

Dear Marco,
many thanks for your revised manuscript. It has been reviewed by two scientists again. It has improved a lot. But I would like you to carry out a few further amendments (see below). I look forward to your revised version of this pioneering article and appreciate your endurance.
Best wishes, Christoph

REVIEWER #1

General comments

The manuscript is overall much better than the first version and, in my opinion, only requires a few minor adjustments before publications.

The introduction reads substantially better than version 1 and, this time, really sets the stage to an interesting paper. One major improvement over the last version is the way the various experiments are used and presented. This is now much more logical and helpful. There are still a few sections where additional explanations would be desirable (see specific comments), but overall, the structure is good.

I still think that reliance on only the Atlantic data for assessing statistical model fit is a poor choice and this leads to some convoluted explanations in the main text, but this is no a major issue at this point. I respect this is the author's choice, even I believe it is a bad choice. I would simply suggest that the authors make it as clear as possible throughout the manuscript that all fits pertain to one single section in the Atlantic only.

One remaining frustrating detail is that the authors seem to have replaced the word "realistic" with the word "accurate". Unfortunately, replacing a word doesn't fix the problem, which is to overly rely on highly subjective statements of quality. Please reduce this further by being more specific or more quantitative in your statements.

Specific comments

P5, l13, remove "on" from "on shipboard"

P11, footnote (2). Just add that to the main text directly.

P14, l24: "recently measurements" change to "recent"

P15, l16: consider rephrasing this "very soon the [Mndiss] reaches the typical background concentration", maybe "as Mndiss approaches near-constant deep background concentrations quickly such that the NADW plume is no longer discernable."

Section 3.3, p17. Expand this section – this is interesting. Please provide an interpretation as to why Atlantic value don't change but Pacific results do.

P18, l15. Table 6 shows a reliability index of 1.88 for NoThresh, not 2.77.

P17, l20: What is your definition of "reasonably accurate"?

P18, l27: overstatement "only the Reference simulation is accurate". What is your definition of accurate?

P19, l4-5: I don't understand the logic behind that argument about "usefulness": "We think that coupling the model to the sediment would ... maybe not that useful, because the sediment source is large at shallow sediments, while most of the Mn burial occurs near the hydrothermal vents"

P19, l10-11: Does the model handle "small" shelf regions well? Can you make any quantitative argument about how well the model resolution handles shelves?

P19, l12: Why only report on Slomp's maximum values?

P19, l16: "out of proportion in some regions of the ocean". Where is that specifically? Is that only in the Arctic, as alluded to in the next sentences?

P19, l20-21: add a comma in "...Pacific Ocean, where..."

P19, l22: "In the East Pacific Ocean the California Current induces Ekman transport and hence equatorial upwelling". Very puzzling bit of physical oceanography? Some references in support of that statement would be very welcome. I believe wind induces Ekman transport, not the California current. I'm also not aware of physical theories of equatorial upwelling that argue the California current induces it.

P19, l23: "upwelling from OMZ sediments". Maybe "upwelling of water that has been in contact with OMZ sediments"

P19, l24: "This is partly captured by our model". Which part?

P19, l25: "In the South Pacific Ocean this effect is more clear in the data of Resing et al. (2015) (Fig. 8a,b, East Pacific around 20S)." ... and in the model?

P19, l27-29: Better would be to provide a back of the envelop estimate of how much bias may come from not representing fluxes from OMZ sediments.

P20, l21: "and especially at low latitudes". Please substantiate this with a few sentences. The previous discussion was all about the Southern Ocean, not on low latitudes"

P20, l23: What is "the most settling Mn"? Do you mean the particulate Mn fraction that contributes most to the sinking Mn flux is from biological particles?

P20, l28: include, not includes

P22, l29: "because Mn redox does not depend on O₂". Rewrite sentence. What is "Mn redox"?

P22, l31-43: "...For this reason we have not included a dependency on [O₂] to the model...." Consider rewriting these few sentences in a less convoluted way.

P24, l6: remove "e.g." – say what you mean in words instead.

P25, l3: "for an accurate simulation of [Mndiss]". What is your definition of accurate? Replacing "realistic" with "accurate" doesn't remove the problem of relying on subjective statements.

Appendix A, p27,l3: do you refer here as the "Pearson correlation coefficient"? please specify.

Appendix A – Table 6. Why are there only errors for the Reference case and not all cases? For comparison purposes, errors should be calculated on all cases.

Figure 7, caption. What do you mean by the word “by in “the by red” or “the by blue” lines? Probably remove this. Also, would be good to make these lines thicker on the figure. They are very thin, even when the figure is full screen.

Figure 8: x-axis labels and sub-plot titles overlap. Fix spacing.

Figure 5 and 9. Choose a consistent name between GIPY5 or GIPY5_e.

Figure 13, explicitly state in the caption if relative difference is $(\text{ref-low hydro})/\text{low hydro}$ or $(\text{ref-low hydro})/\text{ref}$?

Figure 16, make the colored lines thicker

REVIEWER #2

The manuscript by Hulthen et al. has changed considerably in the revised version, with the two most important changes being that

- a biological cycling of Mn has now been implemented into the model, as a response to remarks by the reviewers. The cycling is parameterized as following the uptake and release of phosphorous from a global biogeochemical model, assuming a constant Mn:P ratio from Twining and Baines (2013). The modeled Mn distribution, as far as I can see, has no effect of phytoplankton growth in the model, i.e. Mn limitation is not included in the model. This is a reasonable first step, but should be mentioned in the model description.
- mostly in response to the second reviewer, the paper now contains a much more detailed description of the manganese observations along the dutch Geotraces section GA02, including a discussion of the methods for these observations.

The inclusion of a biological cycling of Mn in the model is reasonable, and I think it strengthens the paper a lot. I have, however, a bit mixed feeling about the new focus of the manuscript on the GA02 section: Reviewer 2 suggested "showcasing these (unpublished?) measurements in more depth early in the manuscript would help build a sense of expectations with regard to what the model is expected to do or not to do. That could also be used as a 'roadmap' to explain how the paper is organized and why." To me this aim has not yet been reached fully, the observational results (3.1) and the modelling results, especially section 3.2 still stand side by side in a too unconnected manner.

One example for this is the elevated value of Mn in the DSOW overflow, which is visible quite clearly in Fig 7 and discussed over a few sentences in section 3.1. In the modelling part, this feature is never mentioned again, and indeed the color scale in Fig. 9 is chosen in

such a way that it is not even visible in the observations anymore. My expectation is that the model does not reproduce this feature, and that is not even bad; it probably just highlights that the model is missing sediment resuspension, a locally important but probably globally unimportant process. In my opinion, often the most important information in model-data comparisons is where the two do not agree because here one learns about processes.

This is just one example, but my general impression is that the present manuscript does not integrate the observational and modelling parts enough and thus misses what the second reviewer had in mind when he suggested to focus more on observations. The model-data comparison could be made much more precise, and I would argue that the authors should try to do that in a second revision.

One aspect to improve is the "over-reliance on subjective statements to describe the simulations and on imprecise/loose statements" that was already mentioned by the second reviewer. I think the present manuscript still contains many too colloquial statements, but now also in the description of the observations. A few examples are that "very soon the Mndiss reaches the typical background concentration" when talking about NADW or "slightly elevated concentrations in the subsurface are also observed in ... AAIW .. but once again concentrations reach the typical background concentration". Both statements in section 3.1 would be much more precise with some indication where something happens. Likewise, "for a small part" later on the same page is not a helpful description. When comparing the NoBio with the reference run, it is stated that "NoBio generally compares better with the observations in the Pacific Ocean especially with the US EPZT transect" Again, it would be helpful if the authors could be a bit more specific: Where is the improvement, and in what aspect? At the surface or at 500m depth?

Minor comments

Section 2.2.7: the value of 0.4 'derived by Middag et al. (2013)' should probably be $0.4 \cdot 10^{-3}$.

Section 3.1: "Whereas: on line 13 should probably be "As"

Page 16, line 17: maybe include "hydrothermal" before "forcing field"

Page 20, line 23ff: To me this is one of the most interesting results of the new model runs including biology; but why just state this here without giving any numbers? In my first review I already suggested that it would be informative to have an idea on the relative magnitude of the sinking fluxes of authigenic and of biologically incorporated Mn; The authors would make that point much stronger when calculating e.g. the globally integrated fluxes of Mn from Mn_{ox} and from the biological compartment, maybe at 100m depth and at some depth deeper in the water column. Just a suggestion..

The authors might consider using the same software for plotting observational and modelling results; Fig 7 is made using ODV, Fig 9 using ferret, I believe.