My major comments:

The authors would like to thank the reviewers for volunteering their time to review this manuscript. Your comments make this manuscript better and better. I have carefully read your valuable suggestions, and the following is my reply.

-The authors plainly describe what they see, not what they learn from the analysis. The manuscript could be further improved if the authors present the results with more scientific insight.

-It is very difficult to verify simulated lifetime and annual emission and removal due to lack of observational evidence. The authors only compare the GEFS-Aerosols results with GEOS4-GOCART from Colarco et al. 2010. How about the AeroCom consensus?

GEFS-Aerosols, a new global aerosol model developed by NOAA, became operational in September 2020, and the last study time of AeroCom Phase III was 2010, so GEFS-Aerosols did not have the opportunity to participate in AeroCom for inter-model comparison. However, GEFS-Aerosols has been participating in GAFIS (global air quality forecasting and information system) (https://community.wmo.int/en/activity-areas/gaw/science-for-services/gafis) since 2022, a WMO (World Meteorological Organization) organized project for models inter-comparison.

Here are some minor comments for the authors to consider:

-Line 36 'As a first step towards this goal, 'GEFS-Aerosols was implemented to replace NGAC. The efforts to enable prognostic aerosol capability toward the goal started with the implementation of NGAC. It is not clear to me why the authors view the GEFS-Aerosols implementation as the "first step."

NGAC is an offline model that was replaced in September 2020 by the online model GEFS-Aerosols. NCEP is continuing to develop the model, such as adding aerosol data assimilation to the system. The goal is to incorporate aerosol components into NOAA UFS (Unified Forecast System). As such, we see this replacement as the first step toward our ultimate goal.

-Line 43-44: 'because these processes occur before the model output and they are the determinants of aerosol concentration." I agree that budget analysis is important to examine model's fidelity/performance. However, the justification "these processes occur before the model output and they are the determinants of aerosol concentration" is very odd and weak. Budget analysis can reveal whether the model have the bulk emission and removal processes right. Whether these tendency diagnostics are model output is totally irrelevant.

I agree with the reviewer's comments. To clarify what I meant, the sentence has been changed to "because these processes are determinants of aerosol concentrations".

-Line 40 'instead of focusing on aerosol concentration and aerosol optical depth (AOD) in a general aerosol evaluation'. Comparing GEFS-Aerosols model output with PM/AOD observations is needed to thoroughly assess the model performance and identify potential model deficit. It is certainly all right for the authors to focus on budget analysis in this manuscript. Since the model vs observation evaluation has been conducted and reported in other papers [Lines 109-111], the authors should briefly describe the efforts.

"Bhattacharjee et al. [2023] evaluated the simulation results of the GEFS-Aerosols model using AOD data derived from satellite retrieval (MODIS and VIIRS), AOD data simulated by other models (MEERA2 and NGAC), and AOD data observed from 50 AERONET stations. The period of evaluation from August 2019 to August 2020 almost coincides with the time period of this study, namely from September 2019 to September 2020. In addition to the regular daily or monthly forecast evaluations of GEFS-Aerosols, three special events were also utilized to evaluate the performance of GEFS-Aerosols. These include dust events in Northwest Africa, agricultural fires in northern India and the August fire complex in northern California. Zhang et al., [2022b] evaluated not only the AOD simulated by GEFS-Aerosols from 5 July to 30 November 2019, but also the aerosol concentrations simulated by GEFS-Aerosols during the 22-month ATOM (Atmospheric TOmography Mission) period from 2016 to 2019." has been added to the text from lines 40 to 48.

-Line53 Eq1: Initial + Emissions + Reactions = Final + Removal

Based on the governing equation, I'll probably present the equation as

Final = Initial + Emissions + Reactions – Removal

The governing equation usually describe how a variable (such as aerosol concentration) changes when other variables change. Since this study focuses on the aerosol mass balance in GEFS-aerosols, the governing equation was converted to mass balance equation.

-Line 78: 2.3 GEFS-Aerosols. Consider presenting this sub-section first in Section 2.

It has been changed

-Line 83: 'GOCART' Please define the acronym

Added at line 24

-Line 104: 'Fire Radiant Power (FRP) '. Fire Radiative Power?

Corrected

-Line 113: 'These processes ultimately define the aerosol concentration and AOD output by the model.' These processes ultimately determined 3-d aerosol

distribution, which in term affect concentration and AOD. But this sentence is somehow odd.

It has been changed

-Line 115 'MERRA2' MERRA-2 is also based on GOCART. Does sea salt emission and removal scheme in GEFS-Aerosols differ from those in MERRA2?

GEFS-Aerosols has the same sea salt emission mechanism as GEOS4-GOCART [Gong, 2003], but MERRA2 is based on GEOS5, which updated sea salt emission scheme [Randles et al., 2017]; the sea salt removal scheme in GEFS-Aerosols is also different from that in MERRA2.

Gong, S. L.: A parameterization of sea-salt aerosol source function for sub-and super-micron particles, Global biogeochemical cycles, 17(4), Doi10.1029/2003GB002079, 2003.

Randles, C. A., A. M., da Silva, V., Buchard, P. R., Colarco, A., Darmenov, R., Govindaraju, A., Smirnov, B., Holben, R., Ferrare, J., Hair, Y., Shinozuka and Flynn, C. J.: The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part 1: System Description and Data Assimilation Evaluation, Journal of Climate, 30(17), 6823-6850, 10.1175/jcli-d-16-0609.1, 2017

-Line 138 'Fig 6' The principal behind the budget analysis is that aerosols net production is approximately equal to net loss when averaged over a long time (say multiple years). It is not clear whether the monthly residual (Left side of Eq1 – Right side of Eq 1) should be interpreted as 'model error'.

Annual aerosol deposition and sedimentation of BC, OC, dust and sea salt are added to Table 2, and as the reviewer states, aerosol emissions are almost equal to their total removals. We assume that the difference between the left side of Equation 1 and the right side of Equation 1 should be zero. If not, it means there is an error inside the model or in our analysis.

-Line 155 "Therefore, the model errors for dust and sea salt are higher than those for BC and OC, while the model errors for dust are the highest.". The text seems indicate that the model errors for dust and sea salt are caused by non-linearity in the emission/removal scheme. This is not necessarily true.

I should say yes, it might be true. If we could calculate dust and sea salt emissions and removals more precisely in our analysis, the model error could be very close to zero and much smaller than the numbers we saw in Fig 6 and Fig 7.

-Line 159 'Global Aerosol Mass'. It is insightful to specify when specific aerosol species reach max and min. For instance, dust loading peaks in June and reached min in Nov. This results are consistent with Africa dust activities. However, it seems unnecessary for the authors to specify the exact date.

I agree and corrected. The total amount of dust is highly correlated with the intensity of dust activity in Africa.

-Line 172 'Annual trend'. How annual trend can be inferred from one-year simulation? Please clarify it.

Corrected to "In the simulated year, the trends for BC and OC masses are decreasing (16.4% and 22.3%, respectively) and the trends for dust and sea salt are increasing (24.9% and 16.0%, respectively); for sulfates the trend is almost constant with only a very slight decrease (8.09%)."

-Line 181-186. The discussions about the partition can be presented in a table.

The partition of dust and sea salt emissions is shown in Table 1.

-Line 189 'Aerosol emissions are directly and indirectly related to their mass in the atmosphere'. Aerosol loading is certainly related to their emissions, and aerosol emissions certainly affect aerosol mass. However, the statement is very awkward.

The "indirect" mentioned in this sentence refers to sulfate, because there is no sulfate emission in GEFS-Aerosols. Sulfate is converted mainly from SO_2, which mainly comes from anthropogenic sources in GEFS-Aerosols.

-Line 224-225: 'the size distribution of aerosol emissions becomes too important for the removal process in GEFS-Aerosols simulations when the aerosol particle size is not changed in the model' Please clarify this sentence.

For example, if the total dust emission is 50kg, of which dust1 emission is 5 kg, dust2 emission is 5 kg, dust3 emission is 10 kg, dust4 emission is 25 kg, dust4 emission is 5 kg, then finally 5 kg of dust is removed as dust1, 5 kg of dust is removed as dust2, 10 kg of dust is removed as dust3, 25 kg of dust is removed as dust4, and 5 kg of dust is removed as dust5. In summary, for each dust size, the amount emitted is the amount removed since the particle size of the dust does not change in the GEFS-Aerosols simulation.

On the other hand, 100% OC is emitted as hydrophobic, but during the removal process, 50.5% OC is removed as hydrophobic and 49.5% OC is removed as hydrophilic, because hydrophobic OC can be converted into hydrophilic OC.

-Line 229 'as they do not undergo a size (bin) change.' The GOCART is a bulk mass scheme. It's not clear to me why the authors expect bin change.

Please refer to the previous reply.

-Line 265 'GOCART' Presume it's GOES4-GOCART. It does not hurt to make it clear.

Corrected

-Line 263-269, The differences in lifetime between GEFS-Aerosols and GEO4-GOCART are attributed to model resolution and simulation period. Both model use GOCART scheme. The differences and similarities between the two GOCART schemes should be considered. The difference between the two host AGCMs should also be discussed. If identical emission and removal scheme are implemented in both GEFS and GEOS4, the emissions and removal fluxes from the two model will still be different. The model with more active moisture process may produce more wet removal. The model with more noisy wind field may produce higher dust emissions.

The similarities and differences between GEFS-Aerosol and GEOS4-GOCART in terms of aerosol deposition and emissions have been discussed in Section 2.1 “GEFS-Aerosol” and in Section 3.11 “Annual Budget”. As for the difference of AGCM (Atmospheric General Circulation Model) in the two models, GEOS4-GOCART uses a dynamic core based on Lin and Rood (1996), and GEFS-Aerosols using the dynamic core FV3 (Finite Volume Scheme with Lagrangian Vertical Coordinate) was also developed based on the work of Lin and Rood (1996). Other configurations in AGCM, such as land models and microphysics, are quite different. Discussing their impact on aerosols is a very large topic that hopefully can be covered in future work.

Lin, S. J., & Rood, R. B. (1996). Multidimensional flux-form semi-Lagrangian transport schemes. Monthly Weather Review, 124(9), 2046-2070.

-Line 310 ‘interannual variations’ It is not clear why the authors attempt to analyze interannual variations with a 15-month data set.

The authors sought to find an answer to the question “Can we use past emissions to predict future emissions, for example, for wildfire emissions?”

The 15-month data show no regularity in the nature sources of aerosol emissions on a global scale.

-Line 332-333: ‘The study of monthly and interannual variations in aerosol mass is important because it determines whether it is appropriate to use aerosol climatology fields rather than aerosol prognostic fields in weather forecasting to save computational resources.’ I thought that the use of climatological, prescribed, or prognostic aerosols in the operational model is largely determined by the resource constraint. The study of monthly and interannual variations is important because it addresses many important aerosol-related scientific questions.

Totally agree with the reviewer's point of view. For example, as NOAA/NCEP/EMC extend the global aerosol forecast from 5 days to 35 days, how to predict fire emissions in the 35-day forecast becomes more and more important.