Editor Review

The specific issues I've identified are as follows:

1) Review #1: Section 2 ("L189-191") mentions issues about the public availability of data that do not appear in the "Data availability" statement. If they do not appear elsewhere in the manuscript, it may be worth mentioning these issues here so that readers understand the limitations.

Answer) We addressed this issue by adding a statement in the "Data availability' section that discharge data for other rivers cannot be made publicly available, as follows.

"Unfortunately, discharge data for other rivers cannot be made publicly available."

2) Review #2: Q6 ("ERA5 forcing has been updated …") has a long response, but it is unclear how much of this has made it into the manuscript.

Answer) We have added the response to Review #2 in the revised manuscript (line 390~392) as follows.

"Although NPRT was conducted with external forcings that include interannual variations, the results may have several uncertainties due to insufficient spin-up time and/or climatological data for river input and atmospheric deposition.**"**

3) Review #2: Q10 ("Describe primary productions …") also has a long response but is unclear how it has changed the manuscript (specifically on primary production – I can see the response for the carbon cycle).

Answer) Reflecting the editor's suggestion, we have included information on net primary production in the revised manuscript (lines 144–150).

4) Review #3: Your response to Q4 ("Overall, I suggest …") would benefit from a bit of expansion so that both I and your referee can understand the changes made to the manuscript that you allude to.

Answer) We revised the manuscript by removing generalized or non-specific sentences, i.e., the first sentence in the abstract, introduction, and conclusions (lines 17, 40–43, and 450–451 in the original manuscript).

5) Review #3: Q13 ("This is an odd description …") is ambiguously answered in your response. It looks like text in the conclusions has been altered, but this unclear. And if the manuscript has not been changed, why has it not been changed. A bit more clarity is needed.

Answer) We have reorganized and rewritten our response to Reviewer #3 (lines 337–345, 549–553 in the revised manuscript).

6) A general comment I'd make is that your responses would be improved by adding the modified text to your response. Where the change is large, this is probably less helpful, but where a short addition or clarifying language has been added, it would be helpful to reproduce it here to spare readers the need to chase it down.

Answer) I would like to express my deep gratitude for your constructive feedback.

7) Ln. 478: This states that you have an alkalinity bias of >400 umol/kg, but the accompanying plot appears to show a much smaller bias, closer to 100 umol/kg. Can you clarify? And possibly just calculate and report the mean bias.

Answer) Fig. 15f in the manuscript showed the distribution of alkalinity bias between GLODAPv2 and NPRT. A significantly positive bias of >400 umol/kg was predominantly distributed in subtropical regions

8) Ln. 1012-1014: Can you please remove the confusing use of bracketed statements here and just write the caption clearly? I should say that I've written things exactly like this myself in the past, but have come to realise how unhelpful it is. The figure itself would also benefit from a legend to identify what the different lines are, but I appreciate that may not be simple to achieve.

Answer) We have revised the related statements based on your suggestions.

9) Finally, a general issue about your figures that should be addressed before publication (i.e. you do not need to do this right now, but it may be better to) is your use of colour scales that are not colourblind-friendly. I should quickly add that this is something I am also "guilty" of in my own publications, although I am adapting my own practise. Please consult our guidance on this here:

https://www.geoscientific-model-development.net/submission.html#figurestables

Answer) We appreciate your pointing out the aspects we overlooked. We realized the need to pay closer attention to the time series plots and have revised them accordingly. If there are any other figures that require modification, we are willing to make the necessary adjustments.

R1

Major comments

I think that the discussion is feeble and I think we need to deepen the discussion with more analysis. In addition, I suggest the authors reconstruct the section, especially after the biological parts (Section 3.2). The production of chlorophyll relies on the amount of nutrients and light. Therefore the turn of the sections for nutrients should be put before the chlorophyll and discussions should be more with the reproducibility of the nutrients. Furthermore, the purpose of this study is to show the reproducibility of the model. Since the ocean condition in the northwest Pacific is deeply affected by regional oceanic conditions seasonally, at least the authors should show and describe the reproducibility regionally and seasonally. How about making Tables to summarize the seasonal reproducibility for each parameter in the season (spring, summer, autumn, winter) at several depths in each region? If you put in such tables, it will help us and the authors to understand the advantages and disadvantages of the model, which will help us to improve it for the next opportunity.

Answer) In response to the reviewer's suggestion, we have revised the order of the sections on nutrients (section 3.2.1) and chlorophyll concentrations (section 3.2.3). Furthermore, we analyzed the seasonal variability in all nutrients (nitrate, phosphate, and silicate) and calculated the monthly TD scores for nutrients (Table 1 in the revised manuscript). In addition, in the NPRT results, we inferred the potential causes of significant biases for nutrients distributed in each region and proposed possible improvements (lines 318–345; Figs. 7–10 in the revised manuscript).

[Analysis of seasonal variability for nutrients]

The simulated surface nutrient concentrations in NPRT exhibited a clear seasonal variability. In winter, high nitrate concentrations distributed in the subsurface are supplied to the surface by vertical mixing, resulting in an increase in the upper–layer nutrient concentrations from winter to spring. In summer and autumn, distinct seasonal variability is observed, with concentrations gradually decreasing owing to enhanced stratification. The temporal correlation coefficients between model results and observations (WOA18) for nutrients (nitrate, phosphate, and silicate) are approximately 0.78–0.98, indicating that the seasonal variability is well represented, except for the silicate concentration in the YECS (temporal correlation coefficient: 0.47). In the subarctic region, all nutrients are consistently underestimated regardless of the season, which is likely attributed to the initial and boundary biogeochemical conditions. The initial conditions for nutrients derived from the MOM5–TOPAZ results are approximately 10 μmol kg^{-1} lower than those derived from the WOA18 data (not shown here).

In addition, we need to pay particular attention to the overestimation of silicate in the YECS. In TOPAZ, silicate is regulated through biogeochemical processes such as dissolution and uptake by large phytoplankton within the mixed layer, as well as particles sinking into the deep ocean (Dunne et al. 2012). In NPRT, the overestimation of the silicate concentration can be attributed to two possible factors. The first possibility is that silicate, unlike other nutrients, is only considered taken up by large phytoplankton. The second factor is that the YECS, which is a shallow marginal sea with a maximum depth of less than 50 m, experiences strong vertical mixing throughout the entire water column due to strong winds, surface heat fluxes during winter, and tidal effects. In shallow marginal seas, rather than a decrease in the upper layer silicate concentration due to particles sinking into the deep ocean, it is speculated that particles remain within the mixed layer, continuously increasing through dissolution. To address the large bias in silicates observed in marginal seas such as the YECS, it is necessary to consider the specific marine environment of each region and adopt accurate parameters and external forcings.

Minor comments

Figures and Tables

1) Figure resolutions are low, and the memory and contour labels are too small. (Figs 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 13, 14 16, 17). The memory size is too small to be easily identified. If printed, the diagram will be even smaller and should be noted in large size.

Answer) Reflecting on the reviewer's opinion, we modified all figures.

2) Fig 9 b, c d: The ranges of the x and y axis are too large, being not appropriate.

Answer) We modified all figures (Fig. 14 of the revised manuscript).

3) Fig.11: The authors examine the reproducibility in the Northwest Pacific, subarctic region, East Sea, and YECS in DIC and chlorophyll, but not in nutrients. The reproducibility of chlorophyll and DIC is connected to that of the nutrients. It is better to add the regional and seasonal variability in the nutrients.

Answer) We have incorporated the reviewer's constructive suggestions and explained the seasonal variability of nutrients in the subtropical and subarctic regions, the East Sea, and the YECS (section 3.2.1 in the revised manuscript).

4) Fig.12: Why don't you include the figure of silicate?

Answer) We have added vertical profiles of the biases for all nutrients (nitrate, phosphate, and silicate) in Fig. 10 of the revised manuscript

Section1

- L57-l58 (However, in the Northwest Pacific …. PDO and ENSO (Jung et al., 2017; Ma et al. 2020).) : I don't think so. I think the sentence could be misunderstood, so I think it should be reworded.

  Answer) Thanks to reviewer's comments, we reviewed and revised the content (lines 50–53 in the revised manuscript).

- L73 (simulation of high biomass)

  Answer) We revised it with a more precise statement in the revised manuscript (line: 68)

- L188 (1/12deg): It is helpful to add the distance as well.

Answer) We added distance information (approximately 8 km) - (line 190 in the revised manuscript)

- L211 (Each model is … 'r1i1p1fl'.) : I could not understand this sentence.

  Answer) In the CMIP6 (Coupled Model Intercomparison Project Phase 6) data, r1i1p1f1 is a specific identifier for a particular simulation or ensemble member. Here's what each part of the identifier means:

  - r1: This refers to the first realization (or ensemble member) of the model run. This is the specific run of the model with initial conditions that may vary slightly for each ensemble member.
  - i1: This refers to the first initialization method used in the simulation. Different initial conditions can be used in the model runs to account for variability.
  - p1: This represents the first model physics version used. A model can be run with different physical parameterizations, and this identifies the specific one.
  - f1: This refers to the first experiment (or forcing) applied in the simulation, such as different scenarios of greenhouse gas emissions or other external drivers of climate change.

  In summary, r1i1p1f1 indicates the first ensemble member (realization) using the first initialization, the first set of physics, and the first forcing scenario in the CMIP6 dataset.

  We added the information for r, i, p, f indices in the revised manuscript, along with reference (Taylor et al. 2018). - (lines 220–221 in the revised manuscript)

[reference]

Taylor, K., Juckes, M., Balaji, V., Cinquini, L., Denvil, S., Durack, P., Elkington, M., Guilyardi, E., Kharin, S., Lautenschlager, M., Lawrence, B., Nadeau, D., and Stockhause, M.: CMIP6 Global Attributes, DRS, Filenames, Directory Structure, and CV's, https://github.com/WCRP-CMIP/WGCM_Infrastructure_Panel/ blob/main/Papers/CMIP6_global_attributes_filenames_CVs_v6.2.7.pdf (last access: 15 January 2021), 2018.

Section 2

- L189-191: One of the interesting points in this model includes the climatological monthly mean discharges of 12 major rivers although the target region in the model is the Northwest Pacific. I checked the data availability of this data (L519-538), but there is no information. It is very helpful to write it.

  Answer) River discharge data from China is mostly confidential, making it difficult to obtain directly. The discharge data used in this study was compiled from various sources cited in existing research. Therefore, the data cannot be made publicly available. However, data for the Yangtze River is partially publicly accessible, and the relevant website has been added to the Data Availability section.

Section 3

- L245-248 (In this study, the water …. WOA18 data): The sentence is not appropriate here. You can move to the section 2 (l211-212 in the current version).

  Answer) We revised the manuscript by incorporating the reviewer's opinions (lines 212–216 in the revised manuscript).

- L261: Since the authors describe the NPIW between density 26.6 and 27.2 σ_θ. It is better to add the density contours, not the salinity contours in Fig. 5.

  Answer) We added density contour in Fig. 5 of the revised manuscript.

- L270-282: Instead of the distributions of the surface salinity in the Northwest Pacific, how about showing the expanded figure targeting the YECS, because the WOA18 does not include the data in the YECS very much (I mean that it is hard to get the data), although some of the readers wanted to know the situation.

  Answer) We have added an expanded figure targeting the YECS in the supplementary material (Supplement Fig. S1 in the revised manuscript).

- L414-415 "In particular, for the surface DIC concentration, regarding the latitude, the air-sea exchange and vertical mixing are balanced": I checked Ishizu et al. (2021) and this sentence comes from the misunderstanding. In the subarctic region, biological processes are also impacted.

  Answer) The sentence has been removed as it may lead to misunderstandings in the revised manuscript.

Section 4

- Since I suggested reconstructing Section 3 by dividing results and discussion separately, in this section conclusions would be written only.
- **Answer)** Reflecting the reviewer's opinion, we integrated the discussion with the results (section 3) and revised section 4 (conclusion) to focus primarily on the NPRT results.

R2

I am writing for Kim et al "Development and assessment of the physical-biogeochemical ocean regional model in the Northwest Pacific: NPRT v1.0 (ROMS v3.9-TOPAZ v2.0)

In general, I'm very grateful to see sophisticated efforts to develop NPRT v1.0 in the early stage of manuscript as a reviewer. Regional setting of physical-biogeochemical model needs numerous efforts to tune up regional characteristics and setting boundary conditions. The authors newly coupled existing physical and biogeochemical models that methods have been described well in section 2.3. They found the better representations of plankton phenology which is not represented in the global model. I would like to recommend the manuscript would be suitable in the journal. However, I can give comments for clarification as below.

1) L17. I think it is unclear sentence with non-scientific words. I would recommend removing it or rephrase it to be an general explanation for Northwest Pacific rather than having unspecified words like "different regions" or "various factors" in scientific writing. But it seems second sentence would be the one to explain them in general.

Answer) We removed this sentence in the revised manuscript (lines 17–18 in the old manuscript).

2) L195. I don't know what "las" stands for.

Answer) We have corrected typographical errors (line 197 in the revised manuscript).

drag las → drag law

3) I think Fig. 3 represent a bias of large meandering of Kuroshio extenstion. Please refer Hayashida et al. (https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2023JC019697) and discuss it with biogeochemical properties of NPRT results.

In detail, based on Hayashida et al., several spin-up would be needed to catch Kuroshio large meandering despite of data assimilation. They found 4th cycles of spin-up initial condition can match large meander reliability. This could be a discussion point of view in physical and biogeochemical cycles in the Kuroshio regions. This can be also nice fit to your next potential topic by using NPRT v1.0 I guess. This is not recommendation for the review step so please don't feel pressure to do it right away at this point.

Answer) The meandering of the Kuroshio Extension is caused by various mechanisms, accurate simulation of its main path difficult. Nonetheless, NPRT effectively captures the interannual variability in Kuroshio meandering, as well as the distinct formation of eddies south of the Japanese coast. This eddy formation can physically influence the development of meandering paths and potentially impact biogeochemical processes, as suggested by Hayashida et al. (2023) and Ding et al. (2022). This aspect is briefly described in the revised manuscript (lines 373–375) and will be addressed in future research. We appreciate the reviewer's constructive comments.

[Reference]

Ding, Y., Yu, F., Ren, Q., Nan, F., Wang, R., Liu, Y., and Tang, Y.: The physical-biogeochemical response to a subsurface anticyclonic eddy in the Northwest Pacific. Front. Mar. Sci., 8, https://doi.org/10.3389/fmars.2021.766544, 2022

Hayashida, H., Kiss, A. E., Miyama, T., Miyazawa, Y., and Yasunaka, S.: Anomalous nutricline drives marked biogeochemical contrasts during the Kuroshio Large Meander. J Geophys Res (Oceans) 128:e2023JC019697, https://doi.org/10.1029/2023JC019697, 2023.

4) I'm afraid to see chlorophyll bias in Yellow sea region but it could be natural due to dissolved organic matter influx. Could you describe them more in the manuscript if type I or type II water could be a matter of chlorophyll bias in the MODIS, the modeled chlorophyll, and CDOM.

Answer) As suggested by the reviewer, we analyzed the CDOM data calculated using MODIS data. The YECS is classified as Type II water, where CDOM is dominantly higher compared to other regions (Fig. A2-1). Therefore, the chlorophyll concentration simulated by NPRT shows a significant negative bias compared to MODIS, which may indicate that the chlorophyll concentration observed by MODIS could be overestimated due to CDOM. Thanks to the constructive comments from the reviewer, we were able to confirm that the YECS, classified as Type II water, chlorophyll concentration may be biased due to CDOM. Additionally, we recognized the important considerations when analyzing chlorophyll concentration derived from satellite data.

The above content and figures have been added to the revised manuscript (lines 448–458; supplementary Figure S2).

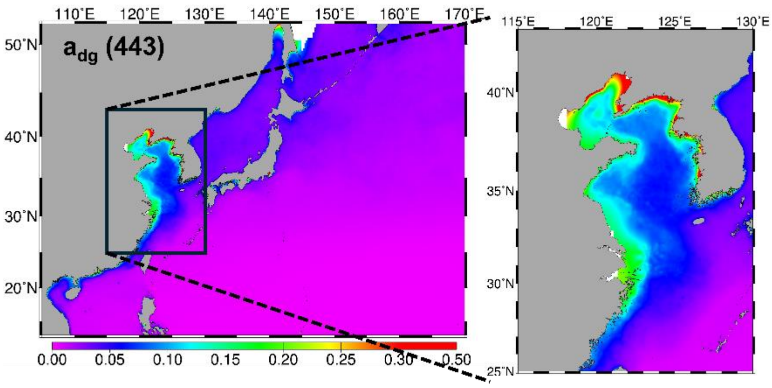

Fig. A2-1. Distributions of the climatological mean absorption coefficient of light for colored detrital matter a_{dg} (443) (unit: m^{-1}) from 2005 to 2014 derived from the MODIS data.

5) CDOM could be very effective on shortwave heating scheme in the regional modeling setup. Author may want to add more discussion about the potential effects of CDOM, e.g., Grace Kim et al. (https://journals.ametsoc.org/view/journals/clim/29/24/jcli-d-16-0053.1.xml) or other potential effects there to be able to develop.

Answer) The absorption coefficient for colored detrital matter (CDM) consists of CDOM and NAP (nonalgal particles). According to Kim et al. (2016), CDOM enhances shortwave heating at the surface while reducing it below the surface. This process can influence penetrating shortwave radiation, mixing, and surface heat flux. Consequently, CDOM can lead to variations in the annual cycle of phytoplankton bloom and decay in the upper layer. In this study, we introduce these findings and briefly describe the potential effects of CDOM (lines 458–462 in the revised manuscript).

[Reference]

Kim, G.E., Gnanadesikan, A., and Pradal, M.A.: Increased surface ocean heating by colored detrital matter (CDM) linked to greater Northern Hemisphere ice formation in the GFDL CM2MC ESM. J. Climate, **29**(24), 9063–9076. https://doi.org/10.1175/ jcli-d-16-0053.1, 2016.

6) ERA5 forcing has been updated in MOM5-TOPAZ. NPRT also have ERA5 forcing? If they have, I think authors should show chlorophyll time series in regional domain and calculate correlation of modeled chlorophyll with satellite ocean color.

Answer) In this study, ERA5 forcing was used to conduct MOM5–TOPAZ and NPRT. Although interannual variability can be examined through atmospheric forcing in MOM5–TOPAZ and NPRT, several uncertainties, as discussed below, still remain regarding the accuracy of interannual variability in chlorophyll concentration simulated by NPRT.

(1) To stabilize biogeochemical variables, a 100-year spin-up time integration was conducted using MOM5–TOPAZ. However, a 100-year spin-up time is not sufficient for biogeochemical variables to reach equilibrium. During the 100-year spin-up period, the time series of surface DIC concentration shows a continuous decreasing trend, indicating that it has not yet reached equilibrium. This decreasing trend is still observed in the MOM5–TOPAZ results using ERA5 data from 2000 to 2014. The results in NPRT, which use biogeochemical variables that have not reached equilibrium as initial and boundary conditions, have limitations in reproducing interannual variability similar to observations. Therefore, this study focused on analyzing seasonal variability rather than interannual variability of biogeochemical variables.

(2) River input and atmospheric deposition data are also applied using climatological data in NPRT, since it is difficult to obtain data with actual interannual variability.

The above content has been added to the revised manuscript (lines 390–392).

7) Temperature, salinity, and/or nutrient timeseries can be compared with KODC datasets as well as.

Answer) As suggested by the reviewer, it is also necessary to evaluate water temperature, salinity, and/or nutrients using observational data provided by the KODC datasets. However, the observational area in the KODC is limited to coastal regions around Korea. To accurately reproduce areas such as marginal seas, it is necessary to set numerical models that account for the specific characteristics of each region. In this study, the focus was primarily on evaluating the open ocean such as the Northwestern Pacific. In future research using this model for marginal seas, we will conduct analyses in conjunction with model experiments.

8) N: P ratio should be compared to see major limiting factors.

Answer) We compared the distribution of the N:P molar ratio in seawater simulated by NPRT with observations (Figure A2-2). Compared with that in the observations, the N:P ratio in WOA18 and NPRT were approximately 3.4 and 4.3 in the subtropical region, respectively. However, NPRT shows a clear positive bias in the Kuroshio Current and its extension region. In the YECS, the N:P molar ratio In WOA18 was approximately 7.3, with the ratio exceeding 16.0, particularly at the Yangtze River estuary, where a distinct phosphate limitation region is observed (Fig. A2-3a). However, the simulated N:P ratio was relatively lower than the observed values in the Yangtze River estuary (Fig. A2-3b). Despite the significant influence of river discharge in regions such as the YECS, the relatively low N:P molar ratio simulated in this study is likely due to inaccuracies in the estimated nutrient input from riverine sources.

The above content has been added to the revised manuscript (lines 365–382; Fig. 11).

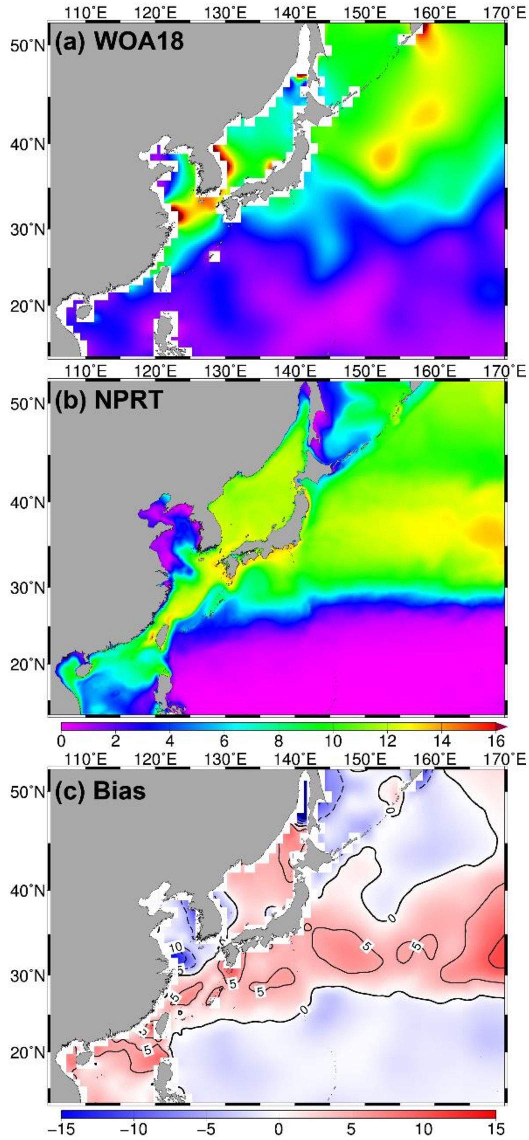

Fig. A2-3. Distributions of the annual mean surface N:P molar ratio in (a) satellite altimeters (CMEMS), (b) NPRT, and (c) biases between NPRT and observations (WOA18) from 2005 to 2014.

9) Please add references which nutrients are limited in those regions.

Answer) The study area, the Northwestern Pacific, can be broadly divided into three limitation regions: the subtropical and subarctic regions and the YECS, which are characterized by nitrogen limitation (Browing et al. 2022), Fe limitation (Watson, 2001; Zhang et al 2021), and phosphate limitation (Lee et al. 2023), respectively.

The above content has been added to the revised manuscript (lines 365–368).

[Reference]

Browning, T. J., Liu, X., Zhang, R., Wen, Z., Liu, J., Zhou, Y., Xu, F., Cai, Y., Zhou, K., Cao, Z., Zhu, Y., Achterberg, E. P., and Dai, M.: Nutrient co-limitation in the subtropical Northwest Pacific, Limnol. Oceanogr. Lett., 7, 52–61, https://doi.org/10.1002/lol2.10205, 2022.

Lee, D.-G., Oh, J.-H., Noh, K. M., Kwon, E. Y., Kim, Y. H., and Kug, J.-S.: What controls the future phytoplankton change over the Yellow and East China Seas under global warming?, Front. Mar. Sci., 10:1010341, https://doi.org/10.3389/fmars.2023.1010341, 2023.

Watson A. J.: Iron limitation in the oceans. In The Biogeochemistry of Iron in Seawater, Turner DR, Hunter KA (eds). Wiley: Chichester; 9–39, 2001.

Zhang, H.-R., Wang, Y., Xiu, P., Qi, Y., and Chai, F. (2021b). Roles of iron limitation in phytoplankton dynamics in the Western and Eastern subarctic pacific, Front. Mar. Sci., 8. https://doi.org/10.3389/fmars.2021.735826, 2021.

10) Describe primary productions and carbon cycles like air-sea co2 flux, pCO2, dissolved inorganic carbon, alkalinity, pH, aragnoite, calcium carbonate saturations and carbon exports (pe-ratio; laufkotter et al 2015, https://bg.copernicus.org/articles/12/6955/2015/bg-12-6955-2015.html).

Answer)

(1) Primary productions

In this study, while various biogeochemical variables are calculated within TOPAZ of NPRT v1.0. In TOPAZ, there exist many secondary diagnostic variables in addition to the primary diagnostic variables (11 fundamental diagnostic variables, see Fig. 1). However, since TOPAZ does not directly compute net primary production (NPP), Laufkötter et al. (2015) calculated NPP $(\mu_i \cdot P_i)$ using the secondary diagnostic variables ($\mu_{max}, T_f^p,\ N_{lim},\ and\ \ L_{lim}$) as follows:

$$\Gamma(P_i) = (\mu_i \cdot P_i) - grazing - sinking - other\ \ losses \qquad (1)$$

where Γ is the sum of the time rate of change and the physical processes of advection, convection, and diffusion. μ_i and P_i indicate the growth rate of phytoplankton and phytoplankton biomass, respectively.

$$\mu = \ \mu_{max} \cdot T_f^p \cdot N_{lim} \cdot L_{lim} \qquad (2)$$

where μ_{max} indicates the maximum growth rate, and $T_f^p,\ N_{lim},\ and\ \ L_{lim}$ represent temperature, nutrient, and light limitation factors, respectively.

In this study, the secondary diagnostic variables related to air-sea gas flux have been stored as NetCDF files in the NPRT output. However, the secondary diagnostic variables related to NPP have not been stored in the model output. In the future, we plan to modify the model to store these variables and directly calculate NPP. Therefore, although we could not analyze NPP directly, we aim to evaluate it indirectly using chlorophyll biomass.

The above content has been added to the revised manuscript (lines 144–150).

(2) Carbon cycle

In this study, the simulated surface distributions of DIC and alkalinity show a predominantly positive bias compared to observations, except in the YECS. We investigated the cause of this issue.

1) The positive bias in surface alkalinity generates favorable conditions for a large carbon accumulation in the ocean from the atmosphere. As a result, the CO_2 flux is predominantly absorbed from the atmosphere to the ocean, leading to a positive bias in surface DIC.

2) In NPRT, not only DIC but also Alk exhibits a predominant positive bias. To investigate the cause of this, we focused on the initial conditions and boundary conditions. The initial and boundary conditions applied in NPRT were derived from the MOM5–TOPAZ model results. The 100-year spin-up period is insufficient for variables such as DIC to reach equilibrium state. As a result, DIC and Alk in the initial and boundary conditions show a positive bias compared to observation, which is believed to influence the NPRT results.

These findings have been incorporated into the revised manuscript to provide a more detailed explanation of the carbon cycle in NPRT (lines 464–493; supplementary S3).

[Reference]

Laufkötter, C., Vogt, M., Gruber, N., Aita-Noguchi, M., Aumont, O., Bopp, L., Buitenhuis, E., Doney, S. C., Dunne, J., Hashioka, T., Hauck, J., Hirata, T., John, J., Le Quèrè, C., Lima, I. D., Nakano, H., Seferian, R., Totterdell, I., Vichi, M., and Völker, C.: Drivers and uncertainties of future global marine primary production in marine ecosystem models, Biogeosciences, 12, 6955-6984, https://doi.org/10.5194/bg-12-6955-2015

11) Please describe differences between biogeochemical states of GFDL-ESM2M in Dunne et al 2013b (https://journals.ametsoc.org/view/journals/clim/26/7/jcli-d-12-00150.1.xml?tab_body=fulltext-display) and this result. ESM2M results published in IPCC AR5 should be enough to compare.

11-1) Please describe oceanic front locations and resolving eddies that would be typically strong enough in regional setting compared to global model ESM2M as well as phenology.

Answer) As suggested by the reviewer, we investigated the biogeochemical state of GFDL-ESM2M and GFDL-ESM2G. Dunne et al. (2013) analyzed the spatial patterns of chlorophyll and nutrient concentrations in GFDL-ESM2M and GFDL-ESM2G data. Unfortunately, we were only able to obtain chlorophyll concentration data. Therefore, we investigated the chlorophyll concentrations of NPRT, MODIS, CMIP6, GFDL-ESM2G, and GFDL-ESM2M datasets (Section 3.2.3; Figs. 13 and 14; Table 3 in the revised manuscript). Furthermore, we were unable to analyze oceanic physical phenomena such as front locations and eddies.

R3

1) The authors develop a high-resolution regional ocean biogeochemical model for the Northwest Pacific Ocean using ROMS and TOPAZ. The model is run for the years 2000-2014 and the physical-biogeochemical results are compared to a number of observational data products. The comparison illustrates variables and regions where the model performs well, but also identifies some model biases, particularly in the biogeochemistry. Overall, I find the model results interesting, and I think the model-data comparison presented is somewhat basic, as each model variable comparison is performed in isolation, with no broader connection of the underlying model biases. There are also some sections that I think would benefit from a more detailed description. I've described these comments in more detail below.

The model comparisons to the observational datasets presented provide an overview of the model biases, but it's a somewhat preliminary comparison without any substantive description of how the various biases connect and what that means for their use in future applications. Because it's a physical-biogeochemically coupled model, the biases in 1 variable will direct translate to other, mechanistically connected variables. Thus, the authors can describe how the bias in something like surface temperature, may impact primary productivity and also surface oxygen due to thermal solubility. I've provided a more detailed example of this later on, when trying to deduce the source of the bias in DIC, but I would encourage the authors to delve a bit deeper into their model results. Especially considering that this is a model descriptive paper in a model descriptive journal, a more thorough description of the model biases and how they support or reinforce biases in other variables is warranted.

Answer) The constructive suggestions from the reviewer provided a valuable opportunity to better understand the causes of bias in our model results. By analyzing the surface temperature bias, we were able to understand why the DO simulated in NPRT predominantly exhibits a negative bias. The simulated sea surface temperature in the model predominantly exhibits a positive bias, with a particularly significant positive bias (> 4°C) observed in the Kuroshio–Oyashio confluence region (Fig. A3-1a). In contrast to the surface temperature bias, the surface DO simulated in NPRT predominantly exhibits a negative bias. A significant negative bias in DO is also distinctly observed in the Kuroshio–Oyashio confluence region. The spatial correlation coefficient between the two variables shows a distinct negative correlation (approximately -0.90; Fig. A3-1). In other words, the negative bias in surface oxygen is likely reproduced due to thermal solubility.

In addition, temperature influences net primary productivity (NPP). However, the NPRT used in this study does not directly calculate the NPP, although it does compute secondary diagnostic variables related to the NPP. According to Laufkötter et al. (2015), the NPP was calculated using these secondary diagnostic variables. Based on this approach, we plan to compute the NPP in future studies and register it as a variable stored in the NetCDF file. We thoroughly analyzed the reasons for the bias in chlorophyll simulated by NPRT by investigating uncertainties in satellite observations, such as the influence of CDOM (Kim et al. 2016).

The above content has been added to the revised manuscript (lines 258–261, 443–460, 505–507, and 517; Figs. 4 and 17 in the revised manuscript).

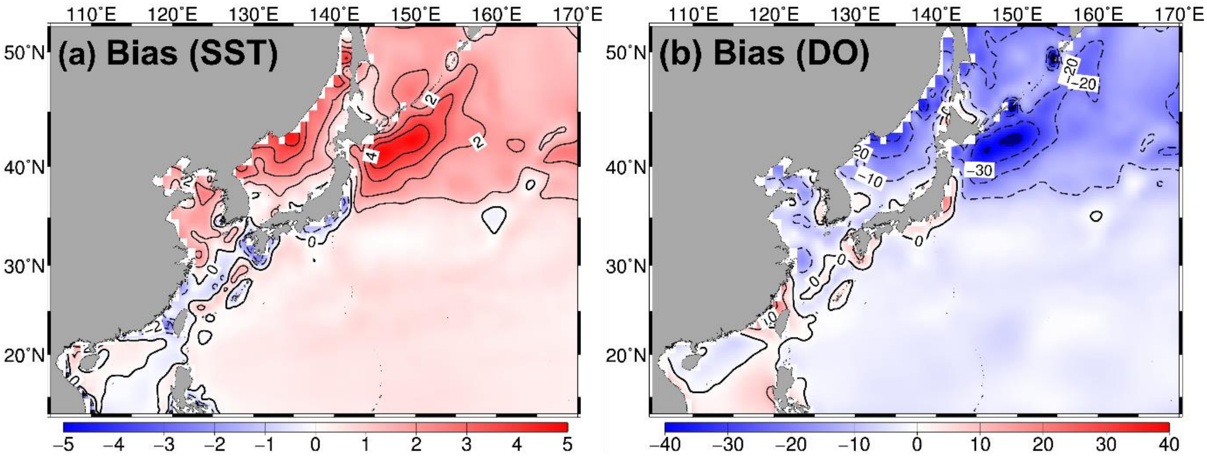

Fig. A3-1. Distributions of annual mean bias in the sea surface temperature (SST) and dissolved oxygen (DO).

[Reference]

Kim, G.E., Gnanadesikan, A., and Pradal, M.A.: Increased surface ocean heating by colored detrital matter (CDM) linked to greater Northern Hemisphere ice formation in the GFDL CM2MC ESM. J. Climate, **29**(24), 9063–9076. https://doi.org/10.1175/ jcli-d-16-0053.1, 2016.

Laufkötter, C., Vogt, M., Gruber, N., Aita-Noguchi, M., Aumont, O., Bopp, L., Buitenhuis, E., Doney, S. C., Dunne, J., Hashioka, T., Hauck, J., Hirata, T., John, J., Le Quéré, C., Lima, I. D., Nakano, H., Seferian, R., Totterdell, I., Vichi, M., and Völker, C.: Drivers and uncertainties of future global marine primary production in marine ecosystem models. Biogeosciences, 12, 6955-6984, https://doi.org/10.5194/bg-12-6955-2015, 2015.

2) The description of the model spin-up for the initial and biogeochemical boundary conditions in section 2.4 is confusing to me. First, the terminology throughout the description appears to be inconsistent and not specific at points. There are various models described as MOM-TOPAZ, MOM5-TOPAZ, and NPRT, which itself is also MOM5-TOPAZ. It's unclear to me if there are either 1, 2, or 3 total models here. My interpretation is that they are running a global simulation of MOM5-TOPAZ for 100 years, which is then used as the initial condition and boundary conditions for the NPRT run. But this is very unclear to me overall and I am likely incorrect in this interpretation. Considering the 100 year spin-up conducted, I think it's appropriate to see some analysis of any drift in biogeochemical variables over this timeframe. As the authors mentioned, running a 100 year spin-up at potentially global resolution is a huge computational effort, and justifies some explanation and description of the results. I'm also curious to see this considering the large DIC bias that I discuss in more detail in the next paragraph. While longer spin-up times are recommended for biogeochemical variables, that doesn't mean that the results will inherently be more accurate than a short spin-up time, given the possibility for significant model drift.

Answer) As noted, the MOM5–TOPAZ global simulation was run for 100 years using the 2000 forcing fields. Then, using the final step of this 100-year simulation as the initial condition, MOM5–TOPAZ was rerun for the 2000–2014 period to generate the biogeochemical variables needed for applying the initial and boundary conditions in NPRT (coupled ROMS–TOPAZ). The term "MOM–TOPAZ" in the main text refers to the MOM5–TOPAZ model, and it has been uniformly revised to "MOM5–TOPAZ" in the revised manuscript (lines 176 and 178). Furthermore, the causes of the DIC bias reproduced in NPRT were analyzed in conjunction with the following questions.

3) I'm a bit concerned with the apparent DIC bias in the model for a couple of reasons. First, it's a relatively large bias, with a R value of 0.41 and a RMSE value of nearly 100 µmol/kg. In fact, the discrepancy is large enough that it's easily visible, even without looking at the difference plot in Figure 13c. It's the large, positive bias for nearly the entire model domain that I'm specifically concerned about, as opposed to the negative bias within the YECS (which will be tougher for the model to capture due to the freshwater influence, and the coarser GLODAP product itself is likely far less reliable there). Second, is that the positive DIC bias does not appear to align with some of the other biogeochemical biases described. For example, my first thought was that it was maybe due to a positive salinity bias, considering how relatively spatially homogenous it is. But, Figure 6 suggests that the model is actually fresher than the observations, which should impart a negative DIC bias. Thus, if Figure 13 was salinity normalized, the positive bias would be even greater. Looking at primary productivity next, it appears that, if anything, the model is biased high with respect to productivity. This is based on higher chlorophyll values (Figure 8a), a greater nutrient drawdown (Figure 11), and greater oxygen concentrations (Figure 16c) than the observations. But again, higher model productivity would generate a negative DIC bias, not positive. So, what exactly is going on here? Are the DIC boundary conditions too high? Are the alkalinity concentrations too high, thereby generating more negative $\Delta pCO2$ concentrations and a large carbon accumulation from the atmosphere over the 100 year spin-up? Without a more mechanistic description and understanding of this bias and additional information regarding the model alkalinity, pCO2, and CO2 flux patterns, I would be really hesitant to use this output for any application involving the carbonate chemistry.

Answer) We investigated the reasons for the significant positive DIC anomaly noted and identified several possible factors, which are as follows:

(1) A 100-year spin-up period is not sufficient for the surface DIC concentration to reach equilibrium. Consequently, the initial and boundary conditions derived from the MOM5–TOPAZ model results exhibit significantly higher DIC concentrations than the observations. In particular, the DIC concentration at the eastern boundary is higher than that observed, which may contribute to the

elevated DIC levels in the subtropical region due to the influence of the North Equatorial Current and the Kuroshio **(**red solid line in Fig. A3-2**).**

(2) The alkalinity in the ROMS–TOPAZ model also exhibited a significant positive bias, except in the Yellow and East China Seas (Fig. A3-3), which may be attributed to a large amount of carbon accumulation from the atmosphere.

The above relevant content has been added to the revised manuscript (lines 464–493 and 561–564; Figs. 15-16, and supplementary Figure S3 in the revised manuscript).

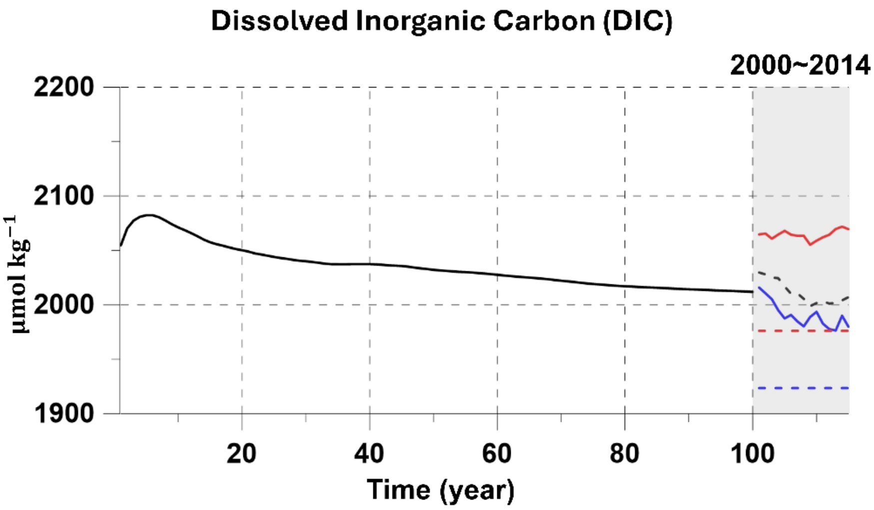

Figure A3-2. Time series of the annual mean surface DIC concentration (μmol kg^{-1}) in the NWP (the subtropical and subarctic regions) during spin-up time using MOM5–TOPAZ. The spin-up process for the first 100 years (black solid line) was conducted under a pCO_2 (369.6 ppm) and atmospheric forcing (ERA5) conditions in 2000. The MOM5-TOPAZ model was then run for the period 2000–2014 (black dashed line; light gray shaded region), using the final time step of the 100-year spin-up as the initial condition. Through these spin-up processes, the initial and boundary conditions for the biogeochemical variables required to conduct NPRT were obtained. After the 100-year spin up time, the red solid and red dashed lines indicate the zonal-averaged surface DIC concentrations applied to NPRT and observations from GLODAPv2 at the eastern boundary conditions, respectively. The blue lines are the same as the red solid and dashed lines, but for the values at the southern boundary conditions.

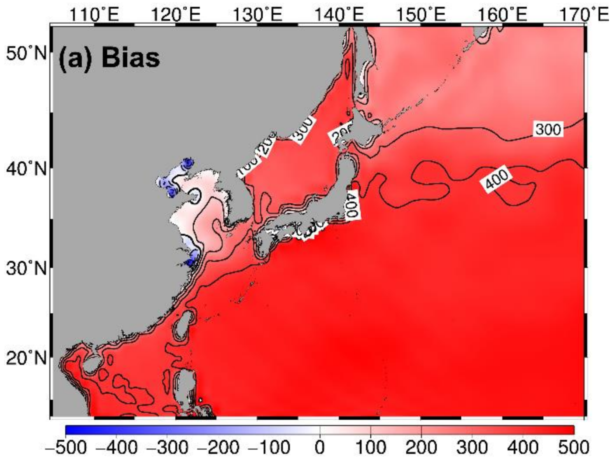

Figure A3-3. Horizontal distributions of total alkalinity bias (μmol kg^{-1}) between NPRT and observation (GLODAPv2).

4) Overall, I suggest that the authors incorporate more specific language and examples into their analysis, and remove a number of generalized sentences that are either too broad and/or non-specific. I'll cite some specific examples under the "specific comments" section below, but a number of occurrences are the first sentence of a new section (e.g. abstract, discussion and conclusion) and they can pretty much all be completely removed.

Answer) As suggested by the reviewer, we revised the manuscript by removing generalized or non-specific sentences, i.e., the first sentence in the abstract, introduction, and conclusions (lines 17, 40–43, and 450–451 in the original manuscript).

5) Figures 3-6 would greatly benefit by having difference plots included, as otherwise it's difficult to see the comparison when the variability across the whole region is relatively large. The authors include difference plots for the remaining model-data comparisons, so it seems like it was missed for the earlier comparisons?

Answer) We have added bias figures (Figs. 3, 4, and 6 in the revised manuscript) as suggested by the reviewer.

Specific comments:

1) The first sentences of the abstract and the introduction are overly generalized to the point of being meaningless.

Answer) We have incorporated the reviewer's feedback by removing the first sentence of both the abstract and introduction.

2) Lines 154-157: mentions a surface flux from the atmosphere for alkalinity? I don't recall this being a common process in NPZD models, so could the authors describe this a little more?

Answer) As pointed out by the reviewer, alkalinity is not a variable that is directly forced via surface flux in TOPAZ.

The equation related to surface alkalinity in the TOPAZ is as follows.

stf_alk = - dry_no3 - wet_no3 (1)

where dry_no3 and wet_no3 represent dry and wet deposition, respectively. As shown in the above equation, surface alkalinity (stf_alk) is a biogeochemical variable influenced by reaction with nitrate (NO_3) entering through the surface flux. Since alkalinity (Alk) is not directly supplied by surface flux, we thought it could confuse. To avoid potential misunderstandings, we removed "Alk" from biogeochemical variables for the atmospheric deposition in the revised manuscript (lines 157–158).

3) Lines 154-157: Where is the climatological data for these air-sea fluxes coming from and what are the seasonal variations? More description is required here.

Answer) The climatological data, provided from MOM's GitHub (https://mom-ocean.github.io/docs/quick-start-guide/), for air-sea fluxes applied in this study are as follows.

- FED: Soluble_Fe_Flux_PI.NC
- LITH: Mineral_Fe_Flux_PI.nc
- NO3/NH4: depflux_total.mean.1860.nc

There is no specific explanation regarding the period provided in the manuals or on the websites offering the climatological data. The data are described only as the pre-industrial simulation results from the Global Chemical Transport Model (GCTM) developed and operated by the National Centers for Environmental Prediction (NCEP). The data is provided as monthly mean data.

The above content has been revised in more detail in the revised manuscript (lines 159–160).

4) Line 266-267: Suggest the authors expand on this statement. How are they attributing the inadequate dispersion to an insufficient model integration time?

Answer) As suggested by the reviewer, we analyzed the bias in physical variables to identify the cause of the overestimated salinity in NPIW. Among various factors, we found that surface salinity is overestimated in the Kuroshio region.

The above content has been revised in more detail (lines 270–272 in the revised manuscript).

5) Line 282: Missing a "respectively" at the end of this sentence.

Answer) done (line 287 in the revised manuscript)

6) Line 297-298: Do the authors have evidence to support the statement that the observational uncertainty for satellite-derived chlorophyll is lower than nutrients? I'm skeptical of this claim.

Answer) Unlike in-situ measurements. satellite observations have the advantage of covering a wide area simultaneously, making them an extremely useful observational tool. However, satellite data are significantly influenced by factors such as atmospheric conditions and reflectance (Clow et al. 2024), resulting in some bias compared to in-situ measurements. In particular, for chlorophyll concentration, satellite observations tend to overestimate concentration in coastal regions where CDOM and suspended particles concentrations are high (Hawes et al. 2000).

In-situ measurements are greater accurate than satellite observation but have the disadvantage of insufficient data coverage at the same time. In addition, when in-situ measurement data are confined to a specific region, they may have limitations in representing the surrounding areas.

Nutrients observations are mostly obtained through in-situ measurements, which results in significant spatial and temporal limitations on data, except in specific regions. We are concerned about the uncertainties associated with nutrient observations due to these limitations.

As pointed out by the reviewer, it cannot be assumed that satellite chlorophyll observation data always have smaller uncertainties than nutrient observations. We decided to remove this sentence in the revised manuscript as it could cause confusion.

[Reference]

Clow, G. L., Lovenduski, N. S., Levy, M. N., Lindsay, K., and Kay, J. E.: The utility of simulated ocean chlorophyll observations: a case study with the Chlorophyll Observation Simulator Package (version 1) in CESMv2.2, Geosci. Model Dev., 17, 975–995, https://doi.org/10.5194/gmd-17-975-2024, 2024.

Hawes, S.K., K.L. Carder, R.H. Evans, "MODIS CDOM and chlorophyll: a first look using SeaWiFS and AVHRR data", in Earth Observing Systems V, W.L. Barnes, Editor, Proc. SPIE 4135, 403-410 (2000).

7) Lines 304-325: This is a confusing paragraph because it appears to be conflating previous literature assessments with the model results. It makes it difficult to understand whether the authors are describing their model's phytoplankton dynamics, or what's already been described in previous studies. Suggest removing the descriptions from prior work and placing it in the introduction and/or discussion.

Answer) As suggested by the reviewer, we removed the discussion on phytoplankton dynamic described in previous studies and focused on the evaluation of NPRT developed in this study. (section 3.2.3 in the revised manuscript)

8) Lines 330-334: The authors here are comparing how riverine nutrient fluxes are handled in global models vs. their model, but it's very confusing what the actual difference is. "Similar to that adopted in the regional biogeochemical model". Which model is this referring to? Here's where more specific language would benefit the description.

Answer) In response to the reviewer's opinion, a more detailed description of the method for applying river effects was provided in the revised manuscript (lines 430–437).

9) Lines 340: Is an R value of 0.45 really sufficient for capturing the annual mean distribution? That seems somewhat low to me, so I would suggest providing more support here.

Answer) The spatial correlation coefficient in NPRT was approximately 0.45, which is somewhat low as pointed out by the reviewer. The spatial correlation coefficient was similar to that of the CMIP6 data, but the TD score was approximately 2.32 in NPRT, which was lower than the TD score of CMIP6 (2.23–3.50). For this reason, it is considered that the spatial pattern of chlorophyll concentration in the YECS reproduced by NPRT was simulated more effectively compared to the CMIP6 data.

The main content has been revised accordingly based on the aforementioned content (lines 424–442).

10) Lines 418-425: I'm not following the last couple sentences here. So, the model has a significant negative bias in DIC compared to GLODAP. But then the next sentence says the model results are consistent with observational data from ships? What observational data from ships is this referring to? The only comparison is to GLODAP, which does incorporate ship data, but the authors phrase this as if they are discussing a completely different product that the model does compare well with. Then the final sentence attributes this to the use of a high-resolution model? Why? Need to be more descriptive here.

Answer) As pointed out by the reviewer, we concluded that the last paragraph caused confusion due to ambiguous expressions, and therefore, we have removed it. Additionally, based on the constructive suggestion regarding DIC analysis, we investigated the causes of significant deviations in DIC (initial and boundary conditions and total alkalinity distribution) and have revised the manuscript to provide more detailed explanations on this matter (section 3.2.4 in the revised manuscript).

11) Lines 450-451: This sentence is overly generalized to the point of being meaningless.

Answer) We removed the sentence from the revised manuscript.

12) Line 472: Replace period with comma

Answer) done (line 557)

13) Lines 485-489: This is an odd description to me, because it starts with a very targeted, tractable method for how model uncertainty within the YECS could be reduced by developing an improved understanding of the riverine runoff and nutrient concentrations. But then, states that this would be too difficult, and instead lists a number of vague strategies that could also apply for essentially any model application. Why would the first method be too difficult? Is it because the authors are already confident with what they're currently using for the river outflow and nutrients here, and thereby think that they will have minimal gains by trying to incorporate more? If so, then this justification would be extremely helpful.

Answer) First, we apologize for any confusion caused by the use of ambiguous expressions as noted by the reviewer. We have removed the ambiguous sentences and described the potential for nutrient biases, particularly the distribution of a positive bias in silicate.

We have presented two possible explanations for the distribution of positive bias.

1) Unlike other nutrients, silicate is taken up only by large phytoplankton. Therefore, its consumption may be significantly lower than that of other nutrients.

2) In the YECS, where vertical mixing is strong due to tidal forcings, silicate may not sink to bottom but instead remain in the mixed layer, leading to silicate accumulation.

For the reasons mentioned above, in marginal seas such as the YECS, it is essential to consider the environmental conditions of specific regions and adopt appropriate parameters and external forcings.

The above content has been revised in more detail in the revised manuscript (lines 337–345, 549–553).

14) Line 642: Should the Jung et al., reference be 2019 rather than 2020? It's listed as 2019 in the main text.

Answer) We revised 'Jung et al. 2019' to 'Jung et al. 2020' in the revised manuscript (lines 154, 158, and 388).