# 1 Review 1

I like this paper. The authors have done a great job at giving a walk through on exactly what is going on in a way that is easy to understand. It goes through each method nicely with no real assumptions made that might make the reader confused.

Major Comments:

None **Thank you for your kind comments and interest in this work.**

Minor Comments:

Line 19 - A comma after "However" is probably needed here. **Done.**

Line 59 - A comma after "In addition" is probably right here **Done.**

Line 59:60 - "that has been developed over thousands of person years" sounds really odd to me. **Changed**

Section 3.1:

a) What is the timestep size? I am guessing it's 0.01 (?), but I think it should be noted. **Now stated explicitly (it was previously implicit).**

b) Why did you not observe z? I am curious about why this choice was made instead of observing all variables.

**The true answer here is that we had some legacy code for L63 set this way. But in fact it reveals some interesting smoothing results because z errors are not strongly correlated with x,y errors climatologically. The full KS manages to deal with this and improves z as it follows how correlations change with time which the simplified smoother does not, thus revealing an aspect of the algorithms which may not have been apparent otherwise.**

c) Why choose x=5 and y=20 to be your observation frequencies?

**This frequency is somewhere near the divergence limit with these error settings etc, as can be seen the RMSEs in Figs 2,4 do rise and fall significantly. It was important to have a system which does not completely converge under the filter leaving room for the smoother to significantly improve results.**

d) Why retain 5%? I know you said it's to avoid divergence, do you have a citation for this or was it tested by you?

**This was obtained by testing in the early stages. With essentially the same idealised system in Dong et al (2021) but using 3DVar we were using climatological error covariances throughout which was working well but the Kalman filter uses flow-dependent error covariances, and this sometimes leads to irreversible error collapse in the very nonlinear L63. The retention of a climatological component prevents this, ensuring new data always gets to be used.**

Points b-c are just me curious about why the choices were made. I don't think more experiments are needed, but maybe add something to give context to why you decided on these parameters. Could be easy like referring to papers whose setup you are using, or just a small sentence here and there. **We have also added the L63 model set up reference at the beginning of the model description.**

Line 183 - I like L=40 is good here, how general do you think that is?

**We found that the full KS RMSEs have mostly converged at L=40 with the timestep, observation frequencies and observation error choices. This will also be a very specific function of the L63 system related to the x,y oscillations and the lobe transition frequency, although we have not investigated in detail.**

Line 198 - What do you mean by "true RMSE" here? I don't understand what these dotted blue/green uncertainty estimates are in agreement with. **Changed to RMSE against truth.**

Line 252 - Say what "i" refers to. **now defined -is the normalised anomaly of the *i*th ensemble member from the mean.**

Line 286 - "Fig. 2, reflecting the improved assimilation approaches." Is this supposed to be a stand alone sentence? It feels like something is missing to make it a complete sentence. **there was a . instead of a , now corrected.**