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 reviewer's comments are reproduced (black) along with our replies and changes made to the text in the revised manuscript.

Anonymous Referee #2

General comments: Aerosol vertical structure is important for investigating global climate change, air pollution transport and control. The authors developed an online data assimilation system for vertical observation by coupling NAQPMS with PDAF, which offers spatiotemporally continuous aerosol vertical profiles. The system can make efficient use of parallel computational resources and produce great improvement in the aerosol vertical structure and surface PM2.5 concentration. Overall, the whole manuscript is within the scope of GMD and well written. I think the research is novelty to impact on other one's research and could be reconsidered after major revisions.

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

Comment 1: In previous assimilation studies of satellite products, 6 hr or 12 hr has been chosen as the assimilation window. The whole manuscript is based the analysis and subsequent 1-hr forecast. So why do authors choose one hour as DA window? What is the difference between 1 hr and 6 hr or 12 or in terms of assimilation effect?

Reply: 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).

Changes in manuscript: Changes have been made in Line 375-390 of the revised manuscript and revised text is “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., 2019b), while a 4D-Var system needs to keep a long window to reduce the effect of the initially specified covariances (Pires et al., 1996). Therefore, we choose 1-hr as assimilation window in NAQPMS-PDAF as other similar studies (Liu et al., 2019b; Ma et al., 2019; Ha et al., 2020) do. The assimilation cycle is set as 1-hr. On one hand, our work focuses on investigating the parallel performance of NAQPMS-PDAF and the improvement of vertical profile simulations after assimilating aerosol extinction coefficient profiles. The performance of ensemble forecast after ensemble filter is not the focus. On the other hand, the 6-hr assimilation cycle used in similar studies (Liu et al., 2011; Ma et al., 2019, 2020; Pang et al., 2018) 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. In order to investigate the difference of different assimilation cycle on the analysis and forecast, an extra experiment assimilating lidar measurements at cycle of 6-hr (NP-LIDAR-6HR) has been performed in the Supplement (Table S1). The RMSE value of the NP-LIDAR and NP-LIDAR-6HR experiment is 0.16 1/km and 0.18 1/km, respectively (Fig. S1). The CORR values of these two experiments is 0.91 and 0.89 (Fig. S1). 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. Other detailed discussion can be found in the Supplement. It can be found that the statistic performance at cycle of 1-hr is better than that at cycle with 6-hr, which supports the setting of cycle of 1-hr in our work.”.

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

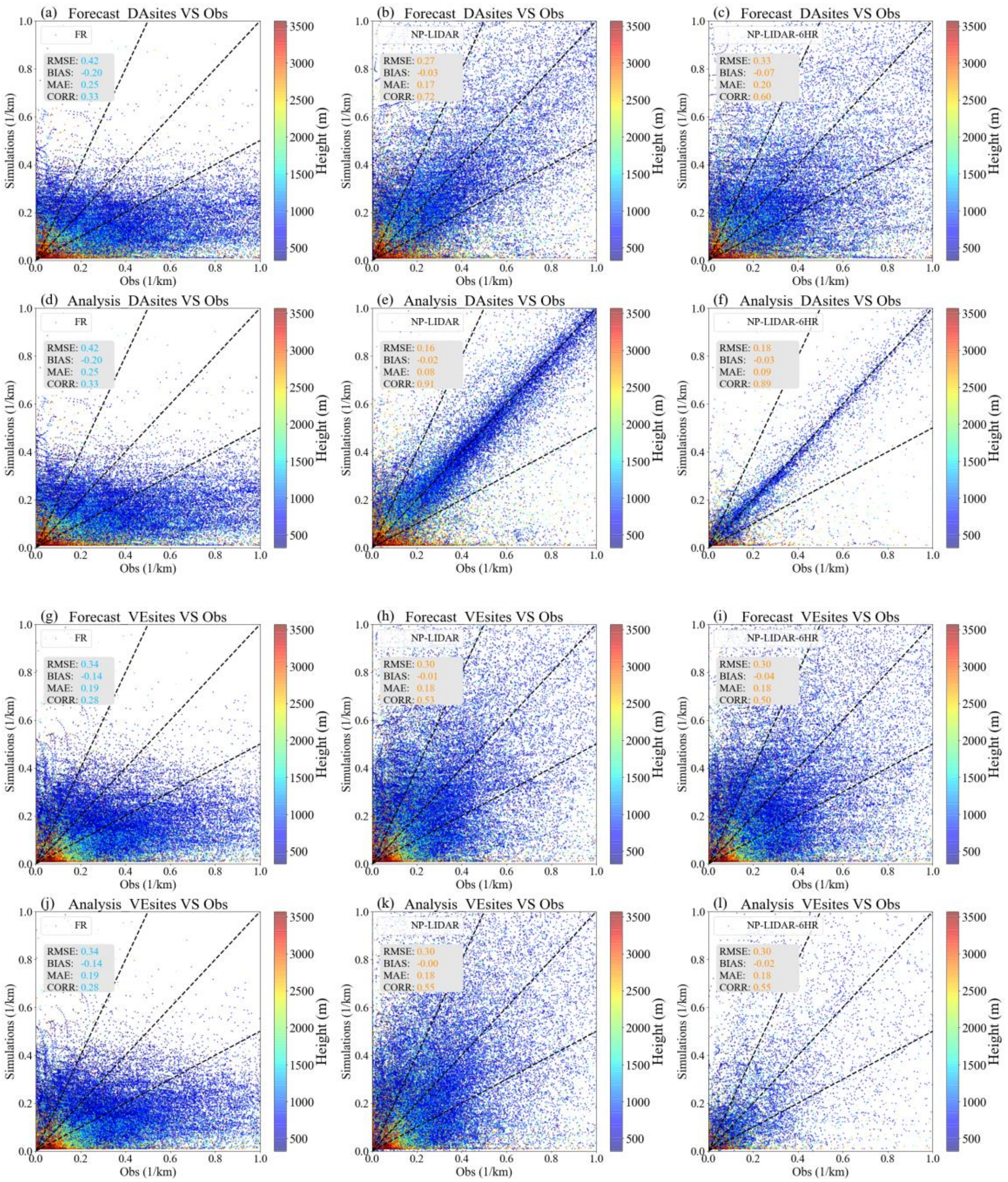

Figure S1. Scatter plots of the modeled hourly extinction coefficients at 550 nm versus the ground lidar hourly aerosol extinction coefficients at 532 nm (1/km) of forecasts of FR (a)/(g), forecasts of NP-LIDAR(b)/(h), forecasts of NP-LIDAR-6HR (c)/(i), analysis of FR (d)/(j), analysis of NP-LIDAR (e)/(k), 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.

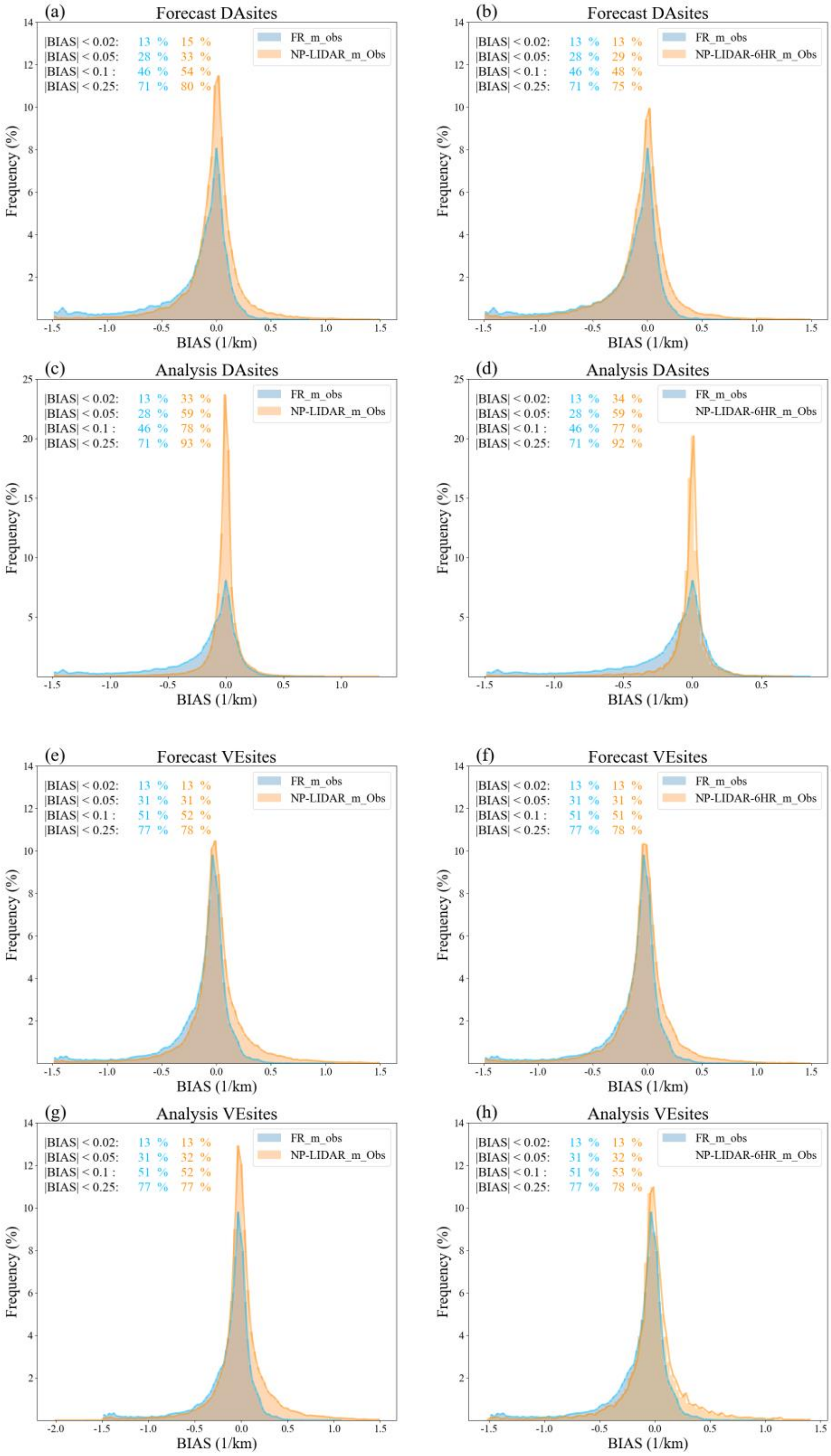

Figure S2. Frequency distributions of BIAS of forecasts of NP-LIDAR versus FR (a)/(e), forecasts of NP-LIDAR-6HR versus FR (b)/(f), analysis of NP-LIDAR versus FR (c)/(g) and analysis of NP-LIDAR-6HR versus FR (d)/(h), which are averaged among DA sites/VE sites.

Comment 2: The paper chooses ESTKF as the assimilation algorithm. What are the improvements or advantages of this Kalman filter algorithms compared to the previous KF algorithms?

Reply:

Firstly, the error-subspace transform Kalman filter (ESTKF, Nerger et al., 2012) is a recently developed Ensemble Kalman filter (EnKF, Evensen, 1994) variant. EnKF originated from the fusion of extended Kalman filter (EKF, Cohn, 1997) and Monte Carlo estimation methods. By providing flow- and location-dependent estimates of first-guess forecast error, the EnKF can potentially provide analysis and forecasts that are much more accurate than data assimilation schemes which assume that the background error does not vary in time (Whitaker and Hamill, 2002).

Secondly, EnKF and its variants can be categorized in deterministic ensemble filter, where the analysis is found through explicit mathematical transformations (SEIK, ETKF, ESTKF and so on), and stochastic ensemble filters, where perturbed forecasted observations are used (original EnKF). As one of ensemble square root filter algorithms, ESTKF is the former. On one hand, the deterministic ensemble filter can only use small ensemble sizes for high-dimensional problems, while stochastic filters need large ensemble sizes (Lawson and Hansen, 2004). On the other hand, the stochastic filters may add another source of sampling error and underestimate the analysis update because observations assimilated is perturbed.

Thirdly, ESTKF is derived from the singular evolutive interpolated Kalman filter (SEIK, Pham et al., 1998) by combining the advantages of the SEIK and the Ensemble Transform Kalman Filter (ETKF, Bishop et al., 2001). These three filters are essentially equivalent apart from computing the ensemble transformation in the error subspace (Vetra-Carvalho et al., 2018). The most significantly difference of ESTKF differs from SEIK and ETKF is that the error-subspace matrix is computed by

$$L = X^f \Omega, \tag{1}$$

where it is a projection matrix of size $N_e \times (N_e - 1)$ given by the set of equations as follows:

$$\Omega_{ij} = \begin{cases} 1-\dfrac{1}{N_e}\dfrac{1}{\dfrac{1}{\sqrt{N_e}}+1} & \text{for } i=j, i<N_e \\ -\dfrac{1}{N_e}\dfrac{1}{\dfrac{1}{\sqrt{N_e}}+1} & \text{for } i\neq j, i<N_e \\ -\dfrac{1}{\sqrt{N_e}} & \text{for } i=N_e. \end{cases} \tag{2}$$

where N_e is the number of ensemble members and $i=1, 2,..., N_e$.

The ESTKF can exhibit better properties than the SEIK filter, like a minimum ensemble transformation as the ETKF (Vetra-Carvalho et al., 2018). Nerger et al. (2012) conducted a series of numerical experiments to compare the performance of SEIK, ETKF and ESTKF using deterministic and random ensemble transformations. They found that the performance for the ESTKF and ETKF are better than SEIK filter with ESTKF having a slightly lower computational cost.

Changes in manuscript: Changes have been made in Line 252-266 of the revised manuscript and revised text is "The error subspace transform Kalman filter (ESTKF, Nerger et al., 2012) used in this study is a recently developed EnKF variant. Firstly, EnKF originated from the fusion of Kalman filter theory and Monte Carlo estimation method. By providing flow-dependent estimates of first-guess forecast error, the EnKF can potentially provide analysis and forecasts that are much more accurate than DA schemes which assume that the background error does not vary in time (Whitaker and Hamill, 2002). Secondly, EnKF and its variants can be categorized in deterministic filter (ETKF, ESTKF and so on) and stochastic filter which assimilates perturbed observations (original EnKF). On one hand, the deterministic filter can only use small ensemble sizes for high-dimensional problems, while stochastic filters need large ensemble sizes (Lawson and Hansen, 2004). On the other hand, the stochastic filters may add another source of sampling error and underestimate the analysis update because observations assimilated is perturbed. Thirdly, ESTKF is derived from the singular evolutive interpolated Kalman filter (SEIK, Pham et al., 1998) by combining the advantages of the SEIK and ETKF.These three filters are essentially equivalent apart from computing the ensemble transformation in the error subspace (Vetra-

Carvalho et al., 2018). The ESTKF can exhibit better properties than the SEIK filter, like a minimum ensemble transformation as the ETKF. Nerger et al. (2012) conducted a series of numerical experiments to compare the performance of SEIK, ETKF and ESTKF using deterministic and random ensemble transformations. They found that the performance for the ESTKF and ETKF are better than SEIK filter with ESTKF having a slightly lower computational cost. The ESTKF is outlined in this section." .

Comment 3: What does "Although the orbits are slightly covered by the model domain, the only difference between the FR and NP-LIDAR experiment is whether ground-based lidar measurements are assimilated (Fig. 12b)" mean? I do not understand this very well. What is the connection between these two sentences?

Reply: We agree with the comment. This sentence is really ambiguous and has been revised.

Changes in manuscript: Changes have been made in Line 572-573 of the revised manuscript and revised text is "Although the orbits are slightly covered by the model domain, the only difference of the averaged profiles between the FR and NP-LIDAR experiment is whether ground-based lidar measurements are assimilated (Fig. 12b)".

Comment 4: L605: Adding "measured by lidar" after "The aerosol vertical profile" for clarity.

Reply: Thanks, we agree with this comment.

Changes in manuscript: Changes have been made in Line 642 of the revised manuscript and revised text is "The aerosol vertical profile from lidar measurements averaged over VE sites shows a similar shape to that over DA sites".

Comment 5: L675: authors listed several reasons to explain that only assimilating lidar measurements has a weaker performance than only assimilating surface PM2.5 measurements. However, these reasons are just a guess without any detailed analysis. So, these reasons should not be listed in conclusion.

Reply: Thanks, we agree with this comment.

Changes in manuscript: We have deleted "This could be explained by the relatively sparser distribution of lidar sites compared with surface $PM_{2.5}$ measurements and the uncertainties in the spatial representation of lidar data, as well as the errors in the lumped variables of extinction coefficients with multiple contributions by different aerosol components. Moreover, the problem can also be attributed to the discordant relationship between aerosol mass concentration and extinction coefficients both in the simulation and measurements”. Please refer to the revised manuscript in Line 713-717.

Comment 6: L685: “a systematic data quality control of lidar measurements is urgently needed to solve this problem in future research” should be deleted. The reason is same as the above comment.

Reply: Thanks, we agree with this comment.

Changes in manuscript: We have deleted “A systematic data quality control of lidar measurements is urgently needed to solve this problem in future research”. Please refer to the revised manuscript in Line 726-727.

Comment 7: Fig. 7: The description of Fig. 7d is missing.

Reply: Thanks, it has been corrected in Line 1272.

Changes in manuscript: Changes have been made in Line 1272 of the revised manuscript and revised text is “Time series of prior RMSE and total spread over all observations for (a) extinction coefficients at 50 m, (b) extinction coefficients at 150 m, (c) extinction coefficients at 502 m, (d) extinction coefficients at 1000 m and (d) the surface $PM_{2.5}$”.

Comment 8: Fig. 12: “(e)” is missing.

Reply: Thanks, it has been added.

Changes in manuscript: Changes have been revised in Line 1312 of the revised manuscript and revised text is “05:00 UTC 18 April 2019 (e)”.

Comment 9: Fig. 15: “2021” should be “2019”.

Reply: Thanks, it has been corrected in Line 1328 of the revised manuscript.

Changes in manuscript: Changes have been made in Line 1328 of the revised manuscript and revised text is “All results are averaged over 1-30 April 2019”.

References:

Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects, 129, 17, 2001.

Cohn, S. E.: An introduction to estimation theory, J. Meteor. Soc. Jap., 75, 257–288, 1997.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10143, https://doi.org/10.1029/94JC00572, 1994.

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.

Lawson, W. G. and Hansen, J. A.: Implications of Stochastic and Deterministic Filters as Ensemble-Based Data Assimilation Methods in Varying Regimes of Error Growth, Mon. Wea. Rev., 132, 1966–1981, https://doi.org/10.1175/1520-0493(2004)132<1966:IOSADF>2.0.CO;2, 2004.

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., Janjić, T., Schröter, J., and Hiller, W.: A Unification of Ensemble Square Root Kalman Filters, Mon. Wea. Rev., 140, 2335–2345, https://doi.org/10.1175/MWR-D-11-00102.1, 2012.

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.

Pham, D. T., Verron, J., and Gourdeau, L.: Singular evolutive Kalman filters for data assimilation in oceanography, C. R. Acad. Sci. Ser. II, 326, 255–260, https://doi.org/10.1016/S1251-8050(97)86815-2, 1998.

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.

Vetra-Carvalho, S., van Leeuwen, P. J., Nerger, L., Barth, A., Altaf, M. U., Brasseur, P., Kirchgessner, P., and Beckers, J.-M.: State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems, Tellus A: Dynamic Meteorology and Oceanography, 70, 1–43, https://doi.org/10.1080/16000870.2018.1445364, 2018.

Whitaker, J. S. and Hamill, T. M.: Ensemble data assimilation without perturbed observations, Mon. Wea. Rev., 130, 1913–1924, https://doi.org/10.1175/MWR3156.1, 2002.

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.