Reviewer 1

This manuscript investigates the long-term PM2.5 trends across 36 cities in Northeast China from 2015 to 2025 using a LightGBM model integrated with the SHAP explainability framework. By employing a meteorological normalization approach, the authors successfully disentangle the contributions of anthropogenic emissions and meteorological conditions to PM2.5 concentrations. The study is further strengthened by validations using independent indicators such as CEDS emission inventories, VIIRS fire counts, and CO observations. The topic is of significant regional interest, and the methodology is robust. The findings provide valuable insights for air quality management in cold, industrial, and agricultural regions. Minor revision is recommended. I also have some suggestions about this manuscript.

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

Major suggestion

(1) The results indicate that while regional observed PM2.5 levels remained essentially stable after 2022, PMemi increased, suggesting that meteorological factors partially offset the rise in emissions. This finding carries significant policy implications. It is recommended to state more explicitly in the abstract, discussion, and conclusion that the stability of observed PM2.5 does not imply a continued decline in the weather-normalized pollution baseline, and that further emission reductions remain challenging.

Response: Thank you for this important suggestion. We agree that the post-2022 divergence between observed $PM_{2.5}$ and PM_{emi} has important implications for interpreting recent air-quality changes in Northeast China.

We have therefore revised the Abstract, Discussion, and Conclusion to state this point more explicitly. After 2022, regional observed $PM_{2.5}$ remained nearly stable, whereas PM_{emi} increased, indicating that the apparent stability of observed concentrations partly reflected offsetting meteorological modulation rather than continued improvement in the weather-normalized pollution baseline. We now emphasize that stable observed $PM_{2.5}$ concentrations should not be interpreted as evidence of continued improvement in the underlying non-meteorological pollution baseline.

We have also strengthened the policy interpretation by noting that regional annual mean $PM_{2.5}$ remained above 25 μg m^{-3} during 2023–2025 and that achieving further $PM_{2.5}$ reductions is becoming increasingly challenging, requiring additional control of persistent anthropogenic sources such as industrial emissions, winter heating, scattered fuel combustion, and other combustion-related sources. To avoid overinterpreting PM_{emi} as a direct measure of emissions, we describe the post-2022 change as an increase in the weather-normalized pollution baseline rather than as a direct increase in emissions (Lines 27–30, 785–791, and 847–849).

(2) PMemi shows a high correlation with the CEDS index; however, this does not mean that PMemi is an emissions inventory. In the comparison section, it should be emphasized that PMemi is not an emissions inventory, but rather the non-meteorological component derived after weather normalization.

Response: Thank you for this important clarification. We agree that the strong correlation between PM_{emi} and the CEDS emission indices should not be interpreted as indicating that PM_{emi} is an emissions inventory or a direct measure of emissions.

We have therefore further clarified this distinction in the CEDS comparison section. PM_{emi} is explicitly defined as the weather-normalized, non-meteorological component of ambient $PM_{2.5}$ obtained after averaging out short-term meteorological variability. The CEDS emission indices are used only as independent emission-related information to evaluate whether the long-term evolution of PM_{emi} is broadly consistent with anthropogenic emission changes.

We have retained the species-specific comparisons and the first-difference analysis, which show that the PM_{emi}–CEDS agreement mainly reflects long-term co-evolution rather than close year-to-year correspondence. The revised manuscript therefore emphasizes that the CEDS comparison serves as an external consistency check on the long-term behavior of PM_{emi} and does not constitute direct validation of PM_{emi} as an emissions inventory (Lines 525–540, and 747–769).

(3) Using the method where grid center points fall within city polygons may introduce bias in areas with complex terrain; it is recommended to discuss this as a limitation.

Response: Thank you for this important comment. We agree that the grid-center-based aggregation of ERA5 meteorological data may introduce spatial-representation uncertainty, particularly in prefecture-level cities with complex terrain or strong elevation gradients.

In our analysis, city-level meteorological predictors were obtained by averaging ERA5 grid cells whose center points fell within each prefecture-level city polygon, with the nearest grid cell to the polygon centroid used when no grid-cell center fell within the polygon. Given the 0.25° spatial resolution of ERA5, this approach may not fully resolve sub-city meteorological heterogeneity and may therefore introduce uncertainty into the meteorological predictors used for model training and subsequent weather normalization.

We have now added this issue to the Discussion as a limitation of the meteorological interpretation. The revised manuscript explicitly notes that the magnitude of this spatial-representation uncertainty may vary among cities, particularly in areas with complex terrain, and that future studies could reduce this limitation by using higher-resolution meteorological datasets, area-weighted spatial aggregation, or additional station-based meteorological observations (Lines 233–239 and 813–819).

(4) Section 4, “Annual trend and fire-activity comparison” resembles an analysis methodology; it is suggested that this section be merged into Section 3 or renamed “Analysis methods”.

Response: Thank you for this helpful suggestion. We agree that the former Section 4 primarily describes analysis methodology rather than results.

To clarify the manuscript structure, we have renamed Section 4 from “Annual trend and fire-activity comparison” to “Analysis methods”. We retained it as a separate section rather than merging it into Section 3 because Section 3 specifically describes the machine-learning framework, including feature engineering, model training and cross-validation, SHAP interpretation, and meteorological normalization, whereas Section 4 describes the subsequent statistical and external-consistency analyses.

The existing subsections, “4.1 Annual aggregation and trend analysis” and “4.2 External consistency analysis”, have been retained under the revised section title (Lines 418).

(5) Line 471, The statement “current annual standard of 25 μg/m³” should specify whether it refers to the Chinese national standard, WHO guidelines, or another standard. The current Chinese annual secondary standard for PM2.5 is 35 μg/m³, whereas the WHO 2021 interim target is 25 μg/m³. If the latter is intended, please state this explicitly.

Response: Thank you for pointing out this ambiguity. We agree that the reference standard should be stated explicitly.

The 25 μg m^{-3} value in our manuscript refers to the tightened Grade II annual $PM_{2.5}$ limit specified in China’s revised national ambient air quality standard (GB 3095–2026), rather than solely to the WHO guideline. Under GB 3095–2026, this 25 μg m^{-3} annual limit will be implemented nationwide from 2031.

We have therefore revised the relevant text in the Results to identify the 25 μg m^{-3} value explicitly as the tightened Chinese national Grade II annual $PM_{2.5}$ limit and removed the ambiguous wording “current annual standard” (Lines 550–556).

(6) CEDS typically provides NOx rather than NO. Please change “NO” to “NOx” and confirm that the selected species match the CEDS variable names.

Response: Thank you for this comment. We have replaced “NO” with “NOx” throughout the manuscript and Figure 4, and confirmed that the selected CEDS species are BC, OC, SO_2, NOx, and NH_3. The relevant text and figure labels have been revised accordingly.

(7) Avoid presenting the correlation between PMemi and CEDS as direct causal evidence of emission-control policy effects. Use more cautious wording such as "consistent with" or "associated with".

Response: Thank you for this comment. We have revised the relevant wording throughout the manuscript to avoid interpreting the correlation between PM_{emi} and CEDS as direct causal evidence of emission-control policy effects. The CEDS comparison is now described in terms of consistency or association with the long-term evolution of PM_{emi} and is explicitly interpreted as an external consistency check rather than direct evidence of causality. The corresponding wording in the Results, Discussion, and Conclusion has been revised accordingly (Lines 530–545, 746–769, and 833–858).

Minor suggestion

(1) Line 139–140, "For cities with multiple monitoring station" should be "For cities with multiple monitoring stations".

Response: Thank you for pointing this out. We have corrected "monitoring station" to "monitoring stations" in the revised manuscript (Lines 152–154).

(2) Line 529, "meteorology more often acted as a cleaner under normal conditions" should be "meteorology more often favored PM2.5 removal under normal conditions".

Response: Thank you for this suggestion. We have replaced "meteorology more often acted as a cleaner under normal conditions" with "meteorology more often favored $PM_{2.5}$ removal under normal conditions" in the revised manuscript (Lines 615–617).

(3) For title, "meteorology" should be "meteorological".

Response: Thank you for this correction. We have changed "meteorology" to "meteorological" in the manuscript title (Line 2).

(4) Line 30, "thermal" should be replaced as "temperature".

Response: Thank you for this suggestion. We have replaced "thermal" with "temperature" in the revised manuscript (Lines 35–36).

(5) Ensure all abbreviations (e.g., BLH, SSRD) are defined at their first appearance. While most seem defined, a final check is recommended.

Response: Thank you for this suggestion. We performed a final check of abbreviations throughout the manuscript and ensured that technical abbreviations are defined at their first appearance. In particular, BLH, SSRD, and the other meteorological variables are already defined when first introduced, and the definitions of LightGBM and VIIRS have been clarified where appropriate (Lines 19–20, 26–27, 35–36, and 175–177).

(6) Please check the formatting consistency in the reference list, particularly regarding journal abbreviations and year placement.

Response: Thank you for this suggestion. We have carefully checked the reference list and corrected inconsistencies in journal abbreviations, year placement, and overall reference formatting to ensure consistency throughout the manuscript.

(7) I suggest that the authors should carefully review the grammar of this manuscript.

Response: Thank you for this suggestion. We have carefully reviewed the manuscript for grammar, wording, and sentence structure, and corrected the identified language issues throughout the revised manuscript.