different ESMs." Is not compelling. I think I understand what the authors are getting at in terms of the difference in physics and chemistry between knowing the structural forms versus knowing the values of parameters, but in ocean biogeochemistry, such a clean distinction does not exist. There is typically no single "correct" structural form, just a bunch of semi-empirical options which must be calibrated and can give similar overall behaviors (e.g. Martin, versus Pace power functions, decreased remineralization timescale versus increased sinking speed, Monod versus Michaelis-Menten for growth.

48 – I think "be" should be "have remained"

146 - The authors mention Guidi et al., 2015 UVP transfer efficiency estimates as being supportive of GFDL CM4, but this is the only place where an independent verification is applied. The authors should conduct a quantitative assessment of models for transfer efficiency based on the available estimates including not only Guidi et al., 2015 but also:

Henson et al 2012 <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2011GB004099>

Weber et al., 2016 <https://www.pnas.org/doi/abs/10.1073/pnas.1604414113>

Devries and Weber, 2017 <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016GB005551>

I would note that while the Henson and Weber estimates have strongly divergent patterns, the Devries and Weber 2017 distinguish between the 1000m and 1000-2000 m transfer efficiency which illustrate very different regional patterns.

Figure 5 – UF, UEP, and UTE should be defined in the caption.

234 – Why is this true “especially in the North Atlantic and the North Pacific”? More clarification is necessary

241 – this assertion is incorrect – sequestration horizons are provided in Nowicki et al., 2022:

<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2021GB007083>

260 – It is not at all clear to me how the authors are able to distinguish between structural and parametric uncertainty. Looking through the methods, it seems that “structural” is being defined as the the Martin shape, and “parametric” as the deviation from that shape, or vice versa, but it seems that the term “TEdeviation” is defined towards the separation whereas “TEcorrection” is used in the definition of the structural uncertainty component, so it is not clear how all the terms are related. In any case, while I think it is potentially useful to frame the distinction between aspects of the transfer efficiency that are representable with the Martin curve in shape and its modulation through the b exponent versus that which is not, the authors need to explain what these definitions mean at the beginning of the section for it to be helpful.

294 – Bringing up the constraints from nutrient and oxygen data begs the question of why the authors did not include these comparisons with observations as a means of reducing uncertainty...

unfortunately, such an analysis is made extremely difficult as these models have radically different initialization/spinup procedures as detailed in Seferian et al., 2016

(<https://gmd.copernicus.org/articles/9/1827/2016/>) for CMIP5 and Seferian et al., 2020

<https://link.springer.com/article/10.1007/s40641-020-00160-0>) for CMIP6. The authors should point out that a more consistent protocol for model initialization/spinup would provide much better opportunity for model assessment in the future.

Reviewer #3 (Remarks to the Author):

Review of Wang & Fennel “Organic matter transfer efficiency is the largest uncertainty for biological carbon pump in CMIP6 models”

In this manuscript, Wang & Fennel analyze earth system model output from the latest CMIP simulations to identify processes that result in uncertainty in the strength of the biological pump. Above 900m, carbon export by sinking particles is the greatest source of uncertainty, whereas, below 900m, uncertainty is related to transfer efficiency, related to the rate of sinking particle remineralization. There are three groups of models, closely related to which biogeochemistry package is used: (1) with largely homogeneous transport efficiency (MARBL/HAMOCC) that is independent of temperature; (2) low transport efficiency in the subtropical gyres and high transport efficiency in the tropics and high latitudes (PISCES/COBALT) due to temperature dependence; and (3) high transport efficiency in the subtropical gyres and low transport efficiency in the tropics and high latitudes (BLING) due to ballasting. Either contrasting pattern from group (2) or (3) can be justified depending on which

observations are considered. Introducing variable reactivity distributions for sinking particles flips the PISCES model from (2) to (3), suggesting that distinguishing between different size and composition particles is important for simulating the future response of the biological pump to climate changes.

I thought this manuscript was interesting, well-written, and easy to follow. I only have a few suggestions:

(1) The authors make a prominent point of defining biomes for their analysis (Figure 1), but don't really apply them to the principal analysis. I think the discussion of results in section 2.1 would benefit from a box-and-whisker-plot for POC flux in each biome, and at each depth, with the location of each model indicated with a symbol. Similarly Figure 3. I am not really a fan of these "postage stamp" plots, I would prefer the authors did something more visual to highlight the patterns they are trying to emphasize, rather than making the reader scrutinize ~30 little panels each time.

(2) Which group does the CMCC-ESM2/BFM model fall in to? Line 128 indicates "no distinct latitudinal pattern" and it is dismissed, but it looks like there is something interesting going on between the Southern Ocean and other regions. Does BFM have any particular mechanisms in common, or different from the three overall groups?

Minor comments/clarifications:

Line 66-67: Cael & Lauderdale (2021) also found a ~20-25uatm sensitivity of changing "b" values in a biogeochemical model.

Line 70: Both export and transfer efficiency have structural and parametric uncertainty, the authors only make the distinction for transfer efficiency in this paper.

Line 131: Add "(high)" for "b" in the subtropical gyres.

Line 170-176: "PISCESv2 simulates two particle size classes with different sinking velocities": yes, but why did it not produce a similar result to GFLD-CM4/BLING or PISCESv2_RC?

Figure 4: did not calculate "b".

Line 240: Not exactly a global sequestration depth map (depends on how long you consider "sequestered"), but Hill et al. (2004, 10.1029/2002JC001598) have maps of residence times of carbon at different depths in the ocean from a model.

Line 278: "structural uncertainty can be significant" see also Lauderdale & Cael (2021).

Line 408 & 409: Define TE_correction. Is that related to TE_deviation?

We thank the reviewers for their constructive comments which we found very helpful while revising the manuscript. We considered all comments carefully and detail our responses and actions below. The complete reviewer comments are shown in black font followed by our responses to each comment in blue font.

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

Review of Wang and Fennel, “Organic matter transfer efficiency is the largest uncertainty for biological carbon pump in CMIP6 models”

This manuscript analyses the suite of CMIP6 models that report data on particulate organic carbon export and 3-D fluxes to identify sources of uncertainty in the modern-day modelled biological carbon pump estimates. The manuscript presents an interesting approach to considering uncertainty and provides useful results to those working in the field of biogeochemical oceanography.

Response: We would like to thank the reviewer for the positive assessment and constructive comments.

The approach to estimating structural uncertainty is not described in the Methods section. In equation 13 and 14, the TE_{correction} term isn't defined. This made it a bit confusing to understand the results.

Response: Sorry for the mistake caused by a mix-up of different manuscript versions. It should be $TE_{deviation}$ in equations 13 and 14:

$$U_{II} = var(\log_{10}(\overline{TE_{deviation}}))$$
$$U_{I\&II} = 2 \times var(\log_{10}(\overline{TE_{martin}}), \log_{10}(\overline{TE_{deviation}}))$$

In addition, we have re-defined the parametric (U_P) and structural (U_S) uncertainty as type I (U_I) and type II (U_{II}) uncertainty. Please see our response to the following comment for more details.

The use of the Martin curve as the “true” attenuation curve against which to assess parameterisation uncertainty makes the assumption that the Martin curve correctly describes the flux attenuation. However, there are many possible mathematical forms to describe the flux attenuation curve (Cael & Bisson, 2018, Frontiers Mar Sci). Although I understand the rationale behind choosing the Martin curve, the authors need to include a discussion of the implications of this assumption for their results.

Response: We appreciate this constructive and helpful comment. We agree with the reviewer that we cannot assume that the Martin curve or any other parameterization as the truth. In our revised manuscript, we re-defined the parametric (U_P) and structural (U_S) uncertainty as type I (U_I) and type II (U_{II}) uncertainty to distinguish between model uncertainty of transfer efficiency that is representable by the Martin curve by changing the exponent b (type I) and that which is not (type II). The motivation, definition, as well as interpretation and discussion of these two types of model uncertainty have been rephrased in our revised manuscript:

Introduction section (changed text is highlighted in bold font here):

“The BCP can be assessed by quantifying carbon export **out of** the euphotic zone and the depth-dependent transfer efficiency^{4–8}, **which measures the fraction of the exported carbon that survives respiration and reaches the deep ocean. Given the complex mechanisms controlling the transfer efficiency, simple parameterizations with different mechanistic underpinnings have been used to represent and to compare the transfer efficiency**^{9,10}. The Martin curve¹¹, **which normalizes the vertical carbon flux to a reference depth, e.g., 100 m, through the power-law equation $F=F_0 \times (z/z_0)^{-b}$, is the most widely used parameterization**^{12–16}. The exponent b is an indicator of the carbon flux attenuation with depth: the larger the exponent b , the faster the attenuation and hence the lower the transfer efficiency. A recent estimate of the exponent b is 0.84 ± 0.14 ¹³, which is equivalent to a range in transfer efficiency from 5% ($b=0.98$) to 12% ($b=0.70$) at 2000 m **and corresponds to an uncertainty in equilibrium atmospheric CO₂ levels of 28^{13,14} or 46⁹ μatm depending on the climate model used. However, this neglects the uncertainty resulting from the choice of the Martin curve over a different parameterization, which has been estimated to contribute an additional uncertainty in atmospheric CO₂ levels of 12 μatm ⁹. Rather than using the simple parameterizations which are compared in ref.⁹, current ESMs simulate the transfer efficiency in a more mechanistic manner by resolving some key controlling processes, such as the ballast protection of organic matter^{17–19}, temperature^{18–20} and oxygen-dependent^{17–21} remineralization, and aggregation/disaggregation²⁰. The extent to which the uncertainty in transfer efficiency in state-of-the-art ESMs can be represented by changing the exponent b of the Martin curve, or not, warrants further evaluation.**

In this study, **the model uncertainty in simulating the BCP is quantified by the variance of averaged vertical carbon flux over different spatial scales across ESMs. We distinguish between the model uncertainty within the euphotic zone and below by partitioning it into the model uncertainty of carbon export and the model uncertainty of transfer efficiency. Motivated by the wide application of the Martin curve, we further decomposed the latter into two components: the uncertainty which can be represented by changing the exponent b of the Martin curve (referred to here as type I) and the uncertainty which cannot (referred to as type II)**”

Section 2.4 (changed text is highlighted in **bold** fond here)

“.....
The Martin curve has been widely used **to fit the sparse observations of deep vertical carbon flux for extrapolating and comparing the transfer efficiency. The exponent b has also been perturbed to assess the sensitivity of atmospheric CO₂ levels to transfer efficiency**^{14,32}. Our results show that the model uncertainty of transfer efficiency is underestimated by using the Martin curve and exponent b . This is consistent with findings in ref.⁹ where six simple alternative parameterizations were compared with the Martin Curve. Therefore, sampling and using the vertical profiles of transfer efficiency is recommended rather than only reporting the exponent b when calibrating models.”

Section 4.3 (changed text is highlighted in **bold** fond here)

“.....
The model uncertainty of the transfer efficiency (U_{TE}) is further decoupled **into type I (U_I , dimensionless) and type II (U_{II} , dimensionless) to distinguish between model uncertainty of transfer efficiency which can be represented by changing the exponent b of the Martin curve or not. To achieve this, the simulated vertical profiles of transfer efficiency from each ESMs were**

fitted to the Martin curve and the variance of the predicted transfer efficiency ($\overline{TE_{martin}(z)}$) is exclusively due to the **exponent b**:

.....”

Please include a table in supplementary material with the key references describing the biogeochemical model structure for each of the CMIP6 models used in the analysis.

Response: Thanks for the suggestion. A table describing the mechanisms included in the different biogeochemical models with references has been added to our revised manuscript:

Table 2. Mechanisms included in different biogeochemical models to simulate the vertical attenuation of vertical carbon flux

Ocean BGC model	Ballast effect			Temperature dependence	Oxygen dependence	Bacteria utilization
	CaCO ₃	Opal	dust			
MARBL ¹⁷	Yes	Yes	Yes	No	Yes	Implicitly
BLINGv2 ¹⁸	Yes	No	Yes	Yes	Yes	Implicitly
COBALTv2 ¹⁹	Yes	Yes	Yes	Yes	Yes	Implicitly
PISCESv2 ²⁰	No	No	No	Yes	Yes	Implicitly
HAMOCC6 ^{21,22}	No	No	No	No	Yes	Implicitly
BFM5.2 ²⁵	No	No	No	Yes	No	Explicitly

Section 2.1: I found this section quite heavy-going as there are lots of numbers being cited. The authors should consider whether the detailed info on the fluxes and multiplication factors for the different biomes can be tabulated (or graphically presented), thus leaving the text to describe the key features.

Response: Thanks for the suggestion. We have added a new figure (see below) in our revised manuscript to better visualize the variability across different ESMs and moved the previous Figure 2 into the supplement. In addition, we have removed the redundant details from this section to focus more on the key messages.

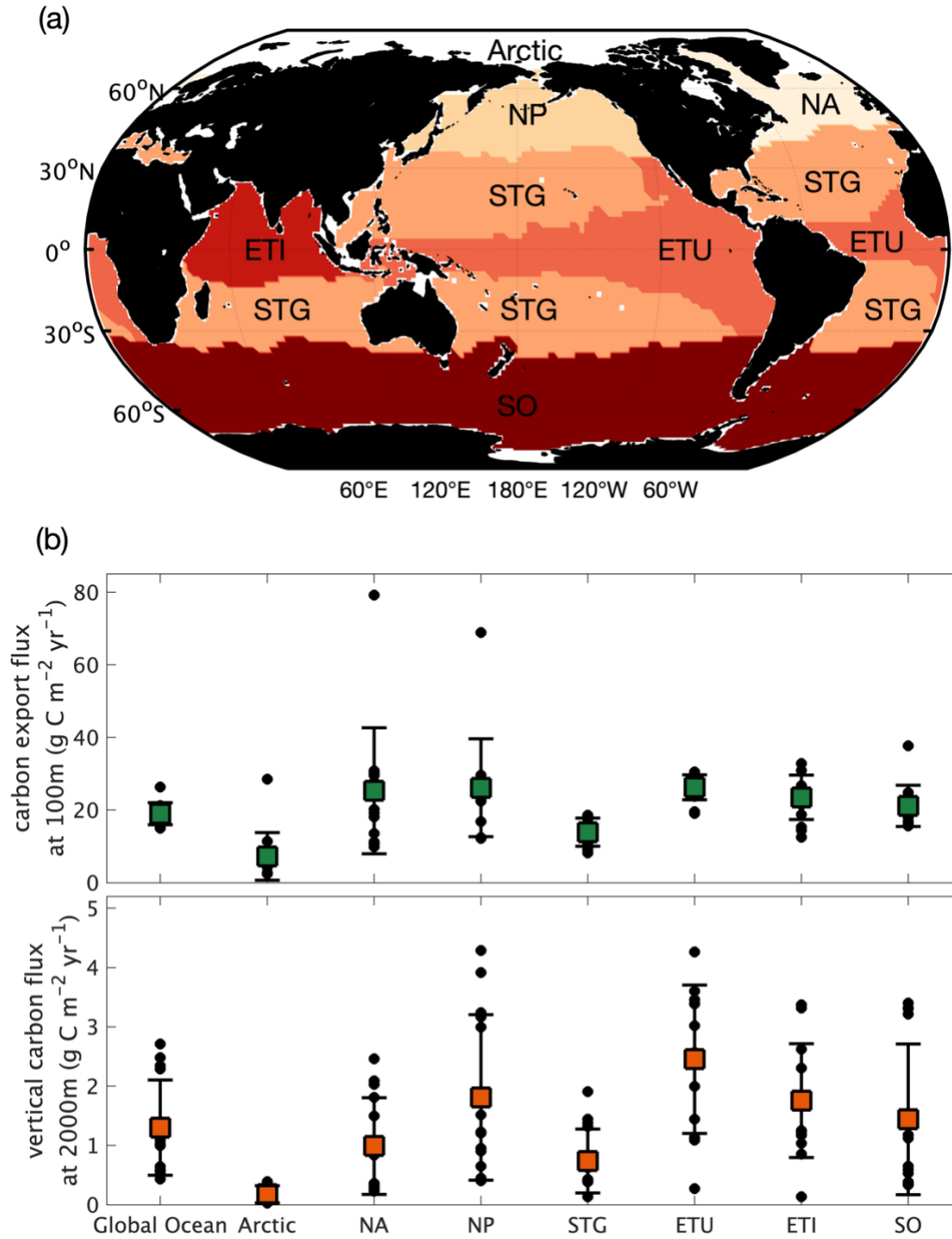


Figure 1. Map of biogeochemical biomes (a) and the vertical carbon flux averaged over each biome at 100 m and 2000 m (b). Errorbars represent the mean and standard deviation of 14 CMIP6 models and black dots represent each single model. The global distributions of all models are shown in Figure S1. Abbreviations: Southern Ocean (SO), Equatorial Tropical Indian (ETI), Equatorial Tropical Upwelling regions (ETU), Subtropical Gyres (STG), North Pacific (NA), North Atlantic (NA), and Arctic Ocean (Arctic).

Line 124: how do these values of b compare to observations-based estimates?

Response: Please see our revised manuscript text (changed text is highlighted in **bold font**):

“The global median of the exponent b varies from 0.63 in IPSL-CM5A2-INCA to 1.01 in CMCC-ESM2. This is higher than the estimates of 0.64 and 0.54 from observations in refs.^{5,15}. However, the global average of the exponent b ranging from 0.70 to 1.03 is consistent with a more recent data-constrained model estimate from 0.70 to 0.98¹³.

.....
In the three IPSL models and EC-Earth3-CC, the transfer efficiency at 2000 m is high in the productive **high-latitude regions and the Equatorial Tropical Upwelling biome** but low in the Subtropical Gyres. **These broad qualitative features can also be found in global patterns derived from free-drifting neutrally buoyant sediment traps (NBSTs)¹², an inverse model²³, and a data-assimilative ecosystem model²⁴, which have also reported the lowest transfer efficiency in Subtropical Gyres. Comparisons show that the transfer efficiency at 1000 m simulated by the four PISCESv2 ESMs is in general higher than that in the inverse model²³ (Figures S3-S5). The simulated transfer efficiency at 2000 m from $10.53 \pm 0.83\%$ to $17.99 \pm 1.26\%$ by the four PISCESv2 ESMs is also higher than the 5% to >8% range from the data-assimilative model²⁴.**

In CMCC-ESM2, the latitudinal patterns of transfer efficiency are similar to the four PISCESv2 ESMs. In general, the simulated transfer efficiency at 2000 m in the CMCC-ESM2 is comparable to the inverse model²³ and the data-assimilative ecosystem model²⁴ except that the transfer efficiency from CMCC-ESM2 is much higher in the North Atlantic and Arctic Ocean.

In contrast, in GFDL-CM4, the transfer efficiency at 2000 m is high in the Subtropical Gyres and low in the productive regions. This is because the ballast effect of CaCO_3 outweighs the temperature dependence of the remineralization rate and the ballast effect by opal is missing in this model¹⁸. General patterns of the fitted exponent b between the productive and oligotrophic regions are similar to two observational estimates^{5,15} (Figures S2-S3) despite regional discrepancies in transfer efficiency (Figures S6-S7). For instance, the elevated exponent b , which is equivalent to reduced transfer efficiency, near the equator is missing in ref.⁵ because it cannot be resolved by sparse concurrent measurements of carbon export and deep vertical carbon flux. In addition, the variation of b in GFDL-CM4 (0.50-1.03) is smaller than that of 0.36-1.09 and 0.42-1.75 from refs.^{5,15} (Figure S3).”

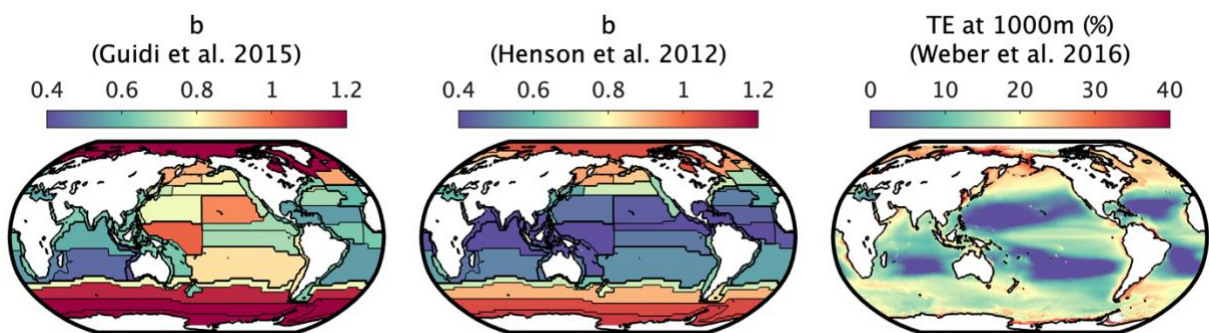


Figure S3. The regionalized exponent b within Longhurst provinces¹ and transfer efficiency at 1000 m which are estimated by refs.²⁻⁴.

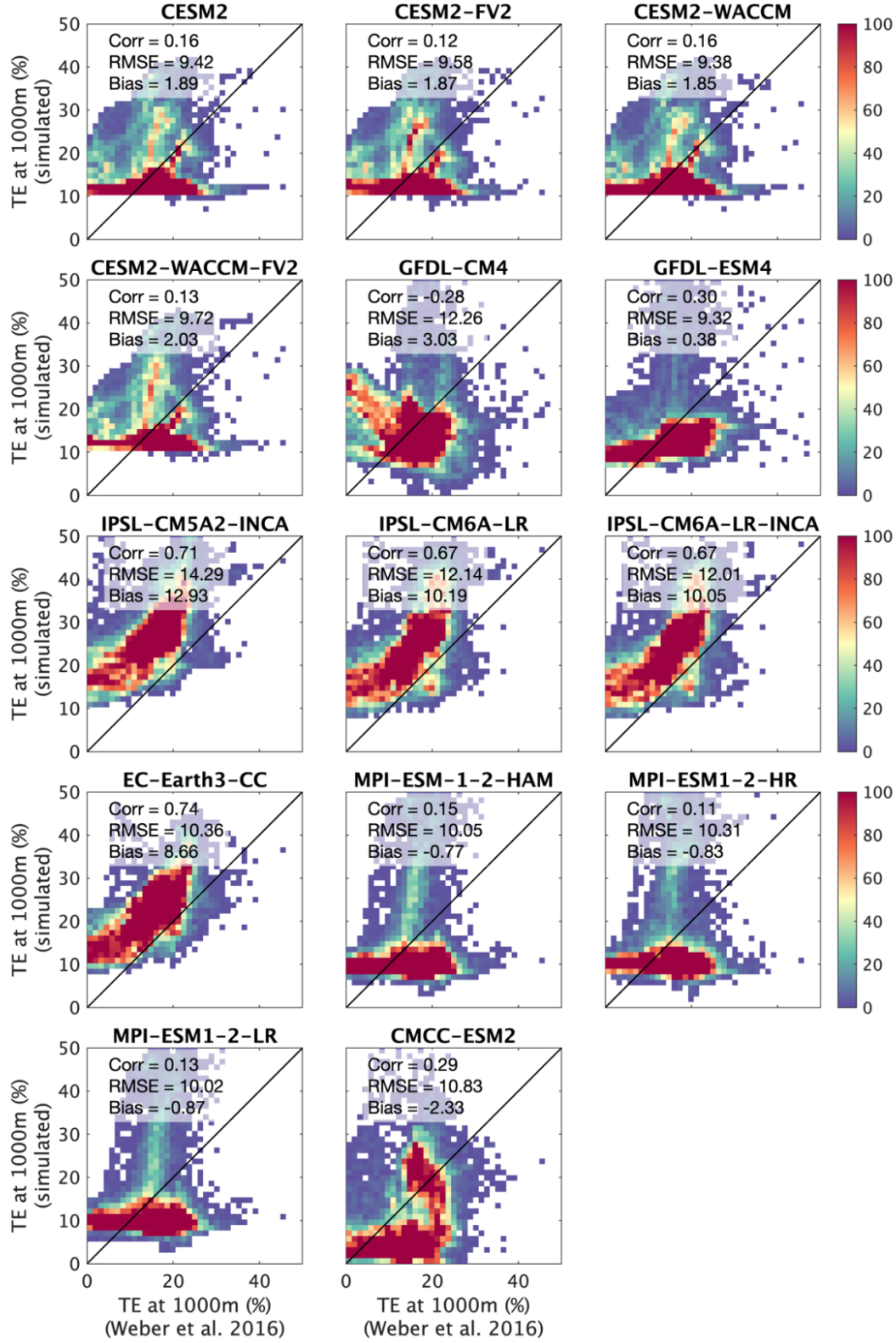


Figure S4. Comparisons of simulated transfer efficiency at 1000 m by CMIP6 models with estimates from ref.⁴

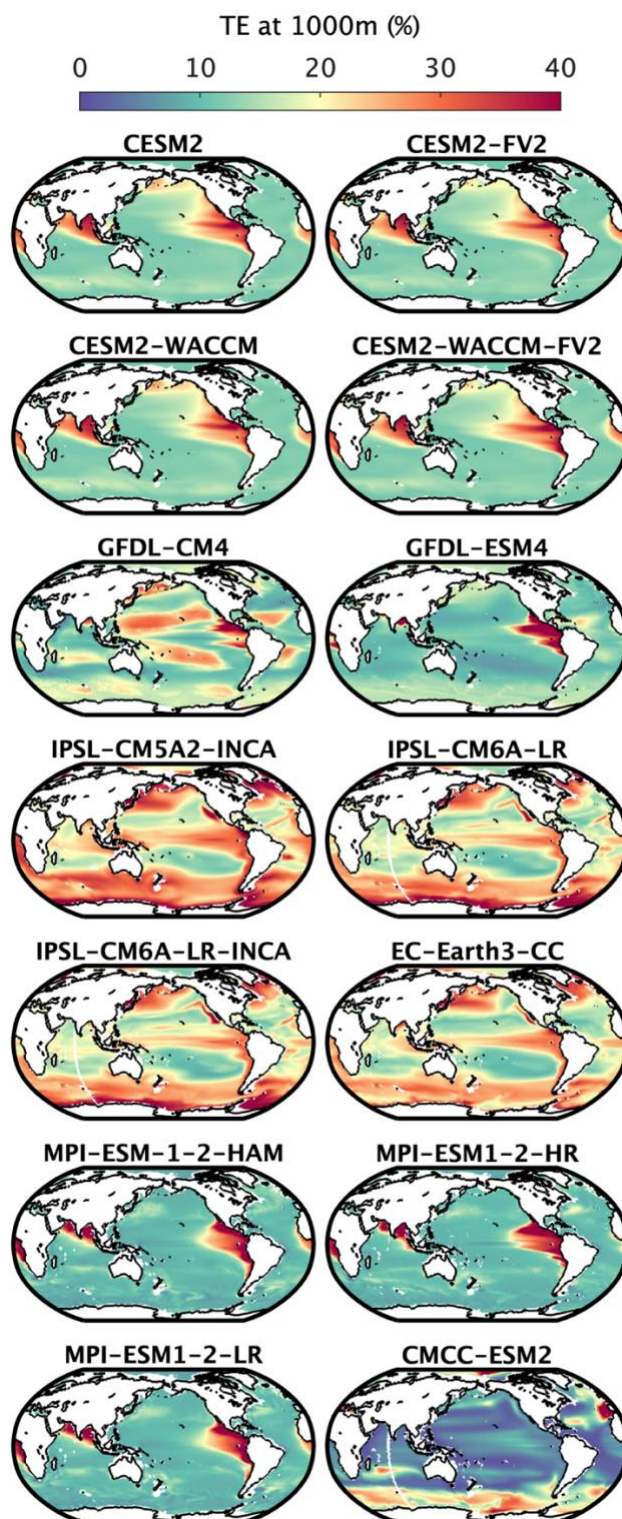


Figure S5. Spatial distributions of the transfer efficiency at 1000 m simulated by the CMIP6 models.

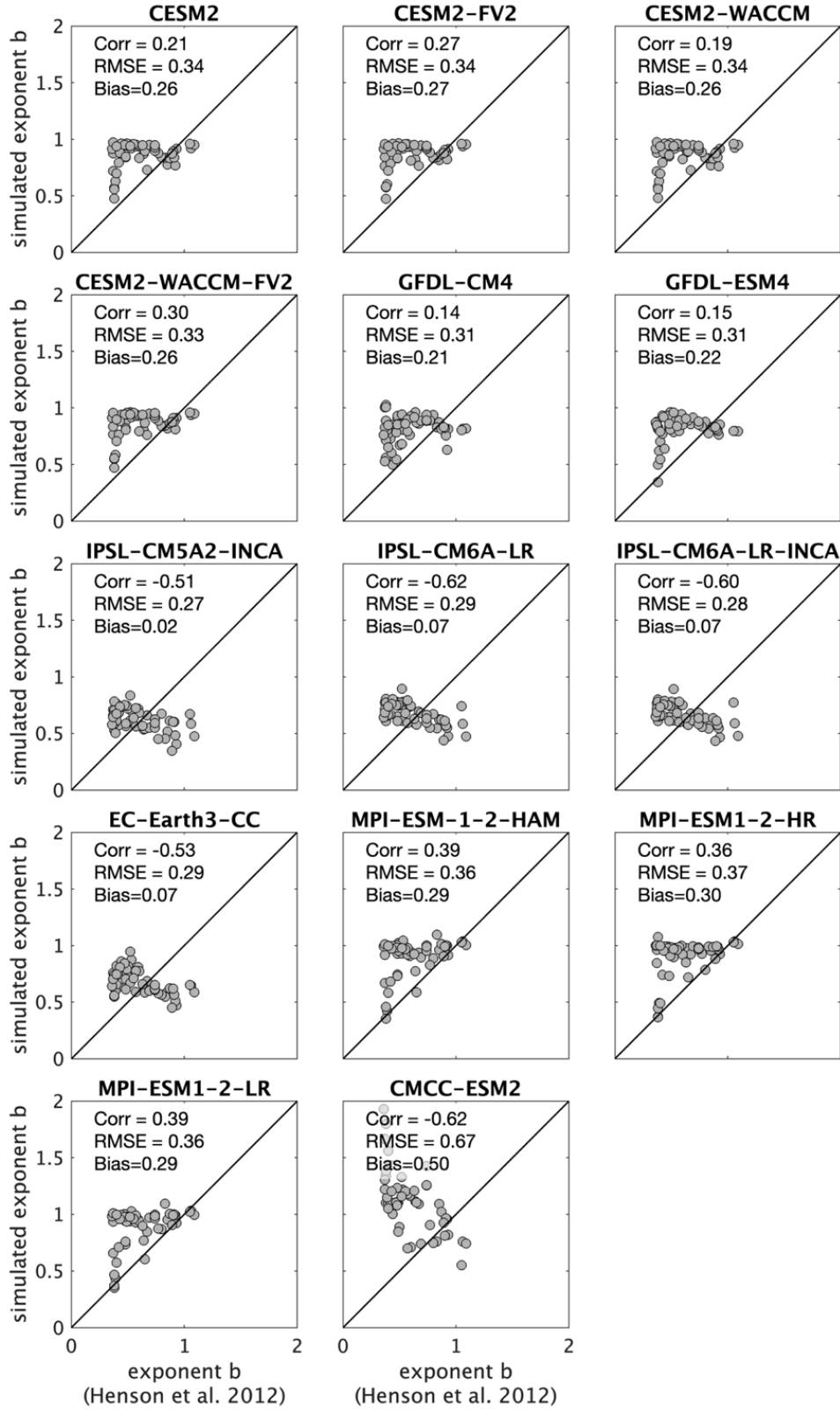


Figure S6. Comparisons of the regionalized exponent b which is simulated by CMIP6 models with estimates from ref.³

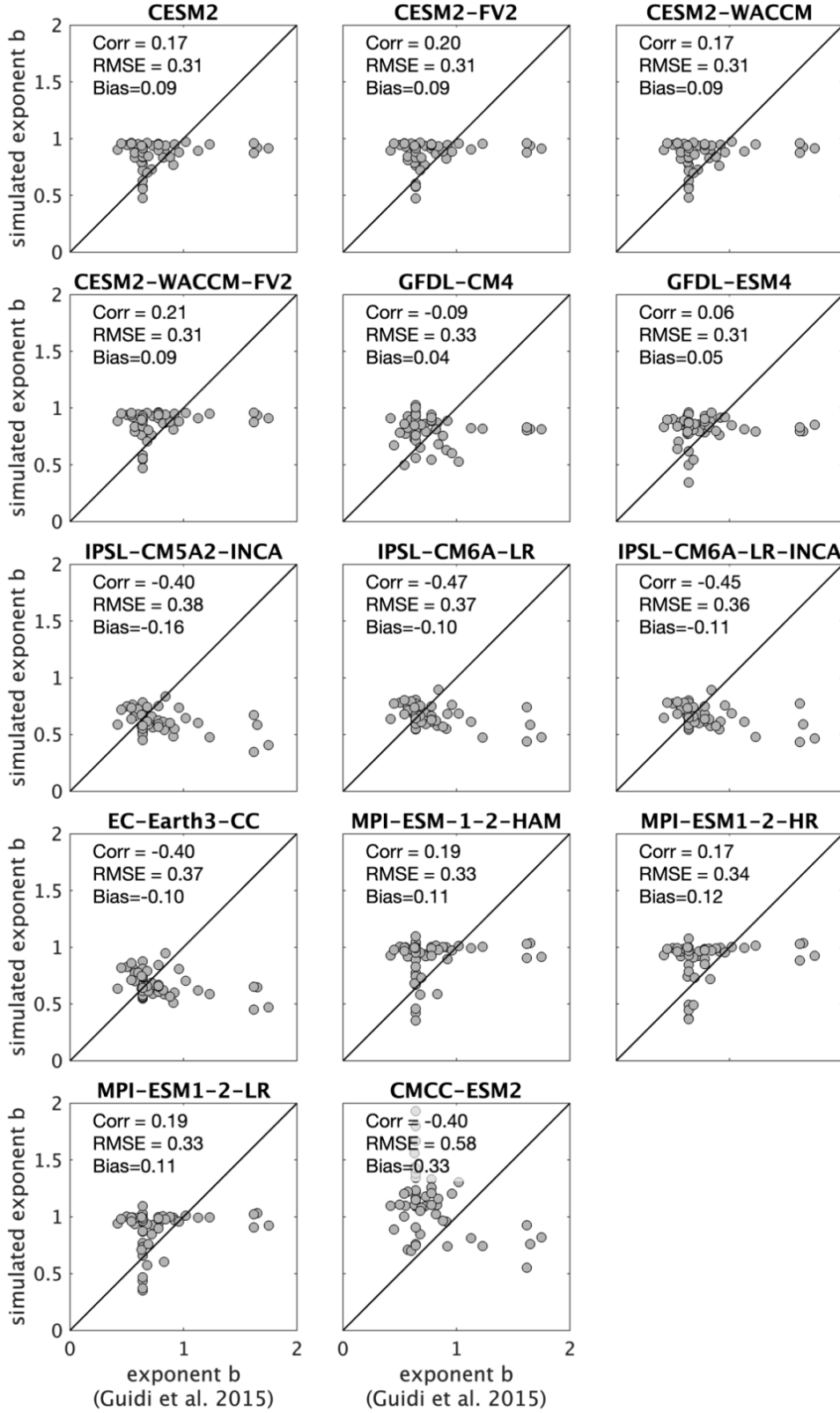


Figure S7. Comparisons of the regionalized exponent b which is simulated by CMIP6 models with estimates from ref.²

Line 125: do none of the models used here have temperature-dependent remineralisation?

Response: In MARBL and HAMOCC6, the temperature-dependence of remineralization is not considered. These two biogeochemical models are used by 4 CESM2 and 3 MPI ESMs. In our original manuscript, we focused on major mechanisms controlling the spatial patterns of transfer efficiency but neglect the discussion of other reasons. In our revised manuscript, we have attempted to make the discussion more comprehensive (changed text is highlighted in **bold** font):

“.....
The simulated transfer efficiency at 2000 m is homogenous in the four CESM2 models, the three MPI models, **and GFDL-ESM4 across latitudes except for the elevated transfer efficiency in Oxygen Minimum Zones (OMZs). This is because the remineralization rate is suppressed under low oxygen^{17,19,22} in these biogeochemical models (Tables 1-2). In GFDL-ESM4, the remineralization rate is dependent on temperature. The high temperature in low-latitude regions facilitates remineralization and should reduce the transfer efficiency. However, this effect is compensated for by the ballast effect of elevated CaCO₃ and lithogenic fluxes in low latitude regions, which enhances the transfer efficiency by protecting organic matter from remineralization or by accelerating the sinking velocity¹⁹. In the four CESM2 models, the ballast effect of CaCO₃ and opal are both simulated and may cancel each other out flattening the latitudinal gradient of transfer efficiency¹⁷, because the former is predominant in the Subtropical Gyres and the latter in productive regions. No temperature dependence is simulated in the four CESM2 and three MPI models^{17,21,22}.**
.....”

Line 142: it's not clear what “neglecting the protections by opal” means here.

Response: Based on the ballast hypothesis, organic matter is protected from remineralization by minerals including CaCO₃ and opal while sinking to the deep ocean. Given the elevated opal export in the productive regions due to diatoms, the ballast protection of organic matter by opal will increase transfer efficiency in productive regions and flatten the global pattern of transfer efficiency reported in GFDL-CM4. However, this mechanism is missing in the GFDL-CM4 model. In our revised manuscript, we have made this clearer (changed text is highlighted in **bold** font):

“In contrast, in GFDL-CM4, the transfer efficiency at 2000 m is high in the Subtropical Gyres and low in the productive regions. This is because the ballast effect of CaCO₃ outweighs the temperature dependence of the remineralization rate and the ballast effect by opal is missing in this model¹⁸.”

Line 162-163: It would be interesting to discuss the observation that the models used here have contrasting latitudinal patterns of transfer efficiency in the context of the ongoing discrepancy in observations-based global patterns of transfer efficiency (e.g. Henson et al., 2012 vs Marsay et al., 2015).

Response: Thanks for the suggestion. We have added some discussion in our revised manuscript (changed text is highlighted in **bold** font):

*“The contrasting global patterns of transfer efficiency have different implications for the future response of the BCP to climate change. Warming-enhanced stratification will likely reduce nutrient supply to the surface ocean, leading to an expansion of oligotrophic ecosystems. **This may amplify or dampen the** currently predicted reduction of transfer efficiency due to increased temperature. **Currently, observation-based estimates of transfer efficiency suggest contrasting patterns with one supporting high transfer efficiency in oligotrophic ecosystems^{5,15} and the other favouring the opposite pattern¹².** This indicates that existing in-situ observations provide insufficient constraints and suggests that an expansion of concurrent measurements of carbon export and deep vertical carbon flux is necessary to reconcile the contrasting patterns or falsify one of them.”*

Line 167: include a reference for “oligotrophic regions tend to export fecal pellets and refractory organic particles”. Also, as written this sentence implies that fecal pellets aren’t exported at high latitudes.

Response: Thanks for the suggestion. The fecal pellets are exported at both low and high latitude regions. In the low latitudes, however, the relative contributions of fecal pellets to the total export are higher than in the high latitudes where diatoms favour the formation of aggregates (Siegel et al., 2014; Nowicki et al., 2022). In our revised manuscript, we have removed this sentence to make the section more concise, but we have changed an earlier sentence in the same section to make the point clearer (changed text is highlighted in **bold** font):

*“**Aside** from ballast effects, this global pattern has **also** been suggested to be driven by particle composition because **of the elevated relative contribution from fast-sinking fecal pellets in oligotrophic regions^{5,26,27}.**”*

Figure 5a: Is the x-axis percentage (/100), as used in lines 198-220? What do negative contributions to uncertainty mean? Define U_f , U_{ep} etc. in the caption.

Response: The x-axis represents the values of different uncertainties. Based on the definition of different uncertainties in the Methods section, the unit for the model uncertainty of the vertical carbon flux (U_F) and the model uncertainty of carbon export (U_{EP}) is $(\log(\text{mg C m}^{-2} \text{ day}^{-1}))^2$, while the model uncertainty of transfer efficiency (U_{TE}) as well as type I and type II uncertainties (U_I and U_{II}) are dimensionless. We have clarified this in the Methods section in our revised manuscript.

The negative contributions of the co-dependence between carbon export and transfer efficiency ($U_{EP\&TE}$) suggest that they are anti-correlated across the different ESMs that we compared. In other words, when the carbon export changes, the transfer efficiency changes in the opposite direction thus reducing differences in the deep vertical carbon flux across different ESMs. This is particularly important in the Arctic Ocean and the North Atlantic on the biome scale (P10 L224-226) and in the Subtropical Gyres on the local scale (P10 L233-236).

As suggested, the definition of different uncertainties has been added to the figure caption in our revised manuscript (changed text is highlighted in **bold** font):

*“**Figure 3. Partitioning of model uncertainty in the vertical carbon flux (U_F) as a function of depth (a) and at 2000 m (b) into the model uncertainty of carbon export (U_{EP}), the model uncertainty of transfer efficiency (U_{TE}), and the co-dependence between carbon export and transfer efficiency ($U_{EP\&TE}$) as described in Methods. The horizontal dashed black lines represent the crossover depth where the uncertainty of transfer efficiency (U_{TE}) starts to dominate the model uncertainty in the vertical carbon flux (U_F).”***

Line 239-241: the new paper by Ricour (2023, Nature Geosci) may provide some useful information on sequestration horizon.

Response: Thanks for the very helpful suggestion. We have revised our manuscript accordingly (changed text is highlighted in **bold** font):

*“With respect to long-term carbon sequestration and global climate, the relative importance of **model uncertainty between carbon export and transfer efficiency** depends on the competition between the crossover depth and the sequestration horizon, **i.e.** the depth where carbon is sequestered for climate-relevant timescales (e.g., >100 years). Typically, the deeper the organic carbon is respired, the longer it is sequestered^{28,29}. The 1000 and 2000 m **depths have been** widely used as static horizons^{5–7,15,23,24}. **However, the sequestration horizon is spatially variable. There have been some published global maps about the carbon sequestration time²⁷ and mean residence time at each depth³⁰. A recent study suggests that carbon sequestration occurs throughout the water column and that about half of the sequestration is above 1000 m³¹. On the global scale, the crossover depth of 900 m underscores the importance of both model uncertainty of carbon export and transfer efficiency for understanding the century-scale sequestration. The importance of model uncertainty of transfer efficiency is expected to increase when considering longer-term sequestration.**”*

Figure 7: Define Up, Us etc. in the caption.

Response: We have renamed these two types of uncertainty as type I and type II uncertainty. Definitions were also added to the figure caption in our revised manuscript (changed text is highlighted in **bold** font):

*“**Figure 5. The partitioning of model uncertainty of transfer efficiency (U_{TE}) as a function of depth (a) and at 2000 m (b) into the type I (U_I) and type II (U_{II}) uncertainties that are representable with the Martin curve and not, as well as their co-dependence ($U_{I\&II}$) as described in Methods.**”*

Line 289: “more similar” than what?

Response: We have rewritten this section and removed this sentence in our revised manuscript.

Section 2.5: this section was interesting, but rather an aside – it didn’t seem to address the main point of the manuscript at all. I got to the end of the section and couldn’t see how it was more than tangentially relevant. I suggest the authors rewrite this section. For example, what are the

implications of the results to our understanding of future project changes to carbon storage? This section gives the impression that the only issue is paucity of observations and increasing the size of in situ datasets will make the models better, but there are also constraints on model parameterisation, complexity, structure etc. too that could be discussed.

Response: We really appreciate this constructive comment and suggestion. This section has been rewritten. Please see details in our revised manuscript.

Figure 8b: how are the match-ups between the models and observations done?

Response: The in-situ observations were compared with the monthly climatology of modelled POC fluxes in the nearest model grid cell of each ESMs. We have clarified this in a new section in our revised manuscript (new text):

“4.2 In-situ observations of the vertical carbon flux

In-situ observations of the vertical carbon flux were compiled by ref.⁵⁹ and are available in the supporting information of ref.¹³. Most of the available in-situ observations (~90%) were collected by sediment traps deployed for less than 30 days and hence are compared to the monthly climatology in the nearest model grid cell of each ESMs.”

Line 339: why do the 14 models produce 2 contrasting patterns of transfer efficiency?

Response: As we discussed in section 2.2, this is because of the differences in model structures and model parameters across different ESMs. We have improved our discussion in section 2.2 in our revised manuscript to make it more comprehensive and revised this sentence to make it clearer (changed text is highlighted in **bold** font):

Section 2.2

“..... The simulated transfer efficiency at 2000 m can be classified into three, qualitatively different latitudinal patterns (Figures 2 and S2): the first is a homogenous transfer efficiency, the other two are opposite latitudinal patterns, either high in the oligotrophic Subtropical Gyre and low in productive regions or vice versa.

The simulated transfer efficiency at 2000 m is homogenous in the four CESM2 models, the three MPI models, and GFDL-ESM4 across latitudes except for the elevated transfer efficiency in Oxygen Minimum Zones (OMZs). This is because the remineralization rate is suppressed under low oxygen^{17,19,22} in these biogeochemical models (Tables 1-2). In GFDL-ESM4, the remineralization rate is dependent on temperature. The high temperature in low-latitude regions facilitates remineralization and should reduce the transfer efficiency. However, this effect is compensated for by the ballast effect of elevated CaCO₃ and lithogenic fluxes in low latitude regions, which enhances the transfer efficiency by protecting organic matter from remineralization or by accelerating the sinking velocity¹⁹. In the four CESM2 models, the ballast effect of CaCO₃ and opal are both simulated and may cancel each other out flattening the latitudinal gradient of transfer efficiency¹⁷, because the former is predominant in the Subtropical Gyres and the latter in productive regions. No temperature dependence is simulated in the four CESM2 and three MPI models^{17,21,22}.

In the three IPSL models and EC-Earth3-CC, the transfer efficiency at 2000 m is high in the productive high-latitude regions and the Equatorial Tropical Upwelling biome but low in the

Subtropical Gyres. **Global patterns in the three IPSL models and EC-Earth3-CC are due to the temperature-dependent remineralization rate and the differentiation of sinking velocities between two size classes of organic particles in their biogeochemical models, i.e., PISCESv2²⁰.** In productive regions where diatoms are dominant, the exported organic particles tend to be large and fast sinking because diatoms aggregate efficiently.

In CMCC-ESM2, the latitudinal patterns of transfer efficiency are similar to the four PISCESv2 ESMs. However, the transfer efficiency is low in the North Pacific and Equatorial Tropical Upwelling biomes. Unlike other models analyzed here, the utilization of organic particles by bacteria is simulated explicitly in CMCC-ESM2. It is dependent on the nutritional content of organic particles when bacteria are abundant and otherwise on temperature and bacterial biomass²⁵. Although the bacterial utilization of organic particles can be modulated by oxygen, it appears that this oxygen-dependence is deactivated in the CMCC-ESM2 because transfer efficiency in OMZs is not elevated.

In contrast, in GFDL-CM4, the transfer efficiency at 2000 m is high in the Subtropical Gyres and low in the productive regions. **This is because the ballast effect of CaCO₃ outweighs the temperature dependence of the remineralization rate and the ballast effect by opal is missing in this model¹⁸.**

Conclusion (new text)

“..... Owing to the discrepancies in model structures and parameters, the simulated transfer efficiency from 14 CMIP6 models can be classified into three broad categories, with the first category features a homogenous global pattern of transfer efficiency and the remaining categories encapsulate two contrasting patterns.”

Conclusion section: the manuscript ends rather abruptly! What are the implications of the results shown for model development?

Response: Thanks for the comment. We have rephrased the conclusion section in our revised manuscript (changed text is highlighted in **bold** font):

“In this study, model uncertainty in simulating the BCP within state-of-the-art **ocean biogeochemical models** was partitioned into different contributors and analyzed. Comparisons across 14 CMIP6 models show large variations in carbon export at 100 m and increased variations at depth due to differences in transfer efficiency. **Owing to the discrepancies in model structures and parameters, the simulated transfer efficiency from 14 CMIP6 models can be classified into three broad categories, with the first category features a homogenous global pattern of transfer efficiency and the remaining categories encapsulate two contrasting patterns.** In addition, uncertainty analysis based on 6 of 14 representative CMIP6 models suggests an initial dominance of model uncertainty of carbon export **which is overtaken** by the model uncertainty of transfer efficiency at 900 m on the global scale. Since a recent study has suggested that the century-scale carbon sequestration is equivalently important above and below 1000 m³¹, model uncertainties of carbon export and transfer efficiency are both important, implying the necessity to improve model representations both within the euphotic zone and below. In addition, an expansion of observations of not only carbon export but also flux attenuation in the deep ocean is important. The model uncertainty of transfer efficiency was further decoupled into uncertainties that are representable by the Martin curve or not, both of which show comparable significance,

suggesting that the usage of Martin curve and exponent b for comparing transfer efficiency underestimates the model uncertainty of transfer efficiency. Therefore, the vertical profile of transfer efficiency is recommended for model validation rather than using the exponent b . ”

Reviewer #2 (Remarks to the Author):

The manuscript “Organic matter transfer efficiency is the largest uncertainty for biological carbon pump in CMIP6 models” by Wang and Fennel analyzes CMIP6 models for their ocean biogeochemical diversity. The manuscript provides a highly useful comparison of the CMIP6 models in their diversity in representing carbon export at 100 and transfer efficiency. However, I have three criticisms of the manuscript in its present form: 1) the authors are overly vague on their definition of “uncertainty”, 2) the authors do not compare the models in a quantitative way to available estimates of transfer efficiency to assess whether any of the models are reproducing observed patterns (with the exception of the qualitative statement comparing to Guidi et al., 2015) and 3) the separation between “structural” and “Parametric” uncertainty is unclear. All these concerns are described below and should be easily rectified in a revised version.

Response: We really appreciate these constructive comments and suggestions.

Regarding the first point, we defined the uncertainty as the variance of the averaged properties of interest (i.e. vertical carbon flux, transfer efficiency) over different spatial scales (i.e., global ocean, biomes) across the different CMIP6 models. We have revised the introduction section in our revised manuscript to make it clearer (new text):

“.....
In this study, the model uncertainty in simulating the BCP is quantified by the variance of averaged vertical carbon flux over different spatial scales across ESMs.....”

Regarding the second point, we have added quantitative comparisons of transfer efficiency in our revised manuscript as suggested. For details, please see our responses to the specific comments below.

Finally, with respect to the separation between parametric and structural uncertainties, we have re-defined these two types of uncertainty as type I (U_I) and type II (U_{II}) uncertainty. Type I represents the uncertainty of transfer efficiency that is representable with the Martin curve by changing the exponent b . Type II is the uncertainty that cannot be represented as such. Since the Martin curve and its exponent b have been widely used to extrapolate and compare the transfer efficiency, we believe that this distinction will be helpful to discuss whether the usage of Martin curve is appropriate. The motivation, definition, as well as interpretation and discussion of these two types of model uncertainty have been rephrased in our revised manuscript. Please see our response to the following specific comments for details.

7, 8, 9, and 10 - The general word “uncertainty” is used where I believe the more specific “model structural diversity” is intended. As the CMIP6 models are all derived from a similar set of scientific literature and often the same algorithms, they cannot be expected to span the true uncertainty range. As such, the analysis of CMIP6 models can only sample the realized structural diversity. On the flip side, it may be that some of this diversity can be falsified by observations such that the differences from observations should be characterized as “bias” rather than “uncertainty”. However, given that the Hawkins and Sutton use of “structural uncertainty” has

become so pervasive, I think it is fine to use “structural uncertainty” throughout as long as the authors add a contextualization statement on the above limitations.

Response: Thanks for pointing this out. We agree and have added a definition and a statement of caveat in the Introduction section of our revised manuscript (new text):

“.....
In this study, the model uncertainty in simulating the BCP is quantified by the variance of averaged vertical carbon flux over different spatial scales across ESMs. ”

“..... *A caveat is that the uncertainty analyzed is sampled over the realized diversity of the selected CMIP6 models which do not necessarily span the real uncertainty range. ”*

9-10 – “but by uncertainty in” should be “and”

Response: Addressed as suggested.

41-43 – The sentence justifying a distinction between “structural” and “parametric” uncertainty - “The former arises from the choice of model structure while the latter is due to the assignment of poorly known parameters and the varying degrees of parameter tuning across different ESMs.” Is not compelling. I think I understand what the authors are getting at in terms of the difference in physics and chemistry between knowing the structural forms versus knowing the values of parameters, but in ocean biogeochemistry, such a clean distinction does not exist. There is typically no single “correct” structural form, just a bunch of semi-empirical options which must be calibrated and can give similar overall behaviors (e.g. Martin, versus Pace power functions, decreased remineralization timescale versus increased sinking speed, Monod versus Michaelis-Menten for growth).

Response: We agree with the reviewer that there is no single “correct” structural form. In our revised manuscript, we re-defined the parametric (U_P) and structural (U_S) uncertainty as type I (U_I) and type II (U_{II}) uncertainty as follows:

Introduction section (changed text is highlighted in **bold** font here):

*“The BCP can be assessed by quantifying carbon export **out of** the euphotic zone and the depth-dependent transfer efficiency⁴⁻⁸, **which measures the** fraction of the exported carbon that survives respiration and reaches the deep ocean. **Given the complex mechanisms controlling the transfer efficiency, simple parameterizations with different mechanistic underpinnings have been used to represent and to compare the transfer efficiency**^{9,10}. The Martin curve¹¹, **which** normalizes the vertical carbon flux to a reference depth, e.g., 100 m, through **the** power-law equation $F=F_0 \times (z/z_0)^{-b}$, **is the most widely used parameterization**¹²⁻¹⁶. The exponent b is an indicator of the carbon flux attenuation with depth: the larger the exponent b , the faster the attenuation and hence the lower the transfer efficiency. A recent estimate of the exponent b is 0.84 ± 0.14 ¹³, which is equivalent to a range in transfer efficiency from 5% ($b=0.98$) to 12% ($b=0.70$) at 2000 m **and corresponds to an uncertainty** in equilibrium atmospheric CO_2 levels of $28^{13,14}$ or $46^9 \mu\text{atm}$ depending on the climate model used. **However, this neglects the uncertainty resulting from the choice of the Martin curve over a different parameterization, which has been estimated to contribute an additional uncertainty in atmospheric CO_2 levels of $12 \mu\text{atm}^9$. Rather than using the simple parameterizations which are compared in ref.⁹, current ESMs simulate***

the transfer efficiency in a more mechanistic manner by resolving some key controlling processes, such as the ballast protection of organic matter^{17–19}, temperature^{18–20} and oxygen-dependent^{17–21} remineralization, and aggregation/disaggregation²⁰. The extent to which the uncertainty in transfer efficiency in state-of-the-art ESMs can be represented by changing the exponent b of the Martin curve, or not, warrants further evaluation.

In this study, the model uncertainty in simulating the BCP is quantified by the variance of averaged vertical carbon flux over different spatial scales across ESMs. We distinguish between the model uncertainty within the euphotic zone and below by partitioning it into the model uncertainty of carbon export and the model uncertainty of transfer efficiency. Motivated by the wide application of the Martin curve, we further decomposed the latter into two components: the uncertainty which can be represented by changing the exponent b of the Martin curve (referred to here as type I) and the uncertainty which cannot (referred to as type II)

Section 2.4 (changed text is highlighted in **bold** fond here)

“.....
The Martin curve has been widely used **to fit the sparse observations of deep vertical carbon flux for extrapolating and comparing the transfer efficiency**. The exponent b has also been perturbed to assess the sensitivity of atmospheric CO₂ levels to transfer efficiency^{14,32}. Our results show that the model uncertainty of transfer efficiency is underestimated by using the Martin curve and exponent b . This is consistent with findings in ref.⁹ where six simple alternative parameterizations were compared with the Martin Curve. Therefore, sampling and using the vertical profiles of transfer efficiency is recommended rather than only reporting the exponent b when calibrating models.”

Section 4.3 (changed text is highlighted in **bold** fond here)

“.....
The model uncertainty of the transfer efficiency (U_{TE}) is further decoupled **into type I (U_I , dimensionless) and type II (U_{II} , dimensionless) to distinguish between model uncertainty of transfer efficiency which can be represented by changing the exponent b of the Martin curve or not**. To achieve this, the simulated vertical profiles of transfer efficiency from each ESMs were fitted to the Martin curve and the variance of the predicted transfer efficiency ($\overline{TE_{martin}(z)}$) is exclusively due to the **exponent b** :
.....”

48 – I think “be” should be “have remained”

Response: Revised as suggested and this sentence has been revised further and moved to section 2.5 (changed text is highlighted as **bold** fond here):

*“Unlike the euphotic zone, model representations of flux attenuation **have remained** relatively simple³⁵ **and there is no agreement on which processes should be included.**”*

146 - The authors mention Guidi et al., 2015 UVP transfer efficiency estimates as being supportive of GFDL CM4, but this is the only place where an independent verification is applied. The authors should conduct a quantitative assessment of models for transfer efficiency based on the available estimates including not only Guidi et al., 2015 but also:

Henson et al 2012 <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2011GB004099>
Weber et al., 2016 <https://www.pnas.org/doi/abs/10.1073/pnas.1604414113>
Devries and Weber, 2017

<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016GB005551>

I would note that while the Henson and Weber estimates have strongly divergent patterns, the Devries and Weber 2017 distinguish between the 1000m and 1000-2000 m transfer efficiency which illustrate very different regional patterns.

Response: Thanks for the suggestion. We have added quantitative comparisons of transfer efficiency between 14 CMIP6 models with previous observational and modelling estimates in our revised manuscript. The estimate of transfer efficiency from Devries and Weber (2017) was not used for point-to-point comparisons because we do not have access to their model output. With respect to the contrasting patterns of transfer efficiency between Zeu-1000m and 1000-2000m, as discussed in Devries and Weber (2017), this is not enough to reverse the latitudinal patterns of transfer efficiency between the different depth horizons. As a result, the pattern of transfer efficiency at 2000 m is similar with that at 1000 m. We have added this discussion in our revised manuscript (changed text is highlighted in **bold** font):

“The global median of the exponent b varies from 0.63 in IPSL-CM5A2-INCA to 1.01 in CMCC-ESM2. This is higher than the estimates of 0.64 and 0.54 from observations in refs.^{5,15}. However, the global average of the exponent b ranging from 0.70 to 1.03 is consistent with a more recent data-constrained model estimate from 0.70 to 0.98¹³.

.....
In the three IPSL models and EC-Earth3-CC, the transfer efficiency at 2000 m is high in the productive high-latitude regions and the Equatorial Tropical Upwelling biome but low in the Subtropical Gyres. These broad qualitative features can also be found in global patterns derived from free-drifting neutrally buoyant sediment traps (NBSTs)¹², an inverse model²³, and a data-assimilative ecosystem model²⁴, which have also reported the lowest transfer efficiency in Subtropical Gyres. Comparisons show that the transfer efficiency at 1000 m simulated by the four PISCESv2 ESMs is in general higher than that in the inverse model²³ (Figures S3-S5). The simulated transfer efficiency at 2000 m from 10.53±0.83% to 17.99±1.26% by the four PISCESv2 ESMs is also higher than the 5% to >8% range from the data-assimilative model²⁴.

In CMCC-ESM2, the latitudinal patterns of transfer efficiency are similar to the four PISCESv2 ESMs. In general, the simulated transfer efficiency at 2000 m in the CMCC-ESM2 is comparable to the inverse model²³ and the data-assimilative ecosystem model²⁴ except that the transfer efficiency from CMCC-ESM2 is much higher in the North Atlantic and Arctic Ocean.

In contrast, in GFDL-CM4, the transfer efficiency at 2000 m is high in the Subtropical Gyres and low in the productive regions. This is because the ballast effect of CaCO₃ outweighs the temperature dependence of the remineralization rate and the ballast effect by opal is missing in this model¹⁸. General patterns of the fitted exponent b between the productive and oligotrophic regions are similar to two observational estimates^{5,15} (Figures S2-S3) despite regional discrepancies in transfer efficiency (Figures S6-S7). For instance, the elevated exponent b , which is equivalent to reduced transfer efficiency, near the equator is missing in ref.⁵ because it cannot be resolved by sparse concurrent measurements of carbon export and deep vertical

carbon flux. In addition, the variation of b in GFDL-CM4 (0.50-1.03) is smaller than that of 0.36-1.09 and 0.42-1.75 from refs.^{5,15} (Figure S3).

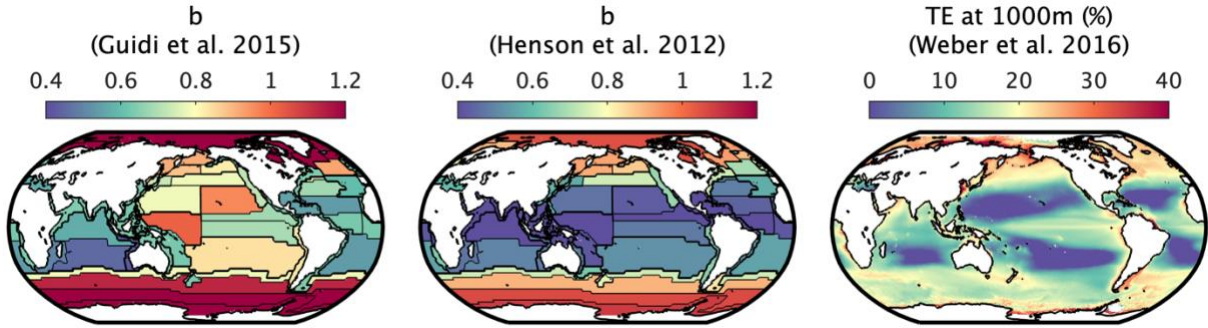


Figure S3. The regionalized exponent b within Longhurst provinces¹ and transfer efficiency at 1000 m which are estimated by refs.²⁻⁴.

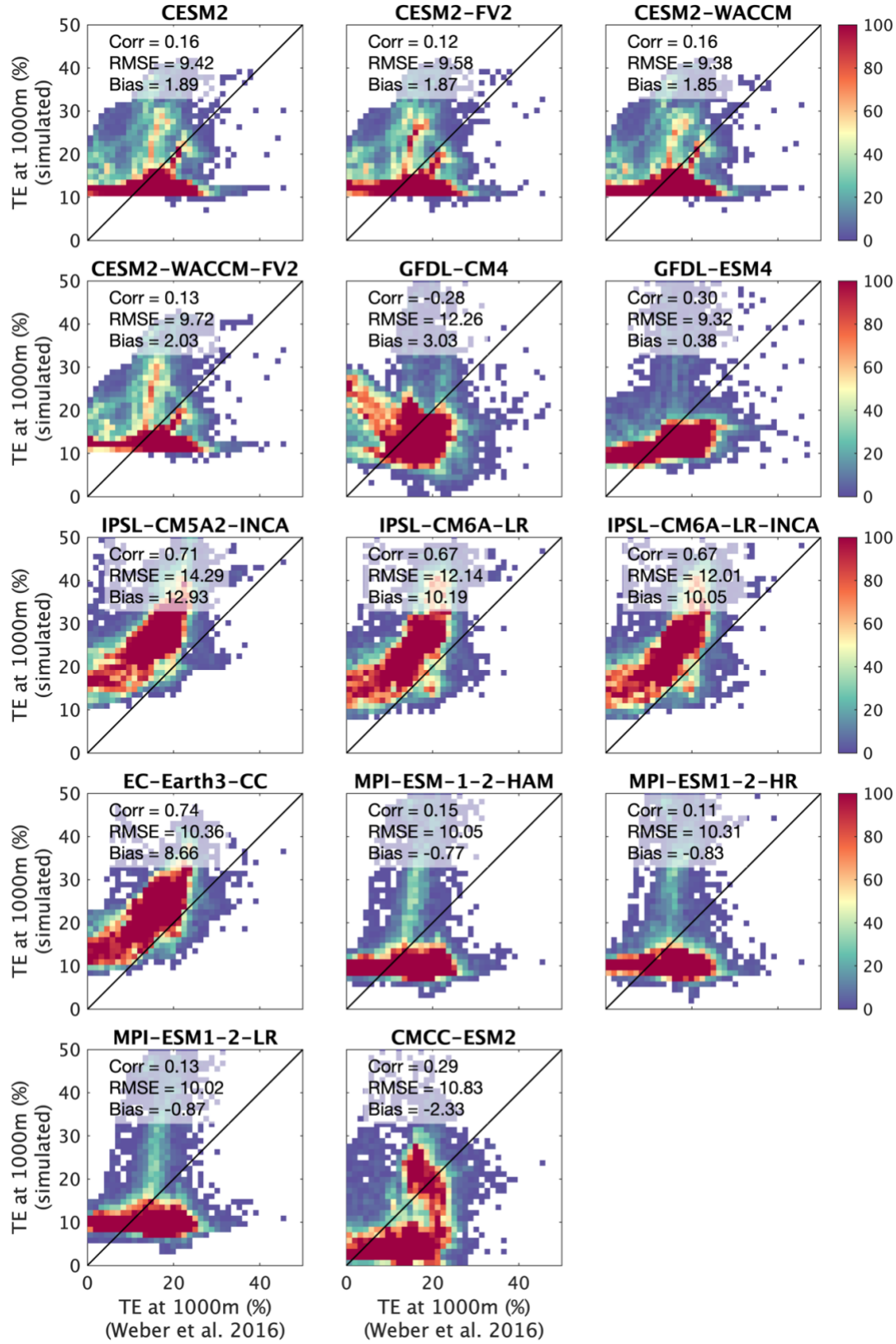


Figure S4. Comparisons of simulated transfer efficiency at 1000 m by CMIP6 models with estimates from ref.⁴

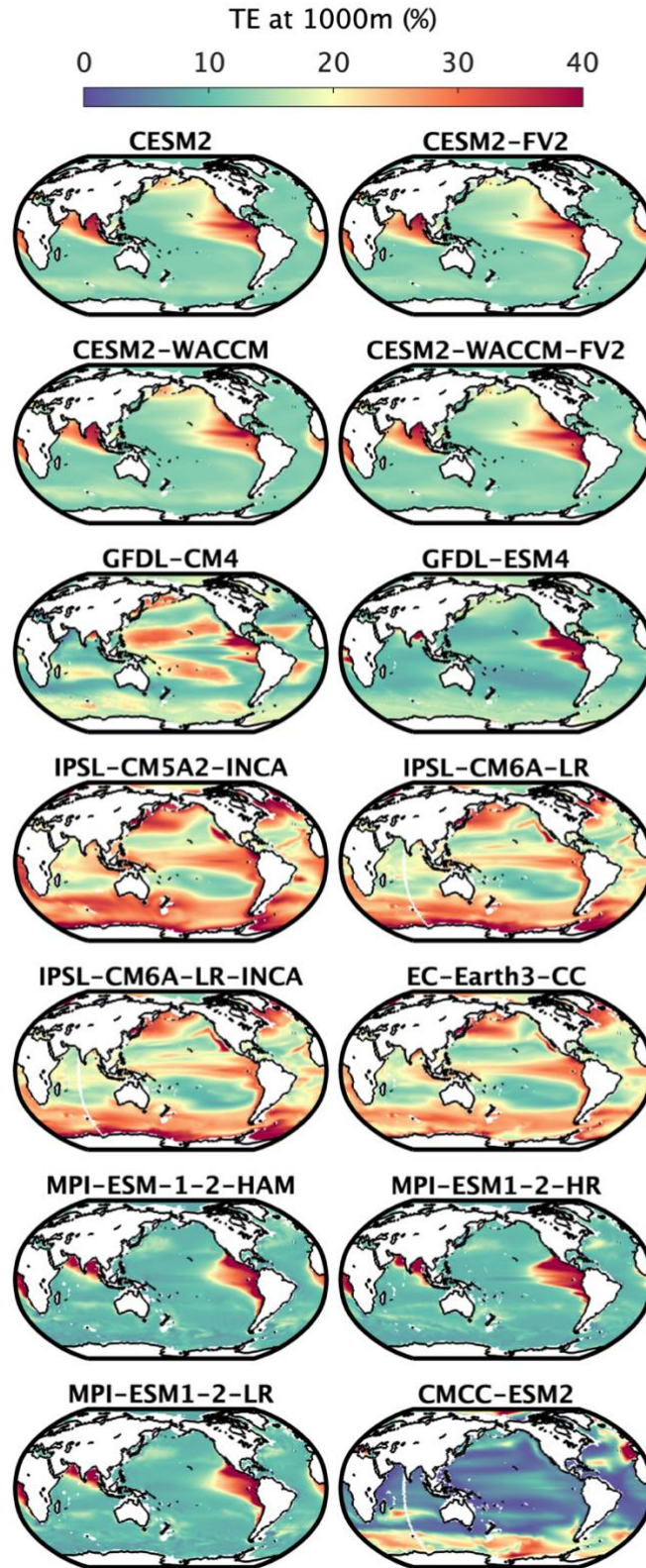


Figure S5. Spatial distributions of the transfer efficiency at 1000 m simulated by the CMIP6 models.

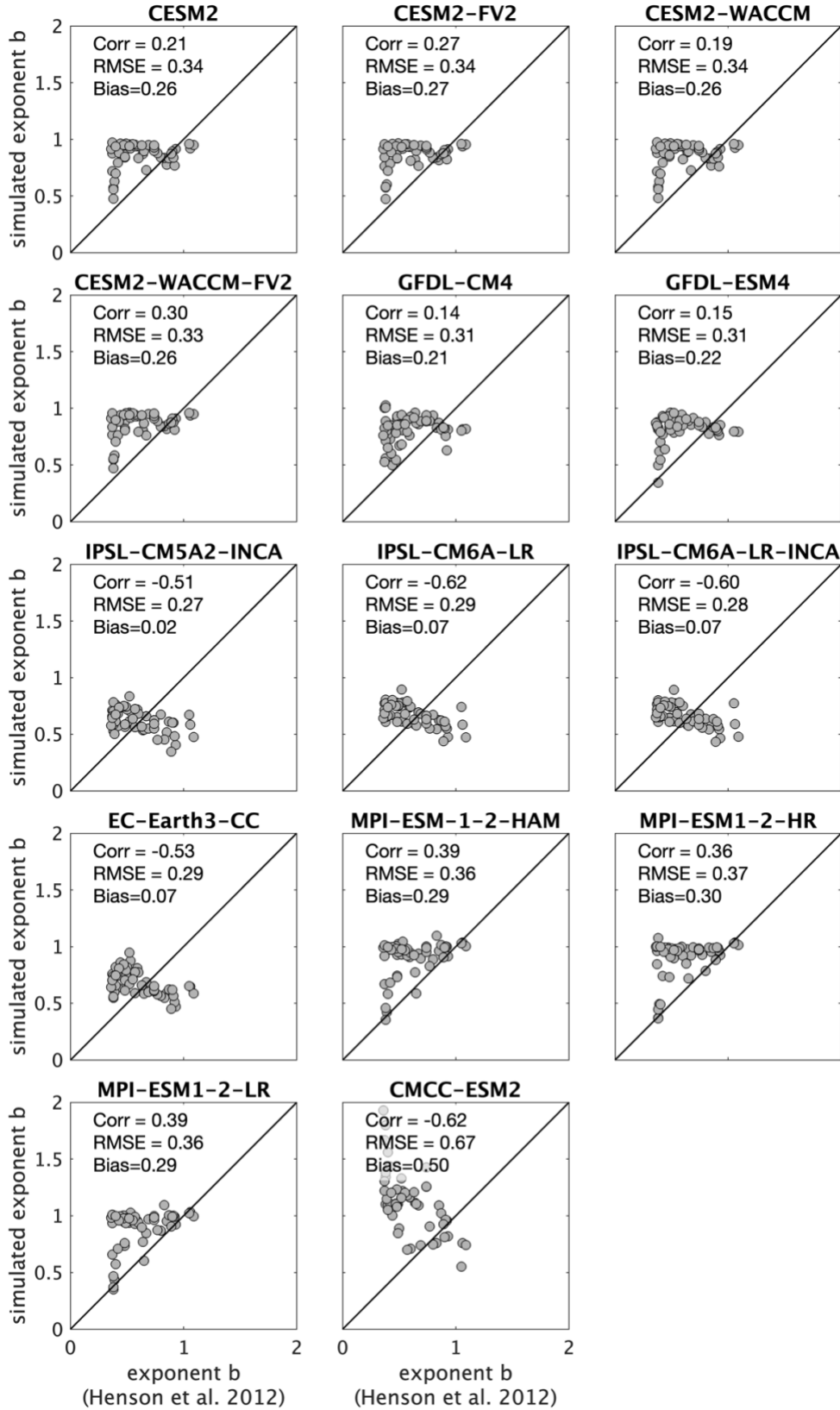


Figure S6. Comparisons of the regionalized exponent b which is simulated by CMIP6 models with estimates from ref.³

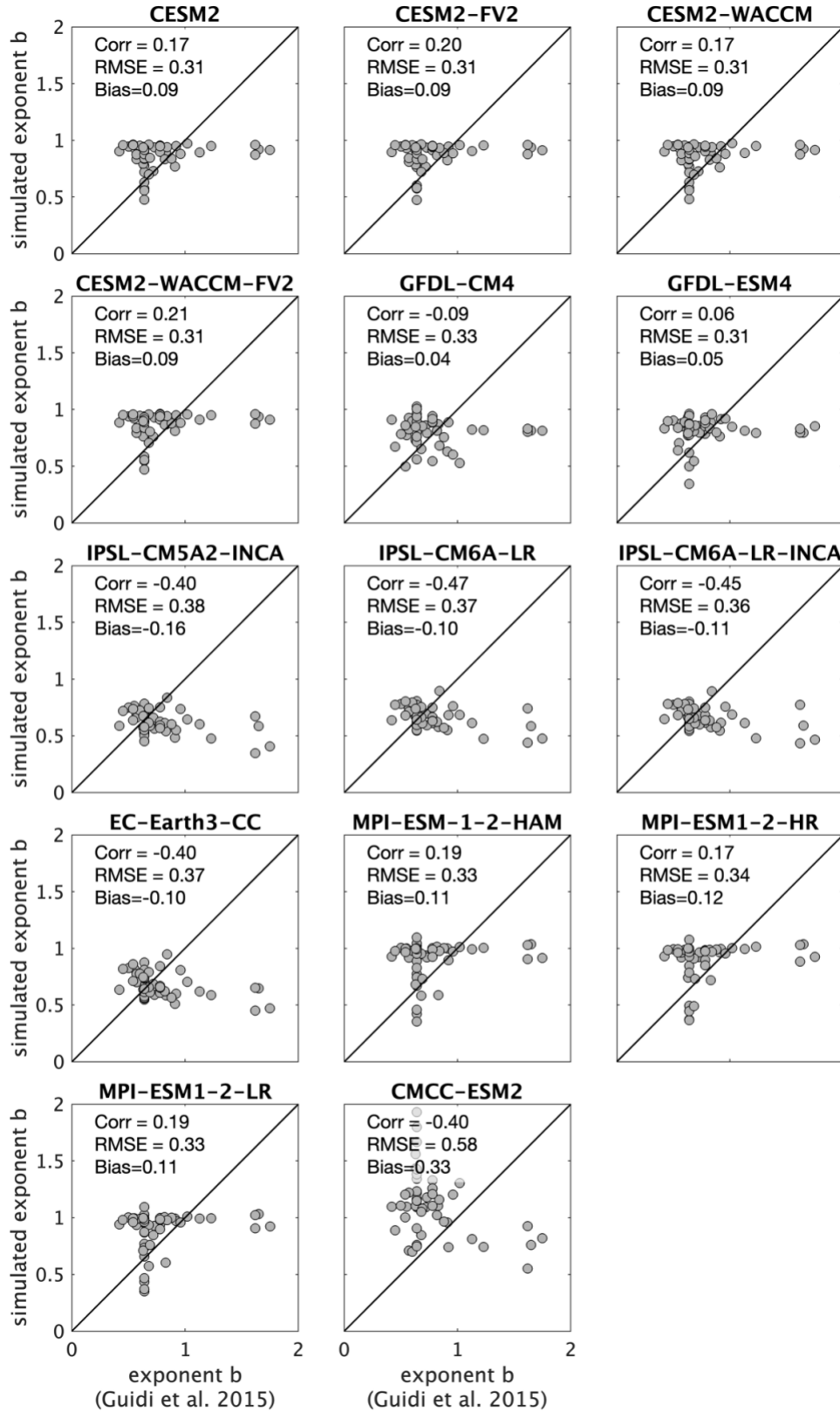


Figure S7. Comparisons of the regionalized exponent b which is simulated by CMIP6 models with estimates from ref.²

Figure 5 – UF, UEP, and UTE should be defined in the caption.

Response: Revised as suggested (changed text is highlighted in **bold** font):

*“**Figure 3. Partitioning of model uncertainty in the vertical carbon flux (U_F) as a function of depth (a) and at 2000 m (b) into the model uncertainty of carbon export (U_{EP}), the model uncertainty of transfer efficiency (U_{TE}), and the co-dependence between carbon export and transfer efficiency ($U_{EP\&TE}$) as described in Methods. The horizontal dashed black lines represent the crossover depth where the uncertainty of transfer efficiency (U_{TE}) starts to dominate the model uncertainty in the vertical carbon flux (U_F).**”*

234 – Why is this true “especially in the North Atlantic and the North Pacific”? More clarification is necessary.

Response: Thanks for pointing out this. This should be in the North Atlantic and the Arctic Ocean. As we discussed in P10 L222-223, the model uncertainty of carbon export is high and remains dominant above 1900 m and 1700 m in these two biomes, respectively. We have corrected this and added explanations in our revised manuscript (changed text is highlighted in **bold** font):

*“..... the model uncertainty of the carbon export is more important when studying mesopelagic ecosystems especially in the North Atlantic and **the Arctic Ocean where the model uncertainty of carbon export is dominant above 1900 and 1700 m, respectively (Figure 3).**”*

241 – this assertion is incorrect – sequestration horizons are provided in Nowicki et al., 2022: <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2021GB007083>

Response: Thanks for the very helpful comment. We have removed this sentence and revised our manuscript (changed text is highlighted in **bold** font):

*“With respect to long-term carbon sequestration and global climate, the relative importance **of model uncertainty between carbon export and transfer efficiency** depends on the competition between the crossover depth and the sequestration horizon, **i.e.** the depth where carbon is sequestered for climate-relevant timescales (e.g., >100 years). Typically, the deeper the organic carbon is respired, the longer it is sequestered^{28,29}. The 1000 and 2000 m **depths have been** widely used as static horizons^{5–7,15,23,24}. **However, the sequestration horizon is spatially variable. There have been some published global maps about the carbon sequestration time²⁷ and mean residence time at each depth³⁰. A recent study suggests that carbon sequestration occurs throughout the water column and that about half of the sequestration is above 1000 m³¹. On the global scale, the crossover depth of 900 m underscores the importance of both model uncertainty of carbon export and transfer efficiency for understanding the century-scale sequestration. The importance of model uncertainty of transfer efficiency is expected to increase when considering longer-term sequestration.**”*

260 – It is not at all clear to me how the authors are able to distinguish between structural and parametric uncertainty. Looking through the methods, it seems that “structural” is being defined as the the Martin shape, and “parametric” as the deviation from that shape, or vice versa, but it

seems that the term “TEdeviation” is defined towards the separation whereas “TEcorrection” is used in the definition of the structural uncertainty component, so it is not clear how all the terms are related. In any case, while I think it is potentially useful to frame the distinction between aspects of the transfer efficiency that are representable with the Martin curve in shape and its modulation through the b exponent versus that which is not, the authors need to explain what these definitions mean at the beginning of the section for it to be helpful.

Response: Thanks for pointing out this. We are sorry for a mistake in the previous version caused by a mix-up of different manuscript versions. It should be $TE_{deviation}$ in equations 13 and 14:

$$U_{II} = var(\log_{10}(\overline{TE_{deviation}}))$$

$$U_{I\&II} = 2 \times var(\log_{10}(\overline{TE_{martin}}), \log_{10}(\overline{TE_{deviation}}))$$

In addition, we have renamed the parametric (U_p) and structural (U_s) uncertainty to type I (U_I) and type II (U_{II}) uncertainty. Type I is the uncertainty of transfer efficiency that is representable with the Martin curve by changing the exponent b. Type II is the uncertainty that cannot be represented by changes in b. The motivation, definition, as well as interpretation and discussion of these two types of model uncertainty have been rephrased in our revised manuscript. Detailed revisions in our manuscript can be seen in our response to a similar comment above. Finally, definitions are also explained in the beginning of the section 2.4, as suggested (changed text is highlighted in **bold** font):

*“As shown in Figure 5a, all sources of uncertainty grow with depth. **The type I uncertainty (can be represented by variations in the exponent b of the Martin curve)** is initially the largest source of uncertainty before **the type II uncertainty (not representable by variations in b)** takes over.”*

294 – Bringing up the constraints from nutrient and oxygen data begs the question of why the authors did not include these comparisons with observations as a means of reducing uncertainty... unfortunately, such an analysis is made extremely difficult as these models have radically different initialization/spinup procedures as detailed in Seferian et al., 2016 (<https://gmd.copernicus.org/articles/9/1827/2016/>) for CMIP5 and Seferian et al., 2020 (<https://link.springer.com/article/10.1007/s40641-020-00160-0>) for CMIP6. The authors should point out that a more consistent protocol for model initialization/spinup would provide much better opportunity for model assessment in the future.

Response: Thanks for the suggestion. We have added this statement and discussion in the section 2.5 in our revised manuscript (new text):

“..... However, assessing model uncertainty in simulating the BCP through comparisons with these biogeochemical measurements is made difficult by the fact that the CMIP6 models are initialized and spun up following different procedures³³. A consistent protocol for model initialization and spin-up would provide better opportunities for model assessment.”

Reviewer #3 (Remarks to the Author):

Review of Wang & Fennel “Organic matter transfer efficiency is the largest uncertainty for biological carbon pump in CMIP6 models”

In this manuscript, Wang & Fennel analyze earth system model output from the latest CMIP simulations to identify processes that result in uncertainty in the strength of the biological pump. Above 900m, carbon export by sinking particles is the greatest source of uncertainty, whereas, below 900m, uncertainty is related to transfer efficiency, related to the rate of sinking particle remineralization. There are three groups of models, closely related to which biogeochemistry package is used: (1) with largely homogeneous transport efficiency (MARBL/HAMOCC) that is independent of temperature; (2) low transport efficiency in the subtropical gyres and high transport efficiency in the tropics and high latitudes (PISCES/COBALT) due to temperature dependence; and (3) high transport efficiency in the subtropical gyres and low transport efficiency in the tropics and high latitudes (BLING) due to ballasting. Either contrasting pattern from group (2) or (3) can be justified depending on which observations are considered. Introducing variable reactivity distributions for sinking particles flips the PISCES model from (2) to (3), suggesting that distinguishing between different size and composition particles is important for simulating the future response of the biological pump to climate changes.

I thought this manuscript was interesting, well-written, and easy to follow. I only have a few suggestions:

Response: We are grateful for the positive comments.

(1) The authors make a prominent point of defining biomes for their analysis (Figure 1), but don't really apply them to the principal analysis. I think the discussion of results in section 2.1 would benefit from a box-and-whisker-plot for POC flux in each biome, and at each depth, with the location of each model indicated with a symbol. Similarly Figure 3. I am not really a fan of these “postage stamp” plots, I would prefer the authors did something more visual to highlight the patterns they are trying to emphasize, rather than making the reader scrutinize ~30 little panels each time.

Response: Following this suggestion, we have added new figures to replace our original Figures 2 and 3 in our revised manuscript and moved the original Figures 2 and 3 into the supplement. We have grouped the simulated transfer efficiency by the 14 CMIP6 models into 3 categories in Figure 2 to highlight the different global patterns of transfer efficiency.

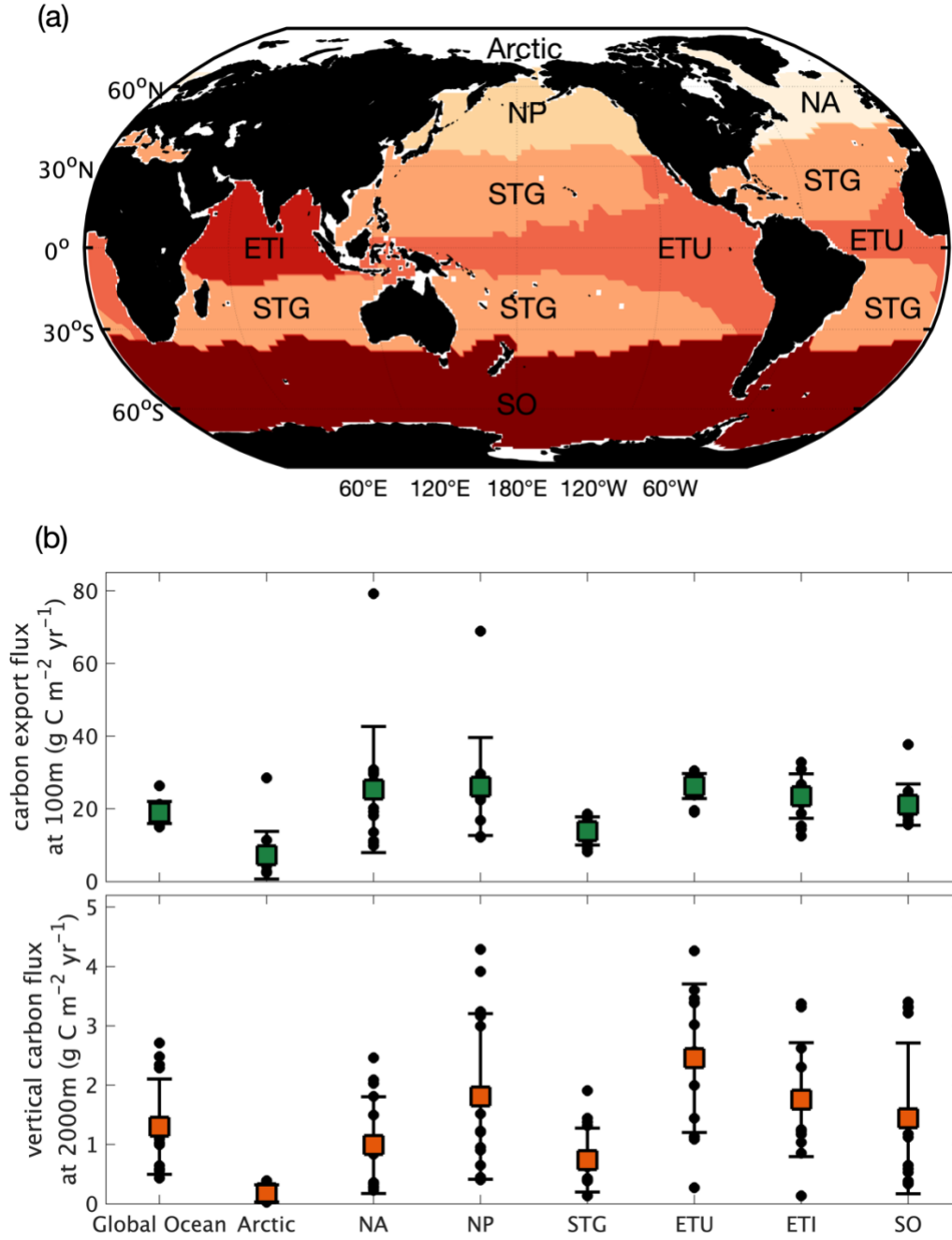


Figure 1. Map of biogeochemical biomes (a) and the vertical carbon flux averaged over each biome at 100 m and 2000 m (b). Errorbars represent the mean and standard deviation of 14 CMIP6 models and black dots represent each single model. The global distributions of all models are shown in Figure S1. Abbreviations: Southern Ocean (SO), Equatorial Tropical Indian (ETI), Equatorial Tropical Upwelling regions (ETU), Subtropical Gyres (STG), North Pacific (NA), North Atlantic (NA), and Arctic Ocean (Arctic).

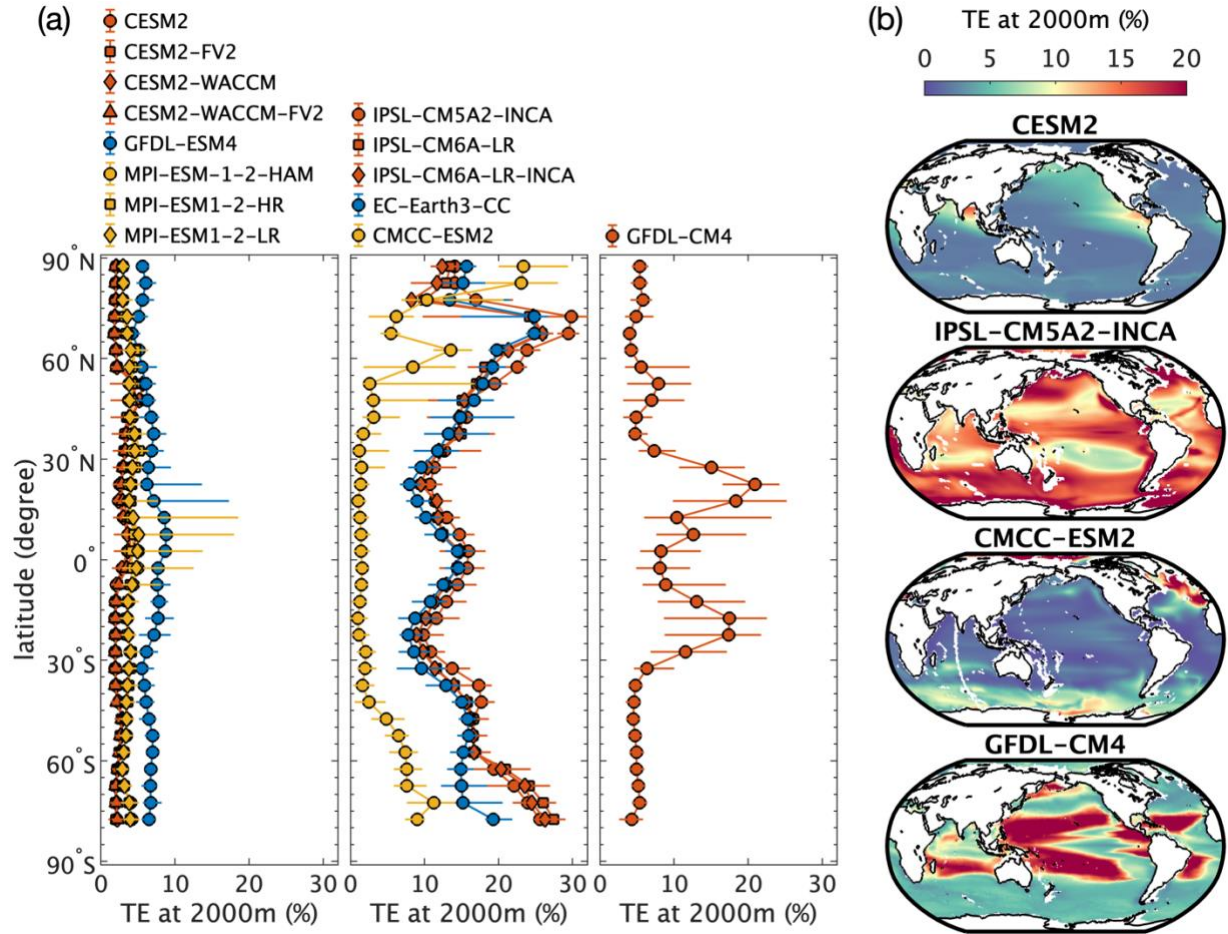


Figure 2. Three different latitudinal patterns of simulated transfer efficiency at 2000 m are displayed by the CMIP6 models (a). Error bars represent the interquartile range within a 5-degree latitudinal band. The global distribution of transfer efficiency at 2000 m from 4 representative models is shown in panels (b). The global distributions of all models are shown in Figure S2.

(2) Which group does the CMCC-ESM2/BFM model fall in to? Line 128 indicates “no distinct latitudinal pattern” and it is dismissed, but it looks like there is something interesting going on between the Southern Ocean and other regions. Does BFM have any particular mechanisms in common, or different from the three overall groups?

Response: The CMCC-ESM2 model falls into the same group as the four PISCESv2 ESMs. However, unlike four PISCESv2 ESMs, the transfer efficiency in CMCC-ESM2 is low in the North Pacific and near the equator. The CMCC-ESM2 is different from all of the other 13 CMIP6 models because it is the only model which simulates the utilization of organic particles by bacteria explicitly. When bacteria are abundant, the utilization rate of organic particles is dependent on the nutritional content of organic particles. In contrast, when bacterial abundance is low, the utilization rate is dependent on temperature and bacterial biomass. In addition, the oxygen dependence can be activated/deactivated depending on the parameterization used. We have included this information in our revised manuscript (new text):

“In CMCC-ESM2, the latitudinal patterns of transfer efficiency are similar to the four PISCESv2 ESMs. However, the transfer efficiency is low in the North Pacific and Equatorial Tropical Upwelling biomes. Unlike other models analyzed here, the utilization of organic particles by bacteria is simulated explicitly in CMCC-ESM2. It is dependent on the nutritional content of organic particles when bacteria are abundant and otherwise on temperature and bacterial biomass²⁵. Although the bacterial utilization of organic particles can be modulated by oxygen, it appears that this oxygen-dependence is deactivated in the CMCC-ESM2 because transfer efficiency in OMZs is not elevated. In general, the simulated transfer efficiency at 2000 m in the CMCC-ESM2 is comparable to the inverse model²³ and the data-assimilative ecosystem model²⁴ except that the transfer efficiency from CMCC-ESM2 is much higher in the North Atlantic and Arctic Ocean.”

Minor comments/clarifications:

Line 66-67: Cael & Lauderdale (2021) also found a ~20-25uatm sensitivity of changing “b” values in a biogeochemical model.

Response: Thank you for pointing this out. Indeed, a sensitivity of -21.6/+24.8 uatm (which combines to 46 uatm) was identified by changing the exponent b from 0.70 to 0.98. We have added this into our revised manuscript as suggested (changed text is highlighted in **bold** font):

*“A recent estimate of the exponent b is 0.84 ± 0.14 ¹³, which is equivalent to a range in transfer efficiency from 5% ($b=0.98$) to 12% ($b=0.70$) at 2000 m **and corresponds to an uncertainty in equilibrium atmospheric CO₂ levels of 28^{13,14} or 46⁹ μatm depending on the climate model used”***

Line 70: Both export and transfer efficiency have structural and parametric uncertainty, the authors only make the distinction for transfer efficiency in this paper.

Response: Yes, this is correct. However, based on comments by other reviewers, we have re-defined the parametric (U_p) and structural (U_s) uncertainty as type I (U_I) and type II (U_{II}) uncertainty. The former is the model uncertainty of transfer efficiency that can be represented with the Martin curve by changing the exponent b while the latter is the remaining uncertainty which cannot be represented that way. Since the Martin curve and its exponent b have been widely used to extrapolate and compare the transfer efficiency, we believe that this distinction will be helpful to discuss whether the usage of Martin curve is appropriate. The motivation, definition, as well as interpretation and discussion of these two types of model uncertainty have been rephrased in our revised manuscript (see responses to the other reviewers for details).

Line 131: Add “(high)” for “b” in the subtropical gyres.

Response: We have decided to remove the parenthetical references to b throughout the manuscript for the sake of simplifying the text. We only describe the patterns in the transfer efficiency, since we state at the beginning that high transfer efficiency corresponds to low b, and vice versa.

Line 170-176: “PISCESv2 simulates two particle size classes with different sinking velocities”: yes, but why did it not produce a similar result to GFLD-CM4/BLING or PISCESv2_RC?

Response: For reasons why patterns in PISCESv2 are different from GFDL-CM4/BLING see our discussion in section 2.2 (changed text is highlighted in **bold** font):

*“In the three IPSL models **and EC-Earth3-CC**, the transfer efficiency at 2000 m is high in the productive **high-latitude regions and the Equatorial Tropical Upwelling biome** but low in the Subtropical Gyres. **Global patterns in the three IPSL models and EC-Earth3-CC are due to the temperature-dependent remineralization rate and the differentiation of sinking velocities between two size classes of organic particles in their biogeochemical models, i.e., PISCESv2²⁰**. In productive regions where diatoms are dominant, the **exported** organic particles tend to be large and fast sinking because diatoms aggregate efficiently.*

*.....
In contrast, in GFDL-CM4, the transfer efficiency at 2000 m is high in the Subtropical Gyres and low in the productive regions. **This is because the ballast effect of CaCO₃ outweighs the temperature dependence of the remineralization rate and the ballast effect by opal is missing in this model¹⁸**.”*

With respect to differences between PISCESv2 and PISCESv2_RC, in our original manuscript we attributed these to the reactivity spectrums of organic particles. However, we double checked the relevant paper (Aumont et al., 2017) and found that the two simulations with and without the reactivity spectrums produced similar global patterns of the exponent b (see Figure 10 from Aumont et al. (2017) inserted below). The authors suggested that the zooplankton activity was a major driver for the high exponent b in the productive regions by running a sensitivity simulation without flux feeding by zooplankton (see Figure S3 from Aumont et al. (2017) inserted below). Therefore, we have removed the paragraph (P7-8 L150-161) from our original manuscript and changed the text to (new text):

“In addition, zooplankton activity can be another reason as high mesopelagic zooplankton abundance can be sustained by carbon export in productive regions and in turn can attenuate the vertical carbon flux efficiently¹⁶.”

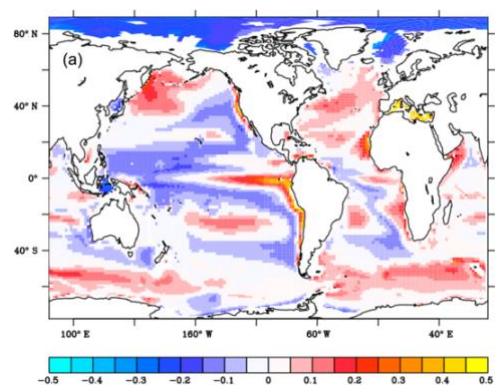


Figure 10 from Aumont et al., (2017). Anomalies of exponent b from the global median value from simulations without (a) and with (b) inclusion of reactivity spectrums of organic particles.

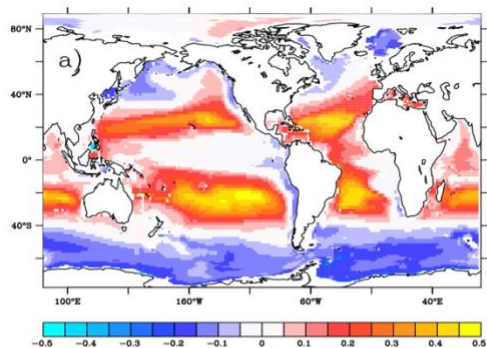
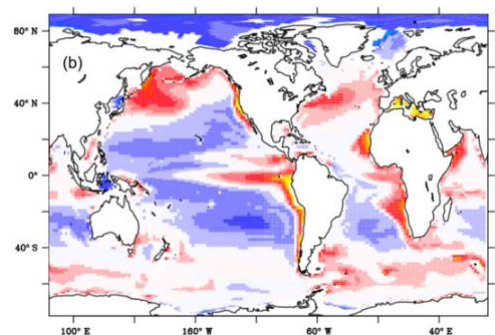


Figure S3 from Aumont et al., (2017). Anomalies of exponent b from the global median value from simulations without flux feeding by zooplankton (a) or reactivity spectrums of organic particles (b)

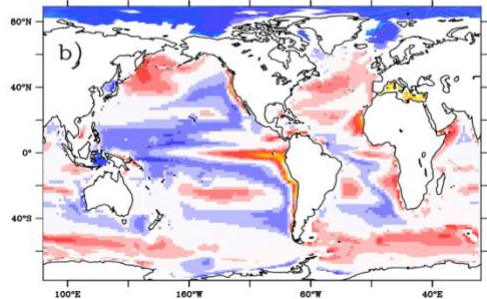


Figure 4: did not calculate “b”.

Response: As stated in the previous response, we have removed this figure in our revised manuscript.

Line 240: Not exactly a global sequestration depth map (depends on how long you consider “sequestered”), but Hill et al. (2004, 10.1029/2002JC001598) have maps of residence times of carbon at different depths in the ocean from a model.

Response: Thanks for the very helpful suggestion. We have removed this sentence and revised our manuscript as follows (changed text is highlighted in **bold** font):

*“With respect to long-term carbon sequestration and global climate, the relative importance of **model uncertainty between carbon export and transfer efficiency** depends on the competition between the crossover depth and the sequestration horizon, **i.e.** the depth where carbon is sequestered for climate-relevant timescales (e.g., >100 years). Typically, the deeper the organic carbon is respired, the longer it is sequestered^{28,29}. The 1000 and 2000 m **depths have been** widely used as static horizons^{5–7,15,23,24}. **However, the sequestration horizon is spatially variable. There have been some published global maps about the carbon sequestration time²⁷ and mean residence time at each depth³⁰. A recent study suggests that carbon sequestration occurs throughout the water column and that about half of the sequestration is above 1000 m³¹. On the global scale, the crossover depth of 900 m underscores the importance of both model uncertainty of carbon export and transfer efficiency for understanding the century-scale sequestration. The importance of model uncertainty of transfer efficiency is expected to increase when considering longer-term sequestration.**”*

Line 278: “structural uncertainty can be significant” see also Lauderdale & Cael (2021).

Response: Yes, thank you. We have added a reference to Lauderdale & Cael (2021) (new text):

“Our results show that the model uncertainty of transfer efficiency is underestimated by using the Martin curve and exponent b . This is consistent with findings in ref.⁹ where six simple alternative parameterizations were compared with the Martin Curve.”

Line 408 & 409: Define TE_correction. Is that related to TE_deviation?

Response: Sorry for mistakes caused by a mixing of different manuscript versions. It should be $TE_{deviation}$ in equations 13 and 14:

$$U_{II} = var(\log_{10}(\overline{TE_{deviation}}))$$
$$U_{I\&II} = 2 \times var(\log_{10}(\overline{TE_{martin}}), \log_{10}(\overline{TE_{deviation}}))$$

As we replied above, we have re-defined the parametric (U_P) and structural (U_S) uncertainty to type I (U_I) and type II (U_{II}) uncertainty.

Partitioning uncertainty of the ocean's biological carbon pump in CMIP6 models

Bin Wang^{1*}, Katja Fennel¹

¹Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, Canada

Corresponding author: Bin Wang (Bin.Wang@dal.ca)

Abstract

Model uncertainty in simulating the biological carbon pump was quantified and partitioned using 14 models from the Coupled Model Intercomparison Project Phase 6 (CMIP6). Uncertainty increases with depth. **On the global scale**, uncertainty in carbon export **dominates** above 900 m **and** uncertainty in transfer efficiency below. Reducing **model** uncertainty in **carbon export and** transfer efficiency offers **similar benefits** for understanding **century-scale carbon sequestration** and climate. **The CMIP6 models produce three different qualitative patterns in transfer efficiency: one where it is globally homogenous and two opposite latitudinal patterns due to different model structures and parameters. The exponent b of the Martin curve, which has long been used to compare different representations of transfer efficiency, is shown here to underestimate uncertainty in transfer efficiency. This highlights the importance of using vertical profiles of transfer efficiency rather than the single exponent b in model validation and intercomparison exercises.**

1. Introduction

The ocean's biological carbon pump (BCP) is a key mechanism for the ocean's uptake and sequestration of CO₂ from the atmosphere. It comprises the fixation of CO₂ through photosynthesis in the surface ocean and the subsequent transport of organic carbon into the deep ocean where it is remineralized and sequestered. Therefore, accurate estimation of the BCP is important for understanding the ocean's role in the global carbon cycle and for predicting how it may be changing as climate warms.

Earth System Models (ESMs) are considered the best tools for predicting the ocean's response to climate change on **decadal** to centennial timescales. Under global warming, ESMs project a decline in carbon export, i.e. the carbon flux leaving the euphotic zone, over the 21st century¹⁻³. However, there are large variations in the underlying drivers and magnitude of the decline among

the different ESMs which implies that uncertainty in these estimates is large. The model-to-model difference is larger than the projected reductions in carbon export during the 21st century¹. Therefore, it is necessary to partition and diagnose the different factors that contribute to model uncertainty in simulating the BCP and to reduce it.

The BCP can be assessed by quantifying carbon export out of the euphotic zone and the depth-dependent transfer efficiency⁴⁻⁸, which measures the fraction of the exported carbon that survives respiration and reaches the deep ocean. Given the complex mechanisms controlling the transfer efficiency, simple parameterizations with different mechanistic underpinnings have been used to represent and to compare the transfer efficiency^{9,10}. The Martin curve¹¹, which normalizes the vertical carbon flux to a reference depth, e.g., 100 m, through the power-law equation $F = F_0 \times (z/z_0)^{-b}$, is the most widely used parameterization¹²⁻¹⁶. The exponent b is an indicator of the carbon flux attenuation with depth: the larger the exponent b , the faster the attenuation and hence the lower the transfer efficiency. A recent estimate of the exponent b is 0.84 ± 0.14 ¹³, which is equivalent to a range in transfer efficiency from 5% ($b=0.98$) to 12% ($b=0.70$) at 2000 m and corresponds to an uncertainty in equilibrium atmospheric CO₂ levels of 28^{13,14} or 46⁹ μatm depending on the climate model used. However, this neglects the uncertainty resulting from the choice of the Martin curve over a different parameterization, which has been estimated to contribute an additional uncertainty in atmospheric CO₂ levels of 12 μatm ⁹. Rather than using the simple parameterizations which are compared in ref.⁹, current ESMs simulate the transfer efficiency in a more mechanistic manner by resolving some key controlling processes, such as the ballast protection of organic matter¹⁷⁻¹⁹, temperature-¹⁸⁻²⁰ and oxygen-dependent¹⁷⁻²¹ remineralization, and aggregation/disaggregation²⁰. The extent to which the uncertainty in transfer efficiency in state-of-the-art ESMs can be represented by changing the exponent b of the Martin curve, or not, warrants further evaluation.

In this study, the model uncertainty in simulating the BCP is quantified by the variance of averaged vertical carbon flux over different spatial scales across ESMs. We distinguish between the model uncertainty within the euphotic zone and below by partitioning it into the model uncertainty of carbon export and the model uncertainty of transfer efficiency. Motivated by the wide application of the Martin curve, we further decomposed the latter into two components: the uncertainty which can be represented by changing the exponent b of the Martin curve (referred to here as type I) and the uncertainty which cannot (referred to as type II). To achieve this goal, we chose the 14 ESMs that are used by the Coupled Model Intercomparison Project Phase 6 (CMIP6)

and report depth-resolved vertical carbon flux (Table 1). We averaged their estimates over the global ocean and different biological biomes prior to the uncertainty analysis (Figure 1a). A caveat is that the uncertainty analyzed is sampled over the realized diversity of the selected CMIP6 models which do not necessarily span the real uncertainty range. However, our results indicate which aspects of ESMs should be prioritized for improvement to reduce model uncertainty and should guide future observational strategies.

Table 1. List of the 14 CMIP6 models compared in this study and the subset of 6 models, shown in bold, that are used in the uncertainty analysis. Acronyms: National Center for Atmospheric Research, USA (NCAR), NOAA Geophysical Fluid Dynamics Laboratory, USA (GFDL), Institut Pierre Simon Laplace, France (IPSL), Max Planck Institute for Meteorology, Germany (MPI-M), Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici, Italy (CMCC)

Institute	Model name	Ocean BGC model	Model Grid	Output frequency
NCAR	CESM2	MARBL	320×384×60	monthly
	CESM2-FV2	MARBL	320×384×60	monthly
	CESM2-WACCM	MARBL	320×384×60	monthly
	CESM2-WACCM-FV2	MARBL	320×384×60	monthly
GFDL	GFDL-CM4	BLINGv2	1440×1080×75	monthly
	GFDL-ESM4	COBALTv2	720×576×75	monthly
IPSL	IPSL-CM5A2-INCA	PISCESv2	182×149×31	monthly
	IPSL-CM6A-LR	PISCESv2	362×332×75	monthly
	IPSL-CM6A-LR-INCA	PISCESv2	362×332×75	monthly
EC-Earth consortium	EC-Earth3-CC	PISCESv2	362×292×75	yearly
HAMMOZ-Consortium	MPI-ESM-1-2-HAM	HAMOCC6	256×220×40	monthly
MPI-M	MPI-ESM1-2-HR	HAMOCC6	802×404×40	monthly
	MPI-ESM1-2-LR	HAMOCC6	256×220×40	monthly
CMCC	CMCC-ESM2	BFM5.2	362×292×50	yearly

2. Results and discussion

2.1 Vertical carbon flux from CMIP6 models

As shown in Figures 1b and S1, the globally integrated carbon export flux at 100 m varies by a factor of 1.8 from 5.4 Pg C yr⁻¹ (15 g C m⁻² yr⁻¹) in CMCC-ESM2 to 9.2 Pg C yr⁻¹ (26 g C m⁻² yr⁻¹) in GFDL-CM4. At 2000 m, the model-to-model variations increase to a factor of 6.2. On biome and local scales, the carbon export flux at 100 m is consistently low in the Subtropical Gyres and high in the Equatorial Tropical Upwelling across different CMIP6 models. Nevertheless, the model-to-model variability in these two biomes reaches factors of 2.3 and 1.6 at 100 m and increases to factors of 14 and 16 at 2000 m, respectively. In addition, there are large disagreements in the spatial patterns across biomes between the different CMIP6 models. At 100 m, the North Atlantic and the North Pacific tend to be the two biomes with the highest carbon export flux in the 4 CESM2 and 2 GFDL models and especially in GFDL-CM4. However, in the 3 MPI models, the average carbon export flux in these two biomes is even lower than in the Subtropical Gyres. As a result, the model-to-model variability in carbon export at 100 m reaches factors of 8.0 and 5.7 in these two biomes. Unlike the Subtropical Gyres and the Equatorial Tropical Upwelling region, the North Atlantic and the North Pacific experience a relatively mild increase of model-to-model variability at 2000 m. In the Southern Ocean, the averaged carbon export flux at 100 m is relatively high in the 2 GFDL models, the 3 IPSL models, EC-Earth3-CC, and CMCC-ESM2. However, the high carbon export flux is confined to the Subantarctic Ocean in the 3 MPI models, and the average carbon export flux is even lower than in the Subtropical Gyres in the 4 CESM2 models. As a result, the model-to-model variability in the Southern Ocean is about a factor of 2.4 at 100 m and grows to a factor of 10 at 2000 m.

In summary, there are substantial differences in the simulated vertical carbon flux across different CMIP6 models on biome and local scales. In addition, the model-to-model variability in vertical carbon flux at 2000 m is much larger than that of the carbon export at 100 m, which suggests large discrepancies in transfer efficiency across different CMIP6 models.

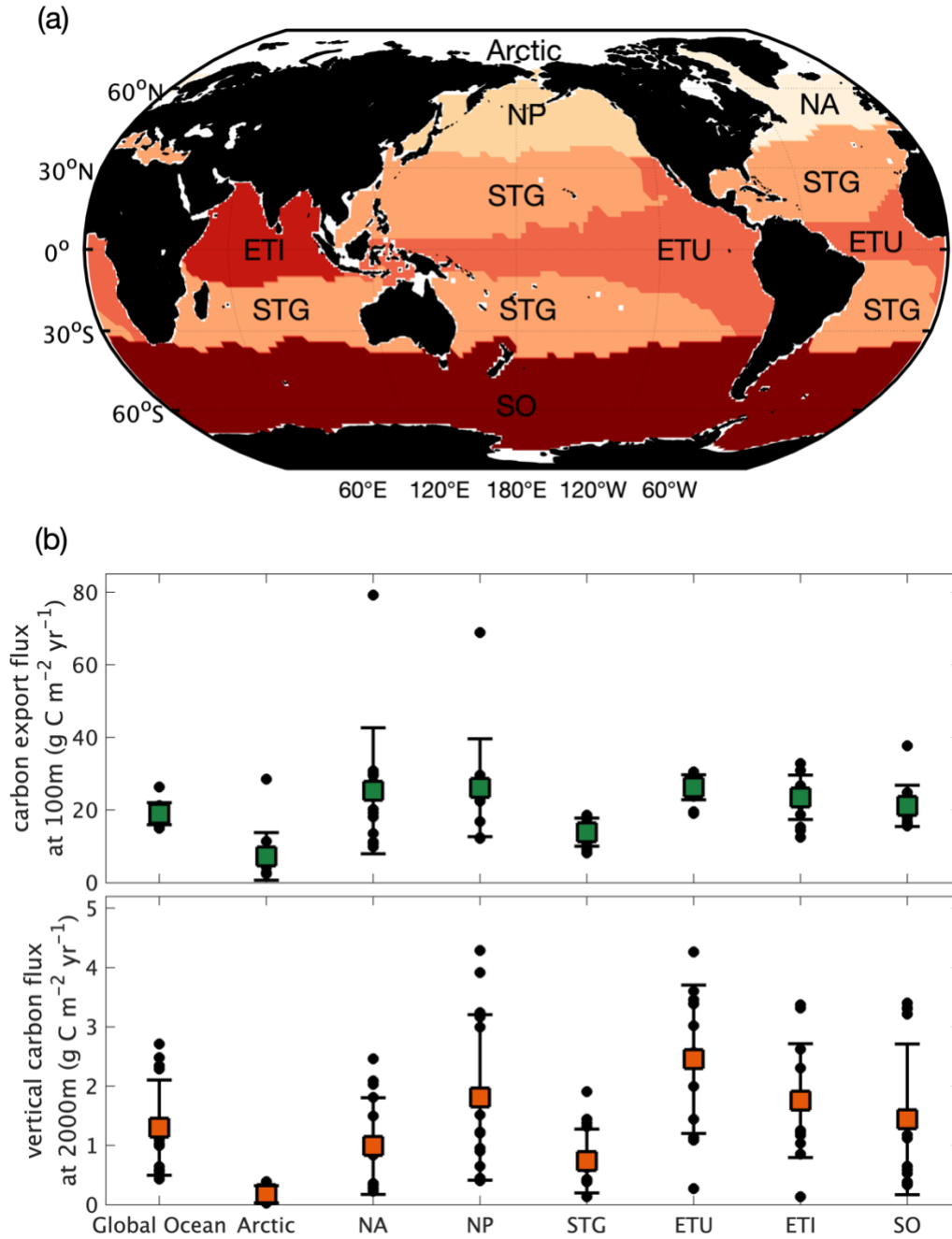


Figure 1. Map of biogeochemical biomes (a) and the vertical carbon flux averaged over each biome at 100 m and 2000 m (b). Errorbars represent the mean and standard deviation of 14 CMIP6 models and black dots represent each single model. The global distributions of all models are shown in Figure S1. Abbreviations: Southern Ocean (SO), Equatorial Tropical Indian (ETI), Equatorial Tropical Upwelling regions (ETU), Subtropical Gyres (STG), North Pacific (NA), North Atlantic (NA), and Arctic Ocean (Arctic).

2.2 Transfer efficiency from CMIP6 models

The global median of the exponent b varies from 0.63 in IPSL-CM5A2-INCA to 1.01 in CMCC-ESM2. This is higher than the estimates of 0.64 and 0.54 from observations in refs.^{5,15}. However, the global average of the exponent b ranging from 0.70 to 1.03 is consistent with a more recent data-constrained model estimate from 0.70 to 0.98¹³. The simulated transfer efficiency at 2000 m can be classified into three, qualitatively different latitudinal patterns (Figures 2 and S2): the first is a homogenous transfer efficiency, the other two are opposite latitudinal patterns, either high in the oligotrophic Subtropical Gyre and low in productive regions or vice versa.

The simulated transfer efficiency at 2000 m is homogenous in the four CESM2 models, the three MPI models, and GFDL-ESM4 across latitudes except for the elevated transfer efficiency in Oxygen Minimum Zones (OMZs). This is because the remineralization rate is suppressed under low oxygen^{17,19,22} in these biogeochemical models (Tables 1-2). In GFDL-ESM4, the remineralization rate is dependent on temperature. The high temperature in low-latitude regions facilitates remineralization and should reduce the transfer efficiency. However, this effect is compensated for by the ballast effect of elevated CaCO_3 and lithogenic fluxes in low latitude regions, which enhances the transfer efficiency by protecting organic matter from remineralization or by accelerating the sinking velocity¹⁹. In the four CESM2 models, the ballast effect of CaCO_3 and opal are both simulated and may cancel each other out flattening the latitudinal gradient of transfer efficiency¹⁷, because the former is predominant in the Subtropical Gyres and the latter in productive regions. No temperature dependence is simulated in the four CESM2 and three MPI models^{17,21,22}.

In the three IPSL models and EC-Earth3-CC, the transfer efficiency at 2000 m is high in the productive high-latitude regions and the Equatorial Tropical Upwelling biome but low in the Subtropical Gyres. These broad qualitative features can also be found in global patterns derived from free-drifting neutrally buoyant sediment traps (NBSTs)¹², an inverse model²³, and a data-assimilative ecosystem model²⁴, which have also reported the lowest transfer efficiency in Subtropical Gyres. However, the transfer efficiency is not elevated in the Equatorial Tropical Upwelling region in the NBST-derived global estimate¹². Global patterns in the three IPSL models and EC-Earth3-CC are due to the temperature-dependent remineralization rate and the differentiation of sinking velocities between two size classes of organic particles in their

biogeochemical models, i.e., PISCESv2²⁰. In productive regions where diatoms are dominant, the exported organic particles tend to be large and fast sinking because diatoms aggregate efficiently. This is consistent with findings from an inverse model analysis suggesting that the transfer efficiency is tightly correlated with changes in temperature and phytoplankton community structure²³. We therefore use the empirical relationship from ref²³ to re-calculate their transfer efficiency at 1000 m for comparison with the CMIP6 models. The ballast effects are neglected in this analysis because model output of the SiO₂:POC ratio in ref²³ is not available. This underestimates their transfer efficiency in the Antarctic Ocean but has minor impacts elsewhere. Comparisons show that the transfer efficiency at 1000 m simulated by the four PISCESv2 ESMs is in general higher than that in the inverse model²³ (Figures S3-S5). The simulated transfer efficiency at 2000 m from 10.53±0.83% to 17.99±1.26% by the four PISCESv2 ESMs is also higher than the 5% to >8% range from the data-assimilative model²⁴.

In CMCC-ESM2, the latitudinal patterns of transfer efficiency are similar to the four PISCESv2 ESMs. However, the transfer efficiency is low in the North Pacific and Equatorial Tropical Upwelling biomes. Unlike other models analyzed here, the utilization of organic particles by bacteria is simulated explicitly in CMCC-ESM2. It is dependent on the nutritional content of organic particles when bacteria are abundant and otherwise on temperature and bacterial biomass²⁵. Although the bacterial utilization of organic particles can be modulated by oxygen, it appears that this oxygen-dependence is deactivated in the CMCC-ESM2 because transfer efficiency in OMZs is not elevated. In general, the simulated transfer efficiency at 2000 m in the CMCC-ESM2 is comparable to the inverse model²³ and the data-assimilative ecosystem model²⁴ except that the transfer efficiency from CMCC-ESM2 is much higher in the North Atlantic and Arctic Ocean.

In contrast, in GFDL-CM4, the transfer efficiency at 2000 m is high in the Subtropical Gyres and low in the productive regions. This is because the ballast effect of CaCO₃ outweighs the temperature dependence of the remineralization rate and the ballast effect by opal is missing in this model¹⁸. General patterns of the fitted exponent *b* between the productive and oligotrophic regions are similar to two observational estimates^{5,15} (Figures S2-S3) despite regional discrepancies in transfer efficiency (Figures S6-S7). For instance, the elevated exponent *b*, which is equivalent to reduced transfer efficiency, near the equator is missing in ref.⁵ because it cannot be resolved by sparse concurrent measurements of carbon export and deep vertical carbon flux. In addition, the variation of *b* in GFDL-CM4 (0.50-1.03) is smaller than that of 0.36-1.09 and 0.42-

1.75 from refs.^{5,15} (Figure S3). Aside from ballast effects, this global pattern has also been suggested to be driven by particle composition because of the elevated relative contribution from fast-sinking fecal pellets in oligotrophic regions^{5,26,27}. Although production of aggregates by phytoplankton and fecal pellets by zooplankton are formulated in the 14 CMIP6 models, there is no differentiation between their remineralization rates and sinking velocities. In addition, zooplankton activity can be another reason as high mesopelagic zooplankton abundance can be sustained by carbon export in productive regions and in turn can attenuate the vertical carbon flux efficiently¹⁶.

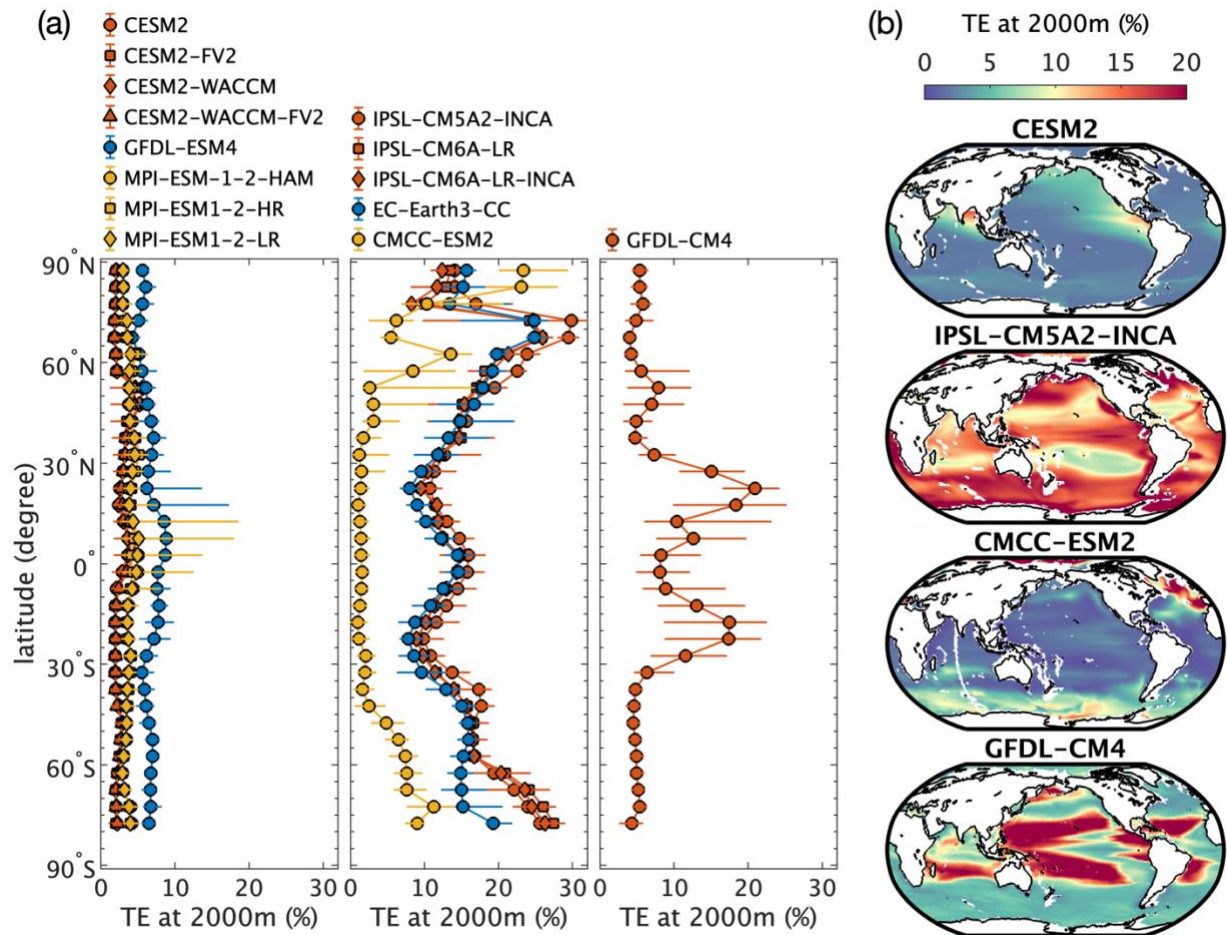


Figure 2. Three different latitudinal patterns of simulated transfer efficiency at 2000 m are displayed by the CMIP6 models (a). Error bars represent the interquartile range within a 5-degree latitudinal band. The global distribution of transfer efficiency at 2000 m from 4 representative models is shown in panels (b). The global distributions of all models are shown in Figure S2.

Table 2. Mechanisms included in different biogeochemical models to simulate the vertical attenuation of vertical carbon flux

Ocean BGC model	Ballast effect			Temperature dependence	Oxygen dependence	Bacteria utilization
	CaCO ₃	Opal	dust			
MARBL ¹⁷	Yes	Yes	Yes	No	Yes	Implicitly
BLINGv2 ¹⁸	Yes	No	Yes	Yes	Yes	Implicitly
COBALTv2 ¹⁹	Yes	Yes	Yes	Yes	Yes	Implicitly
PISCESv2 ²⁰	No	No	No	Yes	Yes	Implicitly
HAMOCC6 ^{21,22}	No	No	No	No	Yes	Implicitly
BFM5.2 ²⁵	No	No	No	Yes	No	Explicitly

The contrasting global patterns of transfer efficiency have different implications for the future response of the BCP to climate change. Warming-enhanced stratification will likely reduce nutrient supply to the surface ocean, leading to an expansion of oligotrophic ecosystems. This may amplify or dampen the currently predicted reduction of transfer efficiency due to increased temperature. Currently, observation-based estimates of transfer efficiency suggest contrasting patterns with one supporting high transfer efficiency in oligotrophic ecosystems^{5,15} and the other favouring the opposite pattern¹². This indicates that existing in-situ observations provide insufficient constraints and suggests that an expansion of concurrent measurements of carbon export and deep vertical carbon flux is necessary to reconcile the contrasting patterns or falsify one of them.

2.3 Model uncertainties of carbon export and transfer efficiency

As shown in Table 1, there are 6 different biogeochemical models that are used by the 14 ESMs analyzed here. Despite other model discrepancies, e.g., resolution, the ESMs that use the same biogeochemical model tend to produce similar magnitudes and global patterns of vertical carbon flux (Figure S1) and transfer efficiency (Figures 2 and S2). Therefore, a subset of 6 representative ESMs, using different biogeochemical models, namely CESM2, GFDL-CM4, GFDL-ESM4, IPSL-CM5A2-INCA, MPI-ESM1-2-HR, and CMCC-ESM2 (Table 1), are used to calculate the different uncertainties that are described in the Methods.

As shown in Figure 3a, the model uncertainty of the vertical carbon flux (U_F) at 100 m is exclusively contributed by the model uncertainty of carbon export (U_{EP}), because at that depth the transfer efficiency is defined as 100% in all CMIP6 models. Thereafter, due to discrepancies in

the transfer efficiency between different CMIP6 models, the model uncertainty of the vertical carbon flux grows quickly with depth following the model uncertainty of transfer efficiency (U_{TE}), which eventually dominates over the model uncertainty of carbon export and the co-dependence between carbon export and transfer efficiency ($U_{EP\&TE}$). The crossover depth is defined as the depth where the model uncertainty of transfer efficiency starts to dominate the model uncertainty of the vertical carbon flux, i.e., it is the largest source of uncertainty and its relative contribution is >50%. On the global scale, the crossover depth is located at 900 m, leading to the dominance of model uncertainty of transfer efficiency at 2000 m which alone can contribute to about 83% of model uncertainty of the vertical carbon flux. On the biome scale, after excluding the Arctic Ocean and the North Atlantic, the crossover depth varies from 500 m in the Equatorial Tropical Upwelling to 1500 m in the Subtropical Gyres, indicating differences in the relative importance of model uncertainties due to carbon export versus transfer efficiency in different biomes. Nevertheless, the model uncertainty of transfer efficiency has become the dominant source of uncertainty at 2000 m, with the relative contributions ranging from 57% in the North Pacific to above 85% in the Southern Ocean and the Equatorial Tropical Upwelling. In the Arctic Ocean and the North Atlantic, the model uncertainty of carbon export is high and remains dominant until 1900 m and 1700 m, respectively. Unlike other biomes, the simulated carbon export and transfer efficiency are weakly anti-correlated ($\text{corr} > -0.5$) across the CMIP6 models and hence their co-dependence counteracts about 23% and 71% of model uncertainty of the vertical carbon flux in these two biomes.

On the local scale, the model uncertainty of the vertical carbon flux at 2000 m is in general dominated by the model uncertainty of transfer efficiency (Figure 3b). The simulated vertical carbon flux at 2000 m is highly uncertain in the Arctic Ocean, the Gulf Stream regions, the western part of equatorial Atlantic Ocean, the subtropical gyre in the South Atlantic Ocean, the OMZs in the equatorial Pacific Ocean and the Indian Ocean, the Pacific Ocean around 40°N and 40°S, and the Indian Ocean around 40°S, primarily because the model uncertainty of transfer efficiency is high in these regions. In addition, the co-dependence can amplify or dampen the model uncertainty of the vertical carbon flux because the simulated carbon export and transfer efficiency can be highly correlated ($\text{corr} > 0.5$) or anti-correlated ($\text{corr} < -0.5$) around the latitude of 40° and in the Subtropical Gyres, respectively (Figures 3b and 4).

The relative importance of the model uncertainty of carbon export versus model uncertainty of transfer efficiency for the BCP varies with the depth scale of interest. Since the BCP provides

energy and material for metabolisms in the mesopelagic zone, the model uncertainty of the carbon export is more important when studying mesopelagic ecosystems especially in the North Atlantic and the Arctic Ocean where the model uncertainty of carbon export is dominant above 1900 and 1700 m, respectively (Figure 3). With respect to long-term carbon sequestration and global climate, the relative importance of model uncertainty between carbon export and transfer efficiency depends on the competition between the crossover depth and the sequestration horizon, i.e. the depth where carbon is sequestered for climate-relevant timescales (e.g., >100 years). Typically, the deeper the organic carbon is respired, the longer it is sequestered^{28,29}. The 1000 and 2000 m depths have been widely used as static horizons^{5–7,15,23,24}. However, the sequestration horizon is spatially variable. There have been some published global maps about the carbon sequestration time²⁷ and mean residence time at each depth³⁰. A recent study suggests that carbon sequestration occurs throughout the water column and that about half of the sequestration is above 1000 m³¹. On the global scale, the crossover depth of 900 m underscores the importance of both model uncertainty of carbon export and transfer efficiency for understanding the century-scale sequestration. The importance of model uncertainty of transfer efficiency is expected to increase when considering longer-term sequestration.

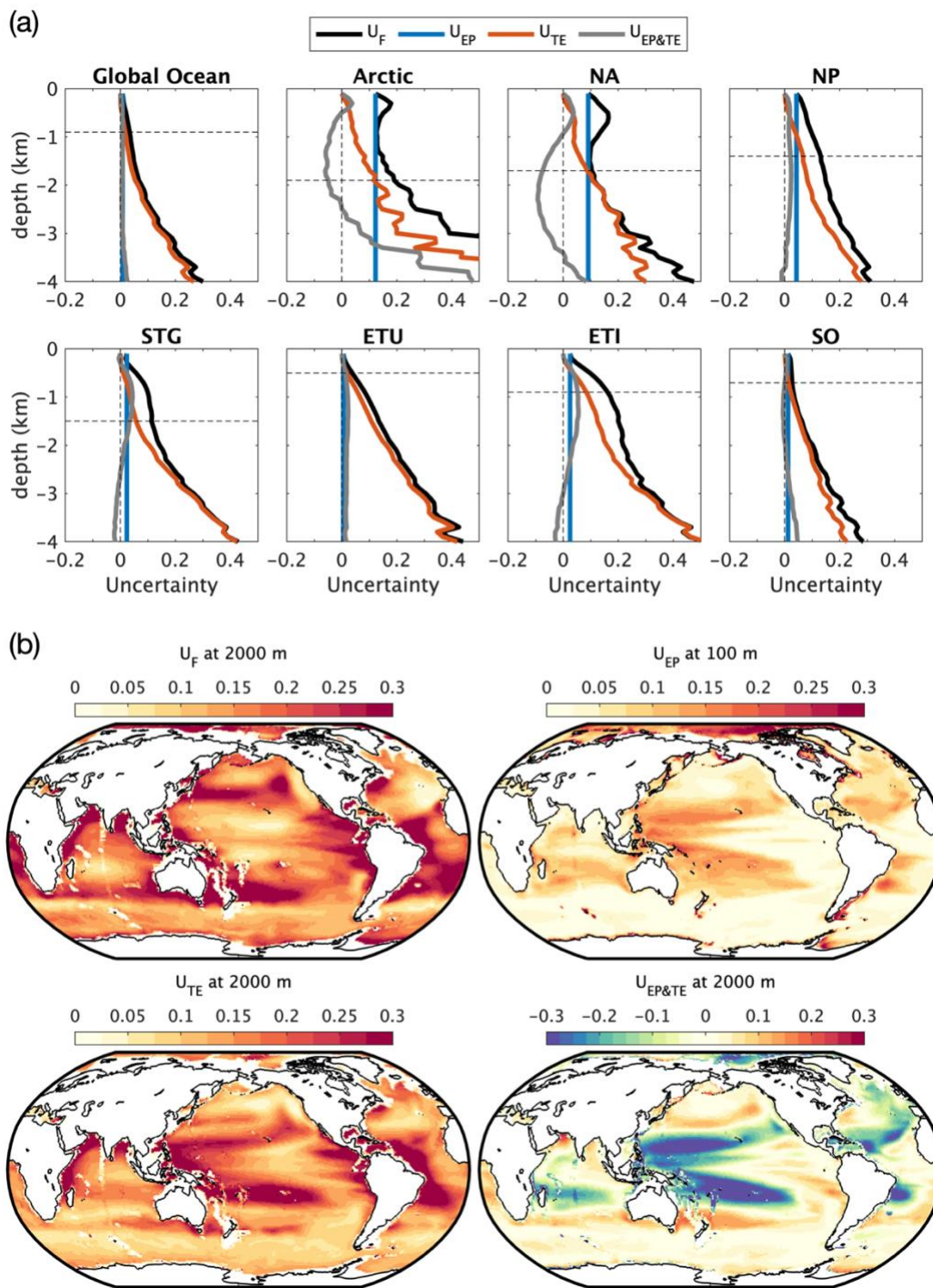


Figure 3. Partitioning of model uncertainty in the vertical carbon flux (U_F) as a function of depth (a) and at 2000 m (b) into the model uncertainty of carbon export (U_{EP}), the model uncertainty of transfer efficiency (U_{TE}), and the co-dependence between carbon export and transfer efficiency ($U_{EP\&TE}$) as described in Methods. The horizontal dashed black lines represent the crossover depth

where the uncertainty of transfer efficiency (U_{TE}) starts to dominate the model uncertainty in the vertical carbon flux (U_F).

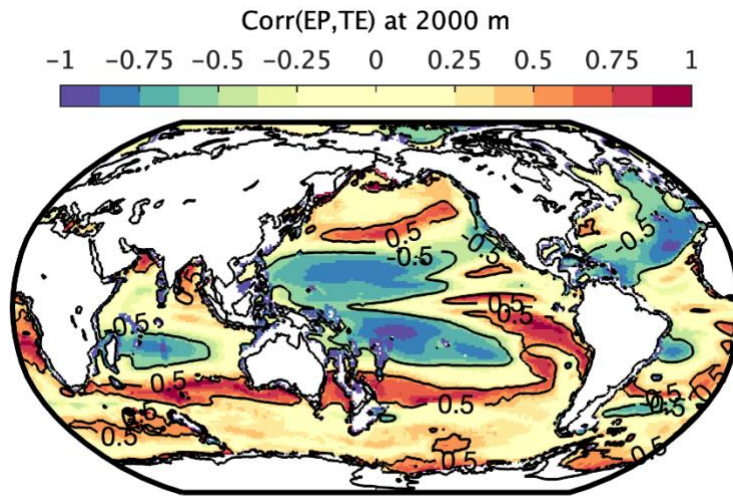


Figure 4. Correlations between the simulated carbon export at 100 m and transfer efficiency at 2000 m. Black contour lines represent the correlation coefficient being ± 0.5 .

2.4 Type I and type II uncertainties

As shown in Figure 5a, all sources of uncertainty grow with depth. The type I uncertainty (can be represented by variations in the exponent b of the Martin curve) is initially the largest source of uncertainty before the type II uncertainty (not representable by variations in b) takes over. Both have significant contributions to the model uncertainty of transfer efficiency. On the global scale, both type I and type II uncertainties at 2000 m contribute about 30% of the model uncertainty of transfer efficiency individually with the remaining 40% attributed to their co-dependence. On the biome scale, the contribution of the type I uncertainty ranges from 24% in the Subtropical Gyres to 47% in the Equatorial Indian biome. The relative contribution from the type II uncertainty varies over a wider range across biomes. It can reach up to 73% in the North Atlantic and can be as low as 16% in the Equatorial Indian biome. On the local scale, all these sources of uncertainty have comparable contributions to the model uncertainty of transfer efficiency, except in the western part of the Equatorial Atlantic, the boundary along the Subtropical Gyres, and the west coast of the Indian Ocean, which are characterized by high type I uncertainty (Figure 5b).

The Martin curve has been widely used to fit the sparse observations of deep vertical carbon flux for extrapolating and comparing the transfer efficiency. The exponent b has also been perturbed to assess the sensitivity of atmospheric CO_2 levels to transfer efficiency^{14,32}. Our results show that the model uncertainty of transfer efficiency is underestimated by using the Martin curve and exponent b . This is consistent with findings in ref.⁹ where six simple alternative parameterizations were compared with the Martin Curve. Therefore, sampling and using the vertical profiles of transfer efficiency is recommended rather than only reporting the exponent b when calibrating models.

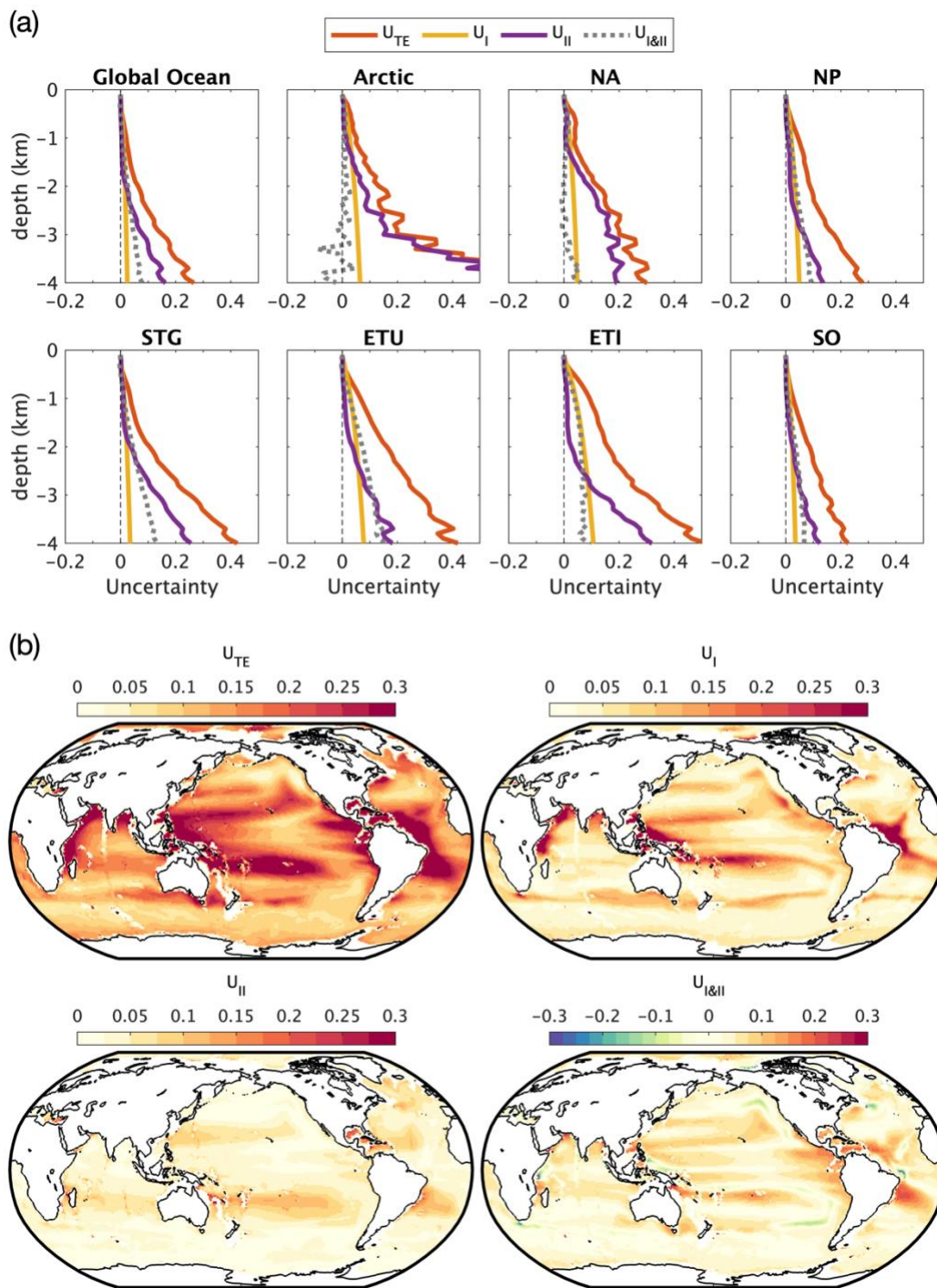


Figure 5. The partitioning of model uncertainty of transfer efficiency (U_{TE}) as a function of depth (a) and at 2000 m (b) into the type I (U_I) and type II (U_{II}) uncertainties that are representable with the Martin curve and not, as well as their co-dependence ($U_{I&II}$) as described in Methods.

2.5 Future implications

In this study, we find large variability in simulating the BCP across different CMIP6 models because of discrepancies in model structures and parameters. This casts doubt on the reliability of future projections. Which model aspects need prioritized improvement depends on the depth and time scales of interest. When the global, century-scale carbon sequestration is of interest, the model uncertainties of carbon export and transfer efficiency are both important.

A reduction of model uncertainty requires a mechanistic understanding of the carbon export and flux attenuation within the euphotic zone and below, and the integration of this knowledge into ESMs. Model skill and model representation in the euphotic zone have seen many improvements between the fifth and sixth phases of Coupled Model Intercomparison Project (CMIP5 and CMIP6), e.g., an increase in the number of phytoplankton/zooplankton functional groups and allowance of variable nutrient stoichiometries^{33,34}. Unlike the euphotic zone, model representations of flux attenuation have remained relatively simple³⁵ and there is no agreement on which processes should be included. For instance, temperature-dependent remineralization^{7,12,35} and particle composition^{5,7,35} are key drivers in setting the global patterns of transfer efficiency, but are not included in all CMIP6 models. A recent study prioritized these two processes along with fragmentation and zooplankton migration as key mechanisms to be included in the next generation of ESMs¹. However, adding a new mechanism does not necessarily bring about improvements in models' predictive skills, as it may introduce additional uncertainty which cannot be constrained by observations. Therefore, comprehensive evaluations of the value added and uncertainty of different mechanisms should be conducted before routine use in ESMs.

Apart from model structures, the divergent degrees of parameter tuning across different CMIP6 models can contribute to model uncertainty. Parameter optimization should be applied to address this issue by searching the optimal values and providing a range of acceptable solutions of parameters that are related to the BCP^{24,27,35–37}.

It is worth noting that both refinement of model structures and optimization of model parameters critically depend on appropriate observations³⁸. Given the fact that the currently available in-situ observations of vertical carbon flux are biased to productive regions, increased availability of observations in Subtropical Gyres would be desirable and should reduce the large model uncertainty there. In addition, models producing opposite global patterns of transfer efficiency may have only minor differences in model's goodness-of-fit when compared with in-

situ observations, e.g., the GFDL-CM4, GFDL-ESM4, and IPSL-CM6A-LR-INCA (Figure S8). This suggests that an increase in observations of not only carbon export but also depth-resolved deep vertical carbon flux is urgently needed.

Over the past decades, the vertical carbon flux has been measured by sediment traps^{39,40} and radioactive tracers such as ^{234}Th - ^{238}U and ^{210}Po - ^{210}Pb but these are limited due to logistical constraints^{41–43}. These constraints are now being alleviated by autonomous profiling platforms. Biogeochemical-Argo (BGC-Argo) floats that are equipped with bio-optical sensors such as backscatter and optical sediment traps can inform us on vertical carbon flux in a cost-effective manner^{44–46}. They also bring about a large expansion of biogeochemical measurements including nutrient and oxygen, which reflect the accumulative effects of production, respiration, and transport of organic matter^{23,24,47–51}. However, assessing model uncertainty in simulating the BCP through comparisons with these biogeochemical measurements is made difficult by the fact that the CMIP6 models are initialized and spun up following different procedures³³. A consistent protocol for model initialization and spin-up would provide better opportunities for model assessment. Since different mechanisms can cancel out each other, leading to the same vertical carbon flux, observations of the vertical carbon flux are insufficient on their own to tease apart different mechanisms^{52,53}. This is a well-known issue of underdetermination and is common in biogeochemical models³⁸. Although independent observations are becoming more readily accessible through BGC-Argo floats, their utilization for model assessment requires the corresponding tracers to be included explicitly in ESMs. For instance, the availability of particle size distribution has increased rapidly via the separation between small and large particles based on high-frequency measurements of backscatter^{54,55} and is expected to further expand due to the emergence of miniaturized imaging sensors⁵⁶. This information can not only provide supplementary constraints on the vertical carbon flux but also gain process-level insights into the distinction between remineralization and fragmentation, which are two complementary processes in setting the vertical carbon flux but leave contrasting effects on the particle size distributions^{52,57}. However, model comparison with these independent observations requires at least two particle size classes to be simulated.

3. Conclusions

In this study, model uncertainty in simulating the BCP within state-of-the-art ocean biogeochemical models was partitioned into different contributors and analyzed. Comparisons

across 14 CMIP6 models show large variations in carbon export at 100 m and increased variations at depth due to differences in transfer efficiency. Owing to the discrepancies in model structures and parameters, the simulated transfer efficiency from 14 CMIP6 models can be classified into three broad categories, with the first category features a homogenous global pattern of transfer efficiency and the remaining categories encapsulate two contrasting patterns. In addition, uncertainty analysis based on 6 of 14 representative CMIP6 models suggests an initial dominance of model uncertainty of carbon export which is overtaken by the model uncertainty of transfer efficiency at 900 m on the global scale. Since a recent study has suggested that the century-scale carbon sequestration is equivalently important above and below 1000 m³¹, model uncertainties of carbon export and transfer efficiency are both important, implying the necessity to improve model representations both within the euphotic zone and below. In addition, an expansion of observations of not only carbon export but also flux attenuation in the deep ocean is important. The model uncertainty of transfer efficiency was further decoupled into uncertainties that are representable by the Martin curve or not, both of which show comparable significance, suggesting that the usage of Martin curve and exponent b for comparing transfer efficiency underestimates the model uncertainty of transfer efficiency. Therefore, the vertical profile of transfer efficiency is recommended for model validation rather than using the exponent b.

4. Methods

4.1 CMIP6 models

To assess different factors that contribute to model uncertainty in simulating the BCP within state-of-the-art biogeochemical ocean models, we chose the 14 ESMs that are part of the CMIP6 and report depth-resolved vertical carbon flux (Table 1). Model outputs from the historical simulations were averaged over 15 years (2000-2014) to produce the climatology of vertical carbon flux. Only one ensemble member (r1i1p1f1) is used for each model. To facilitate comparisons between models, the model outputs were first remapped onto a common 1°×1° regular grid with a vertical resolution of 100 m and then compared with different spatial averaging, i.e., from single grid cells to the global average. On biome scales, the simulated vertical carbon flux is averaged over different biomes which encapsulate the large-scale variability in biogeochemical properties (Figure 1a). This map of biomes is adapted from ref.⁵⁸ by combining the subtropical regions of the Pacific Ocean and the Atlantic Ocean in the northern and southern

hemisphere into the Subtropical Gyres (STG). In addition, the equatorial upwelling regions of the Pacific Ocean and the Atlantic Ocean are combined into the Equatorial Tropical Upwelling (ETU).

4.2 In-situ observations of the vertical carbon flux

In-situ observations of the vertical carbon flux were compiled by ref.⁵⁹ and are available in the supporting information of ref.¹³. Most of the available in-situ observations (~90%) were collected by sediment traps deployed for less than 30 days and hence are compared to the monthly climatology in the nearest model grid cell of each ESMs.

4.3 Different sources of model uncertainty in simulating the BCP

To diagnose the model uncertainty in simulating the BCP, we describe the vertical carbon flux at depth z ($\overline{F(z)}$) as a product of the carbon export ($\overline{EP}$) and transfer efficiency ($\overline{TE(z)}$):

$$\overline{F(z)} = \overline{EP} \times \overline{TE(z)}, \quad (1)$$

where the carbon export is defined as the vertical carbon flux at 100 m and the overbar represents the average over different spatial scales. Although the depth horizon of 100 m does not necessarily separate the upper layer from the lower layer where no organic particles are produced, it facilitates comparisons between different models.

The model uncertainty in simulating the BCP is calculated based on the variance of the vertical carbon flux across different ESMs and hence is referred to as the model uncertainty of the vertical carbon flux (U_F , in unit of $(\log(mg\ C\ m^{-2}\ day^{-1}))^2$). In addition, prior to the calculation of variance, a log-transformation is applied to the vertical carbon flux which varies by orders of magnitude:

$$\log_{10}(\overline{F(z)}) = \log_{10}(\overline{EP}) + \log_{10}(\overline{TE(z)}), \quad (2)$$

$$U_F = var(\log_{10}(\overline{F_z})). \quad (3)$$

The model uncertainty of the vertical carbon flux can then be decomposed into the model uncertainty of carbon export (U_{EP} , in unit of $(\log(mg\ C\ m^{-2}\ day^{-1}))^2$), the model uncertainty of transfer efficiency (U_{TE} , dimensionless), and the co-dependence between carbon export and transfer efficiency ($U_{EP\&TE}$, in unit of $(\log(mg\ C\ m^{-2}\ day^{-1}))^2$):

$$U_F = U_{EP} + U_{TE} + U_{EP\&TE}, \quad (4)$$

$$U_{EP} = var(\log_{10}(\overline{EP})), \quad (5)$$

$$U_{TE} = var(\log_{10}(\overline{TE})), \quad (6)$$

$$U_{EP\&TE} = 2 \times var(\log_{10}(\overline{EP}), \log_{10}(\overline{TE})). \quad (7)$$

The model uncertainty of the transfer efficiency (U_{TE}) is further decoupled into **type I** (U_I , dimensionless) and **type II** (U_{II} , dimensionless) to distinguish between model uncertainty of transfer efficiency which can be represented by changing the exponent b of the Martin curve or **not**. To achieve this, the simulated vertical profiles of transfer efficiency from each ESMs were fitted to the Martin curve and the variance of the predicted transfer efficiency ($\overline{TE_{martin}(z)}$) is exclusively due to the **exponent b** :

$$\overline{TE_{martin}(z)} = \left(\frac{z}{100}\right)^{-b}. \quad (8)$$

The simulated transfer efficiency is represented as:

$$\overline{TE(z)} = \overline{TE_{martin}(z)} \times \overline{TE_{deviation}(z)}, \quad (9)$$

where the $TE_{deviation}(z)$ represents the deviations of the simulated transfer efficiency from the corresponding Martin curve. By applying a log-transformation, equ. (9) becomes:

$$\log_{10}(\overline{TE(z)}) = \log_{10}(\overline{TE_{martin}(z)}) + \log_{10}(\overline{TE_{deviation}(z)}). \quad (10)$$

Therefore, the model uncertainty of the transfer efficiency (U_{TE}) can be expressed as the sum of the **type I** (U_I), **type II** (U_{II}), and **their** co-dependence ($U_{I\&II}$, dimensionless):

$$U_{TE} = U_I + U_{II} + U_{I\&II} \quad (11)$$

$$U_I = var(\log_{10}(\overline{TE_{martin}})) \quad (12)$$

$$U_{II} = var(\log_{10}(\overline{TE_{deviation}})) \quad (13)$$

$$U_{I\&II} = 2 \times var(\log_{10}(\overline{TE_{martin}}), \log_{10}(\overline{TE_{deviation}})) \quad (14)$$

Acknowledgement

This study was supported by Katja Fennel's NSERC Discovery Grant and the Northwest North Atlantic Biological Carbon Pump (NWA BCP) project of the Ocean Frontier Institute.

Data Availability Statement

CMIP6 model outputs are available from the Earth System Grid Federation (ESGF) data repository at <https://esgf-node.llnl.gov/search/cmip6/> (last access: September 2021).

Code Availability Statement

The codes for data analysis will be publicly available at the Zenodo repository once the manuscript is accepted.

Author Contribution Statement

B.W. and K.F. conceived the study. B.W. carried out the analysis. B.W. and K.F. discussed the results and wrote the paper.

Conflict of interest statement

The authors declare that they have no conflict of interest.

Reference

1. Henson, S. A. *et al.* Uncertain response of ocean biological carbon export in a changing world. *Nat. Geosci.* **15**, 248–254 (2022).
2. Wilson, J. D. *et al.* The biological carbon pump in CMIP6 models: 21st century trends and uncertainties. *Proc. Natl. Acad. Sci.* **119**, e2204369119 (2022).
3. Laufkötter, C. *et al.* Projected decreases in future marine export production: the role of the carbon flux through the upper ocean ecosystem. *Biogeosciences* **13**, 4023–4047 (2016).
4. Buesseler, K. O., Boyd, P. W., Black, E. E. & Siegel, D. A. Metrics that matter for assessing the ocean biological carbon pump. *Proc. Natl. Acad. Sci.* **117**, 9679–9687 (2020).
5. Henson, S. A., Sanders, R. & Madsen, E. Global patterns in efficiency of particulate organic carbon export and transfer to the deep ocean. *Global Biogeochem. Cycles* **26**, 1–14 (2012).
6. Lima, I. D., Lam, P. J. & Doney, S. C. Dynamics of particulate organic carbon flux in a global ocean model. *Biogeosciences* **11**, 1177–1198 (2014).
7. Cram, J. A. *et al.* The Role of Particle Size, Ballast, Temperature, and Oxygen in the Sinking Flux to the Deep Sea. *Global Biogeochem. Cycles* **32**, 858–876 (2018).
8. Claustre, H., Legendre, L., Boyd, P. W. & Levy, M. The Oceans' Biological Carbon Pumps: Framework for a Research Observational Community Approach. *Front. Mar. Sci.* **8**, (2021).
9. Lauderdale, J. M. & Cael, B. B. Impact of Remineralization Profile Shape on the Air-Sea Carbon Balance. *Geophys. Res. Lett.* **48**, e2020GL091746 (2021).
10. Cael, B. B. & Bisson, K. Particle Flux Parameterizations: Quantitative and Mechanistic

Similarities and Differences. *Front. Mar. Sci.* **5**, (2018).

11. Martin, J. H., Knauer, G. A., Karl, D. M. & Broenkow, W. W. VERTEX: carbon cycling in the northeast Pacific. *Deep Sea Res. Part A. Oceanogr. Res. Pap.* **34**, 267–285 (1987).
12. Marsay, C. M. *et al.* Attenuation of sinking particulate organic carbon flux through the mesopelagic ocean. *Proc. Natl. Acad. Sci.* **112**, 1089–1094 (2015).
13. Gloege, L., McKinley, G. A., Mouw, C. B. & Ciochetto, A. B. Global evaluation of particulate organic carbon flux parameterizations and implications for atmospheric pCO₂. *Global Biogeochem. Cycles* **31**, 1192–1215 (2017).
14. Kwon, E. Y., Primeau, F. & Sarmiento, J. L. The impact of remineralization depth on the air–sea carbon balance. *Nat. Geosci.* **2**, 630–635 (2009).
15. Guidi, L. *et al.* A new look at ocean carbon remineralization for estimating deepwater sequestration. *Global Biogeochem. Cycles* **29**, 1044–1059 (2015).
16. Aumont, O. *et al.* Variable reactivity of particulate organic matter in a global ocean biogeochemical model. *Biogeosciences* **14**, 2321–2341 (2017).
17. Long, M. C. *et al.* Simulations With the Marine Biogeochemistry Library (MARBL). *J. Adv. Model. Earth Syst.* **13**, e2021MS002647 (2021).
18. Dunne, J. P. *et al.* Simple Global Ocean Biogeochemistry With Light, Iron, Nutrients and Gas Version 2 (BLINGv2): Model Description and Simulation Characteristics in GFDL’s CM4.0. *J. Adv. Model. Earth Syst.* **12**, e2019MS002008 (2020).
19. Stock, C. A. *et al.* Ocean Biogeochemistry in GFDL’s Earth System Model 4.1 and Its Response to Increasing Atmospheric CO₂. *J. Adv. Model. Earth Syst.* **12**, e2019MS002043 (2020).
20. Aumont, O., Ethé, C., Tagliabue, A., Bopp, L. & Gehlen, M. PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. *Geosci. Model Dev.* **8**, 2465–2513 (2015).
21. Mauritsen, T. *et al.* Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO₂. *J. Adv. Model. Earth Syst.* **11**, 998–1038 (2019).
22. Tjiputra, J. F. *et al.* Ocean biogeochemistry in the Norwegian Earth System Model version 2 (NorESM2). *Geosci. Model Dev.* **13**, 2393–2431 (2020).
23. Weber, T., Cram, J. A., Leung, S. W., DeVries, T. & Deutsch, C. Deep ocean nutrients

- imply large latitudinal variation in particle transfer efficiency. *Proc. Natl. Acad. Sci.* **113**, 8606–8611 (2016).
24. DeVries, T. & Weber, T. The export and fate of organic matter in the ocean: New constraints from combining satellite and oceanographic tracer observations. *Global Biogeochem. Cycles* **31**, 535–555 (2017).
25. Vichi, M. *et al.* Biogeochemical Flux Model (BFM): Equation Description and User Manual. BFM version 5.2. *BFM Rep. Ser. N. 1, Release 1.2, June 2020, Bol. Italy, <http://bfm-community.eu>* 104 (2020).
26. Siegel, D. A. *et al.* Global assessment of ocean carbon export by combining satellite observations and food-web models. *Global Biogeochem. Cycles* **28**, 181–196 (2014).
27. Nowicki, M., DeVries, T. & Siegel, D. A. Quantifying the Carbon Export and Sequestration Pathways of the Ocean’s Biological Carbon Pump. *Global Biogeochem. Cycles* **36**, e2021GB007083 (2022).
28. Boyd, P. W., Claustre, H., Levy, M., Siegel, D. A. & Weber, T. Multi-faceted particle pumps drive carbon sequestration in the ocean. *Nature* **568**, 327–335 (2019).
29. Siegel, D. A., DeVries, T., Cetinić, I. & Bisson, K. M. Quantifying the Ocean’s Biological Pump and Its Carbon Cycle Impacts on Global Scales. *Ann. Rev. Mar. Sci.* **15**, 329–356 (2023).
30. Hill, C., Bugnion, V., Follows, M. & Marshall, J. Evaluating carbon sequestration efficiency in an ocean circulation model by adjoint sensitivity analysis. *J. Geophys. Res. Ocean.* **109**, (2004).
31. Ricour, F., Guidi, L., Gehlen, M., DeVries, T. & Legendre, L. Century-scale carbon sequestration flux throughout the ocean by the biological pump. *Nat. Geosci.* **16**, 1105–1113 (2023).
32. Wilson, J. D., Barker, S., Edwards, N. R., Holden, P. B. & Ridgwell, A. Sensitivity of atmospheric CO₂ to regional variability in particulate organic matter remineralization depths. *Biogeosciences* **16**, 2923–2936 (2019).
33. Séférian, R. *et al.* Tracking Improvement in Simulated Marine Biogeochemistry Between CMIP5 and CMIP6. *Curr. Clim. Chang. Reports* **6**, 95–119 (2020).
34. Laurent, A., Fennel, K. & Kuhn, A. An observation-based evaluation and ranking of historical Earth system model simulations in the northwest North Atlantic Ocean.

Biogeosciences **18**, 1803–1822 (2021).

35. Dinauer, A., Laufkötter, C., Doney, S. C. & Joos, F. What Controls the Large-Scale Efficiency of Carbon Transfer Through the Ocean’s Mesopelagic Zone? Insights From a New, Mechanistic Model (MSPACMAM). *Global Biogeochem. Cycles* **36**, e2021GB007131 (2022).
36. Wang, B., Fennel, K., Yu, L. & Gordon, C. Assessing the value of biogeochemical Argo profiles versus ocean color observations for biogeochemical model optimization in the Gulf of Mexico. *Biogeosciences* **17**, 4059–4074 (2020).
37. Stukel, M. R., Décima, M. & Landry, M. R. Quantifying biological carbon pump pathways with a data-constrained mechanistic model ensemble approach. *Biogeosciences* **19**, 3595–3624 (2022).
38. Fennel, K. *et al.* Ocean biogeochemical modelling. *Nat. Rev. Methods Prim.* (2022) doi:10.1038/s43586-022-00154-2.
39. Francois, R., Honjo, S., Krishfield, R. & Manganini, S. Factors controlling the flux of organic carbon to the bathypelagic zone of the ocean. *Global Biogeochem. Cycles* **16**, 20–34 (2002).
40. Honjo, S., Manganini, S. J., Krishfield, R. A. & Francois, R. Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: A synthesis of global sediment trap programs since 1983. *Prog. Oceanogr.* **76**, 217–285 (2008).
41. Giering, S. L. C. *et al.* Sinking Organic Particles in the Ocean—Flux Estimates From in situ Optical Devices. *Front. Mar. Sci.* **6**, (2020).
42. Le Moigne, F. A. C. *et al.* Export of organic carbon and biominerals derived from ²³⁴Th and ²¹⁰Po at the Porcupine Abyssal Plain. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **72**, 88–101 (2013).
43. Buesseler, K. O. *et al.* An assessment of particulate organic carbon to thorium-234 ratios in the ocean and their impact on the application of ²³⁴Th as a POC flux proxy. *Mar. Chem.* **100**, 213–233 (2006).
44. Lacour, L., Llort, J., Briggs, N., Strutton, P. G. & Boyd, P. W. Seasonality of downward carbon export in the Pacific Southern Ocean revealed by multi-year robotic observations. *Nat. Commun.* **14**, 1278 (2023).
45. Estapa, M. L., Feen, M. L. & Breves, E. Direct Observations of Biological Carbon Export

- From Profiling Floats in the Subtropical North Atlantic. *Global Biogeochem. Cycles* **33**, 282–300 (2019).
46. Dall’Olmo, G. & Mork, K. A. Carbon export by small particles in the Norwegian Sea. *Geophys. Res. Lett.* **41**, 2921–2927 (2014).
 47. Wilson, J. D., Ridgwell, A. & Barker, S. Can organic matter flux profiles be diagnosed using remineralisation rates derived from observed tracers and modelled ocean transport rates? *Biogeosciences* **12**, 5547–5562 (2015).
 48. Hennon, T. D., Riser, S. C. & Mecking, S. Profiling float-based observations of net respiration beneath the mixed layer. *Global Biogeochem. Cycles* **30**, 920–932 (2016).
 49. Arteaga, L. A., Pahlow, M., Bushinsky, S. M. & Sarmiento, J. L. Nutrient Controls on Export Production in the Southern Ocean. *Global Biogeochem. Cycles* **33**, 942–956 (2019).
 50. Quay, P., Emerson, S. & Palevsky, H. Regional Pattern of the Ocean’s Biological Pump Based on Geochemical Observations. *Geophys. Res. Lett.* **47**, e2020GL088098 (2020).
 51. Huang, Y., Fassbender, A. J., Long, J. S., Johannessen, S. & Bernardi Bif, M. Partitioning the Export of Distinct Biogenic Carbon Pools in the Northeast Pacific Ocean Using a Biogeochemical Profiling Float. *Global Biogeochem. Cycles* **36**, e2021GB007178 (2022).
 52. Weber, T. & Bianchi, D. Efficient Particle Transfer to Depth in Oxygen Minimum Zones of the Pacific and Indian Oceans. *Front. Earth Sci.* **8**, 376 (2020).
 53. Wang, B. & Fennel, K. An Assessment of Vertical Carbon Flux Parameterizations Using Backscatter Data From BGC Argo. *Geophys. Res. Lett.* **50**, e2022GL101220 (2023).
 54. Wang, B. & Fennel, K. Biogeochemical-Argo data suggest significant contributions of small particles to the vertical carbon flux in the subpolar North Atlantic. *Limnol. Oceanogr.* **67**, 2405–2417 (2022).
 55. Briggs, N., Dall’Olmo, G. & Claustre, H. Major role of particle fragmentation in regulating biological sequestration of CO₂ by the oceans. *Science (80-.).* **367**, 791–793 (2020).
 56. Picheral, M. *et al.* The Underwater Vision Profiler 6: an imaging sensor of particle size spectra and plankton, for autonomous and cabled platforms. *Limnol. Oceanogr. Methods* **20**, 115–129 (2022).
 57. Cram, J. A. *et al.* Slow Particle Remineralization, Rather Than Suppressed Disaggregation, Drives Efficient Flux Transfer Through the Eastern Tropical North Pacific Oxygen Deficient Zone. *Global Biogeochem. Cycles* **36**, e2021GB007080 (2022).

- 601 58. Wang, W., Moore, J. K., Martiny, A. C. & Primeau, F. W. Convergent estimates of marine
602 nitrogen fixation. *Nature* **566**, 205–211 (2019).
- 603 59. Mouw, C. B., Barnett, A., McKinley, G. A., Gloege, L. & Pilcher, D. Global ocean
604 particulate organic carbon flux merged with satellite parameters. *Earth Syst. Sci. Data* **8**,
605 531–541 (2016).
- 606

Decision letter and referee reports: second round

4th Jun 24

Dear Dr Wang,

Your manuscript titled "Partitioning uncertainty of the ocean's biological carbon pump in CMIP6 models" has now been seen by our reviewers, whose comments appear below. In light of their advice we are delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment under the open access CC BY license (Creative Commons Attribution v4.0 International License).

We therefore invite you to revise your paper one last time to address the remaining concerns of our reviewers. At the same time we ask that you edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table".

*****Please take care to match our formatting and policy requirements. We will check revised manuscript and return manuscripts that do not comply. Such requests will lead to delays. *****

Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files as a Related Manuscript file.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at <https://www.nature.com/documents/commsj-file-checklist.pdf>.

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication under a [CC BY license](#) (Creative Commons Attribution 4.0 International License). This license allows maximum dissemination and re-use of open access materials and is preferred by many research funding bodies.

For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit <https://www.nature.com/commsenv/article-processing-charges>

At acceptance, you will be provided with instructions for completing this CC BY license on behalf of all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:
[redacted]

** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Michael Stukel, PhD
Editorial Board Member
Communications Earth & Environment

Martina Grecequet, PhD
Associate Editor,
Communications Earth & Environment
@CommsEarth

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

Review 2 of Wang et al. "Partitioning uncertainty of the ocean's biological carbon pump in CMIP6 models"

I find the manuscript to be clearer and easier to follow in its revised form. I have just a couple of comments to further improve clarity.

Throughout: I find the term "vertical profile of transfer efficiency" distracting. I know what the authors mean (after thinking about it), but it's not a phrase used regularly by either observationalists or modellers, as far as I'm aware. One possibility would be to replace it with "vertical profile of flux", where appropriate, as this is what is actually both measured in situ and modelled.

Line 42: Change to: "A recent global mean estimate of the exponent..."

Lines 286-288: Is this a suggestion to observationalists to collect/report depth profiles of flux, or to modellers to save the 3D output of flux so they can assess depth profiles?

Table 2: Caption has 'vertical' in it twice. Also suggest adding "of respiration" after "Temperature dependence" and "Oxygen dependence" to avoid any confusion.

Figure 2 caption: Define in the caption what the 3 different latitudinal patterns are.

Reviewer #2 (Remarks to the Author):

I am satisfied with the author's revisions and recommend publication.

We thank the reviewers for their constructive comments which we found very helpful while revising the manuscript. We considered all comments carefully and detail our responses and actions below. The complete reviewer comments are shown in black font followed by our responses to each comment in blue font.

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

Review 2 of Wang et al. "Partitioning uncertainty of the ocean's biological carbon pump in CMIP6 models"

I find the manuscript to be clearer and easier to follow in its revised form. I have just a couple of comments to further improve clarity.

Response: We would like to thank the reviewer for the positive assessment and constructive comments.

Throughout: I find the term "vertical profile of transfer efficiency" distracting. I know what the authors mean (after thinking about it), but it's not a phrase used regularly by either observationalists or modellers, as far as I'm aware. One possibility would be to replace it with "vertical profile of flux", where appropriate, as this is what is actually both measured in situ and modelled.

Response: Revised as suggested.

Line 42: Change to: "A recent global mean estimate of the exponent...."

Response: Revised as suggested.

Lines 286-288: Is this a suggestion to observationalists to collect/report depth profiles of flux, or to modellers to save the 3D output of flux so they can assess depth profiles?

Response: This is a suggestion to both observationalists who collect depth profiles and modellers who compare model output with observed profiles. We have rephrased this sentence to make it clearer:

"Therefore, we recommend observationalists to sample the vertical profiles of carbon flux rather than only reporting the exponent b and modellers to compare the profiles during model calibration and intercomparisons."

Table 2: Caption has 'vertical' in it twice. Also suggest adding "of respiration" after "Temperature dependence" and "Oxygen dependence" to avoid any confusion.

Response: Revised as suggested.

Figure 2 caption: Define in the caption what the 3 different latitudinal patterns are.

Response: Revised as suggested (changed text is highlighted in **bold** font):

"Figure 2. Three different latitudinal patterns of simulated transfer efficiency at 2000 m displayed by the CMIP6 models. (a) Three latitudinal patterns of simulated transfer efficiency at

2000 m with one latitudinal homogenous pattern and two opposite patterns, either high in the oligotrophic Subtropical Gyres and low in productive regions or vice versa. Error bars represent the interquartile range within a 5-degree latitudinal band. (b) The global distribution of transfer efficiency at 2000 m from 4 representative models. The global distributions of all models are shown in **Supplementary Figure 2.**”

Reviewer #2 (Remarks to the Author):

I am satisfied with the author's revisions and recommend publication.

Response: Thank you for the positive assessment.