The authors appreciate the reviewers for the effort to review our manuscript and to provide constructive comments. As suggested, we carefully revised the manuscript thoroughly according to the valuable advices. Listed below are our point-by-point responses in blue to the reviewer's comments in black. The modifications corresponding to the comments and the revised language and grammars in the manuscript are marked in red.

Anonymous Referee #1

General comments: This paper describes a chemistry-transport model that is implemented with a data assimilation compartment using the Parallel Data Assimilation Framework. First of all, I would admit that I don't expertise in the atmosphere chemistry area, all my comments are from the data assimilation with PDAF. As for as I know, the online data assimilation approach with PDAF is first thoroughly described in Nerger et al., (2019) GMD paper, where the structure, the algorithm, the implementation are shown based on a climate model AWI-CM. Actually, such implementation has been widely used also for other ocean models such as FESOM, MITgcm etc. Another group like C-Coupler (Liu et al., 2020 GMD) provides similar data assimilation functions as well. I found this research is comprehensive with both technical and experimental perspectives. The results are clearly present and well organized. I would recommend acceptance after minor revisions. In general, I think the paper is well written with adequate evidence for their results.

Reply: We thank the reviewer for the positive assessment and constructive suggestions of our manuscript.

Comment 1: As for the paper structure, I would suggest the authors rephrase 2.3.1 and 2.3.2, i.e., the technical parts. Most of these implementation details are already well described in Nerger et al., 2019. I didn't see too many differences compared to Nerger's. Currently, it's rather a repetition. Please cite this paper directly and show your differences to condense the context. For 2.3.2, again, not necessarily a repetition of these algorithm details, which are well-known in amount of studies. The authors should concentrate on things that are distinguishable from others' work. For example,

localization radius, whether it is sensitive to your configuration (I see the authors cite two other researches, but actually these are different stories if the configuration changes based on my practices), are you using same localization radius for different observations; forgetting factor, which value is set, why is that; things like that.

Reply:

Firstly, we agree with that our manuscript should focus on the differences compared to other studies. The first part of Sec. 2.3.1 is the implementation of two level of parallelization based on MPI. We have cut out redundant descriptions but not deleted the content altogether because it is the first time to introduce PDAF into the study area of CTM. The second part of Sec. 2.3.1 is the main program flow. The third part of Sec. 2.3.1 is the description of the dimension of the state vectors used in PDAF which is specially redesigned in our work. The fourth part of Sec. 2.3.1 is two modules dealing with the data transfer before and after the time loop of the NAQPMS-PDAF which are also designed in our work. In Sec. 2.3.2, we have cited the literature about description of forecast step. The analysis step of ESTKF is retained because ESTKF is recently developed especially used in the area of CTM and the difference between ESTKF and other deterministic ensemble filter is mainly the computation of the ensemble transformation in the error subspace. Another anonymous reviewer wants more information about the PDAF and the ESTKF algorithm, so the revision above is the result of both considering two reviewers' comments.

Secondly, two kinds of observations (surface PM2.5 mass concentration and vertical profiles of aerosol extinction coefficients measured by ground-based lidar) are assimilated into NAQPMS-PDAF in this study. We set the localization radius as 200 km for both observations. For surface PM2.5 concentration, we follow Kong et al. (2020) and set the localization radius as 200 km, because the kind of observation, the atmospheric chemistry-transport model (NAQPMS) as well as the ensemble filter algorithm of this study is same as their work. Therefore, we here focus the localization radius of ground-based lidar and forgetting factor of the data assimilation system, which are set as 200 km and 0.96 in this study. The setting of forgetting factor is omitted in the original manuscript and have been revised.

The several sensitivity tests have been made to supplement the setting of these two data assimilation parameters in this study. We refer to Gillet-Chaulet (2020) about the period chosen for sensitivity tests when assimilating real observation under limited computational resources. For sensitivity tests, we choose the study period from 00:00 UTC 23 April to 05:00 UTC 23 April 2019 with abundant pollution plume measured by the ground-based lidar. A series of sensitivity tests are performed with the localization radius (5 km, 50 km, 100 km, 150 km, 200 km, 250 km, 300 km and 400 km) and forgetting factor (0.6, 0.7, 0.8, 0.9, 0.92, 0.94, 0.96, 0.98 and 1.0). The configuration of data assimilation is same as the NP-LIDAR-PM25 experiment in the manuscript expect for the study period. The results of sensitivity tests are evaluated by the VE sites (the ground-based lidar measurements not assimilated) of model domain. Following Nerger (2015) and Nerger (2021), Figure S1 and Figure S2 show the time-mean RMSE and Pearson correlation coefficient for the sensitivity tests, respectively. As we can see in Figure S1, RMSE of aerosol extinction coefficients converges for all combinations of localization radius and forgetting factor. The minimum RMSE of 0.36 1/km is obtained for localization radius of 200 km and forgetting factor of 0.96, 0.98 and 1.0. The maximum Pearson correlation coefficient of 0.75 is obtained for localization radius of 150 km and 200 km and is not sensitive to forgetting factor when forgetting factor is larger than 0.9. Forgetting factor is used to inflate the forecast covariance matrix to reduce under-sampling issues, especially in the long run (Pham et al., 1998). Although the statistical results vary slightly with forgetting factor due to relatively short run time, the combined results of RMSE and Pearson correlation coefficient can provide the optimal parameters in the series of sensitivity tests. To sum up, the forgetting factor of 0.96 and localization radius of ground-based lidar of 200 km is the most optimal parameters. Moreover, the setting of localization radius is same as Cheng et al. (2019) performing ensemble filter to assimilating lidar measurements, and is also close to Ma et al. (2020) performing ensemble filter to assimilating aerosol extinction coefficient profiles measured by ground-based lidar.

Changes in manuscript: Changes have been made in Sec. 2.3.1 in Line 176-183 of the revised manuscript and revised text is “As the NAQPMS described in Section

2.1 is well written and its source code is available, this study chooses the online method to couple the PDAF with the NAQPMS in order to gain the best performance. The core modification in the coupling is parallelization for ensemble simulations. Message Passing Interface standard (MPI; Gropp et al., 1994) both used in NAQPMS and PDAF allows each process to handle distributed parts of a program and data exchange. The communicator MPI_COMM_WORLD is used in NAQPMS as one-level parallelization to improve computational efficiency and the distribution of processes is exemplified in Fig. 1a.".

Also, changes have been made in Line 306-307 and revised text is "We set the horizontal localization radius and forgetting factor to 200 km and 0.96 according to a series of sensitivity tests in the Supplement and other related studies (Kong et al., 2020; Zhao et al., 2020; Cheng et al., 2019b; Ma et al., 2020)".

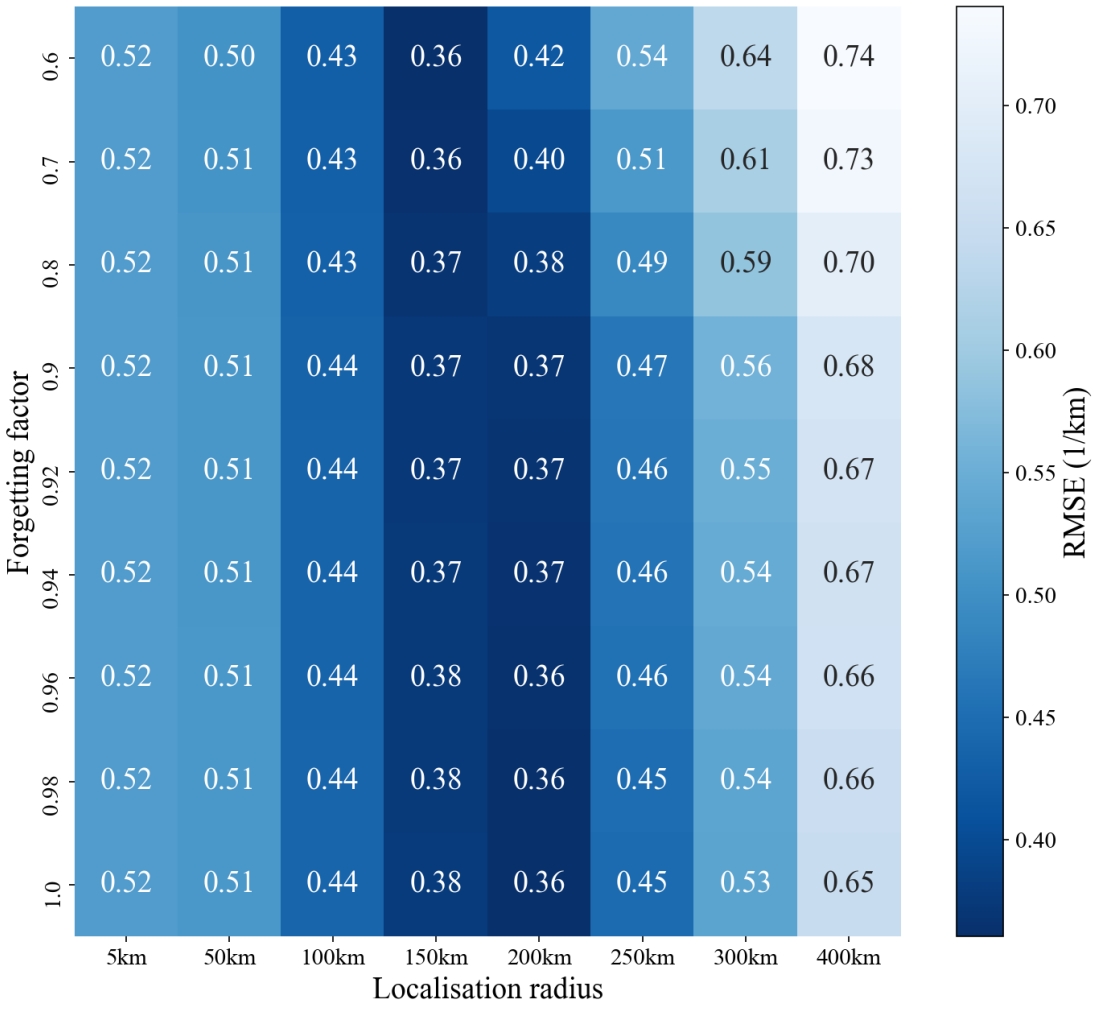

Figure S1. RMSE for sensitivity tests with localization radius of ground-based lidar (5 km, 50 km, 100 km, 150 km, 200 km, 250 km, 300 km and 400 km) and forgetting factor of NAQPMS-PDAF (0.6, 0.7, 0.8, 0.9, 0.92, 0.94, 0.96, 0.98 and 1.0).

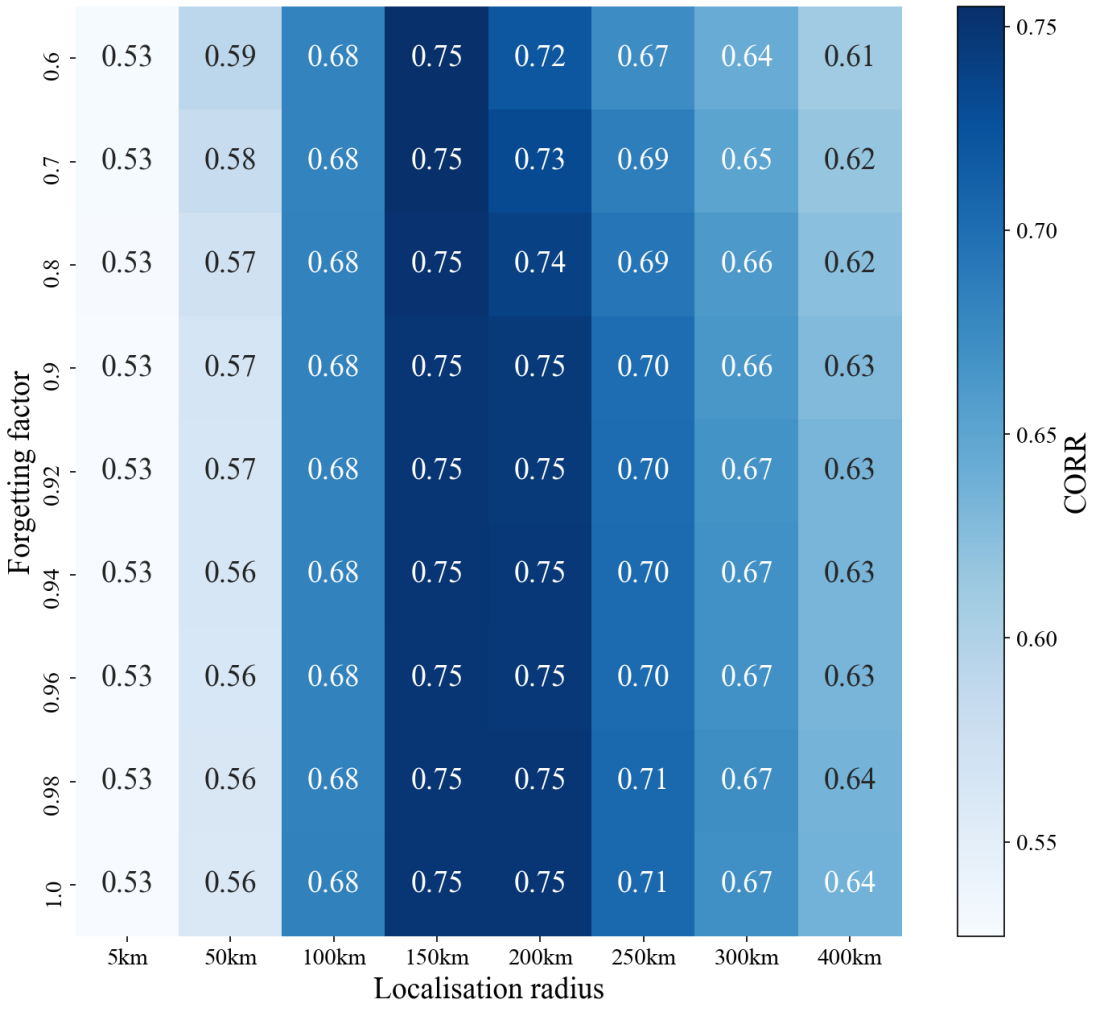

Figure S2. Same as Figure S1 but with Pearson correlation coefficients which denoted by CORR.

Comment 2: Please use a larger fontsize for Figure 10 and 11. The caption tells "the increments in the RMSEs of the surface PM2.5 forecasts (g, h, i) and the increments in the RMSEs of the surface PM2.5 forecast (j, k, i)". Something goes wrong there?

Reply: We agree with the comment. The size of font in Figure 10 and Figure 11 is small as well as the dpi, and these two figures have been revised. "increments in the" in the caption is superfluous and has been removed.

Changes in manuscript: Changes have been made in Line 1291-1295 of the revised manuscript and revised text is "the RMSEs of the surface $PM_{2.5}$ forecasts (g, h, i) and the increments in the RMSEs of the surface $PM_{2.5}$ forecast (j, k, i)".

Comment 3: Taking Figure 14 and 15, the authors found that the system seems not well constrained by DA in high level. Could the authors add some discussions about the physical reason or other aspects behind this problem?

Reply: Atmospheric chemistry-transport model (CTM) is an approximate representation of the evolution of air pollutants, which contains several physical and chemical processes such as direct emission, advection, diffusion, dry deposition, wet deposition, aqueous chemistry, gas-phase chemistry, heterogeneous chemical processes and so on. The concentration variability of gases and aerosols are not only affected by the above processes and are also constrained by meteorological input, initial conditions and boundary conditions. Emission is one of the most significant uncertainty sources. Studies on the analysis and forecast of air pollutants with CTM usually perturbed the initial emission to create initial ensemble (Tang et al., 2011; Kong et al., 2020; Dai et al., 2019; Cheng et al., 2019). The inversion estimation of emissions with CTM data assimilation is even a research hotspot (Kong et al., 2019; Wu et al., 2020; Feng et al., 2020; Tang et al., 2013; Ma et al., 2019; Dai et al., 2021), which is not the focus in our work.

Emission which is one of input of CTM can be divided into emission from agriculture, biomass burning, industry, power plant, resident, transportation. However, the most kinds of emissions mainly concentrated around the surface. Only biogenic emission, industrial emission and emission from power plant can emit air pollutants at

a certain altitude. As a result, after perturbing initial emission to create initial ensemble, the error character (represented by the ensemble spread) of extinction coefficient profiles on the background simulations with a clear decreasing from the surface to a certain altitude (blue curve in Figure 15a). It means that the analysis increment of each assimilation cycle tends to apportion more aerosol concentration (which can transform to extinction coefficient with observation operator) near surface. Therefore, the significant adjustment of aerosol extinction coefficients mainly occurs below the altitude of 3 km (Figure 14). The averaged extinction profiles (red curve in Figure 14) show a maximum value around the altitude of 600 m, which is in the planetary boundary layer (PBL). The most polluted plume occurs in PBL (usually 1 ~ 2 km) and affects the concentration of surface $PM_{2.5}$ (Yang et al., 2010; Lei et al., 2021).

In summary, the limited altitude of emission which is the perturbed to create ensemble limits the constraint of DA in high level. However, it is well constrained the aerosol profiles in the PBL which can significantly affect the surface.

Changes in manuscript: Changes have been made in Line 676-680 of the revised manuscript and revised text is “It means that the analysis increment of each assimilation cycle tends to apportion more aerosol concentration near the surface”.

References:

Cheng, Y., Dai, T., Goto, D., Schutgens, N. A. J., Shi, G., and Nakajima, T.: Investigating the assimilation of CALIPSO global aerosol vertical observations using a four-dimensional ensemble Kalman filter, Atmos. Chem. Phys., 19, 13445–13467, https://doi.org/10.5194/acp-19-13445-2019, 2019.

Dai, T., Cheng, Y., Suzuki, K., Goto, D., Kikuchi, M., Schutgens, N. A. J., Yoshida, M., Zhang, P., Husi, L., Shi, G., and Nakajima, T.: Hourly Aerosol Assimilation of Himawari-8 AOT Using the Four-Dimensional Local Ensemble Transform Kalman Filter, J. Adv. Model. Earth Syst., 11, 680–711, https://doi.org/10.1029/2018MS001475, 2019.

Dai, T., Cheng, Y., Goto, D., Li, Y., Tang, X., Shi, G., and Nakajima, T.: Revealing the sulfur dioxide emission reductions in China by assimilating surface observations in WRF-Chem, Atmos. Chem. Phys., 21, 4357–4379, https://doi.org/10.5194/acp-21-4357-2021, 2021.

Feng, S., Jiang, F., Wu, Z., Wang, H., Ju, W., and Wang, H.: CO Emissions Inferred From Surface CO Observations Over China in December 2013 and 2017, J. Geophys. Res. Atmos., 125, https://doi.org/10.1029/2019JD031808, 2020.

Gillet-Chaulet, F.: Assimilation of surface observations in a transient marine ice sheet model using an ensemble Kalman filter, The Cryosphere, 14, 811–832, https://doi.org/10.5194/tc-14-811-2020, 2020.

Kong, L., Tang, X., Zhu, J., Wang, Z., Pan, Y., Wu, H., Wu, L., Wu, Q., He, Y., Tian, S., Xie, Y., Liu, Z., Sui, W., Han, L., and Carmichael, G.: Improved Inversion of Monthly Ammonia Emissions in China Based on the Chinese Ammonia Monitoring Network and Ensemble Kalman Filter, Environ. Sci. Technol., 53, 12529–12538, https://doi.org/10.1021/acs.est.9b02701, 2019.

Kong, L., Tang, X., Zhu, J., Wang, Z., Li, J., Wu, H., Wu, Q., Chen, H., Zhu, L., Wang, W., Liu, B., Wang, Q., Chen, D., Pan, Y., Song, T., Li, F., Zheng, H., Jia, G., Lu, M., Wu, L., and Carmichael, G. R.: A Six-year long (2013–2018) High-resolution Air Quality Reanalysis Dataset over China base on the assimilation of surface observations from CNEMC, Earth Syst. Sci. Data, 13, 529–570, https://doi.org/10.5194/essd-13-529-2021, 2020.

Lei, L., Sun, Y., Ouyang, B., Qiu, Y., Xie, C., Tang, G., Zhou, W., He, Y., Wang, Q., Cheng, X., Fu, P., and Wang, Z.: Vertical Distributions of Primary and Secondary Aerosols in Urban Boundary Layer: Insights into Sources, Chemistry, and Interaction with Meteorology, Environ. Sci. Technol., 55, 4542–4552, https://doi.org/10.1021/acs.est.1c00479, 2021.

Ma, C., Wang, T., Mizzi, A. P., Anderson, J. L., Zhuang, B., Xie, M., and Wu, R.:

Multiconstituent Data Assimilation With WRF-Chem/DART: Potential for Adjusting Anthropogenic Emissions and Improving Air Quality Forecasts Over Eastern China, J. Geophys. Res. Atmos., 2019JD030421, https://doi.org/10.1029/2019JD030421, 2019.

Ma, C., Wang, T., Jiang, Z., Wu, H., Zhao, M., Zhuang, B., Li, S., Xie, M., Li, M., Liu, J., and Wu, R.: Importance of Bias Correction in Data Assimilation of Multiple Observations Over Eastern China Using WRF-Chem/DART, J. Geophys. Res. Atmos., 125, https://doi.org/10.1029/2019JD031465, 2020.

Nerger, L.: On Serial Observation Processing in Localized Ensemble Kalman Filters, 143, 15, 2015.

Nerger, L.: Data assimilation for nonlinear systems with a hybrid nonlinear Kalman ensemble transform filter, Quart J Royal Meteoro Soc, qj.4221, https://doi.org/10.1002/qj.4221, 2021.

Pham, D. T., Verron, J., and Christine Roubaud, M.: A singular evolutive extended Kalman filter for data assimilation in oceanography, Journal of Marine Systems, 16, 323–340, https://doi.org/10.1016/S0924-7963(97)00109-7, 1998.

Tang, X., Zhu, J., Wang, Z. F., and Gbaguidi, A.: Improvement of ozone forecast over Beijing based on ensemble Kalman filter with simultaneous adjustment of initial conditions and emissions, Atmos. Chem. Phys., 11, 12901–12916, https://doi.org/10.5194/acp-11-12901-2011, 2011.

Tang, X., Zhu, J., Wang, Z. F., Wang, M., Gbaguidi, A., Li, J., Shao, M., Tang, G. Q., and Ji, D. S.: Inversion of CO emissions over Beijing and its surrounding areas with ensemble Kalman filter, Atmospheric Environment, 81, 676–686, https://doi.org/10.1016/j.atmosenv.2013.08.051, 2013.

Wu, H., Tang, X., Wang, Z., Wu, L., Li, J., Wang, W., Yang, W., and Zhu, J.: High-spatiotemporal-resolution inverse estimation of CO and NOx emission reductions during emission control periods with a modified ensemble Kalman filter, Atmospheric Environment, 236, 117631, https://doi.org/10.1016/j.atmosenv.2020.117631, 2020.

Yang, T., Wang, Z., Zhang, B., Wang, X., Wang, W., Gbauidi, A., and Gong, Y.: Evaluation of the effect of air pollution control during the Beijing 2008 Olympic Games using Lidar data, Chin. Sci. Bull., 55, 1311–1316, https://doi.org/10.1007/s11434-010-0081-y, 2010.