Authors' Response to Reviews of

Valid time shifting Ensemble Kalman filter (VTS-EnKF) for dust storm forecasting

Mijie Pang, Jianbing Jin*, Arjo Segers, et al.
Geoscientific Model Development Discussions, 10.5194/gmd-2023-219

RC: ***Reviewers' Comment***, AR: Authors' Response, □ Manuscript Text

1. Overview

Response to Referee #1: We would like to thank the referee for the careful review throughout the paper and the in-depth comments that help to improve our paper.

2. Grave concerns

RC: ***Grave concern 1: Contrary to the authors' claims, the Neighbouring Time approach to aggregating ensemble members is not novel.***

The authors claim that their neighbouring time approach to increase ensemble sizes is novel. Unless I am missing some detail in their manuscript, that claim is incorrect. In fact, that method has been extensively tested. The common name for that method is "valid time shifting" (VTS). A similar, but extremely popular, variant of this method is "time-lagged ensembles". This concern is particularly grave for this study because this ensemble size increasing method is a huge part of this study's supposed novelty.

Here is a sampling of papers that employed those methods.

Gasperoni, N. A., X. Wang, and Y. Wang, 2023: Valid Time Shifting for an Experimental RRFS Convection-Allowing EnVar Data Assimilation and Forecast System: Description and Systematic Evaluation in Real Time. Mon. Wea. Rev., 151, 1229–1245, https://doi.org/10.1175/MWR-D-22-0089.1.

Huang, B., and X. Wang, 2018: On the Use of Cost-Effective Valid-Time-Shifting (VTS) Method to Increase Ensemble Size in the GFS Hybrid 4DEnVar System. Mon. Wea. Rev., 146, 2973–2998, https://doi.org/10.1175/MWR-D-18-0009.1.

Xu, Q., L. Wei, H. Lu, C. Qiu, and Q. Zhao, 2008: Time-expanded sampling for ensemble-based filters: Assimilation experiments with a shallow-water equation model. J. Geophys. Res., 113, D02114, https://doi.org/10.1029/2007JG000450.

Van den Dool, H. M., and L. Rukhovets, 1994: On the weights for an ensemble-averaged 6–10-day forecast. Wea. Forecasting, 9, 457–465, https://doi.org/10.1175/1520-0434(1994)009<0457:OTWFAE>2.0.CO;2.

To address this concern, please remove all claims that the Neighboring Time approach is novel in your manuscript.

AR: Thanks for the comment. We have removed all the improper statement in the paper. The corresponding papers

are cited. Details are below:

> In this paper, the standard EnKF assimilation is coupled with a valid time shifting (VTS) method (Xu et al., 2008; Lu et al., 2011; Zhao et al., 2015; Huang and Wang, 2018) for better resolving the position error in long-distance dust storm transport simulation. This assimilation methodology is referred to as VTS-EnKF throughout this paper. For assimilation analysis at a given time, the background error covariance of the simulated dust plume is calculated using not only the original ensemble simulation, but also the same ensemble simulations at neighboring moments (a few hours earlier and later) (Gasperoni et al., 2022, 2023). These extra ensemble members represent the potential position spread of the actual dust plume, effectively accounting for transport errors. The resampled ensemble members quantify the complex covariance that captures both intensity and position error dynamics, without requiring additional processing on observations, meteorological fields, or other physical parameters. We tested the VTS-EnKF on two severe dust storm events that occurred in 2021. Our results show superior assimilation performance compared to the standard EnKF, particularly when position errors are present in the simulated dust plume.

RC: ***Grave concern 2: Their EnKF's struggle with positioning error is highly contrived.***

The EnKF's struggle with positioning errors is likely simply due to their choice of meteorological forcing. Specifically, they failed to account for uncertainties in meteorological forcing. This could have been avoided by using the ECMWF's ensemble forecasts instead of the operational forecast. With an ensemble of forecasts, there should be more ensemble spread in the positioning of the dust storms, thus ameliorating the EnKF's issue with positioning error.

To address this concern, the authors need to rerun all of their experiments using the ECMWF's ensemble forecast data. This will likely take months of effort. The ECMWF has archived some of its ensemble forecasts on MARS. The ERA5's 10-member ensemble is also available through the Climate Data Store.

AR: Thanks for the in-depth comment. Now we have re-design our assimilation system: Each of our ensemble simulation (N=32) is driven by the perturbed emission and perturbed meteorology input. European Center for Medium-ranged Weather Forecast (ECMWF) ensemble forecast (totally 51 ensembles) are used, and 32 of them are randomly picked out for driving our ensemble simulation. It turns out that only trivial wilder ensemble spread found and the results are quite similar to the original EnKF experiments. We believe that the position error is mainly caused by the meteorology while ECMWF's ensemble forecast is not sufficient to account for the position error. By our method, the error can be alleviated.

Below are the new descriptions about the consideration of both emission and meteorology uncertainties:

> Another source of the uncertainties arises from the meteorological field. In our previous papers, uncertainties from meteorology and the position error were neither taken into account (Jin et al., 2022; Pang et al., 2023). In this paper, European Center for Medium-ranged Weather Forecast (ECMWF) ensemble forecast (totally 51 ensembles) are used. Each one of the model ensembles is driven by one unique ensemble meteorology field. 32 ensemble meteorological fields are randomly selected. Its grid resolution is about 14 km. The 6-hourly short-term meteorological forecast field is interpolated to hourly values. The grid resolutions are also averaged to fit the model resolution.
>
> In general, we assign the dust simulation uncertainty to both emission and meteorology. Ensemble emission field $[\boldsymbol{f}_1, ..., \boldsymbol{f}_N]$ are generated randomly following the emission uncertainty choice $\boldsymbol{f}_{\text{priori}}$ and $\mathbf{B}$ in Jin et al. (2022). Meteorologic field $[\boldsymbol{w}_1, ..., \boldsymbol{w}_N]$ are randomly selected from the total 51

ensemble meteorology. They are used to forward the LOTOS-EUROS model $\mathcal{M}$ for the ensemble dust simulations $[\boldsymbol{x}_1, ..., \boldsymbol{x}_\mathrm{N}]$ as:

$$[\boldsymbol{x}_1,\ \ldots,\boldsymbol{x}_\mathrm{N}]\ =\ [\mathcal{M}(\boldsymbol{f}_1,\boldsymbol{w}_1),\ \ldots,\ \mathcal{M}(\boldsymbol{f}_\mathrm{N},\boldsymbol{w}_\mathrm{N})]$$

N refers to the total ensemble number.

Below are the new results on assimilation analysis:

4.1 Impact on assimilation analysis

Figure 4 displays the spatial distribution of ground BR-PM_{10} observations (scatter) and dust field forecasts from the average of the ensembles (panel a.1), the posteriori from EnKF analysis (panel a.2) and EnKF with localization (panel a.3), the average of the enlarged ensembles (panel b.1), the posteriori from VTS-EnKF analysis (panel b.2) and VTS-EnKF analysis with localization (panel b.3) at 11:00, 15th March, 2021 China Standard Time (CST). It should be noted that the average dust concentrations in panel b.1 are calculated from the 160 ensemble simulations used in VTS-EnKF, which slightly differ from the average of 32 ensembles. In DSE1, the RMSE and NMB from the pure ensemble model simulation are as high as 856.36 μg m^{-3} and -78.31 %. Both EnKF and LEnKF assimilation analyses achieve very limited improvement in estimating the dust state field. As shown in panel a.2 and panel a.3, the RMSE and NMB remain high at 819.04 μg m^{-3} and -75.65 % in *EnKF*, and 782.57 μg m^{-3} and -73.52 % in *L500*. The main reason for this is the imbalanced uncertainty between the ensemble simulations and the observations, as described in Sect. 3.2. As observed in the light blue box in panel a.1, the simulated dust plume is located farther southeast compared to the PM_{10} measurements. This snapshot exhibits an apparent position error. After EnKF analysis, the simulated dust plume in the light blue box barely changes, as depicted in panel a.2. Numerous ground stations in this area report high PM_{10} concentrations, but the assimilated dust field fails to resolve most of them. The localization method offers limited assistance in this situation, as illustrated in panel a.3. With the unresolved positional error, the EnKF, which focuses more on intensity correction, is much less effective.

When it comes to the VTS-EnKF analysis result, an improved dust field can be noticed. Concerning the Root Mean Square Error (RMSE) and Normalized Mean Bias (NMB), the two priors depicted in panels a.1 and b.1 exhibit highly similar performances. However, slight differences do exist. For instance, the average of the expanded 160-member ensemble used in VTS-EnKF displays a marginally broader spread. The increased ensemble size provides more room for representing background uncertainties. The enhanced capacity for this is best illustrated in Fig. 6 (a), which exhibits the uncertainty quantified by the enlarged ensemble simulations in VTS-EnKF formulations. High uncertainty values are seen in pixels where large model-minus-observation errors are present, such as within the light blue box. This allows the posterior to be adjusted in order to better conform to the observations. In contrast, the relatively low uncertainty over these areas depicted in Fig. 1 (b.2) suggests that the EnKF method is highly confident in the absence of aerosols and does not require any modification. The observations are effectively assimilated in the VTS-EnKF analysis. As displayed in panel b.2, the dust plume within the light blue box is adjusted to better match the observations. In particular, the dust to the east of the marked region is well represented in comparison to the posteriori of *EnKF*. The RMSE and NMB are reduced to 742.33 μg m^{-3} and -68.21 %. Moreover, the posteriori of *VTS-L500* yields an improved dust field with the RMSE and NMB further reduced to 696.1 μg m^{-3} and -63.93 %. The implementation of the localization method eliminates spurious correlations and generates a background error covariance that more accurately describes the model uncertainties. Despite the noticeable improvements achieved

in DSE1, the residual errors, as indicated by the RMSE and NMB metrics, remain relatively high. This is mainly due to some observations with extremely high value (exceeding 5000 μg m^{-3}), which is far higher than the surrounding stations and hard for the EnKF to adapt. In particular, the western extent of the dust plume is covered by the insufficient stations, which results in an inadequate representation of the dust load. By incorporating neighboring ensembles, the dust plume is extended wilder, as can't be verified by the observations.

Figure 5 presents the spatial distribution of ground-based BR-PM$_{10}$ observations (scatter) and dust concentration forecasts from the average of model ensembles (panel a.1), EnKF (panel a.2), and LEnKF analysis (panel a.3), as well as the average of the enlarged model ensembles (panel b.1), VTS-EnKF (panel b.2), and VTS-EnKF with localization analysis (panel b.3) at 11:00, March 28th, 2021 CST. During this assimilation snapshot in DSE2, the model-simulated dust field is observed to have moved further southeast, as depicted in panel a.1. As illustrated by the light blue box in panel a.1, the model-simulated dust plume missed most of the observations with high PM$_{10}$ concentrations. Consequently, although the EnKF analysis remains effective in this case, dust in light blue box is nearly unchanged. The RMSE and NMB are reduced to 348.13 μg m^{-3} and -45.96 $\%$ in the *EnKF* scenario, with further reductions to 301.38 μg m^{-3} and -39.12 $\%$ when the localization method is employed in the *L500* case.

For the enlarged ensembles, the RMSE and NMB of the priori for *VTS-EnKF* are 433.08 μg m^{-3} and -8.93 $\%$. With VTS-EnKF assimilation, the RMSE of the posterior further decreases to 246.23 μg m^{-3}, and the NMB is -31.61 $\%$ in *VTS-EnKF*. Unlike the *EnKF*, the dust plume in light blue box is noticeably tuned to better fit the observations These error and bias values are significantly lower than those obtained with the *EnKF*, thanks to the better-scaled background covariance displayed in Fig. 6. Moreover, by incorporating localization, the RMSE and NMB are further reduced to 221.15 μg m^{-3} and -27.23 $\%$ in *VTS-L500*. The dust load within the light blue box (panel b.3) is accurately reproduced within its actual range (2000$\sim$3000 μg m^{-3}).

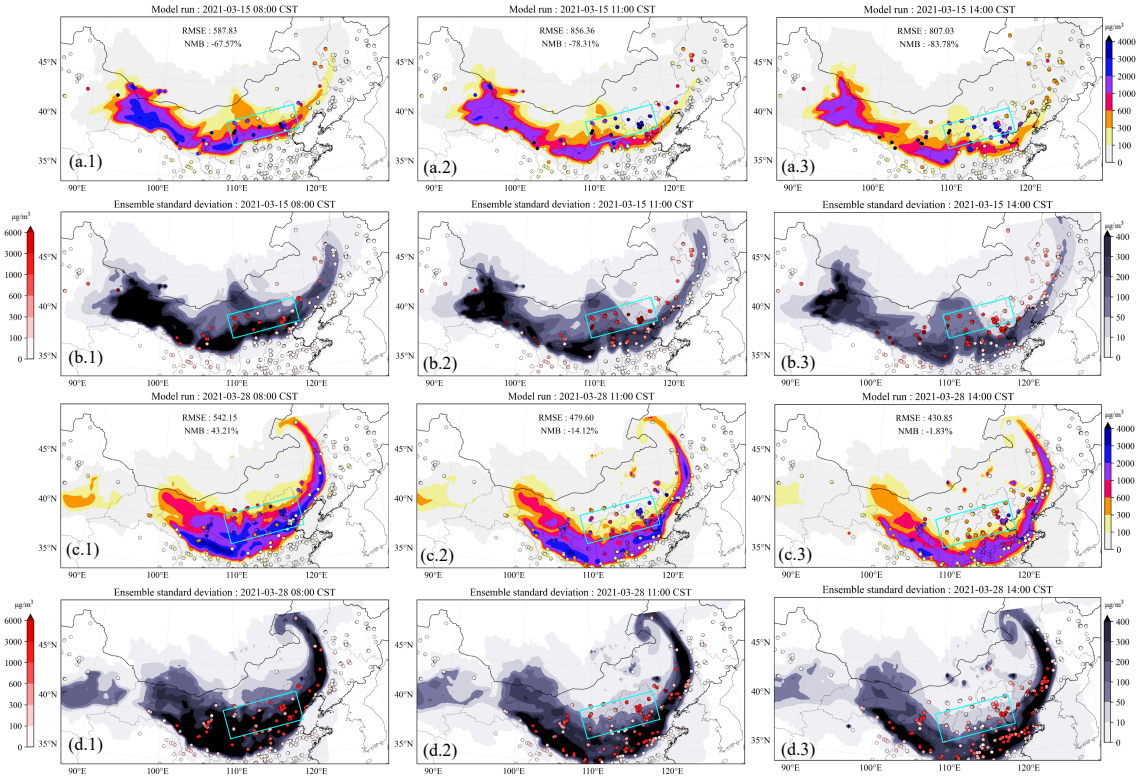

Figure 1. Evolution of the simulated dust plume from average of 32 model ensembles with scatter of ground BR-PM_{10} observations (**a.1-3**). Their corresponding standard deviation from model ensembles with scatter of the model-minus-observation differences (absolute value) (**b.1-3**) at 08:00, 11:00 and 14:00 15th March, 2021, respectively. Figures below are the same except the time is at 05:00 (**c.1** and **d.1**), 08:00 (**c.2** and **d.2**), 11:00 (**c.3** and **d.3**) 28th March, 2021, respectively. BR-PM_{10}: baseline-removed PM_{10}. The colorbar in panel **a** and **c** represents the concentrations, and the colorbar in panel **b** and **d** represents the model-minus-observation differences (left) and standard deviation (right).

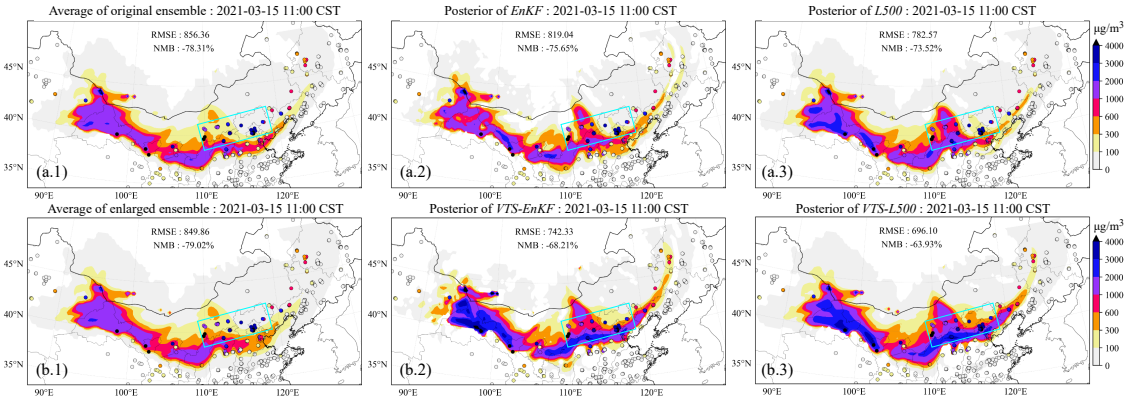

Figure 4. Spatial distribution of ground-based BR-PM_{10} observations (scatter) and simulated dust plume (SDP) on surface from central time ensemble model mean (**a.1**), the posteriori SDP updated by EnKF (**a.2**), the posteriori SDP updated by EnKF with localization (**a.3**), central and neighboring time ensemble model mean (**b.1**), the posteriori SDP updated by VTS-EnKF (**b.2**), the posteriori SDP updated by VTS-EnKF with localization (**b.3**) at 11:00, 15th March 2021 (CST).

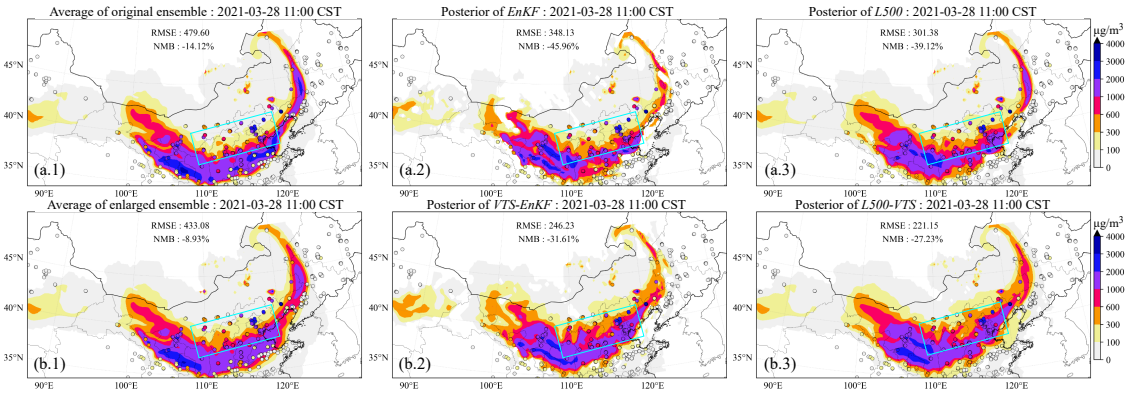

Figure 5. Spatial distribution of ground-based BR-PM_{10} observations (scatter) and simulated dust plume (SDP) on surface from central time ensemble model mean (**a.1**), the posteriori SDP updated by EnKF (**a.2**), the posteriori SDP updated by EnKF with localization (**a.3**), central and neighboring time ensemble model mean (**b.1**), the posteriori SDP updated by VTS-EnKF (**b.2**), the posteriori SDP updated by VTS-EnKF with localization (**b.3**) at 11:00, 28th March 2021 (CST).

RC: ***Grave concern 3: The authors did not satisfactorily demonstrate that the NTEnKF's improved performance over EnKF is purely due to NTEnKF's ability to handle positioning errors. The fact that the NTEnKF has less sampling error-related under-dispersion is likely playing a role.***

To address this concern, run an experiment with the NTEnKF that uses the same number of members as

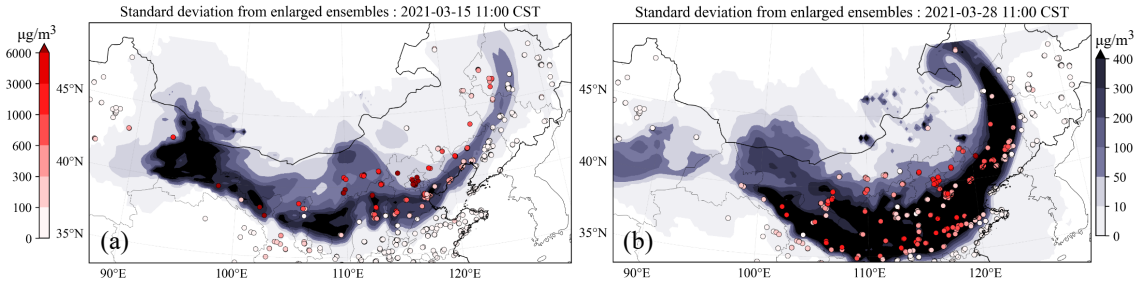

Figure 6. Spatial distribution of standard deviation from model ensembles with scatter of model-minus-observation differences (absolute value) at 11:00 in DSE1(**a**) and 08:00 in DSE2(**b**). The initial assimilation analysis is performed at these time. Colorbar left is for model-minus-observation differences and right is for standard deviation.

the EnKF (i.e., run a 32-member NTEnKF experiment). I suspect that the NTEnKF's performance will be comparable to the EnKF's in such a situation. Remember to use the ECMWF ensemble forecasts as your meteorological forcings.

AR: Agree with the referee. We now run experiments with the VTS-EnKF with same ensemble number as EnKF, which is 32. We separated the original 32 ensemble into [6,6,8,6,6] and choose the neighbouring time following the original time set (-2, -1, 0, +1, +2 hour). We found that by this limited ensemble number, VTS-EnKF provides slightly better performance than the EnKF. The sampling error indeed plays a role under the smaller ensembles. However, by applying the localization, this error can be noticeably reduced. It proves that our VTS-EnKF's superiority in handling the position error.

4.2 Assessment of smaller ensembles

To further assess the performance of VTS-EnKF, VTS-EnKF experiments with same ensembles as the EnKF are designed. They are referred to as *VTS-EnKF-small* and *VTS-L500-small*, respectively. The total 32 ensembles are composed of 8 central ensembles and 4×6 ensembles from neighboring ±1 and ±2 hours. Figure 9 displays the time series of RMSE and NMB on a 24-hour dust forecast after three assimilation analyses in DSE1. In terms of RMSE, *VTS-EnKF-small* only shows slightly better performance than the EnKF. This mostly caused by the sampling error arises from limited ensembles resampled from the central ensembles (only 8 ensembles). However, by applying the localization, the RMSE is noticeably reduced by 100 $\mu g\ m^{-3}$. The performance is comparable to the *VTS-L500* (red dash line) with totally 160 ensembles. By mitigating the sampling error, the VTS-EnKF's capability of handling the position error can be revealed, which can be noticed by comparison with *L500* and *VTS-L500-small*. This improvement can be better seen in NMB. NMB of *VTS-L500-small* is much lower than the *EnKF* and *L500*. Its performance is also comparable to the *VTS-L500* with 160 ensembles.

Same experiments on DSE2 are also carried out. Results can be found in Fig. S2 in supporting information. Similar to DSE1, the *VTS-EnKF-small* achieves slightly better RMSE and NMB than *EnKF* and *L500*. While in *VTS-L500-small*, noticeable improvements can been found especially for the forecast after the second and last assimilation. Reduction of 100 $\mu g\ m^{-3}$ in RMSE and 20% in NMB are obtained.

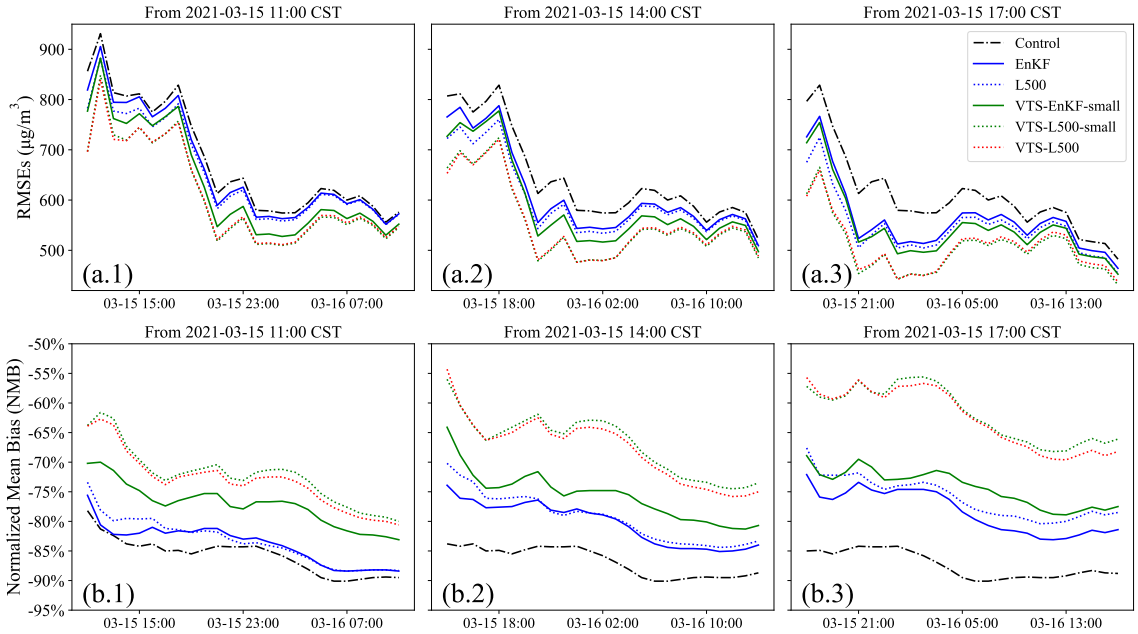

Figure 9. Time series of 24-hour Root Mean Square Error (RMSE) on the dust forecast starting from 11:00 (**a.1**), 14:00 (**a.2**), 17:00 (**a.3**) and normalized mean bias (NMB) starting from 11:00 (**b.1**), 14:00 (**b.2**), 17:00 (**b.3**) on 15th March 2021.

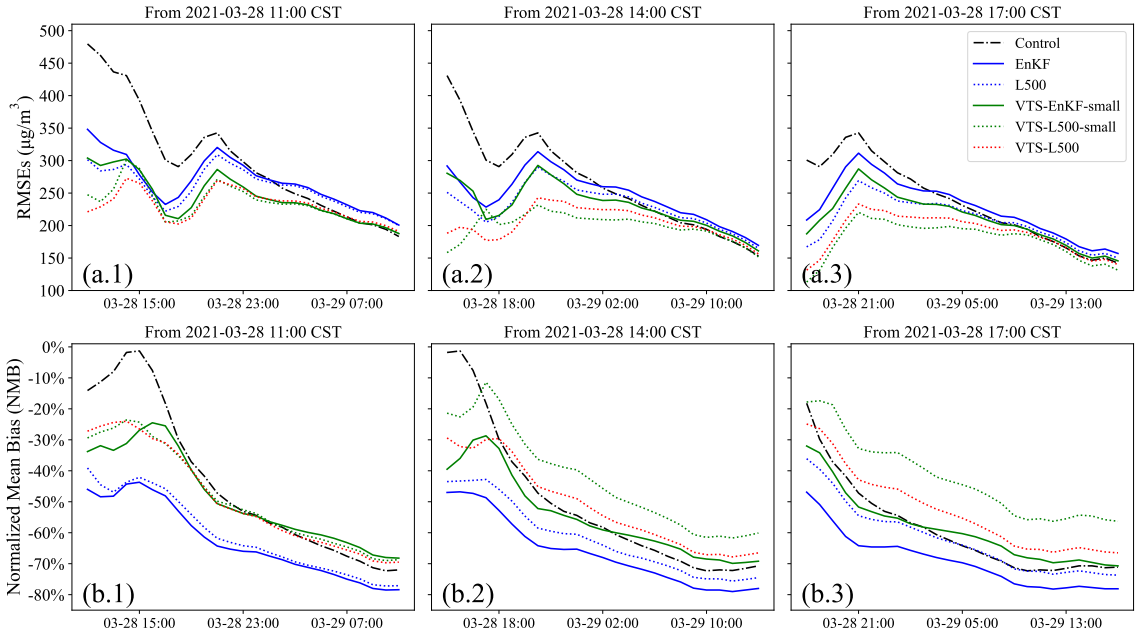

Figure S2. Time series of 24-hour Root Mean Square Error (RMSE) on the dust forecast starting from 11:00 (**a.1**), 14:00 (**a.2**), 17:00 (**a.3**) and normalized mean bias (NMB) starting from 11:00 (**b.1**), 14:00 (**b.2**), 17:00 (**b.3**) on 28th March 2021.

On the other hand, we have also tested the standard EnKF with more ensemble members, but very limited improvements were obtained.

> In EnKF-based experiments, *EnKF* and *L500*, the ensemble number N is set to 32, which is found to be sufficient to represent the uncertainty in the dust simulation while remaining computationally affordable. Testing with N greater than 32 shows only limited improvements.

3. Major concern

RC: ***The authors did not explore the statistical problems that surround the use of Neighbourhood time ensembles. The primary issues are***

the ensemble members are correlated with each other (i.e., the ensemble is no longer i.i.d.), causing the estimated ensemble variance to be biased from the true forecast variance, and, the ensemble becomes non-Gaussian, especially if time points far apart are used, strengthening the possibility that the EnKF creates suboptimal biased analyses

AR: The choice of neighbouring time ensembles definitely affects the analysis. An improper choice of interval can lead to undesirable analysis, such as less-effective ensemble members (interval too small) or unrepresentative ensemble covariances (interval too large). We didn't explore the choice of neighbouring time interval in our original manuscript. Here, we design new experiments that use different time intervals to tell the impact of the choice of neighbouring time. Below are the new experiment settings and results:

Table 1. Experiment settings.

Name	Running ensemble number	Initial assimilation time set (hour)	Ensemble set	Localization distance (km)
Control	32	None	[32]	None
EnKF	32	t	[32]	None
L500	32	t	[32]	500
VTS-EnKF	160	$t-2, t-1, t, t+1, t+2$	[32,32,32,32,32]	None
VTS-L500	160	$t-2, t-1, t, t+1, t+2$	[32,32,32,32,32]	500
VTS-EnKF-small	32	$t-2, t-1, t, t+1, t+2$	[6,6,8,6,6]	None
VTS-L500-small	32	$t-2, t-1, t, t+1, t+2$	[6,6,8,6,6]	500
VTS-EnKF-t1	96	$t-1, t, t+1$	[32,32,32]	None
VTS-EnKF-t2	96	$t-2, t, t+2$	[32,32,32]	None
VTS-EnKF-t3	96	$t-3, t, t+3$	[32,32,32]	None
VTS-EnKF-t4	96	$t-4, t, t+4$	[32,32,32]	None
VTS-EnKF-t5	96	$t-5, t, t+5$	[32,32,32]	None
VTS-EnKF-t6	96	$t-6, t, t+6$	[32,32,32]	None

4.4 Sensitivity of time interval

Previous researches have found that an improper neighboring time interval τ can lead to undesirable results, such as less-effective ensemble members (interval too small) (τ too small) or ensemble member clustering and unrepresentative ensemble covariances (τ too large) (Xu et al., 2008; Gasperoni et al., 2022, 2023). To explore the sensitivity of the choice of neighboring time interval, series of VTS-EnKF experiments with different neighboring time interval were carries out. Time intervals ranging from 1 to 6 hour were tested. As shown in Fig. 10, snapshots from 6 experiments on DSE1 clearly depicts the trend. In general, all the VTS-EnKF experiments show better performance than EnKF. While in terms of specific time interval, different patterns can be noticed. For short intervals including 1 and 2 hour, there is not sufficient ensemble spread to account for the position error. Thus there are still position error remaining and RMSE is still high. For long intervals including 5 and 6 hour, dust plume is clustered away from central dust plume. Three dust branches are noticed in *VTS-EnKF-t5* and an overly backwards dust plume is noticed in *VTS-EnKF-t6*. In this case, 3-hour interval is the best choice with the lowest RMSE (696.11 $\mu g\ m^{-3}$) and NMB (-63.5 $\%$).

Same experiments on DSE2 are also performed and snapshots are shown in Fig. S3. Similar patterns are found on DSE2. Lowest RMSE and NMB are achieved in *VTS-EnKF-t4*. Too short interval leads to inability in position error correction and too long interval leads to excessive dust plume. Considering both cases, 3-hour interval is the preferred choice which holds the capability to handle position and not creates excessive clustered dust plume.

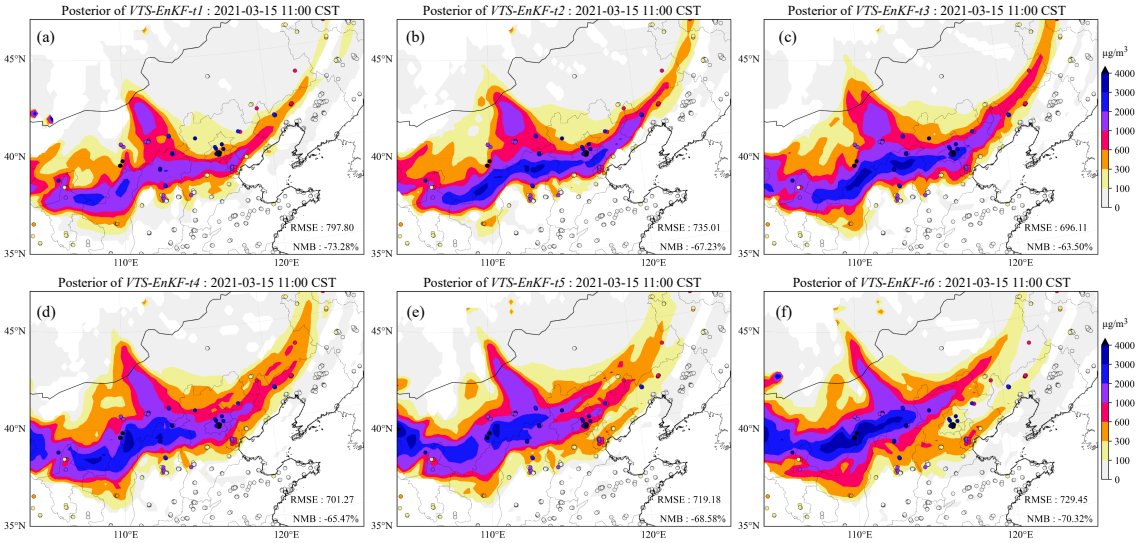

Figure 10. Spatial distribution of ground-based BR-PM_{10} observations (scatter) and simulated dust plume (SDP) on surface from the posteriori SDP updated by VTS-EnKF-t1 (**a**), the posteriori SDP updated by VTS-EnKF-t2 (**b**), the posteriori SDP updated by VTS-EnKF-t3 (**c**), the posteriori SDP updated by VTS-EnKF-t4 (**d**), the posteriori SDP updated by VTS-EnKF-t5 (**e**), the posteriori SDP updated by VTS-EnKF-t6 (**f**) at 11:00, 15th March 2021 (CST).

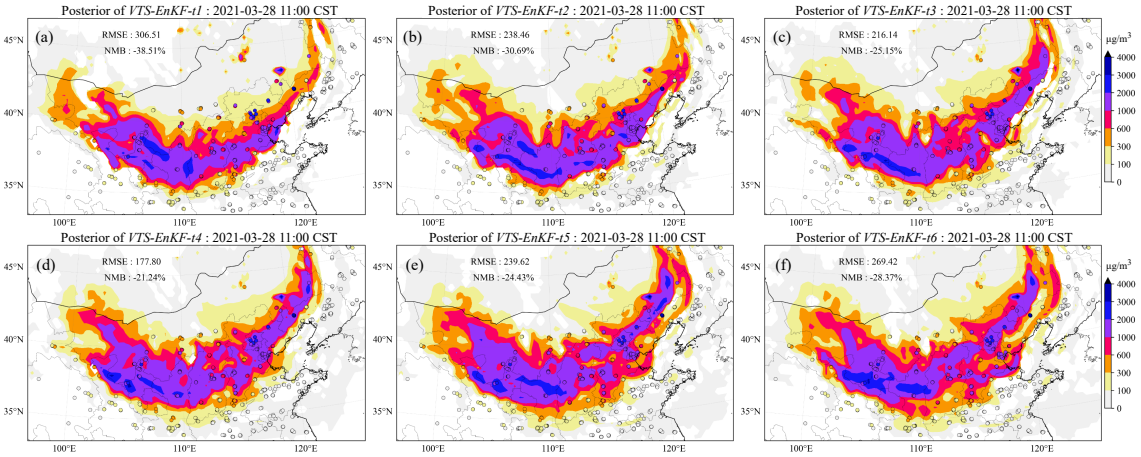

Figure S3. Spatial distribution of ground-based BR-PM_{10} observations (scatter) and simulated dust plume (SDP) on surface from the posteriori SDP updated by VTS-EnKF-t1 (**a**), the posteriori SDP updated by VTS-EnKF-t2 (**b**), the posteriori SDP updated by VTS-EnKF-t3 (**c**), the posteriori SDP updated by VTS-EnKF-t4 (**d**), the posteriori SDP updated by VTS-EnKF-t5 (**e**), the posteriori SDP updated by VTS-EnKF-t6 (**f**) at 11:00, 28th March 2021 (CST).

4. Minor comments

RC: ***1) The authors' writing seem to imply that the Pf matrix does not normally account for position errors. That is incorrect. The Pf matrix accounts for both intensity and position uncertainties if the forecast ensemble has both kinds of uncertainty. However, note that the Pf matrix only adequately represents position uncertainties if it is sufficiently small – position uncertainties result in non-Gaussian statistics if those uncertainties are large.***

AR: Thanks for the comment. We didn't clearly explain the constitutions of the $\mathbf{P}^f$. We agree that the $\mathbf{P}^f$ accounts for both intensity and position errors. While when there are significant position errors, non-Gaussian statistics can be aggravated. Which will mislead EnKF that relies on Gaussian distribution of errors. Descriptions are made in Line 490-493. Details are below:

> The background error covariance of EnKF is generally designed to represent the intensity and position uncertainty. However, when the position error is sufficiently large, the background error covariance can't adequately represent the position error, which is highly non-Gaussian. In the case of the long-distance dust storm tracking, the EnKF is incapable of thoroughly resolving the observations. Observations over low model uncertainty pixels are 'ignored' by the EnKF algorithm.

RC: ***2) Line 64: Please acknowledge that the EnKF is suboptimal for non-Gaussian problems. Though the EnKF can be employed in such situations, the EnKF is probably injecting some kind of bias because it is designed specifically for Gaussian problems.***

AR: Thanks for your comment. We agree that EnKF is Gaussian-dependent. It is added in Line 66-68.

> Meanwhile, inherited from Kalman filter, EnKF relies on Gaussian distribution of error statistics (Amezcua and Van Leeuwen, 2014). For non-Gaussian problems, EnKF can create suboptimal results (Lei et al., 2010).

RC: ***3) Given the centrality of the EnKF to the paper, it seems unusual that only 3 papers are cited between lines 60-67. In particular, the sentences in lines 66 and 67 are missing supporting references. Here's a good review paper about the EnKF that you can use to find more EnKF references: Houtekamer, P. L., and F. Zhang, 2016: 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.***

Also, the stochastic EnKF scheme you are using is not the one that Geir Evensen formulated. It is the one Burgers formulated. Here's the paper: Burgers, G., P. Jan van Leeuwen, and G. Evensen, 1998: Analysis Scheme in the Ensemble Kalman Filter. Mon. Wea. Rev., 126, 1719–1724, https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Peter Jan van Leeuwen of Colorado State University (Evensen's good friend), recently published a much more satisfactory explanation of the stochastic EnKF than Burgers et al (1998): van Leeuwen PJ. A consistent interpretation of the stochastic version of the Ensemble Kalman Filter. QJR Meteorol Soc. 2020; 146: 2815–2825. https://doi.org/10.1002/qj.3819

AR: Thanks for these strong supporting reference. We have added more references when describing the EnKF in Line 58-68:

Filtering methods, on the other hand, assimilate observations sequentially and are more efficient for operational forecasting systems. Various filtering approaches, such as Kalman Filter (Kalman, 1960), Extended Kalman Filter (Brunner et al., 2012), and Particle Filter (Leeuwen et al., 2019), have been developed. Among all the filtering methods, the Ensemble Kalman Filter (EnKF) is the most popular filtering method due to its ability to handle high-dimensional models, easy parallelization (Evensen, 1994; Katzfuss et al., 2016; Houtekamer and Zhang, 2016). It uses limited ensembles to estimate the background error covariance statistics of the model (Hamill, 2006; Houtekamer et al., 2014). Its advantages include handling non-linearity, not requiring explicit calculation of tangent linear operators, and computational efficiency (Bannister, 2017). EnKF has been successfully applied in various disciplines, e.g.,weather forecasting (Houtekamer et al., 2005) and hydrology (Reichle et al., 2002). Meanwhile, inherited from Kalman filter, EnKF relies on Gaussian distribution of error statistics (Amezcua and Van Leeuwen, 2014). For non-Gaussian problems, EnKF can create suboptimal results (Lei et al., 2010).

As to the Peter Jan van Leeuwen's new explanation of the stochastic EnKF, we are quite interested in this interpretation and will examine the impact in the future. Thanks for the recommendation!

RC: ***4) Eq. 9 – The notation can be mistaken as summing up matrices containing the ensemble members at different time points. Please find another way to mathematically express the idea that you are concatenating ensembles across time. Perhaps you can refer to the valid time shifting papers that I referenced earlier.***

AR: Thanks for the comment. We have used better expression of the idea. Details are below:

$$\mathbf{X}^{f,new} = [\boldsymbol{x}_{t-\tau}^{f,1}, \boldsymbol{x}_{t-\tau}^{f,2}, \cdots, \boldsymbol{x}_{t-\tau}^{f,\mathrm{N}}, \boldsymbol{x}_{t}^{f,1}, \boldsymbol{x}_{t}^{f,2}, \cdots, \boldsymbol{x}_{t}^{f,\mathrm{N}}, \boldsymbol{x}_{t+\tau}^{f,1}, \boldsymbol{x}_{t+\tau}^{f,2}, \cdots, \boldsymbol{x}_{t+\tau}^{f,\mathrm{N}}]$$

References

Amezcua, J. and Van Leeuwen, P. J.: Gaussian Anamorphosis in the Analysis Step of the EnKF: A Joint State-Variable/Observation Approach, TELLUS A, 66, 23 493, , 2014.

Bannister, R. N.: A Review of Operational Methods of Variational and Ensemble-Variational Data Assimilation, Q. J. R. Meteorolog. Soc., 143, 607–633, , 2017.

Brunner, D., Henne, S., Keller, C. A., Reimann, S., Vollmer, M. K., O'Doherty, S., and Maione, M.: An Extended Kalman-filter for Regional Scale Inverse Emission Estimation, Atmos. Chem. Phys., 12, 3455–3478, , 2012.

Evensen, G.: Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model Using Monte Carlo Methods to Forecast Error Statistics, J. Geophys. Res., 99, 10 143, , 1994.

Gasperoni, N. A., Wang, X., and Wang, Y.: Using a Cost-Effective Approach to Increase Background Ensemble Member Size within the GSI-Based EnVar System for Improved Radar Analyses and Forecasts of Convective Systems, Mon. Weather Rev., 150, 667–689, , 2022.

Gasperoni, N. A., Wang, X., and Wang, Y.: Valid Time Shifting for an Experimental RRFS Convection-Allowing EnVar Data Assimilation and Forecast System: Description and Systematic Evaluation in Real Time, Mon. Weather Rev., 151, 1229–1245, , 2023.

Hamill, T. M.: Ensemble-Based Atmospheric Data Assimilation, in: Predictability of Weather and Climate, edited by Palmer, T. and Hagedorn, R., pp. 124–156, Cambridge University Press, 1 edn., , 2006.

Houtekamer, P. L. and Zhang, F.: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation, Mon. Weather Rev., 144, 4489–4532, , 2016.

Houtekamer, P. L., Mitchell, H. L., Pellerin, G., Buehner, M., Charron, M., Spacek, L., and Hansen, B.: Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations, Mon. Weather Rev., 133, 604–620, , 2005.

Houtekamer, P. L., Deng, X., Mitchell, H. L., Baek, S.-J., and Gagnon, N.: Higher Resolution in an Operational Ensemble Kalman Filter, Mon. Weather Rev., 142, 1143–1162, , 2014.

Huang, B. and Wang, X.: On the Use of Cost-Effective Valid-Time-Shifting (VTS) Method to Increase Ensemble Size in the GFS Hybrid 4DEnVar System, Mon. Weather Rev., 146, 2973–2998, , 2018.

Jin, J., Pang, M., Segers, A., Han, W., Fang, L., Li, B., Feng, H., Lin, H. X., and Liao, H.: Inverse Modeling of the 2021 Spring Super Dust Storms in East Asia, Atmos. Chem. Phys., 22, 6393–6410, , 2022.

Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, J. Basic Eng., 82, 35–45, , 1960.

Katzfuss, M., Stroud, J. R., and Wikle, C. K.: Understanding the Ensemble Kalman Filter, The American Statistician, 70, 350–357, , 2016.

Leeuwen, P. J., Künsch, H. R., Nerger, L., Potthast, R., and Reich, S.: Particle Filters for High-dimensional Geoscience Applications: A Review, Q. J. R. Meteorolog. Soc., 145, 2335–2365, , 2019.

Lei, J., Bickel, P., and Snyder, C.: Comparison of Ensemble Kalman Filters under Non-Gaussianity, Mon. Weather Rev., 138, 1293–1306, , 2010.

Lu, H., Xu, Q., Yao, M., and Gao, S.: Time-Expanded Sampling for Ensemble-Based Filters: Assimilation Experiments with Real Radar Observations, Adv. Atmos. Sci., 28, 743–757, , 2011.

Pang, M., Jin, J., Segers, A., Jiang, H., Fang, L., Lin, H. X., and Liao, H.: Dust Storm Forecasting through Coupling LOTOS-EUROS with Localized Ensemble Kalman Filter, Atmos. Environ., 306, 119 831, , 2023.

Reichle, R. H., McLaughlin, D. B., and Entekhabi, D.: Hydrologic Data Assimilation with the Ensemble Kalman Filter, Mon. Weather Rev., 130, 103–114, , 2002.

Xu, Q., Wei, L., Lu, H., Qiu, C., and Zhao, Q.: Time-Expanded Sampling for Ensemble-Based Filters: Assimilation Experiments with a Shallow-Water Equation Model, J. Geophys. Res., 113, , 2008.

Zhao, Q., Xu, Q., Jin, Y., McLay, J., and Reynolds, C.: Time-Expanded Sampling for Ensemble-Based Data Assimilation Applied to Conventional and Satellite Observations, Weather Forecasting, 30, 855–872, , 2015.