Reviewer 1

This manuscript reports and interprets the reductions in ground-level PM2.5 and PM10 as observed by the air quality surveillance network in China during 2015-2020, using a machine learning approach to attribute these changes to drivers of emissions and meteorology. A key finding is that anthropogenic emissions are dominant in the observed changes. While I find the scope fits ACP well, I cannot recommend acceptance of this paper at its present form. The main concern is the severely lack of novelty in all aspects (data, method, and insights from the analysis) among a wealth of literature.

Response: We thank this reviewer for his comments. We will respond to his comments point by point as shown below.

Main comments:

1) Method: the inclusion of concentrations of PM2.5 (PM10), SO2, NO2, O3 and CO in the machine learning (ML) model of PM10 (PM2.5) is very confusing (and inadequate in my opinion). The ultimate aim of this approach is to separate contributions from emissions and meteorology to the changes in PM2.5 and PM10. Meanwhile, these pollutant concentrations themselves are jointly determined by both factors. In Line 193-205, the authors fix emissions in 2015 in the trained ML model to separate the two contributions, so the variations and trends driven by these pollutant concentrations (and these variations are in the top-7 ranks according to their importance scores) are attributed to "meteorology", which is essentially incorrect.

Response: We sincerely thank the reviewer for this insightful comment. In response to this concern, we have substantially revised our machine learning model in the updated version of the manuscript. Specifically, we have now exclusively included emissions, meteorological factors, and temporal descriptors as input variables, while removing the concentrations of $PM_{2.5}$, PM_{10}, SO_2, NO_2, O_3, and CO from the model. This adjustment ensures a clearer and more appropriate attribution of contributions from emissions and meteorology to the changes in PM levels. We are pleased to note that the revised model still demonstrates strong performance, effectively capturing the variations in PM concentrations without the inclusion of other pollutant variables. This modification also aligns more logically with our study's objective of separating emission-driven and meteorology-driven influences. The relevant sections of the manuscript, including the Methods (See Lines 180–264) and Results (See Lines 348–377), have been updated accordingly to reflect these changes.

2) There are many existing papers that used statistical and machine learning models to attribute the changes of air pollution in China into emission and meteorological contributions. I list several examples below.

a. https://acp.copernicus.org/articles/19/11031/2019/

b. https://www.sciencedirect.com/science/article/pii/S0160412023006347

c. https://acp.copernicus.org/articles/19/11303/2019/

d. https://pubs.acs.org/doi/full/10.1021/acs.est.2c06800

e. https://acp.copernicus.org/articles/21/9475/2021/

and many more. The method of this paper exhibits no significant improvement/novelty relative to the above papers. The data locations and time period are also well covered by these papers. Results and insights from this manuscript, without a process-based model or framework, are overall shallow based on the ML model alone. There are few novel insights or analyses compared to the above papers.

Response: We thank the reviewer for raising this important point regarding the novelty of our work and for providing the list of relevant literature. We agree that several excellent studies have explored the separation of emission and meteorological contributions to air pollution in China. However, our manuscript provides significant advancements and novel insights in the following key aspects, which distinguish it from the existing body of work: First, **extended and more dataset:** Our analysis extends the investigation to the period **2015–2022**. This more recent timeframe captures the crucial later phases of China's Air Pollution Prevention and Control Action Plan, as well as the unique emission variations associated with the COVID-19 pandemic and subsequent economic recovery, which are not covered by the cited studies (which end in 2020 or earlier). Second, **comprehensive Analysis of Both $PM_{2.5}$ and PM_{10}:** The studies listed by the reviewer primarily focus on either $PM_{2.5}$ or PM_{10}. A key novelty of our work is the **simultaneous and comparative analysis of both particulate matter species** within a unified methodological framework. This allows for a direct investigation of the differing drivers and behaviors of fine and coarse particles, providing a more holistic understanding of particulate air pollution. Third, **Application of SHAP for long-term attribution and physical consistency evaluation. Although machine-learning models have been used previously, the application of SHAP-based feature attribution to multi-year PM variability is still very limited. Our study extends SHAP usage from short-term prediction contexts to long-term trend attribution, allowing transparent quantification of how key predictors contribute to interannual PM changes.**

Importantly, we also demonstrate that emission features and their SHAP attributions exhibit strong temporal consistency ($R \approx 0.89–0.95$) across species such as SO_2, NOx, and BC, providing a physically interpretable connection between emission evolution and model-inferred contributions. This strengthens confidence in the mechanistic fidelity of ML-derived conclusions. Fourth, **Cross-validated, out-of-sample year-by-year reconstruction. Unlike many studies that train and evaluate models within the same period, our leave-one-year-out design produces true out-of-sample predictions for each individual year, enabling more credible reconstruction of meteorology-only and emission-only scenarios for trend decomposition.**

Overall, rather than proposing a completely new attribution paradigm, our contribution lies in extending the observational period, integrating $PM_{2.5}$ – PM_{10} analyses, and introducing SHAP-based long-term mechanistic attribution with strict temporal cross-validation. These elements together provide new quantitative evidence on how anthropogenic precursors and meteorology shaped PM evolution during the most recent decade.

3) The section of "4. Discussions" introduces new analysis of the correlations of PM2.5/PM10 vs. the other observed concentrations of CO, NO2 and SO2. This piece

emerges randomly and doesn't fit well within the story of machine-learning interpretation of PM trends. It is also unusual to introduce new results in the "Discussion" section.
Overall, the manuscript reads to me a shallow analysis of air quality trends in China, a well-covered topic in existing work. This work does not offer a substantial contribution beyond the existing literature.

Response: We sincerely thank the reviewer for raising this important structural concern. Following your suggestion, we have substantially revised the Discussion section to ensure clear logical integration with the LightGBM–SHAP interpretation framework.
Instead of presenting correlation analysis as an isolated new result, we now：(1) Use Fig. 8 (ambient CO/NO_2/SO_2 distributions) and Table 1 (correlations) only as context to support the SHAP-based emission attribution results. (2) Remove unrelated exploratory statistics, and connect all correlation patterns directly back to the ML-derived mechanisms discussed earlier. (3) Ensure that no new analysis beyond what is necessary for interpreting SHAP contributions is added in the Discussion.The revised Discussion now functions as an interpretation chapter rather than a results chapter, fully aligned with the narrative of SHAP-based attribution.
This improved version appears in Section 5 (See Lines 439–489) of the revised manuscript.

Other comments:

1) Line 21: The PM2.5 and PM10 trends appear very small to me. Check if correct.

Response: Thank you for pointing this out. Upon re-examining our calculations, we identified that the previously reported trend values were affected by an issue in the annual-mean preprocessing step. This has now been fully corrected. The revised $PM_{2.5}$ and PM_{10} interannual trends—together with their 1-σ uncertainties, relative (% yr^{-1}) trends, and associated p-values—are presented in the updated Table S3. The corrected results are also reflected in the rewritten Section 4.1, where we now report the updated magnitudes, uncertainties, and statistical significance of the trends for all cities. Please see lines 314–347 in the revised manuscript for the updated trend analysis and discussion.

2) Line 31: Throughout the paper, there is little explanation of the so-called "co-emission-chemical transformation-meteorological synergy". Also, if this topic is not a core finding from the work, it might not be suitable in the abstract.

Response: Thank you for your insightful comment. In the earlier version of the manuscript, the term “co-emission–chemical transformation–meteorological synergy” appeared in the abstract and main text without a clear definition, which may have overstated the conceptual scope of our findings. Following your suggestion, we have removed this terminology entirely. In the revised manuscript, we no longer use any “synergy” wording. Instead, the Discussion section provides a purely descriptive summary based on the observed correlations and SHAP-derived interactions. Specifically, the results indicate that PM variability reflects the combined influence of emissions, secondary chemical processes, and meteorological modulation. To avoid introducing any new conceptual framework, we refer to this only as an “emission–chemical transformation–meteorological coupling” pattern in the Discussion, where it serves solely as a concise summary of the empirical relationships revealed by the analysis (see lines 440–489).

Importantly, this description is not presented as a theoretical mechanism or central conclusion, but simply consolidates what is directly supported by the data and SHAP interpretation. The term does not appear in the abstract or earlier sections of the manuscript.

Line 51-57: I suggest to move these descriptions to follow the first introduction of PM2.5 and PM10 (Line 39).

Response: Thank you for the suggestion. The descriptions originally placed in Lines 51–57 have been moved to follow the first introduction of $PM_{2.5}$ and PM_{10}, as recommended. See revised manuscript, Lines 40–52.

3) Line 60-61: VOC is also a very important category of PM precursors.

Response: Thank you for pointing this out. In the previous version, VOCs were not explicitly mentioned when introducing PM precursors. We have now revised the Introduction to include a clear statement acknowledging the essential role of VOCs in secondary aerosol formation. Specifically, we added the following sentence:

"In addition to these inorganic precursors, volatile organic compounds (VOCs) also play an important role in secondary aerosol formation, particularly through pathways leading to secondary organic aerosols, as recognized in numerous atmospheric chemistry studies."

This revision appears in the Introduction (Lines 62–65) of the revised manuscript.

4) Line 66-67: Secondary aerosols can be formed in both the boundary layer and free troposphere. I do not know the purpose of emphasizing "free atmosphere" here.

Response: Thank you for this helpful suggestion. We agree that the term *"free atmosphere"* was unnecessary and could be misleading, since secondary aerosol formation occurs throughout the atmospheric column. Accordingly, we have revised the text by replacing "free atmosphere" with "atmospheric" to ensure clarity and accuracy. This modification has been implemented in the revised manuscript (Lines 69–70).

5) Line 75: The paper (Zhang et al. 2016) is not a "conventional linear modeling approach". Please cite adequate papers.

Response: Thank you for pointing this out. We agree that *Zhang et al. (2016)* is not representative of traditional linear regression–based PM prediction methods. In the revised manuscript, we have replaced this citation with an appropriate reference illustrating a conventional multivariate linear regression approach. The correction has been implemented in Lines 78 of the revised manuscript.

The new citation is:

Zhao, R., Gu, X., Xue, B., Zhang, J., and Ren, W.: Short period $PM_{2.5}$ prediction based on multivariate linear regression model, PLOS ONE, 13, e0201011, https://doi.org/10.1371/journal.pone.0201011, 2018.

6) Line 103-104: Besides the table, should provide a map of these cities. Without a map it is very hard to locate them.

Response: Thank you for the helpful suggestion. We have added a dedicated city-distribution map to improve geographic clarity. The map has been included as Fig. S1 in the Supplementary Material, and the manuscript now references this figure in Lines 106–108 of the revised version.

7) Line 114-115: The "reference state" of air pollutant measurements was at 273 K before September 2018, and at 298 K afterwards. Is this factor considered?

Response: Thank you for pointing this out. Yes, this reference-state change has been fully considered. All pollutant concentration records were used in their standardized, quality-controlled form as provided by the national monitoring network, in which the conversion between the 273 K and 298 K reference states is already applied during data processing. Therefore, no inconsistency was introduced into our model input.

8) Line 118-119: Why GEOS-FP is chosen while more stable met fields (e.g., MERRA2) are available?

Response: Thank you for raising this question. We agree that MERRA-2 is a widely used and dynamically consistent reanalysis product. In our study, we chose GEOS-FP primarily because of its finer native resolution (0.25° × 0.3125°) and its near–real-time availability for 2015–2022, which is well suited to resolving mesoscale temperature, humidity and wind gradients over the BTH and YRD regions and to matching our analysis period. In addition, recent studies have shown that GEOS-FP provides meteorological fields of sufficient quality for air-pollution and hydrometeorological applications. Chen et al. (2023) evaluated four meteorological reanalysis datasets, including both GEOS-FP and MERRA-2, for satellite-based $PM_{2.5}$ retrieval over China and confirmed that GEOS-FP can be reliably used as meteorological input for $PM_{2.5}$ estimation over our study domain. Huang et al. (2023) used GEOS-FP numerical weather prediction fields as part of a global flood-forecasting system and demonstrated that GEOS-FP precipitation forcing can support skillful hydrological predictions. On this basis, we consider GEOS-FP an appropriate choice for our work, mainly because it offers higher spatial resolution and up-to-date coverage while having been successfully applied in related studies. We do not claim that GEOS-FP is universally superior to MERRA-2; rather, it provides a practical and well-validated meteorological driver for our specific application.

References:

Chen, Z., Chen, J., Zhang, Y., Jiang, Y., Liu, M., Liu, H., Zhao, W., and Yan, X.: Evaluation of four meteorological reanalysis datasets for satellite-based $PM_{2.5}$ retrieval over China, Atmos. Environ., 305, 119795, https://doi.org/10.1016/j.atmosenv.2023.119795, 2023.

Huang, Z., Wu, H., Gu, G., Li, X., Nanding, N., Adler, R. F., et al.: Paired satellite and NWP precipitation for global flood forecasting, J. Hydrometeorol., 24, 2191–2205, https://doi.org/10.1175/JHM-D-23-0044.1, 2023.

9) Line 135: "Paraffinic reactive primary emissions" is not a conventional term. Could you please change it to "VOC emissions" and list the VOC species you used?

Response: Thank you for pointing out this issue. We agree that "paraffinic reactive primary emissions (PRPE)" is not a conventional term for representing the broader VOC spectrum. In the CEDS inventory, PRPE corresponds to the paraffinic fraction of NMVOCs and is primarily designed for use in chemical transport models (e.g., GEOS-Chem). During our data extraction, PRPE was the only VOC-related category with complete sectoral and temporal coverage for 2015–2022, while other VOC sub-species showed incomplete availability across sectors or years. Accordingly, PRPE was adopted as the VOC-related proxy consistently provided by CEDS within the constraints of our analysis framework. We also note in the Conclusions (See lines 517-523) that future studies should incorporate more comprehensive VOC categories as higher-resolution or more complete emission inventories become available.

10) Equations 1-5 and associated text: are these very conventionally accepted concepts really worth such detailed discussion in the main text?

Response: Thank you for this helpful suggestion. We agree that the original level of detail was unnecessary for concepts that are already well established in the literature. Accordingly, we have removed Equations 1–5 and substantially shortened the accompanying explanations. A concise descriptive summary is now provided in the revised manuscript (Lines 197–208), improving readability while retaining the essential methodological information.

11) Section 3.1: Again, these trends (<0.1 ug/m3/yr for most cases) appear too small to me according to my understanding of air quality changes in China.

Response: Thank you for pointing this out. We re-examined our computations and confirmed that the originally reported trend values were affected by an error in the annual-mean preprocessing step. This issue has now been corrected. In the revised manuscript, the $PM_{2.5}$ and PM_{10} interannual trends have been fully recalculated, and the results—including the absolute trends, their 1-σ uncertainties, the relative (% yr^{-1}) trends, and their corresponding uncertainties—are all updated in Table S3. Section 4.1 has also been revised accordingly to reflect the corrected magnitudes and statistical significance of the trends, and now reports the complete set of updated indicators. The revised discussion of these results can be found in lines 315–347.

12) Line 239: The scatter plots in Figure 2 have too many overlapping points, and should be converted to colored 2-d histogram density plot.

Response: Thanks for your comments. We have modified the corresponding figures (Please see figure2).

13) Figure 4: Why are the meteorology-driven changes overall opposite for PM2.5 (positive) and PM10 (negative)? What is the key parameter causing this?

Response: Thank you for the comment. In the revised analysis, the updated LightGBM model yields meteorology-driven changes that are consistent for both $PM_{2.5}$ and PM_{10}, rather than opposite as in the previous version. The discrepancy observed in the earlier manuscript has

therefore been corrected. The physical interpretation of the revised results is provided in the main text (Lines 379–404).

14) Figures 6 and 7: Instead of showing the trends of these parameters, it might be more straightforward to support the analysis by showing the contributions of each parameter to PM2.5 and PM10 trends?

Response: Thank you for the suggestion. In the revised manuscript, we use Fig. 6 and Fig. 7 to support the contribution analysis. These figures demonstrate that each emission or meteorological variable shows strong temporal consistency with its SHAP-derived contribution ($R \geq 0.89$ for emission precursors and $R \approx -0.96$ for meteorological drivers). This confirms that the SHAP-based attribution responds coherently to the real temporal evolution of each driver, thereby strengthening the credibility of the derived emission- and meteorology-related influences on $PM_{2.5}$ and PM_{10}.

15) Line 343: Clarify if the "correlations" are calculated based on hourly or daily data

Response: Thank you for the comment. The correlations reported in this section are calculated using hourly observations of $PM_{2.5}$/PM_{10} and the corresponding hourly concentrations of CO, NO_2, and SO_2.

17) Section 4: Based on these correlations alone, no conclusive argument can be made, as also indicated by many conjecturing text in this section. I find it hard to understand the purpose of this section and this analysis.

Response: Thank you very much for this constructive comment. We appreciate the reviewer's concern that, in the earlier version, the discussion section relied on several exploratory correlations that were not fully integrated with the machine-learning–based attribution framework. We agree that this issue could reduce the clarity and focus of the interpretation. In the revised manuscript, we have reorganized the content accordingly. As part of the restructuring of the Results and Discussion sections, the original Section 4 has now become Section 5. The section has been rewritten to function strictly as an interpretive extension of the SHAP-derived attribution results, rather than as a set of independent analyses. Specifically: (1) Figure 8 is now used solely to provide spatial context for CO, NO_2, and SO_2 emission patterns in the BTH and YRD regions. Its purpose is to support the interpretation of why these precursors exhibit the correlation strengths listed in Table 1, without drawing conclusions from the correlations themselves. (2) Table 1 is presented only as supporting observational evidence, demonstrating that the empirical co-variation among $PM_{2.5}$/PM_{10} and their precursors is consistent with the species identified as influential by the SHAP analysis. No inference is made based solely on correlation statistics. (3) Figure 9 (the T2M–NO_x interaction) is now explicitly linked to the SHAP-based findings and serves only to illustrate the temperature-dependent modulation of NO_x-related $PM_{2.5}$ formation. This interpretation aligns with established atmospheric chemistry understanding and does not introduce new stand-alone results. Importantly, the revised Section 5 no longer introduces additional analyses. It functions entirely

as an interpretive and mechanistic discussion of the machine-learning attribution outputs. We sincerely thank the reviewer for raising this point, which has substantially improved the coherence and clarity of the manuscript.