

Response to Anonymous Referee #1 (RC1)

Manuscript: egusphere-2026-3299

Decadal transition in the source structure of PM_{2.5} at a coastal background site in North China: weakening combustion and nitrate amplification during high-PM_{2.5} conditions

Overall assessment. This manuscript presents a valuable long-term record of PM_{2.5} composition and source-related changes at Tuoji Island, a coastal background site in the Bohai Strait. The integration of chemical composition, positive matrix factorization (PMF), high-PM_{2.5} sample analysis, ion balance, and ISORROPIA-II diagnostics is well aligned with the scope of ACP. The long observation period and regional-background setting distinguish the study from many urban or episode-based source-apportionment studies.

The principal conclusions are generally supported by the measurements. EC and sulfate decreased across the observed stages, whereas nitrate became relatively more important. The high-PM_{2.5} analysis further shows that nitrate was preferentially amplified in the later stage. The revised PMF treatment is substantially more transparent than a conventional factor interpretation alone: the manuscript reports factor-number testing, an unrotated Base solution, Fpeak sensitivity, Bootstrap mapping, DISP results, species-level residuals, and observed–predicted relationships. The Supplement also makes clear that the measurements included major water-soluble ions, OC/EC, and acid-digested trace elements.

I consider the paper suitable for ACP after revision. The remaining issues mainly concern the comparability of the PMF subset, uncertainty in the high-PM_{2.5} PMF results, the distinction between observational evidence and mechanistic interpretation, and the source-profile constancy assumption underlying the pooled decadal PMF. The revised manuscript should address these points explicitly.

Response. We thank the referee for this positive evaluation of the long-term record, the multi-evidence design, and the transparency of the PMF diagnostics. All four remaining areas of concern, namely the comparability of the PMF subset, the uncertainty of the high-PM_{2.5} PMF results, the separation of observational evidence from mechanistic interpretation, and the source-profile constancy assumption of the pooled decadal PMF, have been addressed in the revised manuscript. The specific changes and their locations are detailed in the point-by-point responses below.

Comment 1. Clarify the relationship between the full composition dataset and the PMF subset. The stage-level composition analysis uses 672 valid samples, whereas PMF uses 313 samples with complete common inputs, and the PMF comparison periods (2012; 2019–2020; 2021–2023) are narrower and of unequal length relative to the composition stages. Sect. 2.5 explains why the two datasets differ. Please state the three PMF-period sample counts explicitly, briefly comment on the comparability of the unequal-length comparison periods (a single early year versus multi-year later

periods), and confirm that the main component changes in the PMF subset are consistent with those in the full dataset.

Response. We agree that the relationship between the two datasets needed to be stated more explicitly. Sect. 2.5 now reports the three PMF-period sample counts (54 for 2012, 81 for 2019–2020, and 178 for 2021–2023, 313 in total) together with the corresponding stage-relative high-PM_{2.5} subsets (11, 17, and 35 samples), and states that these non-contiguous windows of unequal duration are used for stage-level comparisons. The revised composition analysis contains 669 valid samples after exclusion of three records without usable chemical concentrations; the PMF matrix remains unchanged at 313 samples.

To confirm that the PMF subset reproduces the main component changes, Text S5 and Tables S8a–S8b compare the direction of stage changes in PM_{2.5}, OC, EC, SO₄²⁻, NO₃⁻, NH₄⁺, and the NO₃⁻/SO₄²⁻ ratio between composition samples from the PMF-covered years (2012, 2019–2020, and 2021–2023; n = 77, 110, and 290) and the PMF subset. Six of the seven indicators change in the same direction in each adjacent-stage contrast; NH₄⁺ is a small, non-monotonic exception, consistent with its behaviour described in Sect. 3.1. From the early to the late stage, all seven indicators agree in direction. This result is summarised in Sect. 3.2.

Revised manuscript text (Sect. 2.5; p. 5, lines 170-180): Only samples with all common PMF input species and corresponding uncertainties were retained, giving 54 samples for 2012, 81 for 2019–2020, and 178 for 2021–2023 (313 in total); the corresponding stage-relative high-PM_{2.5} subsets contained 11, 17, and 35 samples (Text S5 and Table S8a). ... These non-contiguous windows of unequal duration were used for stage-level comparisons.

Revised manuscript text (Sect. 3.2; p. 9, lines 310-320): The direction of change between adjacent stages agreed between the composition samples from the PMF-covered years (2012, 2019–2020, and 2021–2023) and the PMF subset for six of seven indicators; NH₄⁺ was the exception, with small, non-monotonic changes consistent with Sect. 3.1 (Text S5 and Table S8b). The main factor patterns were also retained in the F_{peak} = 0.5 sensitivity solution (Table S3b and Fig. S3). Excluding 2020 retained the decline in mixed combustion and the increase in secondary nitrate, but changed the quantitative middle-period balance between the port-related mixed and secondary nitrate factors (Text S5 and Table S8c).

Revised Supplement text (Text S5; p. 14): The 313-sample PMF matrix comprised 54 samples from 2012, 81 from 2019–2020, and 178 from 2021–2023; the corresponding stage-relative high-PM_{2.5} PMF subsets contained 11, 17, and 35 samples (Table S8a). The PMF windows are unequal and non-contiguous and are used for stage contrasts. To assess subset representativeness, stage changes in PM_{2.5}, OC, EC, SO₄²⁻, NO₃⁻, NH₄⁺, and the NO₃⁻/SO₄²⁻ ratio were compared between all composition samples from the PMF-covered years (2012, 2019–2020, and 2021–2023; n = 77, 110, and 290) and the PMF subset. Six of seven indicators had the same direction in each adjacent-stage contrast; NH₄⁺ was the exception, with small changes of opposite sign consistent

with its non-monotonic behaviour in the main text. All seven indicators agreed from the early to late stage (Table S8b).

Comment 2. Reflect the sample size in the presentation of the high-PM_{2.5} PMF comparison. The high-PM_{2.5} PMF results are based on 11, 17, and 35 samples. Please state these numbers prominently (including in Fig. 4 and its caption), note that the early-stage estimates are based on smaller subsets, and report the PMF percentages with precision appropriate to the sample size.

Response. Thank you for this suggestion. The subset sizes are now stated prominently: Sect. 3.2 opens the high-PM_{2.5} PMF comparison with the counts 11, 17, and 35, and the Fig. 4 caption lists $n = 54, 81, \text{ and } 178$ for all PMF samples and $n = 11, 17, \text{ and } 35$ for the high-value subsets. The text also states that, because the early-stage subset contains 11 samples, the comparison emphasises factor ranking and broad stage differences. In line with Technical correction 3, all PMF percentages are reported to one decimal place, and this precision is stated in the Fig. 4 caption.

Revised manuscript text (Sect. 3.2; p. 9, lines 300-315): The stage-relative high-PM_{2.5} PMF comparison used 11, 17, and 35 samples remaining after mineral-dust-outlier removal from the PMF subset in 2012, 2019–2020, and 2021–2023, respectively (Fig. 4; Table S8a). ... Because the early-stage high-PM_{2.5} subset contained 11 samples, the comparison emphasises factor ranking and broad stage differences.

Revised manuscript text (Fig. 4 caption; p. 10, lines 320-325): Sample counts are $n = 54, 81, \text{ and } 178$ for all PMF samples and $n = 11, 17, \text{ and } 35$ for the high-value subsets. Values are group-mean relative contributions and are reported to one decimal place; no error bars are shown.

Comment 3. Briefly explain the PMF fit diagnostics. Table S3a reports $Q(\text{True})/Q(\text{expected}) = 13.79$ for the selected seven-factor solution. Fig. S2b and Table S3c show that the major inorganic ions are reproduced well, whereas several weak trace-element variables have lower observed–predicted agreement. Please explain in Sect. 2.5 or Sect. 3.2 that the main nitrate and sulfate factors are supported by well-reproduced strong variables and interpret the metal-related factors cautiously.

Response. We agree that the elevated Q ratio required an explicit explanation in the main text. Sect. 2.5 now interprets $Q(\text{True})/Q(\text{expected}) = 13.79$ directly: it distinguishes the well-reproduced strong secondary-ion variables that support the nitrate and sulfate factors from the weaker fits of OC, EC, and several trace elements, and it states that the metal-related factors are interpreted from tracer profiles and broad stage patterns rather than from small contribution differences. Text S2 and Table S3c provide the species-level R^2 values, scaled-residual fractions, and $Q(\text{True})$ contributions behind this statement, and Figs. S2a–S3 show the corresponding diagnostics.

Revised manuscript text (Sect. 2.5; p. 5, lines 180-185): The seven-factor $Q(\text{True})/Q(\text{expected})$ value was 13.79, indicating that some residual structure remained relative to the prescribed uncertainties. However, the core secondary-ion

variables were well reproduced: observed–predicted R^2 values were 0.979 for NH_4^+ , 0.957 for NO_3^- , and 0.846 for SO_4^{2-} , and 98.1 %, 97.4 %, and 93.3 % of their scaled residuals, respectively, lay between -3 and $+3$. OC and EC contributed most to $Q(\text{True})$, and several weak trace-element variables had lower fits. The metal-related factors are therefore interpreted mainly from their tracer profiles and broad stage patterns rather than from small differences in contribution percentages.

Revised Supplement text (Text S2; p. 3): For the major inorganic ions, 93.29 %–100 % of scaled residuals were within -3 to $+3$; the corresponding proportions were 83.39 % for OC and 84.98 % for EC. Observed–predicted R^2 values were 0.979 for NH_4^+ , 0.957 for NO_3^- , 0.846 for SO_4^{2-} , 0.771 for OC, and 0.615 for EC, whereas values were lower for several weak trace-element variables. EC and OC made the largest individual contributions to $Q(\text{True})$, at 15.26 % and 12.81 %, respectively, and no single species dominated total $Q(\text{True})$.

Comment 4. Describe the high- $\text{PM}_{2.5}$ groups consistently as stage-relative high values. The manuscript already defines the high- $\text{PM}_{2.5}$ samples as the highest 20 % within each stage, and the fixed $75 \mu\text{g m}^{-3}$ comparison in Text S3 is a helpful sensitivity test. Because the actual thresholds differ among stages, please make the stage-relative definition and the differing thresholds explicit in Fig. 2 and its caption so that these samples are not read as pollution episodes of identical absolute severity.

Response. This point is well taken. The stage-relative character of the definition is now explicit wherever these samples are introduced or used. Sect. 2.6 defines them as the highest 20 % selected independently within each stage after mineral-dust-outlier screening and adds that they represent each stage's upper tail rather than identical absolute pollution severity. Fig. 2a now marks the concentration cutoff for selecting high- $\text{PM}_{2.5}$ samples in each stage separately with red dashed segments and labels the sample counts, and the caption describes these elements explicitly. The fixed $75 \mu\text{g m}^{-3}$ comparison is retained in Text S3 as a sensitivity check on this definition.

Revised manuscript text (Sect. 2.6; pp. 5-6, lines 200-205): After screening, samples were ranked by $\text{PM}_{2.5}$ concentration within each stage, and the highest 20 % were selected. These stage-relative high- $\text{PM}_{2.5}$ samples represent each stage's upper tail and do not imply identical absolute pollution severity across stages.

Revised manuscript text (Fig. 2 caption; p. 7, lines 255-260): Red dashed segments mark the stage-specific concentration cutoffs for selecting high- $\text{PM}_{2.5}$ samples, and labels give the high-value and screened sample counts.

Comment 5. Keep the thermodynamic interpretation at the level supported by reverse-mode diagnostics. The manuscript already states that particle-phase stoichiometry does not represent gas-phase NH_3 and that pH is a model diagnostic. This is appropriate. Similar care should be applied to ALWC and NH_4NO_3 -related solid-salt results, because they are calculated using the observed particle-phase ions. These results support the internal consistency of a nitrate-rich aerosol state, but they should

not be described as independent proof of formation rates, gas–particle partitioning, or persistence.

Response. We agree with this reading of the thermodynamic evidence and have applied the same care to all reverse-mode outputs. ALWC is consistently described as model-diagnosed. Section 3.3 now states that the ion-balance indicators, reverse-mode ALWC, and stable-mode salt allocation all derive from the same particle-phase ion measurements. Their agreement describes related aspects of the same nitrate-rich aerosol state but does not provide independent confirmation of formation rates, gas–particle partitioning, or persistence. Because concurrent gaseous NH_3 , HNO_3 , and HCl were not measured, the diagnostics are used for relative stage comparison only. Stable-mode NH_4NO_3 allocation describes the salt distribution calculated from the same ion inputs; Text S4 and the Conclusions explain the relationship between these diagnostics. Reverse-mode pH is used only to examine sensitivity to ion balance, not to infer ambient aerosol acidity or stage differences in acidity (Sect. 2.4 and Text S4; Song et al., 2018).

Revised manuscript text (Sect. 3.3; p. 11, lines 375-385): However, the ion-balance indicators, reverse-mode ALWC, and stable-mode salt allocation are all derived from the same particle-phase ion measurements. Their agreement therefore describes a coherent nitrate-rich aerosol state but does not provide independent confirmation of the underlying mechanism. Because concurrent gaseous NH_3 , HNO_3 , and HCl were unavailable, these diagnostics are used for relative stage comparison rather than direct inference of gas–particle partitioning (Fountoukis and Nenes, 2007; Song et al., 2018).

Revised manuscript text (Sect. 3.3; p. 12, lines 385-390): This allocation shares the ion inputs used for the other thermodynamic diagnostics and describes another aspect of the same inorganic-ion shift.

Revised manuscript text (Sect. 4; p. 13, lines 410-420): Taken together, these related diagnostics support the high-value nitrate signal without independently resolving its formation mechanism.

Revised Supplement text (Text S4; p. 11): Metastable ALWC, stable-mode salts, and particle-phase ion-balance quantities all depend on the same measured ion set. They therefore describe related aspects of the same aerosol state and should be interpreted together rather than as independent diagnostics. Concurrent gaseous NH_3 , HNO_3 , and HCl were unavailable, so these outputs are used for relative stage comparison rather than as direct gas–particle partitioning evidence. Reverse-mode pH is sensitive to ion measurement errors and charge balance (Song et al., 2018). Here, it is used only to examine the model response to ion balance, not to infer ambient aerosol acidity or stage differences in acidity.

Comment 6. Maintain the present caution for mixed PMF factors. The secondary nitrate and secondary sulfate factors are well identified. The biomass/coal combustion and port-emission factors represent mixed signals. Retain the explanation that the port factor combines port activity, shipping, coastal aerosol, and regional metal inputs. If

the current factor names are retained, state their mixed character consistently in the Methods, Results, and figure captions.

Response. We appreciate the referee's endorsement of the mixed-factor interpretation. The comment offered the option of keeping the original names provided their mixed character is stated consistently; we chose the stronger alternative and renamed the factors themselves, so that the names carry the mixed character wherever they appear. F2 is now the mixed biomass-burning and coal-combustion factor and F6 the port-related mixed factor, applied consistently across the methods, results, conclusions, and all figures and tables, including the Abstract, Sect. 3.2, Fig. 4, Text S2, Fig. S3, and Table S8c. Sect. 3.2 explains the mixed composition of each factor: F2 combines biomass-burning and coal-combustion tracers, and F6, whose profile V/Ni ratio of 0.10 is inconsistent with pure ship exhaust, reflects port and shipping activity, coastal aerosol, and regional metal inputs.

Revised manuscript text (Sect. 3.2; p. 8, lines 280-285): The resolved factors were F1 secondary sulfate, F2 mixed biomass burning and coal combustion, F3 mineral dust, F4 secondary nitrate, F5 sea salt, F6 a port-related mixed factor, and F7 industrial metals.

Revised manuscript text (Sect. 3.2; pp. 8-9, lines 285-290): F2 combined K^+ , OC, and EC with Pb, As, and SO_4^{2-} , supporting a mixed biomass-burning and coal-combustion interpretation. ... F6 contained Na^+ , Fe, Ni, and Cu but had a profile V/Ni ratio of 0.10, inconsistent with a pure ship-exhaust signature; it was therefore interpreted as a mixed factor reflecting port and shipping activity, coastal aerosol, and regional metal inputs.

Revised manuscript text (Abstract; p. 1, lines 25-30): the mixed biomass-burning and coal-combustion factor decreased from 43.4 % to 13.7 %, indicating weakened primary-combustion signals. ... In 2021–2023, the port-related mixed factor was the largest stage-average PMF factor (27.1 %)

Revised Supplement text (Fig. S3; p. 8): F2 and F6 are interpreted as the mixed biomass-burning and coal-combustion factor and the port-related mixed factor, respectively.

Comment 7. Revise the title to reflect the single-site design. "Over North China" may be read as a multi-site regional analysis. Replace it with wording such as "at a coastal background site in North China", or otherwise make the single-site spatial scope explicit in the title, to match the spatial scope of the evidence.

Response. Agreed. The title has been revised following the wording suggested by the referee and now makes the single-site scope explicit.

Revised manuscript text (Title; p. 1): Decadal transition in the source structure of $PM_{2.5}$ at a coastal background site in North China: weakening combustion and nitrate amplification during high- $PM_{2.5}$ conditions

Comment 8. Justify the source-profile constancy assumption of the pooled decadal PMF. Combining samples from 2012 to 2023 in a single PMF run assumes that factor

profiles remain reasonably invariant over a decade of substantial emission change. Please add a brief justification (for example, reference to pooled long-term PMF practice and to the Base, Fpeak, Bootstrap, and DISP stability diagnostics already reported) and acknowledge this assumption as a methodological limitation where quantitative factor comparisons are discussed.

Response. We agree that the pooled decadal PMF rests on a source-profile constancy assumption that must be stated openly. Sect. 2.5 now justifies the pooled design by reference to a previous PMF application that combined multiple sites and time periods (Yang et al., 2023) and states the assumption directly: the pooled PMF assumes that the chemical composition of each factor remains broadly stable across periods. We deliberately present the Base, Fpeak, Bootstrap, DISP, residual, and observed–predicted diagnostics as evidence of numerical stability and model fit rather than as proof of period-invariant profiles, and cross-period interpretation is restricted to chemically coherent factor profiles and broad stage differences. The same assumption is stated in the Conclusions and Text S2 (Xie et al., 2022).

Revised manuscript text (Sect. 2.5; p. 5, lines 175-180): A pooled solution provided a common factor-profile framework across periods, consistent with a previous PMF application that combined multiple sites and time periods (Yang et al., 2023).

Revised manuscript text (Sect. 2.5; p. 5, lines 185-195): A single pooled solution was used to compare relative factor contributions under a common tracer framework across the three PMF periods. The pooled PMF assumes that the chemical composition of each factor remains broadly stable across periods. Bootstrap, DISP, Fpeak, residual, and observed–predicted diagnostics were used to assess numerical stability and model fit. Cross-period interpretation focuses on chemically coherent factor profiles and broad stage differences in their relative contributions.

Revised manuscript text (Sect. 4; p. 13, lines 425-430): The pooled PMF assumes that the chemical composition of each factor remains broadly stable across periods (Xie et al., 2022).

Revised Supplement text (Text S2; p. 3): The pooled solution provides a common profile framework for stage comparison, following a PMF application that combined multiple sites and time periods (Yang et al., 2023). The pooled PMF assumes that the chemical composition of each factor remains broadly stable across periods (Xie et al., 2022).

Technical correction 1. Use either "enrichment factor" or "enrichment ratio" consistently (both currently appear, e.g., in Sect. 3.1 and Sect. 3.2) and state the calculation once.

Response. Done. The manuscript now uses enrichment ratio throughout, and the calculation is stated once, in Sect. 3.1, in a single definition that covers both of its usages: chemical components and PMF factors. The Fig. 3b axis label and the Fig. 3 caption were updated to the same term.

Revised manuscript text (Sect. 3.1; p. 8, lines 260-265): For chemical components, the enrichment ratio is the mass fraction in stage-relative high-PM_{2.5} samples divided by the corresponding screened stage-average fraction; for PMF factors, it is the relative factor contribution in the high-value subset divided by its screened stage-average contribution. Ratios were calculated from unrounded values.

Revised manuscript text (Fig. 3 caption; p. 8, lines 275-280): (b) Enrichment ratios relative to screened stage-average fractions; the dashed line marks a ratio of 1.

Technical correction 2. Define all error bars and statistical summaries directly in the relevant figure captions.

Response. Thank you for pointing this out. Every figure caption now defines the statistics and uncertainty elements shown, and captions state explicitly where no error bars or uncertainty intervals are displayed. The Fig. 1 caption defines the violin, box, whisker, and point elements and the bootstrap 95 % confidence intervals; the Fig. 2, 3, and 4 captions state the statistics plotted and note the absence of error bars where applicable; the Fig. 5 caption defines the medians and interquartile bars; and the Fig. S9 caption states that no uncertainty intervals are shown.

Revised manuscript text (Fig. 1 caption; p. 6, lines 210-215): (a) PM_{2.5} distributions; violin widths show distributions, boxes show interquartile ranges, centre lines show medians, whiskers extend to 1.5 times the interquartile range, grey points show a subsample of observations for visualisation, and coloured points and labels show means. (b, c) Bars show stage means, and error bars show bootstrap 95 % confidence intervals for the means. (d) Ratios were calculated from stage-mean concentrations.

Revised manuscript text (Fig. 3 caption; p. 8, lines 275-280): Panel (a) shows means of sample-level mass fractions, panel (b) shows ratios of these mean fractions, and panel (c) shows ratios of group-mean NO₃⁻ and SO₄²⁻ concentrations. No error bars are shown.

Revised manuscript text (Fig. 5 caption; p. 12, lines 385-390): Points and labels show medians, and horizontal bars show interquartile ranges.

Revised Supplement text (Fig. S9; p. 18): Frequencies are based on individual trajectories rather than daily modal assignments; equal-season-standardised stage frequencies are reported in Text S6. No uncertainty intervals are shown.

Technical correction 3. Review the number of decimal places used for PMF contributions based on small subsets.

Response. Agreed. All PMF contribution percentages are now reported to one decimal place, a precision matched to the subset sizes, and the Fig. 4 caption states this convention explicitly.

Revised manuscript text (Fig. 4 caption; p. 10, lines 320-325): Values are group-mean relative contributions and are reported to one decimal place; no error bars are shown.

References cited in this response

- Fountoukis, C., and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K^+ – Ca^{2+} – Mg^{2+} – NH_4^+ – Na^+ – SO_4^{2-} – NO_3^- – Cl^- – H_2O aerosols, *Atmos. Chem. Phys.*, 7, 4639–4659, <https://doi.org/10.5194/acp-7-4639-2007>, 2007.
- Song, S., Gao, M., Xu, W., Shao, J., Shi, G., Wang, S., Wang, Y., Sun, Y., and McElroy, M. B.: Fine-particle pH for Beijing winter haze as inferred from different thermodynamic equilibrium models, *Atmos. Chem. Phys.*, 18, 7423–7438, <https://doi.org/10.5194/acp-18-7423-2018>, 2018.
- Xie, M., Lu, X., Ding, F., Cui, W., Zhang, Y., and Feng, W.: Evaluating the influence of constant source profile presumption on PMF analysis of $\text{PM}_{2.5}$ by comparing long- and short-term hourly observation-based modeling, *Environ. Pollut.*, 314, 120273, <https://doi.org/10.1016/j.envpol.2022.120273>, 2022.
- Yang, Z., Islam, M. K., Xia, T., and Batterman, S.: Apportionment of $\text{PM}_{2.5}$ Sources across Sites and Time Periods: An Application and Update for Detroit, Michigan, *Atmosphere*, 14, 592, <https://doi.org/10.3390/atmos14030592>, 2023.