Review for Pan et al.: The Sensitivity of Smoke Aerosol Dispersion to Smoke Injection Height and Source-Strength in Multiple AeroCom Models

Last update on Oct 31, 2025, by Xiaohua Pan

Review #1,

General Remarks: This study presents a case study of a Siberian wildfire event in April 2008 using four models and three different experiments using those models (with an additional no biomass burning simulation). The methods section could be rearranged some to improve flow, and I have a few questions regarding the methodology. The text of the results section is well written; however, the discussion could be expanded. Lastly, I believe there is room for improvement with some of the figures. Although many of the results in this paper have been documented previously (importance of biomass burning injection height, uncertainty in wet removal, and variability in biomass burning emission strength between inventories), I find there still to be novel aspects of this paper. The primary novel aspect of this paper is unlike most biomass burning studies, it evaluates model-observation agreement across multiple models. I believe that this paper could be an appropriate fit for ACP after major revisions to address the following concerns.

Response: We thank the reviewer for their thoughtful and constructive feedback. We appreciate the recognition of the novelty of our study, particularly the coordinated multi-model evaluation of biomass burning aerosol simulations using harmonized experimental design and satellite-based constraints. Although the individual sensitivities to injection height and emission strength have been explored in prior studies, our present work is to systematically assess these sensitivities across multiple global models by harmonizing the biomass burning emission inventory or the biomass burning injection height and comparing the results from simulations using the models' default biomass burning emission schemes.

We acknowledge the reviewer's suggestions regarding the structure of the methods section, the need for a more extensive discussion and improvements to figure clarity. In response, we have significantly expanded the Introduction and Discussion sections to better contextualize our findings within the broader literature. As for the methods described in Section 2, we have revised the structure to improve the logical flow, including clearer subheadings and a more concise description of the model configurations and experimental design. In addition, we have revised most figures to improve clarity and readability, including enhanced color scales, clearer region labels, and more consistent formatting across panels. You can find the revised figures at the end of this document.

We believe these revisions address the reviewer's concerns and significantly strengthen the manuscript.

Specific Comments:

Introduction: I find the literature review of this study to be too brief. Given the number of studies that have reported on the impact of biomass burning plume injection height; the effects of BBIH on model-observation agreement/air quality found in prior studies should be discussed more. I think that at a minimum the impacts of BBIH on air quality and the ways different models simulate or assume BBIH could be separate more in-depth paragraphs.

***Response**:* We added significantly to the literature review in Introduction, as indicated below, from three aspects 1) impacts of BB emission injection height on air quality, 2) the ways different models simulate or assume BB emission injection height, and 3) different biomass burning inventories. Here is the revised text:

The impact of smoke aerosols on environments near the source and downwind depends not only on the emitted mass amount (or source strength), but also on factors such as injection height, chemical transformation, removal processes, and transport after emission (Kahn et al., 2008; Paugam et al., 2016; Wilmot et al., 2022). This is especially true for large boreal forest fires that often emit smoke above the planetary boundary layer (PBL) into the free troposphere, and sometimes even into the lower stratosphere, where long-distance transport is more efficient (e.g., Val Martin et al., 2010, 2018; Peterson et al., 2018). Previous studies have demonstrated that biomass burning (BB) emission injection height has a substantial influence on surface-level air quality and on the agreement between model simulations and observations, particularly during intense wildfire events. Numerous modeling studies have shown that adjusting injection heights can significantly alter simulated surface aerosol and trace gas concentrations, thereby affecting air quality assessments, model accuracy, and radiative forcing estimates (e.g., Li et al., 2023; Feng et al., 2024; June et al., 2025). When smoke remains within or near the planetary boundary layer (PBL), it contributes primarily to elevated regional pollution, including increased surface-level particulate matter and ozone concentrations (Kahn et al., 2008; Val Martin et al., 2010; Petrenko et al., 2012). In contrast, smoke injected into the free troposphere is generally transported more efficiently, with reduced surface deposition near-source, enabling long-range and even intercontinental impacts on air quality and visibility (e.g., Sessions et al., 2011; Sofiev et al., 2012). Intercomparison efforts, such as those produced by the AeroCom community, have consistently identified plume-rise representation as a key factor driving variability in simulated aerosol burdens and transport efficiency (Rémy et al., 2017; Zhu et al., 2018). Uncertainty in modeling the vertical smoke aerosol distribution in models has been reported in many studies, and the issue persists (e.g., Koch et al., 2009, Chen et al., 2009, Koffi et al., 2012, Paugam et al., 2016, Vernon et al., 2018, Zhu et al., 2018; Tang et al., 2022, Li et al., 2023.)

Current atmospheric models employ a range of approaches for parameterizing smoke injection height, from simple assumptions to physically based schemes. Common approaches include: 1) Prescribed injection heights that vary with altitude and latitude (e.g., Dentener et al., 2006; Matsui, 2017; Matsui and Mahowald, 2017; Horowitz et al., 2020; Xie et al., 2020). 2) Emission placement within the PBL or at a fixed altitude (e.g., Chin et al., 2002; Colarco et al., 2010; Takemura et al., 2005, 2009). 3) Climatological or seasonally averaged satellite-derived heights, e.g., from the Multi-angle Imaging SpectroRadiometer (MISR) and/or Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP). 4) Daily satellite plume height retrievals, that constrain model emissions using observed vertical profiles (e.g., Val Martin et al., 2010; Rémy et al., 2017; Vernon et al., 2018; Zhu et al., 2018). 5) Dynamic plume-rise models, that simulate plume rise in real time based on fire radiative power, estimated heat flux, burned area, boundary-layer depth, buoyancy, and/or meteorological conditions (e.g., Freitas et al., 2007; Sofiev et al., 2012; Veira et al., 2015a, b; Paugam et al., 2016, Lu et al., 2023). Each of these approaches has advantages and limitations; for example, the climatological schemes (i.e. scheme 1-3) may present statistical conditions and are easier to implement in models, but they will not capture the highly variable nature of fire emission on daily and sub-daily bases, whereas the more dynamic schemes capture event-to-event variability but may be limited by either satellite coverage (scheme 4) or the accuracy of the input data, and they are sensitive to the parameterizations of atmospheric stability structure, entrainment, and turbulence (scheme 5). These different fire injection representations, along with various fire emission estimates, can lead to a wide range in simulated trace gases and aerosol amounts in the atmosphere, their vertical distributions, long-range transport, surface concentrations, and other environmental impact (e.g., Petrenko et al., 2017; Pan et al., 2020; Parrington et al., 2025).

Our project, named Biomass Burning Emission Injection Height (BBEIH), is a part of the international initiative AeroCom Phase-III study (https://aerocom.met.no/experiments/BBEIH/). It is designed primarily to assess the impact of the smoke emission vertical profile, while also examining the impact of emission source strength. We address two key questions in this study: **1)** How sensitive are simulated near-source and downwind plume characteristics—including vertical aerosol distribution, near-surface concentration, and Aerosol optical depth (AOD)— to the injection height of biomass burning emissions? and 2) To what degree does the choice of biomass burning emission inventory affect smoke dispersion? Unlike previous studies that typically rely on a single model, the novelty of the current work lies in its multi-model comparative analysis of BB plume representations. Specifically, there are two parts to the project: the first part is BBIH, in which we evaluate the default vertical distribution schemes implemented in each participating model (corresponding to Schemes 1 and 2 described above), and then uniformly apply Scheme 3 across all models to assess its impact; the second part is BBEM, in which we compare the model simulations using two emission datasets obtained with different methods: the Global Fire Emissions Database (GFED) that estimates fire emissions using burned area, fuel load, and combustion completeness (Giglio et al., 2013; van der Werf et al., 2017; Randerson et al., 2018), and the Fire Energetics and Emissions Research (FEER) dataset that derives emissions empirically from satellite-observed fire radiative energy (FRE) (Ichuko and Ellison, 2014). We focus on the boreal fire case over Siberia and Kazakhstan in April 2008, which was the largest fire event in Russia during 2000-2008 estimated from MODIS satellite observations in terms of total burned area (Vivchar, 2011). Long-range transport of this Siberia/Kazakhstan smoke was detected over Alaska during the NASA ARCTAS (Arctic Research of the Composition of the Troposphere from Aircraft and Satellites) and NOAA ARCPAC (Aerosol, Radiation, and Cloud Processes affecting Arctic Climate) field campaigns in April 2008, with CO and aerosol concentrations enhanced above background levels by 100-300% (Warneke et al., 2009, 2010).

- Feng, X., Mickley, L. J., Bell, M. L., Liu, T., Fisher, J. A., and Val Martin, M.: Improved estimates of smoke exposure during Australia fire seasons: importance of quantifying plume injection heights, Atmos. Chem. Phys., 24, 2985–3007, https://doi.org/10.5194/acp-24-2985-2024, 2024.
- Freitas, S. R., Longo, K. M., Chatfield, R., Latham, D., Silva Dias, M. A. F., Andreae, M. O., Prins, E., Santos, J. C., Gielow, R., and Carvalho Jr., J. A.: Including the plume rise of vegetation fires in numerical weather prediction models, Atmos. Chem. Phys., 7, 3385–3398, https://doi.org/10.5194/acp-7-3385-2007, 2007.
- June, N. A., Ford, B., Croft, B., Chang, R. Y.-W., and Pierce, J. R.: Inclusion of biomass burning plume injection height in GEOS-Chem-TOMAS: Global-scale implications for atmospheric aerosols and radiative forcing, J. Adv. Model. Earth Syst., 17, e2024MS004849, https://doi.org/10.1029/2024MS004849, 2025.
- Kahn, R. A., Li, W.-H., Moroney, C., Diner, D. J., Martonchik, J. V., and Fishbein, E.: Aerosol source plume physical characteristics from space-based multiangle imaging, J. Geophys. Res., 112, D11205, https://doi.org/10.1029/2006JD007647, 2007.

- Li, Y., Tong, D., Ma, S., Freitas, S. R., Ahmadov, R., Sofiev, M., Zhang, X., Kondragunta, S., Kahn, R., Tang, Y., Baker, B., Campbell, P., Saylor, R., Grell, G., and Li, F.: Impacts of estimated plume rise on PM2.5 exceedance prediction during extreme wildfire events: a comparison of three schemes (Briggs, Freitas, and Sofiev), Atmos. Chem. Phys., 23, 3083–3101, https://doi.org/10.5194/acp-23-3083-2023, 2023.
- Parrington, M., Whaley, C. H., French, N. H. F., Buchholz, R. R., Pan, X., Wiedinmyer, C., Hyer, E. J., Kondragunta, S., Kaiser, J. W., Di Tomaso, E., van der Werf, G. R., Sofiev, M., Barsanti, K. C., da Silva, A. M., Darmenov, A. S., Tang, W., Griffin, D., Desservettaz, M., Carter, T., Paton-Walsh, C., Liu, T., Uppstu, A., and Palamarchuk, J.: Biomass burning emission estimation in the MODIS era: Lessons learned, challenges and recommendations, Elementa: Sci. Anthropocene, 13, 00089, https://doi.org/10.1525/elementa.2024.00089, 2025.
- Petrenko, M., Kahn, R., Chin, M., Soja, A., Kucsera, T., & Harshvardhan. (2012). The use of satellite-measured aerosol optical depth to constrain biomass burning emissions source strength in the global model GOCART. Journal of Geophysical Research: Atmospheres, 117(D18), D18212. https://doi.org/10.1029/2012JD017870
- Petrenko, M., R.A. Kahn, M. Chin, and J.A. Limbacher, 2017. Refined use of satellite aerosol optical depth snapshots to constrain biomass burning emissions in the GOCART model. J. Geophys. Res. 122, doi:10.1002/2017JD026693.
- Rémy, S., Veira, A., Paugam, R., Sofiev, M., Kaiser, J. W., Marenco, F., Burton, S. P., Benedetti, A., Engelen, R. J., Ferrare, R., and Hair, J. W.: Two global data sets of daily fire emission injection heights since 2003, Atmos. Chem. Phys., 17, 2921–2942, https://doi.org/10.5194/acp-17-2921-2017, 2017.
- Sessions, W.R., H.E. Fuelberg, R.A. Kahn, and D.M. Winker, 2011. An investigation of methods for injecting emissions from boreal wildfires using WRF-Chem during ARCTAS. Atmosph. Chem. Phys. 11, doi:10.5194/acp-11-5719-2011.
- Sofiev, M., Ermakova, T., and Vankevich, R.: Evaluation of the smoke-injection height from wild-land fires using remote-sensing data, Atmos. Chem. Phys., 12, 1995–2006, https://doi.org/10.5194/acp-12-1995-2012, 2012.
- Val Martin, M., Logan, J. A., Kahn, R. A., Leung, F.-Y., Nelson, D. L., and Diner, D. J.: Smoke injection heights from fires in North America: analysis of 5 years of satellite observations, Atmos. Chem. Phys., 10, 1491–1510, https://doi.org/10.5194/acp-10-1491-2010, 2010.
- Veira, A., Kloster, S., Schutgens, N. A. J., & Kaiser, J. W. (2015a). Fire emission heights in the climate system – Part 1: Global plume height patterns simulated by EMAC. Atmospheric Chemistry and Physics, 15(13), 7155–7171. https://doi.org/10.5194/acp-15-7155-2015.
- Veira, A., Kloster, S., Schutgens, N. A. J., & Kaiser, J. W. (2015b). Fire emission heights in the climate system – Part 2: Impact on transport, black carbon concentrations, and radiation. Atmospheric Chemistry and Physics, 15(13), 7173–7193. https://doi.org/10.5194/acp-15-7173-2015.
- Vernon, C.J., R. Bolt, T. Canty, and R.A. Kahn, 2018. The impact of MISR-derived injection-height initialization on wildfire and volcanic plume dispersion in the HySPLIT model. Atmosph. Meas. Tech. 11, 6289–6307, doi:10.5194/amt-11-6289-2018.

Section 2.1: Could the authors explain the reasoning behind not including an experiment that uses the FEERv1.0-G1.2 and MISR plume injection height (a BBIH+BBEM simulation)? If this simulation also does not reproduce observations, I think it would strengthen the conclusion that changing the emissions and injection height are not enough to accurately simulate biomass burning plumes pointing towards biases in transport and/or deposition.

***Response**:* A simulation combining both the FEERv1.0-G1.2 emissions and MISR-based injection heights (i.e., a BBIH+BBEM experiment) would indeed provide an additional perspective on whether improving both source strength and vertical distribution is sufficient to reproduce observed aerosol distributions. Our main goal in this study, however, was to disentangle the individual impacts of biomass burning injection height (BBIH) and emission strength (BBEM). The BBIH–BASE and BBEM–BASE comparisons were specifically designed to isolate these two factors. The impact of combined BBIH+BBEM experiment could be estimated from the BASE, BBIH, and BBEM experiments.

Accordingly, we have added a paragraph below to the Discussion, i.e., Section 4.2 (Discrepancies between model and satellite observations):

We did not conduct a simulation combining both MISR-based injection heights and the FEERv1.0-G1.2 emissions (i.e., a BBIH+BBEM experiment), as our main goal in the current study is to disentangle the individual impacts of biomass burning injection height and emission strength. The impact of combined BBIH+BBEM experiment could be estimated from the BASE, BBIH, and BBEM experiments, with the assumption that the effects of injection height and emission strength are approximately multiplicative and independent, such that $BBAOD_{BBIH+BBEM} = BBAOD_{BBEM} (1 + BBAOD_{BBIH} / BBAOD_{BASE})$. However, given the small differences between the BASE and either the BBIH or BBEM results downwind and in the free troposphere, we do not expect that the BBIH+BBEM experiment would produce substantially better agreement between model and satellite data.

***Page 6, Line 34**:* "It is assumed that, within each land cover region, the sampled plume profiles are representative of the entire region. This assumption is supported in part by statistical consistency across multiple cases within most land cover types." I think at least some discussion on the limitations of using a monthly data set is warranted, given that fire strength impacts plume height and varies with time and space.

***Response**:* We agree that using a monthly, regionally averaged dataset to represent plume injection height introduces limitations, particularly given the known variability in fire intensity and meteorological conditions that influence plume rise on shorter timescales, like sub-daily or daily.

To address this, we have added a brief discussion acknowledging this limitation in Section 2.2.3 (MISR plume heights for 2008) as below:

Although the MISR-based monthly and regionally averaged plume-height used in the BBIH runs offers a valuable constraint on vertical smoke injection, it does not capture short-term variability in fire intensity or meteorological conditions. Recent work by Noyes and Kahn (2025), which analyzed MISR-derived plume heights over Siberia from 2017 to 2021, provides a statistical assessment of plume-height variability in Siberia, stratified by month, ecosystem, and whether plumes were confined to the PBL or entered the free troposphere (FT). They found that approximately 80% of 117 April fire plumes remained within the PBL. For these PBL-confined plumes, the median height was about 1 km ±0.2 km above sea level, whereas FT plumes reached a median height of about 2 km±0.5 km. Although these results support the use of monthly mean profiles as a first-order approximation, such monthly and regional averages smooth over high plume events and diurnal variability. For example, although the Val Martin et al. (2018) plume-height included monthly plumes from 2008, the plume injection heights during intense events, such as the strong April 2008 Siberian wildfires examined in this study, may be underestimated.

Reference:

Junghenn Noyes, K. T. and Kahn, R. A.: Siberian wildfire smoke observations from space-based multi-angle imaging: A multi-year regional analysis of smoke particle properties, their evolution, and comparisons with North American boreal fire plumes, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2025-395, 2025.

Page 7, Line 7: The final two paragraphs of this section feel as though they come in the wrong place. I think this section could be arranged in this general order: introduce MISR, introduce MINX, introduce how it is applied to the models used in this study, then discuss Figure 2.

Response: This suggestion is reasonable. We rearranged the content as suggested in Section 2.2.3 (MISR plume heights for 2008).

Figure 2a: The boxes here are difficult to read since they are similar colors to the map below? Could the boxes just be black (or another color with large contrast to the map underneath)?

Response: We replotted Figure 2a, changing the box color to black.

General Methods question: How is missing data from the observations handled when regionally and taking the April average of the model simulations? Are days with missing CALIOP data excluded from the model for that comparison? Are the latitudes North of the MODIS and MISR boundary excluded in the regional average? Is model data averaged at the time of the satellite overpass? How are the differences in satellite data availability incorporated into the Table 3 presentation where the median of all satellites is presented? These discussion points could be a part of Section 2.3.

Response: We acknowledge this approach has limitations and have noted it in Section 2.3 "Model Evaluation Datasets MODIS, MISR, and CALIOP" as below:

We evaluated the simulated monthly AOD at 550 nm wavelength against three satellite datasets: MODIS, MISR, and CALIOP. They each provide spatial and temporal coverage, but with different sampling, across the source and downwind regions, which aligns with the AeroCom Phase III BBEIH experiment design. We computed monthly mean values for each observational dataset and each model within the focus regions, using only the valid data available from each source. Due to the logistical challenges of aligning model output with multiple satellites, each with distinct overpass timing and data gaps, we did not strictly synchronize model sampling with satellite observations. Although this approach introduces some temporal mismatch, it is commonly adopted in multi-model and multi-satellite intercomparison studies to reduce complexity and ensure broader spatial and temporal coverage; it is usually unavoidable in statistically based analyses of this type (e.g., Kim et al., 2019).

Figure 3a & 3b have separate captions and should be separate figure numbers. Same comment for *Figure 5a & 5b*.
Response*:* We separated two panels in both Figure 3 and Figure 5, with changing Figure 3a to Figure 3, Figure 3b to Figure 4, Figure 5a to Figure 5, and Figure 5b to Figure 6.

General Figures comment*:* Switch to a different sequential color map that is not jet based for Figures 2a and 3a, the colormap used for Figure 8 is ok. Switch to a diverging colormap for Figures 5 and 6 (e.g. goes from blue to red with white in the middle without the green/yellow/orange colors).
Response*:* We changed the colors scales used in the difference maps in Figure 5 and 6 (now Figure 8) to blue-white-red without green colors. This change improved the color contrast from the total maps. Thanks for your suggestions.

Page 9, Table 3: Include BBEM in this table with a reminder in the caption that BBEM doesn't include GFDL. Either remove the "BASE/Satellites" and "BBIH/Satellites" rows or update the table caption to include this information.
Response: we added BBEM to Table 3 (now Table 4), following review's suggestion using available models, i.e., CAM5, GEOS, and SPRI, but no GFDL. We also added a note on that. We kept BASE/Satellites" and "BBIH/Satellites and add "BBEM/Satellite" for a straight-forward comparison.

Table 4. The medians of regional mean AOD from satellites and model simulations

Median	**KAZA**	**RUS1**	**RUS2**	**RUS3**	**PAC**	**ALA**
Satellites	0.390	0.505	0.543	0.471	0.421	0.337
BASE	0.198	0.583	0.631	0.371	0.242	0.101
BBIH	0.196	0.565	0.641	0.371	0.247	0.104
BBEM*	0.328	0.660	0.641	0.398	0.219	0.066
BASE/Satellites	51%	115%	116%	79%	57%	30%
BBIH/Satellites	50%	112%	118%	79%	59%	31%
BBEM*/Satellites	84%	131%	118%	85%	52%	20%

Section 3.4: The summary vertical profile metrics of Z_a and F_{2km} are useful. However, I think a non-normalized metric would also be helpful since in terms of Z_a and F_{2km} the nobb simulations are often the closest to CALIOP.
Response: We appreciate this suggestion, but we note that the nobb simulations of Za and F2km are not often the closest to CALIOP. In our analysis, Za and F2km are derived from the total aerosol extinction profiles, which include contributions from both biomass burning (BB) and non-BB sources. This is necessary when comparing with the CALIOP profiles, as they also contain all aerosol types in the atmospheric column. We "normalize" the model results by subtracting the nobb simulations only when we aim specifically to isolate the biomass burning contributions.

General model-observation comparison*:* From the introduction and the discussion in Section 3.4 it seems that part of the decision to focus on April 2008 was the ARCTAS and ARCPAC field campaigns showing that smoke originating from the Boreal Asia fires impacted Alaska. I think the study could be improved by including model-observation comparisons to these field campaigns.

Response: We selected April 2008 for this study because it had the highest biomass burning emissions of the year (Fig. A1). It also marked the largest fire event in Russia during 2000–2008 in terms of total burned area, as estimated from MODIS satellite observations (Vivchar, 2011). The significant atmospheric impacts during this period including to the downwind region over Alaska were further corroborated by the ARCTAS and ARCPAC field campaigns (*Warneke et al., 2009, 2010*). This rationale is discussed in the Introduction.

We appreciate the reviewer's suggestion that incorporating more detailed comparisons with in-situ observations from these campaigns would strengthen the evaluation of model performance, particularly in downwind regions such as Alaska. We do mention that the smoke was observed by these campaigns as far away as Alaska. However, due to the lack of hourly outputs along the aircraft tracks from these multi-models, we were unable to include these comparisons in the current study. Our focus was on satellite-based evaluation (MODIS, MISR, CALIOP), which provides consistent spatial and temporal coverage across the source and downwind regions, and aligns with the AeroCom Phase III BBEIH experiment design. Future work could incorporate these observations to further constrain model performance.

Page 14, Table 4 & 5: Both BASE and BBIH use the GFED4.1 emission inventory, for a given model (ex. GEOS), shouldn't the total emission be the same between BASE and BBIH? Additionally, correct the column headers of emibboa and emibbbc to use words and acronyms that are defined.
***Response**:* Right. The total emission of OC and BC is the same between BASE and BBIH, and it is consistent with the GFED4s emission dataset for April 2008, as shown in Tables 4 and 5. The small difference between emission in BBIH and BASE is due to the way the models implement the BB emission vertical distribution, not the BB emission amount. We combined Tables 4 and 5 into Table 3, corrected the column headers and added this note in the text describing this table in Section 2.2.2.

Table 3. Global biomass burning emissions of OA and BC for April 2008. Unit: Tg mon^{-1}.

Models	OA/OC ratio	Emission BASE & BBIH		Emission BBEM	
		OA (OC)	BC	OA (OC)	BC
CAM5	1.4	2.31 (1.65)	0.150	5.12 (3.66)	0.384
GEOS	1.8	2.95 (1.64)	0.150	6.62 (3.68)	0.385
GFDL	1.6	2.59 (1.62)	0.148	-	-
SPRI	2.6	4.23 (1.63)	0.148	9.54 (3.67)	0.383

General Discussion: There is little discussion on how this study fits into prior studies of biomass burning emission and biomass burning injection height. I think additional discussion on this front would be beneficial to understanding how this study expands the knowledge presented by prior studies. For example, Zhu et al. (2018) used GEOS-Chem and the MISR-based injection heights.
***Response**:* That is a good point. We added your suggested discussion to the introduction section.

Technical Corrections:

Page 1, Line 17: "Each model performed four simulations: (1) BASE, using a common emission inventory with default injection height; (2) BBIH, with vertical distribution adjusted using MISR plume heights; (3) BBEM, with an alternative emission inventory; and (4) NOBB, excluding biomass burning emissions." For (2) and (3) I would switch the word with to using.

Response*:* We changed it in the revised version as suggested:

Each model performed four simulations: (1) BASE, using a common BB emission inventory with default injection heights; (2) BBIH, using monthly MISR plume heights; (3) BBEM, using an alternative BB emission inventory; and (4) NOBB, excluding BB emissions.

Page 1, Line 29: "These findings suggest that increasing emission strength alone is insufficient; improving vertical injection near-source to loft more smoke above 3 km in Siberia and reducing excessive aerosol wet removal during transport are critical." Clarify that "insufficient" and "critical" are referring to improving model-observation agreement of the biomass burning event. Also, I would argue that all three of these things need to be improved to simulate the Siberian wildfire event, the wording of this sentence makes the emission strength seem unimportant.

Response*:* Reworded as follows:

These results suggest that although using monthly MISR injection heights and increasing emission strength improve model performance, they are insufficient to fully reconcile model–observation discrepancies. Lofting injected smoke to higher elevation in Siberia and reducing the model aerosol wet removal rate warrant further exploration.

Page 3, Table 1: Since the column header is "Default emission altitude in boreal Eurasia" and the text explains that CAM5 and GFDL assume different emission altitudes outside this region, I think the table could be made easier to read by removing the "30° - 60°N:" and "Global:" parts of the default emission altitude in boreal Eurasia column.

Response: We modified the Table 1 as indicated below:

Table 1. List of models and their default BB emission altitudes in Boreal Eurasia (50°-60°N)

Model Name (abbreviation)	**lon°×lat°× #lev**	**Default BB emission altitude scheme**	**Meteorology**	**References**
CAM5-ATRAS (CAM5)	2.5°×1.9°×30	Dentener scheme: 20% within 0-0.1 km 20% within 0.1-0.5 km 20% within 0.5-1 km 40% within 1-2 km	Free running with T and winds Nudged to MERRA-2 in free troposphere	Matsui, 2017; Matsui and Mahowald, 2017
GEOS-i33p2 (GEOS)	0.5°×0.5°×72	PBL scheme: Uniformly distributed between surface and PBLH	Replay with MERRA-2 meteorology	Chin et al., 2002; Colarco et al., 2010
GFDL-AM4 (GFDL)	1.25°×1°×49	Dentener scheme: 20% within 0-0.1 km 20% within 0.1-0.5 km 20% within 0.5-1 km 40% within 1-2 km	Nudged to NCEP meteorology	Horowitz et al., 2020 ; Xie et al., 2020
MIROC-SPRINTARS (SPRI)	0.56°×0.56°×40	Fixed altitude scheme:	Free running with Ps, T, and winds nudged to ERA5	Takemura et al., 2005, 2009

		Uniformly distributed between surface and sigma level of 0.74 (~3 km)		

Page 5, Line 10*:* "(used in the BASE run)" GFED4.1s is also used in the BBIH simulations.
Response*:* We added BBIH as suggested.

Page 6, Line 7*:* "Pan et al (2020)" should read Pan et al. (2020).
Response: We changed it to Pan et al. (2020).

Review for Pan et al.: The Sensitivity of Smoke Aerosol Dispersion to Smoke Injection Height and Source-Strength in Multiple AeroCom Models

Last update on Oct 31, 2025, by Xiaohua Pan

Review #2

Overall Notes: The study investigates the sensitivity of biomass burning aerosol dispersion to injection height and source strength at four models participating in the AeroCom Phase III intercomparison. Particular focus is on an intense event of Siberian wildfires in April 2008. Simulations employing the default representation of plume injection height are compared with those using MISR satellite-derived plume heights, and results based on two different emission inventories are analyzed. The model outputs are evaluated against multiple active and passive satellite remote sensing datasets.

The manuscript explores an important and timely topic: the representation of biomass burning aerosol, in particular wildfire smoke, in climate and Earth system models. In light of the extreme fire events observed worldwide in recent years, and the expected increase in their frequency and intensity under climate warming, the study is highly relevant.
The manuscript is well structured and clearly written. The language is at a good level; however, most of the figures require revision. The multi-panel maps and vertical profiles are too small and hardly legible. Some overlaid boxes and certain legends are also difficult to read.
***Response**:* We thank the reviewer for their constructive feedback on the manuscript, particularly regarding the relevance of the topic, the clarity of the writing, and the overall structure. We appreciate the reviewer's suggestion regarding the figures. In response, we have revised all multi-panel maps and vertical profile plots to improve their readability. Specifically, we have improved the color contrast in the difference maps, increased the character size of the vertical profiles, and adjusted the layout to ensure that each subplot is more legible. The colors of overlaid boxes and legends have been changed for clarity across figures. We believe these changes significantly enhance the visual quality and interpretability of the figures, and we hope the reviewer finds them satisfactory. You can find the revised figures at the end of this document.

There are also some concerns about the content. While the approach with multiple sensitivity simulations is appropriate and has been well implemented on a seasonal basis, the focus on one month of a single past event seems too narrow. The available climate models and simulation results would allow a more comprehensive and statistically robust analysis. For example, the sensitivity of smoke to injection height and source strength in the models could be analysed for average and extreme events over a multi-annual period and for different vegetation types and climate regions. In their current form, the results do not differ significantly from what would be expected and is already known from previous studies. What is certainly new here is that several models were tested. By describing the model uncertainties in more detail and with more specificity, the paper could be improved.
Response: A multi-year, multi-event analysis across different vegetation types and climate regimes would provide a more comprehensive and statistically robust assessment of model sensitivity to biomass burning injection height and source strength. However, the scope of the current study is limited to a single, well-documented period, April 2008, due to the availability of coordinated model simulations from multiple modeling groups as part of the AeroCom Phase III BBEIH experiment. The high-latitude Siberian boreal region provides highest percentage of

smoke injections above the PBL, favoring long-range transport (*Val Martin et al*., 2018, Fig. 11), a focus of this study. Further, the April 2008 fires in Siberia are among the largest early season boreal fire episodes in recent decades and coincided with extensive satellite observations and field campaigns (e.g., ARCTAS and ARCPAC), offering a unique opportunity to evaluate model performance under real-world conditions.

Noyes & Kahn (2025) examine MISR injection heights and plume-particle-property evolution for about 3,600 wildfire plumes in Siberia over a five-year period (2017-2021). The results are stratified by month and biome type and were analyzed in conjunction with Reanalysis meteorology. We briefly summarize here the observations presented in that paper most relevant to the fires included in the current study. Most of the April 2008 burning occurred in a region classified as Mixed Forests by the MODIS International Geosphere-Biosphere Program (IGBP) Land Cover Type (Friedl and Sulla-Menashe, 2019). In the 2017-2021 data set from *Noyes & Kahn* (2025), about 7% of the April plumes in Mixed Forest injected into the free troposphere, and a similar proportion persisted through the rest of the burning season (May-September). The PBL height in April in this region was just under 1 km in both studies, and Mixed Forest plumes injected into the FT concentrated around 2.7±0.5 km ASL in 2017-2021, similar to the ~3 km height for the thicker parts of the 2008 plumes, to which the MISR height retrievals are sensitive.

Some of our results align with those of previous studies, which lends confidence to our conclusions and suggests greater applicability than just for the cases included here. We emphasize that this is the first coordinated multi-model intercomparison that systematically isolates the effects of injection height and emission strength using a harmonized experimental design and satellite-based constraints. The novelty of this work lies in the cross-model comparisons, the quantification of inter-model variability, particularly in the vertical aerosol distribution and long-range transport, and the identification of relative differences in underlying model attributes.

To address the reviewer's suggestion, we rewrote the discussion section to more explicitly discuss the discrepancies among models and from the satellite observations in Section 4 as below.

4. Discussion

4.1. Sources of aerosol discrepancies among models

A key contribution of the current study is the ability to intercompare model performance in simulating smoke-transport. To this end, we investigated the sources of discrepancies among models by examining the model-simulated OA, the major BB aerosol component, averaged over four source-to-downwind areas, RUS1, RUS2, RUS3, and PAC, for April 2008. This analysis includes five key variables from the BASE runs by the four models: (1) total emission from biomass burning and anthropogenic sources, (2) loss frequency due to wet and dry deposition, (3) column mass load, (4) effective mass extinction efficiency (MEE), and (5) AOD. Here, the loss frequency is calculated as the ratio of column mass load to total (wet+dry) deposition rate, and MEE is the ratio of AOD to column mass load. Results are summarized in Table 5 for the individual models, along with the multi-model median, inter-quartile range (IQR) normalized by the median (expressed as a percentage to indicate inter-model spread), and the ratio of maximum to minimum values among the models. Figure 10 further illustrates the model diversity, expressed as the percentage deviation of each model from the multi-model median for each variable. For clarity, the deposition residence time in Figure 10 is calculated as the reciprocal of the loss frequency, to highlight whether shorter residence time leads to lower mass load, as expected).

Table 5. Total emission, area-mean deposition loss frequency, column mass load, MEE, and AOD for OA averaged over RUS1, RUS2, RUS3, and PAC for April 2008 from model BASE simulation, along with associated statistical values.

	Emission (Tg)		Loss frequency (day^{-1})		Load	MEE	AOD
	Total	(BB, Anthro)	Total	(Wet, Dry)	($g\ m^{-2}$)	($m^2\ g^{-1}$)	
CAM5	1.32	(1.32, 0.003)	0.54	(0.53, 0.01)	0.021	4.29	0.09
GEOS	1.60	(1.59, 0.003)	0.21	(0.18, 0.03)	0.028	9.88	0.25
GFDL	1.57	(1.56, 0.002)	0.22	(0.17, 0.05)	0.023	8.89	0.20
SPRI	2.26	(2.26, 0.003)	0.74	(0.72, 0.02)	0.022	26.7	0.46
Median	1.58	(1.58, 0.003)	0.38	(0.35, 0.03)	0.022	9.39	0.22
IQR/Med%	16.2	(16.2, 5.25)	98.2	(114, 70.8)	9.33	67.7	58.6
Max/Min	1.71	(1.71, 1.27)	3.53	(4.29, 4.64)	1.32	6.23	4.98

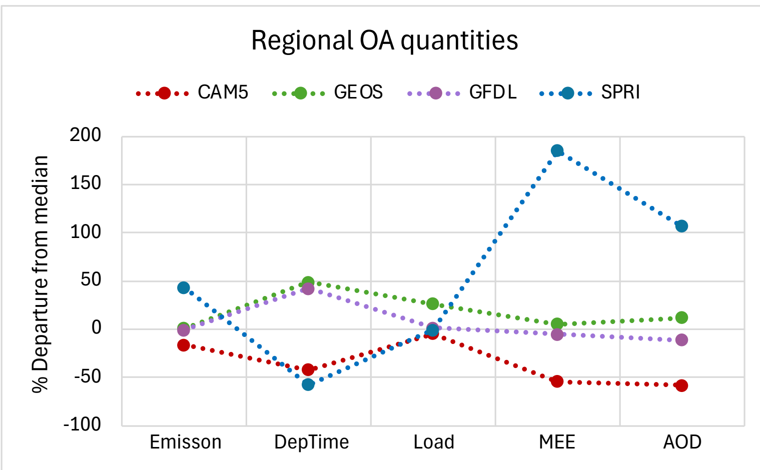


Figure 10. Comparisons of model-simulated key variables determining OA AOD in each model for April 2008, averaged over four regions from RUS1 to PAC. Colored symbols represent the percentage deviation of each model from the multi-model median. The actual values from individual models, along with the multi-model statistics (median, IQR/median, and max/min), are listed in Table 5.

Fundamentally, sources and removal rates determine the mass load, and the mass load and MEE together determine the AOD. In this study region and period, BB emission is the predominant source of OA, accounting for more than 99% of the total OA emission. For the OA loss due to deposition, all models agree that wet deposition is the major removal process, with the loss frequency 3 to 50 times higher than that of dry deposition (Table 5). Interestingly, despite significant differences in OA emissions and deposition rates among the models, the disparity of the resulting OA loads is surprisingly small. The inter-model spread in OA mass load, indicated by the IQR divided by the median, is only 9.3%, compared to 16% for emissions and 98% for loss frequency. This small spread in OA mass load is mainly due to the compensating effects of emission and removal frequency. For example, SPRI has the highest OA emission (because of its assumed highest OA/OC ratio among models as 2.6; Table 3) but also the fastest removal rate (i.e. the shortest deposition residence time), whereas GFDL has much lower emission but a significantly slower removal rate (i.e., longer deposition residence time). As a result, they end up with very similar OA mass load despite contrasting parameter choices. Note that this analysis does not account for OA inflow and outflow due to transport, nor for any secondary OA formation from volatile organic compound oxidation in the regional source/sink budget. Therefore, mass is not strictly conserved within the study region. Nevertheless, we are considering by far the dominant controls on OA in this case, so the key findings regarding the inter-model diversity remain robust.

Although OA mass loads are relatively consistent across models (max/min = 1.3 and IQR/median = 9.3%), the differences in OA AOD are very large (max/min = 5 and IQR/median = 59%). This large spread in AOD is primarily attributable to substantial differences in MEE (max/min = 6.2 and IQR/median = 68%). For instance, SPRI exhibits an extremely high MEE at 26.7 $m^2\ g^{-1}$, whereas CAM5 has the lowest value of 4.3 $m^2\ g^{-1}$ (Fig.10 and Table 5). This large contrast in MEE results in the large difference in OA AOD. Theoretically, MEE depends on aerosol optical and microphysical properties, including particle refractive indices, size distribution, dry density, and hygroscopic growth under ambient humidity (e.g., Hess et al., 1998; Chin et al., 2002). The results in Fig.10 indicate that SPRI assumes remarkably strong hygroscopic growth for OA particles, making MEE about three times the multi-model median value, whereas CAM5 assume much lower water vapor uptake ability, producing a MEE value roughly half the multi-model median. The global spatial distribution of OA mass load, OA AOD, and OA MEE are shown in supplemental Fig. A2.

Clearly, using the remotely sensed AOD as a constraint is necessary to produce realistic model simulations, but by itself, it is insufficient for evaluating the underlying factors that contribute to model AOD diversity. To improve future aerosol modeling and AeroCom intercomparisons, this study—along with Petrenko et al. (2025)—strongly recommend constraining MEE values (ranging from 4.3 to 26.7 in this study) and OA/OC ratios (ranging from 1.4 to 2.6 in this study). Unfortunately, there are no statistically robust observational constraints for MEE, emission, deposition, and mass load covering the major aerosol types, key variables that each play a critical role in determining AOD (e.g., Kahn et al., 2023). Further, the OA/OC ratio does exhibit a wide range in nature that depends on many factors, including the burned vegetation type, chemical structure of OA compounds, formation of OA from different precursors, aging of the airmass, and meteorological conditions in the environment. Although the range of OA/OC ratio in this study are within the observed values (e.g., Malm et al., 1994; Aiken et al., 2008; Hodzic et al., 2020), more systematic measurements of this ratio are highly desirable to obtain robust statistics for the most probable values under various conditions.

4.2. Discrepancies between model and satellite observations

As presented in Section 3, the models show a stronger meridional decline in AOD from the source regions to the downwind regions, compared to satellite data (e.g., Fig. 4, Fig. 7, and Table 4). The models also significantly underestimate the aerosol extinction in the middle to upper troposphere compared to CALIOP lidar data. These discrepancies persist across all experiments and models. Possible explanations include: a) excessively rapid aerosol wet removal along the transport pathways, b) underestimated BB injection height (with both model default assumptions and monthly MISR values lower than actual plume height in our study area), and c) insufficient vertical mixing. Below, we evaluate each explanation in turn.

Excessive wet removal: Our model budget analysis indicates that wet deposition is the dominant removal process for OA across all models (Table 5). This is expected, given the submicron size and hygroscopic nature of OA smoke particles. Among the models, Figure 11 and Table 5 show that CAM5 and SPRI exhibit significantly higher wet depositional loss rates than GEOS and GFDL, and their average deposition residence times over the four regions from RUS1 to PAC are ~50% lower than the multi-model median, whereas the GEOS and GFDL are 50% higher. This behavior is consistent with the steeper meridional reduction of AOD from RUS1 to PAC in CAM5 and SPRI than in other two models (Fig. 7a). The inter-model differences likely stem from differences in model representations of precipitation amount and wet scavenging parameterization, among other factors. A recent paper by Zhong et al. (2022) analyzing biomass burning aerosol lifetimes in the AeroCom global models found that the BB aerosol lifetime is strongly correlated with precipitation, indicating that wet deposition is a key driver for BB aerosol burden. Notably, however, even with much smaller loss frequency in the GEOS and GFDL models, their AOD decrease from RUS1 to PAC remain far more rapid than indicated by the satellite-retrieved AOD,

Although the dominance of wet deposition is not surprising, the degree to which it varies among models—and its potential role in the underestimation of downwind AOD and vertical aerosol extent—warrants further investigation. Future AeroCom experiments might consider performing additional sensitivity studies that involve changing the removal rates and/or implementing standardized diagnostics and tracer experiments to better quantify and compare aerosol removal pathways across models. In addition, improved wet removal metrics should be considered. Recent work (Hilario et al., 2024) suggests that precipitation intensity and relative humidity are more robust indicators of wet-scavenging efficiency, implying that models may benefit from incorporating these meteorological controls into wet-deposition parameterizations.

Underestimated BB injection height: As shown in Section 3 (Fig. 5 and Fig. 6), the change of model simulated AOD in BBIH from BASE depends on BB injection profile differences between the default used in BASE and the MISR scheme in BBIH. Figure 2b shows that the GEOS default injection height (PBL scheme) is much lower than MISR, SPRI (fixed altitude scheme) is much higher than MISR, whereas GFDL and SPRI (Dentener scheme) are similar to MISR. As a result, GEOS gains the most notable improvement in BBIH. For example, in RUS1, the fraction of AOD below 2 km (F_{2km}) improved significantly in BBIH, decreasing from 87% in BASE to 68% in BBIH, closer to the CALIOP-observed value of 51%. This improvement reflects a shift from all BB emissions being confined within the PBL in the BASE run to 55% of BB emissions being injected above the PBL in BBIH. In comparison, the default biomass burning injection heights in CAM5 and GFDL are relatively close to those retrieved by MISR, such that the differences between the BASE and BBIH simulations are minimal for these two models. In SPRI, however, which used a fixed altitude scheme in BASE that distributed emissions uniformly up to 3 km, the BBIH scheme degrades agreement with observed AOD. This is because its default BB injection height is higher than MISR; using the MISR injection height puts more emission in the PBL (45-55%) than the default (22-25%), with increasing the fraction below 1 km from 30% to 70%. Although the changes in BBIH are still too small to substantially improve the agreement between models and satellites, these results demonstrate that the model simulations do respond to changes in injection height, and shifting the injection profile to place more smoke above 3 km would help.

We did not conduct a simulation combining both MISR-based injection heights and the FEERv1.0-G1.2 emissions (i.e., a BBIH+BBEM experiment), as our main goal in the current study is to disentangle the individual impacts of biomass burning injection height and emission strength. The impact of combined BBIH+BBEM experiment could be estimated from the BASE, BBIH, and BBEM experiments, with the assumption that the effects of injection height and emission strength are approximately multiplicative and independent, such that $BBAOD_{BBIH+BBEM} = BBAOD_{BBEM} (1 + BBAOD_{BBIH} / BBAOD_{BASE})$. However, given the small differences between the BASE and either the BBIH or BBEM results downwind and in the free troposphere, we do not expect that the BBIH+BBEM experiment would produce substantially better agreement between model and satellite data.

Regarding the injection height, the monthly and regional-mean MISR plume height is broadly representative of typical plume injection behavior (Val Martin et al., 2018; Noyes and Kahn, 2025), but this approach might underrepresent extreme events or diurnal variability in plume rise, such as the strong April 2008 Siberian wildfires we focus on the current study. In addition, MISR observations (Val Martin et al., 2018), taken in the late morning (~10:30 a.m. local time), tend to underestimate typical peak daytime plume heights, as only about 20% of plumes rise above the boundary layer at that time, compared to ~55% by late afternoon (Ke et al., 2021). Future modeling should consider how injection profiles might be adjusted to address this limitation and better represent plume rise above 3 km. Providing observations to adequately constrain aerosol transport models in this respect might require applying the combination of near-source injection height from multi-angle imaging (e.g. MISR and follow-on multi-angle satellite imagers), and downwind aerosol-plume vertical distribution (e.g., CALIOP and subsequent space-based aerosol lidars) (Kahn et al., 2008).

***Insufficient vertical mixing*: ** Underestimation of aerosol extinction at higher altitudes by the models may also indicate insufficient vertical mixing or turbulent mixing. It is difficult to attribute the difference between CALIOP and the models and among different models to the transport and/or removal processes without having adequate diagnostic tools. In that regard, implementing common tracers for transport and removal would be highly desirable to more precisely diagnose and attribute the causes responsible for these discrepancies. The models use different advection schemes, vertical diffusion parameterizations, and convective transport treatments, all of which can affect the vertical distribution of aerosols. However, a comprehensive evaluation of these processes is beyond the scope of this study.

Reference

Friedl, M. and Sulla-Menashe, D.: MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MCD12Q1.006, 2019.

(1) In the introduction, a more detailed discussion of the options for parameterizing smoke injection heights in models would be useful, as well as a clearer explanation of the range of the different emission inventories (including their rationale and uncertainties), since this is ultimately one of the main motivations of the study.

***Response**:* We have expanded the Introduction to include a more detailed discussion of the available approaches for parameterizing smoke injection heights in atmospheric models. Additionally, we have clarified the range, rationale, and uncertainties associated with the emission inventories used in this study. See below:

Current atmospheric models employ a range of approaches for parameterizing smoke injection height, from simple assumptions to physically based schemes. Common approaches include: 1) Prescribed injection heights that vary with altitude and latitude (e.g., Dentener et al., 2006; Matsui, 2017; Matsui and Mahowald, 2017; Horowitz et al., 2020; Xie et al., 2020). 2) Emission placement within the PBL or at a fixed altitude (e.g., Chin et al., 2002; Colarco et al., 2010; Takemura et al., 2005, 2009). 3) Climatological or seasonally averaged satellite-derived heights, e.g., from the Multi-angle Imaging SpectroRadiometer (MISR) and/or Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP). 4) Daily satellite plume height retrievals, that constrain model emissions using observed vertical profiles (e.g., Val Martin et al., 2010; Rémy et al., 2017; Vernon et al., 2018; Zhu et al., 2018). 5) Dynamic plume-rise models, that simulate plume rise in real time based on fire radiative power, estimated heat flux, burned area, boundary-layer depth, buoyancy, and/or meteorological conditions (e.g., Freitas et al., 2007; Sofiev et al., 2012; Veira et al., 2015a, b; Paugam et al., 2016, Lu et al., 2023). Each of these approaches has advantages and limitations; for example, the climatological schemes (i.e. scheme 1-3) may present statistical conditions and are easier to implement in models, but they will not capture the highly variable nature of fire emission on daily and sub-daily bases, whereas the more dynamic schemes capture event-to-event variability but may be limited by either satellite coverage (scheme 4) or the accuracy of the input data, and they are sensitive to the parameterizations of atmospheric stability structure, entrainment, and turbulence (scheme 5). These different fire injection representations, along with various fire emission estimates, can lead to a wide range in simulated trace gases and aerosol amounts in the atmosphere, their vertical distributions, long-range transport, surface concentrations, and other environmental impact (e.g., Petrenko et al., 2017; Pan et al., 2020; Parrington et al., 2025).

Our project, named Biomass Burning Emission Injection Height (BBEIH), is a part of the international initiative AeroCom Phase-III study (https://aerocom.met.no/experiments/BBEIH/). It is designed primarily to assess the impact of the smoke emission vertical profile, while also examining the impact of emission source strength. We address two key questions in this study: **1)** How sensitive are simulated near-source and downwind plume characteristics—including vertical aerosol distribution, near-surface concentration, and Aerosol optical depth (AOD)— to the injection height of biomass burning emissions? and 2) To what degree does the choice of biomass burning emission inventory affect smoke dispersion? Unlike previous studies that typically rely on a single model, the novelty of the current work lies in its multi-model comparative analysis of BB plume representations. Specifically, there are two parts to the project: the first part is BBIH, in which we evaluate the default vertical distribution schemes implemented in each participating model (corresponding to Schemes 1 and 2 described above), and then uniformly apply Scheme 3 across all models to assess its impact; the second part is BBEM, in which we compare the model simulations using two emission datasets obtained with different methods: the Global Fire Emissions Database (GFED) that estimates fire emissions using burned area, fuel load, and combustion completeness (Giglio et al., 2013; van der Werf et al., 2017; Randerson et al., 2018), and the Fire Energetics and Emissions Research (FEER) dataset that derives emissions empirically from satellite-observed fire radiative energy (FRE) (Ichuko and Ellison, 2014). We focus on the boreal fire case over Siberia and Kazakhstan in April 2008, which was the largest fire event in Russia during 2000-2008 estimated from MODIS satellite observations in terms of total burned area (Vivchar, 2011). Long-range transport of this Siberia/Kazakhstan smoke was detected over Alaska during the NASA ARCTAS (Arctic Research of the Composition of the Troposphere from Aircraft and Satellites) and NOAA ARCPAC (Aerosol, Radiation, and Cloud Processes affecting Arctic Climate) field campaigns in April 2008, with CO and aerosol concentrations enhanced above background levels by 100-300% (Warneke et al., 2009, 2010).

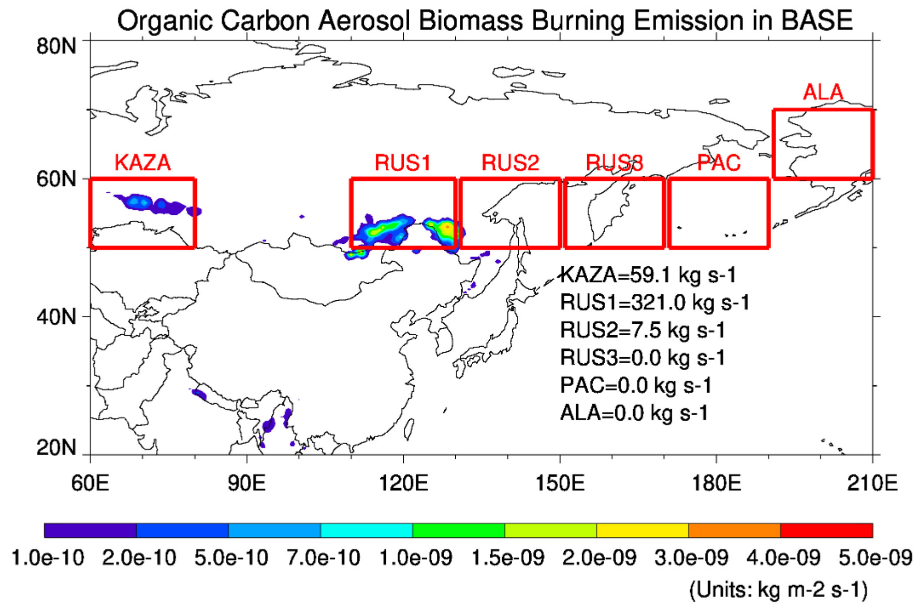
(2) Neither in the model description nor in the description of the emission datasets is there any reference to the specific biomass-burning aerosol species that are modeled. Are there differences in the emission composition across the inventories, and if so, how might these affect the results?

***Response**:* The predominant species determining the biomass burning aerosol extinction and AOD is organic aerosol (OA), that is OC multiplied by the OA/OC ratio, as we show in the manuscript. We have added below in Section 2.2.2 (BB emission inventories: GFED4.1s and FEER1.0):

This study employs two BB emission inventories—GFED4.1s (used in the BASE and BBIH run) and FEERv1.0-G1.2 (or FEER1.0, used in the BBEM run)—to assess the sensitivity of aerosol distributions to differences in source strength and spatial allocation. Both GFED4.1s and FEER1.0 provide biomass burning emissions of primary aerosols and aerosol precursor gases such as organic carbon (OC), black carbon (BC), sulfur dioxide (SO_2), nitrogen oxides (NO_x), and ammonia (NH_3), and non-methane volatile organic carbon (NMVOC) gases (van der Werf et al., 2017; Ichoku and Ellison, 2014). The predominant species determining the biomass burning aerosol extinction and AOD is organic aerosol (OA), equal to OC multiplied by an OA/OC ratio. All models participating in the BBEIH include aerosol-related emissions of OC, BC, and SO_2, although the CAM5 and GFDL models include additional NMVOCs, NO_x, and NH_3 aerosol precursor gases. In all cases, OA is the predominant species for BB aerosol mass and AOD.

(3) The section on dry and wet deposition is interesting and relevant, but too brief. That wet deposition constitutes the dominant removal pathway for smoke aerosol is not surprising, given the typical particle size of smoke aerosol compared to, for instance, desert dust or volcanic ash.
***Response**:* We have substantially revised the discussion section 4 to enhance the analysis of inter-model differences (please see our response to the second "overall notes"). The updated section now includes a more detailed examination of emissions, aerosol removal frequencies due to wet and dry deposition, aerosol mass loading, mass extinction efficiency, and aerosol optical depth (AOD), including a new Table 5 and a new Figure 10. Compared to the previous version, this revision offers deeper insights into the processes and parameters across different models that contribute to these differences. Additionally, we have incorporated discussion of model discrepancies and likely causes for these discrepancies, taking advantage of the multi-model aspect of our study. In short, the new analysis points out that despite significant differences in OA emissions and deposition rates (dominated by wet deposition) among the models, the disparity of the resulting OA loads is surprisingly small. This small spread in OA mass load is mainly due to compensating effect among the emission and deposition factors. Despite the relatively similar OA mass load in the region, the differences in OA AOD are very large, primarily attributable to substantial differences in effective mass extinction efficiency (MEE).

Revised figures



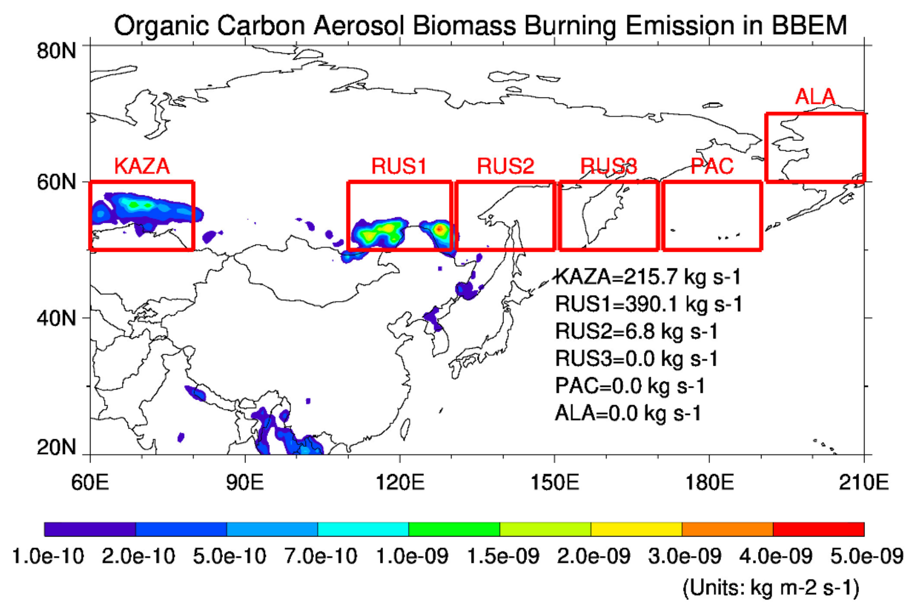


Figure 1. Biomass burning emissions from two inventories. Top: Monthly mean spatial distribution of organic carbon (OC) emissions from biomass burning in April 2008, based on the GFED4.1s inventory (used in the BASE run), in units of kg m^{-2} s^{-1}. Bottom: Same as top, but from the FEERv1.0-G1.2 inventory (used in the BBEM run). The six focus regions—KAZA, RUS1, RUS2, RUS3, PAC, and ALA—are outlined and labelled with total emissions.

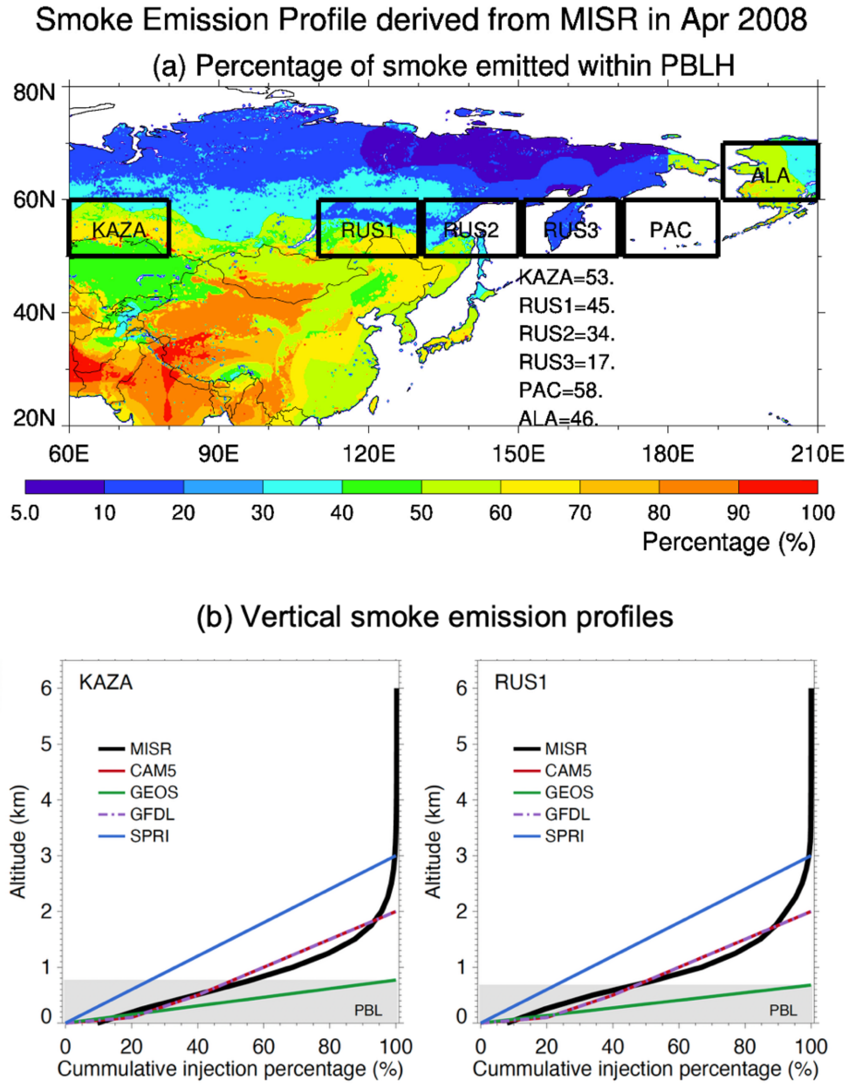


Figure 2. (a) Spatial distribution of the percentage of smoke emitted within the planetary boundary layer (PBL) in April 2008, derived from the MISR- based plume height (units: %), with regional mean values of the six focus regions listed below (over land only).
(b) Cumulative vertical smoke emission profiles over KAZA and RUS1, with the black thick curve representing the MISR-based plume height used in the BBIH run and the colored curves representing the model default vertical profiles from the models' BASE runs. The PBL layer is shaded in grey.

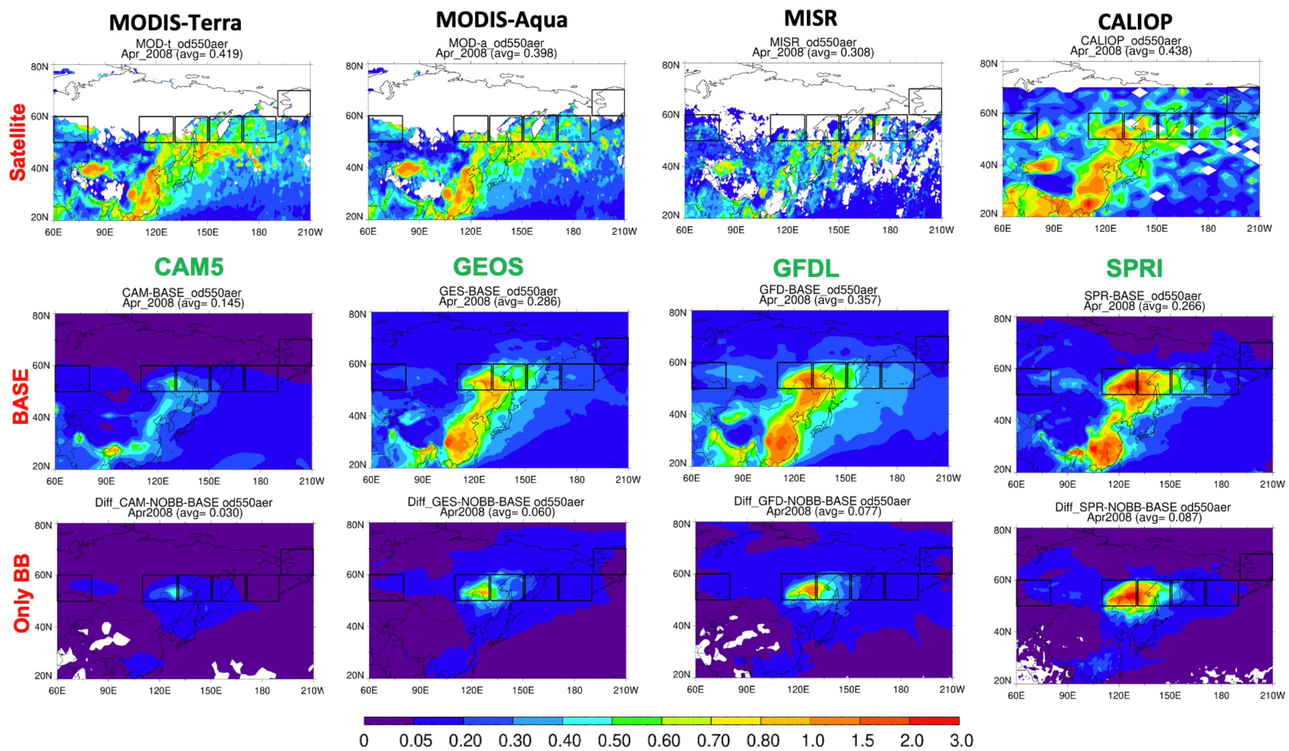


Figure 3. Spatial distribution of AOD at 550 nm in April 2008, from four satellite instruments (MODIS-Terra, MODIS-Aqua, MISR, and CALIOP) (Row 1); from four model BASE simulations (CAM5, SPRI, GEOS, and GFDL) (Row 2), and from biomass burning only AOD (BASE minus NOBB) (Row 3). Black boxes indicate the six focus regions.

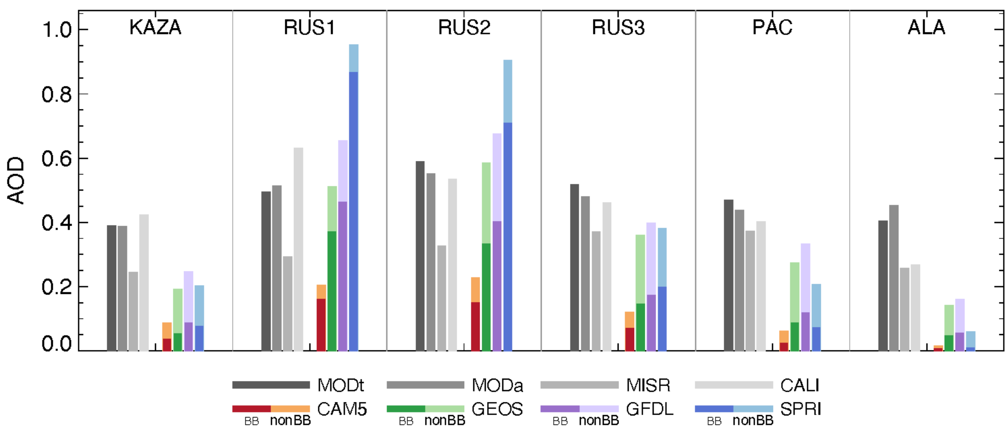


Figure 4. Regional mean AOD at 550 nm in April 2008 over the six focus regions (KAZA, RUS1, RUS2, RUS3, PAC, and ALA), derived from four satellite datasets where valid (MODIS-Terra, MODIS-Aqua, MISR, and CALIOP), and from four BASE model simulations (CAM5, SPRI, GEOS, and GFDL). Model AOD values are separated into contributions from biomass burning (BB; darker color) and non-biomass burning (nonBB, from NOBB runs; lighter color).

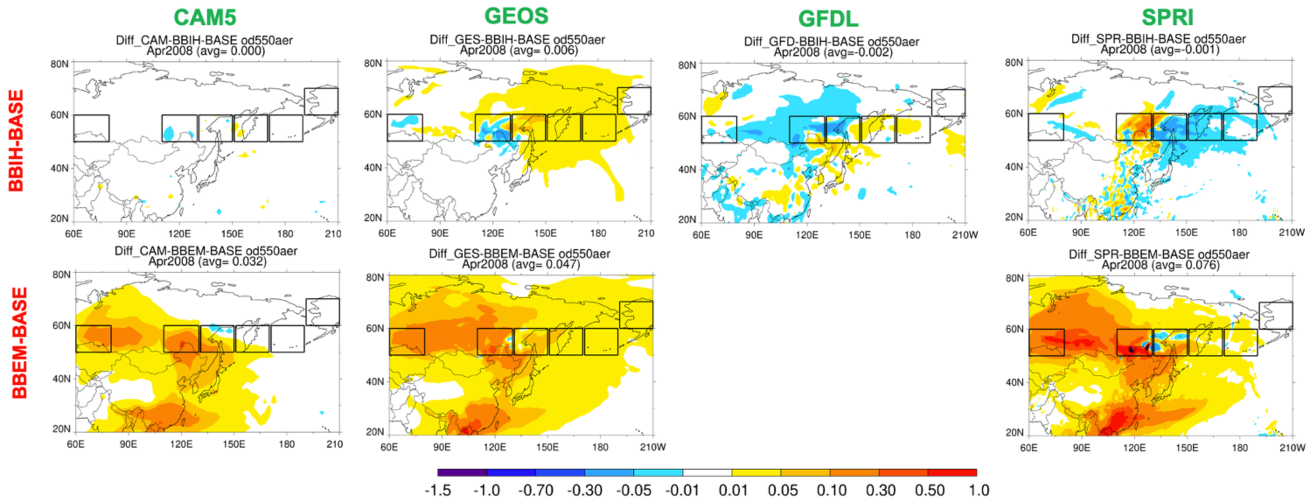


Figure 5. Spatial differences in AOD at 550 nm between BBIH and BASE (Row 1) and between BBEM and BASE (Row 2), simulated by the four models for April 2008. Only three models—CAM5, GEOS, and SPRI—submitted BBEM simulations. Focus regions are outlined in black.

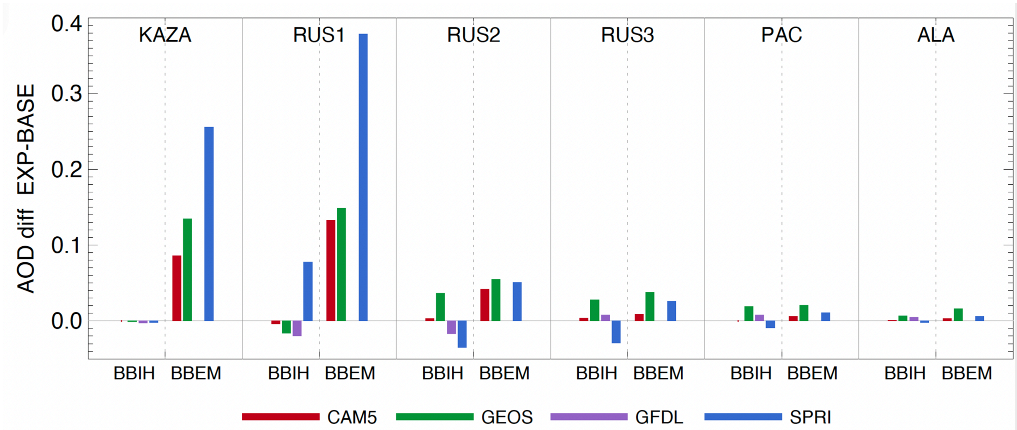


Figure 6. Regional mean differences in AOD at 550 nm for April 2008 across the six focus regions (KAZA, RUS1, RUS2, RUS3, PAC, and ALA), as simulated by four models (CAM5, SPRI, GEOS, and GFDL). **Left in each panel:** BBIH minus BASE; **Right in each panel:** BBEM minus BASE. Only three models, CAM5, GEOS, and SPRI, submitted BBEM simulations.

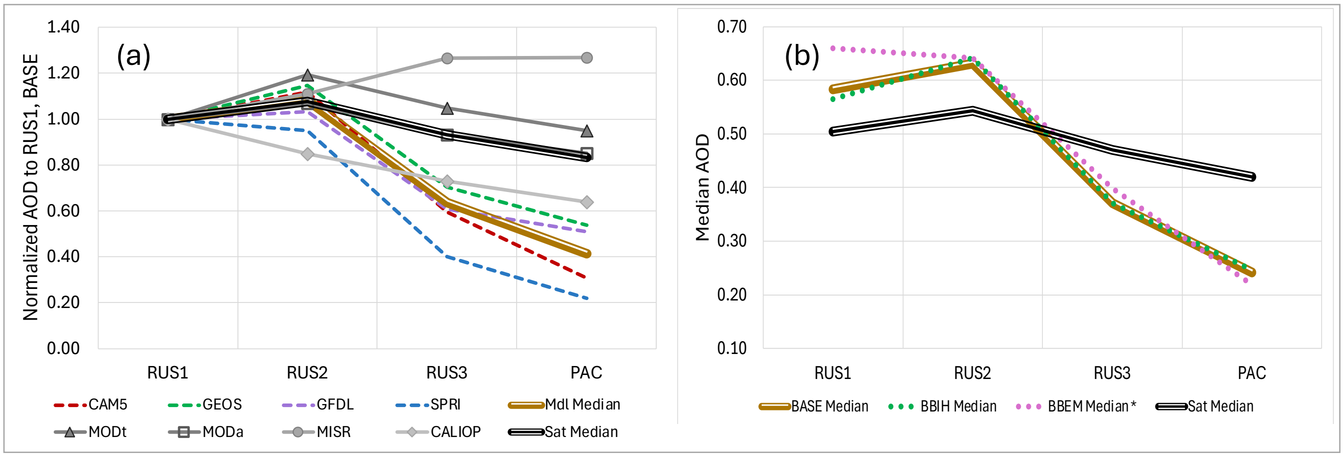


Figure 7. (a) The normalized 550 nm AOD gradient (relative to RUS1) from the BB source region RU1 to three downwind regions, based on satellite observations and the BASE simulations. (b) Comparison of the model median AOD values for four regions from the BASE, BBIH, and BBEM experiments, along with the satellite median values.

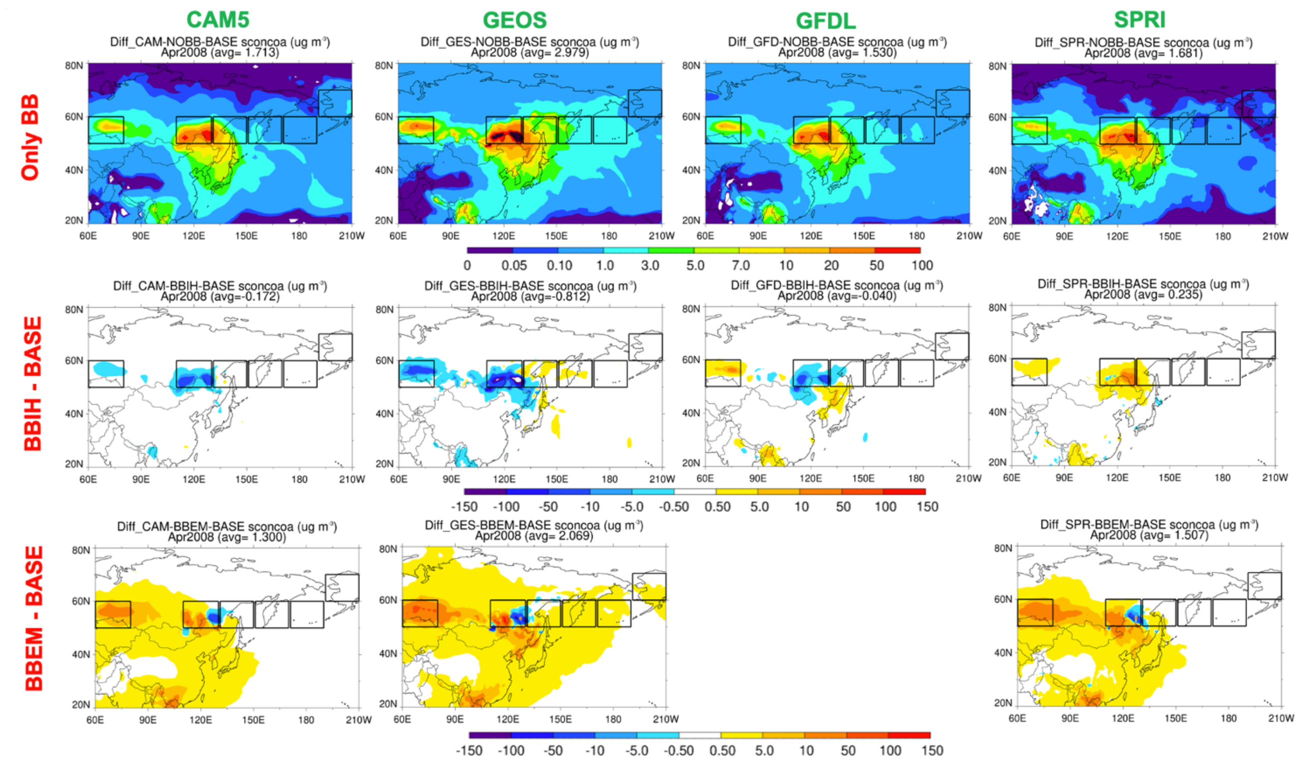


Figure 8. Spatial distribution of differences in surface OA concentrations for April 2008 across four models: CAM5, SPRI, GEOS, and GFDL. **Row 1:** Only BB (BASE minus NOBB). **Row 2:** BBIH minus BASE. **Row 3:** BBEM minus BASE. Note that only CAM5, GEOS, and SPRI provided BBEM simulations.

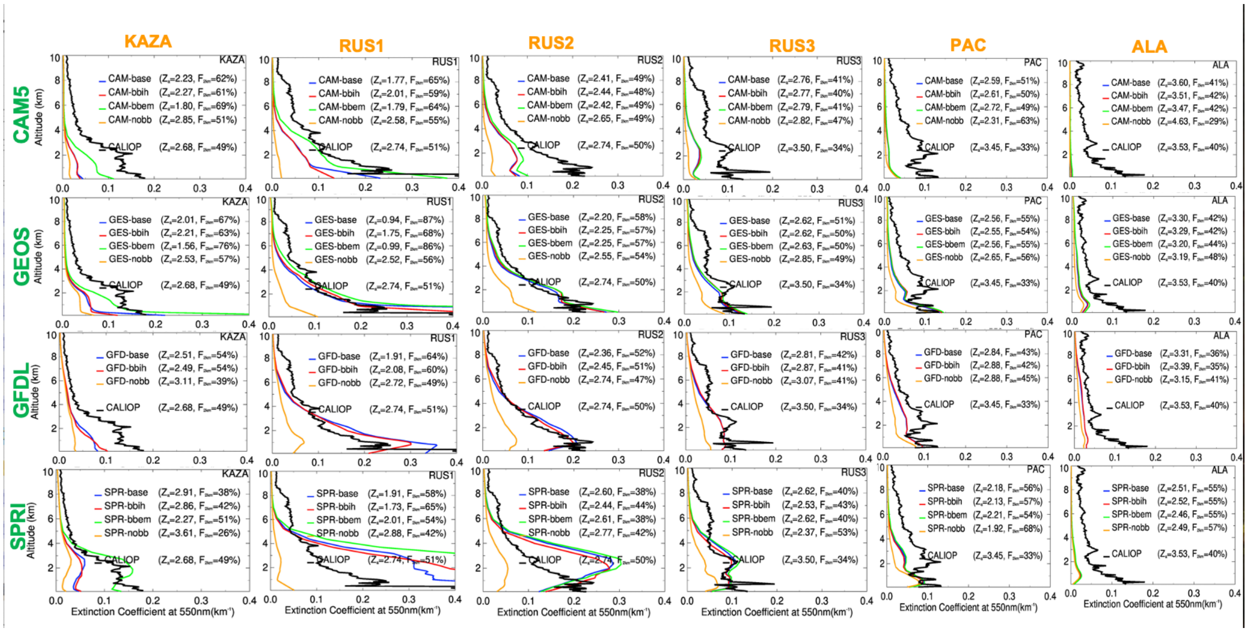


Figure 9. Vertical profiles of aerosol extinction in source and downwind regions. Aerosol extinction profiles for April 2008 from four models (CAM5, GEOS, GFDL, and SPRI), averaged over six regions. **Column 1-2:** KAZA and RUS1 (source regions); **Columns 3–6:** RUS2, RUS3, PAC, and ALA (downwind regions). Each panel includes CALIOP observations (thick black curves) and model outputs from four experiments—BASE, BBIH, BBEM, and NOBB—shown as colored curves. Summary statistics are listed beside the legend: Z_a (mean aerosol layer height) and F_{2km} (fraction of AOD within the lowest 2 km.)

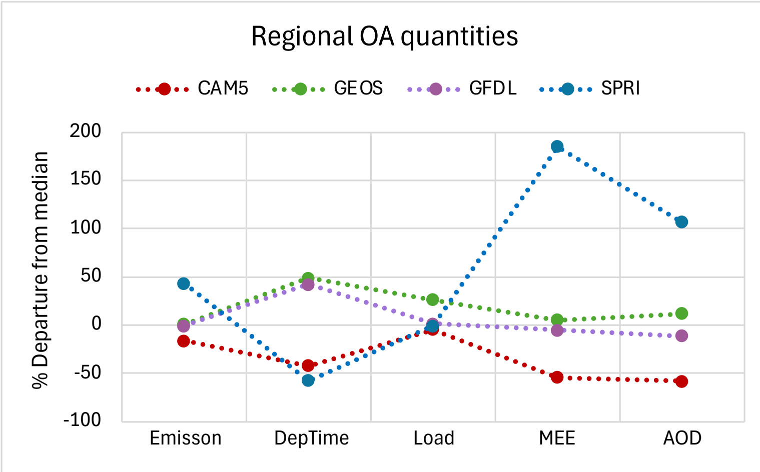


Figure 10. Comparisons of model-simulated key variables determining OA AOD in each model for April 2008, averaged over four regions from RUS1 to PAC. Colored symbols represent the percentage deviation of each model from the multi-model median. The actual values from individual models, along with the multi-model statistics (median, IQR/median, and max/min), are listed in Table 5.

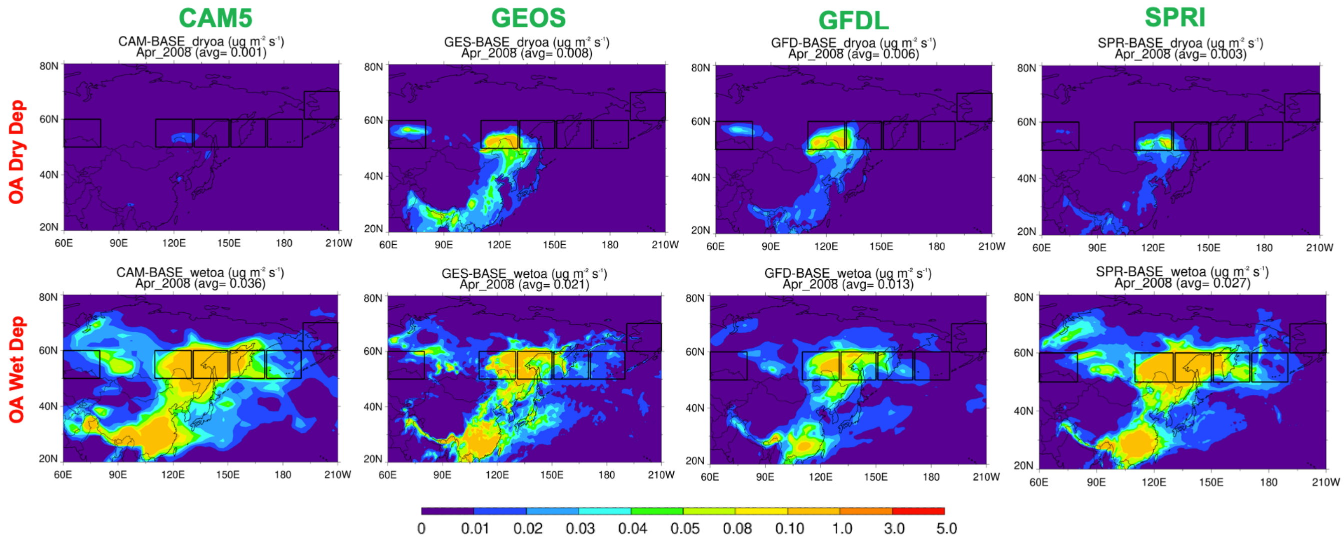


Figure 11. Spatial distribution of OA dry deposition (units: µg m^{-2} s^{-1}) and wet deposition for April 2008, as simulated by four models (CAM5, GEOS, GFDL, and SPRI) in their BASE runs.