

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2020-444-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2020-444

Anonymous Referee #1

Referee comment on "A new exponentially decaying error correlation model for assimilating OCO-2 column-average CO₂ data using a length scale computed from airborne lidar measurements" by David F. Baker et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2020-444-RC1>, 2021

The authors study the correlations between errors in XCO₂ satellite retrievals, based on reference lidar measurements, and discuss various ways to account for them in atmospheric inversions. The paper looks a bit like the clean minutes of a brainstorming meeting: every sentence is well written but the logical flow is curvy and difficult to follow. The authors have not done enough to make their thoughts accessible and to take the text beyond elaborate speculation, perhaps simply because their thinking is not yet ripe for publication. Maybe it doesn't matter, as the paper will be cited anyway given the role of this activity for the OCO-2 team, but for the few who will bother to read it, it may be a daunting task, perhaps in the end wasted. I am listing a number of comments here to help clarify the presentation.

- Footnote 1, p. 25: the disclaimer here is a bit hidden, but it is actually essential. Basically, if the "good reasons" listed here are correct, all results of the paper can be ignored. This observation could be fatal for the patient reader who painfully reaches this page... In the end, nothing is given to convince the reader that the MFL-OCO-2 differences do indeed represent OCO-2 errors, that the two scales of correlation lengths found (10 and 20 km) should be used at all in OCO-2 error models. It's embarrassing.
- There is good and interesting math here, but the authors belittle it by arbitrarily rejecting certain math results: why should the negative weights not be physical (l. 20) or considered undesirable (l.331)? They simply follow from the authors' correlation model: if the authors are not satisfied with this consequence, they should change the model rather than fooling the math.
- l.35-37: the sentence seems general, but does not apply to GOSAT in practice.
- l.43: why would it make little sense to assimilate the measurements individually? From the text, it is obvious that it is so much easier than trying averages. So, this can make a lot of sense. Personally, I would still prefer the averaging but for reasons that are not discussed here (numerical stability, but this is only a feeling).
- l.48: interesting comment... The dependence of the correlations on the scene questions the representativeness of the ACT measurements used here. This key element is only briefly touched on in the warning in line 181.
- l.123: OCO-2 was already defined in l. 39. Same comment for ASCENDS a bit later.
- l.168: the motivation behind removing the linear trend is obscure. For the constant value removal, I do not see how this affects the calculation of the autocovariances.

- I.208 and 219: Kalman filters are used to control fluid state variables, but not boundary conditions such as surface fluxes. They are outside the scope of the discussion, unless the authors refer to simplifications like the Kalman smoother, but in this case the assimilation window covers periods much larger than the observation error correlations lengths which are discussed here.
- Section 3.1.3: Some main elements (the tridiagonal influence matrix, the statistically-optimal inflation factor) have been shown years ago by Chevallier et al (<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2007GL030463>) in a short paper. It may have been too brief, but was much more accessible, I think.