Response to reviewer comments

Reviewer #2

This manuscript proposes a PMF-based reconstruction method (PMFr) for imputing missing values in $PM_{2.5}$ speciation datasets by explicitly using source-receptor relationships, rather than relying solely on conventional statistical or machine-learning approaches. The study is interesting and potentially valuable, particularly because it does not evaluate imputation quality only in terms of reconstruction error, but also examines whether the reconstructed dataset preserves consistency in subsequent source apportionment results. This is a meaningful strength of the work. The comparison with several benchmark methods under multiple missing-data scenarios is also generally appropriate.

However, the manuscript still requires revision before it can be considered for publication. The paper shows clear promise, but some of the claims are broader than what is fully supported by the presented analysis, and several parts of the interpretation would benefit from more careful and balanced framing. In particular, the manuscript should more clearly distinguish the conditions under which PMFr performs especially well from those under which its advantage is limited, and it should present the generalizability of the method more cautiously in light of the assumptions underlying PMF-based reconstruction.

Major Comments:

Comment #1:

The method is evaluated using data from a single urban site over a limited observational period and for a specific set of $PM_{2.5}$ chemical species. However, the conclusion extends the potential applicability of PMFr to other atmospheric datasets, including VOC-related contexts. While this extension may be reasonable as a future possibility, the current manuscript does not yet demonstrate such breadth. I suggest that the authors moderate the scope of their claims and more clearly state that the present findings support the method under the conditions tested in this study.

Response:

We agree that our original conclusion overstated the breadth of the PMFr method by extending its applicability to other datasets (e.g., VOCs) without empirical demonstration in the current study. To address this, we have carefully revised the Conclusion section to strictly constrain the scope of our claims.

Our revised Conclusions section is detailed below (Line 330-341):

"We developed a physically interpretable imputation method (PMFr) for reconstructing missing $PM_{2.5}$ speciation data by leveraging source-receptor relationships encoded in key chemical species. Benchmarking against commonly used imputation techniques, including Mean, LI, KNN, BPCA, and a deep learning predictive model, demonstrates

that PMFr achieves improved accuracy and robustness while preserving physical and chemical interpretability, especially for key marker species. Crucially, the PMFr-completed dataset is better suited for subsequent PMF source apportionment because it preserves source-profile composition and source-contribution temporal features. Nevertheless, the advantage of PMFr may become less substantial when source-related constraints are weakened, such as when all key tracers for a specific source factor are simultaneously missing, or when baseline methods can already capture stable co-variation patterns for certain species. These chemically consistent and physically meaningful estimates also rely on the temporal stability of source chemical compositions. Recognizing the limitations of such static assumptions for long-term datasets, we highlight the necessity of systematically verifying source stability in extended applications. Therefore, this work offers a simple and generalizable solution that strengthens the reliability of real-world speciation datasets and enhances their suitability for source apportionment and policy-relevant analyses."

Comment #2:

Several explanations offered for species-specific performance differences are plausible and scientifically sensible, but they are still interpretive rather than directly demonstrated. For example, the discussion of lower sulfate performance, or the explanation of differential behavior for OC/EC and NH_4^+/NO_3^-, seems to go beyond the evidence shown in the main performance metrics. These interpretations should be framed more cautiously, using language such as "may reflect," "is likely associated with," or "is consistent with," unless additional analysis is provided to directly support those mechanistic explanations.

Response:

Following the reviewer's suggestion, we have carefully reviewed the discussion section and moderated the scope of our claims. We have replaced definitive causal language with more cautious and nuanced phrasing to accurately reflect that these explanations are plausible hypotheses rather than directly proven mechanisms.

The corresponding modifications in the revised manuscript are detailed below:

1. Original text: The absence of other cations like Na^+ and Mg^{2+} impact the imputation efficiency when missing SO_4^{2-} concentration values are high. The formation of NH_4NO_3 dominates nitrate, while $(NH_4)_2SO_4$ account for only part of sulfate.

   Revised Text (Line 188-190): The absence of other cations like Na^+ and Mg^{2+} may impact the imputation efficiency when the missing SO_4^{2-} concentrations are high. This difference is likely because the formation of NH_4NO_3 typically dominates nitrate, while $(NH_4)_2SO_4$ accounts for only a portion of the total sulfate.

2. Original Text: EC is primarily emitted from motor vehicles, whereas OC consists of both primary organic carbon (POC) and secondary organic carbon (SOC); POC is directly emitted, while SOC forms in the atmosphere through secondary processes. POC can partially originate from motor vehicles, whereas SOC is associated with secondary sources such as SS and SN

   Revised Text (Line 203-206): EC is primarily emitted from motor vehicles, whereas OC encompasses both directly emitted primary organic carbon (POC) and secondary organic carbon (SOC) formed through atmospheric processes. The behavior of POC is consistent with partial origins from vehicular emissions, while the variations of SOC are likely associated with secondary sources such as SS and SN.

3. Original Text: The decline is attributable to the absence of key tracers, consistent with the tracer-dependent variability observed at the NEPB site—where the strong OC-EC correlation reflects their common origin in motor-vehicle emissions. PMFr is affected because PMF overestimate the loading of OC and EC in the OT factor, thereby underscoring their contributions from other sources.

   Revised Text (Line 220-223): The decline is likely attributable to the absence of key tracers, consistent with the tracer-dependent variability observed at the NEPB site—where the strong OC-EC correlation may reflect their common origin in motor-vehicle emissions. The performance of PMFr may be impacted because PMF tends to overestimate the loading of OC and EC in the OT factor, thereby obscuring their contributions from other sources.

4. Original Text: This improvement arises because Ti is predominantly emitted from dust sources, enabling PMFr to estimate missing values using the characteristic Ti-Ca-Si ratios in source profiles once the CD factor is identified

   Revised Text (Line 233-235): This improvement is likely associated with the predominant emission of Ti from dust sources, enabling PMFr to estimate missing values by leveraging the characteristic Ti-Ca-Si ratios in source profiles once the CD factor is identified.

5. The performance of DBN declines for ionic species due to insufficient valid training samples and variables caused by long missing gaps and increasing number of missing specie (Figures S22, S23, S24, and S25). NH_4^+ and NO_3^- are strongly correlated due to the predominance of NH_4NO_3 during fall in NEPB site.

   Revised Text (Line 247-249): The performance of DBN declines for ionic species, which may be attributed to insufficient valid training samples and variables caused by long missing gaps and an increasing number of missing species (Figures S22–S25). Furthermore, NH_4^+ and NO_3^- are strongly correlated, a pattern consistent with

the predominance of NH_4NO_3 during the fall at the NEPB site.

Comment #3:

One of the strengths of the manuscript is that it does not hide the fact that PMFr is not uniformly superior in every situation. There are cases in which performance is weaker, or where the gap between PMFr and alternative methods narrows. These include certain species such as sulfate, situations involving high Fe concentrations, and some scenarios involving medium gaps or instrument-failure-type missingness. These limitations are scientifically important and should be more explicitly synthesized in the discussion. A dedicated paragraph or subsection on the strengths and limitations of PMFr across species types and missingness patterns would make the manuscript more informative and more credible.

Response:

To address this comment, we have added a dedicated subsection, Section 3.4 "Applicability and Limitations of PMFr", to synthesize the strengths and limitations of PMFr across species types and missingness patterns. In this section, we clarify that PMFr is most applicable when source-related chemical structures can be constrained by at least one tracer for each source factor, and when the imputed dataset is intended for subsequent source apportionment analysis.

We further discuss the situations in which the advantage of PMFr becomes less substantial. First, when the missing pattern shifts from MCMI to MCMS, source-related observational constraints become less robust because multiple species are absent at the same timestamp. Under this condition, the corresponding source contribution vector (G) is less directly constrained by observed species, and the reconstruction should be interpreted with caution. We therefore added quantitative examples showing that the performance of PMFr declines from MCMI to MCMS for NH_4^+ and NO_3^-, while also noting that sensitivity analysis indicates that PMFr can still outperform baseline methods when the pre-imputation step provides a reasonable estimate of the general temporal variation of the missing species.

Second, we discuss species for which the advantage of PMFr is less substantial, including SO_4^{2-} and crustal elements such as Ca, Si, and Fe. For crustal elements, baseline methods can become competitive because these species are primarily emitted directly and often exhibit relatively stable inter-variable correlations.

Importantly, we also emphasize that comparable concentration-level performance does not necessarily imply that baseline methods are equally reliable for further source apportionment. To demonstrate this point, we added a downstream PMF evaluation section, **"3.3 Assessing the Impact of Imputation on PMF Source Apportionment".** In this section, we selected two representative cases where the advantage of PMFr

became less pronounced: SO_4^{2-} missingness in Case 2 and high-concentration Fe missingness in Case 5. By comparing PMF-resolved source profiles and source contributions derived from different imputed datasets with those from the complete dataset, we showed that PMFr better preserved source-profile composition and source-contribution temporal patterns. These results clarify that the value of PMFr lies not only in direct concentration reconstruction, but also in preserving source-related chemical and temporal structures required for physically interpretable PMF analysis.

The following text has been added to Section 3.4 of the revised manuscript:

"Applicability and Limitations of PMFr (Line 292-318):

PMFr is applicable when source-related chemical structures can be constrained by at least one tracer for each source factor, and the completed dataset is suitable for subsequent source apportionment analysis. One limitation of PMFr is related to missing patterns in which source-related constraints become insufficient. As shown in Table S7, the performance of PMFr declines when the missing pattern shifts from MCMI to MCMS. For NH_4^+ at a 10% missing rate, the MAPE increases from 9.57% under MCMI to 20.67% under MCMS, and the IoA decreases from 0.98 to 0.95. For NO_3^- at a 10% missing rate, the MAPE increases from 14.82% under MCMI to 23.92% under MCMS. At a 20% missing rate, the MAPE increases from 13.63% to 25.87% for NH_4^+, and from 22.81% to 28.46% for NO_3^-. As shown in Table S6, when OC and EC are simultaneously missing, the performance of PMFr becomes comparable to that of baseline methods. For instance, at a 10% missing rate, the R^2 values for OC are 0.73 for PMFr, 0.74 for DBN, 0.68 for KNN, and 0.66 for BPCA. For EC, R^2 values are 0.84 for PMFr, 0.85 for BPCA, 0.80 for DBN, and 0.79 for KNN. Fundamentally, PMFr assumes that the source contribution vector (G) can be sufficiently constrained by observed species, which requires at least one key tracer for each factor. The key tracers used for imputation and source identification are shown in Table S13. If all key tracers for a specific source are simultaneously missing, the corresponding source contribution vector G is less directly constrained by observed species and should be interpreted with caution. Nevertheless, sensitivity analysis indicates that PMFr can still outperform baseline methods when the pre-imputation step provides a reasonable estimate of the general temporal variation of the missing species (Text S5 and Table S14). The numerical advantage of PMFr is less substantial for certain species such as SO_4^{2-} and crustal elements. For crustal elements, baseline methods can become competitive because these species are primarily emitted directly and usually exhibit relatively stable inter-variable correlations. As shown in Table S5, when imputing Ca, Si, and Fe at a 15% missing rate, several statistical or machine-learning methods perform comparably to PMFr. For Ca, the R^2 values are 0.93 for PMFr, 0.91 for BPCA, and 0.90 for DBN. For Si, PMFr achieves an R^2 of 0.82, which is matched by DBN and closely followed

by KNN (0.79). For Fe, the R^2 values are 0.83 for PMFr, 0.86 for DBN, 0.84 for KNN, and 0.84 for BPCA, with DBN and KNN achieving slightly higher IoA values than PMFr. This reduced separation suggests that statistical or machine-learning methods can capture stable co-variation patterns among some primary species, thereby reducing the relative advantage of the source-constrained PMFr method for these specific cases. However, comparable concentration-level performance does not necessarily imply that baseline methods are equally reliable for source apportionment. The PMF evaluation results showed that PMFr better preserved source-profile composition and source-contribution temporal patterns, even in representative cases where direct imputation metrics became comparable among methods."

Comment #4:

The manuscript provides multiple reasons for adopting the 7-factor solution, including interpretability, residual behavior, and diagnostic stability. However, the current presentation reads more as a list of supporting points than as a clearly structured argument. The authors should revise this section so that the decision logic becomes easier to follow. For example, the discussion could more explicitly distinguish why the lower-factor solutions were insufficient, why the higher-factor solutions were over-resolved or physically less meaningful, and why the selected solution best balanced interpretability and statistical diagnostics.

Response:

We sincerely thank the reviewer for this constructive suggestion. In the revised manuscript, we have reorganized Section 3.1 to explicitly follow a decision logic that contrasts the 7-factor solution with alternative scenarios. Specifically:

1. **4-6 factor solutions:** We now explicitly state that these were statistically insufficient, as evidenced by a sharp decline in the Q/Q_{exp} ratio (11.2%) when increasing to 7 factors, and physically inadequate due to the improper lumping of distinct sources like traffic and secondary sulfate.
2. **8-9 factor solutions:** We clarify that these led to statistical over-resolution with diminishing returns in Q/Q_{exp} improvements and significant instability (high BS unmapped rates). Physically, these solutions resulted in uninterpretable splitting of stable sources like coal combustion.
3. **Optimal 7-factor solution:** We demonstrate how this solution best balances robust statistical diagnostics with clear, physically meaningful source identification.

The revised text is provided below (Line 130-166):

“PMF solutions were explored with four to nine factors using datasets containing 10% missing values. The best-fitting solution was selected by the model performance, including the interpretability of the factor profiles, which is the key basis for determining the optimal factor number and imputation, and the distributions of scaled residuals (Figures S2 and S3). Bootstrapping (BS), displacement (DISP), and combined BS-DISP analyses were also performed for these solutions. Four-to-six factor solutions were statistically insufficient to fully explain the variance in the input data matrix. When the factor number increased from six to seven, the Q/Q_{exp} ratio experienced a decline of 11.2%. This drop indicates that the 6-factor model leaves a substantial amount of residual variance unexplained. Because of this lack of statistical resolution, these lower-factor solutions failed to effectively decouple distinct emission sources. Specifically, the 5-factor solution improperly lumped on-road traffic (OT) emissions with metal smelting (Figure S5). In the 6-factor solution, sulfate and nitrate were mixed together as a single identified secondary inorganic aerosol factor (Figure S6). Eight and nine factor solutions demonstrated statistical over-resolution with diminishing returns. As the factor number increased from seven to eight, the Q/Q_{exp} ratio dropped less dramatically (8.5%) compared to the previous step. Furthermore, the 8-factor solution exhibited a high unmapped rate during the BS analysis, highlighting severe statistical instability. From a physical perspective, these higher-factor solutions over-resolved the data into physically meaningless components. For instance, the 8-factor solution isolated a Cu-high loading factor that lacks a clear chemical profile (Figure S7), while the 9-factor solution further fragmented the coal combustion source into two unidentifiable sources (Figure S8). For the 7-factor solution, the model predicted concentrations of tracers such as Ca, V, NH_4^+, and NO_3^- correlated with the observed values with coefficients of determination R^2 of 0.92, 0.91, 0.98, and 0.88, respectively (Table S4).

The high R^2 values of bulk species indicate that the 7-factor model fits well for the data. For tracers like Si, Mn, Se, and Cu, the scaled residuals follow a normal distribution with a mean of 0 and a variance of 1. For bulk species like NH_4^+, NO_3^-, and SO_4^{2-}, the scaled residuals exhibit a light-tailed distribution, with the highest frequency concentrated near 0 and ranging from -2 to 2. Additionally, the scaled residuals of the bulk species OC and EC follow a normal distribution with a mean of 0 and a variance of 1. The distribution of scaled residuals demonstrates the validity of our solution. Physically, the 7-factor solution successfully decouples all distinct emission sources without redundant splitting.”

Comment #5:

An important contribution of the study is not only that PMFr often performs well numerically, but also that it preserves source-related structure in a way that may be

more physically meaningful for subsequent source apportionment. At present, however, these two strengths are sometimes discussed together as if they were the same claim. The paper would be stronger if it clearly distinguished between them. For example, there may be situations in which another method performs competitively on certain numerical metrics, whereas PMFr retains greater interpretive consistency in the source-apportionment context. Making this distinction explicit would sharpen the central message of the paper.

Response:

Indeed, numerical imputation accuracy and preservation of source-related structure are two distinct aspects of PMFr and should be clarified. Numerical metrics such as R^2, IoA, and MAPE evaluate concentration-level reconstruction accuracy, whereas source-structure preservation evaluates whether the completed dataset remains suitable for physically interpretable PMF source apportionment.

To address this issue, we revised the manuscript to define two complementary validation endpoints: direct reconstruction accuracy and physical source-feature preservation. We further added a subsequent PMF evaluation to explicitly assess the second endpoint. Two representative cases were selected for this analysis: SO_4^{2-} missingness in Case 2 and high-concentration Fe missingness in Case 5. These cases were chosen because the numerical advantage of PMFr over baseline methods became less substantial, allowing us to test whether PMFr could still retain greater interpretive consistency in downstream PMF analysis under such narrowed performance gaps.

The results showed that PMFr better preserved the source-profile composition of the secondary sulfate and crustal dust factors, including the SO_4^{2-}/ NH_4^+, Fe/Ca, and Ca/Si relationships, and achieved the highest agreement with the complete-dataset source contributions. PMFr also more consistently reproduced the temporal features of these sources, including daytime SS enhancement and CD peaks associated with daytime dust-related activities. This clarification sharpens the central message that PMFr is valuable not only because it reconstructs missing concentrations accurately, but also because it better preserves source-related chemical and temporal structures required for subsequent PMF analysis.

Revised Section 3.3 "Assessing the Impact of Imputation on PMF Source Apportionment" as below (Line 260-290):

"Results showed that the advantage of PMFr over baseline methods narrowed mainly under two challenging conditions: instrument-failure-type missingness and missingness of specific species such as SO_4^{2-} and crustal elements such as Fe. Accordingly, two representative cases were selected for downstream PMF evaluation: SO_4^{2-} missingness in Case 2 at a 10% missing rate and high-concentration Fe missingness in Case 5 at a 20% missing rate. The SS and CD factors were used to assess whether these imputation

differences propagated into PMF-resolved source profiles and source contributions. For the SS factor (Figure S32), the SO_4^{2-}/ NH_4^+ mass ratio derived from the PMFr-completed dataset was 3.39 (NH_4^+ associated with NH_4NO_3 removed), close to that from the complete observed dataset (3.44). BPCA also produced a comparable ratio of 3.33, whereas LI (3.93), KNN (2.91), DBN (2.55), and Mean (3.83) showed larger deviations. This indicates that PMFr better preserved the SO_4^{2-}- NH_4^+ relationship in the SS profile, which is critical for maintaining the chemical interpretability of the SS factor. For the CD factor (Figure S33), PMFr also reproduced the crustal elemental ratios consistently. The Fe/Ca and Ca/Si ratios from the complete observed dataset were 0.93 and 1.64, respectively, while PMFr yielded corresponding values of 0.95 and 1.62. In contrast, larger deviations were observed for several baseline methods, such as DBN for Fe/Ca (1.21) and BPCA or LI for Ca/Si (1.43 and 1.44, respectively). These results suggest that inappropriate imputation can alter the resolved source-profile composition, whereas PMFr maintains the physical consistency of source profiles.

For the SS factor contributions, PMFr achieved the highest Pearson's correlation coefficient (r) of 0.943, followed by KNN (0.914), BPCA (0.913), LI (0.900), Mean (0.802), and DBN (0.743). For the CD factor contributions, PMFr also showed the highest temporal agreement, with an r of 0.954, followed by Mean (0.948), KNN (0.926), BPCA (0.925), LI (0.796), and DBN (0.706). These results indicate that competitive concentration-level imputation does not necessarily guarantee equivalent preservation of PMF-resolved source-contribution patterns.

As shown in Figure S34a,b, the PMFr-derived SS contribution closely reproduced the diurnal pattern from the original complete dataset, particularly during daytime periods when secondary sulfate formation is expected to be enhanced. Similarly, PMFr captured the diurnal variation of CD more consistently than baseline methods, especially around the daytime peak likely associated with dust resuspension and other daytime dust-related activities. The selected time-series episodes showed the same behavior (Figure S35a,b). For Case 2 at a 20% missing rate, PMFr achieved the highest r of 0.985 for SS, compared with BPCA (0.981), KNN (0.980), DBN (0.954), LI (0.936), and Mean (0.705). For the high-Fe missing case, PMFr also showed the highest agreement for CD, with an r of 0.951, followed by Mean (0.928), BPCA (0.888), KNN (0.881), DBN (0.713), and LI (0.683). Therefore, the advantage of PMFr is not limited to pointwise concentration accuracy; it also better preserves the chemical and temporal source structures needed for physically interpretable PMF source apportionment. These results indicate that inaccurate imputation may propagate into PMF analysis and introduce source-apportionment biases, potentially making the imputed dataset less reliable than one processed using conventional PMF missing-value treatments."

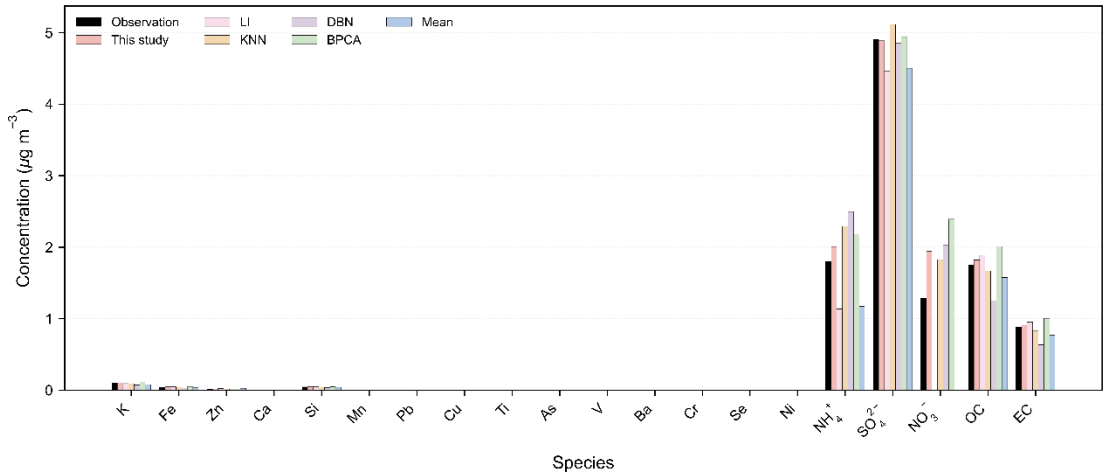

Figure S32. Comparison of PMF-resolved secondary sulfate source profiles derived from datasets completed using different imputation methods and from the complete non-missing dataset under Case 2.

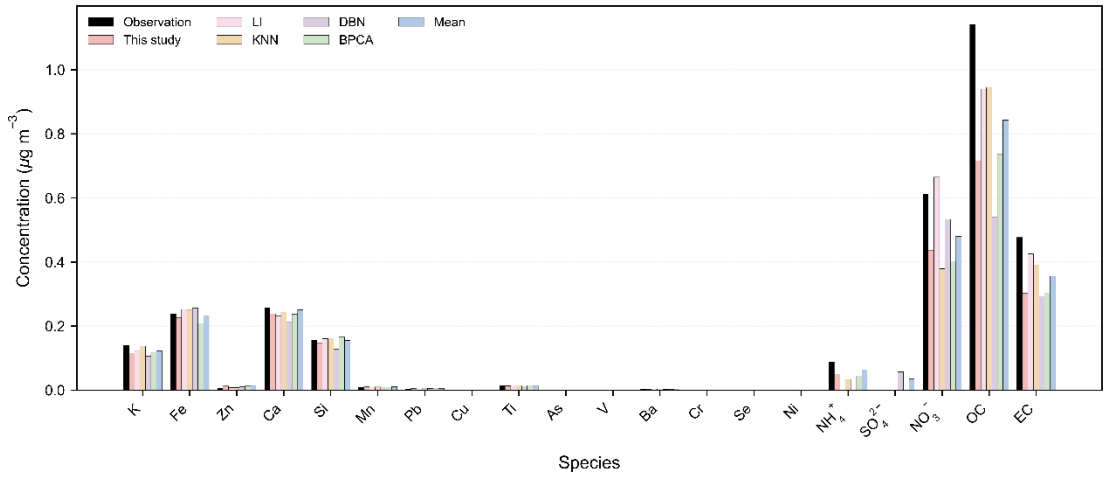

Figure S33. Comparison of PMF-resolved crustal dust source profiles derived from datasets completed using different imputation methods and from the complete non-missing dataset under Case 5.

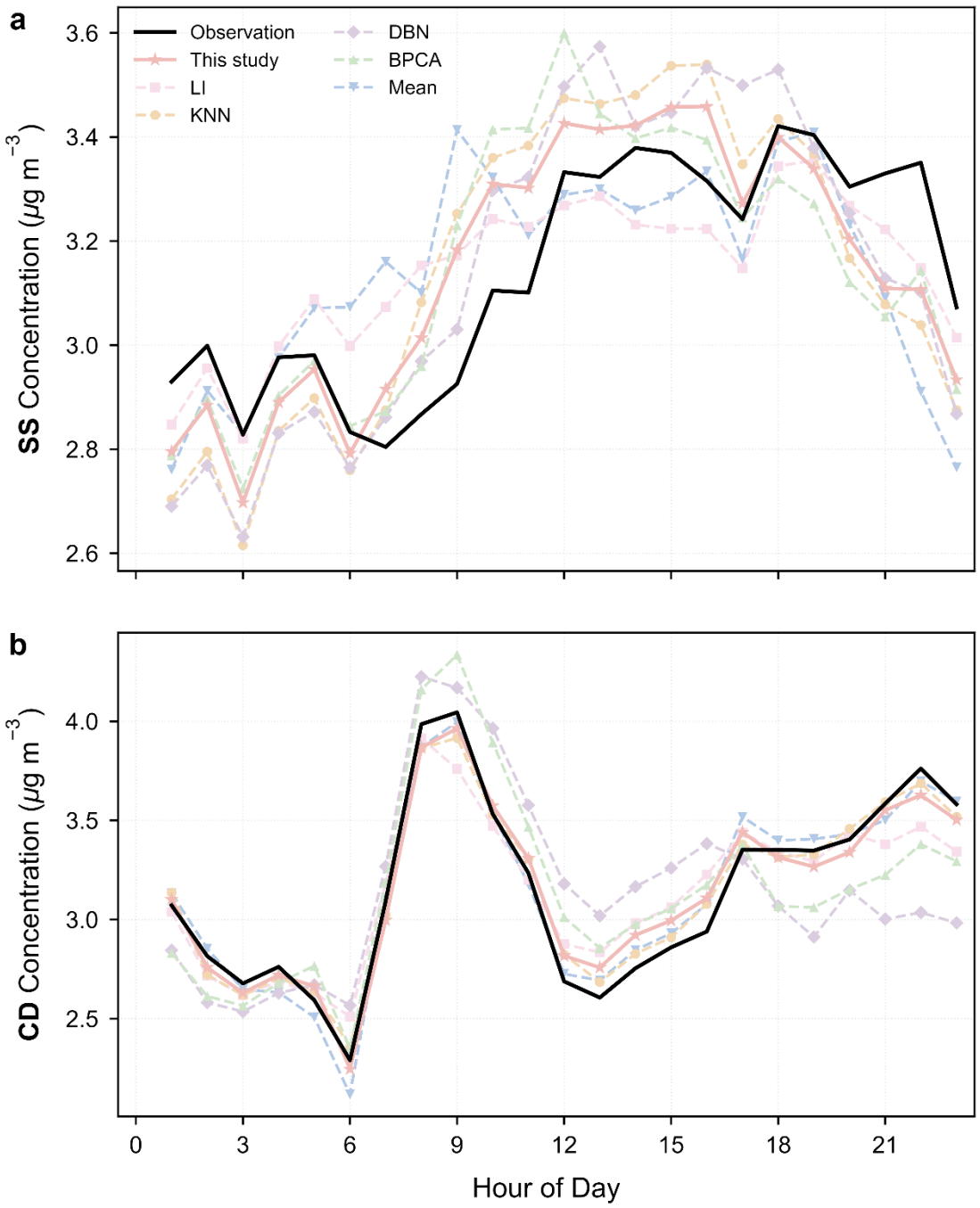

Figure S34. Diurnal variations of PMF-resolved source contributions after imputation for representative cases: **(a)** secondary sulfate (SS) contribution under Case 2 and **(b)** crustal dust (CD) contribution in Case 5.

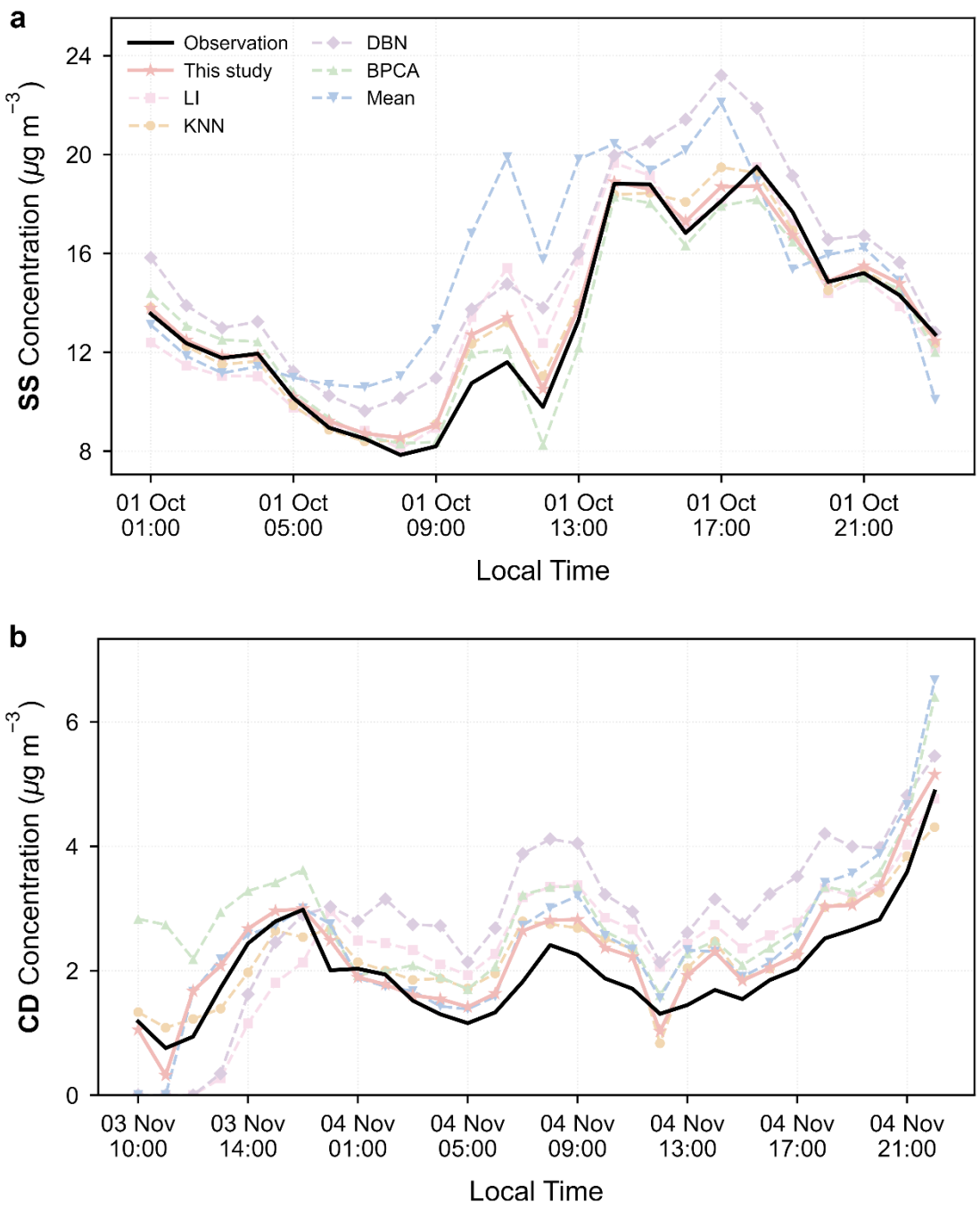

Figure S35. Selected time-series episodes of PMF-resolved source contributions after imputation: **(a)** secondary sulfate (SS) contribution under Case 2 and **(b)** crustal dust (CD) contribution under Case 5.

Minor Comments:

Comment #1:

The manuscript contains a number of grammatical and stylistic issues, including subject-verb agreement errors, awkward phrasing, incorrect verb forms, article usage problems, and inconsistent spacing around parentheses.

(e.g., “by multipling”, “The best-fitting solution were selected”, “Potassium (K) was

treat as missing", "PMFr achieve/decline/capture", and "~of scaled residuals (Reff et al., 2007;~")

Response:

We sincerely apologize for the grammatical and stylistic errors in the original manuscript. We have performed a comprehensive professional proofreading of the entire text to ensure clarity and technical accuracy.

1. Changed "by multipling the estimated source-specific $PM_{2.5}$ mass" to "by multiplying the estimated source-specific $PM_{2.5}$ mass by the resolved source profiles." **(Line 45-46)**
2. Changed "Potassium (K) was treat as missing" to "Potassium (K) was treated as missing." **(Line 94)**
3. Changed "distributions of scaled residuals(Reff et al., 2007; Brown et al., 2015)" to "distributions of scaled residuals (Reff et al., 2007; Brown et al., 2015)" **(Line 133)**
4. Changed "both with low standard deviation" to "both with low standard deviations." **(Line 180-181)**
5. Changed "The best-fitting solution were selected" to "The best-fitting solution was selected." **(Line 131-132)**
6. Changed "with DBN exhibiting lower standard deviation" to "with DBN exhibiting a lower standard deviation." **(Line 181)**
7. Changed "according to the trend indicator R^2, 0.96 and 0.91, respectively." to "For inorganic ions, PMFr performs best when imputing NH_4^+ and NO_3^- with R^2 values of 0.96 and 0.91, respectively." **(Line 182)**
8. Changed "The performance of PMFr decline" to "The performance of PMFr declines." **(Line 184)**
9. Changed "Nevertheless PMFr still outperforms" to "Nevertheless, PMFr still outperforms." **(Line 185)**
10. Changed "struggle to impute high SO_4^{2-} concentrations accurately." to "struggle to accurately impute high SO_4^{2-} concentrations." **(Line 188)**
11. Changed "typically dominates nitrate" to "typically dominates the nitrate fraction." **(Line 190)**
12. Changed "but also higher MAPE" to "but also a higher MAPE." **(Line 193)**
13. Changed "element concentration fluctuates" to "element concentrations fluctuate." **(Line 196-197)**
14. Changed "missing data(Junninen et al., 2004)" to "missing data (Junninen et al., 2004)." **(Line 202-203)**
15. Changed "caused by random missing." to "caused by random missingness." **(Line 208)**

16. Changed “when the missing percentage are 10% and 20%” to “when the missing percentages are 10% and 20%.” **(Line 213-214)**
17. Changed “owing to constructed source-receptor relationships” to “owing to the constructed source-receptor relationships.” **(Line 216-217)**
18. Changed “the difficulty machine-learning methods face” to “the difficulties that machine-learning methods face.” **(Line 217)**
19. Changed “shows best agreement” to “shows the best agreement.” **(Line 209)**
20. Changed “PMFr capture the temporal variability” to “PMFr captures the temporal variability.” **(Line 228-229)**
21. Changed “under Case 6-8” to “under Cases 6-8.” **(Line 236-237)**
22. Changed “the degradation being substantial in Case 6” to “with the degradation being substantial in Case 6.” **(Line 241-242)**
23. Changed “proved by Lee et al. (Lee et al., 2023)” to “have also been reported by Lee et al. (2023)” **(Line 254)**

Comment #2:
Table 1 should be checked for ion notation consistency. Several ionic species are listed without charge notation (e.g., NH_4, SO_4, NO_3, Ca, K), whereas the main text uses formal ionic expressions.

Response: Modified.

Comment #3:
The conclusion section does a good job of emphasizing the promise of PMFr, but it would be stronger if it also briefly acknowledged the conditions under which the method appears less robust. A short, balanced statement about both the advantages and the observed limitations would make the paper’s ending more convincing and scientifically grounded.

Response:
We appreciate the reviewer’s valuable suggestion, which helped us make the Conclusion section more balanced and scientifically grounded. To address this comment, we revised the Conclusion section to include a balanced statement on the observed limitations of PMFr. Specifically, we now state that the advantage of PMFr may become less substantial when source-related constraints are weakened, such as when all key tracers for a specific source factor are simultaneously missing, or when baseline methods can already capture stable co-variation patterns for certain species. We also added that the chemically consistent and physically meaningful estimates

produced by PMFr rely on the temporal stability of source chemical compositions, and that source stability should be verified before applying PMFr to long-term datasets.

The revised Conclusion now reads (Line 330-341):

"We developed a physically interpretable imputation method (PMFr) for reconstructing missing $PM_{2.5}$ speciation data by leveraging source--receptor relationships encoded in key chemical species. Benchmarking against commonly used imputation techniques, including Mean, LI, KNN, BPCA, and a deep learning predictive model, demonstrates that PMFr achieves improved accuracy and robustness while preserving physical and chemical interpretability, especially for key marker species. Crucially, the PMFr-completed dataset is better suited for subsequent PMF source apportionment because it preserves source-profile composition and source-contribution temporal features. Nevertheless, the advantage of PMFr may become less substantial when source-related constraints are weakened, such as when all key tracers for a specific source factor are simultaneously missing, or when baseline methods can already capture stable co-variation patterns for certain species. These chemically consistent and physically meaningful estimates also rely on the temporal stability of source chemical compositions. Recognizing the limitations of such static assumptions for long-term datasets, we highlight the necessity of systematically verifying source stability in extended applications. Therefore, this work offers a simple and generalizable solution that strengthens the reliability of real-world speciation datasets and enhances their suitability for source apportionment and policy-relevant analyses."