Reply to anonymous referee #2

Review of "Development and Evaluation of the Aerosol Forecast Member in NCEP's Global Ensemble Forecast System (GEFS-Aerosols v1)" by Zhang et al. for publication in Geoscientific Model Development

The paper presents a description of the new GEFS-Aerosols modeling capability that is part of the FV3-based ensemble forecasts of the Global Forecast System (GFS). A number of experiments are performed with this system, and results are explicitly shown evaluating different biomass burning emissions assumptions and impacts of model horizontal resolution. Model results are compared to MODIS and VIIRS observations, AERONET and ATom data, and results from the GEOS-FP, MERRA-2, ICAP, and NGACv2 model-derived products. The model is shown to have considerably better performance relative to its predecessor NGACv2 system when compared to data sets and independent model products. Residual differences in the GEFS-Aerosols performance versus observations and models are speculated at.

The paper is overall well organized and the figures are for the most part clear (I detail some places below where I have suggestions to improve). I recognize here this is a significant update to the modeling capabilities for this major meteorological forecasting system, and I appreciate the progress the authors are making on this work. I nevertheless have a number of concerns about the paper as prepared here that I wish to see addressed before it can be published in a final form. I have many minor suggestions articulated below, but I here will lay out a few more major points.

Reply: We really appreciate the reviewer's very helpful comments and suggestions. The paper has been revised throughout based on all the general and specific comments listed by the reviewers, including the text, figures, references, etc.

First, the model description is lacking in some significant respects. In particular, there is no description of loss processes in the aerosol scheme and how they impact the simulation. This is unfortunate because in a number of places it is asserted that uncertainties in wet removal schemes explain differences between the model and observations. A general description of the approach would be helpful here, and it would be useful also to see differences in the large-scale and convective-scale precipitation between the different resolution runs as a means to explore these differences.

Reply: We thank the reviewer's suggestion. We have added the model descriptions about sink and source processes, including the convective wet scavenging, large scale wet removal, dry deposition etc. in Section 2.1.1 and Section 2.1.2. According to the other reviewer's suggestion, we have removed the discussion and comparison of different resolution in Section 5, so we did not include the comparison of precipitation between different resolution.

More generally, a budget analysis for a new modeling system is a useful add (see e.g., Textor et al. 2006, www.atmos-chem-phys.net/6/1777/2006/) for some inspiration. It is helpful to see how the lifetime of your model is similar to and different than other systems.

Reply: We thank the reviewer's suggestion. We have the other group leading by Li Pan is working on the budget analysis and aerosol lifetime in GEFS-Aerosols. They just finished the draft and plan to submit it soon:

Pan, L., P. S. Bhattacharjee, L. Zhang, R. Montuoro, B. Baker, J. McQueen, G. A. Grell, S. A. McKeen, S. Kondragunta, X. Zhang, G. J. Frost, F. Yang, I. Stajner: Analyzing GEFS-Aerosols

annual budget to understand simulated BC, OC, Dust, Sea salt and Sulfate results in the model, to be submitted, 2022.

Second, the comparisons between the GEFS-Aerosol simulation and the comparison datasets is in most cases only qualitative. There are any number of places where the performance is described as "very good" or "better" than this or that. For the most part these are not very helpful qualifiers, and in some cases I can't reconcile the assertions with the graphics presented, or at least I don't know what exactly is being highlighted. Better is something like the presentation in Figures 10 and Table 2, which are at least quantitative (well, semi-quantitative in Figure 10). These provide more objective measures of quality. Please address this in the revisions.
Reply: We have revised the manuscript throughout with more quantitative statements and descriptions in the evaluation instead of using "very good" or "better". We also added the correlation and the RMSE values for GEFS-Aerosols, ICAP and NGACv2 with respect to AERONET observation in Table 1. Our descriptions and discussions have been modified to include these statistical results in Fig. 9, Fig. 10, Fig. 11, and Fig. 12 in the revised manuscript.

Third, and related, where discrepancies within the comparisons are noted there are appeals to wet removal schemes, plume rise model, dust emissions, and the like. Mostly these assertions are not grounded in anything presented in the paper. A compositional analysis that links underestimates in Europe to Saharan dust emissions (is that really the culprit?) would be helpful. Something similar (sensitivity tests?) to the points about wet removal too. I note a reference below that is relevant, but in particular it is pretty clear that this model suffers somewhat from a common problem in aerosol models with insufficient scavenging of especially black carbon in convective updrafts. Further expansion on this point should be included.
Reply: We have revised the descriptions and assertions. About the European AOD underprediction, this was a mistake in previous descriptions, it is not related to dust, however the sulfate AOD. We have modified it. Also, in several places about the assumptions related to wet removal, we have revised them and emphasized that it needs further investigations. We also added the references related to the black carbon wash out issues that also indicated in other models.

Finally, also noted below, the authors have chosen to evaluate the model performance with a focus on a perturbed period following the June 2019 Raikoke eruption. I note there is no indication of whether the model includes volcanic emissions at all, and Raikoke is evidently not in the simulation. If other pre-COVID periods were available for this evaluation I would prefer that, but at the least I think some acknowledgement of this state would be important to introduce as a caveat, probably most relevant to discussion of high northern latitude biomass burning.
Reply: The model has the capability to include the volcanic emission for SO_2 which is based on the estimate of injection height and eruption time. While the prediction results in the paper, we have not included the volcanic emission into the model for the 2019 Raikoke eruption. We have emphasized it in the revised manuscript as "While for the predicted results in the paper, we have not included the volcanic emission into the model for the June 2019 Raikoke eruption, it may partially impact on the underprediction over high northern latitude.".
For the pre-COVID periods, other than 2019, we only the ATOM-1 evaluation is based on 2016 summer. Because the GBBPEx operational data for GEFS-Aerosols model only launched from the 2019 summer for operational prediction other than the ATOM-1 periods, so we can not run the

retrospective experiments before 2019 summer except the ATOM-1 periods. We have added some acknowledgements about the Raikoke eruption in Section 6.

Page 4, Line 24: EMC = Environmental Modeling Center (https://www.emc.ncep.noaa.gov/emc_new.php)
Reply: Revised

Page 4, Line 26: I don't see it explicitly, but I presume in the GEFS-Aerosols member the aerosols are not in any way interactive with the radiation, clouds, etc. Please clarify that's the case. Also, assuming so, how does GEFS-Aerosols differ from other GEFS members except for the prognostic aerosols? Is it meteorologically equivalent to another member of the ensemble?
Reply: Yes. The aerosol feedback has not been included there are not in any way interactive with the radiation, clouds. We have clarified it. The only differences of between the GEFS-Aerosols other GEFS members is the prognostic aerosols, and the meteorological parts of GEFS-Aerosols are the same to other ensemble members. We have added these descriptions in the revised manuscript in Section 2.1.3 and Section 6.

Page 5, Line 11: Citations for FV3? I think it has quite a literature.
Reply: Cited.

Page 5, Line 17: I have no context to understand what GFSv15 and GEFSv12 mean. Please clarify.
Reply: The atmospheric model of FV3GFS include the dynamical core of FV3 and physical scheme of GFS scheme, here the GFSv15 is the version of physical scheme, and GEFSv12 is the version of FV3GFS with ensemble members. We have clarified it in Section 2.1.1.

Page 5, Line 17: Here or somewhere nearby it would be relevant to state the model resolution of your simulations, including also the vertical coordinate. The horizontal resolution is referred to finally in the paper much later, but I don't see the vertical resolution discussed at all.
Reply: We have added the model vertical resolution descriptions in Section 1 and Section 2.

Page 5, Line 25: Please clarify if you are in fact getting DMS emissions from Lana et al. (2011). If so, that is a departure from GOCART, which uses the DMS seawater concentrations and then determines emissions dynamically based on surface wind speeds. If using DMS direct from Lana et al. (2011), for what year and seasonal variability are you assuming?
Reply: We thank the reviewer's very good comments. We have modified it as "The marine dimethyl sulfide (DMS) emission is calculated as a product of sea water DMS concentration and sea-to-air transfer velocity as described by Chin et al., [2000]."

Page 5, Line 27: In the abstract you refer to a HRRR-based plume rise model, but here you say WRF-Chem. On page 11 you refer again to HRRR-based model heritage from the WRF-Chem. Please clarify this consistently throughout.
Reply: Yes, the plume-rise module in HRRR-Smoke is originally from WRF-Chem, with some modifications and application of the FRP in HRRR-Smoke. We have clarified it.

Page 6, Line 2: The GOCART model referred to here with the 5-bin sea salt is in Colarco et al. (2010), doi: 10.1029/2009jd012820

Reply: We thank the reviewer's very good suggestion. We have updated the citation.

Page 6, Line 6: The "S" and "A" terms are not obviously defined in the text. I cannot find a reference for the FENGSHA scheme here, or at least the Tong et al. 2017 citation is missing in the references. Please state what "S" and "A" are (where they derive from) and add the citation.
Reply: We have the descriptions of "S" and "A" terms as "A represents particle supply limitation (availability), ρ is air density, g is gravitational acceleration, S is the soil erodibility potential". Also, the citation of Tong et al., 2017 has been added: "Tong, D. Q., Wang, J. X. L., Gill, T. E., Lei, H., and Wang, B. (2017), Intensified dust storm activity and Valley fever infection in the southwestern United States, *Geophys. Res. Lett.*, 44, 4304– 4312, doi:10.1002/2017GL073524."

Page 7, Line 1: Later in the text wet removal is appealed to in various places to explain the agreement (or lack thereof) with ATom data. I note there is no mention of loss processes and how parameterized in the model. Are the loss processes also in the same sequence as the emissions in GEFS-Aerosols? What is the process order?
Reply: We have added descriptions about wet removal and other chemical processes about source and loss in Section 2.1.2 as "The metrological fields (such as land use and other climatological surface fields, vegetation type etc.) are imported from the FV3 atmospheric model to the chemical model to drive the aerosols components. They are consistent in the FV3 atmospheric model and chemical model. Other than the aerosols convective wet scavenging, all the chemical related processes of source and sink, such as emission, dry deposition, settling, large-scale wet deposition, chemical reactions are handled by the chemical model. The large-scale wet deposition and dry deposition modules are from WRF-Chem for GOCART aerosols scheme, which are column model driven by meteorological input from atmospheric model. Large-scale wet removal of aerosols includes below-cloud removal (washout) following Easter et al. [2004] and the details of below-cloud wet scavenging via interception and impaction can be found in Slinn [1984]. The dry deposition is the same as Chin et al. [2002]. After updating the chemical tracers in chemical model, they are passed back to FV3 atmospheric model for transport and advection."
We also described the chemical sequences in Section 2.1.3 as "All aerosol composition and emission-related processes are computed in GEFS-Aerosols after the atmospheric physics has been advanced and passed to the chemical model following the sequences as emission, settling of dust and sea salt, plume-rise of fire emission, dry deposition, large-scale wet deposition, chemical reactions and carbonaceous aerosol updating."

Page 7, Lines 12 - 16: This text just reads out of place here as it is descriptive of the GEFS configuration and not the aerosols themselves. This belongs I think in Section 2.1.1.
Reply: We have moved this part to Section 2.1.1

Page 7, Lines 22-23: "aerosol optical properties from NASA" is not terribly descriptive. If from GEOS/GOCART please cite appropriate sources (e.g., Colarco et al. 2010, Colarco et al. 2014).
Reply: Revised by adding the citations.

Page 8, Line 10: MERRA-2 do not provide forecasts, or anyway not in some form readily accessible. It is a reanalysis and I suspect you are looking at those products, which might just be described as state snapshots or averages.
Reply: Revised is as "AOD product". Thanks for the comment.

Page 8, Line 32: The GEOS system I think referred and used here is the near-real time GEOS-Forward Processing (GEOS-FP) system. Suggest that terminology. And my understanding is the "branding" is no longer using "GEOS-5" but simply "GEOS".
Reply: According to the other review's suggestion. We have removed the comparison with GEOS-5 and all use the MERRA-2 product. Thanks for the review's comment all the same.

Page 10, Line 6: What is the spatial resolution of the CEDS inventory used here? And are you in fact using the earlier CEDS inventory cited here and not more recently available versions that go through 2019?
Reply: The original spatial resolution of CEDS 2014 is 0.5x0.5 degree. When we started the real-time run from 2019 summer, at that time, the most recent CEDS emission is 2014 version. The 2019 version CEDS is available since 2021. The CEDS 2019 will be updated in our next operational updating this summer, which is still under evaluation now.

Page 10, Line 19: What does "GOCART model background fields" refer to? I infer the oxidants. Please clarify. "Does "NASA GEOS/GMI" model refer to the MERRA-2 GMI version (https://acd-ext.gsfc.nasa.gov/Projects/GEOSCCM/MERRA2GMI/), or something else?
Reply: Yes. The "GOCART model background fields" refer to H_2O_2, OH and NO_3. We have rewritten this part and moved to Section 2.1.2. The GMI model link has been added. Thanks for the reviewer's information.

Page 10, Line 23: I find this description confusing and am not sure what is being described versus shown in Figure 4. GBBEPx is stated to blend emissions from several sources…is that really what it is doing, is blending QFED with other emission sources? QFED I think would not be referred to as "MODIS QFED" like here as it is not a MODIS product, but derived from MODIS observations. Second is a reference to 3BEM emissions which is merged with WF_ABBA. But Figure 4 calls this "MODIS" which I don't understand. Finally, the plume rise model is mentioned to take input from FRP data. How does this relate to either of the emission products mentioned here?
Reply: According to the other reviewer's suggestion, we have removed this part and Figure 4-5. But we would also like to answer the above questions. We did not use the QFED fire emission. The Experiment 1 is using an online fire emission module in PREP-CHEM-SRC tool, which need the input data of MODIS and WF_ABBA fire product to calculate the aerosol emission based on different aerosol scheme, including GOCART scheme. It also includes a simple vertical profile to redistribute the fire emission vertically. The Experiment 2 is only using the GBBEPx fire emission at the surface. But it is not based on any observation data. The Experiment 3 is using the GBBEPx fire emission and FRP product as the input for the plume-rise module from HRRR-smoke to generate 3D fire emission. The Experiment 3 is the fire configuration in the operational GEFS-Aerosols.

Page 11, Line 13: It is really hard to read Figure 4, even blown up on a screen, in relation to the comments made about it. I can clearly see more fire spots across the northern latitudes in the GBBPEx emissions, but I cannot tell if the magnitudes are different or not in general because the points are too small to see. It is certainly not evident that emissions are greater in southern Africa (Line 15). My suggestion would be to show a temporal average (a month, a season) to make this point, and you can numerically refer to the relative number of fires observed if you need to.

Reply: According to the other reviewer's suggestion, we have removed the discussion about Experiment 1 and Experiment 2, also removed corresponding figures of Figure 4 and Figure 5.

Page 11, Line 21: What is different in Experiment 1 (prescribed parameters) versus Experiment 3 (real-time FRP data) regarding the plume rise? What are the prescribed parameters?
Reply: According to the other reviewer's suggestion, we have removed the discussion about Experiment 1 and Experiment 2, also removed corresponding figures of Figure 4 and Figure 5. But we would like to answer the above questions. The Experiment 1 is using an online fire emission module in PREP-CHEM-SRC tool, which need the input data of MODIS and WF_ABBA fire product to calculate the aerosol emission based on different aerosol scheme, including GOCART scheme. It also includes a simple vertical profile to redistribute the fire emission vertically. But it is not based on any observation data. The Experiment 3 is using the GBBEPx fire emission and FRP product as the input for the plume-rise module from HRRR-smoke to generate 3D fire emission. The Experiment 3 is the fire configuration in the operational GEFS-Aerosols.

Page 11, Line 31: For here and elsewhere, when you are showing ICAP MME are you withholding NGACv2 from the ensemble mean, or including it? if the former, do you see a problem in how the clearly biased NGACv2 results shown later might confound the interpretation of comparisons?
Reply: Only four models are used to compute ensemble mean in ICAP for total AOD calculations (NASA,NRL, JMA and ECMWF). NGAC is not used for total AOD in ICAP. So, NGAC is not withheld from ICAP in the total AOD comparison. We have clarified it in Section 2.2.3.

Page 12, Line 1: I cannot tell what you mean by saying GEFS-Aerosols are under predicted in eastern Europe. Do you mean Russia at about 60 East?
Reply: Yes. But according to the other reviewer's suggestion, we have removed the part related to this discussion in Section 3.3 and Figure 4-5.

Page 12, Line 2: It is really not clear how to say one of these models is better than the other. Some numerical statistics need to presented in terms of biases, correlations. It is also not apparent from a single day comparison that would be the case.
Reply: According to the other reviewer's suggestion, we have removed this part in Section 3.3 and Figure 4-5. We have used more statistics values to describe the model performance.

Page 12, Line 23: MERRA-2 "reanalysis"
Reply: We have revised it throughout the manuscript.

Page 13, Line 2: Maybe instead of "screening by" something like "…due to the presence of a stable stratiform cloud deck over the ocean that confounds the aerosol retrievals…"
Reply: We thank the reviewer's suggestion. According to the other reviewer's suggestion, we have removed this part in Section 3.3 and Figure 4 and Figure 5.

I also want to point out here (and later in relation to Figure 7) that you have chosen an interesting period for analysis owing to the June 22, 2019, eruption of Raikoke in the northwest Pacific, which was a significant perturbation to the high latitude aerosol environment. There is no mention of volcanic emissions in GEFS-Aerosols until the conclusions where it seems like a future extension, so I presume Raikoke is omitted from the analysis. Likely the ICAP models (and for sure MERRA-

2) do not explicitly account for the eruption, but they could catch some aspect of it through aerosol data assimilation. This needs to be noted somewhere in the discussion. Look especially at the high latitude MODIS points in Figure 7.

Reply: We would like to thank the reviewer for bringing this up. In GEFS-Aerosol, though it has the capability to include the volcanic emission during eruption time, but for the operational and real-time forecast, which need to catch up the real date, it does not include the volcanic emission. Because we can not predict when the volcanic eruption would happen. We have emphasized it in Section 2 and Section 6 in the revised manuscript. Also, we have looked into individual figures of MODIS from July 5 to November 30, 2019 (the period used in the Figure 6 in the revised manuscript). Though the AOD over high latitude is available in MODIS during this time period, the enhanced AOD may also come from Siberian fire, long-range transport from other sources, or from volcanic emission. We have added the discussion into the summary as "While for the predicted results in the paper, we have not included the volcanic emission into the model for the June 2019 Raikoke eruption, it may partially impact on the underprediction over high northern latitude."

Page 13, Line 22: Please clarify here and elsewhere what we're looking at and how it's done. Figure 8 refers to Day 1 AOD forecast biases. I *think* what you are doing is running a 1 day forecast of the aerosol and then resetting the meteorology to the new analysis and making another 1 day forecast, and so on. So you are showing in Figure 8 the ~4 month mean of those 1 day forecast outcomes? How is that compared to the GEOS analyses mentioned here? Are you also looking at GEOS forecast outputs? Or the analysis itself? Are they compatible with what you are doing? Does it matter? Is this just a simple difference of the multi-month means?

Reply: We compared each day model forecast hours with same day GEOS5 analysis or other or reanalysis data (we have used only FP GEOS5 analysis, not GEOS5 forecast) and computed the AOD statistics (Bias, RMSE, correlation etc.) for each grid for each pair of model and analysis for that model forecast hour. We then computed that for the entire 4 months of the study period and averaged it over the entire 4 months for each grid points. We have use DTC MET Tool to calculate these statistical values, this method gives an overall estimate of systematic bias of the model in spatial and temporal scale. We have clarified it in Section 2.2, similar way is also applied to the MERRA-2.

Page 14, Line 1: How might you expect emissions to differ in the 2019 simulation years versus the 2014 valid year for the CEDS inventory used here?

Reply: We would expect some overpredictions over China by using the CEDS 2014 emission for the 2019 simulation. It well known that strong actions have been taken to improve the worsening atmospheric environment in the last 10 years in China, including cutting down the pollutant emissions with forced installation of catalytic converters on vehicles, building clean-coal power generation systems, prohibiting open burning of crop, etc. (Chen et al., 2017; Zhang et al., 2012; Liu et al., 2016). Currently, the $PM_{2.5}$ pollution occurrence has reduced to meet the goals in the Air Pollution Prevention and Control Action Plan (issued by the State Council of China, http://www.gov.cn/zwgk/2013-09/12/content_2486773.htm). Considering the decreasing emission trend over China, the CEDS 2014 anthropogenic would result in the overprediction in 2019. We have added this discussion in Section 3.1 as "It should be noted that these anthropogenic emissions data are not impossible to catch up the date of real-time forecast. And it normally has time lag and represents the emissions of a different previous years. The inconsistency may have

some impact on the predictions in 2019. But that is the most recently available version of anthropogenic emission. It well known that strong actions have been taken to improve the worsening atmospheric environment and decrease the emission over China in the last 10 years (Chen et al., 2017; Zhang et al., 2012; Liu et al., 2016). Considering the decreasing emission trend over China, the CEDS 2014 anthropogenic would result in some overprediction after 2014."

Page 14, Line 28: Suggest adding some statistics of the comparisons tabularly in Table 1. It is hard to read the colors in Figure 10 quantitatively.
Reply: Thanks to the reviewer. We have provided correlation and RMSE statistical values for all the 60 sites (including the one shown in previous Figure 10 in tabular) in Table 1, also added the RMSE figures in Fig.5 in the revised manuscript.

Page 15, Line 20: I don't see what you are referring to here, and if anything ICAP looks closer to the AERONET points in Figure 11b at the time indicated.
Reply: Revised. "GEFS-Aerosols is able to predict the two AOD enhancements in mid-October and early November, which is quite comparable as ICAP. The correlation (RMSE) is 0.856 (0.15) and 0.936 (0.09) for GEFS-Aerosols and ICAP with respect to AERONET at the site of Itajuba, only 0.451 (0.22) for NGACv2."

Page 15, Line 32: I'm not sure what is meant by saying GEFS is both comparable to but slightly lower than ICAP.
Reply: Revised as "the GEFS-Aerosols prediction is slightly lower than ICAP by about 5-10%."

Page 16, Line 7: Only seven sites are shown.
Reply: Revised.

Figure 14: What is the shading shown on the map in the top left corner?
Reply: That is dust AOD on a specific day which is not necessary to show, we have modified the figure to only show the location of the sites.

Page 17, Line 25: There is nothing here that supports the assertion that a low bias over Europe is caused by underestimates in dust emissions. Please explain further.
Reply: Revised. It was a mistake in the previous description which is not related to dust emission. We have modified it as "The large absolute low biases from August to early October 2019 and March 2020 in Europe which are associated with GEFS-Aerosols underestimates of sulfate AOD (Fig. 8)".

Page 18, Line 6: There is nothing that supports the statement that under predictions are due to errors in emission or wet removal processes. Or, put differently, this statement doesn't really explain anything in the nature of the figure comparison shown.
Reply: Thanks for the comments. Yes. The biases need further investigation. We have revised it as "Both the Eastern and Western US regions exhibit GEFS-Aerosols low biases of about 5-30%, with the largest differences in Eastern US occurring in August 2019. However, the trends of total AOD temporal variations with low in summer and high in winter in the GEFS-Aerosols prediction and the MERRA-2 reanalysis are quite consistent over Eastern and Western US. The minor under predictions by GEFS-Aerosols need further investigation."

Page 18, Line 24: I think I know what "log-Z AGL" means, but please explain.
Reply: Explained.

Page 19, Line 28: I note that Figure 18 does not include any labeling Pacific versus Atlantic. Further, this is a very challenging figure to read without blowing up quite large on the monitor. I suggest you split into one figure for each species to allow more room. Finally, what you call in the text here "bias" is presented in the figure as a ratio. Please use consistent terminology.
Reply: Revised. We have added the label of Pacific and Atlantic in Fig. 16 and Fig. 17 in the revised manuscript.

Page 20, Line 3: Not sure about "all of the three experiments." I see two experiments.
Reply: Revised.

Page 20, Line 29: How do you quantify "very good" performance?
Reply: Revised. "The model results show similar pattern as the ATOM-1 in reproducing the profiles of OC even using log scale".

Page 20, Line 32: I don't follow that the model is able to capture variations in the latitude-height profiles. Figure 18 shows that BC is overestimated in the models by a factor of 10 at higher altitudes. This is by the way a known problem in many models that they do not adequately scavenge BC (see e.g., Wang et al. 2014, doi: 10.1002/2013JD020824 and later).
Reply: We thank the reviewer's very good comments. We have revised it. "Overall, predicted BC (middle column of Figure 20) is able to capture the decreasing trend with increasing altitude in the latitude-height profiles, however they are underpredicted in the biomass burning plumes near the tropics from the surface to 5 km height in both model experiments, which have been seen in other models due to insufficient scavenge (Wang et al., 2014, Choi et al., 2020)."

Page 21, Line 4: Here injection height and scavenging are again appealed to for explanations for differences. What is the impact of the injection height parameterization, and how is that evaluated in this model?
Reply: The injection height parameterization is based on the plume-rise module in Grell et al., 2011, a 1-D, time-dependent cloud model with predicted Met. Fields were used online to calculate the injection heights as well as the vertical distribution of the emission rates. So the injection heights is based on dynamical calculation of weather conditions. The mainly impact of applying the plume-rise module it that the fire emission is 3-D distribution based on the online calculating injection height. But this injection height is difficult to evaluated because there is no accurate observation to compare. What we may validate is the vertical profile of the aerosols from fires. We thank the reviewer' suggestion, we have modified the descriptions as "It appears the model does not reproduce the enhancements of BC at 1-4 km height very well over this area. It may be possibly due to relative weak convection or a low modeled injection height that the fire emission has not been lifted enough to this altitude, which need further studies."

References:

Chen, J. M., Li, C. L., Ristovski, Z., Milic, A., Gu, Y. T., Islam, M. S., Wang, S. X., Hao, J. M., Zhang, H. F., He, C. R., Guo, H., Fu, H. B., Miljevic, B., Morawska, L., Thai, P., Fat, L., Pereira, G., Ding, A. J., Huang, X., and Dumka, U. C.: A review of biomass burning: Emissions and impacts on air quality, health and climate in China, Sci. Total Environ., 579, 1000–1034, https://doi.org/10.1016/j.scitotenv.2016.11.025, 2017.

Grell, G., Freitas, S. R., Stuefer, M., and Fast, J.: Inclusion of biomass burning in WRF-Chem: impact of wildfires on weather forecasts, Atmos. Chem. Phys., 11, 5289–5303, https://doi.org/10.5194/acp-11-5289-2011, 2011.

Liu, F., Zhang, Q., van der A, R. J., Zheng, B., Tong, D., Yan, L., Zheng, Y., and He, K.: Recent reduction in NO_x emissions over China: synthesis of satellite observations and emission inventories, Environ. Res. Lett., 11, 114002, https://doi.org/10.1088/1748-9326/11/11/114002, 2016.

Zhang, Q., He, K., and Huo, H.: Cleaning China's air, Nature, 484, 161–162, https://doi.org/10.1038/484161a, 2012.