

Response to the Comments of Referees

We thank the reviewer for the careful re-evaluation of our manuscript and for the constructive comments. These comments have helped us identify an inconsistency in the reported heating-rate enhancement and clarify the interpretation of the optical assumptions, the respective roles of the LETKF and bypass updates, and the specification of observation uncertainty. We have made revision based on the critical comments and suggestions of the referees. The referee's comments are reproduced (*black*) along with our replies (*blue*) and changes made to the text (*red*) in the revised manuscript. All the authors have read the revised manuscript and agreed with submission in its revised form.

Anonymous Referee #1

Comment NO.1: *In my previous Major Comment 3, I asked the authors to clarify how the mismatch between the internally mixed CALIOP-MODIS absorbing component and externally mixed GOCART BC propagates into the diagnosed heating. Specifically, I asked whether the assimilation compensates by altering BC mass and whether the reported peak heating of 3.0 K d⁻¹ should be regarded as an upper bound, a lower bound, or a bracketed estimate.*

I appreciate that the authors have now carried this issue into Section 4.4. However, the response appears to replace the requested assessment with two assumptions. The authors state that the internally mixed CALIOP-MODIS particle has a higher MEE than externally mixed GOCART BC, and then assume $SSA_{int} < SSA_{ext}$, from which they conclude that the 3.0 K d⁻¹ heating is a conservative lower bound.

The algebra that follows is valid only if those optical-property relationships hold for the specific CALIOP-MODIS and CRTM/GOCART representations used here. Neither relationship is demonstrated in the response. In particular, stating that internal mixing enhances BC absorption per unit BC mass does not by itself demonstrate that it has a larger absorption fraction per unit extinction. Once extinction is constrained by the assimilation,

$$\beta_{abs} = \beta_{ext}(1 - SSA) = \beta_{ext} \frac{MAC}{MEE}$$

Therefore, the relevant comparison is MAC/MEE, or equivalently 1-SSA, between the two optical representations. The statement in the response that the “underestimation of the mass absorption cross-section dominates” is precisely the point that requires quantitative support rather than assumption.

The lensing effect can certainly enhance BC absorption per unit BC mass, i.e. increase MAC. However, because the DA constrains extinction rather than BC mass, this does not by itself determine the sign of the heating bias. The corresponding change in MEE must be considered simultaneously. For the author's lower-bound interpretation to hold, the relative enhancement in MAC must exceed the relative enhancement in MEE:

$$\frac{MAC_{int}}{MAC_{ext}} > \frac{MEE_{int}}{MEE_{ext}}$$

Only under this condition does MAC/MEE increase and $SSA_{int} < SSA_{ext}$ follow.

I therefore suggest the authors demonstrate this relationship directly for the optical schemes actually used in the study. At 532 nm, and preferably for representative RH conditions, please provide the MEE and SSA (or MAC) for the CALIOP-MODIS core-grey-shell representation and for CRTM/GOCART BC. In particular, comparison of $(1-SSA)_{int}$ and $(1-SSA)_{ext}$ in the form

$$R_{abs} = \frac{(1 - SSA)_{int}}{(1 - SSA)_{ext}}$$

would directly establish the sign of the effect once extinction is constrained. $R_{abs} > 1$ would support the author's interpretation that the externally mixed calculation underestimates absorption and that 3.0 K d^{-1} value is a lower bound; $R_{abs} < 1$ would imply the opposite.

Because the internally mixed representation referred to here originates from the Kudo et al. (2023) core-grey-shell retrieval, the comparison should use the optical properties of the actual absorbing component employed in that retrieval rather than a generic statement about internally mixed BC.

Response: We sincerely thank the reviewer for this rigorous and insightful comment. You are absolutely correct. Our previous assumption that internal mixing inherently leads to a higher absorption fraction per unit extinction was flawed because we did not quantitatively evaluate the simultaneous changes in the Mass Extinction Efficiency (MEE) and Mass Absorption Cross-section (MAC) between the two specific optical schemes used in our study. We deeply appreciate you pointing out the fundamental relationship $\beta_{abs} = \beta_{ext}(1 - SSA)$ and the requirement to evaluate $R_{abs} = \frac{(1-SSA)_{int}}{(1-SSA)_{ext}}$. Following your excellent suggestion, we have extracted the exact optical parameters from the literature describing the CALIOP-MODIS retrieval (Kudo et al., 2023) and the CRTM/GOCART forward model (Lu et al., 2022) to quantitatively calculate R_{abs} at 532 nm.

(1) Internally Mixed Representation in CALIOP-MODIS Retrieval: According to Table 1 in Kudo et al. (2023), the absorbing component is modeled as an internal mixture of black carbon (BC) and water-soluble substances using a core-grey shell model. The Single Scattering Albedo (SSA_{int}) at 532 nm for this internally mixed BC component is 0.44 under dry conditions ($RH = 0\%$) and 0.64 under wet conditions ($RH = 80\%$). The thick water-soluble shell significantly enhances the scattering cross-section, leading to a relatively high SSA despite the internal mixing.

(2) Externally Mixed Representation in CRTM/GOCART Model: According to Table 1 in Lu et al. (2022), the CRTM GOCART module treats BC as pure, externally mixed hydrophobic/hydrophilic particles with an effective radius of $0.036\mu\text{m}$ and a refractive index of $1.75 + 0.44i$. Based on Mie scattering theory for such small, highly absorbing pure carbon particles, scattering is extremely weak compared to absorption. The SSA at 532 nm for this externally mixed pure BC (SSA_{ext}) is approximately 0.20.

(3) Calculation of the Absorption Ratio (R_{abs}): Once the data assimilation system constrains the total extinction (β_{ext}), the ratio of the absorption fraction between the internal and external models is:

$$R_{abs} = \frac{(1-SSA)_{int}}{(1-SSA)_{ext}}$$

Dry conditions (RH = 0%): $R_{abs} = \frac{1-0.44}{1-0.20} = \frac{0.56}{0.80} = 0.70$

Wet conditions (RH = 80%): $R_{abs} = \frac{1-0.64}{1-0.20} = \frac{0.36}{0.80} = 0.45$

In both dry and wet conditions, $R_{abs} < 1$. This proves that your hypothesis was entirely correct, the externally mixed pure BC model in GOCART actually has a higher absorption fraction per unit extinction than the internally mixed core-grey shell model used in the CALIOP-MODIS retrievals. Because our DA system forces the model's externally mixed BC to match the retrieved extinction, we are inadvertently allocating too much of that extinction to absorption. Therefore, our previous statement was wrong. The estimated shortwave heating rates do not represent a lower bound.

We have completely rewritten the relevant paragraph in Section 4.4 to reflect this quantitative assessment and correct our physical interpretation.

Changes in Manuscript: In Lines 581-592, we have replaced “The CALIOP-MODIS retrieval assumes an internal core-shell mixing state, whereas the GOCART optical scheme assumes externally mixed BC. Because external mixing yields a lower mass extinction efficiency, the DA system must inflate the BC mass to match the observed extinction. However, internal mixing significantly enhances absorption, resulting in a lower single scattering albedo. Consequently, while the DA system successfully constrains the total extinction, applying the externally mixed optical scheme underestimates the absorption fraction per unit extinction. Therefore, the true shortwave heating induced by these elevated BC plumes could be even stronger than our current estimates.” with “It is important to acknowledge how the discrepancies in aerosol mixing state assumptions between the retrieval algorithm and the forward model affect the diagnosed heating rates. The CALIOP-MODIS retrieval assumes an internal core-grey shell mixing state for black carbon (BC) (Kudo et al., 2023), which yields the SSA at 532 nm ranging from 0.44 (dry conditions) to 0.64 (wet conditions at RH=80%). In contrast, the GOCART optical scheme in CRTM assumes externally mixed, pure BC particles (Lu et al., 2022), which possess a much lower SSA of approximately 0.20 due to the lack of a scattering shell. When the data assimilation system constrains the total extinction to match the observations, the absorption coefficient is determined by the fraction $(1-SSA)$. The ratio of the absorption fraction between the internally mixed retrieval and the externally mixed model, defined as $R_{abs} = \frac{(1-SSA)_{int}}{(1-SSA)_{ext}}$, ranges from 0.45 to 0.70. Because $R_{abs} < 1$, applying the externally mixed optical scheme in the model overestimates the absorption fraction per unit extinction.”

Comment NO.2: So upon closely going through the result again, I have a follow up question for the 3.0 K d^{-1} value! In Lines 563-567 state that Figure 13 shows the mean vertical profiles of SW heating rate along the CALIOP tracks and that, relative to FR, the DA experiment produces a “maximum heating rate increase of approximately 3.0 K day^{-1} ” at 4-5 km. The caption likewise defines Fig. 13a as the mean SW heating profiles and Fig. 13b as the vertical profile of the DA-minus-FR heating-rate difference, with both profiles averaged along the CALIOP track. However, the plotted values do not appear to support a mean difference of 3.0 K d^{-1} . In Fig. 13a, the DA

heating rate near 4-5 km is approximately 0.8 K d^{-1} , while FR is approximately 0.5 K d^{-1} , implying a difference of roughly 0.3 K d^{-1} . Consistently, Fig. 13b peaks only slightly above $+0.2 \text{ K d}^{-1}$, approximately $0.25\text{-}0.3 \text{ K d}^{-1}$. Thus, the stated 3.0 K d^{-1} value appears to be about an order of magnitude larger than the mean DA and FR difference actually shown in Fig. 13.

A heating-rate increment approaching 3 K d^{-1} does appear in Fig. 14, but there it corresponds to the upper end of the individual assimilation samples, particularly in the strongly absorbing $\text{SSA} < 0.85$ regime, rather than to the mean vertical-profile difference. Figure 14 is explicitly described as the relationship between assimilation-induced extinction increments and heating-rate increments for individual samples.

This distinction is important because the same 3.0 K d^{-1} value is subsequently used at Lines 571-578 as the basis for the new “conservative lower bound” discussion.

So is the 3.0 K d^{-1} value actually represents the maximum of the mean DA-FR vertical profile shown in Fig. 13b; or the maximum among the individual samples shown in Fig. 14; or a quantity derived using some other spatial/temporal averaging or normalization that is not currently described?

Response: Thank you for pointing out this error. We agree that the maximum track-averaged DA-minus-FR heating-rate difference in Fig. 13b is approximately 0.3 K day^{-1} , not 3.0 K day^{-1} . The value of approximately 3.0 K day^{-1} refers to a local instantaneous heating-rate increment in the case shown in Fig. 15 and was incorrectly cited in the discussion of Fig. 13. We have corrected the value in the discussion of Fig. 13 and placed the description of the 3.0 K day^{-1} increment in the discussion of Fig. 15, explicitly distinguishing the track-averaged result from the case-specific local value. We have also removed the associated statement describing the peak heating as a conservative lower bound. We apologize for the error and the resulting confusion.

Changes in Manuscript:

(1) In Lines 575-577, to accurately reflect the results shown in Figure 13b, we corrected the maximum value of the track-averaged heating rate difference by replacing “..., with a maximum heating rate increase of approximately 3.0 K day^{-1} at 4-5 km” with “The track-averaged DA-minus-FR heating-rate difference reaches a maximum of approximately 0.3 K d^{-1} at 4-5 km (Fig. 13b).”

(2) To avoid confusion and maintain consistency with the corrected values, we deleted the statement interpreting the incorrectly cited 3.0 K day^{-1} value as a lower bound: “It is important to note that the estimated 3.0 K day^{-1} peak heating likely represents a conservative lower bound due to discrepancies in aerosol mixing state assumptions.”

(3) In the caption of Figure 13, to explicitly state the statistical meaning of the profiles and the averaging procedure, we replaced “Figure 13. Vertical profiles of the shortwave (SW) heating rate. (a) Comparison of the mean SW heating rate (units: K day^{-1}) between the FR and DA experiment. (b) The vertical profile of the heating rate difference (DA minus FR). The profiles are averaged along the CALIOP track.” with “Figure 13. Track-averaged vertical profiles of shortwave (SW) heating rate (K d^{-1}). (a) Mean profiles for the FR and DA experiments. (b) Mean profile of the paired

DA-minus-FR heating-rate differences. Both panels use the same CALIOP-track sampling and averaging procedure.”

(4) In Lines 618-620, to introduce the $\sim 3.0 \text{ K d}^{-1}$ value into its correct context and clarify that it represents a local instantaneous heating-rate increment associated with the specific case study in Figure 15, we added the following sentence: “The enhanced SW heating is spatially consistent with this BC increment, with local instantaneous DA-minus-FR heating-rate increases reaching approximately 3.0 K d^{-1} .”

Comment NO.3: *In response to comment 1 and comment 5 of the previous review, I appreciate that the authors have responded constructively by adding the DA_NoBypass experiment and the new discussion in Section 5.2. They report that the bypass is triggered at approximately 28.6% of assimilated grid points, including 43.6% for dust, 17.8% for water-soluble aerosol and 24.5% for BC, predominantly at 4-7 km. They further show that, without the bypass, the BC profile CORR increases from 0.26 to 0.79, and interpret this as evidence of “genuine flow-dependent ensemble covariance skill.” They also acknowledge that DA_NoBypass still underestimates strong aerosol layers, whereas the full DA further reduces BIAS and RMSE, and argue that the independent AERONET/AD-Net improvements reflect dynamical propagation rather than forced local agreement. These additions are useful, but Fig. 16 reveals an important distinction that I think should now be made more explicit.*

For the domain-averaged BC profile, BIAS/RMSE/CORR change from approximately $-0.026/0.043/0.26$ in FR, to $-0.027/0.038/0.79$ in DA_NoBypass, and to $-0.021/0.031/0.95$ in DA. Thus, the pure LETKF clearly improves the vertical-profile structure and reduces RMSE, but does not reduce the negative bias of the domain-averaged BC profile. The additional reduction in negative bias and strengthening of the elevated layer appear only in the full DA. This is consistent with the authors’ own statement that DA_NoBypass captures structural variation but still underestimates the magnitude of strong aerosol layers.

I therefore have the following remaining questions.

3.1 The species-specific triggering fractions omit sea salt, although Fig. 16a shows a noticeable divergence between the DA and DA_NoBypass sea-salt profiles, particularly aloft. The denominator of “28.6% of total assimilated grid points” is also not defined, and the vertical distribution is given only qualitatively as predominantly 4-7 km. Please report the sea-salt triggering fraction, define precisely what constitutes an “assimilated grid point,” and provide the trigger frequency as a function of altitude for each species. A simple temporal summary would also help establish whether triggering is broadly distributed through the assimilation period or concentrated in a small number of major plume events.

3.2 Fig. 16 suggests that the standard LETKF substantially recovers the vertical shape of the BC profile, whereas the bypass provides an important additional correction to the magnitude of the elevated layer when the background spread collapses. I suggest stating this distinction explicitly in Section 4.1, Section 5.2, and where the elevated BC reconstruction is highlighted in the Abstract

and Conclusions. At present, statements such as the “successful reconstruction” of the lofted BC plume do not distinguish between these two contributions.

3.3 Lines 665-669 argue that, because simple bias-rescaling should not significantly alter “the temporal or spatial variation patterns,” the increase in BC CORR from 0.26 to 0.79 demonstrates genuine flow-dependent skill. However, the Fig. 16 caption states that these statistics are calculated from domain-averaged vertical profiles over East Asia. The CORR therefore measures similarity in vertical-profile shape; it does not demonstrate temporal variability or horizontal spatial variability. Because the bypass is preferentially triggered at 4-7 km, it can also modify the domain-mean vertical-profile shape itself.

I suggest providing, for DA_NoBypass, the same point-wise N, BIAS, RMSE and CORR diagnostics used in Fig. 3, where all valid space-time-height observations are included. This would show whether the LETKF-only improvement also holds across the full spatiotemporal sample rather than only in the shape of the domain-averaged profile.

3.4 Lines 674-678 state that the AERONET and AD-Net improvements reflect “genuine dynamical propagation of the assimilated fields rather than the forced local agreement of the bypass mechanism.” I do not think the present experiments isolate that conclusion. The independent validations use the full DA experiment, which includes the bypass, and aerosol increments introduced by the bypass can themselves propagate dynamically away from the CALIOP tracks.

Since DA_NoBypass has already been performed, I strongly suggest evaluating it against the same AERONET AOT/SSA and AD-Net extinction observations. A simple FR / DA_NoBypass / DA comparison, either in the existing figures or in a compact table, would directly show how much independent improvement remains when the empirical bypass is removed. In my view, this is the most informative remaining diagnostic for separating the two mechanisms.

3.5 Please clarify whether the bypass remained active in DA_Obserr. Section 5.3 attributes the relatively weak sensitivity to increasing the observation-error baseline from 10% to 50% partly to the dominance of representativeness error and partly to the robustness of the flow-dependent covariance. However, Section 5.2 explicitly acknowledges that the background correction ‘bypasses the rigorous uncertainty weighting of the standard Kalman filter’ and that ignoring observational uncertainty during the correction step remains a limitation. More importantly, the definition of Rcorr in Lines 297-300 contains the localization weight f_i , the background ensemble mean and the observed extinction, but no weighting by the prescribed observation-error covariance. The triggering criteria and subsequent rescaling likewise depend on Rcorr, the number of local observations, the 3-sigma screening and the empirical thresholds; when triggered, the rescaled ensemble directly serves as the analysis and the LETKF step is bypassed.

Thus, for a given bypass event, the correction magnitude is not directly affected by increasing the prescribed observation-error baseline from 10% to 50%. The altered observation error could still affect subsequent background states through earlier LETKF updates and thereby indirectly change

when the bypass triggers, but the bypass update itself does not contain the observation-error weighting being tested in DA_Obserr.

If the bypass remained active in DA_Obserr, then the weak overall sensitivity of that experiment cannot be attributed solely to the robustness of the flow-dependent covariance. I therefore suggest qualifying this interpretation and reporting the bypass triggering frequency in DA_Obserr so that the contribution of this error-insensitive update pathway can be assessed.

Response: We sincerely thank the reviewer for this thorough and highly constructive evaluation of the DA_NoBypass experiment. The reviewer is entirely correct in pointing out that while the pure LETKF successfully recovers the vertical structural shape, the empirical bypass mechanism is indispensable for correcting the absolute magnitude of the elevated aerosol layers when the background ensemble spread collapses. We agree that making this distinction explicit significantly improves the transparency and physical interpretation of our assimilation system. We have addressed each of your specific points below:

(1) Response to Comment 3.1: We have clarified that the assimilated grid point is defined as the total number of grid cells where valid CALIOP observations were available and assimilated across all assimilation cycles during the study period. Based on this definition, the bypass triggering fraction for sea salt is approximately 42.9%. Regarding the vertical distribution, the triggering frequency is not uniformly distributed but exhibits significant unevenness across different altitudes, predominantly occurring in the free troposphere where the model struggles to reproduce the magnitude of lofted plumes.

(2) Response to Comment 3.2: We fully agree with this distinction. We have revised Section 4.1, Section 5.2, and the conclusions to explicitly state that the standard LETKF substantially recovers the vertical shape and relative partitioning of the aerosol profiles, whereas the bypass mechanism provides the critical magnitude correction for elevated layers (such as the lofted BC plumes) where the model's background spread collapses due to severe transport or emission deficiencies.

(3) Response to Comment 3.3: To demonstrate the genuine flow-dependent skill of the LETKF across the full spatiotemporal domain, we have computed the point-wise statistics (N, BIAS, RMSE, and CORR) for the DA_NoBypass experiment using all valid space-time-height observations, identical to the methodology used for Fig. 3. These results are now provided in Figure S3 in the Supplement, confirming that the pure LETKF indeed yields broad spatiotemporal improvements over the FR experiment, even without the bypass.

(4) Response to Comment 3.4: We agree that increments introduced by the bypass can propagate dynamically, which conflates the evaluation of the LETKF's independent predictive skill. To assess how much independent improvement remains without the bypass mechanism, we have evaluated the DA_NoBypass experiment against the independent AERONET AOT observations. These comparisons are now presented in Figure S4 in the Supplement. The results clearly demonstrate that DA_NoBypass retains substantial independent improvements over the FR experiment, confirming

the LETKF's genuine dynamical propagation skill, while the full DA provides the optimal overall performance.

(5) Response to Comment 3.5: We confirm that the bypass mechanism remained active in the DA_Obserr experiment. The reviewer correctly identifies that because the localized weighted bias ratio (R_{corr}) does not incorporate the prescribed observation-error covariance (R), the bypass updates are insensitive to the inflated observation error baseline. Consequently, the overall robustness of the DA_Obserr experiment is partially attributable to this error-insensitive bypass pathway. We have revised Section 5.3 to explicitly qualify our interpretation, acknowledging that the bypass mechanism contributes to the system's insensitivity to the inflated observation error.

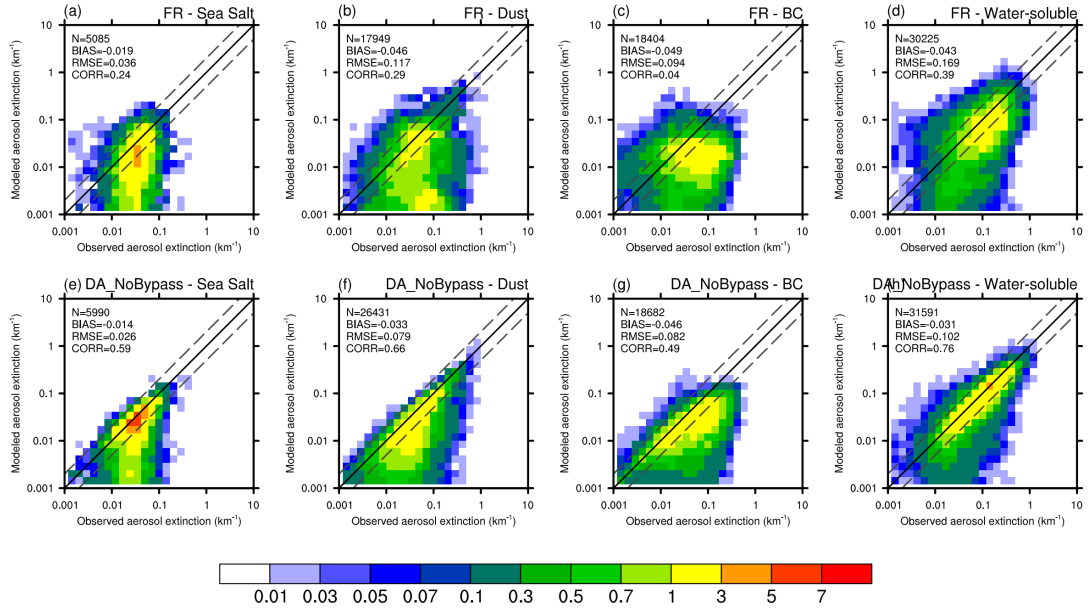


Figure S3. Same as Fig. 3 but for DA_NoBypass experiment.

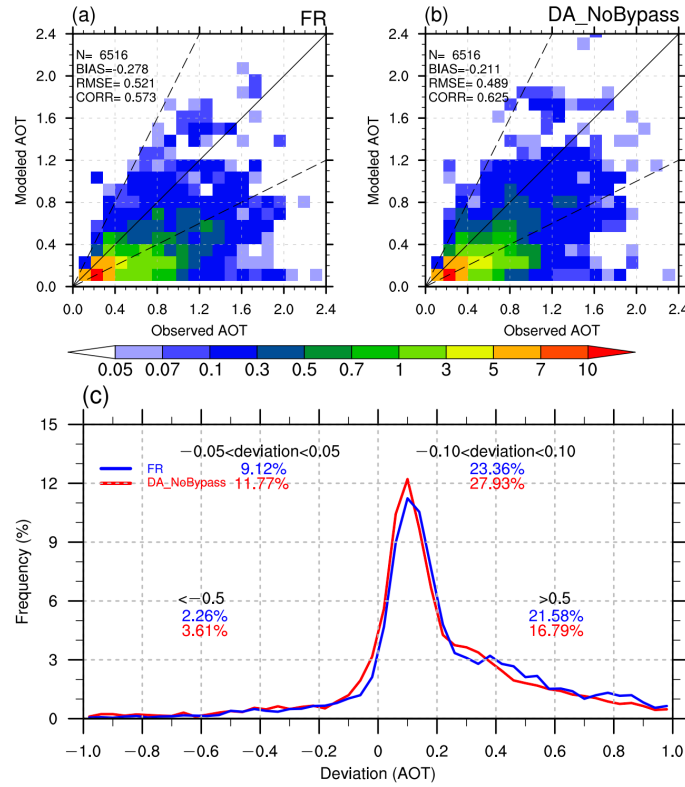


Figure S4. Same as Fig. 6 but for DA_NoBypass experiment.

Changes in Manuscript:

(1) In Lines 404-407, we added the following clarification regarding the elevated BC plume: "It is important to distinguish the respective roles of the assimilation components in this reconstruction. As further detailed in Section 5.2, the standard LETKF update substantially recovers the vertical shape of the BC profile, whereas the empirical background correction provides the critical magnitude adjustment for the elevated layer where the background ensemble spread approaches zero."

(2) In Lines 674-678, we updated the triggering statistics: "The assimilated grid point is defined as the total number of grid cells containing valid CALIOP observations across all assimilation cycles. Disaggregated by species, it triggered in 43.6% of cases for dust, 17.8% for water-soluble aerosols, 24.5% for BC, and approximately 42.9% for sea salt. Vertically, the triggers were unevenly distributed across different altitudes, predominantly occurring in the free troposphere."

(3) In Lines 690-697, we provided the new supplementary figures and the related description for the DA_NoBypass evaluation: "To verify the spatiotemporal skill of the pure LETKF, point-wise statistics for DA_NoBypass across all valid space-time-height observations are presented in Fig. S3, confirming genuine flow-dependent improvements beyond just the domain-averaged vertical shape. Furthermore, to rigorously isolate the dynamical propagation of the LETKF from the forced local agreement of the bypass, we evaluated the DA_NoBypass experiment against independent AERONET AOT observations (Fig. S4). The results show that DA_NoBypass still significantly reduces the RMSE and BIAS relative to the FR experiment, confirming that the flow-dependent

covariance can improve the downstream aerosol transport, although the full DA experiment yields the best overall performance.”

(4) In Lines 730-735, we qualified the interpretation of the DA_Obserr sensitivity: "It is important to note that the bypass mechanism remained active in the DA_Obserr experiment. Because the background correction criteria (R_{corr}) do not incorporate the prescribed observation-error covariance, these specific bypass updates are insensitive to the inflated observation error. Therefore, the relatively weak sensitivity of the overall assimilation performance to the 50% error baseline is attributable not only to the robustness of the flow-dependent covariance and the dominance of representativeness error, but also to the continued operation of this error-insensitive background correction pathway."

(5) In Lines 646-648, we refined the summary statement: "...the standard LETKF effectively constrained the simulated vertical structural shapes, while the integrated background correction successfully restored the magnitude of completely missed elevated black carbon (BC) plumes."

Comment NO.4: *In response to Minor comment 1, the authors clarify that the temporal localization scale $\sigma_t=0.01$ is expressed in units of assimilation time slots and that this value restricts the update to observations at the exact current time step, effectively excluding asynchronous observations within the assimilation window (Lines 274-278). I appreciate this clarification; however, I have one question regarding the description of the assimilation framework.*

The defining feature of a 4D-LETKF, compared with a 3D-LETKF, is that observations distributed within an assimilation window are compared with ensemble members at their corresponding observation times and their information is propagated through the time-evolving ensemble trajectories to constrain the analysis state. With the current temporal localization setting, observations separated from the analysis time receive negligible weight, meaning that the asynchronous observation component appears to be effectively disabled in the present experiments. This interpretation also appears inconsistent with Lines 329-330, where the authors attribute improved aerosol-field evolution to asynchronous assimilation. If asynchronous observations are effectively excluded by the chosen temporal localization, this statement should be reconsidered or clarified.

Also, If the present configuration intentionally uses only observations at the analysis time, then the assimilation behaves effectively as a 3D-LETKF for these experiments, despite being implemented within a 4D-LETKF framework, isn't it?

Response: Thank you for this clarification. We agree with the reviewer that the assimilation configuration used in the present experiments effectively corresponds to a synchronous, three-dimensional LETKF (3D-LETKF), although it is implemented within a 4D-LETKF framework. Specifically, the temporal localization scale is set to $\sigma_t=0.01$ in assimilation time-slot units, restricting the analysis update to observations assigned to the current analysis time slot. Observations in other time slots are effectively excluded, and the capability of the underlying framework to assimilate asynchronous observations is therefore not utilized in these experiments.

We acknowledge that our original description did not adequately distinguish the underlying framework from the configuration actually used in this study.

We also agree that the attribution of improved aerosol-field evolution to “asynchronous assimilation” was incorrect. We have removed relative sentences and revised the text to describe the evaluation of aerosol spatial distributions and temporal evolution without invoking asynchronous assimilation. The spreading of observational information through spatial localization and the subsequent propagation of updated aerosol fields through model integration should not be interpreted as evidence that asynchronous observations were assimilated.

Accordingly, we have clarified the synchronous configuration in Sections 3 and 3.3 and revised the terminology in the Abstract, section heading, figure caption, Results, and Conclusions. References to “4D-LETKF” are retained only when describing the underlying implementation framework, the configuration employed in the experiments is identified as LETKF. These revisions correct the methodological description and interpretation of the existing experiments and do not involve changes to the numerical implementation or simulation results.

Changes in Manuscript:

(1) In Lines 200-204, to distinguish the underlying 4D-LETKF framework from the synchronous LETKF configuration employed in this study, we replaced the section heading and opening description “3 Component-resolved 4D-LETKF system In this study, we developed a component-resolved four-dimensional local ensemble transform Kalman filter (4D-LETKF) data assimilation system with spatial and observational constraints for WRF-Chem (Fig. 2).” with “3 Component-resolved LETKF assimilation system In this study, we developed a component-resolved aerosol data assimilation system with spatial and observational constraints for WRF-Chem (Fig. 2). The system is implemented within a four-dimensional local ensemble transform Kalman filter (4D-LETKF) framework but is configured to perform LETKF analyses in the present experiments. Specifically, the prescribed temporal localization restricts the observations used in each analysis to the current analysis time slot, so the LETKF update effectively corresponds to a three-dimensional LETKF (3D-LETKF), as detailed in Section 3.3.”

(2) In Lines 280-285, to explicitly state that the present configuration excludes observations assigned to other time slots, we replaced “The temporal localization scale is in units of time slots (where one slot corresponds to the temporal interval of the assimilation window). Physically, this scale acts as a strict temporal filter. Specifically, setting this value to 0.01 means that only observations at the exact current time step are utilized for the assimilation update, effectively excluding any asynchronous observations within the assimilation window.” with “The temporal localization scale, $\sigma_t = 0.01$, is expressed in assimilation time-slot units, with Δt expressed in the same units. This setting acts as a strict temporal filter that only observations assigned to the current analysis time slot contribute to the analysis update, whereas observations in other time slots are effectively excluded. Consequently, although implemented within a 4D-LETKF framework, the

LETKF update in the present experiments operates in a 3D-LETKF configuration. The capability to assimilate asynchronous observations is not utilized.”

(3) In Lines 336-338, we removed the incorrect attribution of improved aerosol-field evolution to asynchronous assimilation and replaced “This step evaluates the system's capability to reproduce aerosol fields, highlighting how asynchronous assimilation improves the spatiotemporal evolution of aerosol distributions.” with “This independent evaluation assesses the ability of the assimilation system to reproduce aerosol spatial distributions and temporal evolution, including at locations and times not directly sampled by the assimilated CALIOP observations.”

(4) In Lines 21-22 of the abstract, to accurately identify the method used as synchronous LETKF, we replaced “four-dimensional local ensemble transform Kalman filter (4D-LETKF)” with “local ensemble transform Kalman filter (LETKF)”.

(5) We have revised the terminology in the figure captions and text to maintain consistency for the assimilation configuration used in this study. Specifically:

In Figure 2 caption, we replaced “4D-LETKF” with “LETKF”.

In Lines 231-232, we replaced “Building upon the 4D-LETKF framework, the efficacy of the assimilation system relies heavily on the precise definition of the state vector and its mapping to observation space.” with “The component-resolved assimilation requires an explicit definition of the aerosol state vector and its mapping to the observed component-specific extinction profiles.”

In Line 331, we replaced “This section presents a comprehensive evaluation of the 4D-LETKF aerosol component assimilation system...” with “This section presents a comprehensive evaluation of the LETKF aerosol component assimilation system...”

In Lines 356-357, we replaced “The implementation of the 4D-LETKF assimilation system (Figs. 3e-h) significantly enhances...” with “The component-resolved assimilation system (Figs. 3e-h) significantly enhances...”

Comment NO.5: *The authors state that the representativeness error ϵ_{rep} is derived from the standard deviation of sub-grid retrievals within each 25 km grid cell and that it typically dominates the observation-error budget, allowing the total observation error to realistically represent large uncertainties (Line 159-162). This point is central to the interpretation of the DA_obserr experiment in Section 5.3. However, the screening procedure described in Lines 147-150 may affect the magnitude of the estimated ϵ_{rep} .*

The authors remove pixel retrievals outside one standard deviation of the grid mean and retain only grid cells with $(N > 3)$ and coefficient of variation < 0.5 . If ϵ_{rep} is calculated from the retained retrievals, the estimated sub-grid variability will be reduced because the extreme values contributing to the representativeness error have already been removed. This is a separate potential source of underestimation from the spatial-clustering issue discussed in Lines 163-168.

Could the authors clarify whether ϵ_{rep} is calculated before or after these screening steps? If it is calculated after screening, please discuss how this filtering influences the estimated observation uncertainty and the statement that ϵ_{rep} ensures a realistic representation of high uncertainty. A

comparison of ϵ_{rep} before and after screening for representative grid cells would help quantify whether this effect is negligible or significant.

The revised manuscript has addressed the major concerns raised during the previous review and presents a valuable contribution to aerosol data assimilation and radiative impact assessment. The remaining issues mainly concern clarification of several methodological assumptions and ensuring that the interpretation of some key results is fully supported by the presented analyses. Addressing these points would improve the transparency and robustness of the manuscript.

Response: Thank you for this insightful observation. We confirm that the representativeness error (ϵ_{rep}) is calculated after the statistical screening steps.

Our rationale for this sequence is rooted in the characteristics of high-resolution satellite retrievals. In synergistic active-passive retrievals like the CALIOP-MODIS product, extreme pixel-level outliers frequently originate from retrieval artifacts, such as residual sub-pixel cloud contamination, complex surface albedo effects, or low signal-to-noise ratios, rather than genuine physical aerosol variability. If ϵ_{rep} were calculated prior to screening, these unphysical retrieval errors would be conflated with true sub-grid representativeness error, leading to an artificially inflated total observation error. This would excessively suppress the Kalman gain and limit the efficacy of the assimilation.

However, we fully agree with the reviewer that this strict filtering inevitably truncates the natural tails of the sub-grid aerosol distribution. By removing extreme values, the calculated ϵ_{rep} represents a conservative, lower-bound estimate of the true physical sub-grid variability. We have added a discussion in Section 2.2.1 to explicitly clarify this calculation sequence and its implications for the estimated observation uncertainty. Furthermore, this context makes the DA_Obserr sensitivity experiment in Section 5.3 even more relevant: increasing the baseline error in DA_Obserr tests sensitivity to larger prescribed observation uncertainties, but does not directly quantify or correct the reduction in estimated sub-grid variability caused by screening. This interpretation is further limited by the continued operation of the bypass mechanism.

Changes in Manuscript: In Lines 162-166, to clarify the calculation of the representativeness error and acknowledge the impact of the screening procedure, we added the following discussion: “It should be noted that ϵ_{rep} is calculated after the statistical screening steps. This filtering is essential to prevent unphysical retrieval artifacts, including residual cloud contamination and surface reflectance issues, from artificially inflating the observation error. However, it inevitably truncates the natural tails of the sub-grid aerosol distribution. Consequently, the calculated ϵ_{rep} represents a conservative, lower-bound estimate of the true physical sub-grid variability.”