A Review of

"Development of the CMA-GFS-AERO 4D-Var assimilation system v1.0- Part 1: System description"

submitted to Geoscientific Model Development by Liu et al. (2024)

Review Decision: Major Revisions

Manuscript type: *Development and technical papers*

General Comments:

This manuscript is the first part of a two-part paper that documents the development of a strongly coupled aerosol-meteorological 4D-Var assimilation system with part I focusing on the system description. This paper lands on an important topic, coupled data assimilation, which has gained more and more attention as the general consensus is to consider different components of the Earth system as a whole. Although the structure of the paper is quite well organized and the topic is very relevant to the geoscientific modeling community, I find the current form of the paper difficult to understand in two aspects: 1) English writing and 2) descriptive but lacks interpretation. For 1), there are many spelling errors, grammar errors, and inadequate use of words. In addition, the writing is unclear to an extent that I am unsure whether my interpretation of the concepts addressed in the paper is correct. For 2), it is important for a research paper to provide a thorough description of the result as well as to provide interpretation of the result (what does it imply, what could have caused the result, or does it make sense or not, etc). This paper did a good job at the descriptive part, but lacks interpretation, especially when discussing results from the single observation experiments and the real case experiments. With that, I recommend major revisions with many comments and questions listed below.

1. As stated earlier, the writing is quite difficult to comprehend, preventing the readers from understanding the many seemingly important concepts and the value of the paper as a whole. Here are a few major concerns:

    I. Too many acronyms are used without being introduced at all. I only listed a few here: AURA/MLS, ARPEGE/MOCAGE, MOZART3, TM5, EUARD, IMAGES, STEM-III, CAMx, CMAQ, GEOS-Chem, GRAPES-CUACE, etc.

Please make sure to introduce them when they were first mentioned and pay attention when referring to them at a later time.

II. The word “field” is spelled incorrectly as “filed” in many places. Also, background field error covariance can be shortened to background error covariance. Furthermore, “feedback” does not always have to go with “effect”.

III. Inadequate use of word: the most frequently mis-used word is “set” in this paper, often times causing confusion and misunderstanding. For example, “the observation is set at 0300 UTC”. A more accurate way could be “the observation is placed at 0300 UTC” or “the observation is assumed to take place at 0300 UTC”. Here are a few more examples of imprecise use of word: “we set five experiments”, “we set the single-point observation ideal experiment for BC”, and “we further set the full observation experiment”, etc.

IV. There is nothing wrong with calling it observation “increment”, but a more appropriate term is observation “innovation”. Also, I believe “single observation experiment” is a common term in data assimilation and there is no need to press on its “idealized” part.

V. There are many grammar errors and sentences that don’t quite make sense. See comments below.

2. Abstract: I am not against calling it a *chemistry* meteorology coupled data assimilation system, however, it makes more sense and less misleading to call it an *aerosol* meteorology coupled one since only black carbon (BC) aerosol is considered so far. Besides, the name of the system, CMA-GFS-AERO, actually already suggests that it is an aerosol-meteorology coupled system. Otherwise, it would be called CMA-GFS-Chem.

3. Introduction: While it makes sense to review the previous efforts on CCMM data assimilation focusing on the variational perspective, given that this study uses a 4DVar approach, it is also important to address the previous efforts on coupled aerosol-atmosphere data assimilation using the ensemble-based approaches. I suggest shortening the description on variational approach in the introduction and include some description of the ensemble approaches and highlighting the pros and cons of a variational choice relative to an ensemble approach. Obviously 4DVar is used since it’s part of the CMA-GFS, it makes sense to extend upon the

CMA-GFS 4DVar framework for aerosol coupling. Nevertheless, it is important to point out to the readers what to expect from coupling under a variational setup as opposed to an ensemble approach. For example, in a variational setup, the modeling of cross-variable component in background error covariance could be difficult, especially for aerosol vs. atmospheric processes, while in an ensemble setup one relies on ensemble estimation for cross-variable correlations. On the other hand, in a variational setup, the TLM and ADM are essential, and this can serve a natural transition to the next paragraph on the importance of ADM starting at line 52.

4. Sections 3.1 & 3.2: It reads like there is a bunch of processes, programs, subroutines, and interfaces, but how they all work together to fulfill a coupled system is unclear. Please consider re-organize/re-write these two sections to increase clarity and make sure to stay consistent with what is being shown in Figs. S2 and S3. The key is to address the main processes in AERO-BC and describe what each process does. With that, it would make the readers easier to follow the subsequent TLM and ADM of AERO-BC section since all the pieces are there in the forward section already. In addition, it is not very clear what the interfaces that connect CMA-GFS with AERO-BC in all three model components (forward, TLM, and ADM) actually do in terms of coupling, other than knowing that they act to “couple” the aerosol with the atmosphere.

5. “Section 4.1 Model setup” should be separated from the Result section since it is not a result but a description of model configuration or model setup. It might be better to consider it as a standalone section or to be included as a sub-section of section 3.

6. Page 2, Lines 32-34: What exactly are these moisture and temperature perturbations? And what these perturbations to dynamics?

7. Page 2, Line 35: “enabling to produce the optimal initial values for …” > “enabling the production of an optimal initial condition for …”

8. Page 2, Line 52: I am not sure whether “international mainstream” is a good way to say it here. How about just “major”? Also, it should be “numerical weather prediction centers”, not “numerical weather centers”.

9. Page 3, Lines 72-74: only the surface temperature? What happened to the 3D temperature field?

10. Page 3, Line 86 & Page 8, Line 211: “adding the control variable of BC into …” > “adding BC as a control variable into …”

11. Page 4, Lines 103-104: I am not sure what is meant by “freely combinable”? These physical parameterization processes are common to many global models. What is more important is which “schemes” are being used in each of these physical processes in CMA-GFS.

12. Page 4, Line 114: what is sectional representation method? And is there a reference for that?

13. Page 5, Lines 125: 137: $\mathbf{M}_{0>i}$ and $\mathbf{M}^{T}_{0>i}$ are actually linear and adjoint “models”, not “operators”.

14. Page 5, Lines 134-135: “after the physical and preconditioning transformation” can be omitted since it has already been stated in line 132.

15. Page 6, Line 159: To be consistent with the wordings at line 155, please consider using “forward model” instead of CCMM.

16. Page 6, Lines 160-164: These are not very relevant information.

17. Page 6, Lines 163-165 and Figure S2: These descriptions are not consistent with what is shown in Fig. S2 (a) and (b). If the idea is to show that bc_driver is part of the CMA-GFS-AERO model and acts as the interface of AERO-BC to CMA-GFS, if can be simply stated without showing Fig. S2a. As for Fig. S2b, while sf_bc, trac_vert_diff, and aerosol_bc are listed, the constant/parameter program (as stated in the texts) is missing. If the subroutines under each program is important for the readers to know and will be used/mentioned in the later part of the paper, then they deserve some explanation (e.g., what is cal_aerosol_prop? some sort of calculation of aerosol optical properties?), otherwise, they need not to be mentioned or shown. For example, the q2rh program seems to be irrelevant to AERO-BC, perhaps it can be omitted to help the readers put their focus on only the relevant parts.

18. Page 6, Line 172: it makes more sense to mention the index for size bin of BC here, instead of later at section 3.3.1, as the idea of 6 diameter bins is introduced here: $\Psi_{bc} > \Psi_{bc}^{n}$ where n = 1, 6.

19. Page 6, Line 173: “water-matter variables”: are these water vapor and hydrometeor habits mass mixing ratios?

20. Page 7, Lines 183-184: this last sentence about TLM and ADM codes being written line-by-line manually doesn't seem quite necessary. Why is it important to mention that the code is written manually without using any automatic differentiation tool?

21. Page 8, Lines 231-233: does this suggest that distribution weight only depends on the size bin, and does not vary spatially? meaning that all grid points use the same distribution weight for a given size bin? If so, is it guaranteed that BC mass conserved after the re-distribution?

22. Page 8, Line 245: what is AE31?

23. Page 9, Lines 246-247: what are the quality control procedures?

24. Page 9, Line 257: According to Table 3 of Elbern et al. (2007), the radius of influence varies with station types, and a radius of 10 km corresponds to a rural station. Since 10 km is selected here, does that mean all 32 CAWNET stations are all rural stations? If not, please provide justifications for using a radius of 10 km.

25. Page 9, Lines 264-268: I have trouble understanding this sentence… what is point jump and what does layer jump mean?

26. Page 10, Line 269: "accumulated" > "summed" ?

27. Page 10, Lines 281-282: what is the physical meaning of such a simplification that assumes correlation coefficient is a product of vertical one times the horizontal one? What does this simplification imply?

28. Page 10, Line 290: what does K_p represent and why set it to 10 here?

29. Page 11, Lines 301-302: "referenced to the relationship between length scale of humidity and the height": I have trouble understanding this one as well. Why is a relationship between humidity length scale and "height" being used for the "horizontal" length scale of BC?

30. Page 12, Lines 313-315: does this mean that BC is not cycled since the model is restarted every 6 h from CMA-GS analysis that does not have BC? But the next sentence seems to indicate that 6-h forecast of BC is used as background for the next cycle… these are conflicting ideas.

31. Page 12, Line 321: what does a global scale actually mean here? Resolution, data coverage, etc?

32. Page 12, Line 329: “an important part of introducing an adjoint model” > “an important part of introducing a new modeling component, such as the AERO-BC module”?

33. Page 13, Lines 345-346: “in an approximately linear way” > “in an approximately linear manner”?

34. Page 14, Lines 377-378: 6-h integration seems a rather long time. Is it possible that the AERO-BC processes are not very nonlinear?

35. Page 15, Lines 391-392: I have trouble understanding this one. Which coupled variable? And which physical process variable? Is it also possible that AERO-BC processes are not very nonlinear such that TL approximation is not too much different from the NL one?

36. Page 18, Lines 453-457: It will be quite helpful to add more texts to address the links between Fig. 5a and Fig. 6a as these two figures are results from the same single observation experiment with observation placed at the beginning of the window (i.e., 0300 UTC) where Fig. 5a shows the initial analysis increment while Fig. 6a shows the propagated analysis increment valid at the end of the window. Same idea for Fig. 5b and Fig. 6b, while the only difference is the timing of the observation.

37. Page 19, Lines 467-469: while I think I understand what the authors are trying to say, it is not entirely correct and perhaps not necessary to end the sentence like this. The way the system is setup (i.e., the CMA-GFS-AERO 4DVar system) by minimizing both BC and atmospheric variables together suggests it is a coupled assimilation. I think what the authors are trying to suggest is that the merits of a coupled data assimilation system cannot be fully manifested or exploited by only assimilating a BC observation at the beginning of the window.

38. Page 19, Lines 467-476: I think it is nice to have a paragraph detailing the processes in the 4DVar component of CMA-GFS-AERO that induces non-zero cross-covariance between the atmosphere and BC variables via evolving the initial background covariance with the TL modeling, even though the initial cross-covariance is zero. The current paragraph is trying to do so but remains rather descriptive and lacks interpretation. For that, I suggest checking out Section 2.1 “Coupled data assimilation” of Smith et al. (2015).

Smith, P. J., Fowler, A. M., & Lawless, A. S. (2015). Exploring strategies for coupled 4D-Var data assimilation using an idealised atmosphere–ocean model. Tellus A: Dynamic Meteorology and Oceanography, 67(1). https://doi.org/10.3402/tellusa.v67.27025

In addition, I do not think “co-correlation” is a proper word.

39. Page 19, Line 487: “in fact” should be “in reality” and one can also go on to say “in reality, unlike the single observation experiment, the BC observation is …” to further distinguish the real case from the single observation case.

40. Page 20, Lines 507-508: “assimilated all observations within the assimilation time window”: How frequent is BC observation available for assimilation? I realized that this is actually mentioned in section 3.3.2 that the BC observations are hourly averaged. However, it still didn’t say how frequent BC observations are assimilated in the real-case experiments.

41. Section 4.4: are BC and atmospheric variables minimized together in EXP1 and EXP2 as well? If so, please consider adding a new column in Table 3 to address whether these variables are minimized together or separately. In addition, it might be a good idea to use names that reflects the design of the experiments instead of calling them in numerical order. For example, EXP1 to EXP4 may be renamed to SCDA_BC, SCDA_MET, WCDA_BC+MET, SCDA_BC+MET where SCDA stands for strongly coupled data assimilation while WCDA refers to weakly coupled data assimilation.

42. Page 21, Lines 517-519: I am not sure if one can really say so without showing results from EXP2.

43. Pages 21-22, Lines 539-541: ok, but why? please consider including some interpretation. Are BC and atmospheric variables minimized together in EXP1 but separately in EXP3? It doesn’t seem quite straightforward and easy to understand, at least to me, why would assimilating only BC observations in a strongly coupled setup leads to similar impact from assimilating both BC and atmospheric observations in a weakly coupled setup? What could be the mechanism that leads to such a consequence?

44. Page 22, Lines 553-556: I am not sure if one can make this statement by comparing the *differences* of analysis increments between EXP4 and EXP2 with *actual* analysis increments from EXP1 or EXP3. In addition, I am puzzled while trying to understand how the feedback of BC assimilation on atmospheric

variables is reduced by having also assimilated atmospheric observations in a coupled setup without actually seeing the analysis increments in EXP2 and EXP4. Some thought processes and reasonings from the authors are definitely required to be stated.

45. Page 22, Lines 556-558: This statement is maybe a little too strong. It sounds like having amplified feedback is not a good thing. Without verifying the analysis with the truth (e.g., re-analysis, or observations that are not assimilated), we do not know if the strongly coupled analysis is actually more accurate than the other ones. Hence, we do not know if amplified feedback is good or not good. Although we'd like to think (or theoretically correct to think) that analysis from a strongly coupled setup is better, we still need some evidence to prove it.

46. Page 23, Line 565: "only 10%": does this mean 10% is not much of an increase? And what is 10% increased computation time relative to? Say, if the microphysics process also takes about 10% computation time, then the readers can have a reference to judge whether 10% is large or small. Without any context, it is just a number.

47. Page 24, Lines 591-592: "three component models" > "three model components"

48. Figure 2: I believe the x-axis is missing a base 10 and a minus sign in the power of 10.

49. Figures S2-S3 and almost all figures: figure captions are rather vague and not very helpful. Both Figs. S2 and S3 present rather complicated ideas and deserve a clearer and informative description.

50. Figure 9: When are these analysis increments valid at? beginning, middle, or the end of the window?