Review of « CITRATE 1.0: Phytoplankton continuous trait-distribution model with one dimensional physical transport applied to the Northwest Pacific » by Bingzhang Chen and S. Lan Smith.

This manuscript presents a newly developed model based on an emergent trait-based approach to simulate phytoplankton traits (size) and associated diversity according to environment factors. The authors chose to construct an adaptative dynamics models that employs moment closure to allow continuous distribution of traits and limit the number of variables. Indeed, the model simulates the characteristics of phytoplankton community in terms of total biomass and mean size while the diversity is approximated by the variance in size. Some important processes for phytoplankton growth, corresponding to physiological adaptation to light and variable C:N ratio are also incorporated.

As such, the model described in this paper is of valuable contribution to the scientific understanding of the plankton diversity which has become a central issue of marine ecosystem management. The manuscript is well-written and gives a general overview on the ability of the model to simulate size-structured distribution of phytoplankton in various environmental conditions using two contrasted stations of the Northwest Pacific. However, I have some questions concerning the methodology which has been applied and whether/how this work can be generalized to other stations or a broader oceanic region (e.g. in the context of 3D modeling setup). Indeed, I am not familiar with the use of DRAM-type algorithm to adjust model's parameters value to observational data and it took some time to me to understand exactly the method that is implemented in this study. Therefore, a more detailed description of what is exactly done by the parameters optimization algorithm and how this will be used to apply the model to other regions would be very useful to help the reader to understand the concept behind this model. It would thus significantly raise the likely impact of this paper. I recommend these questions, detailed in the 'general comments' section below, to be addressed before the manuscript could be considered to be published in 'Geoscientific Model Development'.

Some minor and more specific comments are also included at the end of this document.

General comments:

1- Technically, the proposed model setup uses a DRAM algorithm to adjust the targeted parameters values and minimize the differences between model outputs and observational data based on two available dataset for contrasted stations of the Northwest Pacific.
In the discussion section (p. 21, l. 3-7), the authors argue that this model would be 'easy to couple with 3D global or regional ocean models' (see also page 5, lines 6-8).
As far as I understand, the idea would be to use the single set of parameters which has been found in this study (i.e. the one which gives the highest likelihood with regards to observational data for both stations) to run the model in other oceanic regions (otherwise, I do not see how the method can be applied to large oceanic system while seeking for a single set of parameters that would lead to the best fit to observational data over the considered region).
This point is not specify in the manuscript. Could you please add further thoughts on that in your discussion section and describe the preconised method to apply this model to larger oceanic regions?

2- In the introduction section (p. 4, l. 16-17), you write that 'relatively few continuous trait-based models have been validated against oceanic observations'. The comparison that is done in results section 3.3 (p. 13-14) does not constitute a 'validation analysis' as you are using the same observational data to contrain the model's parameters and to 'validate' the

results. Indeed, the specific aim of the method which is used in this study is to provide the best fit between model outputs and measured data. Therefore, the main outcome of this study is actually the parameters set you obtained after running an ensemble of 10 000 simulations.

A validation analysis should involve totally independent data for model parameterization and for validation and could only be carried out if you have runned the model for a different region using the same set of parameters. Here again, a more explicit description of the aim of your method (i.e. seting up a set of parameters which can be subsequently used to run the model in other regions ?) would have been useful to avoid the confusion.

3- How do you convert the mean size and size variance into four size classes fraction? I guess the calculation is done by comparing the occurrence of each size classes from the size distribution (Gaussian distribution of the log biovolume) other time (e.g. seasonal average in fig. 8 and 9?) but this is unclear. Could you please specify the method that has been used in your method section?

4- In the introduction section, you say that functional groups (PFT) models, representing a defined and limited number of plankton types, generally 'underestimate local diversity'. You argue p.3, l. 24-26 that the main reason for that is their inability to resolve the trait space combined with their failure of representing the appropriate mechanism sustaining high level of diversity. Although these considerations are correct, I think they are not specific to aggregated models but can also apply to the model presented here. Indeed, as you point out in the discussion section, you choose to consider the size as master trait but ignore some other major traits (temperature optima, mixotrophy, grazing resistance etc.) which may also vary among planktonic organisms of same size and enable coexistence by achieving similar fitness between different adaptation strategies (i.e. mechanisms for sustaining diversity). In that sense, I would say that the two techniques have a similar bias of taking into account a limited number of traits and mechanisms to explain the huge plankton diversity. Please modify the introduction to consider this point.

Moreover, another difference between the two methods is that the measure of diversity that is provided by moment-closure models corresponds to a relative measure of diversity (variance in size in this case) which only allow a relative and comparative analysis of the phytoplankton diversity (in time and space) but does not provide any absolute measure of diversity (number of taxons or species) to compare with observational data.

p. 4, l. 9, you argue that 'the factors controlling diversity can be directly quantified and better understood' with the continuous trait-based models. This sentence is not unclear. Could you specify how and why are the factors (which factors?) controlling diversity better characterized using the latter method ?

Please put 'et al' in italic while citing referenced publications throughout the manuscript.

Model description

P5 L12: Add the unit of P
P5 L18: Figure 1: What is the inset in the box on the top left (with probability axis)? Please add a description in the figure caption.
P5 L22: Please provide more explanations on the role of the trait diffusion parameter
P6 L7-11: Please provide the references for the growth dependences to light, nutrient concentration and temperature.
P6 L17: Eq. (5A) should be Eq. (4A)
Section 2: I would suggest to separate the description of the biogeochemical equations (section 2.2) and the 1D implementation (section 2.3). Therefore the paragraph l. 9-13 on page 7 should be moved to the section 2.3 and the name of the section 2.2 should mention only 'Biogeochemical model (nutrient, zooplankton etc.)'
P7 L13: The sentence 'the 1D model contains only biological tracers' is unclear. It should be replaced by 'the biological model is runned offline' or something similar
P8 L5: Please replace 'water depth' by 'the depth of the water column depth'.
P8 L11: Please verify the equation for detritus (- -).
P9 L12-14: Do you assume that the surface mixed layer has a depth of 100 m? (the explanation for the use of the threshold of $Kv > 10^{-3}$ $m^2.s^{-1}$ is unclear). Why do you use a different parameterization for the MLD calculation for phytoplankton growth and MLD showed on fig. 2 from observational data (page 12, line 15)?
P11 L8: 'and both model'? Please check the sentence.
Fig. 2 caption: Add the description of the white scatter plots (MLD) in the legend. Change '…at station S1. (f-i) The same for station K2'
Fig. 2: Check x-axis thick labels (subplots b, c, f and g).
Fig. 3: Add x-axis labels.

Results

In general, there is some discussion points that are included in the results section and that should rather be discussed in section 4.

P12 L18-25: This section describes the physical forcing and does not concern a result of the simulation. I would suggest to move this part in section 2.3 (method).
P12 L23: 'with the model estimates of MLD consistent with those measured from *in situ* temperature and salinity profiles': it is not clear what you are comparing exactly. (Please also add a reference to the figure showing that. What are white scatter plots on fig. 2 b and f?).
Fig. 4 caption: Remove the 's' in 'log-likelihood'. Replace (b-j) by (b-i)
P13 L6: The SSqE of the smallest size fraction (fig. 4, q) also increases with time at S1
P13 L9: The figure 5 is not commented in the text**.** Please add a sentence to describe the trend**.**
P13 L 11-16: The discussion on the value of the trait diffusion parameter should appear in the discussion section.
Fig. 6 and 8 captions: Complete the caption with the position of the different variables.
P13 L20: 'the higher surface concentrations'
P13 L22-24: Isn't it in apparent contradiction with the fact that you argued that the modeled MLD is in agreement with observational data (page 12, line 23)?
P13 L22-24: discussion
P14 L5: The observational data on the size fractions are relatively noisy. Could you please provide

more details on how these data were obtained (sampling methods, size measurments, sampling frequency) in section 2.5? At station K2, the size distribution in unclear in data and the model overestimate the proportion of 3-10 µm size class.

At station S1, observational data show the dominance of smaller cells but do not show the vertical structure of the size distribution that is simulated by the model with smaller cells at the surface. Please add a comment on this.

P14 L7-9: 'At station S1'. Do you mean station 'K2'?

P14 L10: discussion

P14 L19: 'following stratification which occurs earlier in S1 than K2'

P14 L24: Show a figure of Chl/C ratio

Fig. 10, g: High growth rate at the surface at K2 despite low TIN and low Chl a concentrations?

P15 L5-6: discussion

P15 L24 – P16 L4: discussion

P16 L3: The 'dynamic equilibrium theory proposed by Huston is only briefly mentionned (see also page 17, l, 1-2). This hypothesis implies that, under non-equilibrium conditions, the outcomes of the competition depend on the timescale of the competitive displacement and the relative rate of change in competitive abilities of each competing species. This point should be further developed and discussed according to your results on diversity in the discussion part of the manuscript.

P16 L6-9: As I said above, the role of the trait diffusion in maintaining diversity and the way it is used in the model is a bit tricky to understand. You could perhaps add a paragraph in the method section to clarify this point which is then only discussed briefly page 17, lines 10-14.

Discussion

P17 L22-27: There is no figures showing the N:C and Chl: C ratio patterns at the two stations.

P18 L3: 'Instead, we employ …': The word 'instead' seems unappropriate as you mention a very different process than in the previous sentence: the trade-offs between maximal growth rate and nutrient affinity in the phytoplankton is not related to the size-dependence of the grazing by predators.

As you mention l. 17-19, the role of grazing in shaping phytoplankton community has been shown to be crucial. In order to take into account a various palatability of phytoplankton for zooplankton feeding according to the cell size, the model should ideally involved a larger number of predator size classes (and/or, at least, an additional mesozooplankton size class) which would lead to much more complexity in the model.

In addition to the predator-prey size-ratio, the predators' feeding mode (Mariani *et al.,* 2013) and the formulation that is used to constrain the herbivorous impact on primary producers community composition are also very important. Please add the information on what kind of predation function you are using in your model in 'Model description'. Also, this points should be mentionned in the discussion section (add other references such as Anderson *et al.,* 2010 ; Prowe *et al.,* 2012).

P20 L12: 'other mecanisms such as vertical migration': I am not sure that vertical migration is a very common process in small phytoplankton populations that are found at the surface of subtropical waters during the summer. What about the nutrient limitation terms? What should be the half-saturation constants for nitrogen/phosphorus uptake in the 1-3 µm size class found in observationnal data?