We are grateful to the editor and referee for their time and energy in providing helpful comments and guidance that have improved the manuscript. In this document, we describe how we have addressed the reviewer's comments. Referee comments are shown in black italics and author responses are shown in blue regular text.

Reviewer #2:

In this paper, Hu et al. documented an important effort of generating a global daily fire-sourced $PM_{2.5}$ dataset from 2000-2023. This dataset is derived using the following steps: 1) They use GEOS-Chem to simulate global daily all-source $PM_{2.5}$ and apply ML to de-bias the all-source $PM_{2.5}$ against surface measurements; 2) they then apply the ratio between simulated fire $PM_{2.5}$ and total $PM_{2.5}$ at the grid level to estimate the fire-sourced $PM_{2.5}$. I think this is an important effort, and the dataset generated will become a valuable asset to the academic communities across different disciplines. I am excited to see such a dataset being made publicly available to the community.

- Thank you for your positive evaluations.

With that being said, I agree with the other reviewer that the paper should do a better job of discussing its difference from the Xu et al., 2023 Nature paper, in addition to the extended temporal period (which is an important update in my opinion). In addition to a transparent and detailed discussion of the differences and contributions, I imagine the project could benefit from several potential analyses to further differentiate this paper from their prior contributions:

1) discuss the recent trends from 2020-2023;

- Thank you for your valuable suggestion. In the revised paper, we added Figure 9 to examine recent changes in fire-related air pollutants over the past four years beyond 2020. We described it as follows: "To better understand recent trends, we examined changes in fire-sourced [$PM_{2.5}$] during the past few years. Relative to 2000-2019, fire [$PM_{2.5}$] decreases across nearly all latitudes from 2020 to 2023 for both inventories (Fig. 9). Regionally, hotspots of increased fire [$PM_{2.5}$] could be found in North America, due to the 2023 Canadian fires, and in the Amazon, due to the 2022 Brazilian fires. Additionally, fire [$PM_{2.5}$] levels increased in central Africa, northern India, and the Indo-China Peninsula, where human-induced agricultural burning is prevalent (Van Der Werf et al., 2017)." (Lines 296-301)

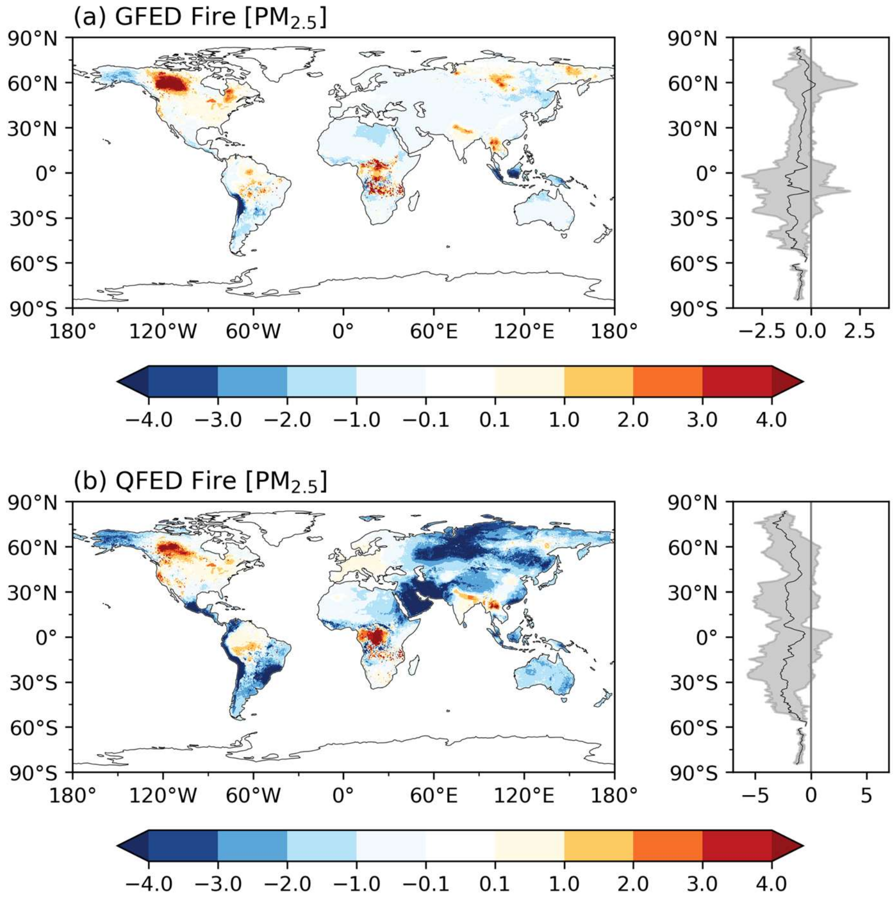

Figure 9. Differences in estimated fire-sourced $PM_{2.5}$ (μg m^{-3}) between 2020-2023 and 2000-2019 derived using (a) GFED and (b) QFED inventories. The zonal averages and one standard deviation are shown alongside each panel.

2) discuss the differences between the GFED and QFED-based dataset (which do not seem to be the focus of their prior work);

- Yes, the inclusion of two inventories is one of the major contributions of our study to the community. In the revised paper, we added Figure S8 to compare the differences in fire [$PM_{2.5}$] between the two inventories at various percentiles. This comparison helps explain why the two datasets exhibited comparable performance for fire episodes, despite a large difference in their mean values: "The probability density distributions of fire-sourced [$PM_{2.5}$] from the two inventories show notable differences (Fig. S8). During 2000-2023, fire [$PM_{2.5}$] from QFED is more than twice that from GFED below the 75th percentile, indicating that QFED predicts significantly higher [$PM_{2.5}$] for low to moderate fire events. However, this difference diminishes above the 90th percentile and becomes particularly constrained at the 99th percentile, where fire-sourced [$PM_{2.5}$] from GFED is 79.29% of that from QFED. It suggests that while both inventories yield comparable estimates for extreme fire episodes, GFED systematically underestimates emissions from smaller fires. This underestimation persists despite improvements in GFED's representation of small fires through additional implementations (Van Der Werf et al., 2017). Consequently, validations in the U.S. reveal substantial low values with GFED relative to observations (Fig. 4), whereas both inventories perform comparably during high-emission fire episodes (Figs. 3 and S7)." (Lines 244-254)

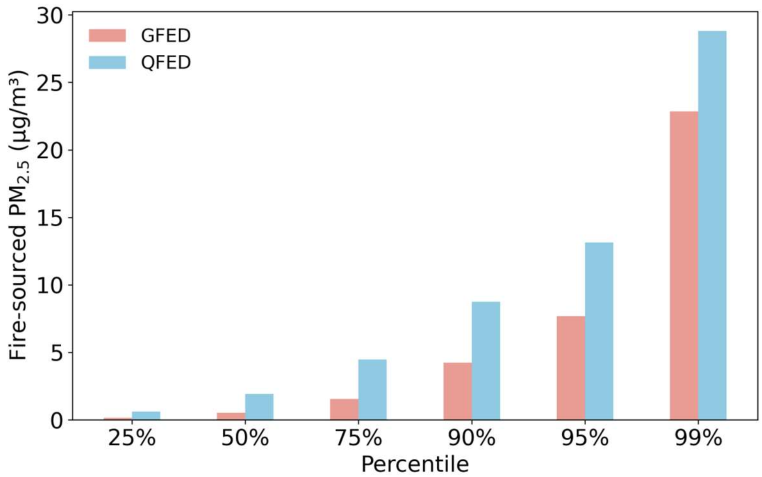

Figure S8. Comparison of daily fire-sourced [$PM_{2.5}$] at different percentiles between simulations with GFED and QFED inventories.

3) a better quantification and discussion of the uncertainty of their datasets.

- In the revised version, we analyzed of the long-term trend of fire-sourced [$PM_{2.5}$] exceeding the WHO health standard with shadings to quantify the uncertainties from two inventories (Figure 8). We also added following discussion to explain the possible causes of the uncertainties in fire emission inventories: "The two datasets derived from different inventories showed discrepancies in both the long-term mean and trend of fire-sourced [$PM_{2.5}$] (Fig. 5). In general, fire-related [$PM_{2.5}$] is much higher when using the QFED inventory compared to GFED, but the long-term trend is more negative with QFED. As expected, these discrepancies can be attributed to differences in the underlying fire emission inventories (Fig. 6), which stem from variations in their estimation methods, data sources, emission factors, and so on (Kaiser et al., 2012; Larkin et al., 2014; Jin et al., 2023). For example, QFED adjusts emission factors based on aerosol optical depth from MODIS (Petrenko et al., 2012; Li et al., 2022), resulting in significantly higher emissions in some regions compared to GFED. In contrast, GFED relies on burning pixels and changes in surface reflectance identified during satellite overpasses under relatively cloud-free conditions, which may lead to underestimating burned areas especially for some small fires (Pan et al., 2020)." (Lines 347-357)

In the revised discussion, we explicitly outlined how our study made further progresses compared to Xu et al. (2023):

"We employed a similar approach to Xu et al. (2023) but incorporated new datasets and perspectives. First, we used global observed $PM_{2.5}$ concentrations from 9541 monitoring sites, significantly more than the 5661 stations used in Xu et al. (2023). The expansion of ground-based stations, particularly in fire-prone regions such as Africa and South America, strengthens the foundation for model training and data validation. Second, we applied two different fire emission inventories. Comparisons showed that fire [$PM_{2.5}$] estimates from these inventories were consistent during extreme fire episodes (Figs 3 and S7). However, for low to moderate fire emissions, fire [$PM_{2.5}$] from GFED was much lower than that from QFED (Fig. S8), suggesting that global population exposure to fire-related air pollution may have been underestimated in Xu et al. (2023) due to the application of GFED. Third, we extended the ending simulation year from 2019 to 2023, capturing an additional four years that included unprecedent fire events, such as the 2023 Canadian fires and the 2022 Brazilian fires. These events provide valuable data for assessing population exposure and associated health impacts. Fourth, we found a global decreasing trend in fire [$PM_{2.5}$] during 2000-2023, which contrasts with the increasing trend reported in Xu et al. (2023). This discrepancy may stem from differences in machine learning approaches (random forest vs. XGBoost in this study), pollution definitions (population-weighted vs. non-weighted), and observational datasets. Despite these differences, both studies identified a turning point in 2017, after which global fire [$PM_{2.5}$] began to increase, with the most pronounced rise observed in boreal regions." (Lines 330-346)

Other comments:

1) One recent study found that GFED emissions inventory had a large positive bias in the western US in 2020. This raises potential concerns about the author's methodology of using model-based fire/total ratios to derive their fire-source $PM_{2.5}$ estimates. I think the paper could benefit from a more detailed discussion of this assumption.

https://pubs.acs.org/doi/10.1021/acs.est.4c05922

➢ In the revised paper, we added following discussion to acknowledge the uncertainties from chemical transport model: “Third, biases in the [$PM_{2.5}$] simulated by the GC model may significantly affect the accuracy of machine learning. Predicting air pollutants involves uncertainties due to variations in meteorological forcing, chemical and physical schemes, initial and boundary conditions, and so on. For example, Qiu et al. (2024) found that the GC model significantly overestimated [$PM_{2.5}$] during extreme wildfire events in 2020 over the western U.S. In contrast, our derived fire [$PM_{2.5}$], using the same GFED inventory, is much lower than the estimates of Childs et al. (2022) for low to median fire events (Fig. 4). These findings suggest that incorporating more validated fire inventories and/or chemical models is necessary to better quantify the uncertainties in derived air pollutant concentrations.” (Lines 378-386)

2) Related to the comment above, I think the paper could benefit from a comparison with the more refined regional fire smoke $PM_{2.5}$ estimates. For example, the authors could compare their two estimates with data from Childs et al (which was recently updated to include 2021-2023) in North America.
https://www.stanfordecholab.com/wildfire_smoke

- ➢ Thank you for your valuable suggestion. In the revised version, we added Figure 4 and related descriptions to validate the derived fire [$PM_{2.5}$] from both inventories against estimates from Childs et al. (2022): “We further compare the fire-sourced [$PM_{2.5}$] data with the estimates by Childs et al. (2022) in the U.S. (Fig. 4). Our estimates show reasonable performance, with correlation coefficients of 0.68 (0.6) and RMSE of 2.79 (2.71) μg m^{-3} using the GFED (QFED) inventory. However, fire-sourced [$PM_{2.5}$] form GFED is overall lower than that of Childs et al. (2022) by -55.04%.” (Lines 239-243)

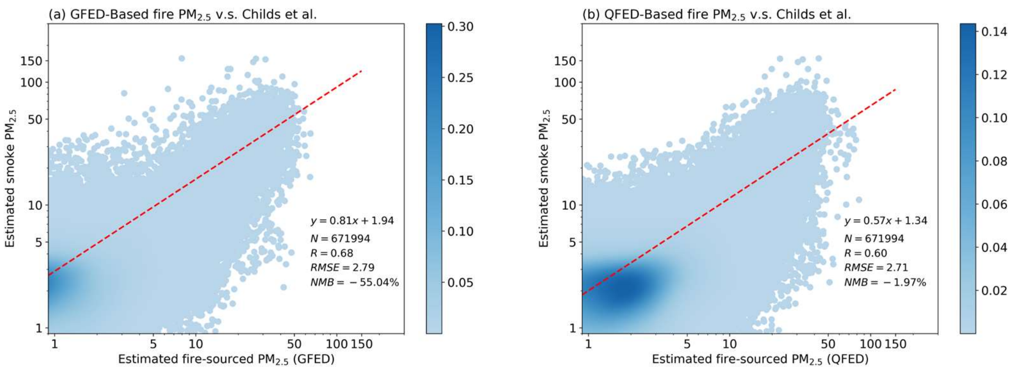

Figure 4. Comparison of fire-sourced $PM_{2.5}$ (μg m^{-3}) estimated using (a) GFED and (b) QFED inventories with smoke $PM_{2.5}$ observed by Childs et al. (2022) at 100156 polygons in U.S. during 2016–2019. Validation metrics of N, regression equation, R^2, RMSE, and NMB are calculated.

3) The difference between GFED and QFED-based estimates is so large that I think it warrants a more in-depth discussion of the potential reasons. Could the authors draw on prior research that evaluates these emission inventories to discuss the potential reasons?

For example: https://doi.org/10.5194/acp-20-969-2020

- In the revised version, we added following discussion to explain the possible causes of the uncertainties in fire emission inventories: "The two datasets derived from different inventories showed discrepancies in both the long-term mean and trend of fire-sourced [$PM_{2.5}$] (Fig. 5). In general, fire-related [$PM_{2.5}$] is much higher when using the QFED inventory compared to GFED, but the long-term trend is more negative with QFED. As expected, these discrepancies can be attributed to differences in the underlying fire emission inventories (Fig. 6), which stem from variations in their estimation methods, data sources, emission factors, and so on (Kaiser et al., 2012; Larkin et al., 2014; Jin et al., 2023). For example, QFED adjusts emission factors based on aerosol optical depth from MODIS (Petrenko et al., 2012; Li et al., 2022), resulting in significantly higher emissions in some regions compared to GFED. In contrast, GFED relies on burning pixels and changes in surface reflectance identified during satellite overpasses under relatively cloud-free conditions, which may lead to underestimating burned areas especially for some small fires (Pan et al., 2020)." (Lines 347-357)

4) Generating these two datasets (GFED and QFED-based) is an interesting contribution that allows researchers to evaluate the potential uncertainty. However, downstream users often just want to use the best available dataset. What can the authors say in terms of which one they recommend more or less? Also, I think the Xu et al. Nature 2023 paper considered other emissions inventories in their sensitivity analyses, why did the authors decide to focus only on GFED and QFED in this work?

- Xu et al. (2023) compared the uncertainties in fire-sourced [$PM_{2.5}$] from different inventories, but only for the year 2012. In contrast, our study provided datasets from both inventories spanning 2000-2023, enabling us to compare their spatiotemporal variations. Initially, we considered using the FINN and GFAS inventories as well. However, these datasets were slower to update and had incomplete emission data for the timeframe of our study. The GFED and QFED inventories, being more efficient in the data update, were ultimately chosen. Based on our evaluations, we find that derived fire [$PM_{2.5}$] from both inventories are reasonable (though with some biases). "Therefore, we recommend using the average of fire-sourced [$PM_{2.5}$] from both inventories to indicate the mean state, while using their difference as the range of uncertainties associated with fire-related air pollutants." (Lines 364-366)

References

Childs, M. L., Li, J., Wen, J., Heft-Neal, S., Driscoll, A., Wang, S., Gould, C. F., Qiu, M., Burney, J., and Burke, M.: Daily Local-Level Estimates of Ambient Wildfire Smoke $PM_{2.5}$ for the Contiguous US, Environmental Science & Technology, 56, 13607-13621, https://doi.org/10.1021/acs.est.2c02934, 2022.

Jin, L., Permar, W., Selimovic, V., Ketcherside, D., Yokelson, R. J., Hornbrook, R. S., Apel, E. C., Ku, I. T., Collett Jr, J. L., Sullivan, A. P., Jaffe, D. A., Pierce, J. R., Fried, A., Coggon, M. M., Gkatzelis, G. I., Warneke, C., Fischer, E. V., and Hu, L.: Constraining emissions of volatile organic compounds from western US wildfires with WE-CAN and FIREX-AQ airborne observations, Atmos. Chem. Phys., 23, 5969-5991, https://doi.org/10.5194/acp-23-5969-2023, 2023.

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J. J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527-554, https://doi.org/10.5194/bg-9-527-2012, 2012.

Larkin, N. K., Raffuse, S. M., and Strand, T. M.: Wildland fire emissions, carbon, and climate: U.S. emissions inventories, Forest Ecology and Management, 317, 61-69, https://doi.org/10.1016/j.foreco.2013.09.012, 2014.

Li, F., Zhang, X., Kondragunta, S., Lu, X., Csiszar, I., and Schmidt, C. C.: Hourly biomass burning emissions product from blended geostationary and polar-orbiting satellites for air quality forecasting applications, Remote Sensing of Environment, 281, 113237, https://doi.org/10.1016/j.rse.2022.113237, 2022.

Pan, X., Ichoku, C., Chin, M., Bian, H., Darmenov, A., Colarco, P., Ellison, L., Kucsera, T., da Silva, A., Wang, J., Oda, T., and Cui, G.: Six global biomass burning emission datasets: intercomparison and application in one global aerosol model, Atmos. Chem. Phys., 20, 969-994, https://doi.org/10.5194/acp-20-969-2020, 2020.

Petrenko, M., Kahn, R., Chin, M., Soja, A., Kucsera, T., and Harshvardhan: 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, https://doi.org/10.1029/2012JD017870, 2012.

Qiu, M., Kelp, M., Heft-Neal, S., Jin, X., Gould, C. F., Tong, D. Q., and Burke, M.: Evaluating Chemical Transport and Machine Learning Models for Wildfire Smoke $PM_{2.5}$: Implications for Assessment of Health Impacts, Environmental Science & Technology, 58, 22880-22893, https://doi.org/10.1021/acs.est.4c05922, 2024.

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, Earth Syst. Sci. Data, 9, 697-720, https://doi.org/10.5194/essd-9-697-2017, 2017.

Xu, R., Ye, T., Yue, X., Yang, Z., Yu, W., Zhang, Y., Bell, M. L., Morawska, L., Yu, P., Zhang, Y., Wu, Y., Liu, Y., Johnston, F., Lei, Y., Abramson, M. J., Guo, Y., and Li, S.: Global population exposure to landscape fire air pollution from 2000 to 2019, Nature, 621, 521-529, https://doi.org/10.1038/s41586-023-06398-6, 2023.