Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-1081-RC3, 2018

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Fundamentals of Data Assimilation applied to biogeochemistry" *by* Peter J. Rayner et al.

Anonymous Referee #3

Received and published: 22 December 2018

The paper intends to review the fundamentals of data assimilation used in biogeochemistry from both theoretical and practical perspectives, and unify the notations at the same time. While such review paper is timely, especially as the atmospheric inversion and data assimilation become more broadly used in biogeochemistry, I find the paper hard to follow and it is not clear whether the paper is targeting general data assimilation researchers or people who work on biogeochemistry data assimilation. If it were the latter, I think it would be more useful to discus how data assimilation has been used in biogeochemistry, what science questions the biogeochemistry community try to address with data assimilation, and the practical differences among different data assimilation methods, challenges, and the future directions. Some part of the paper is rather too general, such as section 2.2. I would suggest a more focused paper. The

following is my specific comments:

1) Section 2.2: Since the paper focuses on Bayesian approach, I don't think the paper needs to discuss non-Bayesian method. 2) Section 3: It is not clear to me what is the difference between "x" (Target variables for assimilation) and "z" (model state variable). It is not explained in the paper. Later on, only "x" is used in the cost function. 3) Section 5: this section has a lot of useful materials. I would suggest discussing how each element is addressed in specific examples. Section 5.1 gave a recipe to decide "target variables". It would be easier to understand if this recipe is discussed within a specific application. 4) Section 6.5: The Kalman filter is discussed in a very general concept here. How Kalman filter has been used in biogeochemistry, and what is the challenge? 5) Section 7 discussed specific Bayesian methods from computation perspective. Again, I found the description is too general. The applications of these methods in specific problem could be very different. For example, in boundary condition estimation, the prior ensembles may not come from the posterior ensemble of previous step, since there is no dynamical model to propagate information forward. As a result, step 5 described in section 7.4 is not applicable.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1081, 2018.