Reply to anonymous referee #1

Overview

The response to my review was often not clear enough about the actual changes to the new version of the manuscript. That made it often difficult to check if the modifications of the manuscript addressed the comments. The main concern of the first review was that the initial manuscript was too complex: various model versions and data sets were used in an often in-consistent way. The revised manuscript has improved on this aspect, but several modifications are still required before it can be recommended for publication.

Reply: We thank the reviewer's very helpful comments and suggestions. The revised manuscript has been revised again taken into account the reviewer's comments as following. We listed all the changes one by one as followings.

Specific comments:

I had asked in my first review to provide more information about the reasons for the differences between GEFS and NGACv2. In the response the authors say that is difficult to specify and refer to the complexity of the different atmospheric modelling approaches. But the manuscript was not updated accordingly. If the authors believe most of the changes are dues to the meteorological modelling they should say so and try to give evidence as good as possible. For example, wet deposition differences as suggested by the other reviewer, are a good candidate. Some of the improvements seems to come from the different emissions, especially for fires. The authors should mention this too in the conclusion section. It is understood that an exact attribution of the causes of model changes is difficult, but the authors should not avoid that important topic completely.

Reply: We thank the reviewer's very good suggestions. We have mentioned these important differences between GEFS-Aerosols and NGACv2 that may contribute to the improvements for GEFS-Aerosols model in the revised manuscript, corresponding discussions are also added into conclusion section. Please see them at: P13: L21-23; P18: L8-9; P22: L30-P23: L1; P23: L6-8; P24: L5-7; P24: L15-19.

The table 2 is a welcome addition but further details and clarification are required: Was the same physics package used (GFSv2015 vs GFSv15 - or is that a typo), which emissions are used for NGACv2 in the result shown in the paper (QFED v2 or GBBEPx) Is the exactly the same GOCART version used for non-dust aerosol etc.

Reply: This is not a typo, GFSv15 was implemented in 2019, while GFSv2015 was implemented in 2015 (corresponding version should be GFSv12, but it did not call GFSv12 at that time). NCEP name the physics package using GFSv15, v16… since the GFS package was implemented into FV3 from 2019. We have revised clearly in **Table 3**.

In 2.2.3 (P11: L6-10), we have described the fire emission used in NGACv2 as "The fire Emissions of carbonaceous aerosols and SO2 are from Global Biomass Burning Emission Product-extended [GBBEPx, Zhang et al. 2014]. GBBEPx emissions are blended from NESDIS's Global Biomass Burning Emission Product from a constellation of geostationary satellites [GBBEP; Zhang et al., 2012] and GMAO's Quick Fire Emissions Data Version 2 from polar-orbiting satellites [QFED2; Darmenov and da Silva, 2015]". We also have clarified them in **Table 2**.

Yes. The GOCART version is not exactly the same version for non-dust aerosol, because one is based on GOCART in WRF-Chem, while the other one is based on the GOCART within GEOS-5. We have clarified them in **Table 2**.

The new manuscript is improved as it focuses more on the comparison of the "new" system "GEFS-Aerosols v1" with the "former" system "NGACv2". However, the nomenclature is still inconsistent. The authors should carefully check text, figures, tables and captions to use always the same short name for the referred to model versions. All the occurrences, also in the figures itself need to be corrected. For example "NGAC" , "NGAC-v2" . NGAC-GOCART" is used in the text and figures. The GEFS aerosol is in some places referred to as "FV3-C384" (Table 3) or GBBEPx (e.g Figure of 18).

Reply: We thank the review's very good suggestions. We have revised it throughout the manuscript, including the tables and figures to use the uniform names.

As pointed out in my review, the authors should clarify that the paper is not about interactive aerosol in the radiation scheme. Still almost half of the introduction section is dedicated to that topic, which is not discussed in the paper. That is misleading and the section should be removed or substantially shortened.

Reply: We really appreciate the reviewer's very helpful comments. We have removed most of this part and shortened it substantially, see P3 L10 to P3 L19.

On the other hand, global aerosol forecasting systems, of which GEFS Aerosol is another example are not sufficiently discussed in the introduction section. For example, the global aerosol forecasting system run by ECMWF/CAMS, JMA, NASA GSFC/GMAO and NRL/FNMOCCAMS or other system participating in the ICAP effort should be discussed and current challenges of aerosol forecasting could be reflected on.

Reply: We really appreciate the reviewer's very helpful comments. We have added other global aerosol forecasting systems, including the ICAP, into the introduction. See P3 L20 to P4 L12.

The figure captions and legends still require a careful review. For example, Figure 16 and 17 do not include the information what is shown. In Figures 18 and 19 it is not clear which species and dates are shown in which panel. All panels have the time label 8/15/16 but the captions say the plots are shown for two dates. Why is only NGACv2 MM shown and not for dust, OC, BC and SO2.

Reply: We have revised the figures from Fig.14 to Fig. 19. We added more clear titles into the figures to show the species, we also revised the figure captions to make them clearly.

For NGACv2 results, there is only the dust concentration being archived in the data storage for 2016 at NCEP, the concentration of other aerosol species are not available for 2016.

Reply to anonymous referee #2

Overview

I appreciate the effort the authors took in revising the manuscript, and in light of the reviewer 1 comments that necessitated a significant restructuring and reorganizing of the paper I find most of my remaining comments are satisfactorily addressed. I especially appreciate the inclusion of statistics in Table 1 and addition of Tables 2 and 3 that further bolster the results with quantification instead of qualification. The paper could use a thorough read again for grammar, but mostly I only have minor comments below and find the paper suitable for publication pending addressing those.

Reply: We thank the reviewer's very helpful comments and suggestions. The revised manuscript has been revised again taken into account the reviewer's comments as following.

Minor comments

Figure 1: What is "GSDCHEM"?

Reply: Revised as "Diagram showing the components within the NEMS infrastructure".

Page 5, Line 13: Don't understand "no execution time" sentence, awkwardly constructed and unclear what is meant.

Reply: In the first review comments, the other reviewer raised a question as "Please indicate the resolution of the 31 non-aerosol members and the resolution of the aerosol member. Are they the same? How does the potentially increased cost and execution time of the aerosol member impact the execution time of the ensemble as a whole?" Then we replied to it as "The resolution of the 31 non-aerosols members and the aerosols member are the same. The aerosol member only performs 120 hours forecast, however, the non-aerosol members perform 384 hours forecast, so the no execution time issue because the aerosol member finishes the forecast before other non-aerosol members, so there is no execution time issue."

We have modified it as "In the operation, there is no execution time by including the aerosols component as one of the ensemble members since this member only performs 120 hours forecast by including aerosol component, which is shorter than other members without aerosol component that perform 384 hours forecast."

Please re-read 2.1.1 carefully to make sure it is clear and logical. It seems a bit jumbled.

Reply: We have rewritten and reorganized the section 2.1.1 as "The global Finite-Volume cubed-sphere dynamical core (FV3) developed by GFDL was chosen by NOAA as the non-hydrostatic dynamical core to be the Next Generation Global Prediction System (NGGPS) of the National Weather Service in the US [Black et al., 2021]. Currently, the FV3 was successfully implemented within the physical scheme of GFS version 15 (named as FV3GFS v15), which became operational on June 2019. It has the capability to provide the metrological basis for coupling with aerosol prediction component. The GEFS is a weather forecast modeling system made up of 31 separate forecasts, or ensemble members, which have the same horizontal (~25 km) and vertical resolution (64 layers from the surface to 0.2 hPa). The GEFS-Aerosols model is only using one of the same weather model as other GEFS members except it includes the prognostic aerosols from the coupling aerosol component. The NCEP started the GEFS to address the nature of uncertainty in weather observations that are used to initialize weather forecast models and uncertainties in model representation of atmospheric dynamics and physics. The aerosol component coupled with

FV3GFS v15 has been merged into the GEFS, as a single ensemble member named as GEFS-Aerosols, for real-time and retrospective forecast that preceded operational implementation, which occurred in September 2020.
In GFS v15, all sub-grid scale transport and convective deposition related to aerosol is handled inside the atmospheric physics routines of simplified Arakawa–Schubert (SAS) scheme. It requires consistent implementation of positive definite tracer transport and wet scavenging in the physics parameterizations, which was implemented subsequent in the forecast system of GEFSv12."

Page 6, Line 20: "capability" instead of "capable"
Reply: Revised.

Page 8, lines 14-16: This sentence indicates a significant change to the values from the optical LUTs, but I have no idea what they actually did. The model is not including nitrate, so they are scaling things up. How is that decided when/where to do that?
Reply: There was a bug in the AOD calculation before when applied the optical LUTs before, we found it and fixed it, there is not necessary to use any scaling and it has been cancelled now. In our next generation forecast system using UFS-Aerosols, the nitrate will be included.

Page 11, Line 21: "It is well known"
Reply: Revised.

Page 12, Line 8: I don't know what "two different fire emission datasets" are referred to here. GBBEPx v3 appears to be what is used here, which is merging several things together. I don't follow if this is differenced from something else somewhere.
Reply: Revised it as "The 1-D cloud module is able to be applied GBBEPx v3 fire emissions datasets to account for plume-rise that distributes the fire emissions vertically and better simulate the fire events and pollution transport of smoke plumes."

Page 14, Line 1: "more comparable to the MODIS observation" than what?
Reply: Revised it as "The high AOD over southern Africa and northern India is more comparable to the MODIS observation that of NGACv2".

Page 14, Line 33: Visual inspection of Figure 8 suggests that both models biases over the Taklimakan Desert are much larger than the 0.1 bias indicated.
Reply: Revised it as "Both GEFS-Aerosols and NGACv2 total AOD have small negative biases (~0.3-0.5) relative to MERRA-2 over the northwestern China dust source region."

Page 16, Line 16: Correlation at Maun Tower is lower in GEFS than for ICAP, per Table 1.
Reply: Revised it as "The correlation coefficients at the sites of Ascension Island and Lubango are much high than those of ICAP (see Table 1)"

Figure 12: I'm surprised to see the correlations for NGAC so poor at Cape Verde and Tenerife. This is large scale dynamics and not so strongly coupled to source processes. Are they using realistic meteorology at all?
Reply: Yes. They are using real-time forecast meteorological data.

Page 19, Line 16: Looks to me like the model “underestimates” rather than “overestimates”
Reply: Revised as “underestimates”.

Figure 14, 15, 18, and 19 would all benefit from titles at the top that explain the quantity shown rather than referring to instrumentation jargon in the legend (e.g., SP2ngm3 in Figure 18).
Reply: Revised.