We thank the reviewers for their careful review of our paper and the helpful comments that improved the manuscript. We give the response in blue and cite the revised text in orange. The original comments are reproduced.

Reviewer 1

I am generally satisfied with the revision of the manuscript, but I still have two comments on the double-space localization and the multiscale localization methods.

1. I understand that the Local DA algorithm is capable of double-space localization. My question is: do we really need to apply two localizations? The results of this study did not demonstrate the substantial benefits of the double-space localizations either. The authors need to comment on and discuss it.

Thanks for your comments. The authors agree that whether to use double-space localization should be explicitly discussed.

Firstly, the performance of double-space localization relies on the model-space localization in the combination. The results show (Figure 11) that double-space localization with a multiscale model-space localization (Ens_5band_DSL) outperforms the one with a fixed localization scale (Ens_noFLTR_DSL). Double-space localization can serve as a supplement to model-space localization.

Whether to use double-space localization depends on the application scenario. If Local DA is extended to a four-dimensional DA scenario, double-space localization is not recommended because observation-space localization does not consider the advection of background error covariance. In the case of a three-dimensional DA scenario, double-space localization is conducive to a small analysis error when the background error covariance is noisy because the localization helps reduce spurious increments. In contrast, if the background error covariance is good, double-space localization is not necessary.

We revised the related text as follows.

In section 4.2.3, line 543-546

“The spurious increment is further reduced in Hybrid_5band_DSL, especially at 850 hPa and 500 hPa, indicating that the positive impact of double-space localization corresponds to less noise in the analysis. According to the above result, double-space localization may serve as a supplement to pure model-space localization which determines the level of analysis error.”

In section 5, line 672-676

“Despite the encouraging results, whether to use double-space localization should be considered case by case. In this study, the background error covariance is noisy, so double-space localization has a positive impact. With a well-sampled ensemble and a well-designed multiscale localization, there is no need to use double-space localization. In the case of applying Local DA in the four-dimensional DA scenario, double-space localization should not be used because observation-space localization does not consider the advection of error covariance."

2. I am glad that the authors provided the algorithms (Eqs. [11-16]) for multiscale localization in Local DA. Those algorithms show that a multiscale localization method neglecting cross-scale covariances was proposed. However, those algorithms seem to be ad hoc for neglecting cross-scale covariances and are unable to be easily extended to include cross-scale covariances. If so, the authors need to clarify this in the manuscript.

Thanks for your comments. The authors agree that the multiscale covariance proposed in this study is difficult to extend to cross-scale covariance.

We revised the related text as follows.

In section 2.2, line 256-260

"Note that the multiscale covariance proposed in this section naturally excludes cross-scale covariance and is hard to incorporate cross-scale localization. How to determine the localization between two scales is also a question. The existing cross-scale localization (e.g., Huang et al., 2021; Wang et al., 2021) is implemented in spectral space and cannot be directly applied in equations (15) and (16). We plan to deal with the cross-scale issue in future work."

In addition, the multiscale localization method without the cross-scale covariances requires retaining the raw ensemble variances from the decomposed ensemble vectors. However, Eq. (11) shows that the decomposed scales are summed to recover the original raw ensemble perturbations. Eq. (11) should be used for the multiscale localization method with the cross-scale covariances. Authors can refer to Buehner (2012) and Huang et al. (2021) for details.

Thanks for pointing out this issue.

We revised the related text as follows.

In section 2.2, line 218-223

"To realize multiscale localization in model space, Local DA first performs scale decomposition with a bandpass filter. The decomposed perturbation, $\mathbf{X}'_{\mathrm{b}}$, is

$$\mathbf{X}'_{\mathrm{b}} = \left(\mathbf{X}^{1}_{\mathrm{b}}, \mathbf{X}^{2}_{\mathrm{b}}, \cdots, \mathbf{X}^{l}_{\mathrm{b}}, \cdots, \mathbf{X}^{N_{\mathrm{b}}}_{\mathrm{b}}\right), \quad (11)$$

where the superscript "l" represents the lth scale and N_{b} is the number of scales. After decomposition, the number of samples becomes N_{b} times as large as the original ensemble size. As a localization approach lacking cross-scale covariance (no $\mathbf{X}_{\mathrm{b}}^{i}\mathbf{X}_{\mathrm{b}}^{j\mathrm{T}}, i \neq j$ term in $\mathbf{X}_{\mathrm{b}}' \mathbf{X}_{\mathrm{b}}'^{\mathrm{T}}$), Local DA computes the STD of the perturbation, *s*, according to"

Reviewer 2

General comments A new algorithm called ”Local DA” is introduced. Compared to the last version, the current manuscript has been considerably improved. Overall, I feel that the paper is lengthy as authors attempts to tackle several problems. But some of them remain still unclear to the end and will be further investigated in the future as authors claims. In my opinion, the most remarkable difference compared to the LETKF is that the Local DA works in columns instead of in points. Therefore, I have one concern in this regard.

1. If I understand correctly (I may get it wrong), the Local DA works in columns. Therefore, I thought there is no need for vertical localization for integrated observations (e.g., PWV). But there is vertical localization applied to PWV. Intuitively, I would have thought that the Local DA should be able to better assimilate PWV data since it does not require the vertical localization, but there is no discussion about this.

Thanks for your comments. Yes, you are right. The Local DA can work in columns, as well as in points. In this study, the Local DA works in 5-column mode.

The vertical localization in the observation space was disabled for all Local DA experiments. It can be seen in the code (src/da_core/da_core_enda.f90) that the vertical distance between an observation and a model grid point is set to zero (dv=0.0) for Local DA (broot_opt==1).

```
if(broot_opt==0) call cal_loc_coef_ll(dh,dv,radii_h_tem,radii_v_tem,envar%local_coef(iobs_sub),coef_mark)
if(broot_opt==1) call cal_loc_coef_ll(dh,0.0,radii_h_tem,radii_v_tem,envar%local_coef(iobs_sub),coef_mark)
```

In the previous revision, we stated that Local DA performs vertical localization in the model space (“Notably, Ens_noFLTR_OL performs vertical localization in model space, identical to Ens_noFLTR.” and “Ens_noFLTR_DSL performs localization in both the model and observation space. In the model space, a fixed localization radius is used, as in Ens_noFLTR, while the localization parameters of Ens_noFLTR_OL are adopted for observation-space localization.”), but we did not explicitly claim that the vertical localization in the observation space is disabled. To make the statement clearer and emphasize no vertical localization in the observation space, we revise the related text as follows.

In section 3.3.2, line 429-431. We state that the vertical localization in the observation space is set in the LETKF experiment (Ens_LETKF).

“**The vertical radius for all observations is 5 km in Ens_LETKF**, where the PWV observations are supposed to be available at 4000 m for LETKF localization.”

We explicitly state that Local DA uses no vertical localization in the observation space at line 434-436

“For convenience, all single deterministic analysis experiments are listed in Table 3, where “M”, “O”, and “M+O” denote model-space, observation-space, and double-space localization, respectively. **The vertical localization in the observation space is disabled for all Local DA**

experiments."

Minor points:

1. Line 55: do not understand "The LETKF, however, performs the analysis in the ensemble space, which implies that a static ensemble is necessary".

The related words are revised as "The large ensemble (>=800) is not always available in practice because of the limited computational and storage resources. However, it is inevitable to use such an ensemble to realize the hybrid analysis in the original LETKF framework because the LETKF works in the ensemble space." (Line 54-57)

2. Section 2.1: It is suggested that dimensions be given here as well.

The dimensions of vectors and matrices in equations (1) and (2) depend on the number of observations used in a local analysis and the complexity of observation operators. Therefore, it is better to give the dimensions and computations of the above variables in the subsequent subsections. We added some words in section 2.1 to state the above circumstance.
Line 115-117
"The dimensions of vectors and matrices in equations (1) depend on the number of observations involved in a local analysis and the complexity of observation operators. We will give the dimensions and computations of the above variables in the following subsections."

3. Line 170: do not understand "Note that the variational DA methods seek the combination of the columns of the square root of the background error covariance matrix, while Local DA combines the columns of the error correlation matrix" since the 3DVAR is also formulated as an optimization problem of control variables.

Thanks for your comments. The variational DA methods and Local DA differ in the control variable transform viewpoint. The control variable transform converts control variables to model state variables. The variational DA methods use the square root of the background error covariance matrix, while Local DA employs the error correlation matrix.
The related words are revised as "Note that the variational DA methods and Local DA differ in the control variable transform viewpoint. The former uses the square root of the background error covariance matrix, while Local DA employs the error correlation matrix." (Line 172-173)

4. Line 210: What is $\overline{\hat{\mathbf{X}}_o}$?

The mean of $\hat{\mathbf{X}}_o$.

The related words are revised as

"Local DA approximates the linear projection $\tilde{\mathbf{Y}} = \mathbf{H}_o\hat{\mathbf{X}}_o$ according to

$$\tilde{\mathbf{Y}} \approx h(\mathbf{x}^f + \hat{\mathbf{X}}_o) - h(\mathbf{x}^f + \overline{\hat{\mathbf{X}}_o}), \quad (10)$$

where h is the nonlinear observation operator, $\mathbf{x}^f$ is the background model state vector, and $\overline{\hat{\mathbf{X}}_o}$ is the mean of $\hat{\mathbf{X}}_o$." (Line 211-213)

5. Line 244: do not change —— > are equal

Our original statement is somehow confusing and we have revised it.
The related words are revised as "Because the multiscale localization does not change the sizes of $\mathbf{C}_{oo}$ and $\mathbf{C}_{mo}$, there is no modification for $\mathbf{v}_o$, $\mathbf{x}^i$, $\mathbf{x}^f$, and $\mathbf{x}^a$." (Line 247)

6. Line 259 : Typo: Coo

Fixed as $\mathbf{C}_{oo}$. (Line 266)

7. Line 283: $\mathbf{x}^f$ is not initial condition

We use "background model state" in the revision. (Line 290)

8. Line 407: Is the fixed multiplication inflation 1.5 employed in all experiments?

Yes.

9. Line 464: "Let us have a quick look at the results", the language is too causal.

The related text is revised as "The domain averaged root mean square root error (RMSE) is examined first. For convenience, the initial condition extracted from GFS analysis is referred to as BAK. All experiments reduce the errors in the observation space after DA" (Line 472-473)

10. Line 511: speculation —— > assumption

Fixed. (Line 519)