Biogeosciences
Discussions

Open Access
EGU

Interactive comment on "Reviews and syntheses: Parameter identification in marine planktonic ecosystem modelling" *by* Markus Schartau et al.

Anonymous Referee #3

Received and published: 5 August 2016

The authors present a summary review of parameter identification for biogeochemical ocean models. This is an important topic, and the manuscript provides a lot of material and some good description of the current state of research and outstanding issues. Yet, while some techniques and issues are discussed in great detail, very little information is provided about others. A less selective, more balanced view is needed to make the manuscript a proper synthesis of parameter identification in marine planktonic ecosystem modelling. My reservations are detailed below.

general comments:

The focus of the manuscript appears to be quite selective and sometimes arbitrary. Some methodologies are described in great detail, while only brief descriptions are given for others. DA techniques that are commonly used are not mentioned at all, for

example Kalman filter-based techniques. If the authors want to provide a synthesis of the current state of the research, these techniques need to be mentioned.

The authors seem to be very focused on mass conservation (is that the reason for not including many ensemble-based techniques?). I do not agree that this is the "one straw that biogeochemical modelers grasp at". In regional models, river inputs routinely break mass conservation, so why should DA techniques not be allowed to create updates to the mass inside the model domain if the data provides evidence for this? At the very least, the authors need to acknowledge that their view on mass conservation is not shared in the entire modelling community.

Given that model complexity and parameter identifiability play an important role in the manuscript (and rightly so), I wonder why there is not more focus on alternatives to the functional group approach. Some approaches, like the "optimal trade-off" are mentioned in section 10.1. Yet there are others which do not require manual parameter selection, like self-selective models (Follows et al., 2007) or the gene-centric approach (Reed et al., 2014), which groups plankton groups based on genetic information. These alternatives to the functional group approach should at least be mentioned, and mentioned earlier than section 10.1.

More focus should be given to the role different data types play in identifying model parameters. Sometimes the manuscript seems to suggest that all is needed is more data in order to identify more parameters, for example in the abstract: "data are often too sparse to constrain all model parameters". Yet more satellite chlorophyll data is probably not helpful in identifying many parameters, other data types and subsurface data are important as well. This is not just true for large-scale models (the issue of different data types is finally discussed in section 9.1 but only in regard to global models). In this context, the authors may also want to discuss the Bio-Argo program which could provide some much needed biogeochemical data products in the near future.

specific comments:

p2 l24: "of fecal pellets" perhaps change to "attached to fecal pellets"

p3 l4: "trophic levels like fish, which would be subject to changes in biomass on multi-annual rather than seasonal time scales": I would argue that the greater challenge with modelling fish is their behaviour and ability to swim, making it impossible to realistically simulate them by tracer variables.

p3 l5: "Every marine planktonic ecosystem model can thus be described as a simplification of the dynamics inherent to a system of nutrients, phytoplankton, zooplankton, detritus, dissolved organic matter, and bacteria". Apart from the fact that some phytoplankton are bacteria, I am wondering why they are listed here if (as stated above) they are often not resolved in models.

p3 l8: "Feedbacks from the ecosystem model ..." Maybe mention that feedback from physical to ecosystem model are essential.

p3 l16: "and the simulated N cycle was shown to already depend on the value assigned to a single parameter, namely the sinking velocity of detritus.": this sentence is not very clear, all other parameters do not affect the nitrogen cycle at all?

p4 l11: "availability of data thus places limitations on the number of model parameters whose values become identifiable." It is not just a numbers game, certain types of parameters may never be fully constrained by certain types of data, even if the model contains just a few parameters.

p4: l18: "Novel DA methods are predominantly devised for improving forecasts ..." While forecasting skill is often used to judge the quality of an assimilation system, many systems are used for hindcasting and creating reanalyses.

p10 l24: "It means that actually the cost function as given by..." this sentence is not very clear

p23 l5: "The third approach leads to more complex representations of growth limitation, as they..." Something may be missing here, the third approach is not described well and

C3

the "they" should be an "it".

p25 l3: the summary of the loss terms here "Cell lysis, excretion and leakage are usually espressed ..." does not agree very well with the summary previously (p2 l22) "... removed by natural mortality (cell lysis due to starvation, senescence, and viral attack)..."

p27 l8: "right and left sides of top and bottom row" it would be useful to have labels (a) - (f) in Figure 2.

p27 l8: "It means that g_m can only be estimated in combination with". It means that estimates of g_m are dependent on the values of the other parameters. If we are certain what the values of phi_agg and gamma_C are, we can still estimate g_m.

p27 l9: "If g_m remains fixed, we do not find such strong collinearity expressed between gamma_C and phi_agg": I would rephrase this, since in this particular experiment only two parameters are varied, i.e. remove the "If g_m remains fixed".

p40 l29: "To account for the lateral flux information was helpful contributed strongly to the emulator accuracy.": something is missing here

Section 9: After discussing methodology, why are global biogeochemical ocean models introduced now? I would move most of this section to the introduction.

technical corrections:

Both "Fig." and "Figure" are used to reference figures.

p15 l17: close parentheses.

p18 l23: "(Fisher, 1922) see also (e.g., Fisher, 1934 ..." move "see also" into parentheses as well.

p36 l30: Citations are now ordered by name, before it was by date.

C4