Supplement of Geosci. Model Dev., 15, 3555–3585, 2022
https://doi.org/10.5194/gmd-15-3555-2022-supplement

Geoscientific Model Development
Open Access
EGU

Supplement of

An aerosol vertical data assimilation system (NAQPMS-PDAF v1.0): development and application

Haibo Wang et al.

Correspondence to: Ting Yang (tingyang@mail.iap.ac.cn) and Zifa Wang (zifawang@mail.iap.ac.cn)

Contents of this file

Text S1: the comparison of assimilation cycle

The assimilation window denotes the time length of the assimilation period (Wu et al., 2008). I guess the "assimilation window" in this comment actually means assimilation cycling. Therefore, we here reply on the setting of 1-hr as assimilation window and continuous 1-hr cycling, respectively.

Firstly, the EnKF system used in our work provides possibilities for using a short assimilation window to have the ensemble perturbations evolve linearly (Houtekamer and Zhang, 2016; Liu et al., 2019), while a 4D-Var system needs to keep a long window to reduce the effect of the initially specified covariances (Pires et al., 1996). So we choose 1-hr as assimilation window in our EnKF system (NAQPMS-PDAF) as other similar studies do (Ma et al., 2020, 2019; Liu et al., 2019; Ha et al., 2020).

Secondly, the assimilation cycling is set as 1-hr in our work. On one hand, the main reason is that our manuscript focuses on investigating the parallel performance of NAQPMS-PDAF which is online coupled and the improvement of vertical profiles after assimilating aerosol extinction coefficient profile. The performance of ensemble forecast after ensemble filter of 1-hr or 6-hr is not the focus. Therefore, we increase the frequency of assimilation from every 6-hr to 1-hr. On the other hand, the 6-hr assimilation cycle used in similar studies (Ma et al., 2020, 2019; Pang et al., 2018; Liu et al., 2011) follows the model configuration of assimilating satellite data with coarse temporal resolution. However, the lidar measurements used in our work can provide large temporal variability with temporal resolution of 1-hr.

Therefore, we perform NP-LIDAR-6HR experiment shown in Table S1 to

compare the assimilation effect between performing 1-hr and 6-hr cycling. Figure S1 and S2 are scatter plots and frequency distribution of extinction coefficients from the model versus the ground-based lidar measurements averaged over 5 DA sites and 6 VE sites of FR, NP-LIDAR and NP-LIDAR-6HR experiment, which are corresponding to Figure 8 and Figure 9 in our manuscript, respectively. As shown in Fig. S1f, extinction coefficient scatters are mainly concentrated around the 1:1 line in the NP-LIDAR-6HR experiment at DA sites. The RMSE and CORR value decreases (increases) from 0.42 1/km (0.33) in the FR experiment to 0.18 1/km (0.89) in the NP-LIDAR-6HR experiment, showing that the effect of assimilating lidar measurement with 6-hr cycling is positive. As shown in Fig. S1e and Fig. S1f, the RMSE value of the NP-LIDAR and NP-LIDAR-6HR experiment is 0.16 1/km and 0.18 1/km, respectively. The CORR value of these two experiments is 0.91 and 0.89. Note that the analysis time period in the NP-LIDAR-6HR is almost 6 times shorter than that in the NP-LIDAR experiment, result in a smaller number of scatters in Fig. S1f than that in Fig. S1e. Fig. S2c and Fig. S2d also show that the performance of BIAS of the NP-LIDAR experiment is slightly better than that of the NP-LIDAR-6HR experiment with 93 % and 92 % scatters within $|BIAS| < 0.25$. It can be found that the statistic performance of the NP-LIDAR experiment is close to that in the NP-LIDAR-6HR experiment, and the performance of the former is slightly better than that in the latter. It means that the performance of assimilating all lidar measurements with 1-hr cycling is slightly better than assimilating the lidar measurements with 6-hr cycling under the current configuration.

As shown in Fig. S1b and Fig. S1c, the RMSE (CORR) value is 0.27 1/km (0.72) and 0.33 1/km (0.60) in the NP-LIDAR and NP-LIDAR-6HR experiment at DA sites, respectively. The frequency of $|BIAS| < 0.25$ is 80 % and 75 % in the NP-LIDAR and NP-LIDAR-6HR experiments at DA sites, which is shown in Fig. S2a and Fig. S2b. It indicates that the statistic performance of the 1-hr forecast in the NP-LIDAR is better than that in the NP-LIDAR-6HR. It can be explained that the performance of NP-LIDAR is much less affected by the attenuation of data assimilation due to 1-hr is less than 6-hr in the NP-LIDAR-6HR experiment. At the VE sites, the statistic performance of extinction coefficients in the NP-LIDAR (Fig. S1h and Fig. S2e) and NP-LIDAR-

6HR (Fig. S1i and Fig. S2f) experiment is nearly close, which both show a significantly improvement than that in the FR experiment (Fig. S1g).

Table S1. Summary of the Experimental design in AC2.

Experiments	$PM_{2.5}$ DA	Ground-based lidar DA	DA cycling
FR	No	No	/
NP-LIDAR	No	Yes	1-hr
NP-LIDAR-6HR	No	Yes	6-hr

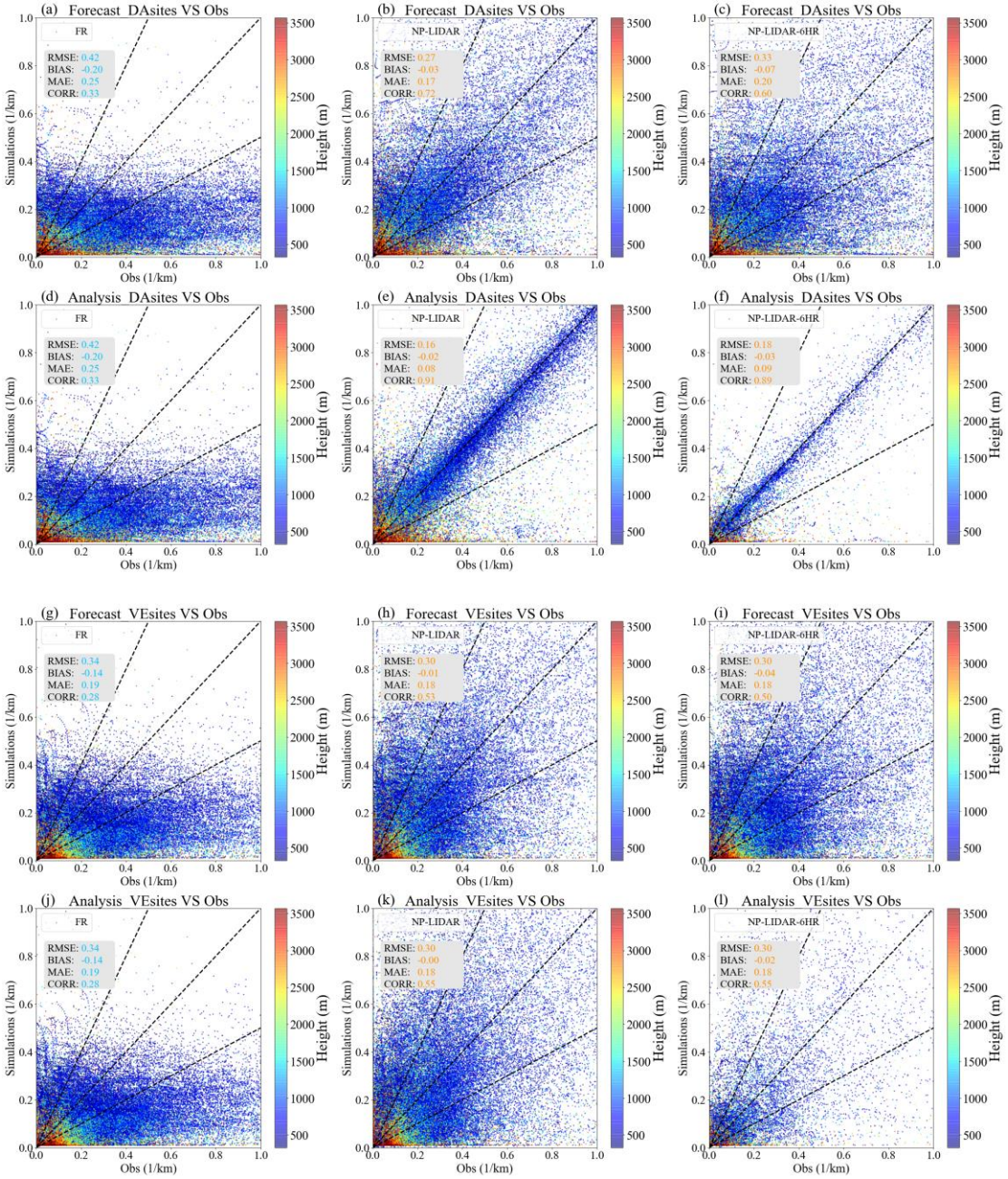

Figure S1. Scatter plots of the modeled hourly extinction coefficients at 550 nm versus the ground lidar hourly aerosol extinction coefficients at 532 nm (1/km) of forecasts of FR (a)/(g), forecasts of NP-LIDAR(b)/(h), forecasts of NP-LIDAR-6HR (c)/(i), analysis of FR (d)/(j), analysis of NP-LIDAR (e)/(k), analysis of NP-LIDAR-6HR (f)/(l), which are averaged among DA sites/VE sites. The three dashed black lines correspond to the 1:2, 1:1 and 2:1 lines in each panel.

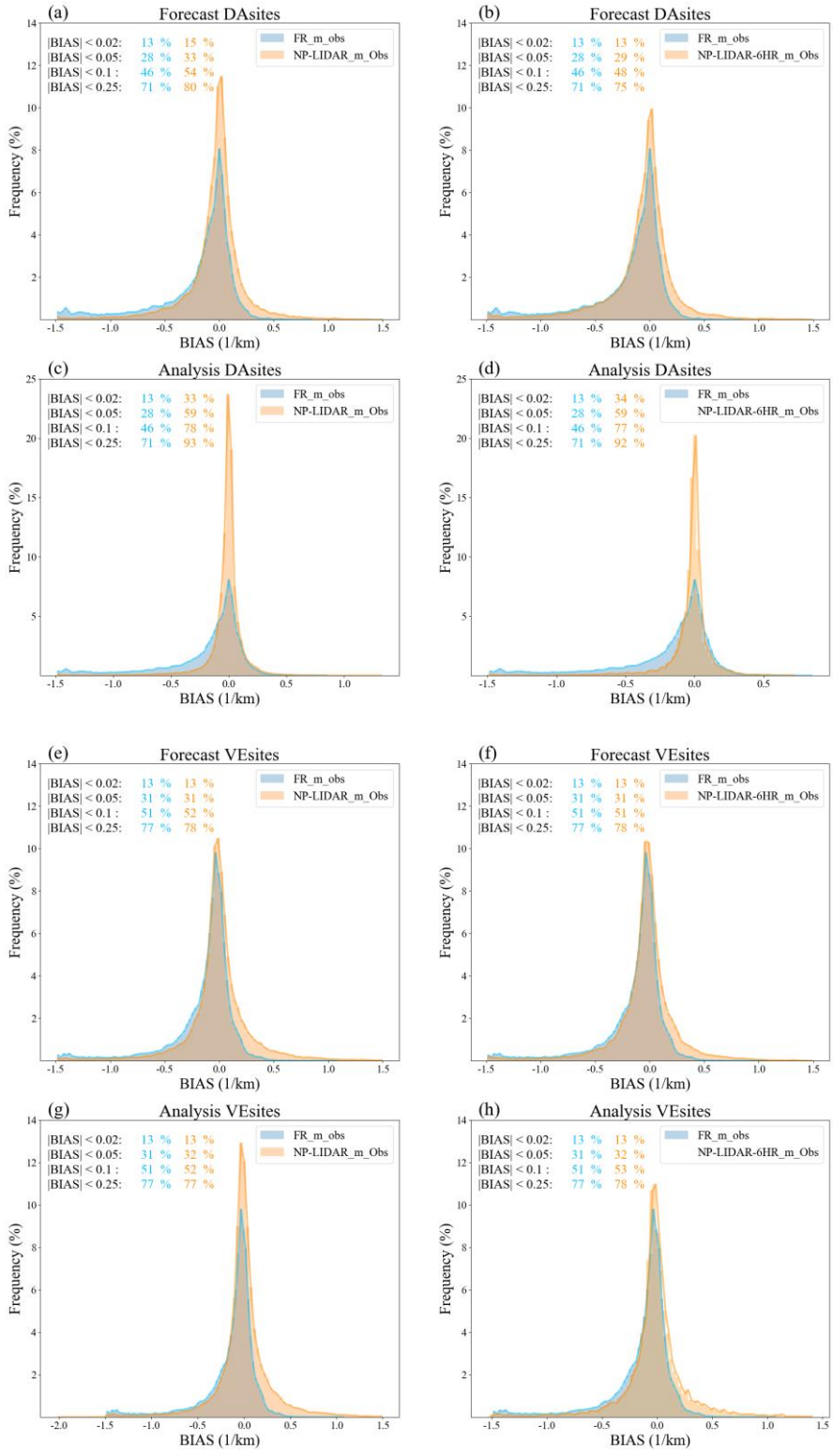

Figure S2. Frequency distributions of BIAS of forecasts of NP-LIDAR versus FR (a)/(e), forecasts of NP-LIDAR-6HR versus FR (b)/(f), analysis of NP-LIDAR versus FR (c)/(g) and analysis of NP-LIDAR-6HR versus FR (d)/(h), which are averaged among DA sites/VE sites.

Text S2: the sensitivity tests on localization radius and forgetting factor

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 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 S3 and Figure S4 show the time-mean RMSE and Pearson correlation coefficient for the sensitivity tests, respectively. As we can see in Figure S3, 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.

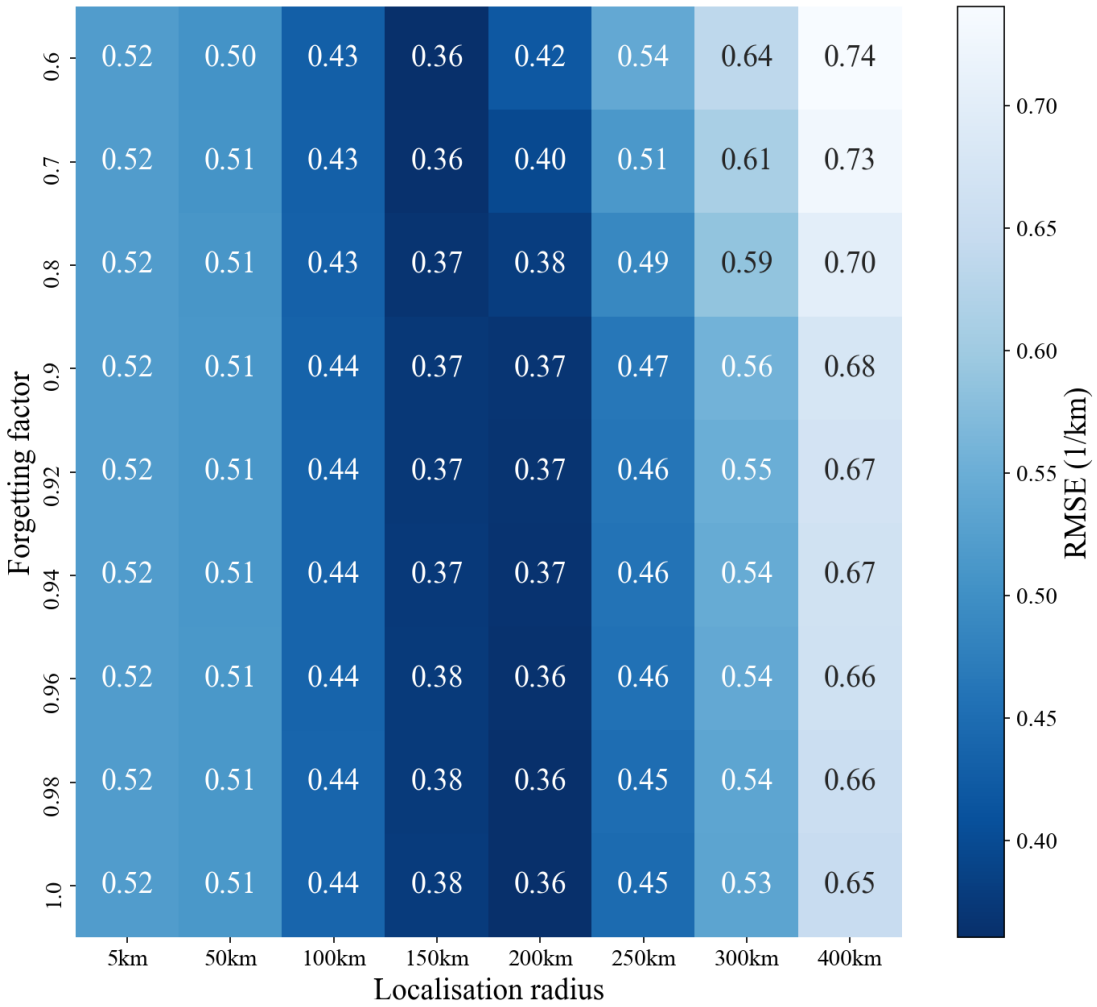

Figure S3. RMSE for sensitivity tests with localization radius of ground-based lidar (5 km, 50 km, 100 km, 150 km, 200 km, 250 km, 300 km and 400 km) and forgetting factor of NAQPMS-PDAF (0.6, 0.7, 0.8, 0.9, 0.92, 0.94, 0.96, 0.98 and 1.0).

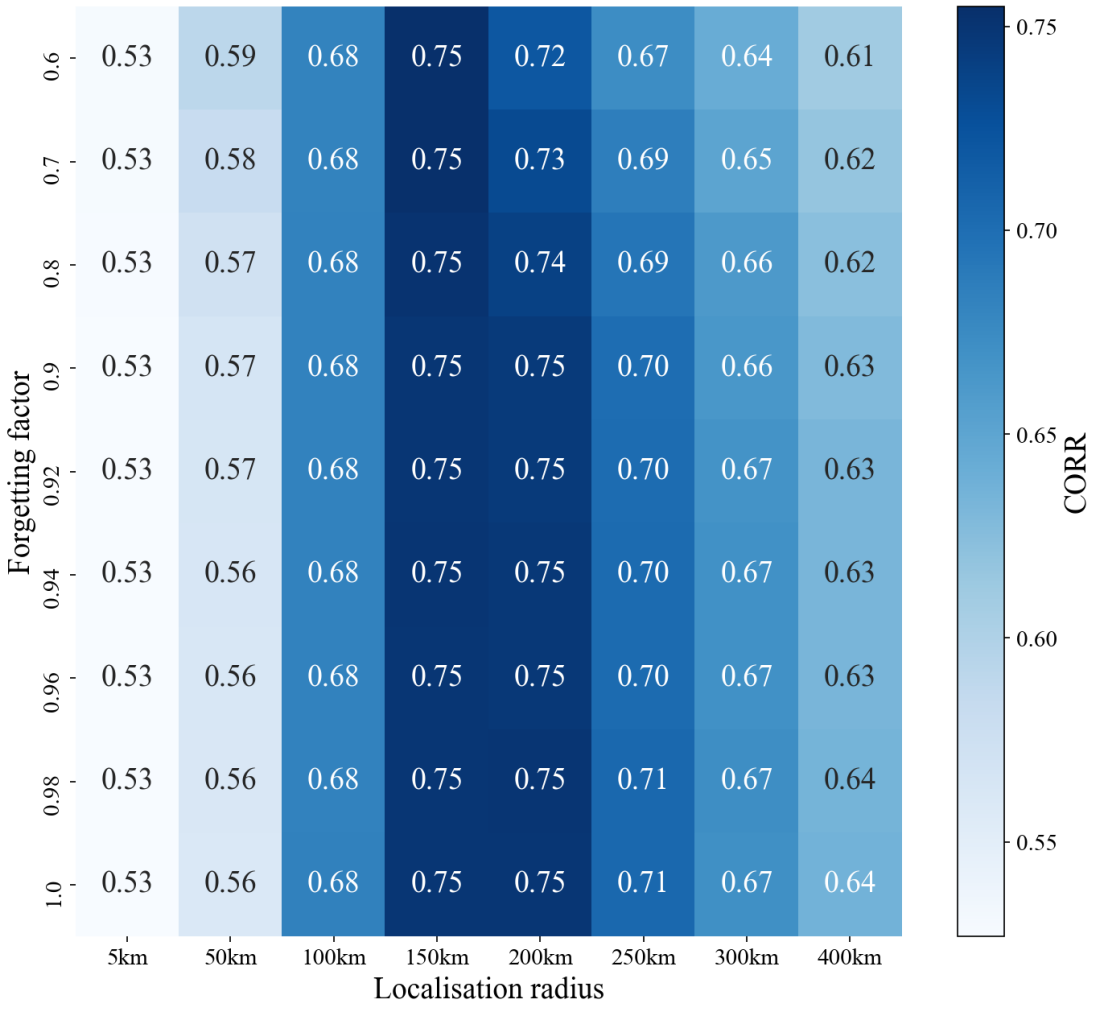

Figure S4. Same as Figure S3 but with Pearson correlation coefficients which denoted by CORR.

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.

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.

Ha, S., Liu, Z., Sun, W., Lee, Y., and Chang, L.: Improving air quality forecasting with the assimilation of GOCI aerosol optical depth (AOD) retrievals during the KORUS-AQ period, 20, 6015–6036, https://doi.org/10.5194/acp-20-6015-2020, 2020.

Houtekamer, P. L. and Zhang, F.: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation, Mon. Wea. Rev., 144, 4489–4532, https://doi.org/10.1175/MWR-D-15-0440.1, 2016.

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.

Liu, Y., Kalnay, E., Zeng, N., Asrar, G., Chen, Z., and Jia, B.: Estimating surface carbon fluxes based on a local ensemble transform Kalman filter with a short assimilation window and a long observation window: an observing system simulation experiment test in GEOS-Chem 10.1, Geosci. Model Dev., 12, 2899–2914, https://doi.org/10.5194/gmd-12-2899-2019, 2019.

Liu, Z., Liu, Q., Lin, H.-C., Schwartz, C. S., Lee, Y.-H., and Wang, T.: Three-dimensional variational assimilation of MODIS aerosol optical depth: Implementation and application to a dust storm over East Asia, 116, D23206, https://doi.org/10.1029/2011JD016159, 2011.

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.

Pang, J., Liu, Z., Wang, X., Bresch, J., Ban, J., Cnen, D., and Kim, J.: Assimilating AOD retrievals from GOCI and VIIRS to forecast surface PM2.5 episodes over Eastern China, 179, 288–304, https://doi.org/10.1016/j.atmosenv.2018.02.011, 2018.

Pires, C., Vautard, R., and Talagrand, O.: On extending the limits of variational assimilation in nonlinear chaotic systems, 48, 96–121, https://doi.org/10.1034/j.1600-0870.1996.00006.x, 1996.

Wu, L., Mallet, V., Bocquet, M., and Sportisse, B.: A comparison study of data assimilation algorithms for ozone forecasts, J. Geophys. Res., 113, D20310, https://doi.org/10.1029/2008JD009991, 2008.

Zheng, H.: Improvement of PM2.5 Forecast by Data Assimilation of Ground and Lidar Observation, doctor, University of Science and Technology of China, 2018.