Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2021-378-RC1, 2022

Comment on gmd-2021-378

Anonymous Referee #1

Referee comment on "Development and evaluation of the Aerosol Forecast Member in the National Center for Environment Prediction (NCEP)'s Global Ensemble Forecast System (GEFS-Aerosols v1)" by Li Zhang et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-378-RC1, 2022

Overview

The paper gives a model description and presents the evaluation results of the aerosol forecast of the GEFS-Aerosols v1 system. This system is a newly developed aerosol module coupled on-line to NOAA's FV3 Global Forecast System (FV3GFS) by means of the National Unified Operational Prediction Capability (NUOPC). The evaluation results are compared against the performance of the previous NGAC v2 aerosol forecast system showing a clear improvement in many aspects of the aerosol forecast.

General remarks

The paper inter-compares several aerosol model/analysis products (ICAP, GEOS5, MERRA, NGAC) with the GFES Aerosol forecast results. However, there is no stringent approach to the choice of these data sets for the different aspects. This makes the paper appear somewhat convoluted and too long. I recommend focussing on the forecast by GEFS-Aerosols v1 and its predecessor NGAC v2 only throughout the paper. These two data sets should be evaluated against observations and observation-based re-analysis data sets such as MERRA. The evaluation results of the two systems should be intercompared for all the discussed topics. If the authors still which to include other forecast or model data sets (ICAP, GEOS5) they need to describe these modelling systems in such a way that the identified differences in the evaluation against observations and observation-based reanalyses can be explained. There is little value in pointing out that GFES Aerosol is higher or lower than ICAP or GEOS5 without saying which one is better, i.e. closer to the observations.

The paper remains too in explaining the reasons for the difference in the evaluation results between GEFS-Aerosols v1 and NGAC v2. It should be stated more clearly what aspects

(emissions, removal processes, aerosol conversion, resolution, transport etc.) is assumed to be the reason for the mainly improved performance of the newer system. Further, I strongly recommend adding a table that summarises the communalities and differences between GEFS-Aerosols v1 and NGAC v2 as the reader is not made familiar with the configuration of NGAC v2.

The evaluation of the forecast consists mainly of the comparisons with respect to observations and analyses of total or speciated AOD. It is an omission of the paper that routine surface PM observations are not included in the evaluation. PM 2.5 observation data sets are widely available, and the forecast of surface PM should be a main objective of any state-of-the-art aerosol forecasting system.

The paper shows detailed comparisons against speciated AOD (BC, OC, SO4/SO2). However, the speciated AOD are model results, i.e. not provided by observation instruments, which mainly observe/retrieve total AOD. Even data assimilation of these observations for the re-analysis (MERRA) is no guarantee that the speciation of the reanalysis is better than the modelled speciation. Therefore, the evaluation with total AOD observations (AERONET) should be given a much larger emphasis in the paper. It is urgently recommended to also include the biases or RMSE (and not only the correlation) against AERONET observations in the paper. At the same time, the applied optimisations of the AOD calculation to account for aerosol species (Nitrates, SOA) not modelled by GEFS-Aerosols needs to be better explained.

The paper also includes and evaluation with flight campaign data (ATOM-1). While this is an interesting aspect of the scientific verification, it seems inconsistent that this section includes a discussion of the impact of spatial resolution which is not discussed before and which is not very large. In the interest of keeping the paper short, I would omit the resolution discussions.

Finally, the paper requires more clarification of the implied benefits of aerosol – weather feedbacks and the relation of this aerosol-aware forecast as part of the NWP ensemble of the NOAA Environmental Modeling System. It remains unclear what benefits were achieved by including the aerosol ensemble member. If no results can be presented as part of the paper, this should be stated more clearly (also in the title) and less emphasis should be given on weather-composition feedbacks as part of the introduction.

Specific comments:

Abstract:

L 11: no need to include references and mentioning of FIM-Chem in the abstract.

L 22: Please mention the main reasons for the improvements in the abstract

P 3 L 10 -22 : The discussion of the various feedbacks would only be justified if the paper reports about modified NWP results because of considering aerosol-weather feedback. This seems not the case and the text should be shortened substantially.

P 4 L 22: Here or elsewhere add the spatial resolution of the NRT GEFS-Aerosols v1 forecasts

P 5 L15: Please clarify how the emissions are added and how this is linked to the diffusion and convection tracer transport parameterisations.

P 5 L 17: Please expand on why wet deposition by large scale and convective precipitation is dealt with in different components.

P 5 L 17: Please comment in this section about the consistency of land use and other climatological surface fields (z0, vegetation type etc.) between the dynamical core and the aerosol model.

P 5 L 21: What is the motivation to include FIM-chem here?

P 5 L 24: Please provide more details on the oxidant fields. Are these statistic climatologies or do they change in space and time because of advection. Is SO2 a tracer?

P 6 L 11: It is not clear from the text what the threshold values are based on … wind tunnel experiments ?

P 6 L 18: please add (BSM)

P 6 L24: What is a 3-year climatology?

P 6 L 25: This section describes more than the coupler – so consider renaming that section or introduce sub-sections.

P 7: A reference to Fig 2a is missing in that section.

P 7 L 11: Please indicate the computational cost of the aerosol module in relation to the cost of the dynamical core.

P 7 L 12: 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?

P 7 L 19: Fig 2b is not clear at all. The names of specific routines such as checkic is not of interest for the reader. Why is re-gridding needed if the aerosol module runs at the same resolution as the core? What are the meaning of the green and yellow boxes. How is Fig 2.b related to Fig 2.a.

P 7 L26: As the AOD evaluation is an important aspect of the paper, more detail (here or elsewhere) needs to be provided to understand the impact of the optimisation of the AOD calculation on the evaluation results.

P 7 L29: This section should be re-arranged to clarify in a better way what the reference data sets are (observations, re-analysis) and what the evaluated forecasts are (GEFS-Aerosols v1 and NGAC v2 and perhaps GEOS5 and ICAP)

P 8 L 12: Please mention the number of stations and comment on the spatial coverage of the AERONET network.

P 8 L 19 / 27: Please comment on the uncertainty of the MODIS and VIIRS retrievals especially with respect to the differences over land and ocean.

P 8 L 32: Please clarify if data assimilation is applied in GEOS5 and how that data set relates to MERRA2.

P 9 L 8: The section on ATOM is very long compared to the other sections. Please consider shorten it to information relevant to the paper.

P 10 L 11: Please provide also numbers in the comparison of the CDES and HTAP 2 emission data. Please comment on the fact that the data represent different reference years and its impact of using the data for simulations in 2019.

P 11 L 9: The section 3.3 remains a bit anecdotic because only plots for selected days are shown. The paper could work without that section and it would just be enough to mention the selected biomass burning data set and injection option. If this section should remain it will require to quantify the mean aerosol biomass burning emissions for the period and to present an evaluation with independent data for the whole period. The comparison with other model and analysis data sets will require a discussion of the underlying vegetation fire data sets, in particular for NCAG v2, which does not seem to capture the fire events.

P 13 L 14: It is not possible to conclude from a map that the temporal variability was captured.

P 13 L 21: Please provide the reasons for that underestimation by NGAC v2.

P 13 L 22: Please clarify if the GEOS5 is a forecast or an analysis (data assimilation of AOD).

P 14 L 13: The comparison with AERONET AOD is more important for the reader than the inter-comparison of various modelled and analysis data sets. The section should therefore start best with the AERONET comparison.

P 14 L 29: Please discuss the biases against AERONET and not only the correlations. Please add a figure for the biases (or RMSE) similar to Fig 10 for the correlation.

P 15 L 11: Please motivate the choice of the selected stations. Why were no North-American or Siberian fire events selected ?

P 16 L 30: Why do you not include the ICAP data in the intercomparison in Fig 14 as you do in Fig 13 and before?

P 18 L 10: Please discuss the reasons for the poorer performance of NGAC v2.

P 18 L 23: Please provide the resolution in km here and before of the “native” grid.

P 18 L 28: Which resolution was used for section 4?

P 19 L 22: Please comment what the impact of the resolution on the dust emissions are. Dust emissions are known to be resolution dependent because of the respective ustar thresholds.

P 22 L 10: Volcanic eruptions have not been mentioned before. Please provide more details. On the other hand, one would expect that topics mentioned in the summary have been dealt with in the paper.

P 22 L 27: Please also mention the biases against Aeronet AOD observations.

P 23 L 11: The paper only contains tests for different resolutions and not for different emissions in section 5.

P 35 Fig 2: Consider introducing two separate Figures (2a = 2, 2b 3). The Fig 2b is not clear and a better caption is required.

P 36 Fig 3: add the different reference years and global total (Tg) in the caption.

P 37 Fig 4: Please add the total in caption.

P 39 Fig 6: “verified” is not the right word. You just show different plots/maps of AOD.

P 40 Fig 7: Please add that you show the temporal mean of the day-1 forecasts etc.

P 49 Fig 16: Why is NGAC not included in that Figure?

P 51 Fig 18: Please add the meaning of red and blue curve in caption.