Response to 'Interactive comment on "Assimilating bio-optical glider data during a phytoplankton bloom in the southern Ross Sea" by Daniel E. Kaufman et al.' by Anonymous Referee #2 that was received and published: 1 September 2017

Dear Reviewer #2,

We greatly appreciate your time and effort spent reviewing our manuscript. According to your constructive feedback, we propose changes to clarify aspects of model setup, optimization method, and conclusions. Please find our responses (in blue) following each of your comments below. All line numbers refer to the original submitted manuscript. Thank you again for your review.

Sincerely,
Daniel Kaufman, Marjorie Friedrichs, John Hemmings, Walker Smith

Review comments for the manuscript: "Assimilating bio-optical glider data during a phytoplankton bloom in the southern Ross Sea (bg-2017-258)" by Daniel E. Kaufman, Marjorie A. M. Friedrichs, John C. P. Hemmings, and Walker O. Smith Jr.

The authors present a data assimilation study that optimizes parameters in a one-dimensional biogeochemical model using glider observations in the southern Ross Sea. They show insensitivity of the result to the geographical location of observations, but the optimizing parameters is sensitive to the sampling frequency.

The paper is overall well-written, but I hope the reviewers be able to address comments that I have.

– The procedure can be clarified more. This study utilizes a one-dimensional model for 3D observations. Does the cost function use all the observations and estimate one set of the parameter? Or is there an optimized parameter set for each location? If the first approach is used, do you expect that the optimized parameter values represent the distribution of those obtained by the second approach?

As described on line 258 and in Table 2, the cost function uses all of the observations for experiment #1A. To make this clearer, we propose modifying the sentence in the abstract on line 16 to read "Assimilation of data from the entire glider track …"

We do find that the optimized parameter values from the first approach (using all observations) represent those obtained by the second approach (using observations from different locations), as described in the latitudinal (Expt. 1b) and longitudinal (Expt. 1c) experiments and shown in Table 4.

– The authors argue that the data assimilation performance is sensitive to the observation sampling frequency due to "mesoscale variability". Mesoscale variability also means the variation in space with the scale of O(100km). But it is odd to see that the geographical region does not show a big impact on the performance. Could the author comment on this?

Here we define mesoscale variability as "days-weeks, 1-10 km" (line 59). We had no a priori expectation that the geographical regions would show minor differences in model solutions, however we believe that the minor differences are reasonably explained in section 4.2, and especially in this section's last paragraph. Although in situ observations from previous studies have shown spatial differences on these scales, it has been unclear whether the differences were due to temporal or spatial variations. The assimilation experiments in this study suggest that variability observed on the mesoscale in this geographical region may be more likely due to temporal patterns than spatial differences. Therefore, one could expect that assimilating these different locations would show a bigger impact if the observation times concurrently varied, such as is demonstrated in the cruise-based and satellite-based assimilation cases. On larger scales, however, it is likely that the importance of spatial variability would be greater. For instance, the distinct spatial differences observed by satellites are generally across scales larger than the 1-10 km discussed here.

– By construction, the role of advection is not considered in this study. Can authors comment on the role of advection in this region? Do authors think the insensitivity of the assimilation performance to the geographical location of observations is related to the omission of advection?

Previous studies have suggested that horizontal transport and eddies may be important near island land masses and the Ross Ice Shelf (Gerringa et al., 2015; Li et al., 2017). In this region of the Ross Sea particularly, moorings and modeling have indicated moderate westward currents close to the ice shelf (Keys et al., 1990); nevertheless, advection appears to be weaker as one moves farther from the shelf edge (Dinniman et al., 2003).

One cannot rule out the possibility that the sensitivity of the optimizations to the observations' location could be affected by adding advection to the model. However, this would likely only be the case if there were, in reality, strong horizontal velocity gradients, i.e. differences in advection between the observation locations. A more thorough examination of the role of horizontal advection on modeled dynamics of the phytoplankton assemblage is beyond the scope of the current study, but would be benefitted greatly by contemporaneous and co-located mooring and/or ship-based current measurements.

– line 85. The effort on estimating biological state variables can be listed here. (e.g., Song, H., C. A. Edwards, A. M. Moore and J. Fietcher, 2016: Data assimilation in a coupled physical-biogeochemical model of the California Current System using an incremental lognormal 4-dimensional variational approach: Part 3, Assimilation in a realistic context using satellite and in situ observations. Ocean Model., 106, 159-172.)

It is a good idea to reference Song et al. 2016 here.

– section 2.1: What is the vertical resolution of the model?

To clarify this, we propose adding details of the model setup to section 2.1, on line 108: “The model is configured to focus on dynamics within the euphotic zone with a vertical resolution of 5 m from the ocean surface to 200 m.”

– line 114: The full name of BCO-DMO can be given.

Absolutely. We will expand the acronym to the full name.

– line 115: 5-m vertical binning is done using averages? or weighted average?

Vertical binning of the glider data was accomplished using averages, and to clarify this we propose modifying the sentence to read: “Data spanning the upper 200 m of the water column were binned by means into hourly, 5-m vertical bins.

– Equation for the cost function shows that the observational error covariance is estimated using the standard deviation of the observations. Is this right? I think using standard deviation may overestimate the observational error if the blooms dominate the chloro- phyll variability. If the error levels of the instruments are known, why not use these values?

The misfit contributions are weighted by using the inverse of the standard deviation, similar to other assimilation efforts (e.g. in Experiment #1 in Hemmings and Challenor, 2012; Friedrichs et al., 2006; Xiao and Friedrichs, 2014). If the aim were to estimate observational error, then the increase in variance due to the bloom would indeed likely lead to over-estimation. However the aim here is to weight the misfit contributions of chlorophyll and POC, and there is less impact of the bloom on these relative weights. Generally, a more sophisticated treatment of uncertainty in both the observations and the model is desirable as indicated by Hemmings & Challenor (2012), but such a treatment is beyond the scope of the present study and may not be practical with the available data. It makes sense therefore to initially employ a simple well-established method as we have done, but we recognize that it does have its limitations.

You may notice that the first reviewer raised a similar question as well, and we there also indicate our proposed text for section 2.3, referencing other assimilation studies that have used standard deviations to weight the misfit contributions. Specifically, we propose modifying the text on lines 126-128 as follows: “where N is the number of observation points, x_i is the simulated value of either chlorophyll or POC at the ith observation point and y_i is its observed value; σ is the standard deviation of the specific observation set assimilated in a particular experiment. Using the standard deviation of the observations to define a characteristic scale of variation for each variable is a technique used in previous studies (e.g. Friedrichs et al., 2006; Xiao and Friedrichs, 2014). It is designed to weight the relative misfit contribution of each variable appropriately when there are insufficient data to define a comprehensive error model. Such a model would require reliable information about the uncertainty associated with observation errors (instrument error and error of representativeness) and non-parametric errors in the simulation such as forcing errors (Schartau et al., 2017). The use of different cost function weighting schemes in plankton modelling including the characteristic scale technique is explored in more detail by Hemmings and Challenor (2012)."

– section 2.4: Personally, it is not easy to digest this method. Maybe a diagram can help me and readers to understand the assimilation procedure better.

We appreciate the difficulty in understanding this section without a high level overview. Although we do not believe a full diagram is necessary, we propose two changes to this section to offer the reader a broader view of the method, rather than its current focus on technical details.

First, to clarify what is being done, rather than how, we propose changing the title of this section from "Implementation of micro-genetic algorithm and direction set algorithm" to "Cost function minimization."

Second, we propose adding a paragraph to the beginning of this section that summarizes the role of the two algorithms:
"Model parameters were optimized in MarMOT by finding the minimum of the cost function (Sect. 2.3) through a combination of the micro-genetic algorithm (μGA) and Powell's non-gradient direction set algorithm. The μGA runs first and identifies sets of parameter values that produce low cost values; this is achieved by "evolving" a population of various parameter sets over successive iterations, called generations. The low-cost parameter sets identified by the μGA are then used as starting points for the direction set method, which performs successive linear searches to identify nearby lower cost solutions."

– lines 244–245: Can you provide the number for the difference? If these two cases (50 m vs 200 m) are not significantly different, I would rather present the one with 200 m. Is it because of the computational time? (Also I hope the authors say something about the speed of this data assimilation calculation).

There is a relatively minor (~14%) difference between the results of the assimilation down to 50 m compared to 200 m. The trends and major conclusions of the study are likely not strongly affected by this choice. Conducting the assimilations for the upper 50 m avoided issues related to assimilating many low values of chlorophyll and POC, and also enabled a direct comparison of these results with the results of Kaufman et al. (2017) who similarly focused on the upper 50m concentrations. Computational time did not play a role in the decision to present results for the upper 50 m.

In further response to your question about computational cost, along with reviewer #1, we propose adding the number of model evaluations conducted for the assimilation experiments to the end of section 4.2, with the text: "The high number of model evaluations in each optimization case (roughly 4000 – 5000) makes such direct optimization impractical for large-scale models; however, the parameters identified in a 1D model by these techniques can be used in larger models, and indeed locally optimized parameters have been previously shown to improve the skill of 3D models in other regions [Oschlies and Schartau, 2005; Kane et al., 2011; McDonald et al., 2012; St-Laurent et al., 2017]."

– section 2.6.2: Are there any changes in spatial coverage between "glider", "cruise" and "satellite" data cases? If they have the same spatial coverages, naming this way may confuse readers because it is obvious that their spatial coverages are significantly different.

As mentioned (on line 418), these cases alter both spatial and temporal resolution, and therefore they don't have identical spatial coverage. As such, we feel these names are appropriate.

– lines 474–476: Do authors have any ideas why satellite-derived data underestimates carbon export?

This is addressed earlier in the manuscript on line 424: "The lower estimates of carbon export occurred because the optimal diatom fraction for fast-sinking detritus obtained via the assimilation of surface-only data (0.62 ± 0.14) was significantly lower than that obtained via the assimilation of data throughout the upper 50 m (Expt. 2a: 0.86 ± 0.05; Expt. 2b: 0.86 ± 0.11)."

– lines 480–483: I think the phrases after "and it is" are not necessary. Please consider to remove them.

Excellent idea. We agree and will take out the phrase starting with "and it is", and we will also remove the unnecessary "Ross Sea" on line 478.

Additional literature cited in responses:

Dinniman, M. S., Klinck, J. M., and Smith, W. O.: Cross-shelf exchange in a model of the Ross Sea circulation and biogeochemistry, *Deep-Sea Res. II*, 50(22–26), 3103–3120, doi:10.1016/j.dsr2.2003.07.011, 2003.

Gerringa, L. J. A., Laan, P., van Dijken, G. L., van Haren, H., De Baar, H. J. W., Arrigo, K. R., and Alderkamp, A.-C.: Sources of iron in the Ross Sea Polynya in early summer, *Mar. Chem.*, *177*, 447–459, doi:10.1016/j.marchem.2015.06.002, 2015.

Kane, A., Moulin, C., Thiria, S., Bopp, L., Berrada, M., Tagliabue, A., Crépon, M., Aumont, O., and Badran, F.: Improving the parameters of a global ocean biogeochemical model via variational assimilation of in situ data at five time series stations, *J. Geophys. Res. Ocean.*, *116*(6), 1–14, doi:10.1029/2009JC006005, 2011.

Keys, H. (J. R.), Jacobs, S. S., and Barnett, D.: The calving and drift of iceberg B-9 in the Ross Sea, Antarctica, *Antarct. Sci.*, *2*(3), 243–257, doi:10.1017/S0954102090000335, 1990.

McDonald, C. P., Bennington, V., Urban, N. R., and McKinley, G. A.: 1-D test-bed calibration of a 3-D Lake Superior biogeochemical model, *Ecol. Modell.*, *225*, 115–126, doi:10.1016/j.ecolmodel.2011.11.021, 2012.

Oschlies, A., and Schartau, M.: Basin-scale performance of a locally optimized marine ecosystem model, *J. Mar. Res.*, *63*(2), 335–358, doi:10.1357/0022240053693680, 2005.

St-Laurent, P., Friedrichs, M.A.M., Najjar, R.G., Martins, D.K., Herrmann, M., Miller, S.K., and Wilkin, J.: Impacts of atmospheric nitrogen deposition on surface waters of the western North Atlantic mitigated by multiple feedbacks. *J. Geophys. Res. Ocean.*, in press September 2017.