Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2018-6-RC2, 2018

Interactive comment on "Inferring particle dynamics in the Mediterranean Sea from in-situ pump POC and chloropigment data using Bayesian statistics" *by* Weilei Wang et al.

Anonymous Referee #2

Received and published: 7 May 2018

Review of Wang et al: This paper uses a Bayesian approach to determine rate parameters for a simple particle model using data collected during the MEDFLUX program at the DYFAMED site in the Mediterranean Sea. I think that the authors have a potentially useful approach here, but the presentation of the manuscript and their ideas requires a lot more work.

P1. Line 18: The authors state that "Particle density and size determine particle sinking speed". This is a very contentious statement. This problem has a very long history, and to date, no relationship between size and settling velocity has even been shown. Density does seem to be a determining factor, but size, not so much.

P1. Lines 21–22: Lots of work was done historically on this problem using isotopes, starting with the work of Bacon and Anderson, and Clegg and Whitfield which the authors cite later, but should also be cited here.

P1. Line 23: The authors state "…much of what is currently known about these processes is from work with particulate thorium." This is a gross simplification and does not represent the myriad techniques that have been and are used to help us understand particle process in the ocean. I would argue that use of thorium tracers is one tool that has been used, but a considerable amount of what we know about aggregation and disaggregation has come from other techniques including optical methods, staining methods, laboratory experimental methods (e.g. rolling tanks), and modeling. Thorium tells us very little about the biological processes that enhance aggregation and disaggregation, or determine the strength of particles. So whilst use of thorium isotopes is a critically important, it is only one of many techniques. One could equally make the case for optical techniques being the main source of information.

P4. Lines 17–18: The authors make two very fundamental, connected assumptions: that sinking speed increases linearly with depth and that the flux attenuation profile is a power law. However, as the authors state on page 8, other observations at the MEDFLUX site show that sinking speeds do not increase with depth. The authors explanation of why they make this assumption does not seem to make sense to me. If you are applying an inverse model to data at a given site, you shouldn't make an assumption that clearly does not hold at that site, unless you have evidence that previous estimates were in error.

Equations 1–6: The authors assume first-order reaction kinetics for aggregation and disaggregation. This is known to be incorrect âĂŤ aggregation is a fundamentally nonlinear process and to assume that it is a linear process depending only on the particle concentration is unphysical. Disaggregation is also not a linear process but depends on environmental process such as turbulence, or factors such as animal abundance. So, the model used by the authors is inherently unrealistic and unphysical from the start âĂŤ to see this, just analytically solve the linear odes with only the aggregation and disaggregation terms and you'll get completely unphysical solutions.

P5: Lines 6–7. The authors assume that chlorophyll is found only in the small particles, and that any chlorophyll found in larger particles comes from aggregation of small particles. If I'm reading the paper correctly, the authors use a size of 70 μm to separate large and small particles. So in this model, there is no photosynthesis in particles greater than 70 μm? This rules out most diatoms and other large phytoplankton. This surely cannot be correct.

P5: Line 20: I must be missing something here. The authors state that the transpose of the vector c is has 48 components, but only 6 are listed.

P6: Lines 5 to end of section: This is very unclear. Why use a Bayesian approach? What do we gain from this? Why won't a more standard approach also work. I'm not averse to using Bayesian approaches, and they are often more informative and successful than standard frequentist approaches. But it's unclear to me why they should be used here. What is more, the explanation of the technique given here is unclear âĂŤ why are two optimizations needed? Why do we need to scale the data and prior precisions? Won't using the logarithms of the parameters bias the end result because you've inherently changed the statistical distribution of parameter uncertainties? (this is similar to the problems incurred by fitting a straight line to log-transformed data that obey a power law or exponential distribution). Also, the authors assume that errors are independent (in order to make their likelihood matrix diagonal). What is the justification for this? Given the data being used, I would have thought that the uncertainties were highly correlated. This whole section needs to be thought out more carefully, and be re-written to be more explanatory.

Table 1: There are no uncertainties in this table. Even if the observation uncertainties are estimated by the limitations/sensitivity of the instruments/methodologies, they should be given! What is more, taking these uncertainties into account will affect the

uncertainties in the parameter estimates given in Table 2.

Figures 2 and 3: Perhaps I missed this, but there seems to be no significant discussion of the fact that their model consistently under-predicts the observations. The quoted R-squared value is obviously being driven by the two clusters of data. This needs to be examined and discussed in detail.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2018-6, 2018.