Response to Reviewer #1 for Geoscientific Model Development:
A revised version of manuscript gmd-2024-148
By Liu et al.

We sincerely thank Reviewer #1 for insightful and constructive feedback on the revised version of our manuscript. We have carefully considered each comment and made further revisions accordingly. The notes below address each comment in detail. Please note that the reviewer's comments are shown in bold type and our responses are in plain type.

Reviewer #1

Review Decision: Minor Revisions
Manuscript type: *Development and technical papers*

General Comments:
The revised manuscript is a huge improvement from the original version in various aspects, including readability, clarity, and coherence. The authors' effort on revising the manuscript is greatly applauded. I've also found the revised manuscript to be very educational in guiding the readers on developing a coupled aerosol-meteorology 4DVar assimilation system step by step with important details. After a thorough review, I only have a few minor comments that I'd recommend the authors to address before the manuscript could be considered for publication.
Response: We sincerely thank the reviewer for the encouraging and constructive feedback. We are very grateful for the positive assessment of our revised manuscript, especially the comments on the improvements in readability and clarity, as well as the educational value of the work. We truly appreciate the recognition of our efforts in developing and describing the coupled aerosol-meteorology 4DVar assimilation system.
We have carefully addressed all the remaining minor comments in the revised manuscript, as detailed below. We also sincerely appreciate the reviewer's time and effort in thoroughly reviewing the revised version and offering valuable suggestions. We hope that the modifications and corrections we have made will meet the standards for publication.

Minor comments:

1. **What is a batch test? My guess is that it is an integrated set of tests that involves many cases, but I am not 100% sure since it sounds more like a test used in software development. Is batch test a common terminology used in operational NWP centers? I wondered whether it is okay to assume that the readers understand what a batch test is. I would rather use more words in plain language to describe what the author intends to do next instead of a terminology that might not be universally recognized.**
   Response: We sincerely appreciate the insightful comments and valuable suggestions. In our study, the term "batch test" refers to a series of cycling assimilation experiments conducted over a certain period (e.g., three months), in which the CMA-GFS-AERO 4D-Var system is run at 6-hour assimilation cycles to evaluate its performance in a realistic forecast-assimilation framework. We acknowledge that "batch test" may not be a widely

recognized term in the NWP community. To enhance clarity, we have revised the manuscript to replace “batch test” with “cycling assimilation experiments”. The specific modifications are as follows:

Abstract:

“…This study focuses on the theoretical architecture and practical implementation of the system, the detailed analysis of a series of cycling assimilation experiments will be described in part 2 of this paper.”

4 Model setup

“…The detailed analysis of the cycling assimilation experiments will be further elaborated in part 2 of this paper.”

5.3 Case study on BC and atmosphere assimilation

“…In the future, we will conduct cycling assimilation experiments using CMA-GFS-AERO 4D-Var to…”

6 Conclusions

“…We intend to explore the impact of assimilating surface BC observations on the forecast fields of BC and atmospheric variables through cycling assimilation experiments…”

2. **Page 1, Line 13: chemical should be replaced by aerosol.**
   Response: Thanks for pointing this out. We have revised “chemical” to “aerosol” as suggested.

3. **Pages 2-3, Lines 36-66: this paragraph is too long. Please consider breaking it into two paragraphs. I recommend that a new paragraph could be started from line 58, focusing on the role of background error covariance in high-dimensional CCMM system.**
   Response: Thanks for the valuable comment. We have revised this section by splitting it into two paragraphs, starting a new one from line 58 as suggested. The new paragraph emphasizes the role of background error covariance in high-dimensional CCMM systems, improving the clarity of the structure.

4. **Page 4, Line 103: chemical should be replaced by aerosol.**
   Response: Thanks again for the valuable suggestion. We have made the change accordingly.

5. **Page 4, Lines 104-105: it should be five sections, not four, as the remaining sections are sections 2, 3, 4, 5, and 6 (total of 5).**
   Response: Thanks for the careful review. We have corrected this mistake and changed “four” to “five” as suggested.

6. **Pages 4-5, Lines 123-125: aren’t those water matter species produced by the double-moment cloud microphysics scheme also considered state variables of the CMA-GFS NLM?**
   Response: Thanks for the insightful comment. In the previous text, we did not explicitly list those water matter species produced by the double-moment cloud microphysics scheme as state variables, primarily because, in operational practice, they are not preserved in the initial fields used for official forecasts.

Specifically, although the analysis fields (produced at 03, 09, 15, and 21 UTC) include these water matter species, the initial conditions used in official forecasts (00, 06, 12, and 18 UTC) are derived from 3-hour forecasts based on those analysis fields. After the 3-hour forecast integration, only a subset of model variables (wind components, potential temperature, specific humidity, and non-dimensional pressure) is saved due to the limitation of I/O time and disk space, while the fields of all hydrometeor contents and cloud cover are discarded (Ma et al., 2021). These variables are then regenerated during model spin-up in the forecast phase. Therefore, these hydrometeor species are generally not considered part of the state variables in the context of the CMA-GFS NLM.

Ma, Z., Zhao, C., Gong, J., Zhang, J., Li, Z., Sun, J., Liu, Y., Chen, J., and Jiang, Q.: Spin-up characteristics with three types of initial fields and the restart effects on forecast accuracy in the GRAPES global forecast system, Geosci. Model Dev., 14, 205-221, https://doi.org/10.5194/gmd-14-205-2021, 2021.

7. **Page 8, Line 236: Upon reading this, I couldn't help but wondering that isn't AERO-BC also 1-D modules? And then, I found the authors actually hinted this at page 7 line 198. I suggest making this point clearer upfront. Also, I am puzzled by the description about "1-D modules with fixed latitude and longitude coordinates"… if it has lat/lon coordinates, then isn't it 2-D?**

   Response: We sincerely appreciate the insightful comments and valuable suggestions. We apologize for any confusion caused by our previous description. The AERO-BC module and its tangent linear (TL) and adjoint versions are designed as one-dimensional (1-D) column modules. Each module operates independently on a single vertical column corresponding to a fixed horizontal location (i.e., fixed latitude and longitude). In this context, "1-D modules with fixed latitude and longitude coordinates" refers to the fact that each 1-D module processes only the vertical profile at a single horizontal grid point, without involving horizontal interactions within the module itself. To form a global 3-D system, these 1-D modules are applied independently at every horizontal grid point, and the interface programs are responsible for passing 3-D meteorological and emission fields from CMA-GFS to each of these 1-D modules. This 3-D extension is also implemented for their TL and adjoint versions.

   Following the reviewer's suggestion, we have revised Section 3.1 to make this point clearer. The updated description now reads:

   "…The AERO-BC module is designed as one-dimensional (1-D) column module, which operates at individual vertical columns corresponding to fixed horizontal locations (i.e., fixed latitude and longitude). In the integration of AERO-BC with CMA-GFS, the interface programs transfer meteorological parameters (e.g., temperature, wind, and humidity) from CMA-GFS to AERO-BC, extend the spatial dimension from 1-D to 3-D, and read emissions for AERO-BC…"

   We have also revised the corresponding sentence in Section 3.2 for clarity:

   "The TL and the adjoint of AERO-BC are 1-D column modules, meaning they operate independently at each fixed horizontal grid point (i.e., fixed latitude and longitude), with vertical variation only. To extend them to 3-D, the tangent linear and the adjoint of the interface programs were also constructed…"

8. **Page 13, Line 362: "at the initial of the assimilation window," it would be better to put "only" right after assimilation window to highlight this very special scenario.**
Response: Thanks for the valuable suggestion. We have revised this sentence by placing "only" right after "assimilation window" to highlight this special scenario.

9. **Page 22, Line 575 and Table 3: I think there is a need to explicitly explain how the two experiments DA_MET_then_BC and DA_MET_BC_simult are conducted differently in terms of their workflows. For DA_MET_then_BC, does that mean the CMA-GFS-AERO 4DVar system is run two times with the same setup, except for the observations being assimilated? In the first run, only the operational meteorological observations are assimilated and then the resulting analysis is used as background for the second run, where only the BC surface observations are assimilated. Is that right?**
Response: Thanks for pointing this out. Yes, the reviewer's understanding is correct. In DA_MET_then_BC, the CMA-GFS-AERO 4D-Var system is indeed run twice. First, we assimilate only operational meteorological observations, and then the resulting analysis is used as the background field for a second assimilation run that includes only BC surface observations. In contrast, DA_MET_BC_simult assimilates both operational meteorological and BC surface observations simultaneously within a single assimilation run. We have revised the text to explicitly describe their workflows, which now reads as follows:
"…These experiments are listed in Table 3…<u>For the DA_MET_then_BC experiment, the CMA-GFS-AERO 4D-Var system was executed twice sequentially within the same assimilation window. In the first step, only operational meteorological observations were assimilated, and the resulting analysis was used as the background field for the second step, in which only BC surface observations were assimilated. Except for the observational datasets, the model configurations and assimilation settings in both steps remained identical. This two-step procedure allows us to separate the effect of BC observations from the influence of meteorological observations and their associated background adjustment, thereby facilitating a clearer attribution of the BC assimilation impact. In contrast, DA_MET_BC_simult assimilated both operational meteorological observations and BC surface observations simultaneously within a single 4DVar run. This one-step assimilation strategy allows all observations to jointly influence the analysis field, reflecting the integrated effect of both meteorological and BC observations</u>…"

10. **Page 23, Line 602: I am not sure what is the purpose of this sentence.**
Response: Thanks for the comment. We agree that the original sentence was unclear. The purpose was to emphasize that, in Fig. 10, we aim to isolate the impact of BC assimilation. Therefore, we show the difference between DA_MET_then_BC and DA_MET, and between DA_MET_BC_simult and DA_MET, in order to remove the contributions from the assimilation of meteorological observations. Considering this comment together with Comment #11, we have revised the text as follows:
"…Figure 10 shows the analysis increments of temperature, pressure, east-west component of horizontal wind, and relative humidity at the first model layer, resulting from BC assimilation in DA_BC, DA_MET_then_BC, and DA_MET_BC_simult. <u>It is worth noting that in DA_BC,</u>

only BC observations are assimilated, so the analysis increments of atmospheric variables purely reflect the response to BC. In contrast, both DA_MET_then_BC and DA_MET_BC_simult assimilate BC and meteorological observations, and thus their analysis increments include the combined effects of both types of observations. To isolate the influence of BC assimilation alone on atmospheric variables, and under the assumption that the contribution of meteorological observations is comparable between DA_MET_then_BC/DA_MET_BC_simult and DA_MET, we calculated the differences between the analysis increments of these experiments and those from DA_MET, which assimilates only meteorological observations. This subtraction effectively removes the contributions from meteorological observations, allowing the resulting increments to be attributed solely to the assimilation of BC observations. In this way, a more direct and fair comparison can be made with DA_BC…"

11. **Page 24, Lines 626-627: ok, but why not just show the analysis increment of DA_MET_then_BC and have it compared with those from DA_BC? I still have trouble understanding using the "difference of analysis increment of two experiment" to compare with the "analysis increment of one experiment". If the concern is because the analysis increments of DA_MET_then_BC and DA_BC includes not only the impact of meteorological observations but also the impact from using different background fields (the nature of DA_MET_then_BC), and the authors would like to isolate the impact of BC from all other factors, then it has to be stated clearly to justify such comparison (which also echoes my minor comment 9). However, I am not sure whether the same reasoning applies to the case for comparing the difference of analysis increment of DA_MET_BC_simult and DA_MET with the analysis increment of DA_BC.**

    Response: We sincerely appreciate the reviewer's insightful comments. The purpose of comparing the difference in analysis increments between DA_MET_then_BC and DA_MET (DA_MET_then_BC – DA_MET) with the increments from DA_BC is to isolate the net impact of BC observations on atmospheric variables, excluding the influence of meteorological observations. In the DA_BC experiment, only BC observations are assimilated, so the analysis increments of atmospheric variables purely reflect the response to BC.

    In contrast, both DA_MET_then_BC and DA_MET_BC_simult assimilate BC and meteorological observations, and thus their analysis increments include the combined effects of both types of observations. By subtracting DA_MET (which only assimilates meteorological observations), we remove the meteorological component, thereby isolating the contribution of BC assimilation. This approach enables a more direct and fair comparison with DA_BC, where only BC observations are assimilated.

    Regarding the reviewer's question on whether the same reasoning applies to DA_MET_BC_simult, we confirm that it does. Although DA_MET_BC_simult assimilates BC and meteorological observations simultaneously, subtracting the analysis increments of DA_MET from those of DA_MET_BC_simult still allows us to isolate the impact of BC assimilation alone-under the assumption that the contribution of meteorological observations is comparable between DA_MET_BC_simult and DA_MET.

    We have clarified this in the revised manuscript as follows:

    "…Figure 10 shows the analysis increments of temperature, pressure, east-west component of horizontal wind, and relative humidity at the first model layer, resulting from BC assimilation in DA_BC, DA_MET_then_BC, and DA_MET_BC_simult. It is worth noting that in DA_BC, only BC observations are assimilated, so the analysis increments of atmospheric variables purely reflect the response to BC. In contrast, both DA_MET_then_BC and DA_MET_BC_simult assimilate BC and meteorological observations, and thus their analysis increments include the combined effects of both types of observations. To isolate the influence of BC assimilation alone on atmospheric variables, and under the assumption that the contribution of meteorological observations is comparable between DA_MET_then_BC/DA_MET_BC_simult and DA_MET, we calculated the differences between the analysis increments of these experiments and those from DA_MET, which assimilates only meteorological observations. This subtraction effectively removes the contributions from meteorological observations, allowing the resulting increments to be attributed solely to the assimilation of BC observations. In this way, a more direct and fair comparison can be made with DA_BC…"

Response to Reviewer #2 for Geoscientific Model Development:
A revised version of manuscript gmd-2024-148
By Liu et al.

We sincerely thank Reviewer #2 for insightful and constructive feedback on the revised version of our manuscript. We have carefully considered each comment and made further revisions accordingly. The notes below address each comment in detail. Please note that the reviewer's comments are shown in bold type and our responses are in plain type.

Reviewer #2

MS type: Development and technical paper
Iteration: Revised submission
2025-03-10

General Comments
I in general welcome the changes in the reviewed version of the manuscript as they provide more clarification and enter a little bit more into explaining the results. I however still have concerns *before* being able to recommend the paper for publication.
Response: We sincerely thank the reviewer for the careful review of our revised manuscript and for acknowledging the improvements we have made. We greatly appreciate the reviewer's continued engagement and constructive feedback, which are invaluable in further improving the clarity and scientific rigor of our work.
In the revised version, we have carefully addressed the remaining concerns and added further clarifications where necessary. Detailed point-by-point responses are provided below. We also sincerely appreciate the reviewer's time and effort in thoroughly reviewing the revised version and offering valuable suggestions. We hope that the modifications and corrections we have made will meet the standards for publication.

At first, I am now confused by the modified pictures in Figures 9 and 10. Following the suggestion of Reviewer 1, the authors have renamed the labels of the experiences in a more explicit manner (Table 3), which is good. However, my understanding is that these panels refer to exactly the same experiments. I would thus expect the pictures to show exactly the same fields based on the same data. Why then are the panels displaying changed fields of increments and differences of increments differing from the original version of the paper? I have no explanation for that myself, and I see none in the answers to the reviewers nor in the paper.
Response: We sincerely apologize for omitting an explanation regarding the updates to Figures 9 and 10 during the revision process and for any confusion this may have caused. We greatly appreciate the insightful comments and careful examination of the figures.
The differences between the revised and original Figures 9 and 10 are due to a modification in the radius of influence for BC observations. Specifically, Reviewer #1 pointed out the importance of considering the station type when determining the radius of influence for BC observations in the calculation of BC observation error. According to Table 3 of Elbern et al. (2007), the radius of

influence differs for urban, rural, and remote stations. Initially, we used a uniform 10 km radius for all 32 CAWNET stations, but after reviewing this suggestion, we realized that this approach was inappropriate. In the revised manuscript, we adopted station-type-specific radii: 2 km for urban stations, 10 km for rural stations, and 20 km for remote stations, following Table 3 of Elbern et al. (2007).

This change is explicitly described in Section 3.3.2 of the revised manuscript as follows:

"The BC observation data were collected from 32 stations (Guo et al., 2020), including 11 urban, 17 rural, and 4 remote stations…"

"…and L is the radius of influence of a BC observation. According to Elbern et al. (2007), L was set to 2 km, 10 km, and 20 km for urban, rural, and remote stations, respectively…"

Since we updated the radius of influence of BC observations, all experiments in Section 5.3 have been redone using the revised settings. Consequently, the corresponding figures and text have been updated to reflect the new results.

We once again appreciate the reviewer's careful examination of Figures 9 and 10 in our revised manuscript.

Regarding Figure 7, can the authors really be explicit and state in the paper whether or not the analysis increments for temperature systematically are positive when only one single BC observation is assimilated? I actually see no reason for it being so, but I would like the authors to provide full clarity on this, based on their experimental results. If their answer is yes (this is systematic), an explanation why would be definitely welcome.

Response: We sincerely thank the reviewer for pointing out the need for clarification. We apologize for any confusion caused by our earlier description. The temperature analysis increment is not systematically positive when a single BC observation is assimilated. Instead, it depends on several factors, including the BC observation innovation, the location of the observation, and the meteorological conditions during the assimilation time window.

In the specific case shown in Figure 7, the BC observation innovations are negative (i.e., -9.5 $\mu g/m^3$ at 0600 UTC and -9.0 $\mu g/m^3$ at 0900 UTC), indicating that the background BC concentration is lower than the observed values. Assimilation of these observations increases the BC concentrations in the analysis, which, under the prevailing meteorological conditions, leads to positive temperature increments near the observation site. As explained in Section 3.3.5, the coupling between BC and atmospheric variables within the system allows this type of feedback to occur.

Therefore, the positive temperature increments observed in Figure 7 are a result of this specific experiment setup and should not be interpreted as a systematic behavior. We have revised the manuscript to make this point clearer. The revised text now reads as follows:

"Figure 7 depicts the analysis increments of temperature at the first model level at the beginning time of the assimilation time window (0300 UTC), with the BC observation placed at 0600 and 0900 UTC, respectively. In this specific case, the analysis increments of temperature are positive, with the value of about 0.02 K near the observation location, when the BC observation is placed at 0600 and 0900 UTC. The temperature analysis increment depends on several factors, including the BC observation innovation, the location of the observation, and the meteorological conditions during the assimilation time window. Here, the positive analysis increments of temperature may be due to the fact that the BC observation innovations at 0600 and 0900 UTC are negative (-9.5

μg/m³ at 0600 UTC and -9.0 μg/m³ at 0900 UTC), indicating that the background BC concentration is lower than the observed values. Assimilation of these observations increases the BC concentrations in the analysis, which, under the prevailing meteorological conditions, leads to positive temperature increments near the observation site. As explained in Section 3.3.5, the coupling between BC and atmospheric variables within the system allows this type of feedback to occur."

We once again appreciate the reviewer's insightful comment and valuable suggestions.

The statements in lines 748-750 are either too strong or totally obscure, and likely a bit of both. I simply cannot understand why the experimental results in Section 5 enable one to state that "meteorological observations can help constrain the uncertainty introduced by BC observations on atmospheric variables, thereby improving the reliability of the assimilation results". My own strong reaction could be due to the fairly unclear formulation in the text: why or how do BC observations "introduce uncertainty" on atmospheric variables ? What is meant by this formulation ? How is "reliability" defined in this context ?

Likewise, I actually dispute that "these results *demonstrate* the successful implementation of the newly developed CMA-GFS-AERO 4D-Var system".

What Is meant by "successful" here ? Should the authors wish to provide proof (a demonstration) of the successful implementation in terms of exactitude of *expected* results, then the only way would be to use their coupled assimilation system in a *fully controlled* modus operandi. One possibility would be OSSE with fully simulated observations of all types, and using prescribed error statistics in the system error covariances and for perturbing the simulated observations from a ground truth. In such experimental mode, the output results should follow some statistical results that separately could be derived from the expected posterior probability density functions (i.e. independently from the experiments).

As this is not what the authors have done, and I am not claiming they should do this, my understanding is that the authors have proven that their 4D-VAR system produces "credible and realistic-looking" results: "credible" in the sense that the results could lead to further explanations with the knowledge the authors have, and "realistic-looking" in the sense that the increments display realistic structures and amplitudes case-by-case.

I therefore strongly invite the authors to consider reformulating their statements at this location in Section 5 (around lines 748-750), and elsewhere where relevant such as in the Conclusion (lines 819-820).

Response: We sincerely appreciate the insightful comments and valuable suggestions. We fully agree that some of the statements in the original manuscript were too strong or not clearly formulated. In the revised version, we have removed ambiguous or overstated terms such as "introduce uncertainty," "improve reliability," and "successful implementation." Instead, we now emphasize what the experimental results actually suggest: the presence of meteorological observations during assimilation may provide additional constraints on the adjustment of atmospheric fields, potentially reducing the degree to which the assimilation of BC observations alone can alter the atmospheric state. In this way, the integration of meteorological observations helps stabilize the adjustment process, supporting more consistent and interpretable assimilation results.

We also agree that the term "successful implementation" may imply a rigorous demonstration of the system's exactitude against a defined ground truth, such as in an Observing System Simulation Experiment (OSSE) framework with prescribed error statistics and known truth states. Since our study is not based on such a controlled setup, we acknowledge that using the term "successful implementation" could be misleading. In the revised manuscript, we have replaced this term with a more accurate description of what the current experiments actually demonstrate: the CMA-GFS-AERO 4D-Var system has been technically implemented and is able to produce credible assimilation increments when assimilating BC and meteorological observations. The increments exhibit realistic structures and amplitudes, which suggests that the system performs as intended under the current configuration and available observations.

We have updated the corresponding statements in Section 5.3 and the Conclusion as follows:

Section 5.3:

"…This suggests that the presence of meteorological observations during assimilation may provide additional constraints on the adjustment of atmospheric fields, potentially reducing the degree to which the assimilation of BC observations alone can alter the atmospheric state. In this way, the integration of meteorological observations helps stabilize the adjustment process, supporting more consistent and interpretable assimilation results. Moreover, the four experiments demonstrate that the CMA-GFS-AERO 4D-Var system has been technically implemented and is able to produce credible analysis increments in both BC and atmospheric fields. These increments display realistic spatial structures and amplitudes, indicating that the system performs as intended under the current configuration and available observations. These results offer practical evidence of the system's functionality and its potential utility for exploring the feedback of BC data assimilation on meteorological forecasts. In the future..."

Conclusion:

"…This demonstrates that the newly developed CMA-GFS-AERO 4D-Var system has been technically implemented and is capable of producing credible assimilation outcomes, highlighting its potential as a useful tool for exploring the feedback of BC data assimilation on meteorological forecasts…"

We once again appreciate the reviewer's guidance, which helped improve the precision and clarity of our manuscript.

I would recommend some English proofreading of the paper.

Response: Thanks for the helpful comment. We have thoroughly proofread the manuscript to enhance the clarity and correctness of the English. We believe that the revised version is now clearer and more readable.

Specific Comments & Typos

line 408: "Specifically, if the BC observation is assumed to take place at the initial of the assimilation window," => "Specifically, if the BC observation is assumed to take place at the ~~initial~~ beginning of the assimilation window,"

(same suggestion holds for other locations in the text such as for instance in the caption of Figure 7)

Response: Thanks for the valuable suggestion. We have revised "at the initial of the assimilation window" to "at the beginning of the assimilation window" in this line. We have also carefully checked the entire manuscript and made similar corrections in other relevant locations, such as the caption of Figure 7.

Figure 10 and everywhere in the paper where "pressure" is referred to: what is exactly this field ? is it "surface pressure" (ie following elevation), is it "mean sea level pressure" ? is it "pressure at the first model level" ? or any other definition of "pressure" ?

Response: We sincerely appreciate the reviewer's insightful comments. The term "pressure" in Figures 8 and 10 and the corresponding text refers to "pressure at the first model level". This has been explicitly stated in the manuscript, as shown in the following examples:

"Figure 8 shows the analysis increments of pressure, east-west component of horizontal wind, and relative humidity at the first model level…"

"…Figure 10 shows the analysis increments of temperature, pressure, east-west component of horizontal wind, and relative humidity at the first model layer…"

To further clarify this, we have revised the manuscript by explicitly specifying "pressure at the first model level" when it is first introduced:

"Figure 8 shows the analysis increments of pressure at the first model level, as well as east-west component of horizontal wind and relative humidity at the same level…"

Regarding the other occurrences of "pressure" in the manuscript, such as the non-dimensional pressure (π) and the unbalanced Exner pressure (π_u), their definitions are already provided in the text, and no further clarification is necessary. We once again thank the reviewer's valuable comments.

Response to Reviewer #3 for Geoscientific Model Development:
A revised version of manuscript gmd-2024-148
By Liu et al.

We sincerely thank Reviewer #3 for insightful and constructive feedback on the revised version of our manuscript. We have carefully considered each comment and made further revisions accordingly. The notes below address each comment in detail. Please note that the reviewer's comments are shown in bold type and our responses are in plain type.

Reviewer #3

The authors have addressed the reviewers' comments and the manuscript has been improved. However, I still have some concerns about the results in Section 5.3. Please find the detailed comments in the attached pdf.

Response: We sincerely appreciate the reviewer's continued efforts in reviewing our manuscript and for pointing out further concerns regarding the results in Section 5.3. We have carefully reviewed the detailed comments provided in the attached PDF and have revised the manuscript accordingly to address each of them. Please find our point-by-point responses below. We hope that our revisions and explanations have adequately addressed the remaining concerns.

1. **Page 3, Line 68: Ensemble Kalman smoothers also use an assimilation window.**

   Response: We sincerely appreciate the insightful comment. We acknowledge that ensemble Kalman smoothers (EnKS) also utilize an assimilation window. To ensure clarity and avoid any misunderstanding that only 4D-Var employs an assimilation window, we have revised the text as follows:

   "…In high-dimensional problems, the limited number of samples may not be able to fully capture all the error characteristics, resulting in inaccuracies in the estimation of background error covariance. Although ensemble Kalman smoothers (EnKS) extend the EnKF framework by incorporating an assimilation window to leverage temporal observational information, they remain constrained by similar limitations in ensemble size. In contrast, 4D-Var explicitly integrates both the complete observational dataset and the full model dynamics within the assimilation window to constrain state evolution, rather than relying solely on ensemble statistics. This generally allows 4D-Var to achieve higher accuracy in high-dimensional problems by making better use of both observational data and model constraints, leading to more precise state estimation…"

2. **Page 3, Line 70: Does the CMA-GFS use an ensemble of variational data assimilations? If so, the background error covariances can be estimated in a similar way as in the EnKFs.**

   Response: We sincerely appreciate the insightful comment. The current version we used in this work, CMA-GFS v4.0, does not employ an ensemble of variational data assimilations; it strictly follows the 4D-Var framework. In later developments, the CMA-GFS En4DVar system has been introduced (e.g., in CMA-GFS v4.2), which incorporates ensemble-based methods to estimate background error covariances.

In coupled chemistry-meteorology models, estimating the cross-variable component of the covariance remains a significant challenge, regardless of whether a 4D-Var or ensemble-based approach is used. This challenge arises from the complex interactions between meteorological and chemical variables, making it difficult to accurately represent their covariance. In future work, we plan to explore hybrid data assimilation approaches within the CMA-GFS En4DVar framework to improve the estimation of background error covariances in coupled systems.

We once again appreciate the reviewer's insightful comment and valuable suggestion.

3. **Page 4, Line101: This sentence sounds like that the surface temperature is not part of the 3D temperature field.**

   Response: We appreciate the reviewer's careful reading. To avoid any potential misunderstanding, we have revised the sentence to clarify that the surface temperature is part of the overall atmospheric temperature field. The revised text now reads as follows:

   "…BC is also the main optically absorbing component of atmospheric aerosols, effectively absorbing solar radiation in the visible to infrared wavelength range, <u>thus affecting the temperature field throughout the atmosphere, including the surface temperature</u>…"

4. **Page 5, Lines 133-134: This is not a must, but it would be better if the authors could add some simple justification for the physical parameterization schemes chosen.**

   Response: Thanks for the valuable suggestion. We have added a brief justification for the selection of physical parameterization schemes. Specifically, we have stated that these schemes are consistent with those used in the operational application of CMA-GFS v4.0 and have been proven to perform well in global numerical weather prediction. The revised text now reads as follows:

   "…The physical parameterization schemes used in this work <u>are consistent with those adopted in the operational application of CMA-GFS v4.0, which have been proven to perform well in global numerical weather prediction. The selected schemes</u> mainly include..."

5. **Page 6, Section 2.2: The authors do not use a consistent rule for denoting matrices. For example, If bolded symbols are for matrices, then the matrix U should also be bold. I would suggest also using bolded but lowercase symbols for vectors. In addition, the symbols for matrices and vectors should not be italicized.**

   Response: We sincerely appreciate the reviewer's careful reading and valuable suggestions. We have revised the manuscript accordingly to ensure consistency in notation: matrices (e.g., **U**, **B**, **R**) are now represented using bold uppercase letters, and vectors (e.g., **w**, **d**, **x**) are represented using bold lowercase letters, as suggested. Additionally, we clarify that in the manuscript, italic and non-italic bold uppercase letters may appear simultaneously: the italic bold symbols denote nonlinear model operators, while the non-italic bold symbols represent their corresponding tangent linear versions. This notation is consistent with conventions commonly used in variational data assimilation literature.

6. **Page 11, Line 305: This is a follow-up question related to Reviewer #3's comment on the**

weights. Estimating the weights based on observations (if possible) rather than on model forecasts seems to be a better idea to me.

Response: We sincerely appreciate the reviewer's insightful follow-up question. We fully agree that estimating the size bin weights (ω^n) based on observations, if available, could potentially improve the accuracy. However, currently available BC observations do not provide the detailed size-resolved distribution of BC. All observational datasets available to us only include the total mass concentration (unit: μg/m³) without distinguishing between size bins, as also mentioned in Section 3.3.1. As a result, a single observation of BC mass concentration cannot be used to determine the weights of the six size-bin variables. Therefore, it is technically infeasible to derive ω^n directly from observations.

Instead, we estimate ω^n based on the distribution of ψ_{bc}^n in the background field, which is generated by the CMA-GFS-AERO model. Our analysis shows that the spatial variability of ω^n over China is relatively small, indicating that using domain-averaged model-based weights does not significantly affect the mass conservation or the quality of the reconstructed ψ_{bc}^n field.

In future work, we agree that if high-resolution size-resolved BC observational data become available, they could be incorporated to refine the estimation of ω^n and potentially improve the overall performance of the system.

We once again appreciate the reviewer's valuable suggestion.

7. **Page 12, Line 338: As the authors stated, representativeness errors are part of the observation error. I understand that this kind of error can be caused by the interpolation, but I do not understand why it can be caused by the forecast forward model.**

Response: Thanks for the valuable comment. Upon further consideration, we realize that our previous revision may have introduced some confusion. The reviewer is absolutely right that representativeness errors are part of the observation error, and they are typically associated with the observation operator, rather than the forecast model itself.

To clarify this point, we have revised the sentence in the manuscript as follows:

"…Representativeness errors reflect the inaccuracies in the observation operator and in the interpolation from the model grid to the observation location…"

We sincerely appreciate the reviewer's insightful feedback, which helped us improve the clarity and accuracy of the manuscript.

8. **Page 15, Lines 408-409: I understand the point of the authors, but I think more care should be taken when writing. Observations of atmospheric variables in the window may affect the model control variable, BC mass concentration. Therefore, in a cycling environment, there is a difference to the 3D-Var system.**

Response: Thanks for the valuable comment. We agree that the original sentence was too absolute and could be misleading. In the revised manuscript, we have clarified that our discussion refers to a single assimilation window without cycling. Specifically, under the assumption that the BC observation is assimilated at the beginning of the assimilation window only and cycling is not considered, the 4D-Var system behaves similarly to the 3D-Var system. We fully agree that in a cycling data assimilation environment, where atmospheric observations are also assimilated, feedback from the atmosphere to BC can occur through model dynamics, which may affect BC concentration. Therefore, the 4D-Var system is indeed different from a standard 3D-Var system in such cases.

The revised text now reads as follows:

"…Specifically, if the BC observation is assumed to take place at the beginning of the assimilation window only, <u>and under the assumption of a single, non-cycling assimilation window, the 4D-Var assimilation behaves similarly to the 3D-Var assimilation.</u> In this case…"

9. **Page 15, Lines 415-416: Again, caution is needed. It is better to refer the readers to the functions in Section 2.2. The matrix $\mathbf{MBM}^T$ is the background error covariance matrix at an observation time. But here what we need should be cross-time error covariances, something like $\mathbf{MB}$.**

   Response: We thank the reviewer for the helpful suggestion. To avoid any potential misunderstanding, we have revised the text as follows:

   "…If the BC observation is assumed to take place at the middle and the end of the assimilation window, $\mathbf{B}$ evolves within the assimilation time window through the TLM $\mathbf{M}_{0\to i}$, obtaining the implicit background error covariance matrix $\mathbf{M}_{0\to i}\mathbf{B}\mathbf{M}_{0\to i}^T$ <u>at the observation time.</u> $\mathbf{M}_{0\to i}\mathbf{B}\mathbf{M}_{0\to i}^T$ includes the cross-covariances information of BC and atmospheric variables, and can realize the feedback of the BC observation to the atmospheric variables through the CMA-GFS-AERO ADM $\mathbf{M}_{0\to i}^T$, further producing analysis increments of atmospheric variables. <u>In other words, the distribution of the analysis increment at the observation time is determined by the cross-time error matrix $\mathbf{M}_{0\to i}\mathbf{B}$."</u>

10. **Page 24, Section 5.3: It will be easier for the readers to follow if each point in Section 5.3 is described and explained in a separate paragraph.**

    Response: We thank the reviewer for the helpful suggestion. In the revised manuscript, we have restructured Section 5.3 by dividing the discussion into separate paragraphs, each focusing on a specific point or aspect of the analysis. This improves the clarity and readability of the section and makes it easier for readers to follow the logic and findings of the case study.

    Specifically, Section 5.3 now consists of the following seven paragraphs:

    (1) Overview of the experimental setup and design.
    (2) Presentation of BC analysis increments.
    (3) General introduction to the impact of BC assimilation on analysis increments of atmospheric variables (introduction of Figure 10).
    (4) Detailed description of the impact of BC assimilation on analysis increments of atmospheric variables in the DA_BC experiment.
    (5) Detailed description of the impact of BC assimilation on analysis increments of atmospheric variables in the DA_MET_then_BC experiment and its comparison with DA_BC.
    (6) Detailed description of the impact of BC assimilation on analysis increments of atmospheric variables in the DA_MET_BC_simult experiment and its comparison with DA_BC and DA_MET_then_BC.
    (7) Concluding remarks on the results of all four experiments.

We believe that this new structure better aligns with the reviewer's suggestion and significantly improves the flow and understanding of the analysis.

11. **Page 25, Line 642: Table 3?**
Response: Thanks for the careful review. We have corrected "Table 1" to "Table 3".

12. **Page 26, Line 675: Did the authors forget to implement their response to Comment 50 from the first reviewer?**
Response: We thank the reviewer for pointing out this oversight. In our response to Comment 50 from the first-round review, we clarified that the analysis increments presented in Figure 9 are valid at the beginning of the assimilation window, which is standard practice in 4D-Var. However, we acknowledge that this clarification was only made in the response letter and not explicitly included in the manuscript. To avoid any confusion for readers unfamiliar with this convention, we have now added a clarification in the revised manuscript as follows:
"Figure 9 presents the analysis increments of BC at the first model layer from the DA_BC, DA_MET_then_BC, and DA_MET_BC_simult experiments. These analysis increments are valid at the beginning of the assimilation window, as is standard in 4D-Var systems…"
We appreciate the reviewer's careful attention to detail.

13. **Page 26, Line 675: Why the new results are different to the previous ones, given that the experimental design are not changed?**
Response: We appreciate the reviewer's question regarding the differences in the new results in Figure 9. We sincerely apologize for omitting an explanation of the updates to Figure 9 and related content during the revision process.
Indeed, the overall experimental design remains unchanged. The differences between the new results and the previous ones are due to a modification in the radius of influence for BC observations. Specifically, Reviewer #1 pointed out the importance of considering the station type when determining the radius of influence for BC observations in the calculation of BC observation error. According to Table 3 of Elbern et al. (2007), the radius of influence differs for urban, rural, and remote stations. Initially, we used a uniform 10 km radius for all 32 CAWNET stations, but after reviewing this suggestion, we realized that this approach was inappropriate. In the revised manuscript, we adopted station-type-specific radii: 2 km for urban stations, 10 km for rural stations, and 20 km for remote stations, following Table 3 of Elbern et al. (2007).
This change is explicitly described in Section 3.3.2 of the revised manuscript as follows:
"The BC observation data were collected from 32 stations (Guo et al., 2020), including 11 urban, 17 rural, and 4 remote stations…"
"…and L is the radius of influence of a BC observation. According to Elbern et al. (2007), L was set to 2 km, 10 km, and 20 km for urban, rural, and remote stations, respectively…"
Since we updated the radius of influence of BC observations, all experiments in Section 5.3 have been redone using the revised settings. Consequently, the corresponding figures and text have been updated to reflect the new results.
We once again appreciate the reviewer's careful examination of our revised manuscript.

14. **Page 29, Lines 718-722: For the given result, I will think that the assimilation of meteorological observations does not update too much the model variables used to calculate the model equivalence to the BC observations. Otherwise, the observation-minus-background innovations in DA_MET_then_BC and DA_BC would be different, leading to differences in the analysis increments.**

    Response: Thanks for the valuable comment. In the DA_MET_then_BC experiment, although the assimilation of meteorological observations does update atmospheric variables (e.g., wind, temperature, humidity), it does not change the background field of BC itself. Thus, the BC assimilation step in both DA_BC and DA_MET_then_BC uses the same BC background. As a result, the observation-minus-background (OMB) values for BC observations are almost identical in both DA_BC and DA_MET_then_BC, except for minor differences caused by the influence of updated meteorological fields on the observation operator. Therefore, the analysis increments of atmospheric variables caused by BC assimilation remain similar between the two experiments.

    We have revised the corresponding text to clarify this point as follows:

    "…This is because, although the DA_MET_then_BC experiment assimilates meteorological observations before BC surface observations, the background field of BC remains unchanged. While the assimilation of meteorological observations updates atmospheric variables, it does not directly alter the BC background field. Therefore, the OMB values for BC observations in DA_MET_then_BC are very close to those in DA_BC, with only minor differences caused by the slight influence of updated meteorological fields on the observation operator. As a result, the analysis increments of atmospheric variables due to BC assimilation are similar between the two experiments…"

15. **Page 29, Line 736: It seems that the BC observations and atmospheric observations are pulling the model trajectory in opposite direction. If this is the case, then the quality of the BC observations should be worried, or the specification of the observation error covariance matrix is poor.**

    Response: We sincerely appreciate the insightful comments. We agree that the simultaneous assimilation of BC and meteorological observations may reflect some competing influences on the model state. In the revised text, we clarified that the reduced BC-induced atmospheric increments in DA_MET_BC_simult are likely due to the stronger constraints imposed by the atmospheric observations. These can moderate the impact of BC observations during the assimilation process. This behavior also highlights the importance of properly specifying the observation error covariance matrix. In future work, we plan to further examine the specification of the BC observation errors and their impact on assimilation performance.

    The revised text now reads:

    "…The differences in analysis increments of the four atmospheric variables caused by BC assimilation between DA_MET_BC_simult and DA_BC/DA_MET_then_BC may be attributed to the stronger constraints imposed by the atmospheric observations. In both DA_MET_then_BC and DA_BC, only BC surface observations are incorporated during the BC assimilation step. At this stage, the system relies solely on BC observations to correct the initial field. In the absence of atmospheric observations, BC observations play a dominant role, leading to larger analysis increments of atmospheric variables. In contrast, in

DA_MET_BC_simult, both operational meteorological observations and BC surface observations are assimilated simultaneously. In this scenario, atmospheric observations may provide additional constraints on the adjustment of atmospheric fields, thereby moderating the impact of BC observations during the assimilation process. As a result, a more balanced adjustment of atmospheric variables is achieved in DA_MET_BC_simult. This behavior also highlights the importance of properly specifying the observation error covariance matrix. In future work, we plan to further examine the specification of the BC observation errors and their impact on assimilation performance."

Once again, we thank the reviewer for the thoughtful comments and constructive suggestions.