Review of Matsumoto et al.: MESMO 3: Flexible phytoplankton stoichiometry and refractory DOM

The manuscript of Matsumoto et al. describes the third version of the MESMO model, a model of intermediate complexity dedicated to studies on ocean biogeochemistry. Major model developments include three phytoplankton functional types, a flexible phytoplankton stoichiometry based on ambient nutrient, temperature and light conditions, as well as an implementation of a refractory DOM pool, in addition to the existing semi-labile pool, also including stoichiometry of C, N, P and Fe. The model is well and comprehensively described and the substantial developments after version 2 fully justify a new version description. The paper is well structured by starting with the developments since version 1, the newly implemented processes and a discussion of the model results. It will be a useful and complete reference for the future use of this model version, and fits perfectly within the scope of the journal. Therefore, I recommend publication after the comments below are addressed:

General comments:

- I think the overall improvement resulting from the changes in the new model version could be made more clear in a quantitative way. This is qualitatively described in paragraph l. 484ff, and on a ‘global-mean’ basis in Table 3. Adding more parameters that are difficult to constrain (such as e.g. the partitioning between DOMsl and DOMr, or uptake stoichiometry) should somehow “pay off” in terms of better agreement with observations or a more flexible representation of processes needed to model biogeochemistry in a changing ocean scenario. It is described that e.g. the nutrient pattern is better in MESMO3 than MESMO2, can you quantify this (with e.g. an RMSE, or a scatter plot/taylor diagram)? Or in Tab. 3, export C:N:P ratio seems to be less close to observations for the new version – what is the cause? Is there a better representation of the spatial pattern instead?
- In combination with the above comment: Could you describe more in detail what your calibration method was (in l. 152f you state that there was a calibration procedure, but you only state the absence of it)? Did you adjust parameters to fit a global mean like in Tab. 3 or a global mean profile of e.g. concentrations, or a specific dataset?
It would be helpful if some runtime information is given for the new model in comparison to the old one. EMICs are great tools for “cost” efficient calculations, so it would be helpful for future users of the model to be able to calculate runtime.

Specific comments:

- 224: “NPP is produced” – maybe revise, e.g. “all NPP is immediately routed to…”
- 227: do you mean “master nutrient variable”? 
- 246: If I understand correctly, your “deep POC split” pathway represents solution of DOM from POM (by e.g. exoenzymes or the like) – is there a reason why this happens only at depth and not at the surface? Or is the relative role of this process in producing DOM larger at the subsurface compared to the surface and therefore it is neglected at the surface?
- 320: Does photodegradation transfer DOMr to DOMsl in your model, or does it just represent a sink for DOMr?
- 463: Can you provide the units for this equation? I'm probably missing sth here, but why is density a factor in the gas-exchange equation (sticking to SI dimensions L=length, T=time, N=chemical amount/mol, M=mass unit):
  \[
  \text{flux} = k \cdot \text{density} \cdot \text{delta conc.} \cdot \text{ice fraction}
  \]
- 489: Based on the changes in the code, can you pinpoint the reason for the improved nutrient fields? A sentence here would be helpful (similar to line 493, where you state why the oxygen minimum depth improved – I find this very helpful).
- 594: Can you give reasons for why the DOM concentration at the surface is so high compared to observations? I was wondering whether this has to do with the model in Hansell, 2012 is formulated, that led to estimating the turnover times: There they scale the DOM production with the square root of NPP, this potentially may lead to lower surface concentrations and – by using their turnover time fitted to this model setup – cause the overestimated concentrations in MESMO3. (eq. 1 in Hansell, D. A., C. A. Carlson, and R. Schlitzer (2012), Net removal of major marine dissolved organic carbon fractions in the subsurface ocean, Global Biogeochem. Cycles, 26, GB1016, doi:10.1029/2011GB004069)

Figures

- Figure 1: The illustration of the sink processes of DOMr might be misleading: does the background decay happen all the time or only at the surface? If it is a true background decay happening everywhere, it might be helpful to write sth like “\(T_{\text{photo}} + T_{\text{background}}\)” instead of only “\(T_{\text{photo}}\).” Or does the background decay only happen at a specific depth level – then this would need to be justified why there is no first-order loss process at the surface?
- Figure 3: The caption beginning with “the effect of…” suggests that an effect is shown, i.e. the difference between a simulation with or without RNPG. Please either show an effect, or adjust the caption accordingly.
- Figure 4: Panel b would benefit from a latitudinal mean from observations for comparison, maybe from the dataset you already cite: Martiny, A. C., Pham, C. T., Primeau, F. W., Vrugt, J. A., Moore, J. K., Levin, S. A., & Lomas, M. W. (2013). Strong latitudinal patterns in the elemental ratios of marine plankton and organic matter. Nature Geoscience, 6(4), 279-283.
- Figure 7: Please define \(\text{DOCt}\) again in the caption