

We thank Reviewer #1 for their constructive comments. We have listed their comments in bold below and our responses in normal formatting.

Reviewer #1

Firstly, the model only infers the atmospheric CO₂ response based on the published relationship between the oceanic inventory change in preformed PO₄ and the atmospheric CO₂ change. The near-linear relationship between the two is only valid if we assume no solubility and air-sea CO₂ disequilibrium effects. I acknowledge that the effects might be implicitly included in the empirical relationship extracted from a few previous global studies. However, I am not sure if the same relationship can be applied to the regional perturbation study. For example, any perturbations in the Southern Ocean (e.g., the ACC band where air-sea CO₂ disequilibrium is large due to the short surface residence time of upwelled waters) might not lead to the atmospheric CO₂ response proportional to the preformed PO₄ response. This point can be especially worrisome because the most sensitive regions turn out to be the ACC band in the study.

Thank you for highlighting this. We have replaced the published preformed PO₄ / CO₂ relationship with one calculated explicitly for this model using an online carbon cycle.

We have tested the impact of CO₂ disequilibrium on the sensitivity results by running an ensemble of experiments where the Martin curve in each region is perturbed individually whilst all others are kept at the control value. The runs have a carbon cycle and so atmospheric CO₂ is predicted as a response to the changing biogeochemistry. The preformed PO₄ and CO₂ from each regional perturbation are plotted over the values from global perturbation experiments (Figure 1). There are minor deviations from the global preformed PO₄ / CO₂ relationship suggesting that disequilibrium effects may be present, but these are relatively minor. The range of CO₂ values predicted by the individual perturbation experiments closely matches the sensitivity patterns from the Latin Hypercube ensemble (Figure 2) further suggesting disequilibrium effects have a minor impact on the results.

We have added these results to the supplementary material that are referenced from a brief discussion of disequilibrium effects in the methods section of the manuscript. We highlight in the discussion that disequilibrium effects may be considered in future analyses.

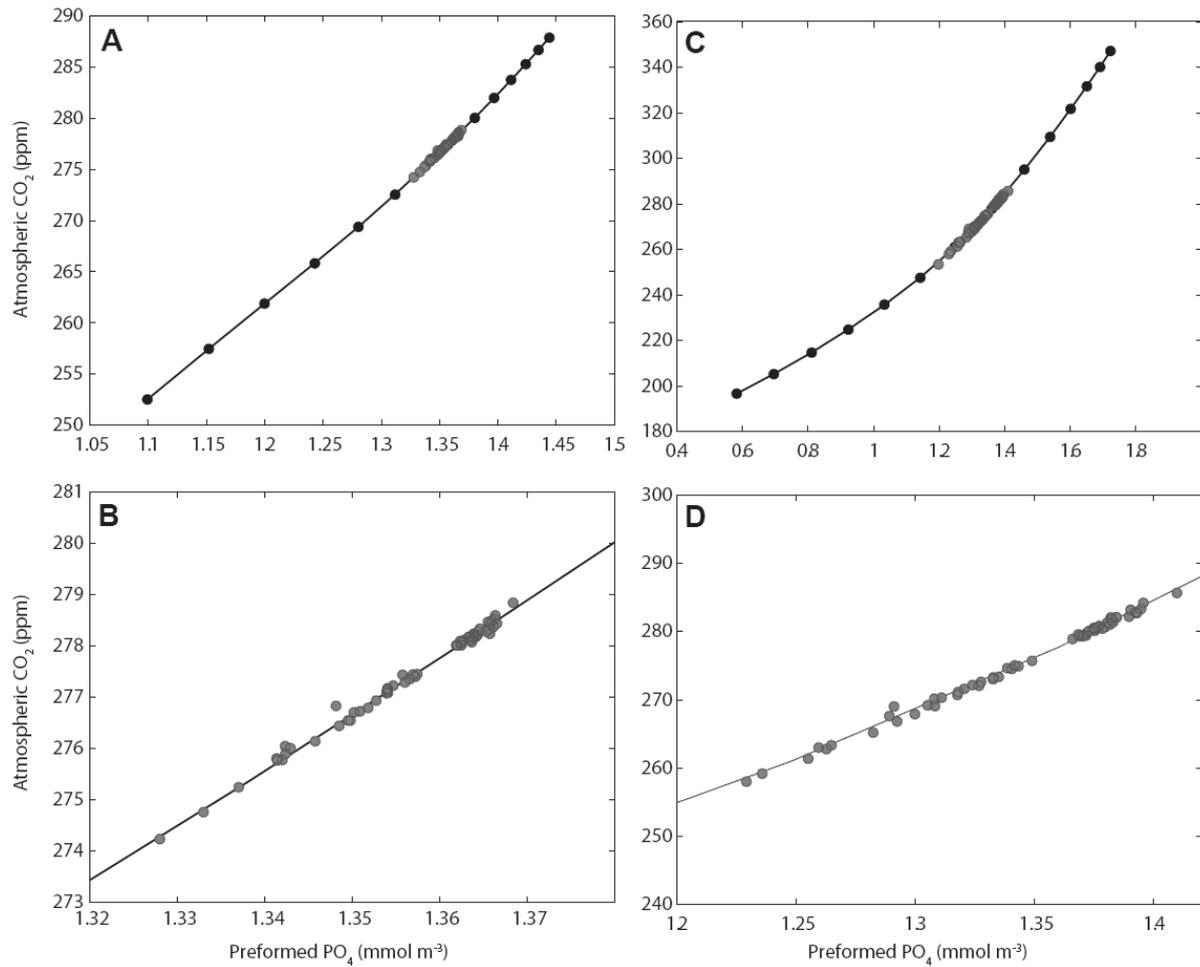


Figure 1. Comparison of CO₂ versus preformed [PO₄] relationships when the Martin curve is varied as a globally uniform value (black line) from -0.4 to -1.6, and when regions are perturbed individually within the same range (grey). Panels (a) and (b) shows results for the constant-export ensemble with the region of interest expanded in (b). Panels (c) and (d) shows results for the nutrient-restoring ensemble with the region of interest expanded in (d).

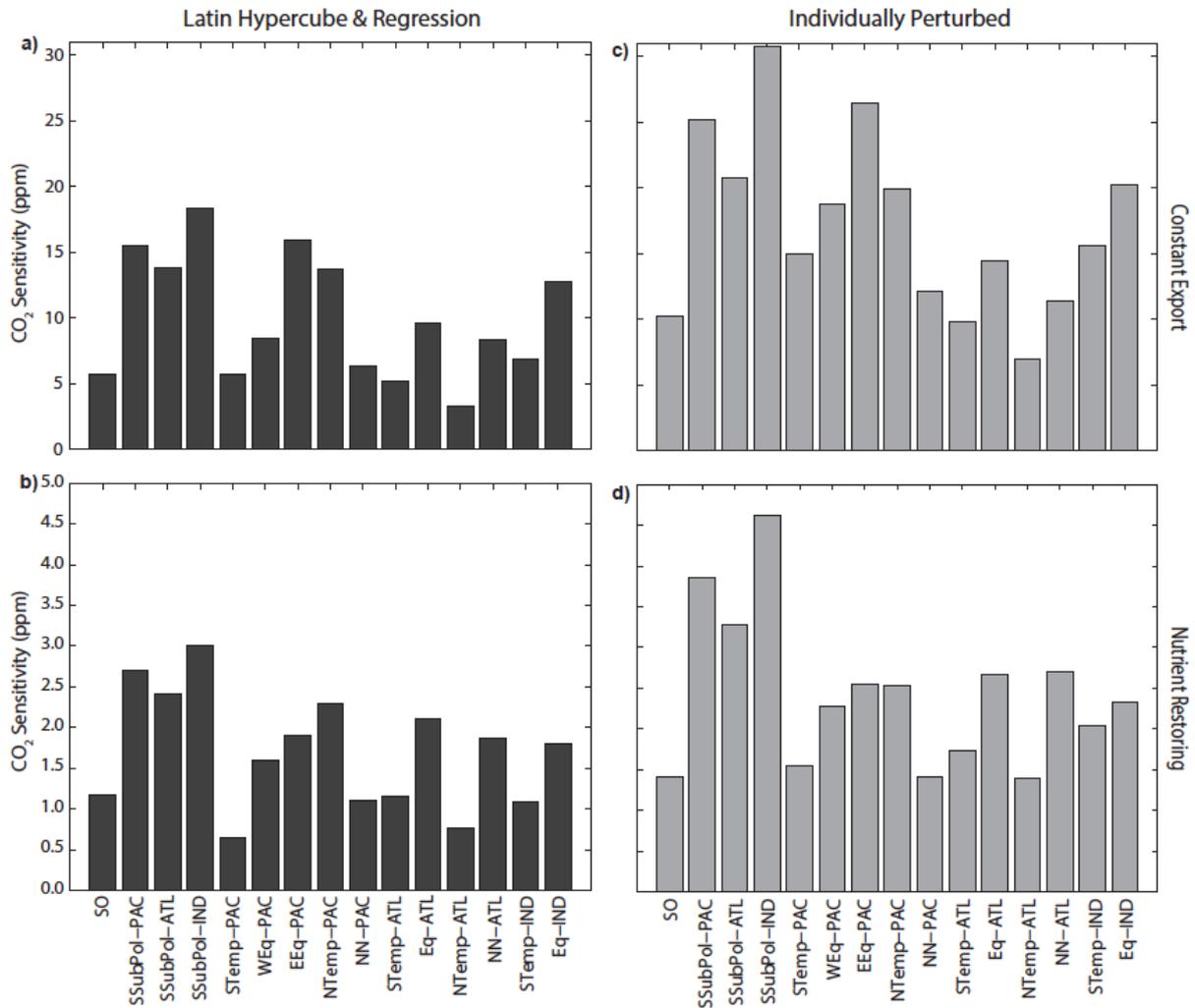


Figure 2. Comparison of CO₂ sensitivity estimates from two methods. Black bars are the Latin hypercube and regression-based sensitivity estimates derived with the statistical relationship between preformed PO₄ and CO₂. Grey bars are the difference between CO₂ when each region is perturbed individually from b=0.2 and b=1.6 and atmospheric CO₂ is calculated explicitly in the model.

Secondly, the sensitivity is estimated using the multiparameter linear regression method applied to the two sets of 200-member ensemble experiments where 15 regional exponents are perturbed simultaneously. Although the method seems sound, I wonder why the sensitivity should be quantified in this way? Are there any merits? Would the sensitivity be the same or different if the authors perturbed the exponent in a region at a time, requiring a total of 15 perturbation experiments for one model scheme? The individual perturbation experiments seem a simpler and cleaner way to quantify the atmospheric CO₂ response to the perturbation and its relations with the export production in each domain.

The reviewer is correct that an alternative approach could be to perturb the remineralisation depth in each region individually (although this would entail at least 4 experiments per region to reasonably characterise the response across the parameter range). The advantage of the Latin hypercube

sampling approach is that the resulting sensitivity combines the direct influence of changing remineralisation in one region of interest plus the joint influence of remineralisation changes in other regions, e.g., Pianosi *et al.*, (2016). Given the uncertainty in the observed spatial distribution of remineralisation and the driving mechanisms, we feel it is important to account for the full range of simultaneous changes in remineralisation depths and keep this as the main result.

As per the previous response, we have run the individual perturbation experiments with an explicit carbon cycle. The Martin curve in each region is perturbed to -0.4, -0.7, -1.3 and -1.6 whilst other regions are maintained at the control value of -1.0. Figure 2 shows ΔCO_2 from the -0.4 and -1.6 experiments for each region. The sensitivity patterns match the results from the Latin hypercube ensemble closely. There are differences in magnitude between the two estimates likely due to the fact that there are minor interactions between regions. However, the patterns between regions are preserved. We have added these results in the supplementary as this will provide important context for the key results and strengthen the analysis and interpretation.

Thirdly, the major novel finding is that the highest sensitivity in atmospheric CO₂ is to the change in remineralization depth in Subantarctic regions due to high export production and the high connectivity to deep water formation regions. I see the reasoning behind it: The export production should be high because the export will determine how much regenerated PO₄ can be affected by the perturbation. The connectivity to deep water formation regions is important because the deep water formation is the main pathway of preformed PO₄ to the ocean's interior and the inventory of preformed PO₄. However, I am not fully convinced by the authors' finding. Both "nutrient restoring" model and "constant export" model show that the sensitivity of atmospheric CO₂ to the remineralization depth change is also high in the "NTemp-PAC" domain (Fig. 3). Yet, the subtropical North Pacific is not a region with high export production nor close to any deep water formation regions. How can it be explained?

Thank you for highlighting this. It is notable that the NTemp-PAC region in general, for this model, does fall along the general trend between export production and sensitivity. In comparison the STemp-PAC region has a much lower sensitivity for a similar export production. This could be related to the age of water masses in the Pacific whereby deeper remineralisation in the STemp-PAC region sequesters organic matter in much younger waters, that will return to the surface ocean faster, thus reducing the sensitivity. We have already noted the range of sensitivity for similar export production in the manuscript, so we have added additional text to discuss this. Overall, because the spatial variability is not overly significant relative to the global uniform variability and because export production is a strong predictor of sensitivity, we have not explored this further.

Similarly, why are the deep water formation regions (i.e., the model NADW and AABW formation regions) not the sensitive regions?

We have added the contribution of preformed PO₄ derived from each region for the control run in Table 1 to help highlight the sensitivity of the deep water formation regions. Kwon *et al.*, (2009) demonstrated that largest changes in preformed PO₄ when changing *b* globally were associated with deep water formation regions. The updated version of Figure 5 (Figure 4 here), shows that this is in response to changes in *b* from globally distributed regions (compare colours across the column for SO and NN-Atl regions).

It may be related to the third point. But I don't understand Figure 5. The sensitivity is normalized to what? What do the authors mean by "mean preformed PO₄ in a region"? Is it the surface PO₄

averaged over each region or the total preformed PO₄ subducted from each region divided by the volume of water subducted from the same region?

The “mean preformed PO₄ in a region” is calculated in a similar way to the global preformed PO₄ concentration: the annual mean surface [PO₄] at the end of a simulation is set as a boundary condition and the transport matrix used to calculate the interior distribution of preformed PO₄. In this case, the annual mean surface [PO₄] is set in the region of interest only and the remaining surface has concentrations of zero. The global mean concentration is then calculated from the interior distribution.

Because the absolute magnitude of regional mean preformed PO₄ varies by an order of magnitude across all regions, it is necessary to normalise the concentrations to allow comparison of regression coefficients using the following relationship:

$$\overline{P_{pre}^{region}} = \frac{P_{pre}^{region} - \min(P_{pre}^{region})}{\max(P_{pre}^{region}) - \min(P_{pre}^{region})}$$

In response to this point and to comments from other reviewers, we have revised the format of the figure (see Figure 4 here). We have also updated the text with the description above and the equation to clarify the method of obtaining preformed PO₄ and normalisation.

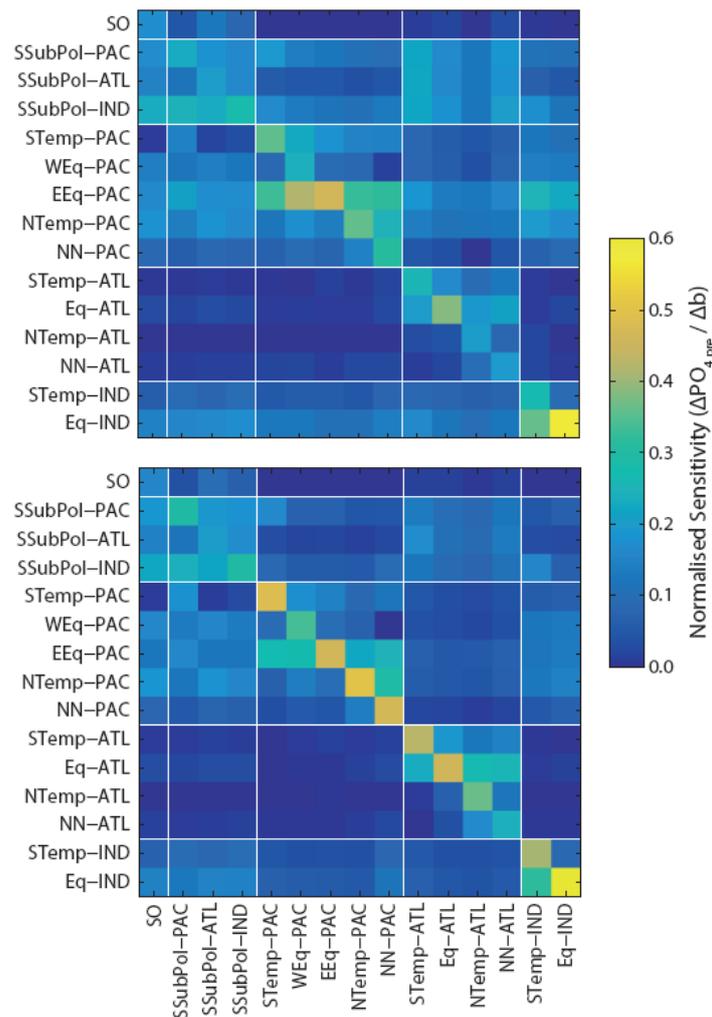


Figure 4. Sensitivity of steady-state normalised mean preformed $[\text{PO}_4]$ exported from each region.

The preformed $[\text{PO}_4]$ from each region is expressed as a function of b using linear regression.

Preformed $[\text{PO}_4]$ is normalised to the range of values in the ensemble to account for large differences in preformed $[\text{PO}_4]$ between regions. The regression coefficients are arranged such that each row shows the impact of changing b in that region on preformed $[\text{PO}_4]$ across other regions. Results from the constant-export and nutrient-restoring schemes are shown in the top and bottom panels respectively.