Interactive comment on "Evaluation of the Transport Matrix Method for simulation of ocean biogeochemical tracers" *by* Karin F. Kvale et al.

Anonymous Referee #2

Received and published: 15 March 2017

A comparison of on-line and offline models, the latter using the Transport Matrix Method, is long overdue and will be a welcome addition to the literature, even just as an example for a single model. It's an easy paper to review as the requirements for such a paper are just 2 things: a good description of the method; a sensible choice of parameters to compare. I'm keen to see the paper published but there are a few details that I'd like to see addressed before then...

Method:

- I don't have an issue with the choice of model as the paper is effectively an example and I'm not sure I see the value in a much longer and exhaustive paper doing the same with a variety of models, particularly as they are all evolving. However, I don't see the value of the comparison of how the model performs with and without the FCT scheme.

For the purposes of the paper all that is needed is a base model – it doesn't matter if it performs a little less well than another version. Hence, I would either cut Section 3.1 or move it into an Appendix. As another option, if it is argued that Section 3.1 is there to allow comparison of the offline run to the FCT case then this needs to be done more rigorously by taking the spun-up UW3 model and running it onwards on-line with FCT now turned on for the comparison.

- The manuscript is a little vague about the details of the runs for the comparison of the on-line models. The starting point is a 13,000 year spin-up of the on-line UW3 model. The TM is then extracted using an extra year run. Are the offline and on-line models then compared purely on the basis of a single extra year run after the 13,000 year spin-up? I would hope not as unless the off-line perfectly mimics the on-line model there is no guarantee that any transient response of the off-line model will be fast, and the differences may be small purely because the two models have had little time to diverge, particularly if it takes 13,000 years to spin-up. The comparison of the difference between the on-line and off-line models should at least state the time over which the models are run for comparison and this should be at least of order 10 years. My recommendation for minor changes is on the assumption that the models were run for longer than 1 year before comparison. If just one year I think the need to re-run the models for longer would constitute major changes

- The manuscript describes the issue with a leapfrog scheme but is a little vague about the compromise made. Is the TM extracted from just one strand of the leapfrog scheme? Is the usual process of blending the 2 strands side-stepped and ignored for the TM?

Comparison

- The authors suggest that some of the largest fractional errors come from differences in small values. In Fig. 7 though it looks like there are significant errors associated with large values for diazotrophs. I'd like to see an extra column of plots showing the

difference between the on-line and off-line models for these fields as elsewhere in the manuscript.

- The relative errors in Fig 13 are very noisy. It would help to additionally have the contours for +/-0.5 relative error marked on the panels for on-line and off-line zonally averaged fields.

- A couple of minor points but the y-axis labelling on Fig 14 needs an extra decimal place for phosphate and it also seems strange that phosphate diverges (then converges) so quickly from the same starting point if it is just a one year run.

C3