

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2021-306-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2021-306

Anonymous Referee #1

Referee comment on "A fast, single-iteration ensemble Kalman smoother for sequential data assimilation" by Colin Grudzien and Marc Bocquet, Geosci. Model Dev. Discuss.,
<https://doi.org/10.5194/gmd-2021-306-RC1>, 2021

In this paper, the authors derive a single iteration ensemble Kalman smoother (SIEnKS) similar in spirit to the ensemble Kalman smoother (EnKS), and iterative ensemble Kalman smoother (IEnKS). They first frame popular smoothing methods as Bayesian maximum a posteriori estimators and then derive the SIEnKS through this perspective. They perform experiments on the 40 variable Lorenz-96 to demonstrate their method and compare it to EnKS, IEnKS, and a linearized IEnKS.

This is a novel and interesting idea, which the paper describes clearly. Reading this manuscript was exciting. I liked the presentation and analysis of each method. Including the pseudocodes was a great idea and helped clear questions about the implementation.

General Comments

1) It is easy to feel lost in the jargons. It would be helpful to the reader if the authors were to explain certain words more explicitly, which I have outlined in the specific comments section.

2) The experiments were done on Lorenz-96 with the 40-variable setting while observing all states. The experiments would be more compelling in an operational sense with a larger test problem and sparse observations. Are you not running these experiments purely because of the difficulty in formulating localization?

Specific Comments

1) Could you specify what is meant by outer-loop? Does this refer to the filtering step which is done first and then the inner loop of smoothing the lagged states?

- 2) What is meant by "online" forecast systems? Does this mean real-time?
- 3) In line 10 of the abstract you write "... prediction/posterior accuracy ...". I feel that the word "posterior" must be replaced by "analysis", since prior/posterior is used with respect to the distributions while forecast/analysis is used with respect to the sample/realization.
- 4) In line 18, you write "four-dimensional ensemble var". If you just mean 4D-EnVar, you should go with writing 4D-EnVar.
- 5) In line 40, the last word "by" should be replaced by "be" to read "... may instead BE dominated by ...".
- 6) In equations 28b) and 28c), you should be having $(I_{\{N_x\}} + \Gamma_1^T \Gamma_1)$ instead of the incorrect $(I_{\{N_x\}} + \Gamma_1 \Gamma_1^T)$ which you have. It should also be replaced in equation 36c.