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 #2: 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. Comments

RC: ***1) As the other referee pointed out, similar techniques have been proposed before as the "valid time shifting" in other applications. The authors should reconsider their novelty about the method and refer to the proper papers.***

AR: Thanks for the comment. We have removed all the improper statement in the paper. The corresponding papers are cited. Details are below: In Line 114-124:

> 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: ***2) The authors have confined their consideration to the issue of dust emission uncertainty in the EnKF's P matrix, as detailed in Section 3.1. However, given that the introduction highlights the emergence of position errors as a result of meteorological input, Why not incorporate meteorological uncertainty within the process of ensemble generation?***

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. 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 $[\boldsymbol{f}_1, ..., \boldsymbol{f}_\mathrm{N}]$ are generated randomly following the emission uncertainty choice $\boldsymbol{f}_{\text{priori}}$ and $\mathbf{B}$ in Jin et al. (2022). Meteorologic field $[\boldsymbol{w}_1, ..., \boldsymbol{w}_\mathrm{N}]$ 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:
>
> $$[\boldsymbol{x}_1, \ ... \ , \boldsymbol{x}_\mathrm{N}] \ = \ [\mathcal{M}(\boldsymbol{f}_1, \boldsymbol{w}_1), \ ... \ , \ \mathcal{M}(\boldsymbol{f}_\mathrm{N}, \boldsymbol{w}_\mathrm{N})]$$
>
> 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}).

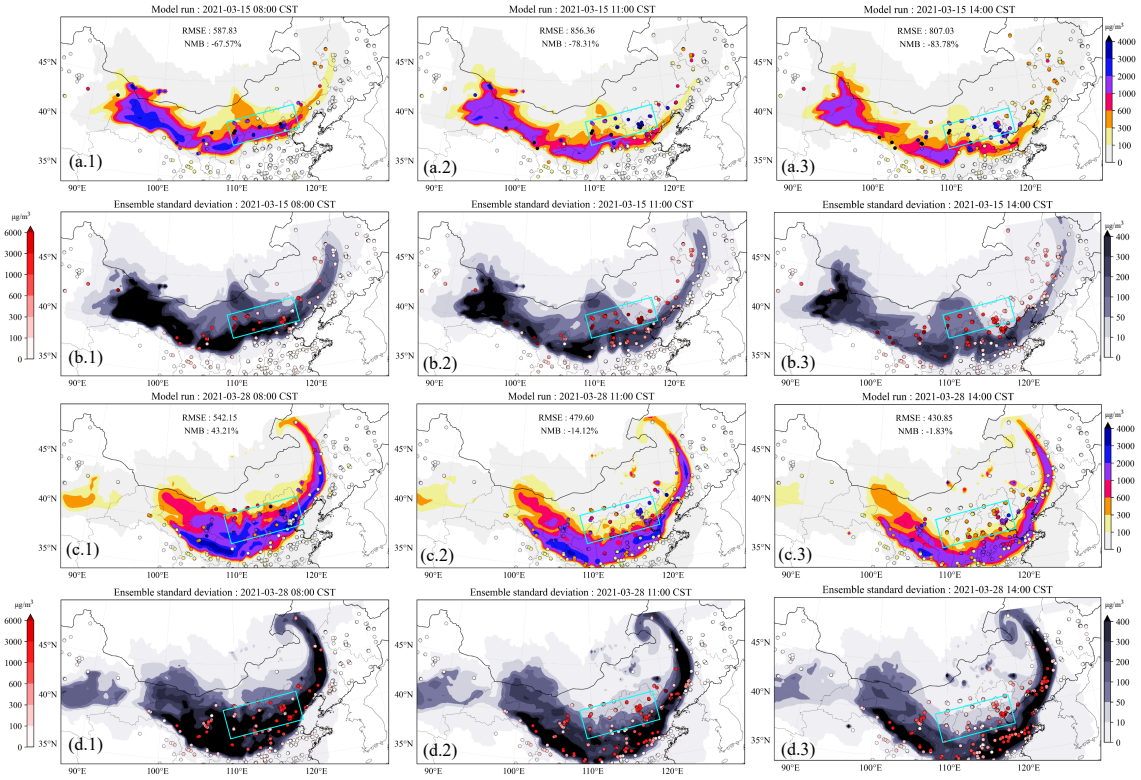

Figure 1. Evolution of the simulated dust plume from average of 32 model ensembles with scatter of ground BR-PM_{10} observations (**a.1-3**). Their corresponding standard deviation from model ensembles with scatter of the model-minus-observation differences (absolute value) (**b.1-3**) at 08:00, 11:00 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).

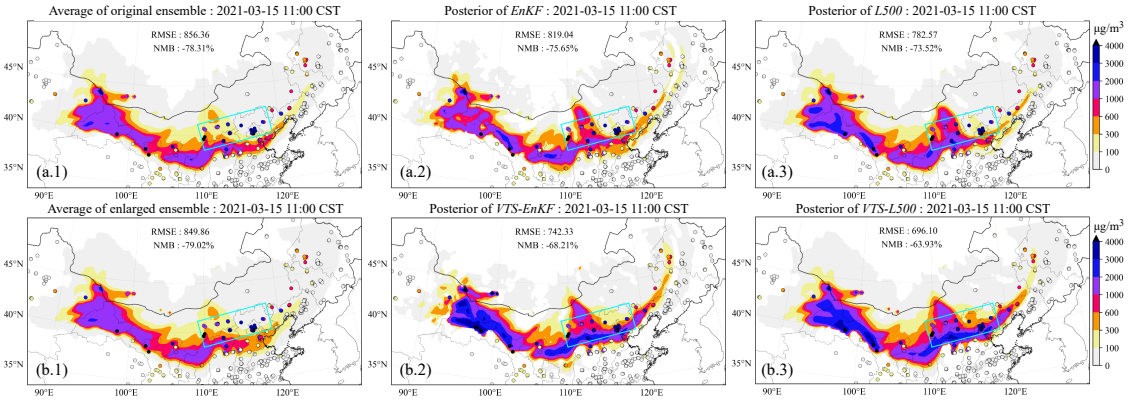

Figure 4. Spatial distribution of ground-based BR-PM_{10} observations (scatter) and simulated dust plume (SDP) on surface from central time ensemble model mean (**a.1**), the posteriori SDP updated by EnKF (**a.2**), the posteriori SDP updated by EnKF with localization (**a.3**), central and neighboring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

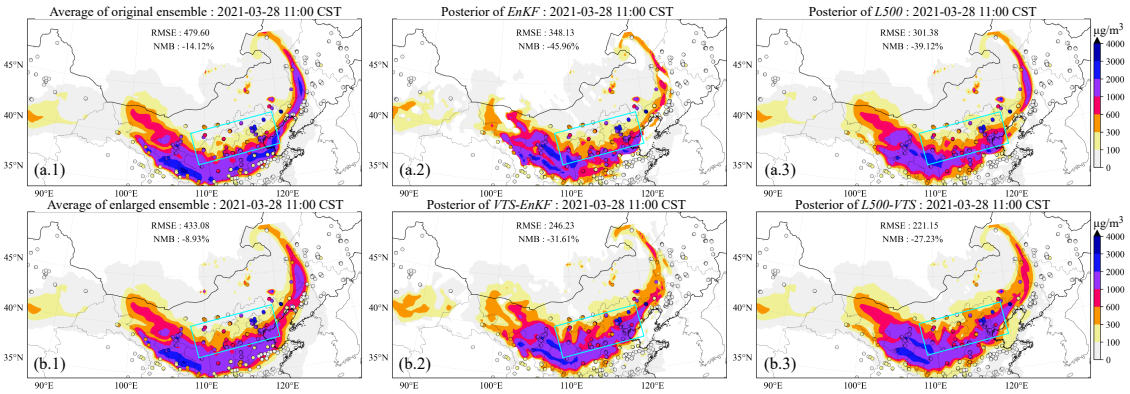

Figure 5. Spatial distribution of ground-based BR-PM_{10} observations (scatter) and simulated dust plume (SDP) on surface from central time ensemble model mean (**a.1**), the posteriori SDP updated by EnKF (**a.2**), the posteriori SDP updated by EnKF with localization (**a.3**), central and neighboring 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: ***3) The decision to merge information from five distinct time points, centering around the central time, is mentioned, yet the rationale behind selecting these specific time points for combination is not fully explained. Could you please elaborate on the relevance of this choice?***

AR: Thanks for the comment. We agree that the choice of neighbouring time ensembles can affect the analysis.

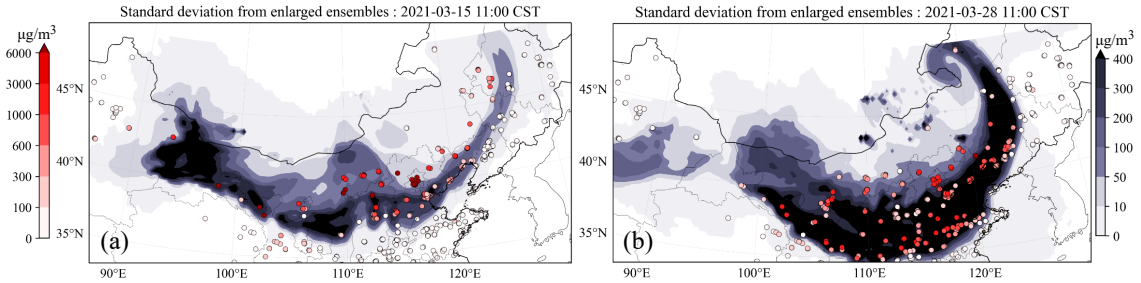

Figure 6. Spatial distribution of standard deviation from model ensembles with scatter of model-minus-observation differences (absolute value) at 11:00 in DSE1(**a**) and 08:00 in DSE2(**b**). The initial assimilation analysis is performed at these time. Colorbar left is for model-minus-observation differences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

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. The experiment settings and new section are added:

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.

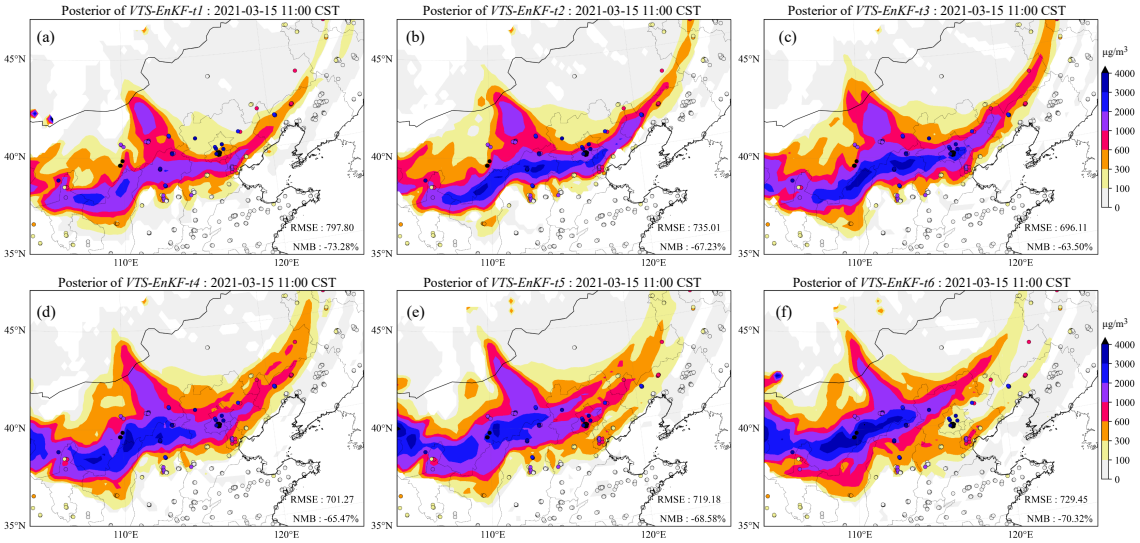

Figure 10. Spatial distribution of ground-based BR-PM_{10} observations (scatter) and simulated dust plume (SDP) on surface from the posteriori SDP updated by VTS-EnKF-t1 (**a**), the posteriori SDP updated by VTS-EnKF-t2 (**b**), the posteriori SDP updated by VTS-EnKF-t3 (**c**), the posteriori SDP 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).

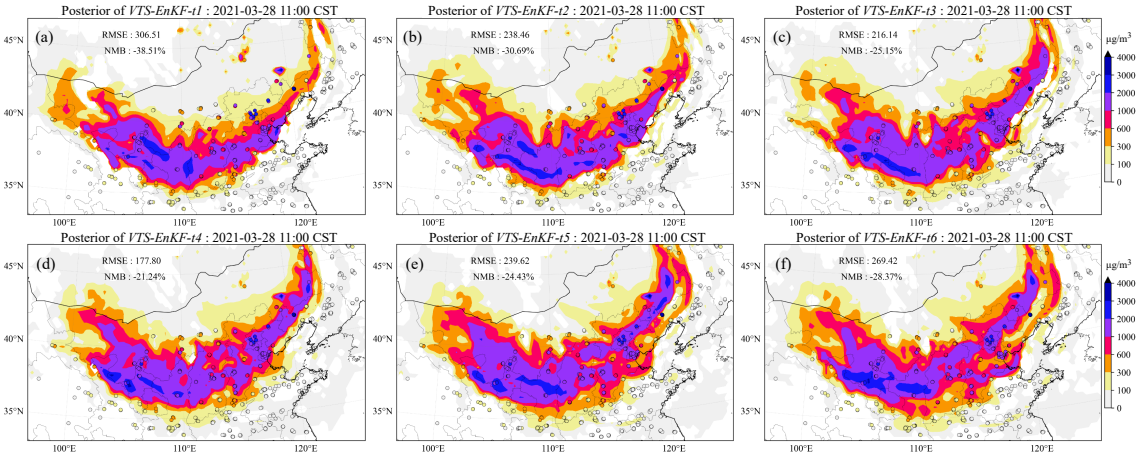

Figure S3. Spatial distribution of ground-based BR-PM_{10} observations (scatter) and simulated dust plume (SDP) on surface from the posteriori SDP updated by VTS-EnKF-t1 (**a**), the posteriori SDP updated by VTS-EnKF-t2 (**b**), the posteriori SDP updated by VTS-EnKF-t3 (**c**), the posteriori SDP 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).

RC: ***4) It would be helpful to clarify the methodology for detecting the occurrence of position error, especially in light of the rapid evolution of dust storms. Is it automatically detected or manually chosen? What criteria do the authors employ to determine the appropriate timing for implementing the NTEnKF?***

AR: Thanks for the comment. In this paper, we decide the timing manually. To better identify the emergence of position error, a simple identification index is designed. Descriptions are made in Supplementary. Details are shown below:

2. Identification of position error

To objectively identify the position error, a simple identification index is designed, which is the interquartile range of $\boldsymbol{\mathcal{H}x} - \boldsymbol{y}$. This index is often used to describe the spread of the data. Here, it depicts the error statistics transiting from Gaussian to non-Gaussian distribution with emergence of position error:

$$\mathrm{IQR} = Q_3(\boldsymbol{\mathcal{H}x} - \boldsymbol{y}) - Q_1(\boldsymbol{\mathcal{H}x} - \boldsymbol{y}) \tag{1}$$

IQR is referred to as interquartile range. Q_3 is the third quartile (75 % of the data) and Q_1 is the first quartile (25 % of the data).

Figure S1 is the time series of the IQR in two cases. It can be clearly seen in both cases that the IQR increases dramatically with the long -term transport of dust. It is a sign that the mismatch between model and observation (position error) is becoming obvious.

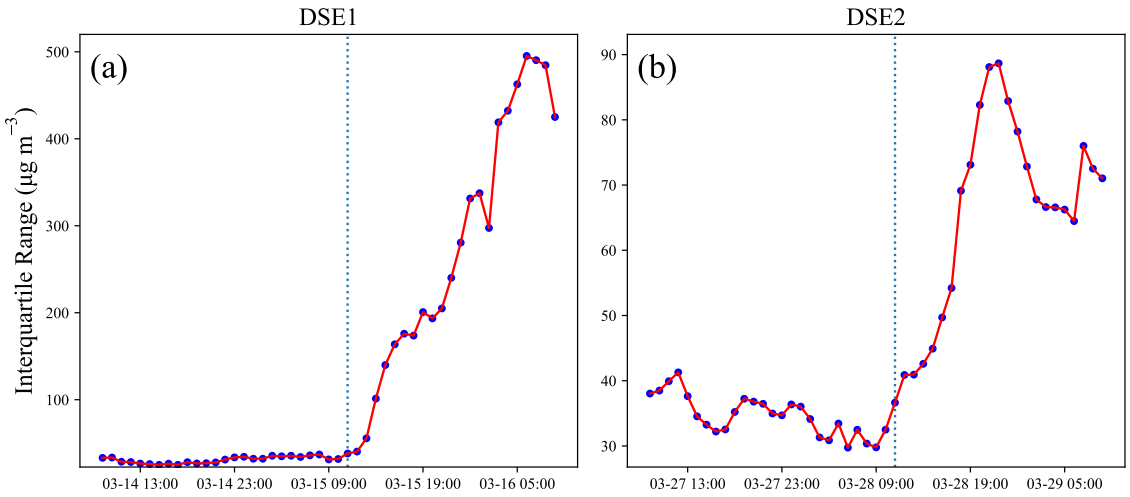

Figure S1. Time series of interquartile range in DSE1 (**a**) and DSE2 (**b**).

RC: ***5) In both Figure 4 (b.2) and Figure 7 (b.2), there are conspicuously high values located to the west of the dust plumes following the NTEnKF analysis. I'm curious if, in the absence or scarcity of observations, applying the NTEnKF could lead to the generation of false or overly extensive dust plumes, potentially exacerbating the inaccuracies of the original model simulation.***

AR: Thanks for the comment. We agree that it is possible that EnKF with VTS exacerbate the dust plume in the absence of surrounding observations. Therefore, we only used neighbouring ± 1 and ± 2 hour to weaken the impact. Sensitivity tests concerning the neighbouring time interval choice is presented in Comment 3).

> 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 neighbouring ensembles, the dust plume is extended wilder, as can't be verified by the observations.

RC: ***6) The scatters depicted in all the figures are too small to recognize, making it challenging for readers to quickly grasp the information being conveyed. Consider magnifying these visuals or narrowing the focus of the map to enhance the visibility of the dust shapes and allow for a more immediate and clear interpretation.***

AR: Thanks for the comment. We have enlarged most of the figures. It should be easier to follow now.

References

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.

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.

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.

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.