Comment on gmd-2021-302

Anonymous Referee #2

Referee comment on "Implementation of an ensemble Kalman filter in the Community Multiscale Air Quality model (CMAQ model v5.1) for data assimilation of ground-level PM$_{2.5}$" by Soon-Young Park et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-302-RC2, 2021

General Comments:

This manuscript describes the initialization of ground-level PM2.5 for a chemical transport model using the ensemble Kalman Filter (EnKF) method. Authors implemented EnKF in the CMAQ model and claim this method improves PM2.5 predictability compared to the 3DVar and No DA. The PM2.5 predictability of South Korea is further improved if the EnKF is applied not only to the nest domain but also mother domain with ground PM2.5 observation.

The manuscript looks like to have a reasonable structure for the paper. It will be available for publication if it is improved with a minor revision.

Specific Comments:

- In the manuscripts, ICs is an abbreviation for both initial conditions and initial concentrations. The same goes for BCs. Authors should only use the abbreviation "C" for either condition or concentration
- What are the background variables that are input to EnKF? I believe the EnKF uses background error covariance between PM2.5 and some meteorological variables. Is it correct that the experiments only update PM2.5, so other meteorological variables are the same as before DA? Adding a list of background and analysis variables is recommended.
- I am wondering about the observation operator for the PM2.5. Is PM2.5 one of the

background variables? If not, the authors need to introduce the observation operator for it to calculate observation operator. Some descriptions for the observation operator would be better to be added in the manuscript.

- Line 174: What is the inflation parameter (alpha) for RTPS?
- Since you have described the parameter for the localization, it would be better to also describe the parameter for RTPS.
- Line 175: Why do you apply the RTPS to ensemble before and after DA? The RTPS compares ensemble spread before and after DA because the amount of inflation in RTPS is proportional to the ensemble spread reduced by the DA. So theoretically, RTPS can be applied only once after DA.
- When you describe DA_icbc, can you show the pm2.5 field for domain1 which contains domain 2? For example, horizontal field of PM2.5 as in figure 4, but with domain 1. The distribution of PM2.5 over domain 1 can be clearer evidence showing the effect of boundary conditions.

Technical Comments:

- Line 245: sate --> state