Review of "*The microbial community model MCoM 1.0: A scalable framework for modelling phototroph-heterotrophic interactions in diverse microbial communities*" by Lücken, Follows, Bragg and Lennart

General comments

This paper presents a very interesting model that is somewhat unique in addressing non-competitive interactions between autotrophic and heterotrophic plankton via metabolites. The model is apparently very scalable and adaptable, and has been applied in the manuscript to a number of case studies - some constrained by observations, others more theoretical.

I appreciated the highly configurable nature of the model, which facilitates its application to a range of hypotheses. On the other hand, I found the model description to be very heavy going, and the use of different subscripts confusing. I felt that the schematic shown in Figure 1 could have been expanded on to reveal how the model works. In general, I did not really understand the bigger picture as I was working through the equations, and I got a bit lost in the technical details at times.

I therefore wonder if this description of the model could be made a little bit friendlier to the reader by having an initial description that explains the main concepts of the model in a less technical way. What are the state variables in the model? How do they interact in a general sense? What is configurable in terms of the model structure and its parameterisation? How would it be set up for a particular case? Etc. If the reader has access to the bigger picture earlier, it may allow them to navigate the equations a bit more easily.

Aside from this issue, I found the case studies to be illustrative, but slightly lacking in terms of rigour. In particular, in Section 4.1 experimental differences in the dynamics of phytoplankton and heterotrophs are hypothesised to be related to the exchange of organic matter and competition for nutrient. MCoM is configured accordingly, but with only one simulation it is not clear what aspects of the model are vital to reproduce the observed dynamics and which are not. There is a similar lack of any control simulations in Section 4.2. I appreciate that the main aim here is to highlight the capabilities of the model, but there is certainly room to explore these limitations in a Discussion section (which is currently not included in the manuscript).

Finally, I am not sure whether the representation of organic compounds is reasonable. In equation 14 it seems that the uptake of each compound is described by a saturating function that is independent of the concentration of all other compounds. This does not seem intuitive to me, as I would expect that uptake of related organic compounds could interfere with each other. Perhaps this is not an issue, if heterotrophs each consume only a very narrow range of organic compounds, but I feel like this issue should be examined. In this regard, it might be useful to compare to how zooplankton grazing is assumed to saturate as a function of all available prey in some models (e.g. Vallina et al. 2014).

Specific comments

Line 21: "Arguably the two most fundamental roles in these ecosystems are primary producers … and their counterparts, heterotrophic organisms". I would argue that primary producers and heterotrophs are actors, not roles. Roles would be primary production and heterotrophy.

Line 68: Define t = time

Equation 1 and others. There is an included term for dispersion, but as far as I can see there is no mention of a spatial component in the model, aside from in equation 32, which relates to a hypothetical influx of nutrients that is not applied here. For the sake of simplicity, can the dispersion terms be removed for this version of the manuscript. It could be mentioned as a straightforward addition in a Discussion section.

Line 99: χ_i^N is mentioned with reference to equation 4, but it does not appear in that equation.

Lines 122-130: Is there any empirical support for these assumptions?

The max function in equation 9 feels a bit clunky. It is needed because some coefficients in $a_{i,j}$ are negative in order to model beneficial metabolite effects. If these beneficial effects appear elsewhere, I wonder if it would be logical to put them in a separate matrix, so that they do not enter into the mortality term (thus removing the need for the max function).

Line 230: Why are these called "pseudo-concentrations"? What are their units?

Section 3.1: The fact that large time-steps may yield inaccurate results seems general to all time-stepped models and integration schemes, so I wonder if this section could be abridged or skipped entirely for the sake of simplicity.

Figure 4: The black line is not mentioned in the legend.

Line 278: The word "polynomial" is redundant.

Figure 5: I was confused by titles in the top row, as I thought they described what was shown in the first row (which is *Synechococcus* in all cases). I would suggest "Axenic *Synechococcus*", "*Synechococcus + R. pomeroyi*" and "*Synechococcus + Tropicibacter sp.*".

Line 288: "*The decisive interaction was hypothesized to be an exchange of organic material, which provided the heterotrophs with energy and organically bound nutrient, and of remineralized nutrient, which could be assimilated by Synechococcus.*" I found this sentence quite hard to understand.

Sections 4.1 and 4.2: How were the model parameters chosen? Were the parameters formally tuned? How was model fit quantified? Can we be confident that the presented results are the best possible fit?

Line 310: : *Tropicibacter sp. … is assumed to be a strong competitor (smaller KN than Synechococcus)*" KN by itself is not sufficient to say which species is the stronger competitor for N (nutrient affinity would be). Also, "*a strong**er** competitor*"?

Line 320: Why "different possible outcomes" and not "different outcomes"?

Line 340: "are caused by the different **net** phytoplankton growth rates"?

Figure 7: This figure could be a lot clearer. In the schematic the different P, DOM and H pools should have unique subscripts (1, 2, 3) and the three time/density plots need to be labelled as it is not clear what each represents.

Line 359: "While such a setup may appear highly artificial, it is robust to parameter variations". It would be interesting to see what assumptions this behaviour is not robust to (i.e. what important mechanism breaks the pattern when removed?

Line 363: "we randomly generated communities". How? And under what constraints?

Line 410: "MCoM is intentionally kept relatively simple". The model does not come across as simple in the 11 pages and 34 equations used to describe it. I appreciate that in some ways it does have some simplicity, but this could maybe be made clearer in the way the model is described. (See also my general comments.)

References

Vallina, S. M., B. A. Ward, S. Dutkiewicz, and M. J. Follows. 2014. Maximal foraging with

active prey-switching: a new "kill the winner" functional response and its effect on

global species richness and biogeography. Prog. Oceanogr. **120**: 93–109.