Reviewer 1

This manuscript demonstrates the application of a method that combines information from surface monitors, satellites, and a 3D model to estimate wildfire smoke impacts over a 17-day time period. There are significant issues with the writing and some aspects of the science, so I feel the manuscript requires major revisions before it can be published.

Major comments

In general, the writing is imprecise and there are quite a few typos. I've pointed out many examples in the specific comments. These issues caused the review to take about twice as long as it would have if the authors had edited their manuscript better. It should not be the reviewers' responsibility to edit writing mistakes and imprecision, and it adds to the stress on the peer review system.

We sincerely apologize for the writing issues and typos in the manuscript, which placed an additional burden on your review. We have carefully revised the manuscript to improve precision and clarity, addressing all specific comments and conducting a thorough proofreading to ensure quality. Thank you for your time and effort in reviewing our work.

Section 4.3 and equation 9. The method will fail when a significant fraction of the smoke impact is out of the PBL. For example, if the AOD without smoke is 0.1 and the AOD with smoke is 1.0, but 100% of the smoke is in the free troposphere, the real impact of removing smoke should have no effect on surface PM2.5. However, equation 9 would attribute 90% of the surface PM2.5 to smoke (as opposed to the correct value, 0%). This erroneous calculation could have a large impact on the results in Table 3, which has some of the main results of the paper that show up in the abstract.

Thank you for raising this point. We agree that the $\frac{PM_{2.5}}{AOD}$ ratio may not accurately represent the smoke contribution to surface PM2.5 when a significant fraction of smoke resides above the PBL. To address this, we have revised our method to avoid directly using the $\frac{PM_{2.5}}{AOD}$ ratio. Instead, we now estimate the surface PM2.5 for the control case using the ratio between the WRF-Chem-simulated surface-level P2.5 concentrations of the experiment run (with Canadian smoke) and control run (without Canadian smoke), as follows:

$$PM_{2.5,control} = PM_{2.5} * \left(\frac{PM_{2.5,WRF,control}}{PM_{2.5,WRF,experiment}} \right)$$

The revised method uses the modeled surface P2.5 concentrations from WRF-Chem for both the control and experiment cases, which inherently accounts for the vertical distribution of smoke. By doing so, it avoids the assumption of directly using AOD to estimate surface-level P2.5. We have updated the Section and equation 9, the results in the abstract have also been updated accordingly.

Sections 5.1-5.4 seem very scattered, and I am not currently seeing how they contribute to the main points of the paper. I recommend removing some of these sections and better integrating the remaining parts with the following sections.

Thank you for your valuable feedback. To enhance the coherence of the manuscript, we have removed Sections 5.3 and 5.4 as suggested and integrated the remaining content more effectively with the subsequent sections.

How is the modeled AOD defined in a cloudy column? Aerosol scattering blows up as f(RH) approaches 100%. How is this handled in a cloudy gridbox and how do the assumptions affect the results?

AOD calculation in WRF-Chem is based on chemical species and does not directly account for the effects of RH or cloud droplets. Aerosol optical properties calculations in WRF-Chem rely on a sectional approach and assume dry aerosol conditions (Barnard et al., 2010). As mentioned in section 3.2, we used four size bins (0.039-0.156, 0.156-0.625, 0.625-2.500, and 2.5-10.0 μm dry diameters) in the simulation. The particle mass and particle number are simulated separately for each bin. Importantly, since the bins are defined based on dry particle diameters, water uptake or loss does not cause particles to transfer between bins. In each bin, the particles are assumed to be spherical and internally mixed. Given this simplification, the conversion from chemical to optical properties follows the following steps:

1. Aerosols are divided into size bins, and the chemical masses $M_{i,j}$ and particle number N_i are calculated for each size bin, i is the bin number and j is the chemical species.

2. The volumes of these chemicals are computed for each bin $V_{i,j} = \frac{M_{i,j}}{\rho_i}$ $(unit: \frac{cm^3}{cm^3\ dry\ air})$

3. The physical diameter of each bin can be calculated using: $D_{p,i} = 2\left(\left(\frac{\sum_j V_{i,j}}{N_i}\right)/\left(\frac{4}{3}\pi\right)\right)^{\frac{1}{3}}$, the aerosol size distribution is defined by N_i amd $D_{p,i}$

4. Now that we have the size distribution, the bulk refractive index of the particles in a bin can be computed, $m_{s,i}$ and $m_{c,i}$ are the bulk complex refractive indices of the shell and core for bin "i". The core refractive index is assigned the value of 1.85+0.17i, and the shell refractive index is given by:

$$m_{s,i} = \frac{\sum_{j=1,j\neq BC}^{j} m_j V_{i,j}}{\sum_{j=1,j\neq BC}^{j} V_{i,j}}$$

5. Shell/core Mie theory is then used to find the absorption $Q_{a,i}$ and scattering efficiency $Q_{s,i}$, then other optical properties such as absorption B_{abs} and scattering coefficient B_{scat} can be calculated to get the AOD values:

$$B_{scat} = \sum_{i=1}^{8\ bins} N_i Q_{s,i} \pi \left(\frac{D_{p,i}}{2}\right)^2$$

$$B_{abs} = \sum_{i=1}^{8\ bins} N_i Q_{a,i} \pi \left(\frac{D_{p,i}}{2}\right)^2$$

$$B_{ext} = B_{scat} + B_{abs}$$

$$AOD = \int B_{ext}(z) dz$$

Details of the "aerosol chemical to AOD" calculation processes can be found in (Barnard et al., 2010).

There are several assumptions can affect the results:

1. The chemical mass inputs for AOD calculation depend on the emission inventory used (e.g., FINN in our study). Errors or biases in the fire emissions inventory can directly propagate into the AOD estimates.
2. The AOD in cloudy grid boxes is effectively a dry AOD because RH effects are not included in the calculations. Post-processing is needed to account for hygroscopic growth, using model output for aerosol water content and the hygroscopicity parameter to estimate how scattering increases with RH. Neglecting hygroscopic growth can lead to an underestimation of AOD, particularly in humid or cloudy regions.
3. The model assumes spherical, internally mixed particles, which may not accurately represent real-world aerosol shapes or external mixing, potentially causing errors in optical property calculations.
4. Due to assumptions about particle size distributions, overestimation of optical properties can occur for coarse-mode particles, particularly in bins with diameters >2.5 µm.

Specific comments

Title: Is "measurements" also applying to "satellite"? If not, then it should be "satellites". If yes, then it's confusing to have "model" in the middle.

We have revised the title for clarity. The new title is: "An Investigation of the Impact of Canadian wildfires on US Air Quality using Model, Satellite and Ground Measurements".

L8: 10-m grid spacing is impressive.

Thank you for catching this error. We have corrected it to 10-km grid spacing.

L16: unclear what these numbers mean. 62 ug m-3 would be far more than a 13% increase considering typical non-smoke concentrations in the US are <10 ug m-3.

To clarify, the number here represents the PM2.5 increase caused by Canadian wildfires only during the study period. There were also local wildfires in the US contributing to the total PM2.5 concentrations, but this value isolates the impact of Canadian wildfire smoke on PM2.5.

Based on the major comments above, we recalculated the surface PM2.5 for the control case. As a result, the PM2.5 increase values have been updated in the revised manuscript to ensure accuracy and clarity.

L22: 92 mortalities over what time period?

The time period is the summer of 2020. We have revised the sentence for clarify as follows: "For regions affected the most by wildfires, like Washington state, an increase in daily PM2.5 of 97.1 $\mu g m^{-3}$ during the summer of 2020 was found, which related to 92 more mortality cases."

L23: I'd say "has been found to be 3-4 times greater" given that this is just one study and smoke and non-smoke composition varies.

Thank you for the suggestion. The sentence has been revised accordingly to reflect this point.

L25-28: Is this sentence about wildfire smoke PM2.5 or PM2.5 in general?

This sentence refers to PM2.5 in general. We have reorganized the paragraph to ensure a more logical flow.

L30-31: Using past tense in a sentence about the future. I'd say "have been projected to increase from…".
Thank you for catching this in consistency. The sentence has been corrected to: "From simulation results, wildfire-related economic costs have been projected to increase from $7 billion per year to $43 billion per year in 2090."

L32 and in general: Focus on making the topic sentence be an intro to the full paragraph. This paragraph moves a long way from FRP and injection heights.

We have revised the first sentence to better introduce all the contents of the paragraph and adjusted the paragraph for clarity and coherence. The revised paragraph now reads:
"
The transport and dispersion of wildfire smoke are influenced by multiple factors, including fire intensity, injection height, atmospheric dynamics, and terrain interactions. Higher fire radiative power (FRP) results in longer smoke transport distances due to higher plume injection heights. A global analysis of over 23,000 wildfires found that significant injection heights into the free troposphere are primarily observed in the boreal forests of North America and Siberia during the northern summer. Smoke remnants from Canadian stand-replacing forest fires have been observed at altitudes exceeding 13 km. Once injected, smoke plumes can descend to the

surface through a combination of subsidence, interception, and diurnal entrainment within the Planetary Boundary Layer (PBL), as observed in the Eastern US for 2002 July Canadian forest fire event (Colarco et al. 2004).

Atmospheric circulation plays a key role in shaping smoke transport. Upper-level winds facilitate long-range horizontal transport, while surface high-pressure systems enhance ground-level pollution through subsidence inversions (Miller et al. 2011). Cyclonic circulation can form a multilayer PBL, characterized by temperature inversions and stable stratification, which trap pollutants in convergent zones (Y. Jiang et al. 2021). Interactions with mountain terrain further modulate smoke dispersion: under stable synoptic conditions, valleys become more stagnant, while unstable conditions promote vertical mixing (Beaver et al. 2010; Lang, Gohm, and Wagner 2015)."

L33-36: Boreal injections have been found to be higher on average than tropics in the MISR plume height climatology (https://www.mdpi.com/2072-4292/10/10/1609), so the contrasting of limited cases seems to give the wrong impression.

Thank you for pointing this out. To avoid confusion, we have included the global analysis of injection heights from (Val Martin et al., 2018) to provide a broader context. Additionally, we used a case study of a Canadian wildfire event to illustrate that boreal wildfire in North America, particularly during the northern summer, generally exhibit higher injection heights. This clarification helps to highlight the regional significance of boreal forest fires without overstating individual cases.

L47: Volatile organic compounds, by definition of being volatile, do not condense. Some of the oxidation products of VOCs can condense. Also, the citations here are not dedicated investigations of the specific processes listed in this sentence, though these studies do exist.

Thank you for your comments. We agree that the original terminology could be misleading. What we intended to convey is that as smoke cools away from the flame front, some semi-volatile gases, including oxidation products of volatile organic compounds (VOCs), can condense onto existing emitted particles, forming organic or inorganic coatings (Junghenn Noyes et al., 2021). These coatings increase particle size and can significantly affect aerosol optical properties. We have revised the sentence for clarity as follows:

During transport, smoke particles undergo physical and chemical transformations that influence their sizes, such as hygroscopic growth (Carrico et al., 2005; Gomez et al., 2018), SOA formation (Ahern et al., 2019) condensation of semi-volatile species, and coagulation process (Aloyan et al., 1997; Sun et al., 2019).

L54-59: Of what data sources?

Spatial interpolation methods rely on ground-based PM2.5 measurements, while linear regression models use satellite-derived aerosol optical depth (AOD) along with surface PM2.5 measurements. Multi-linear regression methods typically combine surface PM2.5 measurements, satellite AOD, and meteorological datasets from models. Geographically weighted regression (GWR) and machine learning approaches draw from similar data sources as multi-linear regression, integrating surface PM2.5, satellite AOD, and meteorological inputs to capture spatial and temporal variations in PM2.5 concentrations.

L59: Which of the approaches above are the traditional ones?

The traditional approaches mentioned here refer to spatial interpolation and linear regression methods, which have been widely used in earlier studies for estimating PM2.5 concentrations. We revised this paragraph to clarify:
To assess surface pollution levels, various approaches have been used for estimating PM2.5 concentrations, each utilizing different data sources. Traditional methods such as spatial interpolation (e.g., inverse distance weighting, ordinary kriging) rely primarily on ground-based PM2.5 measurements, while linear regression methods combine satellite-derived AOD with surface PM2.5 measurements (Hoff and Christopher 2009). More advanced methods, like multi-linear regression, typically incorporate surface PM2.5 data, satellite AOD, and meteorological datasets from models (Gupta and Christopher 2009b). Geographically weighted regression (GWR) and machine learning methods use similar data sources as multi-linear regression but offer better spatial resolution and adaptability to local variations (Xue, Gupta, and Christopher 2021; Ma et al. 2014; Bai et al. 2016; Song et al. 2014). Linear mixed-effect models add temporal variability by including both fixed and random effects (Ma et al. 2016; Lee et al. 2011), while chemistry transport models (CTMs) leverage detailed atmospheric chemistry and physics simulations using emissions inventories, meteorological fields, and chemical species distributions (Geng et al. 2015; Xue et al. 2019). Traditional approaches, such as spatial interpolation and linear regression, are limited in their ability to combine different mechanisms and add various variables with spatial-temporal information to improve their prediction accuracy, which newer techniques address (Zhang, Rui, and Fan 2018). Due to the growth of computing power, machine learning (or artificial intelligence) has become a major focus for estimating the spatial-temporal dynamic distribution of surface PM2.5 concentrations (Zhang, Rui, and Fan 2018; Sayeed et al 2022).

L63: Strange to say that AOD is an *essential* indicator of surface pollution since it is a column measurement. There are much more direct ways to get surface PM (e.g., an in situ measurement), so AOD is not essential. The PM2.5:AOD ratio should be discussed in this paragraph.

We agree that AOD is not a direct indicator of surface PM2.5 and have revised the text to clarify that AOD serves as a valuable proxy, particularly in regions or periods with limited in situ measurements. We have also expanded the discussion to include the PM2.5:AOD relationship, emphasizing how it helps link columnar AOD measurements to surface PM2.5 concentrations.

L85-87: This sentence makes it seem like there are 2 different WRF-Chem domains, implying 2 different simulations. In the later WRF-Chem description, it seems like this is not the case.

Thank you for pointing out this potential confusion. We have revised the sentence to clarify that the study area corresponds to the inner domain, while the larger outer domain includes remote fire sources, both within a single WRF-Chem simulation.

"The study area (inner domain) focuses on the US (25–50°N, 64–125°W), while the outer domain of the same WRF simulation extends to Canada (25–67°N, 70–140°W) to account for Canadian fire emissions and their contribution to US pollution from remote fire sources."

L116: T2M, U10M, and SP are not directly related to the AOD:PM2.5 ratio. I would guess that including them could lead to overfitting.

The selection of predictor variables is based on our previous study (Xue et al., 2021), where we successfully estimated surface PM2.5 using the same set of variables (AOD, BLH, T2M, U10M, RH, SP) for the same time period (August 9th to 25th, 2018). In that study, we performed leave-one-out cross-validation (LOOCV) to assess overfitting. The difference in R^2 between the fitting and validation was minimal (0.037), suggesting that overfitting was not a concern. Therefore, we assumed the same approach would be valid for this study.

L128: Should “in-line” be “online”?
L130: Both aerosols and clouds have microphysical processes (condensation, coagulation, etc.), so please say “aerosol-cloud interaction” here. Similarly, since aerosols have chemical and physical processes, what does “interactions between chemistry, aerosols, and physics” mean here?

The original sentence, “WRF-Chem is an in-line atmospheric chemistry model” Is correct. As referenced in (Powers et al., 2017): “WRF-Chem is a WRF-based in-line atmospheric chemistry model.”
To clarify and avoid any potential misunderstanding, we have revised the sentence to: “WRF-Chem is an atmospheric chemistry model that fully integrated with the meteorological framework of WRF, enabling the simulation of various chemical and physical processes related to aerosol transport, including dispersion, aerosol-cloud interactions, and other key mechanisms.”

L140: 2-moment

Corrected.

L146: “All the key species… are included” is a bold statement. Are you sure?

To avoid potential misunderstanding, we have revised the sentence to: “Key species associated with wildfires are included…”

Table 1 or accompanying texts: What are the lat-lon limits of the domain?

We have added a description to clarify the lat-lon limits of the two nested domains in the accompanying text:
“The model simulation was conducted over two nested domains: an outer domain covering the Canadian region (25-67˚N, 70-140˚W) and an inner domain focused on the CONUS region (25-50˚N, 64-125˚ W). ”

L176: Double “the”

Corrected.

Section 4.1.2: How is AOD defined in the CTM in a cloudy column (see major comment)?

Please refer to our reply to the major comments above for a detailed explanation.

L180-182: This sentence is confusing.

We revised the sentence to:
” We applied the Kriging method for interpolation in areas with sufficient AOD information, while using CTM interpolation where valid AOD retrievals were unavailable, to minimize uncertainties associated with small-scale "missingness" of AOD. These uncertainties arise because CTM relies on fire inventories derived from satellite fire detection products, which may fail to capture small-scale fires due to the spatial resolution of satellite observations and fractional fire coverage within a pixel (Fu et al., 2020). Such undetected fire sources can lead to inaccuracies in CTM outputs. To better represent the AOD distribution in regions with small-scale "missingness," we prioritize the Kriging method over CTM interpolation.”

Section 4.2: Please explain how variability in plume height is captured in the 2 methods in this section.

In our study, plume height is indirectly accounted for through the inclusion of input variables. GWR method captures spatial variability by weighting observations based on their geographic proximity. While GWR does not explicitly account for plume height variability, it reflects its effects spatially through the relationship between AOD and PM2.5.
RF method, on the other hand, captures the non-linear relationships between input variables and surface PM2.5. It incorporates the influence of plume height variability on PM2.5

concentrations, as AOD is a columnar measurement linked to the vertical distribution of aerosols.

In both methods, Plume height is not directly modeled as a standalone variable, but its effects are incorporated through the chosen predicters.

Section 4.3: See major comment.

Please refer to our reply to the major comments above for a detailed explanation.

L264-265: Is this other study investigating the exact same time period? If not, is the comparison relevant?

The other study referenced here investigates a wildfire event from June 6–12, 2015, while our study focuses on August 2018. Although the time periods differ, the comparison remains relevant as both studies explore the impact of Canadian wildfire smoke on the US. The comparison highlights the consistent role of long-range transported Canadian smoke as a significant pollution source during wildfire seasons. By examining AOD values in different years, we provide a broader context for understanding the recurring influence of Canadian wildfires on US air quality.

L268-272: These several sentences were all confusing.

We have revised these sentences to clarify that the observed AOD increases in the US caused by Canadian smoke are recurring phenomena during wildfire seasons. While the magnitude of AOD increases varies across events and years, these examples serve to provide important background context and underscore the widespread and significant impact of Canadian wildfire smoke on air quality in the US. The revised text now better reflects this focus.

Figure 3. It's very hard to read the text on this figure. Can we see well enough what we should be taking from the maps? Would something like 500 hPa geopotential heights be better for showing the steering of smoke? Should the maps include Canada?

Thank you for your feedback. To address this, we have changed the surface weather maps in Figure 3 for easier visualization and improved clarity. We hope the updated figure better conveys the key information.

L281: Is there an estimate of dry dep? Normally dry dep is fairly slow for accumulation-mode sizes.

Estimation of dry deposition or sensitivity tests for its parameterization is beyond the scope of this study. In our simulations, aerosol dry deposition follows the scheme by Binkowski and Shankar (Binkowski & Shankar, 1995), where dry deposition velocity is calculated using gravitational settling velocity (V_g), aerodynamic resistance (R_a), and surface resistance (R_s).

This parameterization has been found to overestimate deposition for accumulation-mode aerosols (Emerson et al., 2020; Ryu & Min, 2022). While several improved schemes address overestimation of deposition velocity, these improvements are more effective for larger aerosols (e.g., 2.5–10 μm in the fourth size bin) and have limited impact for smaller aerosols (diameter <2.5 μm) (Ryu & Min, 2022).

Section 5.3 and 5.4. These just show special cases for specific locations while the rest of the paper shows smoke affects more broadly. I don't think these sections add much, and the manuscript may benefit from cutting them.

Thank you for your feedback. To enhance the coherence of the manuscript, we have removed Sections 5.3 and 5.4 as suggested and integrated the remaining content more effectively with the subsequent sections.

Figure 6 needs axis labels. Also, neither the caption nor the text says which day it's for. Figure 10 gets discussed before Figure 9.

We have corrected the order of the figures in the manuscript so that Figure 9 now appears before Figure 10.

L365: "AOD is crucial" same comment about the use of "essential" earlier.

We have revised the sentence to:
Analyzing daily AOD coverage is essential, as satellite-derived AOD serves as a columnar indicator of pollution with extensive spatial coverage and high-resolution data, making it a valuable predictor for estimating surface PM2.5 concentrations alongside other variables.

L372-374: It's not clear to me that these are actually good-enough performance values, particularly the RMSEs. Can you give more context to these values?

The comparison between model AOD and satellite AOD in our study focuses on a high-pollution period (wildfire events), a broad spatial domain (entire CONUS region), and fine temporal resolution (daily). These factors inherently introduce additional challenges in achieving high agreement, as model AOD often struggles to accurately capture satellite-observed AOD due to various uncertainties in model simulations, including fire emissions, aerosol properties, and transport processes. This is why we do not directly rely on model-simulated AOD for surface PM2.5 estimations in our analysis.

For context, a study of 550 nm AOD comparing MODIS satellite data with WRF-Chem AOD for eastern North America during August 2008, using a 12 km spatial resolution, reported an R value of 0.689 (Crippa et al., 2017). In contrast, our R values range from 0.3 to 0.63. While this may seem lower, it is important to note that our comparison is performed under more challenging conditions, including daily temporal resolution and the presence of extreme events such as wildfire smoke. These factors naturally lead to increased variability, and thus slightly lower R values are expected.

Additionally, in a study evaluating wildfire smoke during the Williams Flats fire (August 4–8, 2019), R values for comparisons between modeled and MAIAC AOD ranged from 0.1 to 0.5, with RMSE values between 0.1 and 0.3 (Ye et al., 2021). However, the Williams Flats fire was relatively small compared to the Canadian wildfires in August 2018 examined in our study. Furthermore, their study domain was limited to a smaller 10° × 5° region, while ours spans the entire CONUS, making our study inherently more complex.

These comparisons emphasize the additional challenges faced in our analysis. While our R values and RMSE reflect the difficulties of modeling AOD during extreme events, they remain consistent with expectations for such scenarios and comparable studies.

Figure 11: Please make missing data white or gray rather than blue, which is also the color of low AODs. It's currently very hard to tell what is missing data vs. low AOD.

Thank you for your suggestion. We have updated Figure 11 to represent missing data in light gray, ensuring a clear distinction from areas of low AOD, which remain blue.

L418: Units of RMSE.

Corrected.

Figure 14 and Tables 2 and 3: Many (most?) readers won't know the EPA regions (at least not precisely). There needs to be a map showing the regions.

Thank you for your suggestion. We have added an EPA region map to Figure 14 to provide clarity and help readers better understand the spatial context of the regions discussed in Tables 2 and 3.

Figure 14: For RF and GWR, are the averages calculated only from the locations of the ground stations (to be consistent with the average of the ground stations)? Or are the averages from RF and GWR calculated from the entire region (in which case, differences with the monitors could be due to differences in the locations used in the averaging)?

The mean surface PM2.5 estimations for RF and GWR in Figure 14 are calculated over the entire region. The differences between our estimations (RF and GWR) and EPA measurements

primarily arise from the spatial distribution of sample points used for averaging. As discussed in the text, RF is more influenced by the spatial distribution of ground stations, with its regional mean values closely following the trends of EPA station measurements. In contrast, GWR is better able to capture PM2.5 variations in regions where stations are more sparsely distributed.

L454-455: Please explain the "FRP < 10000 MW" threshold. MW is an energy rate. How is a rate summed over a year (should the integral of a rate over time be in Joules, not Watts)? Or is it an average rate over the season?

Thank you for pointing this out. You are correct that FRP represents an energy rate, and when summed or integrated over time, it should be expressed in energy units. In this context, the threshold should be described fire radiative energy (unit: MJ), representing the time-integrated fire radiative energy rather than the instantaneous power. We have updated the text to clarify this and corrected the unit to MJ accordingly.

L458-459: I'm not sure how a 17-day investigation alone can give insight into what is happening in the long-term multi-year trend. There is no investigation of how Canadian smoke has changed during the multi-year time period.

To avoid misunderstanding, we revised this sentence to:
Compared with Table [A1], the 17-day investigation highlights how long-range transported smoke from Canada temporarily offsets the descending trend in surface PM2.5 during the study period. For regions 8 to 10, wildfires (including contributions from both local and remote fires) increase the August mean surface PM2.5 by 0 – 97%. While this study focuses on a short-term event, it demonstrates the significant seasonal impact of Canadian smoke on air quality, emphasizing the need for multi-year investigations to assess long-term trends in Canadian smoke contributions.

L467-481: A very important thing missing from these limitations is uncertainties in the plume heights. (1) It affects horizontal transport speed and direction in WRF-Chem, causing errors in the gap filling. (2) It affects the PM:AOD ratio, and if there's variability in plume height between monitors, this will cause errors in the inferred surface concentrations. (3) It will certainly have issues within equation 9 in getting the smoke fractional contribution to PM, as discussed earlier.

Thank you for pointing this out. We have added the uncertainties related to the injection height to section 5.9:
"Uncertainties in the injection height of smoke plumes in WRF-Chem can impact simulation accuracy. Biases in injection height influence horizontal transport speed, direction, and pollution residence time, potentially introducing errors in the AOD gap-filling process. Moreover, uncertainties in the vertical distribution of aerosols can affect the AOD-PM2.5 relationship, with variations in plume height between monitors leading to inaccuracies in estimated surface PM2.5 concentrations."

L476: Figure 4 seems like more monitors than I would expect from FRM alone. Are you sure you aren't using FRM+FEM monitors?

Thank you for pointing this out. Upon reviewing our data, we confirmed that the monitors used correspond to code 88101, which includes both Federal Reference Methods (FRMs) and Federal Equivalent Methods (FEMs). We also clarified this in section 2.2 of the manuscript as follows:

"We obtained the daily surface PM2.5 concentration product that uses Federal Reference Methods and Federal Equivalent Methods (FRMs and FEMs, with code 88101) from US Environmental Protection Agency (EPA) for CONUS within the study period."

L506: particulate *matter*

Corrected.

L522: I don't see how a 17-day study confirms fires becoming the dominant source.
We have revised this to avoid overstating:
Our study highlights the significant contribution of wildfires to particulate pollution during the study period, aligning with prior research that suggests wildfires are becoming an increasingly important source of particulate pollution as industrial pollution declines due to stringent regulations (Xue et al., 2021). However, further multi-year investigations are needed to robustly confirm this trend on a broader temporal scale.