Reviewer #2 (anonymous)

R: The authors examine the effects of "reciprocal bias compensation", (compensating for errors in the physical model component of an earth system model with changes to the biogeochemical component) using a set of experiments with an earth system model of intermediate complexity. Overall, the manuscript makes an important point that is based on a simple experiment which is easy to follow. However at times it is difficult to follow the manuscript, some important information is missing, while other points are mentioned several times.

A: The reviewer made constructive suggestions targeted at making the paper less repetitive and, further, he asks for (some) missing information. We will follow his suggestions in a revised version of the manuscript (which we deem as being, due to the constructive nature of the reviewer's comment, rather straight forward). We thank the reviewer for his/her time and effort!

R: general comments:
At times, the manuscript seems very repetitive and I have tried to point out most of those instances in my specific comments below. This aspect mostly affects the second half of the manuscript; in fact, in the current version, the "Summary and Conclusions" section is followed by another "Conclusions" section. Section 6 repeats points that have been made in Section 5 which repeats many aspects described in Section 4. A thoughtful restructuring of the last sections would make the manuscript more concise and much easier to follow.

A: We agree with the reviewer and will go through the text and exclude repetitions.

The phrase "reciprocal bias compensation" is used a lot in the manuscript but the actual bias is never examined. Beside 4 RMSD values in Table 4, differences between the configurations are not quantified. I would suggest to include bias values and more than just 2 variables in Table 4. It should include at least those variables examined in the manuscript (including physical variables) and could further include values for the projection into the future (4 columns: "MIX+ (historical)", "TUNE (historical)", "MIX+ (future)", "TUNE (future)"; variables appear in rows). Such a table would show the effects of reciprocal bias compensation on the different variables in one place, which would be very beneficial to the reader.

A: This is a very good suggestion and we will extend the respective Table as suggested.

How were the parameters chosen for the spinup experiment used to determine the parameters for TUNE? I am not against a simple way to determine these parameter values but it would be good to know what ranges were considered and how the values were determined ... grid search, randomly chosen from an interval, latin hypercube sampling? More information is needed.

A: Agreed! We will add the respective information.

One interesting experiment would be to include the nkappa_h (vertical diffusion) parameter in an updated TUNE experiment, just to see if in this very simple case (all error is one parameter) the original value of nkappa_h could be recovered or if corrections to the biological parameters and thus reciprocal bias compensation would prevail. No further experiments (generating new projections into the future) would be necessary for this experiment.

A: Agreed! We would, however, like to combine these types of experiments with a new cost function which gives some measure of the realism of the effective mixing. Our current research indicates that the saturation state of noble gases such as argon may be suited to provide such a measure. We will add a respective outlook to the discussion section of our revised manuscript.

R: specific comments
A: We thank the reviewer for investing all this time! All the comments are very constructive and helpful! We will follow all suggestions made by the reviewer.