

Firstly, we would like to thank all referees for their important comments, and we have revised our manuscript accordingly. The original comments from referees are in black, our replies are in blue and the changes in original manuscript are in red.

Reviewer #1:

This study presents a comprehensive analysis of aerosol evolution in the Paris urban plume using innovative coordinated airborne and ground-based observations. The research addresses an important scientific question regarding how aerosol composition and optical properties evolve during 2-6 hours of atmospheric aging, with significant implications for improving aerosol parameterizations in models and remote sensing retrievals. The manuscript is generally well-written with rigorous methodology, but several aspects of the analysis and interpretation require strengthening to fully support the conclusions. I recommend minor revision before publication.

We thank the reviewer's positive comments.

Specific Comments

1. Organic aerosol characterization needs strengthening: The manuscript would benefit from more detailed analysis of organic aerosol sources and oxidation states. While the study correctly identifies organic aerosol as the dominant PM₁ component, the lack of source apportionment (e.g., PMF analysis) and oxidation state analysis (O:C ratios or f₄₄/f₄₃ ratios from mAMS or ToF-ACSM data) limits the mechanistic understanding of the observed changes. These analyses would better support the conclusions about secondary organic aerosol formation and its role in aerosol evolution.

We appreciate the reviewer's suggestion regarding source apportionment analysis. However, PMF generally requires long, compositionally diverse datasets with sufficient temporal variability to derive statistically robust and physically meaningful factors. In the present study, the airborne AMS measurements were acquired during relatively short flights (approximately 2 hours each). Furthermore, airborne observations were specifically designed to follow the evolution of the same urban plume over relatively short transport times. As a result, the organic aerosol composition exhibited only limited spectral variability. Under these conditions, PMF analysis is unable to provide robust or physically interpretable factors. In addition, direct comparison with the PMF or mass spectral results between the ground-based ACSM and airborne AMS may not be straightforward because the mAMS and ACSM differ in instrument design and measurement characteristics. These instrumental differences may influence the resulting factor profiles and their relative contributions.

Following the reviewer's suggestion, we have added the f₄₄/f₄₃ ratio as a qualitative indicator of the relative contribution of more oxidized Org. Combining all the reviewers' comments, we have revised the whole section:

3.2.2 Variation of aerosol components

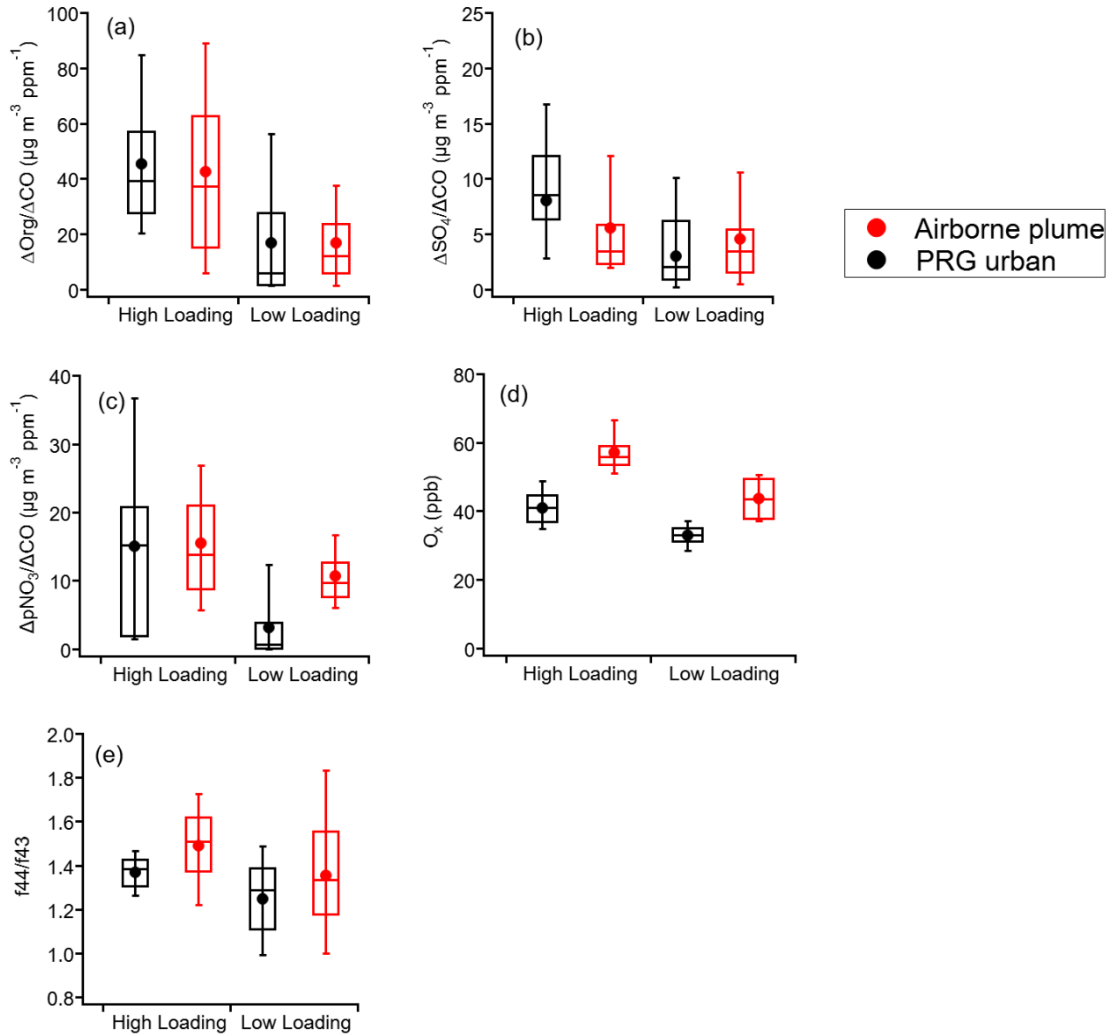


Figure 6. Relative enhancements of (a) organic aerosol (Org), (b) sulfate (SO₄), (c) nitrate (NO₃); (d) O_x (NO₂ + O₃) concentration, (e) f44/f43 measured at the PRG urban site and in the downwind urban plume from airborne observations during heavy and light pollution periods. Results are normalised to ΔCO to evaluate the relative enhancement of each species with respect to primary urban emissions. Boxes show the interquartile range (25th–75th percentiles), whiskers indicate the 10th–90th percentiles, the horizontal line represents the median, and the dots denote the mean values.

Figures 6 and 7 show the normalised NR-PM₁ component concentrations and O_x (NO₂ + O₃), as well as the mass fractions of NR-PM₁ components measured at the PRG urban site and in the downwind urban plume from airborne observations during heavy and light pollution periods. Results in Figure 6 are normalised to ΔCO to evaluate the relative enhancement of each species with respect to primary urban emissions, using CO as a relatively conserved tracer of anthropogenic emissions during short-

range plume transport.

The average O_x concentrations at the PRG urban site differed only slightly between the two aerosol-loading regimes, although the t-test indicated a statistically significant difference ($p < 0.05$). This relatively small difference suggests that the initial photochemical conditions were broadly comparable between the two regimes. Flight A027b, conducted around noon under potentially stronger photochemical conditions, was included in the analysis. Excluding A027b changed the mean downwind O_x concentration by only $\sim 5\%$ and did not alter the overall urban-to-downwind pattern, indicating that this flight had a limited influence on the observed O_x differences. The downwind airborne measurements showed higher average O_x concentrations than those observed at the PRG site (t-test, $p < 0.05$). As O_x is presented here as an absolute concentration rather than normalized to ΔCO , this difference is interpreted as reflecting different photochemical conditions between the urban and downwind air masses rather than as a quantitative measure of photochemical O_x production during transport. The higher average f_{44}/f_{43} ratio observed downwind (t-test, $p < 0.05$) was consistent with a shift toward a more oxidized Org signature. However, these compositional changes were not accompanied by a substantial increase in aerosol mass relative to CO during the early stages of plume aging across different aerosol loading regimes. The average $\Delta Org/\Delta CO$ showed limited urban-to-downwind changes across different aerosol loading regimes after ~ 2 – 6 h of transport, suggesting limited net enhancement of secondary aerosol mass relative to CO. The mass fraction of aerosol components also shows limited variations, and Org was the dominant NR- PM_{10} aerosol components both at the urban site and downwind. Within the airborne measurements alone (as shown in Figure S5 in the Supplementary Information and also in Velazquez-Garcia et al. (2026)), $\Delta Org/\Delta CO$ exhibited a tendency to increase with transport time, suggesting progressive Org mass concentration enhancement within the downwind plume. Nevertheless, the magnitude of this increase was relatively small compared with the overall urban-to-downwind variability and did not result in a clear enhancement relative to the PRG observations. No consistent enhancement in $\Delta SO_4/\Delta CO$ was observed from the PRG site to the downwind plume. The average $\Delta SO_4/\Delta CO$ was lower downwind under high aerosol loading, whereas only a limited difference was observed under low aerosol loading, indicating no clear net accumulation of sulfate relative to CO during the investigated transport period. This may be due to limited precursor availability within the urban plume, as SO_x emissions have declined across European urban areas over the past decades (Ciarelli et al., 2019; Tsimpidi et al., 2025). $\Delta pNO_3/\Delta CO$ showed limited difference between the PRG and downwind measurements under high aerosol loading, whereas higher values were observed downwind under low aerosol loading. This may be partly attributed to enhanced evaporation of particulate nitrate to the gas phase during plume aging under high aerosol loading conditions, when temperatures were higher than during the low aerosol loading periods. The average temperature measured onboard the aircraft was higher during the high aerosol loading periods (18.2 ± 1.6 °C) compared to the low aerosol loading periods (15.7 ± 1.7 °C) (t-test, $p <$

0.05). Although the average RH measured onboard the aircraft was also higher during the high aerosol loading periods ($58 \pm 11\%$) than during the low aerosol loading periods ($52 \pm 5\%$) (t-test, $p < 0.05$), the higher temperatures likely played a key role in promoting the volatilization of ammonium nitrate, thereby reducing $p\text{NO}_3$ formation efficiency. This is similar to prior airborne studies, such as Decarlo et al. (2008), that observed decreases of $p\text{NO}_3$ as the urban plume over Mexico City was transported away from the city core.

We have also added the evolution of $\Delta\text{Org}/\Delta\text{CO}$ within the airborne measurements only in the Supplementary Information:

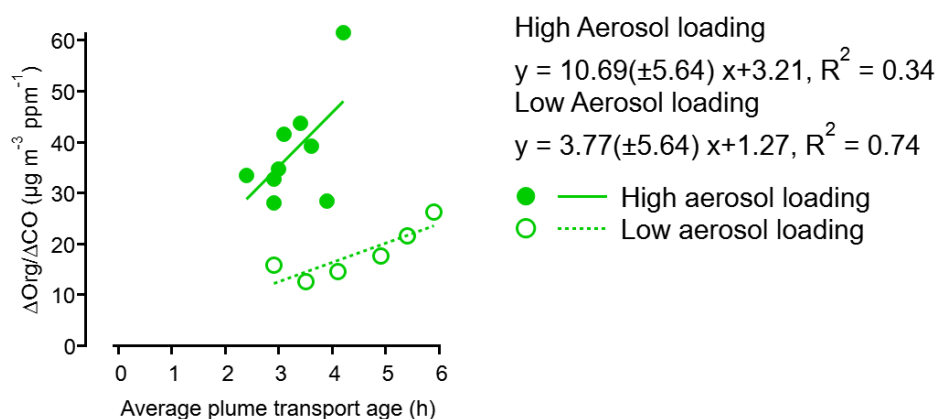


Figure S5. Evolution of $\Delta\text{Org}/\Delta\text{CO}$ as a function of plume transport age

- Quantitative comparison of organic aerosol contributions: At line 393, when comparing with other aircraft studies showing organic aerosol dominance in urban plumes, provide specific percentage contributions of organic aerosol in-plume versus out-plume (background) conditions to facilitate direct comparison and intuitive understanding.

Accept, we have included the organic aerosol mass fraction at Line 393, Section 3.1:

Compared with previous airborne measurements over other urban regions, observations across Europe (McMeeking et al., 2012; Freney et al., 2014; Morgan et al., 2010), Asia (Nault et al., 2018), and the Americas (Shingler et al., 2016; Vu et al., 2016; DeCarlo et al., 2008; Shilling et al., 2018) consistently indicate a substantial contribution of Org to total PM_{10} mass within downwind urban plumes. In our study, Org accounted for approximately 56.4–70.2% of PM_{10} mass under in-plume conditions, compared with 60.6–68.9% under out-plume conditions, indicating the dominant contribution of Org to PM_{10} in both the Paris urban plume and the surrounding regional background.

- SSA wavelength dependence interpretation requires revision: At line 418, there appears to be confusion between AAE and SAE in the interpretation. More importantly, the attribution of $\text{SSA}(450\text{nm}) < \text{SSA}(630\text{nm})$ solely to BrC influence is problematic. Figure 4 shows this pattern occurs in only 3 of 6 flights (A032, A036, A033), while the other three flights (A027a, A031, A034)

show the opposite pattern. The interpretation should consider that SSA wavelength dependence results from the combined effects of SAE (size-dependent scattering) and AAE (composition-dependent absorption). For example, when BC dominates ($AAE \approx 1$), SSA increases with wavelength if $SAE > 1$ (characteristic of small particles). The analysis should examine SAE and AAE variations together and provide flight-specific explanations considering particle size effects.

We appreciate the reviewer's comments. We have added a comparison of the average in-plume SAE and AAE for individual flights to the Supplementary Information to better illustrate the flight-specific variations in SSA. Since the wavelength dependence of SSA is governed by the relative spectral dependence of scattering and absorption, the comparison between SAE and AAE provides a physical basis for interpreting the observed differences in SSA between 450 and 630 nm. During the low aerosol loading periods (A031, A033, and A034), however, the optical signals approached the instrumental measurement limits, resulting in greater relative uncertainties in the derived SSA. Therefore, the apparent wavelength dependence of SSA under these conditions should be interpreted with caution.

We have added the comparison of AAE and SAE as Figure S4 in the Supplementary Information:

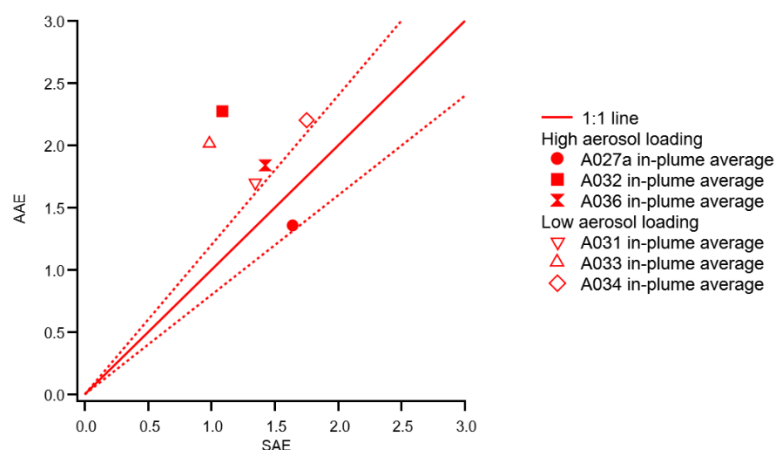


Figure S4: Comparison of the average in-plume absorption Ångström exponent (AAE) and scattering Ångström exponent (SAE) for individual flights. The solid line indicates the 1:1 relationship between AAE and SAE, while the dashed lines represent the estimated $\pm 20\%$ measurement uncertainty in AAE.

We have revised the corresponding discussion in Section 3.1 (Lines 417–423) accordingly:

Comparison of the average in-plume AAE and SAE for individual flights is presented in Figure S4 in the supplementary information. For the high aerosol loading periods flights, Flights A032 exhibited lower average SSA at 450 nm than at 630 nm and exhibited higher AAE than SAE, with the differences exceeding the estimated 20% uncertainty in AAE. This is consistent with the higher SSA observed at 630 nm, reflecting a stronger wavelength dependence of absorption than scattering. For Flights A027a and A036, the differences between SAE and AAE were closer to the $\sim 20\%$ measurement uncertainty of AAE, indicating comparable wavelength dependences of scattering and

absorption and corresponding to relatively small differences in SSA between 450 and 630 nm. During the lower aerosol loading periods, the wavelength dependence of SSA varied among individual flights. Flights A031 and A034 exhibited lower SSA at 630 nm than at 450 nm, whereas Flight A033 showed the opposite pattern. These spectral patterns were qualitatively consistent with the relative magnitudes of AAE and SAE; however, under low aerosol loading conditions, the optical signals approached the instrumental detection limits, resulting in greater relative uncertainties in the derived optical parameters.

4. Particle size distribution interpretation needs stronger evidence: Section 3.2 attributes increased 80–200nm particles in-plume to secondary aerosol formation, but this attribution lacks strong support. The cited literature compares urban versus downwind suburban sites, which is not directly analogous to the in-plume versus out-plume comparison in this study. Since urban emissions naturally contain high concentrations of 80–200nm particles, the observed increase could reflect primary urban emissions. A more appropriate comparison would be with PRG urban site measurements, though instrument detection limits may constrain this approach.

We appreciate the reviewer's suggestion. Although a direct comparison with the PRG measurements is not feasible because the sizing instruments deployed at the aircraft and the PRG site have different measurement size ranges, particle number concentrations in the overlapping 80–200 nm size range were of a similar order of magnitude between the PRG site and the downwind urban plume despite plume dilution. We have therefore revised the discussion in Lines 443–450 to address the reviewer's concern and provide a more balanced interpretation:

Downwind aircraft measurements showed an enhancement in particle number concentrations in the 80–200 nm size range during in-plume periods compared to out-plume periods. Particle number concentrations in the overlapping 80–200 nm size range remained of a similar magnitude between the PRG site and the downwind urban plume despite plume dilution, indicating that the 80–200 nm size range remained a substantial fraction of the particle population during plume transport. In the Greater London region, enhanced contributions from particles in the 100–200 nm size range have been observed at downwind suburban sites (Beddows et al., 2015; Harrison et al., 2019). Likewise, Cirino et al. (2018) reported enhanced particle number concentrations in the 100–200 nm size range in aged urban plumes over downwind suburban areas in South America.

5. Clarify particle versus gas-phase pollution discrepancies: At line 455, the statement that particle mass increases are less pronounced than gas-phase pollutant increases is puzzling given that Figure 2a shows substantial rBC and NO₂ enhancements (approximately doubling in-plume). This discrepancy warrants discussion of potential mechanisms such as deposition processes, dilution effects, or measurement artifacts, and an explanation of why rBC shows strong enhancement while other particle components do not.

We appreciate the reviewer's suggestion. While dilution and deposition may have contributed to the observed differences among pollutants, we consider that the dominant factor is the substantial contribution of the regional background to the total PM₁ aerosol mass. Consequently, the addition of urban emissions resulted in only a modest increase in PM₁ mass concentration within the plume. In contrast, the regional background contained relatively low levels of the pollutants rBC and NO₂, making their plume-related enhancements much more pronounced. This is the primary reason why rBC and NO₂ were selected as plume tracers in this study. Furthermore, although rBC mass concentrations increased significantly within the plume, rBC constituted only a small fraction of the total PM₁ mass concentration (less than 5%). As a result, the pronounced enhancement in rBC had only a limited influence on the bulk PM₁ mass concentration, which remained largely dominated by the regional background aerosol. We have revised the relevant descriptions and included additional discussion to explain the relatively modest enhancement in PM₁ mass concentration relative to the gas- and particle-phase urban plume tracers.

This modest increase in particle volume was consistent with the PM₁ mass concentrations shown in Figure 3, which also exhibited only a limited enhancement between in-plume and out-of-plume periods despite the statistically significant increases in the urban plume tracers, including NO₂, CO mixing ratios and rBC mass concentration. Although rBC exhibited the largest relative in-plume enhancement among the measured PM₁ components, it accounted for only a small fraction of the total PM₁ mass and was characterized by a relatively low background concentration. In contrast, PM₁ mass was dominated by Org and secondary inorganic species, which exhibited much smaller relative enhancements because of their larger regional background contributions. Consequently, the overall increase in PM₁ mass was considerably smaller than that observed for the gas-phase tracers and rBC.

6. Ensure consistent size ranges for volume-effective radius: At line 457, clarify whether the volume-effective radius calculation uses consistent size ranges for aircraft (80-1000 nm) and ground (30-1000 nm) measurements, and discuss the potential impact of any differences on comparability.

Accept. We have also corrected a typographical error here: the reported values correspond to the volume-effective diameter, not the volume-effective radius. We have revised volume-effective diameter results presented here by deriving the volume-effective diameter results using the overlapping particle size range (80 – 1000 nm) measured by both the aircraft and the PRG site. Although the absolute values of volume-effective diameter have been changed, the relative differences were not statistically significant (confirmed by t-test, $p > 0.05$). We have revised the corresponding results in the manuscript:

Using the overlapping particle size range common to both the PRG site and the downwind airborne measurements (80–1000 nm), the average volume-effective radius at PRG was 314 (±31) nm during

high aerosol loading periods and 323 (± 37) nm during low aerosol loading periods, compared with 318 (± 63) nm and 308 (± 103) nm, respectively, in the downwind plume. These differences were not statistically significant (t-test, $p > 0.05$), indicating that no statistically significant change in the volume-effective size was detected within the 80–1000 nm size range between the PRG site and the downwind plume.

7. Refine Ox/ Δ CO ratio analysis: At lines 474-477, specify whether Flight A027b (noon flight with potentially stronger photochemistry) was excluded from the Ox/ Δ CO analysis, and compare in-plume versus out-plume Ox/ Δ CO values to better characterize photochemical processing differences.

We appreciate the reviewer's suggestion. In the revised manuscript, we replaced the Ox/ Δ CO analysis with a comparison of absolute Ox concentrations between the urban and downwind measurements. This is because the downwind out-plume Ox concentration associated with the first plume intercept, when used as the background, can be higher than the Ox concentrations measured at PRG and within the plume, resulting in negative Δ Ox values. In the analysis, Flight A027b, conducted in the afternoon under potentially stronger photochemical conditions, was included. We repeated the analysis after excluding A027b and found only minor changes ($\sim 5\%$) in the overall Ox statistics, indicating that this flight does not materially affect the observed urban-to-downwind differences or their interpretation. The manuscript and corresponding figure have been revised accordingly. The revised Figure 6 and the revised context is presented in our response to Comment #1.

8. Validate complex refractive index values: The CRI real parts as low as 1.3 in Figure 8 are unexpectedly low given the chemical composition ($\sim 60\%$ organic and 40% inorganic components with CRI=1.53). This suggests either methodological issues or exceptionally low organic aerosol CRI values that require explanation. Correlation with O:C ratios from AMS data would help validate these results and provide insights into organic aerosol aging.

We appreciate the reviewer's comment. The unusually low retrieved CRI real parts were observed during periods of low aerosol loading, when the weak scattering signal results in greater variability and uncertainty in the CRI retrieval. This increased variability is reflected by the 25th–75th percentile ranges shown in Figure 8. The ACSM measurements also showed lower f₄₄/f₄₃ ratios during these periods, indicating relatively less oxidized organic aerosol, which is qualitatively consistent with the lower retrieved real CRI. Nevertheless, given the increased retrieval uncertainty at low aerosol loadings, we avoid interpreting these unusually low absolute values as intrinsic aerosol refractive indices and focus instead on the relative variations in CRI.

At the PRG site, no statistically significant difference was observed in the average SSA at 630 nm

between aerosol loading regimes (t-test, $p > 0.05$, Figure 8(b)), whereas the average SSA at 450 nm was slightly lower during low aerosol loading periods (t-test, $p < 0.05$, Figure 8(a)). Consistent with this behaviour, the imaginary part of the CRI increased at both wavelengths during low aerosol loading periods compared to high aerosol loading periods (Figure 8(e)(f)). However, the lower aerosol loadings during these periods resulted in weaker optical signals and consequently greater uncertainty and variability in the CRI retrieval, which may partly contribute to the relatively high imaginary and low real CRI values. This increased variability is reflected in the broader 25th–75th percentile ranges of the retrieved CRI during low aerosol loading periods, as shown in Figure 8. The PRG measurements also showed lower f_{44}/f_{43} ratios during the low aerosol loading than during the high aerosol loading periods, indicating relatively less oxidized organic aerosol. This is qualitatively consistent with the lower real part of the CRI. The tendency toward a higher imaginary CRI may also reflect a greater relative influence of local emissions at PRG. The proximity of the PRG site to emission sources and the inclusion of morning rush-hour emissions (04:00–08:00 UTC) in the selected plume emission period are consistent with the diurnal analysis presented in Di Antonio et al. (2025). During low aerosol loading periods, the reduced influence of regionally transported NR-PM₁ may increase the relative contribution of local traffic emissions, potentially contributing to the higher imaginary part of the CRI despite the overall cleaner conditions

Reviewer #2:

This study reports on airborne measurements of aerosol properties and gases at various downwind distances from Paris and compares them to a ground site within the city. The topic is well suited for this journal and is of scientific interest. Overall, the findings are that there are relatively minor changes in species concentrations with plume age, and also in comparison to the urban site, which makes the analysis and interpretation of the data more difficult (ie, testing if there are real differences). A major issue is that the interpretation tends to assume a Lagrangian-type of experiment, which is not proven – specifically the comparisons with the ground based urban site.

Although the data are self-consistent, (OA/dCO increases, BrC increases), all these changes are small. The authors use statistical test to decide if there is a difference or not, which is commendable, but it would also be good to include an uncertainty analysis to help with the interpretation. For example, error bars (sensitivity analysis) in Figure 10 added to the CRI (real and imaginary) for this data would be useful. It would also support (or not) some of the conclusions, eg, that including plume changes are needed to correctly assess optical properties of aerosols. Also an uncertainty analysis when comparing parameters from various measurement locations would be helpful, especially for the inferred optical properties, these could be fairly high. Another weakness was specifying the particle size ranges covered for each aerosol measurement, which seemed to be handled rather haphazardly throughout. This includes inlets and instrument measurement size ranges. Maybe this could be added to Table S1.

Overall, these are difficult measurements, and the authors have done a reasonable job presenting them, although the findings are not striking, they are important.

We thank the reviewer for the comments.

First, we acknowledge that the coordinated ground-based and airborne measurements in our study do not constitute a strict Lagrangian experiment. Rather, the measurement strategy provides a quasi-Lagrangian framework to characterise urban-to-downwind variations in aerosol physicochemical and optical properties within air masses influenced by the Paris urban plume. We have therefore revised the manuscript to avoid overinterpreting the ground-to-aircraft differences as direct temporal evolution of the same air mass. In the revised manuscript, we describe our results in terms of urban-to-downwind variations in aerosol physicochemical and optical properties, from urban conditions with a greater contribution from fresh emissions to downwind conditions with an increased contribution from secondary aerosol, with atmospheric processing considered as a potential contributor to the observed differences. The title of the manuscript has been revised to “Urban-to-downwind variations in aerosol composition and optical properties in the Paris urban plume from coordinated airborne and ground-based observations”, and the manuscript has been revised accordingly to reflect this change in emphasis. Following the reviewer’s suggestion, we have also revised the aerosol component analysis by subtracting the out-of-plume background concentrations from both the aerosol species and CO and reporting the results as $\Delta\text{species}/\Delta\text{CO}$, thereby reducing the influence of background concentrations on the comparison.

Secondly, the A2S2 was evaluated through laboratory intercomparison experiments before the experiment, as reported in Yu et al. (2024). The uncertainties in the A2S2 extinction and scattering coefficients were estimated to be within 1–2%, while the uncertainties in the absorption coefficients derived using the EMS method were approximately 2% at 450 nm and 7% at 630 nm based on comparison with the AE33. The relatively small experimentally determined uncertainties in the measured extinction, scattering, and absorption coefficients are not expected to materially affect the main results. We have assessed the uncertainties associated with the derived AAE, for which uncertainty propagation is more relevant. We have added Figure S8 to the Supplementary Information to illustrate the estimated measurement uncertainty ranges. We have also revised the discussions regarding the variation of AAE and SAE. Please refer to our response to Comment #12 for the revised context.

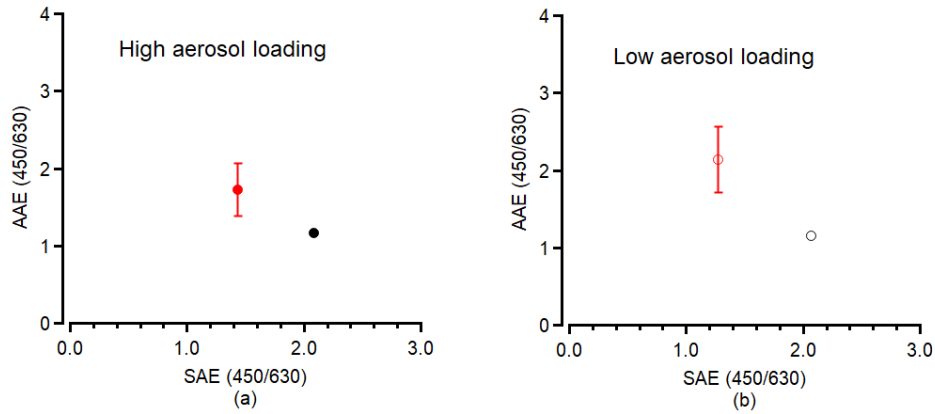
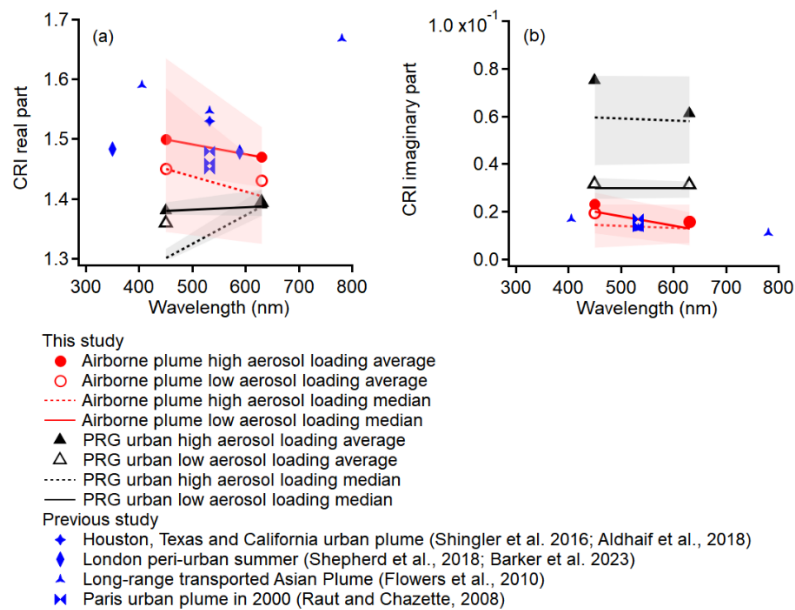


Figure S8. Comparison of average AAE and SAE between the PRG site and the airborne measurements. Error bars represent the estimated A2S2 measurement uncertainties of approximately 20% for AAE.

We have also added the 25th–75th percentile ranges to Figure 10 to illustrate the variability in the retrieved CRI. We acknowledge that the retrieved CRI exhibits relatively large variability during low aerosol loading periods. Nevertheless, the mean values show a consistent tendency toward a larger real part and a smaller imaginary part under downwind conditions. We have revised the discussion accordingly to avoid overinterpreting these differences.



Thirdly, we have followed the reviewer’s comments to add the instrument measurement size range into Table S1.

S1 Summary of instruments and their availability during the campaign

Parameter	Instrument	Frequency	Detection limit	Detection size range
SAFIRE ATR-42 research aircraft				

Non-refractory Aerosol Particle Composition	mAMS (Aerodyne)	Measured at 1 s (FastMS mode), final data averaged over 15 s	Org: 0.08 $\mu\text{g m}^{-3}$ SO ₄ : 0.007 $\mu\text{g m}^{-3}$ NO ₃ : 0.008 $\mu\text{g m}^{-3}$ Chl: 0.008 $\mu\text{g m}^{-3}$ NH ₄ : 0.034 $\mu\text{g m}^{-3}$	~ 70 nm to 700 nm vacuum aerodynamic diameter
Refractory Black Carbon (rBC)	SP2 (DMT)	1 s	< 10 ng m ⁻³	Total particle diameter: up to 1 μm rBC core mass-equivalent diameter: 90 – 580 nm
Aerosol extinction and scattering coefficient at 450 nm and 630 nm	A2S2 (dual CAPS-PM _{SSA} , Aerodyne)	Measured at 1 s, final data were smoothed using a 10-s moving-average window to reduce high-frequency fluctuations	10s integration: 0.69 Mm ⁻¹ for $\sigma_{\text{ext},450 \text{ nm}}$ 0.45 Mm ⁻¹ for $\sigma_{\text{scat},450 \text{ nm}}$ 0.21 Mm ⁻¹ for $\sigma_{\text{ext},630 \text{ nm}}$ 0.12 Mm ⁻¹ for $\sigma_{\text{scat},630 \text{ nm}}$	Particle transmission efficiency exceeds 90% for particles with diameters up to 4 μm . Inlet size cut up to 10 μm
Particle Size Distribution (~100 - 1000 nm)	UHSAS (DMT)	1 s	-	-
Particle Size Distribution (~1 – 10 μm)	Sky-OPC (GRIMM)	6 s	-	-
NO ₂	TD-LIF for most of the flights and CAPS-NO ₂ (T500U, Teledyne) for Flight A036	1 s for TD-LIF 1 min for CAPS-NO ₂	9.8 ppt for TD-LIF < 40 ppt for CAPS-NO ₂	-
CO	Gas analyser (G2401-m, PICARO)	1 s	15 ppb	-
O ₃	Ozone analyser (Model 49i, Thermo Scientific)	5 s	0.5 ppb	-
PRG ground site				
Non-refractory Aerosol Particle Composition	Time-of-Flight Aerosol Chemical Speciation Monitor (ToF-ACSM, Aerodyne)	~6 min	Org: 0.31 $\mu\text{g m}^{-3}$ SO ₄ : 0.04 $\mu\text{g m}^{-3}$ NO ₃ : 0.09 $\mu\text{g m}^{-3}$ Chl: 0.06 $\mu\text{g m}^{-3}$ NH ₄ : 0.19 $\mu\text{g m}^{-3}$	~ 70 nm to 700 nm vacuum aerodynamic diameter
Aerosol absorption coefficient at	Aethalometer (AE33, Magee Sci.)	1 min	-	The instrument has no intrinsic particle size cut. Inlet size cut up to 1 μm

370, 470, 520, 590, 660, and 880 and 950 nm				
Aerosol scattering coefficient at 450, 550, and 700 nm	Nephelometer (Ecotech Aurora 4000)	1 min	< 0.3 Mm ⁻¹ for all wavelengths	The instrument has no intrinsic particle size cut. Inlet size cut up to 1 µm
Particle Size Distribution (~0.02 – ~1 µm)	SMPS (DMA, TSI model 3081 and CPC, TSI model 3775)	~2.5 min	-	-
NO ₂	AC32M monitor (Environment SA)	1 s	0.4 ppb	-
O ₃	41M monitor (Environment SA)	1 s	1 ppb	-
CO	AIRPARIF network in central Paris	1 h	-	-

Table S1: List of instruments at PRG site and aboard the SAFIRE ATR-42 aircraft

Specific comments.

1. I would just note that calling the plumes heavy and light pollution could lead to some confusion. Ie, the term light pollution has a specific meaning that differs from how it is used here.

We appreciate the reviewer's comment. The terms "heavy" and "light" were intended to distinguish relatively higher and lower PM₁ levels within the airborne measurements, rather than to represent absolute pollution categories. To avoid potential confusion, we have revised these terms with "high aerosol loading" and "low aerosol loading" throughout the manuscript.

2. Discuss limitations of one ground site being representative of the Paris metropolitan region emissions to which the downwind data are compared. Ie, how do aircraft data closest to the city compare?

We acknowledge that a single ground site cannot fully capture the spatial heterogeneity of emissions across the Paris urban area. In this study, PRG was used as an urban background reference site to characterize conditions within Paris, rather than to represent the full spatial variability of emissions across the Paris metropolitan area. We have revised the corresponding description in Line 143 – 149 in Section 2.1 accordingly:

The ground-based site used in this study is the Paris Rive Gauche (PRG) site (48.83° N, 2.38° E), located at Université Paris Cité in the central part of the Greater Paris Region (Di Antonio et al.,

2025). The PRG measurement site was located a few hundred meters from major traffic emission sources, including train stations and urban highways. The measurements were conducted on the roof of the university building at approximately 30 m a.g.l. In this study, observations at PRG were used as an urban background reference for comparison with the Paris urban plume sampled downwind by the aircraft, rather than to represent the spatial variability of emissions across the entire metropolitan area. The periods selected from the PRG measurements for comparison with the airborne observations were determined using back-trajectory analysis, as described in Section 2.5.

In this study, only the first two plume intercepts downwind, which were generally the closest aircraft measurements to the city, were considered. Therefore, the mass fractions presented in Figure 7 are interpreted as a comparison between the urban background conditions observed at PRG and the near-downwind Paris urban plume sampled by the aircraft. The NR-PM₁ mass fractions measured by the aircraft, were generally comparable to those observed at PRG, indicating a similar bulk relative aerosol composition between the PRG urban background and the near-downwind Paris plume. The revised discussions of Figure 7 are presented in our response to Comments #9.

3. State why is AMS NO₃- referred to as pNO₃, while none of the other species use a similar type of acronym? (Ie, since not classifying as ions, NO₃ could be interpreted as the nitrate radical)?

We appreciate reviewer's concern. We referred AMS/ACSM measured NO₃ as particulate NO₃ (pNO₃) to distinguish it from the total NO₃ (gas and particulate phases) measured by the TD-LIF instrument deployed aboard the aircraft. Since this study focuses exclusively on the particulate phase, we use pNO₃ for the AMS/ACSM measurement results to avoid ambiguity. We have stated this in Line 223 – 225, Section 2.2.2:

The instrument detects NO₂ via laser-induced fluorescence following excitation at 532 nm, while additional total gas-phase nitrogen species (NO_y) species are quantified through controlled thermal dissociation in parallel heated channels. To distinguish these gas-phase reactive nitrogen measurements from particulate nitrate measured by the mAMS/ACSM, the latter is denoted as pNO₃ throughout this study.

4. Line 171, what is the implication of the 0.75 slope of volume from AMS vs number size distribution and density. The slope and correlation is given but there is no interpretation. Does this main AMS mass may be systematically off by 25%?

We appreciate the reviewer's comment. The slope of 0.75 does not necessarily indicate that the AMS mass concentrations were systematically underestimated by 25%, because the two volume estimations do not represent exactly the same aerosol fraction. The AMS-derived volume was calculated from the measured NR-PM₁ components using composition-dependent densities, whereas the UHSAS-derived volume represents the optically detected particle population within its measurement size range and may also include refractory components that are not detected by the AMS. Therefore, the slope of 0.75 is interpreted as reasonable agreement between two independent

measurements, rather than direct evidence of a 25% systematic bias in the AMS mass measurements. Additionally, we have revised the manuscript to clarify this interpretation:

The comparison is presented in Figure S1 in the Supplementary Information, showing a slope of 0.75 and a coefficient of determination (R^2) of 0.82. The strong correlation indicates consistent variability between the two independent measurements. A slope below unity is expected because the AMS-derived volume represents only NR-PM₁, whereas the volume derived from the UHSAS number size distribution may also include refractory aerosol components that are not detected by the AMS. Additional uncertainties associated with the assumed aerosol densities may also contribute to the differences.

5. Give particle size ranges over which the SP2 and A2S2 can measure.

Following Reviewer #2's general comment, we have added the measurement size ranges of the instruments to Table S1. The A2S2 (or CAPS-PM_{SSA}) does not have an intrinsic particle size cut. Characterization of the modified A2S2 system showed that particle losses were less than 10% for particles up to 4 μm in diameter.

6. In equations 4 and 5 explicitly state where the absorption coefficients at 630 and 450 comes from – what instrument. Does it come from the measured extinction minus the scattering coefficients. I suspect there is significant uncertainty associated with the calculated absorption coefficients.

We appreciate the reviewer's concern. To meet the requirements of fast-response of airborne characterisation study, we developed the A2S2 instrument based on the original design of CAPS-PM_{SSA}. Prior to the campaign, the A2S2 was validated through laboratory intercomparison experiments against an AE33 for aerosol absorption measurements (Yu et al. 2024). The measurement uncertainties in the A2S2 absorption coefficients were estimated to be 6% at 450 nm and 2% at 630 nm, respectively, while the uncertainty in the AAE, estimated from comparison with the AE33, was approximately 20%. However, under low aerosol loading conditions, where the absorption coefficients approach the instrumental detection limit, the relative uncertainty in the derived absorption coefficients may increase.

We have added the discussions of absorption measurement uncertainties in Line 194, Section 2.2.1: The absorption coefficient (σ_{abs}) is derived from the Extinction-Minus-Scattering (EMS) method (Weber et al., 2022). Based on laboratory intercomparison experiments (Yu et al. 2024), the measurement uncertainties in the A2S2 absorption coefficients were estimated to be 2% at 450 nm and 7% at 630 nm, respectively. The uncertainty in the AAE derived from the A2S2 absorption coefficients at 450 and 630 nm was estimated to be approximately 20% based on comparison with the AE33.

We have added the description that the absorption coefficient at 450 nm and 630 nm was from A2S2 measurements in Line 312- Line 316, Section 2.4. We have also added a discussion of the

uncertainties associated with the estimated BrC absorption contribution following the reviewer's comment #7.

The BrC absorption coefficient is determined in the assumption that absorption at 630 nm is contributed by BC only and the AAE for BC is 1. Henceforth, the σ_{abs} contributed by BC at 450 nm can be derived as:

$$\sigma_{\text{abs BC},450} = \sigma_{\text{abs},630} \cdot \left(\frac{450}{630}\right)^{-AAE_{450/630}} \quad (4)$$

Then the σ_{abs} of BrC at 450 nm can be estimated as:

$$\sigma_{\text{abs BrC},450} = \sigma_{\text{abs},450} - \sigma_{\text{abs BC},450} \quad (5)$$

Where the $\sigma_{\text{abs},450}$ and $\sigma_{\text{abs},630}$ were derived using the EMS method with the A2S2 for the airborne measurements and were obtained from the AE33 for the PRG measurements.

$$\sigma_{\text{abs BrC},450} = \sigma_{\text{abs},450} - \sigma_{\text{abs BC},450} \quad (6)$$

To assess the uncertainty associated with the assumed BC AAE, the BrC absorption fraction was also calculated using BC AAE values ranging from 0.8 to 1.2 to provide a sensitivity range. Previous studies also report strongly absorbing BrC species (e.g., tar balls or dark BrC) may also importantly contribute to the absorption at long wavelengths. However, these long-wavelength-absorbing BrC species have been reported predominantly in biomass-burning aerosols (Xu et al., 2026), and no evidence of biomass-burning influence was observed during the selected measurement period of the ACROSS campaign in this study. Therefore, although the assumption of negligible BrC absorption at 630 nm may slightly underestimate the BrC contribution to aerosol absorption at 450 nm, it is unlikely to substantially affect the observed relative variations or the main conclusions of this study.

7. Give a possible range of calculated BrC absorption coefficient since it is determined by difference between measurements with associated uncertainties and depends on aerosol properties that can be variable (BC AAE=1). Also, the assumption is made that there is no BrC light absorption at 630 nm, which may not be true (see publications on tar BrC or dark BrC (d-BrC)). What are the implications of this assumption?

We acknowledge that strongly absorbing BrC species (e.g., tar balls or dark BrC) may exhibit measurable absorption at 630 nm. However, these particles have been reported predominantly in biomass-burning aerosols and are therefore expected to be much less abundant in the summertime anthropogenic urban plume investigated in this study, where no evidence of biomass-burning influence was identified during the flight periods. Therefore, we consider the assumption of negligible BrC absorption at 630 nm to be reasonable for this study. We have included the related discussions in Line 312- Line 316, Section 2.4. Please refer to our response to Comment #6 for the corresponding revisions.

We also revised the BrC absorption fraction results to include a sensitivity range associated with

varying the assumed BC AAE from 0.8 to 1.2 in Line 581 - 583, Section 3.3.2:

The average BrC absorption fraction was 15.7 (± 17.7) % under heavy pollution conditions and 8.9 (± 6.9) % under light pollution conditions at the PRG site, with sensitivity ranges of 11.9–18.1% and 6.9–12.5%, respectively, when the assumed BC AAE was varied from 0.8 to 1.2. to 39.1 (± 20.8) % and 42.0 (± 31.1) %, respectively, within the downwind plume. Within the downwind plume, the average BrC absorption fractions increased to 39.1 (± 20.8) % and 42.0 (± 31.1) %, with corresponding sensitivity ranges of 36.8–41.1% and 45.0–47.2%, respectively, when the assumed BC AAE was varied from 0.8 to 1.2.

8. Discussion on volume increases between lines 450 and 455. Specify what the diameter range is associated with the discussed volume increase.

We have addressed this comment by revising the discussion in Lines 450–455, Section 3.2.1 following reviewer's comments:

The enhancement in particle number concentration under in-plume conditions relative to out-of-plume conditions was largely confined to the 80–200 nm size range and remained relatively modest. As a result, the corresponding increase in particle volume concentration integrated over the UHSAS measured size range (80–1000 nm) was also modest, although statistically significant (confirmed by t-test, $p < 0.05$). The average volume concentration integrated over the UHSAS measured size range (80–1000 nm) was $5.79 (\pm 3.34) \mu\text{m}^3 \text{cm}^{-3}$ for in-plume period and $4.54 (\pm 2.43) \mu\text{m}^3 \text{cm}^{-3}$ for out-plume period during the heavy pollution period and was $2.47 (\pm 0.86) \mu\text{m}^3 \text{cm}^{-3}$ for in-plume period and $1.93 (\pm 0.72) \mu\text{m}^3 \text{cm}^{-3}$ for out-plume period during the light pollution period.

9. Line 478 and on. Is it possible to make such a definite statement that because Org/dCO in the plume was higher than at the Paris ground site that it must be due to SOA formation (30%). Same applies for other species discussed in a similar way. This is not a true Lagrangian experiment.

Combining all the reviewers' comments, we have revised the discussions accordingly to avoid overinterpreting the observed urban-to-downwind differences of aerosol component concentrations. Specifically, we have avoided the interpretation of the variations in the organic aerosol mass fraction as evidence of SOA formation here. The revised context is presented in our response to Reviewer #1's Comment #1.

10. What is the effect of not considering background concentrations when normalizing by CO (in minus out of plume CO, ie y/dCO , where y is the variable of interest)? Normally one also subtracts the out of plume concentrations for the species in both the numerator and denominator (dy/dCO). In this case it is only done for the denominator. A good way to do this is from the slope of the species of interest (y) vs CO. It removes the backgrounds (intercept) and shows how consistent the excess mixing ratio is. Eg, if the correlation is poor, it provides insight on if this number makes much sense if calculated only based on averages.

We appreciate the reviewer's suggestion. However, Org showed only limited variability within these relatively short plume interpreting periods. Therefore, the regression-derived slopes were not considered sufficiently robust for quantitative comparison. We have revised the analysis using the $\Delta\text{species}/\Delta\text{CO}$ approach, with background concentrations determined from the background periods adjacent to each plume period. PRG background concentrations were determined using the airborne background values measured closest to the urban area. A comparison of the f44/f43 ratios between PRG and the downwind airborne measurements has also been added to provide qualitative information on the relative importance of SOA. We have carefully interpreted these results to avoid directly attributing the observed changes to an absolute increase in SOA downwind; instead, they are used to compare the relative contribution of SOA. Combine Reviewer #2's and Reviewer #3's comments we have revised Section 3.2.2. The revised context is presented in our response to Reviewer #1's Comment #1.

11. How was the AAE determined? Seems highly uncertain if just based on absorption measured at two different wavelengths. Could this account for substantial scatter in Fig 9? I suggest that a discussion is needed on how AAE varies with choice of wavelengths.

We appreciate the reviewer for this comment. As described in the Line 188 – 194, Section 2.2.1 in our original manuscript, the A2S2 measured aerosol extinction and scattering coefficients at two wavelengths (450 and 630 nm), and the absorption is therefore derived from the Extinction-Minus-Scattering (EMS) method. The uncertainty in the AAE derived from the A2S2 absorption coefficients at 450 and 630 nm was estimated to be approximately 20% based on comparison with the AE33. We have added discussions of the AAE uncertainties in the method section, and the results have been smoothed over 10 seconds to reduce the variability due to low aerosol loading. Please refer to our response to Comment #6 for the revised context regarding the measurement uncertainty of A2S2. During periods of low aerosol loading, the absolute aerosol light absorption coefficient value was close to the detection limit of the instrument (as presented in Figure 4), resulting in relatively larger variability in the measured absorption coefficients. We have added the discussions regarding the variability of AAE due to low aerosol loading in Line 568 – 575, Section 3.3.2. The revised contexts is presented in our response to Comments #12.

12. Line 574 and on. Is there a plot that can show that AAE increased with plume age?

We have added the AAE results as a function of plume age into the supplementary,

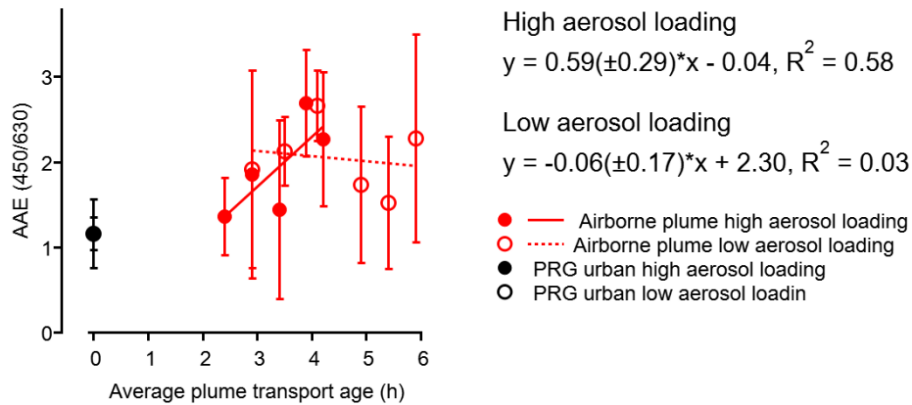


Figure S6. Evolution of AAE as a function of plume transport age

We have revised the discussions in Line 568 – 575 in Section 3.3.2:

Figure 9 compares the AAE and SAE derived from airborne and PRG ground-based measurements. Owing to the lower aerosol concentrations in the airborne observations, particularly during the low aerosol loading periods, the airborne AAE and SAE results exhibited larger variability compared to the PRG urban measurements. Therefore, the impact of differing measurement size cutoffs was likely minor. The spectral dependence of both absorption and scattering varied between the PRG site and the downwind airborne measurements. As shown in Figure 9, the average AAE at the PRG site was close to 1, while the average SAE was approximately 2 under both aerosol-loading regimes. The mean values were of comparable magnitude between the two aerosol loading regimes, while the t-test did not detect a statistically significant difference ($p > 0.05$).

For the downwind airborne measurements, the average AAE was approximately 2 and the average SAE was around 1.5 under both aerosol-loading regimes. The observed difference in average AAE between the PRG site and the downwind airborne measurements exceeded the A2S2 measurement uncertainty in AAE (~20%), while the t-test also indicated a statistically significant difference ($p < 0.05$). Within the downwind plume sampled by the aircraft, the spectral dependence of aerosol absorption and scattering evolved with plume aging. As shown in Figure S6 in the Supplementary Information and in Velazquez-Garcia et al. (2026), AAE exhibited an increasing trend with estimated plume transport age, particularly during the higher aerosol loading periods, indicating a progressively stronger wavelength dependence of aerosol absorption during downwind transport. During the lower aerosol loading periods, AAE showed substantially greater variability, while the average values remained higher than those observed at the PRG urban site.

13. Line 649 states the results show a clear evolution with downwind distance of the Paris plume. Seems an overstatement without a full uncertainty analysis and possible limitations with how excess mixing ratios were determined. This applies to the complete final paragraph.

We have revised the final concluding paragraph to adopt more cautious language and explicitly

acknowledge the uncertainties associated with the study. Following the reviewer's general comments, we have also added an uncertainty analysis to Figure 10 to better quantify the uncertainties in the derived CRI and assess the robustness of the observed trends, the revised Figure 10 is presented in our response to reviewer's general comments.

In summary, our observations reveal systematic urban-to-downwind differences in aerosol physicochemical and optical properties, consistent with a greater influence of secondary atmospheric processing in the downwind plume compared with the urban site. Such differences may not be fully represented in some atmospheric models and satellite retrieval algorithms that rely on fixed CRI values and simplified aerosol optical assumptions. While recent model and satellite retrieval studies (e.g. (Drugé et al., 2022; Jiang et al., 2024)) have introduced dynamic parameterizations to account for atmospheric processing of aerosol optical properties, substantial uncertainties remain in key optical parameters such as SSA (Li et al., 2022a). In-situ measurements are therefore essential to provide observational constraints, particularly in the context of recent changes in global anthropogenic emissions (Li et al., 2023). To the best of our knowledge, this study represents one of the few post-2017 airborne investigations of aerosol physicochemical properties over a European megacity. The relative variations in the derived CRI and observed SSA provide useful observational constraints for improving the representation of aerosol optical properties in chemical transport models and remote-sensing retrievals.

Reviewer #3:

This manuscript investigates the evolution of aerosol composition, size distribution, and optical properties in the Paris urban plume using coordinated aircraft and ground-based measurements from the ACROSS campaign. The study compiles a comprehensive suite of measurements, providing a potentially valuable dataset for examining urban plume evolution. However, the current analysis has substantial methodological and interpretive limitations, particularly regarding the assumed Lagrangian framework, the treatment of statistical significance, and the interpretation of OA evolution. I therefore recommend the manuscript for publication following major revision.

We thank the reviewer for these constructive comments. In response to the methodological concerns raised, we have added additional analyses and supporting results to strengthen our methodology and conclusions. Please refer to our detailed responses below and the corresponding revisions in the manuscript.

Major Comments

1. The primary analysis relies on a Lagrangian assumption that PRG represents the fresh urban plume and that the same air masses are subsequently sampled by the aircraft. However, the trajectory analysis only shows that the sampled air masses passed over or near Greater Paris. It does not establish that the aircraft sampled the same air parcels measured at PRG. The ground site may be

influenced by local emissions, while the aircraft may integrate emissions from the broader metropolitan area.

We appreciate the reviewer's concern. Following the comments from Reviewers #2 and #3, we have revised the Results and Discussion to avoid interpreting the observed differences as direct temporal evolution of the same air mass. Instead, we now describe our results in terms of urban-to-downwind variations in aerosol physicochemical and optical properties, from urban conditions with a greater contribution from fresh emissions to downwind conditions with an increased contribution from secondary aerosol, while discussing atmospheric processing as a potential contributor to the observed differences. We have revised the manuscript accordingly and have avoided overstating the evolution of the urban plume.

Following Reviewer #2's suggestion, we have also revised the aerosol component analysis by subtracting the out-of-plume background concentrations from both the aerosol species and CO and reporting the results as $\Delta\text{species}/\Delta\text{CO}$, thereby reducing the influence of background concentrations on the comparison.

2. CO at PRG is obtained from an AIRPARIF site approximately 6 km away, while the background CO concentration is assumed to be identical to that derived from the corresponding flight. This inconsistency makes the comparison based on ΔCO normalization less convincing. In addition, ΔCO normalization accounts for dilution only when CO and the normalized species share the same emission sources. This assumption requires further justification.

Regarding the reviewer's concern about the CO background, we have revised our approach by replacing the previously used flight-specific 5th percentile with the out-plume CO concentration measured closer to the city during each corresponding flight. These out-plume measurements were conducted within the boundary layer and were considered more representative of the regional background immediately surrounding the urban plume during each flight period. The same background was therefore used for both the ground-based and airborne analyses to ensure a consistent comparison. Using this common boundary-layer background also provides a consistent reference for comparing the ground-based and airborne observations. Given the uncertainties associated with the Lagrangian assumptions, we have revised our discussion to avoid interpreting the normalized aerosol composition as evidence of aerosol production or loss.

We have revised Line 236:

The regional background concentrations of CO, NO₂, and rBC were defined as the 5th percentile of the measurements from each flight and were used for plume identification. Background concentrations of aerosol and gas-phase species were determined from the nearest out-plume measurements, as shown in Figure 2(a).

We have revised Line 298, Section 2.3

The background CO concentration at PRG for each flight period was defined using the average out-plume CO concentration measured closer to the city during the corresponding flight, as these measurements were conducted within the boundary layer and were considered representative of the near-surface background surrounding the urban plume. As a sensitivity test, the analysis was repeated using the 5th percentile of the daily PRG measurements as the PRG background, with the corresponding results presented in the Supplementary Information (Figure S7). Although the absolute values changed with the choice of background, the overall patterns of variation between the PRG and downwind measurements remained.

We have also added the results using different background values (with the PRG background defined as the 5th percentile of the day) in the Supplementary Information to assess the sensitivity of our results to the choice of background.

S10 Analysis using PRG-derived background concentrations

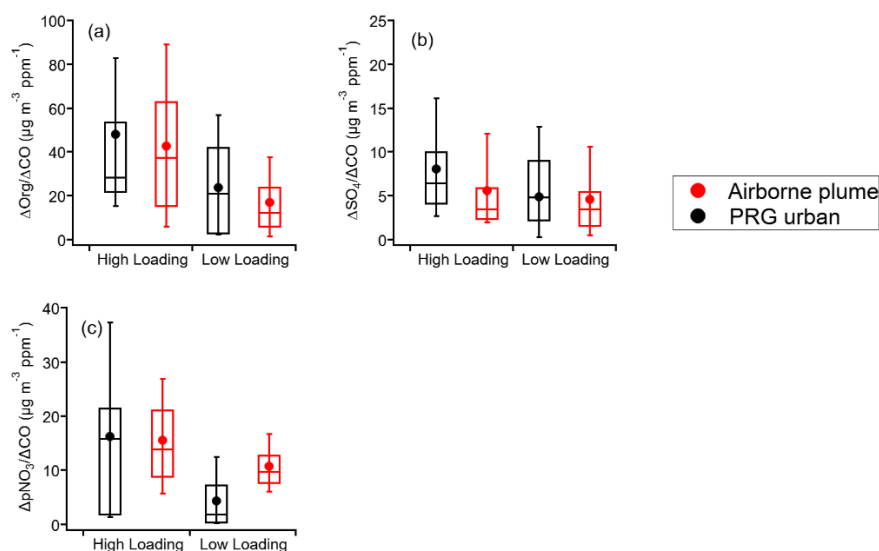


Figure S7. Same as Figure 6, but with the PRG background defined as the 5th percentile of the daily measurements.

Regarding the reviewer's concern for the use of ΔCO normalization. In this study, the analysis specifically focuses on the evolution of the same Paris urban plume, which is dominated by anthropogenic urban emissions. Under these conditions, we consider CO to be a reasonable conservative tracer of primary urban emissions. We have revised the manuscript to clarify that ΔCO normalization helps reduce the influence of dilution, thereby better isolating changes associated with secondary formation, rather than completely accounting for dilution effects. The revised context is presented in our response to Reviewer #1's Comment #1.

3. Statistical significance needs to be interpreted with caution. The manuscript applies t-tests to aircraft

and ground-based measurements that are temporally autocorrelated. Treating individual measurements as independent samples inflates the effective sample size and can produce artificially small p-values. A t-test is also not an appropriate primary criterion for identifying plume periods. In addition, statements that interpret $p > 0.05$ as evidence that two quantities are statistically similar are inappropriate. Failure to reject the null hypothesis does not demonstrate equivalence.

We appreciate the reviewer's concern regarding the interpretation of statistical significance. We acknowledge the limitation of t-test due to temporal autocorrelation. We have revised the manuscript to avoid relying on the t-test results as the primary evidence for the observed differences. Instead, the t-tests are retained as a supplementary statistical comparison, while our interpretation is based primarily on the observed distributions, mean and median values, and the consistency of the differences across individual flights and aerosol loading regimes. We have also revised the corresponding statements throughout the manuscript to avoid interpreting $p > 0.05$ as evidence of statistical similarity. Instead, these results are now described as showing no statistically significant difference between the groups compared.

Regarding reviewer's concern to the plume identification, we would like to clarify that the t-test was not used as the sole criterion for plume identification. The in-plume periods were identified primarily based on continuous enhancements in NO_2 and CO concentrations relative to the regional background, supported by back-trajectory analysis indicating influence from the Greater Paris region. The t-test was used as a supplementary statistical check to assess whether the identified plume periods exhibited higher tracer concentrations than the corresponding out-plume periods. To avoid any further misunderstanding, we have revised the description in the method section to clarify the plume identification procedure. Figure 2 has been revised as:

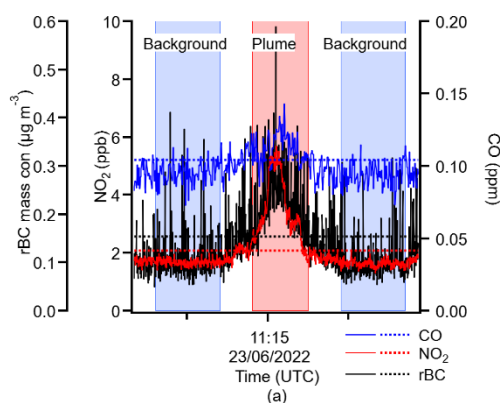


Figure 2. (a) An example of an airborne measured Paris urban plume identified by measured concentrations of NO_2 and CO from the Flight A028 on Jun 23rd, 2022. The dashed line indicates the threshold for each plume tracer, defined as the 5th percentile concentration plus twice the standard deviation of the measurements for each flight.

Line 335 – 340, Section 2.5 has been revised as:

Paris urban plume intercepts were identified in the airborne measurements based on concurrent and persistent enhancements in both NO₂ and CO mixing ratios above a threshold defined as the flight-specific regional background plus two standard deviations. When SP2 measurements were available, concurrent enhancements in rBC mass concentration within the identified plumes further supported the validity of the identification approach. The t-test was also applied as a supplementary statistical check to assess whether the identified plume periods exhibited higher tracer concentrations than the corresponding out-plume periods. The background period for each plume was defined as the 5-min interval before and after the plume period, with a 1-min transition interval applied to avoid contamination from the plume. Figure 2(a) presents an example of an identified plume and background period from the Flight A028 on 23rd June, 2022.

4. The interpretation of OA evolution relies mainly on bulk OA concentrations, OA fractions, and OA/ Δ CO ratios. These metrics alone are insufficient to distinguish POA from SOA or to attribute downwind changes to SOA formation. Source apportionment of the AMS and ACSM mass spectra, such as positive matrix factorization, would provide more direct evidence.

We appreciate the reviewer's suggestions for the source apportionment. However, as indicated in our response to Reviewer #1's Comment #1. PMF analysis is unable to provide robust or physically interpretable factors in our study due to airborne AMS measured organic aerosol composition exhibited limited spectral variability during relatively short measurement periods. Furthermore, the mass spectral results between the ground-based ACSM and airborne AMS may not be straightforward because the mAMS and ACSM differ in instrument design and measurement characteristics. We have added the f₄₄/f₄₃ ratio as an indicator of the oxidation characteristics of organic aerosol. We interpret changes in f₄₄/f₄₃ as a complementary indicator to characterize differences in Org composition and to provide qualitative information on the relative importance of SOA. The revised context is presented in our response to Reviewer #1's Comments #1.

Minor Comments

1. Line 165. Additional justification is needed for applying a constant CE = 0.5.

We have revised the related discussions:

A constant collection efficiency (CE) of 0.5 was applied, as ammonium nitrate accounted for less than 20% of the NR-PM₁ mass and fewer than 5% of the measurements were classified as acidic across all selected flights. These conditions indicate that composition-dependent effects on particle collection were limited, supporting the use of the standard CE value of 0.5 (Middlebrook et al., 2012).

2. Line 269. The use of a constant factor of two to convert eBC to rBC requires additional justification, as this relationship is not universally applicable.

We acknowledge that the empirical conversion from AE33-derived eBC to rBC is not universally applicable. Therefore, we have clarified in the manuscript that the converted AE33-derived rBC-equivalent mass concentration is used only for relative comparisons in this study, rather than for quantitative interpretation. Previous studies showed that eBC mass concentrations measured by the AE33 were approximately twice the rBC mass concentrations measured by the SP2 when the default wavelength-independent multiple-scattering correction factor was applied. In this study, a higher wavelength-independent correction factor of 2.14 was applied to the AE33 measurements, which reduces the derived eBC mass concentration. Therefore, applying the same eBC-to-rBC conversion factor of 2 is expected to provide a conservative estimate of the rBC-equivalent mass concentration. Nevertheless, the resulting rBC-equivalent mass fraction at PRG remained higher than that observed in the downwind urban plume. Therefore, the uncertainty associated with this empirical conversion is unlikely to substantially affect our conclusions. We have revised Line 270 – 272, Section 2.3:

A simplified factor of 2 was applied to convert the AE33-derived eBC mass concentrations to rBC-equivalent mass concentrations, based on previous AE33–SP2 intercomparisons (Laing et al., 2020; Tinorua et al., 2024b). These studies reported AE33-derived eBC concentrations approximately twice the SP2-measured rBC concentrations when using the manufacturer-recommended multiple-scattering correction for AE33. In this study, a higher wavelength-independent multiple-scattering correction factor (2.14) was applied, resulting in lower derived eBC mass concentrations and thus applying a conversion factor of 2 likely providing a conservative estimate of the rBC-equivalent mass concentration when applying the same conversion factor. The converted AE33-derived rBC-equivalent mass concentration results were only used to illustrate the relative changes in the rBC mass fraction (presented in Figure S2 in the Supplementary Information).

3. Line 290. Clarification is needed on whether the calculated inlet losses were applied to the measurements or only estimated.

We have revised the description to state that the calculated inlet losses were used only to evaluate the performance of the inlet system

Losses by diffusion were calculated individually for each instrument using the Particle Loss Calculator (PLC; developed by the Max Planck Institute; Von der Weiden et al., 2009) as a function of the sampling-line geometry (e.g., curvature and tube dimensions) and aerosol

properties (Baron and Willeke, 2001). The calculated losses were used to assess the performance of the modified inlet system and were not applied to correct the measured aerosol properties, as the estimated losses were generally below 10% over the particle size range considered.

4. Line 374. The division into light and heavy pollution using the median PM₁ concentration of 3 µg/m³ appears arbitrary and does not represent physically meaningful pollution regimes. It would be useful to assess whether the conclusions are sensitive to the choice of threshold.

We thank the reviewer for this comment. We agree that the median PM₁ concentration of 3 µg m⁻³ does not define a physically distinct boundary between pollution regimes. This threshold was used solely to provide a descriptive grouping of the sampled conditions according to their relative aerosol loading. We have therefore replaced “heavy pollution” and “light pollution” with “high aerosol loading” and “low aerosol loading” throughout the manuscript. Only one flight (A032) was close to this threshold. Only one flight (A032) was close to the median PM₁ threshold of 3 µg m⁻³. A modest adjustment of this threshold would therefore primarily affect the classification of this flight. However, given its limited aerosol loading, reclassifying A032 is not expected to substantially alter the contrasts between the groups or affect the main conclusions of this study.

5. Line 385. The discussion of northwest air masses during the southwest flight is confusing and should be clarified.

We have corrected the typo regarding the air mass direction in the discussion here:

The Southwest (SW) Flight A027 exhibited the highest NR-PM₁ concentrations among all flights, consistent with the elevated PM₁ concentrations measured at the Paris suburban site (Rambouillet Forest) during the period dominated by the same north-easterly air mass (Yu et al., 2025), suggesting that north-easterly air masses carried higher particle loadings into the Paris region during this flight.

6. Figs. 1 and 2. Consider presenting the standard deviation in addition to the mean concentrations.

We think the reviewer refers to Figure 3 and 4. We have added the error bars ($\pm$ standard deviations) to Figure 3 and 4. We have also fixed an error in Figure 4.

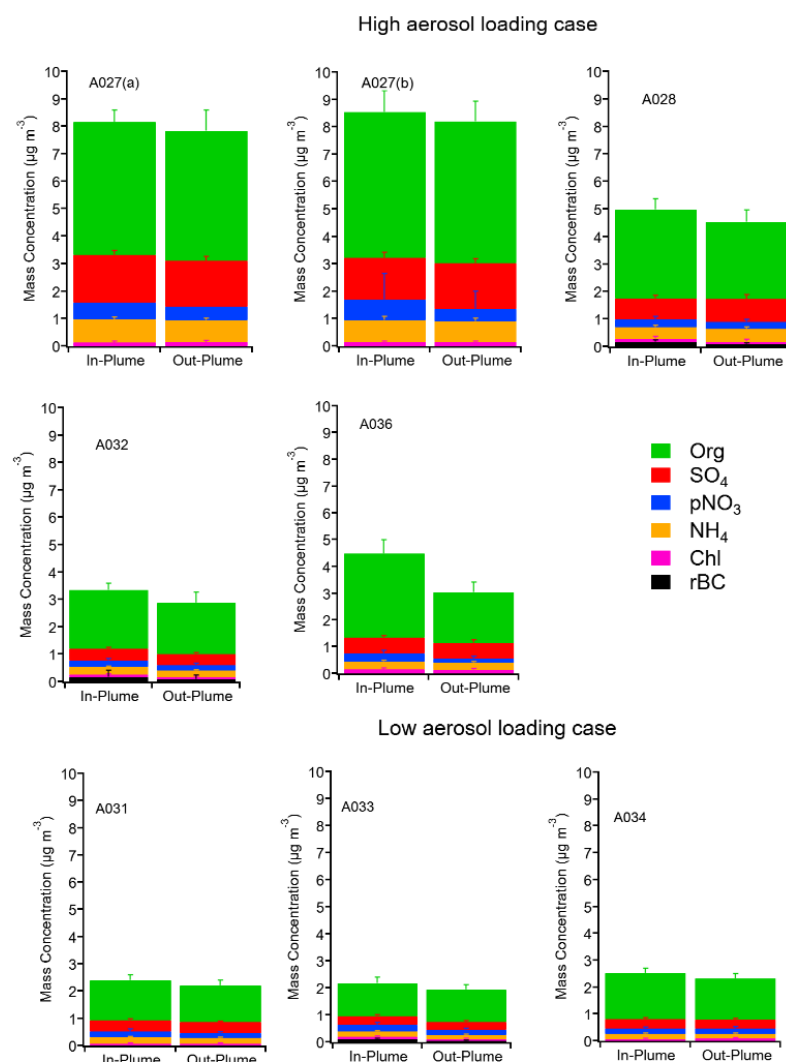


Figure 3. Comparison of average PM₁ or NR-PM₁ component mass concentrations ($\pm$ standard deviation) during the in-plume period and the nearest out-plume background period. rBC data from the SP2 measurements were available only for Flights A028, A032, and A033; therefore, results were reported as PM₁ for these flights and as NR-PM₁ for the remaining flights. Details of the average mass concentrations of each aerosol component, along with the standard deviation for each flight, are presented in Table S4 in the Supplementary Information.

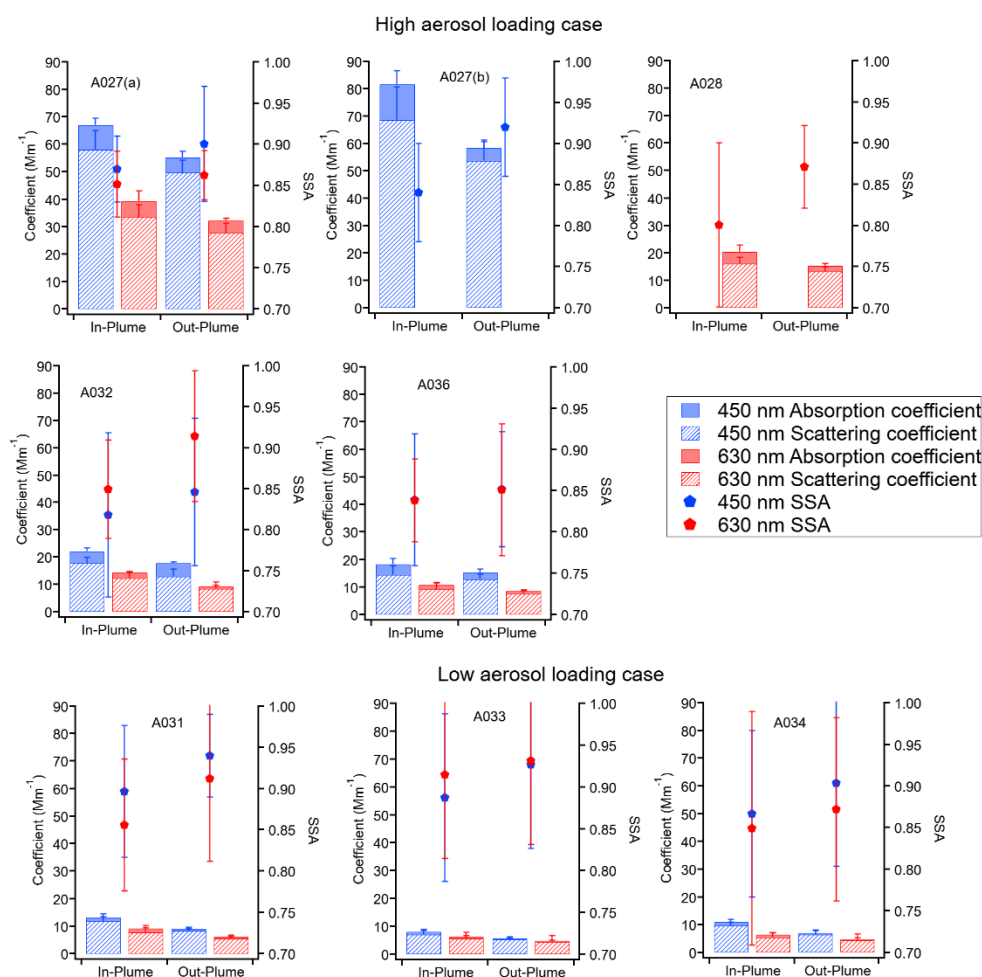


Figure 4. Comparison of average aerosol scattering and absorption coefficients ($\pm$ standard deviation) at 450 and 630 nm for the in-plume period and the nearest out-plume background period. The corresponding SSA at both wavelengths is also presented for in-plume and out-plume periods. Details of the average σ_{ext} , σ_{scat} , and SSA at both 450 nm and 630 nm, along with their standard deviations for each flight, are presented in Table S5 in the Supplementary Information.

7. Line 415. BrC is inferred from AAE, not SAE.

We have revised the sentence combining both Reviewer #1 and Reviewer #3's comments. The revised contexts were presented in our response to Reviewer #1's Comment #3.

8. Line 479. The increase in average OA fraction is relatively small (62.7% to 65.5% and 61.8% to 63.9%) and does not provide convincing evidence for SOA formation.

We have revised the sentence combining all the reviewers' comments. We have added the variation in the f44/f43 ratio as an indicator of the oxidation level of organic aerosol. The results

are presented in our response to Reviewer #1, Comment #1. We have also revised the description here and the revised contexts are presented in our response to Reviewer #2's Comment 9.

9. Fig. 10. A consistent color scheme throughout the manuscript (e.g., red for airborne observations and black for PRG measurements) would improve readability.

Combined Reviewer #2's general comments, we have revised Figure 10 and updated the colour scheme accordingly.

Technical Corrections

1. Line 127. The grammar in "Table S1 in the Supplementary" should be corrected.

We have fixed the related descriptions:

A summary of the instruments used in this study and their availability together with their measurement frequencies is provided in Table S1 and Table S2 in the Supplementary Information.

2. Line 243. The grammar in "Ammonium nitrate and ammonium sulfate was used" should be corrected.

We have fixed the related descriptions:

Ammonium nitrate and ammonium sulfate were used for instrument calibration.

3. Lines 301. The text refers to Eqs. (3) and (4), but the displayed equations are numbered Eqs. (2) and (3). Please correct the cross-references. There also appears to be a typo in the exponent notation in Eq. (3).

We have fixed the equation number issues:

For comparisons at the consistent wavelengths with the airborne A2S2 wavelengths, the PRG NEPH and AE33 results have been scaled using the Ångström exponent approach using equation (2):

$$xAE = -\frac{\ln(\sigma_{\lambda_1}/\sigma_{\lambda_2})}{\ln(\lambda_1/\lambda_2)} \quad (2)$$

Where xAE is the Scattering or Absorption Ångström Exponent (SAE, AAE), σ_{λ_1} and σ_{λ_2} represent the scattering or absorption coefficient at wavelengths λ_1 and λ_2 respectively. The

absorption or the scattering coefficient (σ_λ) at a given wavelength (λ , 450 or 630 nm) can be derived through equation (3):

$$\sigma_\lambda = \sigma_{\lambda_0} \cdot \left(\frac{\lambda}{\lambda_0}\right)^{-x_{AE}} \quad (3)$$

Where σ_{λ_0} is the absorption or scattering coefficient at the wavelength λ_0 . In this study, the σ_{sca} at 630 nm for the NEPH is derived through measurements at 700 nm, and the σ_{abs} at 450 and 630 nm for the AE33 is derived through absorption measurements at 470 and 660 nm, respectively.

There is no typo issue in the exponent notation in the equation. Where x_{AE} is the Scattering or Absorption Ångström Exponent (SAE, AAE) which has already been defined in Line 304 in our original manuscript.

4. Line 337. The grammar in “validity of the is identification approach” should be corrected.

Accept. We have revised the description here:

When SP2 measurements were available, concurrent enhancements in rBC mass concentrations within the identified plumes further supported the validity of the plume identification approach.

5. Line 381. A period is missing.

Accept. We have revised the description here:

According to the flights with available SP2 measurements (A028, A032, and A033), rBC contributed less than 5% to the total PM₁ mass concentration.

6. The Supplement contains inconsistencies between section headings and table/figure captions that should be corrected.

Accept. We have revised the section headings and captions in the supplementary.

References

Ciarelli, G., Theobald, M. R., Vivanco, M. G., Beekmann, M., Aas, W., Andersson, C., Bergström, R., Manders-Groot, A., Couvidat, F., Mircea, M., Tsyro, S., Fagerli, H., Mar, K., Raffort, V., Roustan, Y., Pay, M. T., Schaap, M., Kranenburg, R., Adani, M., Briganti, G., Cappelletti, A., D'Isidoro, M., Cuvelier, C., Cholakian, A., Bessagnet, B., Wind, P., and Colette, A.: Trends of inorganic and organic aerosols and precursor gases in Europe: insights from the EURODELTA multi-model experiment over the 1990–2010

period, *Geosci. Model Dev.*, 12, 4923–4954, 10.5194/gmd-12-4923-2019, 2019.

DeCarlo, P. F., Dunlea, E. J., Kimmel, J. R., Aiken, A. C., Sueper, D., Crounse, J., Wennberg, P. O., Emmons, L., Shinozuka, Y., Clarke, A., Zhou, J., Tomlinson, J., Collins, D. R., Knapp, D., Weinheimer, A. J., Montzka, D. D., Campos, T., and Jimenez, J. L.: Fast airborne aerosol size and chemistry measurements above Mexico City and Central Mexico during the MILAGRO campaign, *Atmos. Chem. Phys.*, 8, 4027–4048, 10.5194/acp-8-4027-2008, 2008.

Middlebrook, A. M., Bahreini, R., Jimenez, J. L., and Canagaratna, M. R.: Evaluation of Composition-Dependent Collection Efficiencies for the Aerodyne Aerosol Mass Spectrometer using Field Data, *Aerosol Science and Technology*, 46, 258–271, 10.1080/02786826.2011.620041, 2012.

Tsimpidi, A. P., Scholz, S. M. C., Milousis, A., Mihalopoulos, N., and Karydis, V. A.: Aerosol composition trends during 2000–2020: in-depth insights from model predictions and multiple worldwide near-surface observation datasets, *Atmos. Chem. Phys.*, 25, 10183–10213, 10.5194/acp-25-10183-2025, 2025.

Velazquez-Garcia, A., Yu, C., Tinorua, S., Bouzidi, H., Aruffo, E., Di Carlo, P., DeCarlo, P., A. Nault, B., Bourriane, T., Tu, K., Pangu, E., L. Pereira, D., Cazaunau, M., Di Antonio, L., Siour, G., Beekmann, M., Bauville, A., Cantrell, C., Formenti, P., Michoud, V., and Denjean, C.: Photochemical ageing drives the browning of urban outflows through secondary brown carbon production, *EGUsphere*, 2026, 1–28, 10.5194/egusphere-2026-4423, 2026.

Xu, L., Lin, G., Wu, C., Chen, X., Wang, H., Chen, C., Chakrabarty, R. K., Gao, M., Vakkari, V., van Zyl, P. G., and Liu, X.: Strong global radiative effects from wildfire dark brown carbon, *Nature Geoscience*, 19, 612–619, 10.1038/s41561-026-01972-9, 2026.