reviewer 2

Review of “Measurement report: Regional transport and vertical variability of aerosol pollution over Beijing using a 528 m meteorological tower” (EGUsphere-2026-801)”

This measurement report presents a valuable multi-year dataset (2020–2024) of vertically resolved aerosol pollution in Beijing, utilizing a unique 528 m meteorological tower. By combining elevated and surface measurements, the study investigates how air mass origin, boundary layer dynamics, and chemical indicators (SOR and NOR) are associated with variability in urban air quality.

The long-term nature of these high-altitude, in situ observations is rare and provides important observational constraints on the interaction between elevated pollution layers and surface conditions. The dataset is therefore of high relevance and well suited for publication as a Measurement Report in ACP.

However, several aspects of the analysis rely on classification-based and correlation-based interpretations that require clearer methodological description and more cautious framing. In particular, the reproducibility of air mass classification, the treatment of measurement and derived uncertainties.

Overall, the dataset is strong and publication-worthy, but the manuscript would benefit from improved methodological transparency and more conservative interpretation aligned with observational evidence.

The manuscript is suitable for publication in ACP after addressing the following key issues:

- Methodological clarity, particularly regarding air mass classification procedures and measurement/instrument calibration records.
- Uncertainty reporting, including variability and limitations in both measured and derived quantities
- Careful limitation of causal interpretation, ensuring that conclusions remain consistent with observational (rather than mechanistic or quantitative attribution) evidence

Thanks for your positive review and valuable suggestions on this work. We have revised thoroughly based on your suggestions. Please allow me to express my respect and gratitude for your valuable insights and suggestions.

I. Main concern/questions

1. Air mass classification methodology and reproducibility

The manuscript attributes differences in aerosol composition to air mass origin (e.g., north, south, northeast). However, the method used to define these air mass categories is not clearly described in the Methods section, and key elements appear only in the Results, the description of the cluster is important.

For example, around L255–L256, a “48-hour trajectory” is mentioned, but the manuscript does not specify:

- The time period over which trajectories were calculated
- The seasonal coverage of the trajectory analysis
- The procedure used to cluster trajectories into air mass categories

Since all conclusions regarding regional transport depend on this classification, the current description is not sufficient for reproducibility.

Required revisions:

- Provide a complete description of the trajectory running/configuration in the Methods section,
- Specify the temporal coverage (dates/seasons/year)
- Describe clearly how air masses clustering was performed and why if applicable.

Thanks for your valuable suggestions. We described the trajectory running/configuration in the Methods section. As follows: The 2-day (48 h) back trajectories were calculated every hour at 500 m height using the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT, NOAA) 4.9 model (Draxler and Hess, 1998), and Global Data Analysis System (GDAS) outputs with 1 °×1 °resolution are used as input (Stein et al., 2015). The trajectories during our observations (shown in Table 1) were then grouped into six clusters using the total spatial variance (TSV) method (Draxler et al., 2012). This method minimizes the inter-cluster differences among trajectories while maximizing the inter-cluster differences, which has been widely used in previous studies (Chen et al., 2015; Kang et al., 2023). The potential source contribution function (PSCF) was used to identify the potential source area (Stohl, 1996). The PSCF method divides the research area into i ×j grids, and the number of air mass trajectories passing through the grid during the research period is counted to calculate the probability of occurrence of pollution trajectories. The time coverage in trajectories clustering is same as observations (Table 1).

1. **Data quality assurance (QA/QC) and instrument calibration**

The measurement report uses reference-grade instruments (such as TOF-ACSM, gas analyzers, TEOM) and includes co-located data from the NZG station. However, QA/QC procedures and calibration details are not described, which is essential for a Measurement Report.

At present, the manuscript does not specify:

- Calibration frequency for each instrument
- Data quality control (DQC) and data quality assurance (DQA) procedures
- How consistency between tower and NZG station data was ensured.

Required revisions:

Discuss more about the QA/QC subsection in the Methods including:

- Calibration frequency and standards for each instrument
- Data filtering and flagging criteria
- Treatment of instrument drift and data gaps
- Impact of the higher altitude on the instruments flow change and the correction applied

Here he author can be mentioned for instance that "The relative ionization efficiency(RIE) calibration of the ACSM was performed before the starting of the campaign and on quarterly basis based on the calibration frequency"

Instrument-specific details:

- TOF-ACSM: the flow calibration, Lense alignment, if done, RIE calibration method, and collection efficiency assumptions the type of vaporizer that is used.
- Gas analyzers: calibration standards and zero/span check frequency
- TEOM: calibration approach and any applied corrections to $PM_{2.5}$
- Confirm that NZG station data follow standard QA/QC procedures and describe how datasets were harmonized

Thanks for your valuable suggestion. We revised the Method section thoroughly based on your suggestion. As follows: "The real-time chemical composition of non-refractory fine particles (NR-$PM_{2.5}$), including SO_4^{2-}, NO_3^-, NH_4^+, Chl^-, and Org, was measured with a ToF-ACSM (Aerodyne Co. Ltd., USA), which relies upon thermal vaporization and 70 eV electron-impact ionization (EI) (Jayne et al., 2000). The ambient particles were sampled into the ToF-ACSM through a $PM_{2.5}$ cyclone (by URG), followed by a Nafion dryer to dry the aerosols. The $PM_{2.5}$ cyclone and the Nafion dryer were connected via a 2-meter-long stainless steel tube with an inner diameter of approximately 7.8 mm (outer diameter: 3/8 inch), and the total sampling flow rate was about 3 L/min. Based on these parameters, the residence time of the sampled air in this tube section is estimated to be approximately 2 seconds. For such a relatively short straight tube and for the predominantly submicron particles ($PM_{2.5}$) in the aerosol, particle losses caused by turbulent deposition, diffusion losses, or gravitational settling are generally negligible. Therefore, no additional correction for particle losses in the sampling line was applied in this study. The dried ambient aerosol particles are focused into the 600 ℃ standard/ captive vaporizer through aerodynamic lenses at a flow of ~ 0.1 L min^{-1}, where they are thermally vaporized to produce gaseous fragments. These vapors are subsequently ionized by electron impact at 70 eV using a tungsten filament, after which the resulting ions are analyzed in a time-of-flight mass spectrometer based on their mass-to-charge ratio. The detected ion signals are then converted into mass concentrations using standard calibration and ionization efficiency procedures, as described in Fröhlich et al. (2013) and Williams et al. (2013). The relative ionization efficiency (RIE) calibration of the ToF-ACSM was performed once two weeks.

The mass concentration of $PM_{2.5}$ was monitored by an R&P model 1400a Tapered Element Oscillating Microbalance (TEOM, Thermo Fisher Scientific Co., USA) at a time resolution of 1 min. This instrument was operated with a hydrophobic filter material to reduce the humidity of the incoming sampled air. The filter was replaced and flow rates of TEOM were checked once two weeks. The collocated gaseous species, including SO_2 (TL43i), NO_x (TL42i), and O_3 (TL49i), were measured by various gas analyzers (Thermo Fisher Scientific Co., USA) at a time resolution of 1 min. Meteorological parameters (T and RH) were measured by an Automatic Weather Station (Huayunshengda, China). All instruments at the CITIC are calibrated once two weeks. Data analysis was performed using Python (version 3.13). All raw observations at the CITIC station were averaged to a temporal resolution of 1 hour. Prior to analysis, quality control (QC) procedures were applied to the dataset, which included corrections based on zero calibrations and the removal of anomalous data.

Additionally, ground level observations at NZG station, including $PM_{2.5}$, SO_2, NO_2, and O_3 with time resolution of 1 hour. The NZG station is located at north of the CITIC with a distance of 4 km (Fig.1). The NZG station (longitude:116.47°, latitude: 39.97°) is one of the national-level ambient air quality monitoring stations operated by the China National Environmental Monitoring Centre (CNEMC). All measurements at the NZG follow mandatory national technical specifications. The QC of observations includes routine zero calibrations, multi-point precision checks, annual accuracy audits, traceable standard-gas transfer through regional QC laboratories, and parallel manual gravimetric comparison for $PM_{2.5}$. The reliability of CNEMC data has been independently verified in peer-reviewed literature through cross-validation against measurements from the U.S. Embassy (Liang et al., 2016). To maintain consistency, observations at the NZG station during the same period listed in Table 1 were analyzed."

1. **Uncertainty reporting and interpretation of observational relationships**

Uncertainty and variability are not consistently reported, and some interpretations extend beyond what can be directly supported by observations.

In particular:

- Derived quantities (SOR, NOR) are reported without uncertainty bounds
- The relationship between RH and SOR/NOR is interpreted as enhanced formation without clearly stating it is an observational association

Required revisions:

- Report mean ± standard deviation (or equivalent) for all grouped analyses
- Include variability or uncertainty ranges in figures
- Rephrase interpretations to indicate association rather than causation for RH–SOR/NOR relationships

Thanks for your suggestions. We added mean ± standard deviation (or equivalent) for all grouped analyses, and variability or uncertainty ranges in Fig.2, Fig.5 and Fig.6. We revised the RH–SOR/NOR relationships as follows: "Notably, the RH in Clusters 2 (54.1%) and 4 (55.4%) was the highest among these six clusters. Further analyses show that both SOR and NOR are closely related to RH (Fig.6). "

II. Minor correction

L63, Please add a reference for this statement.

Thanks and added

L77: Add reference.

Thanks and added.

L79, L88, please remove the hyphen "-"and do the same for the whole document. This sounds like it is AI generated text that might be used for paraphrasing.

Thanks and removed.

L82, L91, L11, L113, add reference

Thanks and added.

L104: Explain the acronym SIAs used here.

Thanks and added.

L118, As the first time using the chemical formula for Nitrogen Dioxide SO_2 and PM, please give the full names for the chemical formulars

Thanks and added.

L124, Please add the latitude and longitude of the tower for the specific geographical location

Thanks and added.

L128, list the mentioned meteorological parameter that were measured

Thanks and added.

L132, Figure 1a, the sampling location is not clearly identifiable within the presented map. It is difficult to precisely locate the site relative to the surrounding residential and commercial areas. I recommend replacing or complementing the current map with a satellite-based map view, which would better illustrate the surrounding urban context and improve spatial interpretability. In its current form, the road network and surrounding neighborhood structure are not clearly visible. For Figure 1b, it should be explicitly stated whether the reported height refers to above sea level (a.s.l.) or above ground level (a.g.l.). For clarity in atmospheric measurements, the altitude should be consistently labeled (e.g., "528 m a.g.l." or "528 m a.s.l." depending on the actual reference used). Regarding Figures 1c and 1d, these figures appear to present results rather than methodological description. I suggest relocating them to the Results section, as they are more consistent with data interpretation than study setup. In addition, the following improvements are recommended to strengthen statistical and graphical clarity:

Thanks for your valuable suggestions. We repotted Fig.1 and Fig.2 by merging Figures 1c and 1d and Fig.2.

- Add error bars (e.g., standard deviation or standard error) to all bar plots showing variability
- Clearly specify the statistical metric used (mean, median, or other) in the caption.
- Explicitly describe the plotted quantity, for example:
  - "Mean $PM_{2.5}$ mass concentration by station (±SD)"
  - or "Median BC concentration with interquartile range"

These improvements will enhance figure readability, statistical transparency, and consistency with ACP presentation standards.

Thanks and revised.

L145: The current description would benefit from explicit explanation, I recommend add this sentence:

"The dried ambient aerosol particles are focused into the 600 °C standard/ captive vaporizer through aerodynamic lenses at a flow of ∼ 0.1 L min−1, where they are thermally vaporized to produce gaseous fragments. These vapors are subsequently ionized by electron impact at 70 eV using a tungsten filament, after which the resulting ions are analyzed in a time-of-flight mass spectrometer based on their mass-to-charge ratio. The detected ion signals are then converted into mass concentrations using standard calibration and ionization efficiency procedures, as described in Fröhlich et al. (2013) and Williams et al. (2013). "

Including a brief and clear description of this sequence would improve methodological transparency and help ensure correct interpretation of the measurement technique.

Thanks and added.

Was the sampling line made of a cyclone connected to the Nafion dryer, or was there a tube in between connecting the cyclone and the dryer? Please explain more clearly how the sampling line setup was configured. If a tube was used, please specify its material (e.g., stainless steel or Teflon) and its length and cross section area. This information will help determine whether particle loss calculations are needed to compensate for particle losses due to wall losses for longer sampling inlet lines.

Thanks for your valuable suggestion. We added the following introduction: " The $PM_{2.5}$ cyclone and the Nafion dryer were connected via a 2-meter-long stainless steel tube with an inner diameter of approximately 7.8 mm (outer diameter: 3/8 inch), and the total sampling flow rate was about 3 L/min. Based on these parameters, the residence time of the sampled air in this tube section is estimated to be approximately 2 seconds. For such a relatively short straight tube and for the predominantly submicron particles ($PM_{2.5}$) in the aerosol, particle losses caused by turbulent deposition, diffusion losses, or gravitational settling are generally negligible. Therefore, no additional correction for particle losses in the sampling line was applied in this study."

L150: precise the model number for the gas measurement instrument mentioned (SO2, NOx and O3, etc).

Thanks and added.

L154: add the latitude and longitude for the NZG station for the transparence in the measurement. L154, I would recommend proving more details about the HYSPLIT model configuration in this section (mentioning the model run details, including dispersion direction: back-trajectory, runtime parameters, as well as the used resolution, period, date, time, season, and year for each run). Similarly, please keep these run parameters in the model output in the Results section.

Thanks and added.

For this Methods section, please consider adding the following:

Add a section detailing the sampling lines for all instruments, as well as the data quality assurance and data quality control procedures that were performed. In addition, include the sampling frequency for each instrument used is useful.

The reader should also be informed about how the data were analyzed and which tools were used to ensure reproducibility. Therefore, please explain the data analysis approach and the software or tools used.

Add a paragraph describing the type of calibration performed, how often calibrations were conducted, how consistent the calibration factors were from one calibration to another, and how calibration was validated and applied to the measurements reported in this study (ACSM, TEOM, and gas measurements). This section is important for ensuring data quality and strengthening the credibility of the results.

In the Methods section, also include a section describing the meteorological characteristics, such as the statistical distribution (e.g., mean, variability) of temperature, relative humidity, wind direction, and wind speed.

Finally, add a section explaining your clustering method, you can refer to the work of Chen et al. (2015) (doi:10.5194/acp-15-12879-2015).

Thanks for your valuable suggestions. We revised the Method section thoroughly based on your suggestions.

L163-L167: add the ±SD at each quantity to show the variability.

Thanks and added.

L175: I would replace during 2020 to 2024 by “from 2020 to 2025 “to indicate the time period

Thanks and changed.

L178: Do you have any PBLH data you can use to support this observation?

We investigated PBLH during 17 October－12 November 2020 based on sounding data, which indicated that the nocturnal PBLH was typically below 300 m, whereas the daytime PBLH was around 1000 m (Ma et al., 2023). Hence, the CITIC station is located above PBL at night while inside PBL during daytime. We added above content in manuscript.

L200–210 **(Figure 2, Panel B):**

The observed daytime increase in organic aerosol and sulfate at the 528 m site may indicate that the station is intermittently located within the evolving planetary boundary layer (PBL). During daytime PBL growth, the upward expansion of the mixed layer can incorporate the site into the PBL, leading to enhanced concentrations of secondary aerosol components such as organic aerosol and sulfate due to stronger vertical mixing and in-situ photochemical production.

However, the observed diurnal behavior of nitrate is opposite to that of organic aerosol and sulfate. This suggests a more complex partitioning process. In particular, it is plausible that a significant fraction of particulate nitrate is present in the form of organic nitrates (PON/$OrgNO_3$), which have different volatility and formation pathways compared to inorganic nitrate. We therefore recommend that the authors quantify the organic nitrate fraction ($POrgNO_3$) if possible, (similar to the observation in this recent paper Habineza et al (2025) DOI: https://doi.org/10.5194/acp-25-15953-2025) or at minimum discuss its potential contribution to the observed nitrate diurnal cycle. The morning increase in nitrate, despite expected boundary-layer growth, remains unclear and requires further examination in terms of partitioning, thermodynamics, and photochemical evolution.

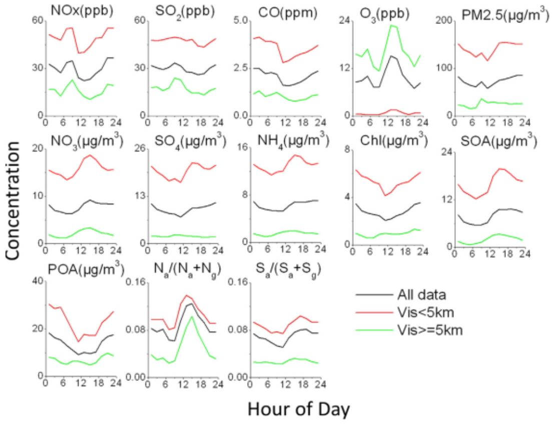
Thanks for your valuable suggestions. We discuss nocturnal nitrate formation via N_2O_5 hydrolysis and organic nitrates. As follows: "Elevated NO_3 levels in the RL (Wang et al., 2018) may also promote organic nitrate formation (Cai et al., 2023), thereby further contributing to the enhanced nighttime nitrate concentrations in this layer."

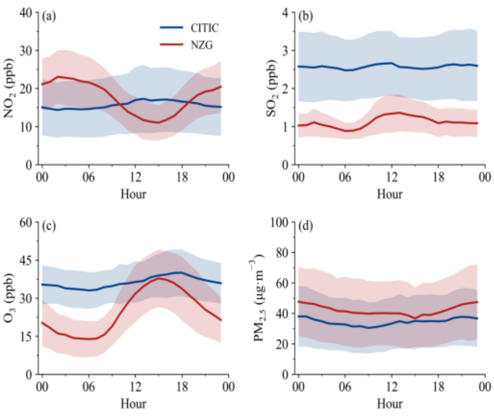
L209–210 (SO_2–NO_2 decoupling):

Similarly, the observed decoupling between NO_2 and SO_2 requires additional justification. In a megacity environment such as Beijing, NO_2 and SO_2 are typically co-emitted from combustion-related sources (traffic, industry, power generation) and therefore often exhibit correlated variability at the surface. The reported vertical and temporal decoupling between these species suggests either (i) distinct source regions and transport pathways, or (ii) differential chemical lifetimes and removal processes. The current explanation invoking a combination of vertical mixing and regional transport is insufficiently constrained without supporting analysis.

Thanks for your comment.

To control air pollution, coal combustion in Beijing is strictly controlled. The comparisons in the following figures support this conclusion. Consequently, ambient SO2 concentrations in Beijing are now predominantly derived from regional transport rather than local sources (Zhang et al., 2019). Therefore, NO_2 and SO_2 represent local and regional transport pollutants, respectively.





Observations in Beijing from Nov. 16, 2012 to Jan. 15,2013 (Quan et al., 2014)

Observations in Beijing from October 2020 to December 2024 (This work)

Quan J.N., X.X. Tie, Q. Zhang, Q. Liu, X. Li, Y. Gao, D. Zhao, 2014. Characteristics of heavy aerosol pollution during the 2012-2013 winter in Beijing, China, Atmos. Environ, 88, 83-89.

Based on the suggestions from you and the other reviewer, we have revised this section thoroughly. We have explained the diurnal variation in three aspects: PBL processes, chemical reactions, and regional transport. Furthermore, we also discuss vertical pollutant mixing in Section 3.4. Details on revision may refer to the secondary version.

We recommend that the authors include boundary layer height (PBLH) data, if available, to better resolve whether the 528 m site is within or above the mixed layer during different periods.

Thanks for your suggestion. We indeed investigate PBLH in previous paper. The following sentence is added: "Our previous observations (17 October – 12 November 2020) indicated that the nocturnal PBLH was typically below 300 m, whereas the daytime PBLH was around 1000 m (Ma et al., 2023)."

For Figures 2(c–d), I recommend replacing the current presentation with a single figure organized by seasonal facets, where each panel displays the four components of NR-$PM_{2.5}$ (e.g., organic aerosol, sulfate, nitrate, and ammonium), similar to the approach used in Fig. 2a of DOI: https://doi.org/10.5194/acp-25-15953-2025. This would improve interpretability by allowing clearer seasonal comparisons of chemical composition and reducing redundancy across panels.

Thanks and revised.

In addition, to strengthen the interpretation of $PM_{2.5}$ variability and its driving factors in this region, I strongly recommend performing a source apportionment analysis (e.g., PMF or equivalent receptor modeling). Such an analysis would provide quantitative insights into the dominant emission sources and their seasonal variability, thereby improving the robustness of the conclusions regarding both primary and secondary aerosol contributions. The correlation between the source profiles and the cluster analysis in section 2.3 Will improve the impact on the measurement report on regional sources and provide a supporting inference on the regional transport's pollution across the used air masses clusters.

Thanks for your suggestions. Following your suggestion, we conducted a source apportionment analysis by SoFi +EPA PMF; however, the results were suboptimal (see the figures below). This is likely attributed to the following factors. First, our measurements captured only a fraction of the total aerosol species, as crustal elements (e.g., Fe, Ti, Ca, Zn, As, Pb, Cu, V, and Ni) were excluded. Additionally, black carbon (BC) data were available for only a limited duration. Second, unlike ground-level aerosols, those in the aloft layer are predominantly composed of secondary components, which inherently complicates traditional source apportionment models. For example, primary aerosol species (including HOA, COA, BBOA, BC, and Cl^-) contribute only 8.0% to the total aerosol mass. Final, the source apportionment cannot pass BS (Bootstrap) error analysis (boot factor 5 <80%). So, we did not add these analyses in revised version.

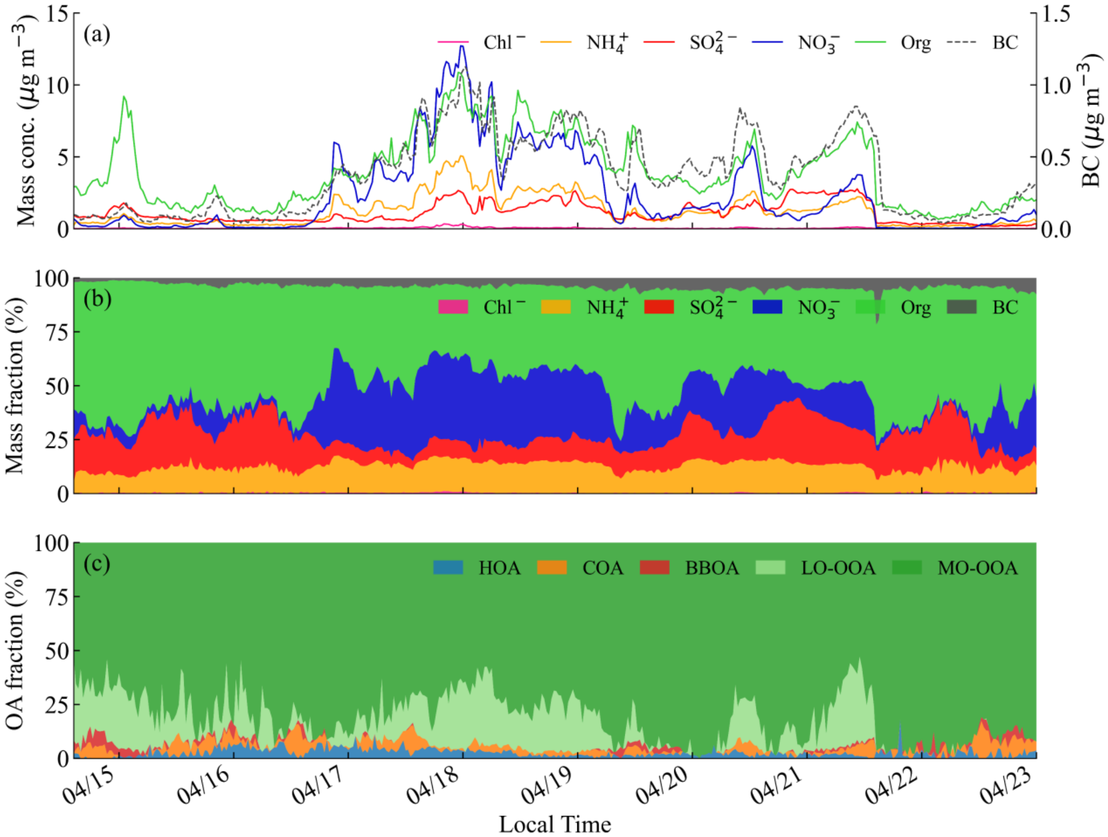


Figure a1. Time series of mass concentrations (a) and fractions (b) of aerosol components, and mass fractions of organics (c) at the CITIC station from April 14 to 23, 2022.

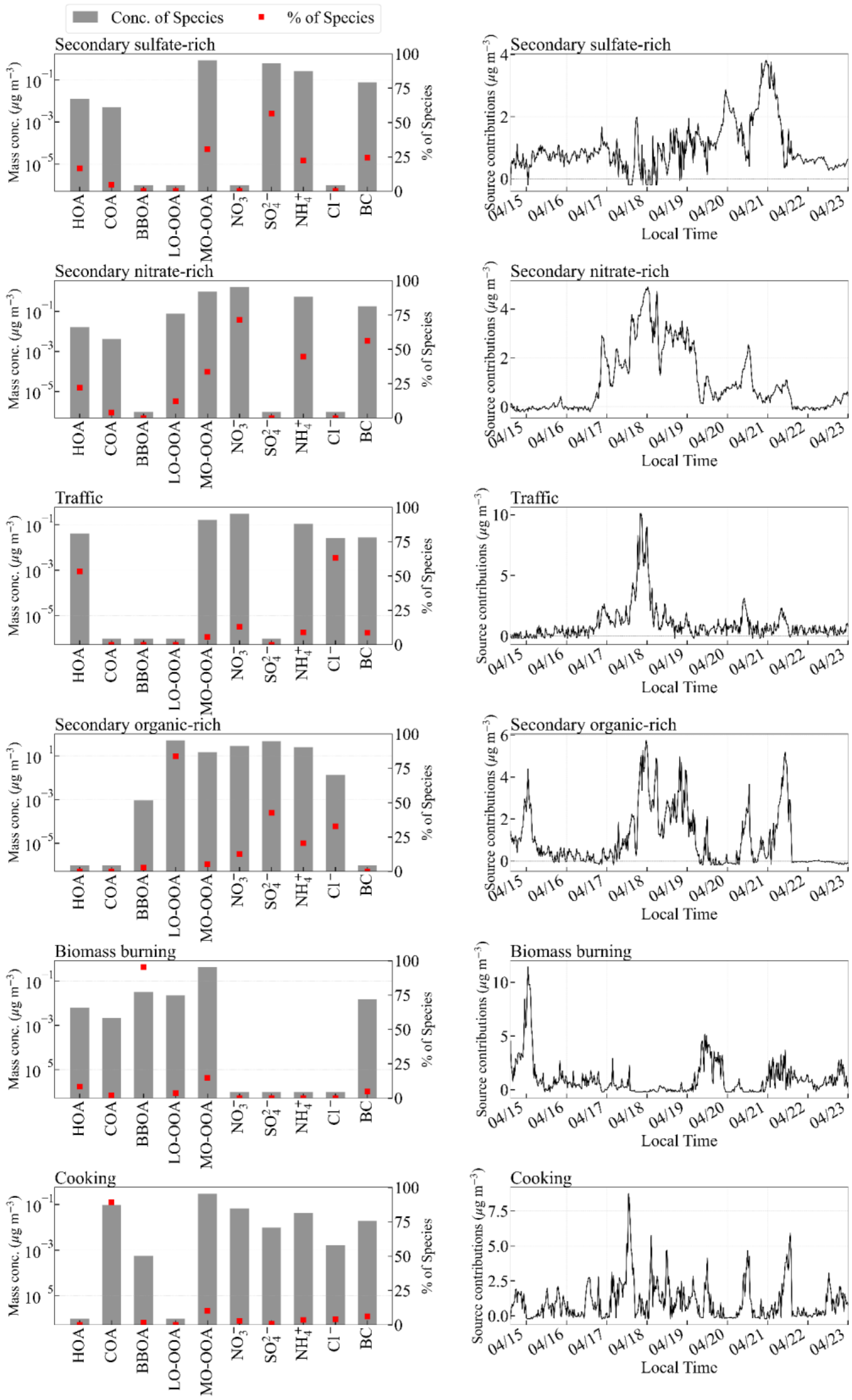


Figure a2. Factor profiles and time series for the 6-factor solution of the $PM_{2.5}$ source apportionment analysis at the CITIC station from April 14 to 23, 2022.

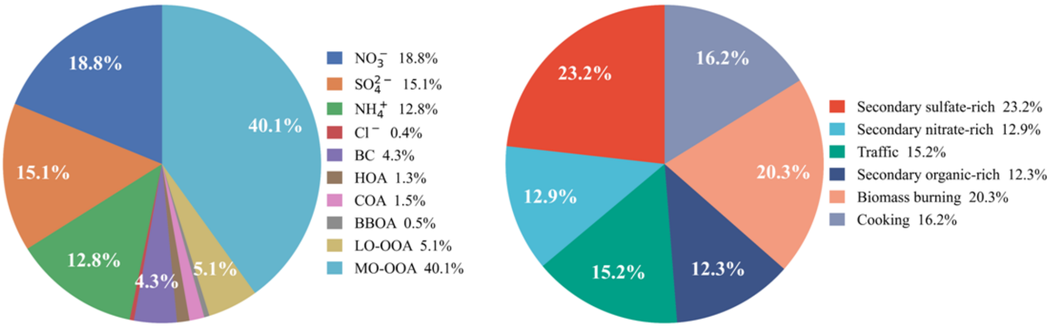


Figure a3. Chemical composition (left) and source apportionment (right) of $PM_{2.5}$ observed at the CITIC station from April 14 to 23, 2022.

Table a1. Mapping of bootstrap factors to base factors

Cluster	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Unmapped
Boot Factor 1	96	3	0	0	0	0	0
Boot Factor 2	0	99	0	0	0	0	0
Boot Factor 3	0	1	97	1	0	0	0
Boot Factor 4	1	5	0	93	0	0	0
Boot Factor 5	0	38	5	0	56	0	0
Boot Factor 6	0	0	0	0	0	99	0

L212: please Link this to the Figure 9, a and b

Thanks and linked.

L 212: Add reference.

Thanks and added.

L222, add reference

Thanks and added.

L223, add reference

Thanks and added.

L262, add reference

Thanks and added.

L264, add reference

Thanks and added.

L 271, Remove the hyphen “-”

Thanks and removed.

L 341 to 346, add the $\pm$SD

Thanks and added.

L361: We recommend performing a correlation analysis between seasonal PM composition and the air mass clusters to assess the relative importance of local and regional sources contributing to PM mass composition at the site. A similar approach applied to PM source profiles would be more useful, particularly if PMF analysis is performed.

Thanks for your suggestions. We conducted PMF analysis, but did not add the result in revision. We explain this in reply to previous question.