Authors' Response to Reviews of

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

Mijie Pang, Jianbing Jin*, Arjo Segers, Huiya Jiang, Wei Han, Batjargal Buyantogtokh, Ji Xia, Li Fang, Jiandong Li, Hai Xiang Lin, and Hong Liao
Geoscientific Model Development Discussions,

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. Major Comments

RC: ***There many glaring language errors throughout this manuscript. These errors are likely egregious enough to distractingly irritate readers with strong English skills. Here is a list of those errors/misunderstandings in the first page alone.***

L3: "model bias" -> "model error" DA corrects errors. Biases are a subset of errors.

L4: "... algorithm that effectively tunes models..." -> "algorithm that effectively improves numerical forecasts ..." When I think of "tunes model", I think of the model parameters (e.g., reaction rate coefficients) being tuned. In other words, your phrasing may mislead readers into thinking that you are doing parameter estimation (which is a separate kind of problem from your study).

L5-6: "However, when the position of the simulation does not align consistently with the observations which is referred to as position error, the EnKF algorithm struggles" -> "However, when the positions of simulated features are inconsistent with those observed (i.e., position errors), the EnKF algorithm struggles."

L7: "EnKF can hardly represent this uncertainty" -> "EnKF cannot adequately treat this error" It is the ensemble's job to represent the uncertainty, not the EnKF.

L8: "standard EnKF" -> "stochastic EnKF" As you are aware, there are many EnKF methods. I am unsure if the EnKF community has settled on whether the Burgers et al EnKF filter is called the "standard" EnKF. The names "stochastic EnKF" or "perturbed observation EnKF" are more precise and less likely to invite consternation.

L9: "ensembles" -> "ensemble members" In the ensemble DA literature, "ensemble" means "a group of model runs", and "an ensemble MEMBER" means "a model run inside the ensemble". Please use "ensemble members" to refer to having multiple model runs, and "ensemble" to refer to a group of model runs. This conflation of "ensemble" and "ensemble members" occurs throughout the manuscript.

L9-10: "In addition to the original ensembles quantifying dust loading variation, this methodology introduces extra ensembles from neighboring time for describing the potential spread of dust position." -> "In addition to the original ensemble members that quantify dust loading variations, this methodology introduces extra ensemble members from neighbouring valid times to incorporate position uncertainties."

L10-11: "allowing observations to be thoroughly resolved into the assimilation calculations" -> "thus enhancing the assimilation of observations"

L12: "that position error" -> "that position errors"

L16: "Dust storms are a natural meteorological disaster (Zhang et al., 2005) whose occurrence is attributed to frequent strong..." -> "Dust storms are natural meteorological disasters (Zhang et al., 2005) attributed to frequent strong..."

It is not the reviewers' job to copy-edit manuscripts. I strongly recommend engaging either an editorial service or a good writer work through the manuscript. If possible, ask a specialist on meteorological EnKFs to look over your manuscript.

AR: Thanks for the comment. We have adopted an editorial service to edit the manuscripts. Below is the polished abstract and introduction:

> **Abstract.** Dust storms pose significant risks to health and property, necessitating accurate forecasting for preventive measures. Despite advancements, dust models grapple with uncertainties arising from emission and transport processes. Data assimilation addresses these by integrating observations to rectify model error, enhancing forecast precision. The Ensemble Kalman Filter (EnKF) is a widely-used assimilation algorithm that effectively optimize model states, particularly in terms of intensity adjustment. However, the EnKF's efficacy is challenged by position errors between modeled and observed dust features, especially under substantial position errors. This study introduces the Valid Time Shifting-Ensemble Kalman Filter (VTS-EnKF) which combines stochastic EnKF with a valid time shifting mechanism. By recruiting additional ensemble members from neighboring valid times, this method not only accommodates variations in dust load but also explicitly accounts for positional uncertainties. Consequently, the enlarged ensemble better represents both the intensity and positional errors, thereby optimizing the utilization of observational data. The proposed VTS-EnKF was evaluated against two severe dust storm cases from spring 2021, demonstrating that position errors notably deteriorated forecast performance in terms of Root Mean Square Error (RMSE) and Normalized Mean Bias (NMB), impeding the EnKF's effective assimilation. Conversely, the VTS-EnKF improved both the analysis and forecast accuracy compared to the conventional EnKF. Additionally, to provide a more rigorous assessment of its performance, experiments were conducted using fewer ensemble members and different time intervals.

RC: ***Your manuscript and results can be misinterpreted to mean "EnKFs cannot handle position errors". My interpretation of your results is "EnKFs can handle position errors (to some degree) IF the ensemble has sufficient spread in feature positions". The issue that you are addressing is with the ensemble used, not the EnKF algorithm itself. Please go over your manuscript and make sure that you have explicitly communicated this distinction. Perhaps all statements that "the EnKF cannot handle position errors" should be replaced with "the EnKF cannot handle position errors if the ensemble is under-dispersive with regards to feature positions". To further limit the potential for misinterpretation, please rename your "EnKF" experiment as the "Basic" or "Naive" experiment.***

AR: Thanks for the comment. We agree that EnKF is unable to accurately account for position errors when the ensemble it uses is under-dispersive in terms of the locations of features. All the statements that can be misinterpreted is replaced (text shown below is in LineXXX). The experiment *EnKF* is renamed as *Basic*.

> The challenge lies in the quantification of position error and its subsequent inaccurate formulation of the background error covariance matrix. Consequently, EnKF calibrates both intensity and position error, while it cannot handle position errors if the ensemble is under-dispersive with regard to position. This deficiency curtails the capacity of current assimilation methodologies to correct position error.

RC: ***The authors used ground station observations to validate their experiments and forecasts. However, as the authors have noted in L396 "the dust plume is covered by the insufficient stations". Have the authors considered using satellite imagery? The authors can acknowledge this possibility in their areas for future research.***

AR: Thanks for the comment. We recognize the current limitation in our study due to the sparse distribution of ground stations in north-west, as mentioned in Section 2.1. One promising direction for future research, as pointed out by the reviewer, is to integrate satellite-derived dust optical depth (DOD) into our analysis. Satellite data, with its wide spatial coverage, can potentially bridge the gaps where ground observations are scarce. We acknowledge this as a valuable addition to our research agenda and plan to explore the utilization of DOD in upcoming studies.

> 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 ensemble, the dust plume is extended wilder. In the future research, assimilating satellite-derived dust optical depth (DOD) observations that have broader coverage may help to better constrain the enlarged ensemble.

RC: ***L420-423: "The EnKF analysis, however, does not improve this dust forecast after the initial assimilation." I disagree with this statement. Figures 7a.1, 7a.2 and 7a.3 clearly show that the EnKF experiment's RMSEs are lower than the Control experiment. I also disagree that "the RMSE and NMB of the dust forecast from the EnKF scenario are nearly identical to the Control Run". Figure 7a.2 and 7a.3 indicate that the EnKF can have up to 100 ug/m3 less RMSE than the Control.***

AR: Thanks for the comment. In the forecast step, we use a cyclic forecast at the interval of 3-hour. In this sentence, we indicate that the EnKF analysis only didn't improve the dust forecast after the **initial** assimilation, which is the forecast starting from 2021-03-15 11:00 in Fig. 7(a.1, b.1). Hence, at the first assimilation time point, we concluded that there are trivial improvements from the EnKF to the forecast performance.

3. Minor comments

RC: ***L37: "uncertain input data". Input data uncertainties are not a part result of numerical approximations. These uncertainties exist naturally on their own.***

AR: Thanks for the comment. We agree that uncertainties exist naturally on their own. This sentence is revised into a reasonable manner:

> However, model forecast skill is limited by the uncertain input data (e.g., wind field and boundary/initial conditions) and numerical approximations (like coarse grid cell and time step) (Mallet and Sportisse, 2006)

RC: ***Caption of Figures 1, 4, 5 and 6: you use the phrasing "with scatters of the model-minus-observation differences" or something similar. Please be more explicit with saying that the colored circles indicate observations/model-minus observations. Readers will easily miss this very important detail. A better phrasing, for example, is "The model-minus-observation differences at various observation sites are indicated by the plotted filled circles".***

AR: Thanks for the comment. We admit that the original captions are implicit. They have been replaced by

> **Figure 1.** Evolution of the simulated dust plume from average of ensemble members (**a.1-3**). Their corresponding standard deviation from ensemble members (**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. The filled circles represent ground BR-PM_{10} observations in (**a**) and (**c**), and the model-minus-observation differences (absolute value) at various observation sites in (**b**) and (**d**). 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). BR-PM_{10}: baseline-removed PM_{10}.
>
> **Figure 4.** Spatial distribution of simulated dust plume (SDP) on surface from average of ensemble members at central time (**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). The filled circles are ground-based BR-PM_{10} observations.

Figure 5. Spatial distribution of simulated dust plume (SDP) on surface from average of ensemble members at central time (**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). The filled circles are ground-based BR-PM_{10} observations.

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

RC: ***L279: What do you mean by "root of error from observations"? Do you mean the "observation error standard deviation"?***

AR: Thanks for the comment. Observation error standard deviation is exactly what we mean.

RC: ***L274-275: "Its mean is 0 and covariance is the root of diagonal from O". This description of covariance is incorrect. While we do take use square-root of O to generate the noise samples, the covariance of those noise samples is still O! You can check this for yourself in Python. The correct sentence is "Its mean is 0 and covariance is O."***

AR: Thanks for the comment. This is apparently a mistake. It is corrected in LineXXX:

> Its mean is 0 and covariance is the diagonal of **O**.

RC: ***It is not clear to me what observations are assimilated. Are you only assimilating the ground-based observation network? Please add a sentence in Section 3.3 that explicitly describes which observations you are assimialting.***

AR: Thanks for the comment. In Section 2.1, we described the observation we assimilate, which is the bias-corrected PM_{10} concentrations from ground monitoring network. But we didn't mention that in Section 3.3. Now an explicit description is made in LineXXX:

> The BC-PM_{10} observations are assimilated.

RC: ***L345: Please state the formula used to calculate NMB.***

AR: Thanks for the comment. We have added the formula for the evaluation metrics we used in Supporting Information.

> **4. Evaluation metrics**
>
> In this paper, Root mean square error (RMSE) and normalized mean bias (NMB) is used to evaluate the performance.
>
> $$RMSE = \sqrt{\frac{\sum_{i=1}^{m}(\boldsymbol{y}_i - \boldsymbol{\mathcal{H}}\boldsymbol{x_i})^2}{m}}$$
>
> $$NMB = \frac{\sum_{i=1}^{m}(\boldsymbol{y}_i - \boldsymbol{\mathcal{H}}\boldsymbol{x_i})}{\sum_{i=1}^{m}\boldsymbol{y}_i}$$
>
> m is the number of observations.

RC: ***Table 1: "Running ensemble member" – what do you mean by "Running"? Do you mean that you, for example, actually time-integrate 160 members? I am guessing that you mean "ensemble size used by the EnKF".***

AR: Thanks for the comment. It refers to both the ensemble size used by EnKF that produces analysis and ran by the model that produce forecast. We have revised it as "Ensemble size used by analysis and forecast"

RC: ***L372: "imbalanced uncertainty". I appreciate that you are trying to use your own words to describe the situation. However, the ensemble DA literature has a term for this: "ensemble underdispersion".***

AR: Thanks for the comment. We have used this proper literature throughout the paper.

RC: ***L430: "a reduction of RMSE is observed". A reduction relative to what? Please state that explicitly.***

AR: Thanks for the comment. We have made a more clear statement:

> By applying the VTS-EnKF analysis, a reduction of RMSE compared to the model run and EnKF can be observed in panel a. There is an approximate decrease of 100 μg m^{-3}in *VTS-EnKF* compared to *Basic*, which indicates that the VTS-EnKF analysis effectively corrects the position error.

RC: ***L426-427: I also disagree that "the localization method is unable to enhance the forecast". Figure 7 clearly shows that localization IS improving the forecast performance, albeit only slightly. The word "unable" is not the same as "slightly able to".***

AR: Thanks for the comment. There is only slight improvements with the localization for sure. We have used a appropriate expression on it in LineXXX:

> Moreover, it has been observed that the localization method only improves the forecast slightly in the presence of position errors.

RC: ***Your use of IQR (supplement) is problematic. A Gaussian distribution with massive spread will exhibit massive IQR. In other words, IQR does not distinguish between Gaussian and non-Gaussian distributions. A simple metric that might work is: |OmB|/(obs error variance + bg error variance). |OmB| is the absolute difference between the observation and background. For Gaussian distributions, that metric rarely exceeds 3. If you consistently see values greater than 3, then you have a clear sign of non-Gaussianity.***

AR: Thanks for the comment. We admit that the IQR cannot fully describe the non-Gaussian error distribution. As the reviewer suggests, the Squared Normalized Innovation is used to better identify the non-Gaussian statistics. Detailed descriptions are made in Supplementary:

> **3. Identification of position error**
>
> To objectively identify the position error, a simple identification index is applied, which is the Squared Normalized Innovation (SNI). SNI is a measure of how well the model forecasts match the observations and is crucial for evaluating the performance of the filter and diagnosing issues related to the error distribution assumptions, particularly the assumption of Gaussian errors (Zupanski and Zupanski, 2006). Here, it depicts the error statistics transiting from Gaussian to non-Gaussian distribution with emergence of position error:
>
> $$\mathrm{SNI} = \frac{\sum_{i=1}^{m} (\boldsymbol{y}_i - \boldsymbol{\mathcal{H}} \boldsymbol{x_i})^2}{trace(\mathbf{O}) + trace(\boldsymbol{\mathcal{H}} \mathbf{P}^f \boldsymbol{\mathcal{H}}^{\mathrm{T}})}$$
>
> m is the number of observations.
>
> If the values consistently diverging from one would indicate non-Gaussian error distributions. Figure S1 is the time series of the SNI in two cases. It can be clearly seen in both cases that the SNI 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.

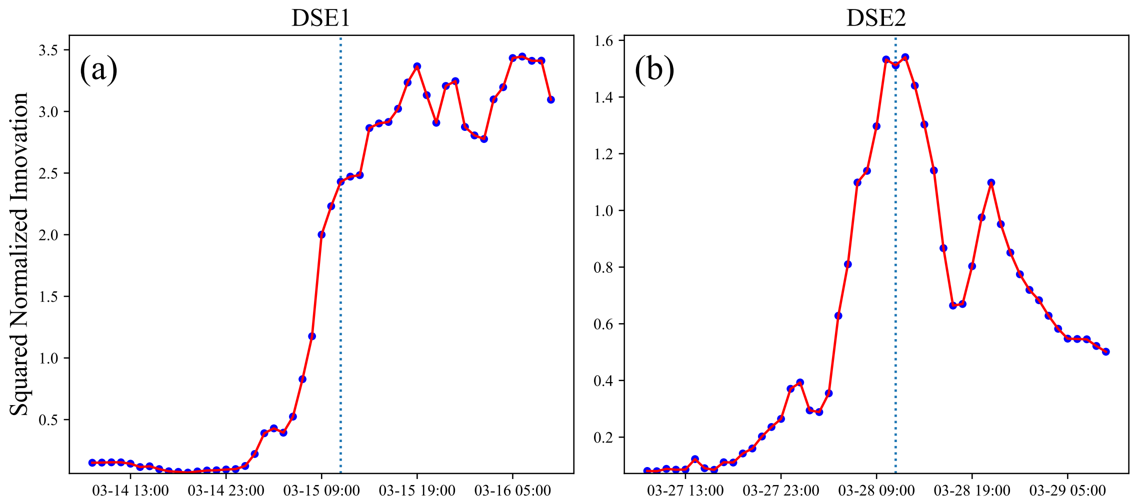

Figure S1. Time series of SNI in DSE1 (**a**) and DSE2 (**b**).

References

Mallet, V. and Sportisse, B.: Uncertainty in a Chemistry-Transport Model Due to Physical Parameterizations and Numerical Approximations: An Ensemble Approach Applied to Ozone Modeling, J. Geophys. Res., 111, , 2006.

Zupanski, D. and Zupanski, M.: Model Error Estimation Employing an Ensemble Data Assimilation Approach, Mon. Weather Rev., 134, 1337–1354, , 2006.

Authors' Response to Reviews of

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

Mijie Pang, Jianbing Jin*, Arjo Segers, Huiya Jiang, Wei Han, Batjargal Buyantogtokh, Ji Xia, Li Fang, Jiandong Li, Hai Xiang Lin, and Hong Liao
Geoscientific Model Development Discussions,

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

1. Overview

Response to Topic editor: We would like to thank the editor for the careful review throughout the paper and the in-depth comments that help to improve our paper.

2. Major comments

RC: ***Based on section 3.1, the error covariance matrices and Kalman gain are directly calculated with the ensemble perturbations. This is different from the common used EnKF algorithms. I expect that the dimension of the dust concentration (model variable) is large given the large domain of the dust model (LOTOS-EUROS). Please clarify how the error covariance matrices are calculated and updated in this study.***

AR: Thanks for the comment. In fact, we didn't directly calculate the prior error covariance matrix since its large size. A more efficient way is adopted as described in Supplementary:

2. Efficient calculation of EnKF and localization

For the sake of computational efficiency, the explicit computation of $\mathbf{P}^f$ is eschewed given its substantial dimensionality, which entails a vast number of state covariance ($\mathbb{R}^{n\times n}$, equating to 39200×39200 in the context of our study). Instead, we opt for deriving a new perturbation matrix $\mathbf{U}$ through the following operation:

$$\mathbf{U} = \boldsymbol{\mathcal{H}}\mathbf{X}^{f\prime}$$

Notably, $\mathbf{U}$ exhibits significantly reduced dimensions ($\mathbb{R}^{n\times m}$, approximately 39200×1000), thereby alleviating the burden compared to the background error covariance matrix $\mathbf{P}^f$.

Subsequently, by incorporating $\mathbf{U}$ into the calculation of the Kalman gain $\mathbf{K}$, we obtain:

$$\mathbf{K} = \frac{1}{\mathrm{N}-1}\mathbf{X}^{f\prime}\mathbf{U}^{\mathrm{T}}\left(\frac{1}{\mathrm{N}-1}\mathbf{U}\mathbf{U}^{\mathrm{T}} + \mathbf{R}\right)^{-1}$$

This formulation enables a scalable reduction in both memory requirements and computational expenses, tailored to the dimensions of the model state and the observational dataset.

To further enhance computational practicality, the localization scheme is applied to $\mathbf{K}$ as follows:

$$\mathbf{K} = \frac{1}{\mathrm{N}-1}\boldsymbol{\rho}_{n\times m}\circ\mathbf{X}^{f\prime}\mathbf{U}^{\mathrm{T}}\left(\frac{1}{\mathrm{N}-1}\boldsymbol{\rho}_{m\times m}\circ\mathbf{U}\mathbf{U}^{\mathrm{T}} + \mathbf{R}\right)^{-1}$$

Here, $\boldsymbol{\rho}_{n\times m}$ signifies the cross-correlation between model states and observations, while $\boldsymbol{\rho}_{m\times m}$ denotes the autocorrelation among observations (Evensen et al., 2022), thereby embodying an advanced strategy for managing the spatio-temporal correlations inherent in complex systems.

RC: ***After the VTS-EnKF update, will the enlarged ensemble size be reduced to the original size? How were the ensemble members selected?***

AR: Thanks for the comment. As Fig. 3 shows, VTS-EnKF is applied only in the initial assimilation analysis. Then the forecasts are made by the enlarged ensemble. The enlarged ensemble size won't be reduced to the original size. The cyclic forecasts are made on these ensemble members. In practice, smaller number of central time ensemble members can be set to be less computation-demanding and achieve adequate performance with aid of localization as described in Sect. 4.3.

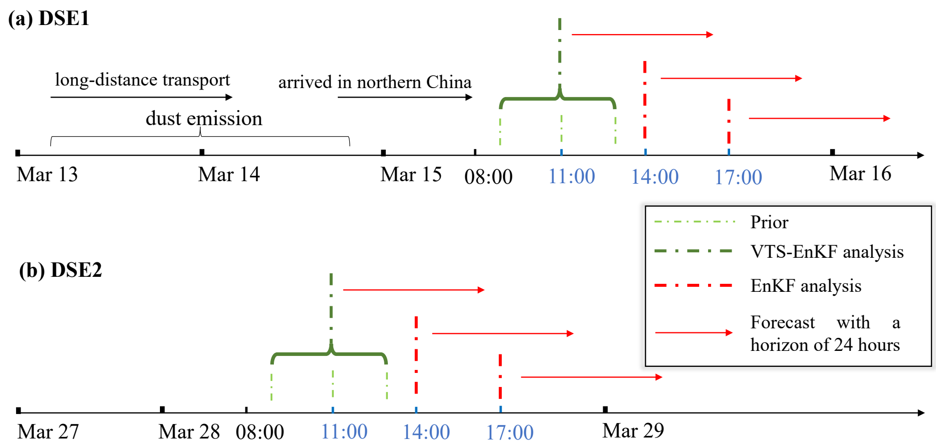

Figure 3. Sequential assimilation time set for DSE1 (**a**) and DSE2 (**b**). Take DSE1 for instance, the assimilation analysis is performed at the intervals of 3 hours from 11:00 to 17:00 and the rolling forecast is made with a horizon of 24 hours based on the assimilation analysis. The EnKF with VTS and EnKF is performed in turn.

RC: ***The emission parameterization should be briefly mention in section 2 and clarify what factors in the parameterization contributes to the uncertainty.***

AR: Thanks for the comment. More detailed description about the emission parameterization is added Supplementary:

5. Emission parameterization

The dust flux rate f is calculated as a function of horizontal saltation F_h, the sandblasting efficiency α, a terrain preference S, and an erodible surface fraction C as:

$$f = F_h \times \alpha \times S \times C$$

The horizontal saltation F_h represents the horizontal flux rate, which is proportional to the third power of the wind friction velocity u_*, as long as this exceeds a certain friction velocity threshold u_{*t}. Explicitly, F_h in a given grid cell is computed from:

$$F_h = \begin{cases} 0, & u_* \leq u_{*t} \\ \frac{\rho_a}{g} u_*^3 (1 + \frac{u_{*t}}{u_*})(1 - \frac{u_{*t}^2}{u_*^2}), & u_* > u_{*t} \end{cases}$$

where g denotes the gravitational constant, and ρ_a represents the atmospheric density. The friction velocity u_* is computed from the ECMWF wind speed at 10 m height assuming neutral atmospheric stability, following a logarithmic profile. The friction velocity threshold (FVT) u_{*t} represents the minimum friction velocity to initiate the movement of soil particles.

Emission errors are also likely to be induced during the formation of the friction velocity (u_*) from meteorology data, the terrain preference (S) from the topography resource, and the erodible surface fraction (C) from the land cover database. Among these factors, the friction velocity threshold (FVT, u_{*t}) is very important and sensitive for the outcome, since it directly influences whether dust saltation will occur and also quantifies the amplitude of the flux rate.

3. Minor comments

RC: ***Please use the formal format for the time (e.g. 1100 UTC).***

AR: Thanks for the comment. China Standard Time (CST) is the local time of where our study mainly focused on. We used it for the consistency with our previous research and better demonstrates the impact to the local region. We have added a description of this time format to the captions to avoid confusion.

RC: ***In the abstract: Even with position error, the perturbations can be Gaussian. I would remove the sentence "EnKF can be bias for the non-Gaussian statistics", since this sentence is irrelevant to this study.***

AR: Thanks for the comment. We admit that this sentence is excessive in the abstract. It is removed in the revised version.

RC: ***L71, Page 3: "For non-Gaussian problem...". Please revise this sentence.***

AR: Thanks for the comment. This sentence is revised as:

> Despite these strengths, the EnKF, as an extension of the Kalman Filter, presumes Gaussian error distributions (Amezcua and Van Leeuwen, 2014). When dealing with non-Gaussian error statistics, EnKF can create suboptimal outcomes for the linearized dynamics or operators and sampling errors caused by finite ensemble members (Lei et al., 2010).

RC: ***L385: This sentence is confusing. Please make it clearer.***

AR: Thanks for the comment. In this sentence, we intend to describe that the uncertainty spread has expanded by applying VTS. The ensemble underdispersion is alleviated, thus enhance the capability of EnKF to deal with the position errors. It is revised in LineXXX:

> This expansion of the uncertainty spread effectively addresses the issue of ensemble underdispersion, thereby boosting the EnKF's capability to handle position errors.

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.

Evensen, G., Vossepoel, F. C., and van Leeuwen, P. J.: Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem, Springer Textbooks in Earth Sciences, Geography and Environment, Springer International Publishing, Cham, , 2022.

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