Respond to the Anonymous Referee #1

In general: Thanks a lot for your comments and suggestions!

To the “Specific comments”:

1. The reference systems, we use here, are the geodetic (geographic, cf. Section 2 in https://arxiv.org/ftp/arxiv/papers/1611/1611.10321.pdf) and geomagnetic (for the rotation, cf. Section 3.1 in https://arxiv.org/ftp/arxiv/papers/1611/1611.10321.pdf). We will add this information to the manuscript.
2. You are right, the storage of 100 ensembles is indeed quite disk space-intensive. One analyzed solution with 100 ensembles needs around 150 MB. However, I think, it is not necessary to store all the ensembles of the analyzed solutions over the whole period of time, one is interested in. Ones we propagated an analyzed ensemble matrix in time, one could calculate the ensemble mean (which represents the reconstructed Ne for this time stamp) and delete the ensemble matrix. Regarding the computational effort, it is not high, and it depends a lot on the propagation method. Because the calculation of the ensembles by the empirical background model is a bit time-consuming and depends on the model, one uses. In general, to apply a relatively easy manageable empirical model, like NeQuick (which needs just F10.7 index as external input), is much easier, than to run a physical model, like TIEGCM https://ccmc.gsfc.nasa.gov/models/modelinfo.php?model=TIE-GCM , with all the inputs one need. Our Analysis Step (data assimilation) plus the Forecast Step cost together around 1 minute on a Linux machine.
3. The formula for ratio in Eq. 6 stands for relative error caused by using of $Rot\left(x^b(t_n)\right)$ instead of $x^b(t_{n+1})$ and represents in this way the relative mean error introduced by approximation of the true state at time t_{n+1} by a simple rotation of the true state at time t_n. Some of the factors in the formulas, like 3 in Eq. 6; 8/10 in Eq. 9 were derived by try-and-error method. This means, we did several runs with different factors and then validate the results. At the end, the factors delivering the best results were chosen for the paper. We will put a corresponding clarifying note in the paper.
4. Lines 268 – 269: it is indeed written a bit confusing in the manuscript. We will correct it.
5. Lines 289 – 291: It is indeed a big difference. The only explanation we see is that the GRACE satellites were flying below the SWARM A, and thus near on the hmF2 value. We checked again the orbit heights of SWARM A and GRACE satellites for the periods of interest within 2015. We found a small but in this context probably important mistake - the GRACE orbit was around 430 km (not 450 km, as was given by us in the paper). We correct the value in the paper.
6. Lines 318 – 319: you right, thanks. This is only true for the SWARM A residuals. We add the correction to the manuscript text.
7. Eq. (8): exact.
8. Eq. (9): please see number 3. above.
9. Line 157: we wrote 0.5/100 and 1/100 to make more clear that we mean one-half and one percent. The factors are chosen by try-and-error.
10. Line 306: The STD value in the Fig. 11 is 7.17 TECU, this is why we round it to 7 TECU.