Biogeosciences Discuss.,
doi:10.5194/bg-2016-242-AC3, 2016

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

Markus Schartau et al.

mschartau@geomar.de

Received and published: 31 October 2016

We thank Referee #3 for reviewing our manuscript. According to the comments provided by Referee #3 we have learned that we should mention filter techniques and also describe aspects of sequential DA approaches. We have done so and, in the end, we introduced changes that correspond to a major revision of our manuscript. The comments provided by Referee 3 are appreciated and we think that the revised manuscript has improved considerably.

Major comments by Referee #3

***Comment* 1**: 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.

Author's response:
During the preparation of our manuscript we concentrated on achieving a balance between biological aspects, problems of parameter identification, and basic DA methodological considerations. It was a difficult and extensive process to elaborate a meaningful structure and work out the content of respective sections of the submitted version of our manuscript. All authors' contributions were mutually revised. We regret that Referee #3 has the impression that the focus of the manuscript is arbitrary and selective. We do not share this view, but we have realised some structural weaknesses and have introduced major changes to the manuscript. Apart from shifting the original Sect. (5) to Sect. (3) (following a suggestion of Referee #4), major changes were done in Sect. (2) and in Sect. (9).

We agree that sequential DA methods, like the ensemble Kalman filter, were under-represented. We now restructured and extended the Theoretical Background section (Sect. 2) and included two subsections under Sect. (2.2 Estimation methods): one about sequential methods (Sect. 2.2.2) and another about variational methods (Sect. 2.2.3). This way we have introduced an explicit representation of sequential DA methods.

We also included two new figures from the publications of Simon and Bertino (2012) (based on a sequential method) and of Losa et al. (2006) (based on results of a weak constraint variational approach), which should further improve the balance with respect to DA methods. Here is a summary of the figures in the revised manuscript (as also given in the response to Referee #2):

1) We have moved the original Fig. (1) with the example of a variable lag fit (VLF) to the Appendix (Fig. A1) and substituted it with a figure from the study of Simon and Bertino (2012, Journal of Marine Systems, 89, 1-18; top two rows of their Figure 3). This figure is a nice example of the improved asymptotic behaviour of a deterministic Ensemble Kalman Filter (DEnKF) when using log-transformed observations and model results to realise the analysis step. Ehouarn Simon and Laurent Bertino kindly provided their results so that we could redraw the figure. The figure is referred to in Sect. (4) about Error Models. The new Fig. (1) is based on results from a sequential, ensemble based, data assimilation (DA) approach, which should further improve balance with respect to ensemble based DA methods.

2) In Sect. (7) we address space-time variations in model parameter estimates and we find it appropriate to include a plot from Losa et al. (2006; top row of their Figure 2), based on results from Losa et al. (2004). It is prominent and illustrative example of variable parameter values in the North Atlantic.

For both new figures and for Figs. (4 and 6) we still have to request permission from Elsevier.

***Comment* 2**: 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.

Printer-friendly version

Discussion paper

Author's response:
We did not state that DA techniques are not allowed to introduce sinks and sources of mass. But we wanted to stress that it is important to explicitly clarify whether mass is conserved or not. Scientists not directly involved in DA applications should recognise this. We confess that our formulation is a bit clumsy and we rephrased it: "So far no fundamental ecophysiological principle has been further exacted beyond the conservation of mass. Whether a balanced mass budget needs to be achieved depends on the scientific problem addressed. Some ecosystem model applications may not critically depend on mass conservation, e.g. when simulating plankton growth to act as food source in regional simulations of fish stock size and recruitment. In biogeochemical models the conservation of mass can be essential, in particular for large-scale or global ocean simulations. A consistent theme running through most ecosystem models is the determination of mass flux of certain biologically important elements, such as nitrogen, phosphorus, iron and carbon (N, P, Fe and C). "

***Comment* 3**: 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.

Author's response:
We appreciate this helpful comment. We confess that we should have considered this aspect. However, we do not think that parameter selection is an issue only for the

functional group approach. The models of Follows et al. (2007) and Reed et al. (2014) also depend on parameters to which fixed values have to be assigned. For example, in Follows et al. (2007), fixed values were assigned to the distributional parameters describing the PAR saturation and inhibition constants (mean and standard deviation) and nutrient half saturation constants (upper and lower limit of uniform distribution) for small and large phytoplankton size classes (Follows et al., Table S1, column “Range”), and fixed values were assigned to various non-stochastic parameters (column “Fixed”). The model proposed by Reed et al. (2014) includes more than 40 parameters and they identified 16 parameters that were of importance for determining the biogeochemical dynamics in their example model setup. However, both studies are important contributions and they provide novel approaches. We extensively discussed whether we can refer to the study of Follows et al. (2007) either in Sect. (6, Model performance as a function of model complexity) or at the end of Sect. (7.1.4 Learning from space and time variation in parameter estimates). In the end we found any discussion on the Darwin model approach inappropriate in Sect. (6) or Sect. (7). We concluded that it does fit to Sect. (10.1 Modelling prospects). We therefore suggest a revision of this paragraph (Sect. 10.1) according to: “ A commonality of new model formulations is to focus on principles, e.g. by considering the adaptation of traits towards optimal trade-offs (e.g., Wirtz and Pahlow, 2010; Dutkiewicz et al., 2009; Smith et al., 2015), or by accounting for allometric relationships in growth and plankton interaction (e.g., Banas, 2011; Acevedo-Trejos et al., 2015), or by using microbial traits in a functional gene approach (Reed et al., 2014). Recent studies have begun to simulate ecosystem complexity and allow the model to “self-organise” according to a relatively simple set of ecological and physiological rules or “trade-offs” (Bruggeman and Kooijman, 2007; Follows et al., 2007). A major advantage of this approach is that the models are able to resolve greater ecological diversity with fewer specified parameters whose values can be assumed to be spatially invariant. This diversity allows the simulated plankton community to reorganise across broad environmental (e.g. spatial) gradients. But the identification of the most important trade-offs governing competition between

Printer-friendly version

Discussion paper

organisms remains a major challenge (Tilman, 1990; Litchman et al., 2007, 2012)."

***Comment* 4**: 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.

Author's response:
Yes, we certainly agree. We want to introduce Sect. (10.2 Examples of recent advancements in data availability). In this short Section we want to briefly refer to different types of data and data products that could be potentially used for DA/ parameter estimation. In this paragraph we mention the Bio-Argo program and added Mignot et al., (2014) and Sauzéde et al., (2015) as references. We also mention new remote sensing products (e.g., net community respiration, Tilstone et al., 2015), new flux estimates based on time series observations (e.g. Emerson, 2014) and latest CO_2 data products by Roedenbeck et al., (2015) and Bakker et al., (2016).

—

Responses to specific comments by Referee #3

***Specific comment* 1**: p2 l24: "of fecal pellets" perhaps change to "attached to fecal pellets"

Author's response:
Aggregated cells can sink without being attached to sinking fecal pellets. Yes, both, cells and fecal pellets, are often incorporated in aggregates that sink. The proposed addition "attached to" would require further explanations and we therefore prefer to leave the sentence as it is.

***Specific comment* 2**: 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.

Author's response:
We changed it to: "These closure assumptions ensure mass conservation while neglecting the actual mass loss to higher trophic levels like fish, which would be subject to fish movements and changes in biomass on multi- annual rather than seasonal time scales. "

***Specific comment* 3**: 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.

Author's response:

We actually meant heterotrophic bacteria and changed the text accordingly.

***Specific comment* 4**: p3 l8: "Feedbacks from the ecosystem model ..." Maybe mention that feedback from physical to ecosystem model are essential.

Author's response:
We think that the original text is fine. It is obvious from the context of ecosystem models being embedded in physical models that feedback from physics to ecosystem is essential. However, we slightly modified the text: "... but are hardly considered in current marine biogeochemical studies. With such resolved, changes in ecosystem components may induce changes in physical environmental conditions, but so far the physical model remains unaffected by ecosystem states in most studies."

***Specific comment* 5**: 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.

Author's response:
We understand the Referee's concern and refined the statement accordingly: "The availability (type and number) of data thus places limitations on the number of model parameters whose values become identifiable, and values of some parameters may never be fully constrained."

***Specific comment* 6**: 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.

Author's response:
We revised the text: "Much of the literature on DA in oceanography is focussed on state estimation (e.g., Allen et al., 2003; Natvik and Evensen, 2003; Dowd, 2007; Nerger and Gregg, 2008; van Leeuwen, 2010). In these studies, the primary objective is to improve hindcasts, nowcasts, or forecasts of time-dependent variables such as chlorophyll *a* (Chl*a*)."

***Specific comment* 7**: p10 l24: "It means that actually the cost function as given by..." this sentence is not very clear.

Author's response:
We corrected this formulation and refined the text (see response to Comment 1 by Referee #1): "The MAP or posterior mode estimator of $\log(\Theta)$ is equivalent here to the posterior median estimate and is obtained by maximising $p(\log(\Theta) \mid \vec{y})$. This leads to a cost function given by Eq. (15) without the second term, $2\sum_{l=1}^{N_\Theta}\log(\Theta_l)$ (cf. , Fletcher, 2010)."

***Specific comment* 8**: 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 the "they" should be an "it".

Author's response:
For clarity we suggest to slightly rephrase it and add one additional sentence: "The third approach involves more complex representations of growth limitation, as it accounts for interrelations between cell quota, N-uptake and the photoacclimation state of the algae (e.g., Geider et al., 1998; Pahlow, 2005; Armstrong, 2006; Wirtz and

Pahlow, 2010). Here, photoautotrophic growth depends on the cellular C:N (or N:C) ratio and the mass distribution of phytoplankton C and N has to be explicitly resolved in the model. Whether the first, second or third approach is considered can be expected to affect estimates of the associated parameter values.”

Some more information is given in the subsections that follow, e.g. in Section (3.5 Algal growth and intracellular acclimation):
“ More complex interdependencies between light and nutrient limitation are resolved by models that account for intracellular acclimation dynamics (e.g., Geider et al., 1998; Pahlow, 2005; Armstrong, 2008; Wirtz and Pahlow, 2010). In these models growth rates become dependent on cell quota, e.g. usually normalised to carbon biomass (N:C), and the amount of synthesized Chl*a* per cell. These approaches involve physiological parameters that are related but not identical to those of classical N- or P-based growth models, which impedes a direct comparison of older estimates of growth parameters with values currently used in models with acclimation processes resolved.”

***Specific comment* 9**: 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)..."

Author’s response:
This formulation is imprecise and we thank Referee #3 for the notice. Cell lysis is associated with cell death whereas exudation and leakage induce a loss of organic mass while the cell is alive and its physiology is fully functional. We rephrased the few sentences for clarification: “Parameterisations of phytoplankton cell losses involve lysis (starvation and/or viral infection), the aggregation of cells together with all other

suspended matter, and grazing by zooplankton. Exudation and leakage are processes of organic matter loss that occur while the physiology of the algae is functional. Cell lysis, exudation and leakage are usually expressed as a single rate parameter and this mass loss of organic matter is assumed to be proportional to the phytoplankton biomass."

***Specific comment* 10**: p27 l8: "right and left sides of top and bottom row" it would be useful to have labels (a) - (f) in Figure 2.

Author's response:
Additional labels (1a/b, 2a/b, 3a/b, and 4a/b) were added to the figure.

***Specific comment* 11**: 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 ϕ_{agg} and γ_C are, we can still estimate g_m.

Author's response:
We added the following sentence for clarification: "Only if Φ_{agg} and γ_C were known, then g_m could be identified in this mesocosm model setup with these available data types."

***Specific comment* 12**: p27 l9: "If g_m remains fixed, we do not find such strong collinearity expressed between γ_C and ϕ_{agg}": I would rephrase this, since in this particular experi- ment only two parameters are varied, i.e. remove the "If g_m remains fixed".

Author’s response:
We have corrected (removed) it as proposed.

***Specific comment* 13**: p40 l29: "To account for the lateral flux information was helpful contributed strongly to the emulator accuracy.": something is missing here.

Author’s response:
We corrected the sentence: “Accounting for the lateral flux information was helpful, contributing strongly to the emulator accuracy.”

***Specific comment* 14**: Section 9: After discussing methodology, why are global biogeochemical ocean models introduced now? I would move most of this section to the introduction.

Author’s response:
The introduction is already extensive and we do not want to further extend it. Although not explicitly highlighted, the underlying structure of the manuscript gradually extends from simulations of algal physiology of simulations of laboratory experiments to global biogeochemical modelling. We have started to revise (and condense) Sect. (9) in order to make it more concise. Some of the methodological aspects addressed before are needed in the text of Sect. (9). It would not be meaningful to consider aspects of parameter identification in global BGC models if they had not been explained before.

We thank Referee #3 for the support. All technical corrections are included in the revised version of our manuscript.

Interactive comment on Biogeosciences Discuss., doi:10.5194/bg-2016-242, 2016.

Printer-friendly version

Discussion paper